NEW DEVELOPMENTS IN QUANTITATIVE TRADING AND INVESTMENT

CHRISTIAN L. DUNIS
PETER W. MIDDLETON
ANDREAS KARATHANASOPOULOS
KONSTANTINOS THEOFILATOS

ARTIFICIAL INTELLIGENCE IN FINANCIAL MARKETS

Cutting-Edge Applications for Risk Management, Portfolio Optimization and Economics

New Developments in Quantitative Trading and Investment

Christian L. Dunis • Peter W. Middleton • Konstantinos Theofilatos
Andreas Karathanasopoulos
Editors

Artificial Intelligence in Financial Markets

Cutting-Edge Applications for Risk Management, Portfolio Optimization and Economics

palgrave
macmillan

Editors
Christian L. Dunis
ACANTO Holding
Hannover, Germany

Peter W. Middleton
University of Liverpool
Liverpool, England

Konstantinos Theofilatos
University of Patras
Patras, Greece

Andreas Karathanasopoulos
American University of Beirut (AUB)
Beirut, Lebanon

ISBN 978-1-137-48879-4 ISBN 978-1-137-48880-0 (eBook)
DOI 10.1057/978-1-137-48880-0

Library of Congress Control Number: 2016941760

Printed on acid-free paper

This Palgrave Macmillan imprint is published by Springer Nature
The registered company is Macmillan Publishers Ltd. London

Preface

The aim of this book is to focus on Artificial Intelligence (AI) and to provide broad examples of its application to the field of finance. Due to the popularity and rapid emergence of AI in the area of finance this book is the first volume in a series called 'New Developments in Quantitative Trading and Investment' to be published by Palgrave Macmillan. Moreover, this particular volume targets a wide audience including both academic and professional financial analysts. The content of this textbook targets a wide audience who are interested in forecasting, modelling, trading, risk management, economics, credit risk and portfolio management. We offer a mixture of empirical applications to different fields of finance and expect this book to be beneficial to both academics and practitioners who are looking to apply the most up to date and novel AI techniques. The objective of this text is to offer a wide variety of applications to different markets and assets classes. Furthermore, from an extensive literature review it is apparent that there are no recent textbooks that apply AI to different areas of finance or to a wide range of markets and products.

Each Part is comprised of specialist contributions from experts in the field of AI. Contributions offer the reader original and unpublished content that is recent and original. Furthermore, as the cohort of authors includes various international lecturers and professors we have no doubt that the research will add value to many MA, MSc, and MBA graduate programmes. Furthermore, for the professional financial forecaster this book is without parallel a comprehensive, practical and up-to-date insight into AI. Excerpts of programming code are also provided throughout in order to give readers the opportunity to apply these techniques on their own.

Authors of this book extend beyond the existing literature in at least three ways. The first contribution is that we have included empirical applications of AI in four different areas of finance: time-series modelling, economics, credit and portfolio management. Secondly, the techniques and methodologies applied here are extremely broad and cover all areas of AI. Thirdly, each chapter investigates different datasets from a variety of markets and asset classes. Different frequencies of data are also investigated to include daily, monthly, macroeconomic variables and even text data from different sources. We believe that the Parts presented here are extremely informative and practical while also challenging existing traditional models and techniques many of which are still used today in financial institutional and even in other areas of business. The latter is extremely important to highlight since all of the applications here clearly identify a benefit of utilizing AI to model time-series, enhance decision making at a government level, assess credit ratings, stock selection and portfolio optimization.

Contents

Part I

Following the introduction, the first part focuses on numerous time-series, which will include commodity spreads, equities, and exchange traded funds. For this part the objective is to focus on the application of AI methodologies to model, forecast and trade a wide range of financial instruments. AI methodologies include, Artificial Neural Networks (ANN), Heuristic Optimization Algorithms and hybrid techniques. All of the submissions provide recent developments in the area of financial time-series analysis for forecasting and trading. A review of publications reveals that existing methodologies are either dated or are limited in scope as they only focus on one particular asset class at a time. It is found that the majority of the literature focuses on forecasting foreign exchange and equities. For instance, Wang et al. [14] focus their research and analysis on forecasting the Shanghai Composite index using a Wavelet-Denoising-based back propagation Neural Network (NN). The performance of this NN is benchmarked against a traditional back propagation NN. Other research is now considered redundant as the field of AI is evolving at a rapid rate. For instance, Zirilli [19] offers a practical application of neural networks to the prediction of financial markets however, the techniques that were used are no longer effective when predicting financial variables. Furthermore, data has become more readily available so input datasets can now be enriched to enable methodologies to capture the relationships between input datasets and target variables more accurately. As a result, more recent research and technological innovations have rendered such methodologies obsolete.

While numerous journal publications apply AI to various assets our search did not uncover recent textbooks that focus on AI and in particular empirical applications to financial instruments and markets. For this reason we believe that an entire section dedicated to time-series modelling, forecasting and trading is justified.

Part II

The second part focuses on economics as a wider subject that encompasses the prediction of economic variables and behavioural economics. Both macro- and micro-economic analysis is provided here. The aim of this part is to provide a strong case for the application of AI in the area of economic modelling and as a methodology to enhance decision making in corporations and also at a government level. Various existing work focuses on agent-based simulations such as Leitner and Wall [16] who investigate economic and social systems using agent-based simulations. Teglio et al. [17] also focus on social and economic modelling relying on computer simulations in order to model and study the complexity of economic and social phenomena. Another recent publication by Osinga et al. [13] also utilizes agent-based modelling to capture the complex relationship between economic variables. Although this part only provides one empirical application we believe that it goes a long way to proving the benefits of AI and in particular 'Business Intelligence'.

With extensive research being carried out in the area of economic modelling it is clear that a whole section should also be devoted to this particular area. In fact we expect this section to draw a lot of attention given its recent popularity.

Part III

The third part focuses on analyzing credit and the modelling of corporate structures. This offers the reader an insight into AI for evaluating fundamental data and financial statements when making investment decisions. From a preliminary search our results do not uncover any existing textbooks that exclusively focus on credit analysis and corporate finance analyzed by AI methodologies. However, the search uncovered a few journal publications that provide an insight into credit analysis in the area of bankruptcy prediction. For instance, Loukeris and Matsatsinis [9] research corporate finance by attempting to predict bankruptcy using AI models. From results produced by these journal publications we believe that corporate finance could benefit from more recent empirical results published in this part.

Earlier research in the area of credit analysis is carried out by Altman et al. [1] who examine the use of layer networks and how their use has led to an improvement in the reclassifying rate for existing bankruptcy forecasting models. In this case, it was found that AI helped to identify a relationship between capital structure and corporate performance.

The most recent literature reviewed in the area of corporate finance applies AI methodologies to various credit case studies. We suspect that this was inspired by the recent global credit crisis in 2008 as is the case with most credit-based research published after the 2008 'credit crunch'. For instance, Hajek [6] models municipal credit ratings using NN classification and genetic programs to determine his input dataset. In particular, his model is designed to classify US municipalities (located in the State of Connecticut) into rating classes based on their levels of risk. The model includes data pre-processing, the selection process of input variables and the design of various neural networks' structures for classification. Each of the explanatory variables is extracted from financial statements and statistical reports. These variables represent the inputs of NNs, while the rating classes from Moody's rating agency are the outputs. Experimental results reveal that the rating classes assigned by the NN classification to bond issuers are highly accurate even when a limited sub-set of input variables is used. Further research carried out by Hajek [7] presents an analysis of credit rating using fuzzy rule-based systems. A fuzzy rule-based system adapted by a feed-forward neural network is designed to classify US companies (divided into finance, manufacturing, mining, retail trade, services, and transportation industries) and municipalities into the credit rating classes obtained from rating agencies. A genetic algorithm is used again as a search method and a filter rule is also applied. Empirical results corroborate much of the existing research with the classification of credit ratings assigned to bond issuers being highly accurate. The comparison of selected fuzzy rule-based classifiers indicates that it is possible to increase classification performance by using different classifiers for individual industries.

León-Soriano and Muñoz-Torres [8] use three layers feed-forward neural networks to model two of the main agencies' sovereign credit ratings. Their results are found to be highly accurate even when using a reduced set of publicly available economic data. In a more thorough application Zhong et al. [20] model corporate credit ratings analyzing the effectiveness of four different learning algorithms. Namely, back propagation, extreme learning machines, incremental extreme learning machines and support vector machines over a data set consisting of real financial data for corporate credit ratings. The results reveal that the SVM is more accurate than its peers.

With extensive research being carried out in the area of bankruptcy prediction and corporate/sovereign credit ratings it is clear that the reader would benefit from a whole section being devoted to credit and corporate finance. In fact the first chapter provides an interesting application of AI to discover which areas of credit are most popular. AI is emerging in the research of credit analysis and corporate finance to challenge existing methodologies that were found to be inadequate and were ultimately unable to limit the damage caused by the 2008 'credit crisis'.

Part IV

The final section of the book focuses on portfolio theory by providing examples of security selection, portfolio construction and the optimization of asset allocation. This will be of great interest to portfolio managers as they seek optimal returns from their portfolios of assets. Portfolio optimization and security selection is a heavily researched area in terms of AI applications. However, our search uncovered only a few existing journal publications and textbooks that focus on this particular area of finance. Furthermore, research in this area is quickly made redundant as AI methodologies are constantly being updated and improved.

Existing journal publications challenge the Markowitz two-objective mean-variance approach to portfolio design. For instance, Subbu et al. [15] introduce a powerful hybrid multi-objective optimization approach that combines evolutionary computation with linear programming to simultaneously maximize return, minimize risk and identify the efficient frontier of portfolios that satisfy all constraints. They conclude that their Pareto Sorting Evolutionary Algorithm (PSEA) is able to robustly identify the Pareto front of optimal portfolios defined over a space of returns and risks. Furthermore they believe that this algorithm is more efficient than the 2-dimensional and widely accepted Markowitz approach.

An older textbook, which was co-authored by Trippi and Lee (1995), focuses on asset allocation, timing decisions, pattern recognition and risk assessment. They examine the Markowitz theory of portfolio optimization and adapt it by incorporating it into a knowledge-based system. Overall this is an interesting text however it is now almost 20 years old and updated applications/methodologies could be of great benefit to portfolio managers and institutional investors.

The Editors

All four editors offer a mixture of academic and professional experience in the area of AI. The leading editor, Professor Christian Dunis has a wealth of experience spanning more than 35 years and 75 publications, both in academia and quantitative investments. Professor Dunis has the highest expertise in modelling and analyzing financial markets and in particular an extensive experience with neural networks as well as advanced statistical analyses. Dr Peter Middleton has recently completed his PhD in Financial Modelling and Trading of Commodity Spreads at the University of Liverpool. To date he has produced five publications and he is also a member of the CFA institute and is working towards the CFA designation having already passed Level I. He is also working in the finance industry in the area of Asset Management. Dr Konstantinos possesses an expertise in technical and computational aspects with backgrounds in evolutionary programming, neural networks, as well as expert systems and AI. He has published numerous articles in the area of computer science as well being an editor for *Computational Intelligence for Trading and Investment*. Dr Andreas Karathanasopoulos is currently an Associate Professor at the American University of Beirut and has worked in academia for six years. He too has numerous publications in international journals for his contribution to the area of financial forecasting using neural networks, support vector machines and genetic programming. More recently he has also been an editor for *Computational Intelligence for Trading and Investment*.

Acknowledgements

We would like to thank the authors of who have contributed original and novel research to this book, the editors who were instrumental in its preparation and finally the publishers who have ultimately helped provide a showcase for it to be read by the public.

Final Words

We hope that the publication of this book will enhance the spread of AI throughout the world of finance. The models and methods developed here have yet to reach their largest possible audience, partly because the results are scattered in various journals and proceedings volumes. We hope that this book will help a new generation of quantitative analysts and researchers to solve complicated problems with greater understanding and accuracy.

France Christian L. Dunis
UK Peter Middleton
Greece Konstantinos Theofilatos
Lebanon Andreas Karathanasopoulos

References

1. E.I. Altman, G. Marco, F. Varetto, Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience), Journal of Banking and Finance 18 (1994) 505±529.
2. Hájek, P. (2011). Municipal credit rating modelling by neural networks. *Decision Support Systems, 51*(1), 108–118.
3. Hajek, P. (2012). Credit rating analysis using adaptive fuzzy rule-based systems: An industry-specific approach. *Central European Journal of Operations Research, 20*(3), 421–434.
4. León-Soriano, R. and Muñoz-Torres, M. J. (2012). Using neural networks to model sovereign credit ratings: Application to the European Union. *Modeling and Simulation in Engineering, Economics and Management: Lecture Notes in: Business Information Processing*, *115*, 13–23.
5. Loukeris, N. and Matsatsinis, N. (2006). *Corporate Financial Evaluation and Bankruptcy Prediction Implementing Artificial Intelligence Methods*. Proceedings of the 10th WSEAS International Conference on COMPUTERS, Vouliagmeni, Athens, Greece, July 13–15, 2006. Pp. 884–888.
6. Osinga, E. C., Leeflang, P. S. H., Srinivasan, S., & Wieringa, J. E. (2011). Why do firms invest in consumer advertising with limited sales response? A shareholder perspective. Journal of Marketing, 75(1), 109–124.
7. QIAO Yu-kun,WANG Shi-cheng,ZHANG Jin-sheng,ZHANG Qi,SUN Yuan (Department of Automatic Control,The Second Artillery Engineering College,Xi'an 710025,Shaanxi,China);Simulation Research on Geomagnetic Matching Navigation Based on Soft-threshold Wavelet Denoising Method[J];Acta Armamentarii;2011-09.
8. Subbu, R., Bonissone, P. P., Eklund, N., Bollapragada, S., and Chalermkraivuth, K. (2005). Multiobjective Financial Portfolio Design: A Hybrid Evolutionary Approach. In *2005 IEEE Congress on Evolutionary Computation (CEC'2005)*, vol. 2. Edinburgh, Scotland: IEEE Service Center, September 2005, pp. 1722–1729.

9. Leitner S.and F. Wall. Multi objective decision-making policies and coordination mechanisms in hierarchical organizations: Results of an agent-based simulation. Working Paper, Alpen-Adria Universit¨at Klagenfurt (in submission), 2013.
10. Teglio, A., Raberto, M., Cincotti, S., 2012. The impact of banks' capital adequacy regulation on the economic system: an agent-based approach. Advances in Complex Systems 15 (2), 1250040–1 – 1250040–27.
11. Zirilli, J. S., 1997: Financial Prediction Using Neural Networks. International Thomson, 135 pp.
12. Zhong, H., Miao, C., Shen, Z., and Feng, Y. (2012). Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings. *Neurocomputing, 128*, 285–295.

Contents

Notes on Contributors

Christian L. Dunis is a Founding Partner of Acanto Research (www.acantoresearch.com) where he is responsible for global risk and new products. He is also Emeritus Professor of Banking and Finance at Liverpool John Moores University where he directed the Centre for International Banking, Economics and Finance (CIBEF) from February 1999 to August 2011.

Christian Dunis holds an MSc, a Superior Studies Diploma in International Economics and a PhD in Economics from the University of Paris.

Peter W. Middleton completed a Phd at the University of Liverpool. His working experience is in Asset Management and he has published numerous articles on financial forecasting of commodity spreads and equity time-series.

Andreas Karathanasopoulos studied for his MSc and Phd at Liverpool John Moores University under the supervision of Professor Christian Dunis. His working experience is academic having taught at Ulster University, London Metropolitan University and the University of East London. He is currently an Associate Professor at the American University of Beirut and has published more than 30 articles and one book in the area of artificial intelligence.

Konstantinos Theofilatos completed his MSc and Phd in the University of Patras Greece. His research interests include computational intelligence, financial time-series forecasting and trading, bioinformatics, data mining and web technologies. He has so far published 27 publications in scientific peer reviewed journals and he has also published more than 30 articles in conference proceedings.

Part I

Introduction to Artificial Intelligence

1

A Review of Artificially Intelligent Applications in the Financial Domain

Swapnaja Gadre-Patwardhan, Vivek V. Katdare, and Manish R. Joshi

1 Introduction

Undoubtedly, the toughest challenge faced by many researchers and managers in the field of finance is uncertainty. Consequently, such uncertainty introduces an unavoidable risk factor that is an integral part of financial theory. The manifestation of risk not only complicates financial decision making but also creates profitable opportunities for investors who can manage and analyze risk efficiently and effectively. In order to handle the complex nature of the problem an interdisciplinary approach is advocated.

Computational finance is a division of applied computer science that deals with practical problems in finance. It can also be defined as the study of data and algorithms used in finance. This is an interdisciplinary field that combines

S. Gadre-Patwardhan (✉)
Institute of Management and Career Courses, Pune, India
e-mail: swapnaja.gadre24@gmail.com

V.V. Katdare
I.M.R. College, Jalgaon, India
e-mail: vvkatdare@rediffmail.com

M.R. Joshi
School of Computer Sciences, North Maharashtra University, Jalgaon, India
e-mail: joshmanish@gmail.com

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets*,
New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0_1

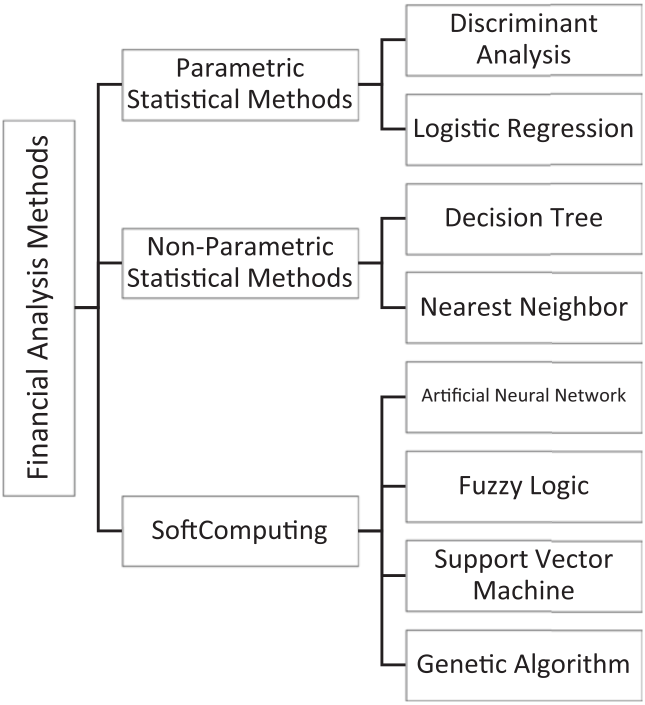


Fig. 1.1 Techniques for analysis of financial applications

numerical methods and mathematical finance. Computational finance uses mathematical proofs that can be applied to economic analyses thus aiding the development of finance models and systems. These models are employed in portfolio management, stock prediction and risk management and play an important role in finance management.

During past few years, researchers have aimed to assist the financial sector through trend prediction, identifying investor behaviour, portfolio management, fraud detection, risk management, bankruptcy, stock prediction, financial goal evaluation, finding regularities in security price movement and so forth. To achieve this, different methods like parametric statistical methods, non-parametric statistical methods and soft computing methods have been used as shown in Fig. 1.1. It is observed that many researchers are exploring and comparing soft computing techniques with parametric statistical techniques and non-parametric statistical techniques. Soft computing techniques, such as, Artificial Neural Network (ANN), Fuzzy Logic, Support Vector Machine (SVM), Genetic Algorithm, are widely applied and accepted techniques in the field of finance and hence are considered in this review.

(A) Parametric statistical methods: Parametric statistics is a division of statistics. It assumes that data is collected from various distributed systems and integrated in order to draw inferences about the parameters of the distribution. There are two types of parametric statistical methods namely discriminant analysis and logistic regression:

(I) Discriminant analysis: Discriminant analysis is a statistical analysis carried out with the help of a discriminant function to assign data to one of two or more naturally occurring groups. Discriminant analysis is used to determine the set of variables for the prediction of category membership. Discriminant function analysis is a type of classification that distributes items of data into classes or groups or categories of the same type.

(II) Logistic regression: Logistic regression is a method of prediction that models the relationship between dependent and independent variables. It the best-fit model to be found and also identifies the significance of relationships between dependent and independent variables. Logistic regression is used to estimate the probability of the occurrence of an event.

(B) Non-parametric statistical methods: These are the methods in which data is not required to fit a normal distribution. The non-parametric method provides a series of alternative statistical methods that require no, or limited, assumptions to be made about the data. The techniques of non-parametric statistical methods follow.

(I) Decision tree: A decision tree is a classifier that is a tree-like graph that supports the decision making process. It is a tool that is employed in multiple variable analyses. A decision tree consists of nodes that a branching-tree shape. All the nodes have only one input. Terminal nodes are referred to as leaves. A node with an outgoing edge is termed a test node or an internal node. In a decision tree, a test node splits the instance space into two or more sub-spaces according to the discrete function.

(II) Nearest neighbour: The nearest neighbour algorithm is a non-parametric method applied for regression and classification. Nearest neighbour can also be referred as a similar search, proximity search or closest-point search, which is used to find the nearest or closest points in the feature space. The K-nearest neighbour algorithm is a technique used for classification and regression.

(C) Soft computing: Soft computing is a set of methods that aims to handle uncertainty, partial truth, imprecision and approximation that are fundamentally are based on human neurology. Soft computing employs techniques like: ANN, fuzzy logic, SVM, genetic algorithm [1].

(I) Artificial neural network: A neuron is a fundamental element of ANN. These neurons are connected to form a graph-like structure, which are also referred to as networks. These neurons are like biological neurons. A neuron has small branches, that is, dendrites, which are used for receiving inputs. Axons carry the output and connect to another neuron. Every neuron carries a signal received from dendrites as shown in Fig. 1.2 [2]. When the strength

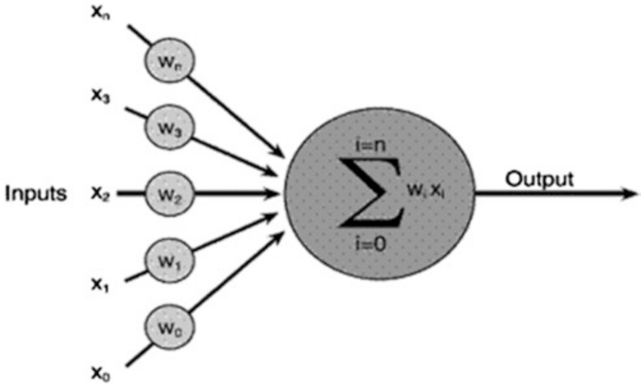


Fig. 1.2 Structure of Artificial Neurons

of a signal exceeds a particular threshold value, an impulse is generated as an output, this is known as the action signal.

Like biological neurons, artificial neurons accept input and generate output but are not able to model automatically. In ANN information or data is distributed and stored throughout the network in the form of weighted interconnections. Simulation of a neuron is carried out with the help of non-linear function. Interconnections of artificial neurons are referred as weights. The diagram below shows the structure of an artificial neuron in which xi is the input to the neuron and wi is the weight of the neuron. The average input is calculated by the formula [2].

$$a = \sum_{i=0}^{n} xiwi \tag{1.1}$$

ANN has a minimum of three layers of artificial neurons: input, hidden and output as shown in Fig. 1.3 [3]. The input layer accepts the input and passes it to the hidden layer. The hidden layer is the most important layer from a computational point of view. All the complex functions reside in this layer.

(II) Fuzzy logic: Fuzzy logic is a type of many values logic that deals with approximate values instead of exact or fixed reasoning. Fuzzy logic is a method of computing based on the degree of truth rather than a crisp true or false value. Its truth value ranges in between 0 and 1.

(III) Support vector machine: SVM is a supervised learning model with related learning algorithms that is used for data analysis and pattern recognition in classification and regression. SVM uses the concept of a hyperplane,

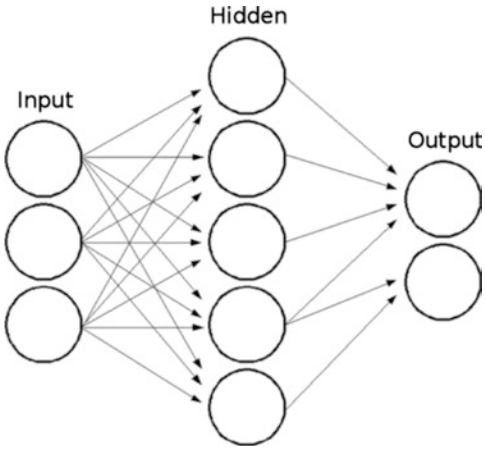


Fig. 1.3 Three layer architecture of ANN

which defines the boundaries of a decision. The decision plane separates the objects based on class membership and is able to handle categorical and continuous variables.

(IV) Genetic algorithm: A genetic algorithm is an artificial intelligence technique that mimics a natural selection process. This technique is mostly used for optimization and search problems using selection, crossover, mutation and inheritance operations.

This chapter emphasizes the application of soft computing techniques namely artificial neural network, expert system (ES) and hybrid intelligence system (HIS) in finance management.

In recent years, it has been observed that an array of computer technologies is being used in the field of finance; ANN is one of these. From the array of available AI techniques, financial uncertainties are handled in a more efficient manner by ANN. These uncertainties are handled by pattern recognition and future trend analysis. The most difficult aspects to incorporate in finance analysis are changes in the interest rates and currency movements. Large 'noisy' data can be handled well by ANN. ANN are characterized as numeric in nature. In statistical techniques, like discriminant analysis or regression analysis, data distribution assumptions are required for input data. However, ANN does not require any data distribution assumptions and hence could be applicable to a wider range of problems than other statistical techniques. Statistical techniques and symbolic manipulation techniques are batch oriented; old and new data are submitted in a single batch to the model and later new mining results are generated. In contrast, in ANN it is possible to add new data to a trained ANN so as to update the existing result. Since financial markets are dynamic in nature, ANN can accommodate new data without reprocessing old data and hence it is used in finance management [4].

An ES is knowledge-based system used to solve critical problems in a particular domain. These are rule-based systems with predefined sets of knowledge used for decision making. Generic ES contain two modules—the inference engine and the knowledge base. The inference engine combines and processes the facts associated with the specific problem using the chunk of the knowledge base relevant to it. The knowledge base is coded in the form of rules, semantic nets, predicates and objects in the system. ES are characterized as efficient, permanent, consistent, timely, complete decision-making systems and hence their use in finance management. ES are characterized as intelligent, capable of reasoning, able to draw conclusions from relationships, capable of dealing with uncertainties and so forth. ES are capable of reproducing efficient, consistent and timely information so as to facilitate decision making [5]. Furthermore Rich and Knight (1991) specified long ago that financial analysis is an expert's task.

HIS are software systems that combine methods and techniques of artificial intelligence, for example, fuzzy expert systems, neuro-fuzzy systems, genetic-fuzzy systems. The integration of various learning techniques is combined to overcome the limitation of an individual system. Because of its facility of combined techniques, it can be used effectively for finance management.

With reference to the financial market, we identified portfolio management, stock market prediction and risk management as the three most important AI application domains. As investment is an important aspect of finance management hence these three cases are considered. In this study, we consider the contribution of researchers in financial domains from the past 20 years in order to study and compare the applications of ANN, ES and HIS with traditional methods. The chapter is organized thus: the second, third and fourth sections deal with the application of ANN, ES and HIS respectively. In the fifth section conclusions are put forth. We enlist popularly used data mining tools as set out in Appendix 1 that includes some sample coding of NN techniques using MATLAB [6] in Finance Management. Code excerpts for implementing typical statistical functions including regression analysis, naïve Bayes classification, fuzzy c-means clustering extracted from different openly available authentic sources [7] are also presented in Appendix 1.

Applications of ANN in Finance

ANN are computational tools and are used in various disciplines for modelling real-world complex problem [8]. ANN resemble biological neurons acting as a source inspiration for a variety of techniques covering a vast field of application [9]. In general, ANN are referred to as information processing systems that which use earning and generalization capabilities, which are adaptive in nature. Due to their adaptive nature, ANN can provide solutions to problems such as forecasting, decision making and information processing. In recent years, ANN have proved to be a powerful tool for handling dynamic financial market in terms of prediction [10], panning, forecasting [11] and decision making [12].

With reference to this various studies have been carried out in order to classify and review the application of ANN in the finance domain [13, 14]. Mixed results have been obtained concerning the ability of ANN in finance domain. It has been observed that financial classification like financial evaluation, portfolio management, credit evaluation and prediction are significantly improved with the application of ANN in the finance domain. We further consider the application of ANN in the finance domain in portfolio management, stock market prediction and risk management. The details of these applications are presented as described previously.

Portfolio Management

The determination of the optimal allocation of assets into broad categories, for example, mutual funds, bonds, stocks, which suits investment by financial institutions across a specific time with an acceptable risk tolerance is a crucial task. Nowadays investors prefer diversified portfolios that contain a variety of securities.

Motiwalla et al. [15] applied ANN and regression analysis to study the predictable variations in US stock returns and concluded that ANN models are better than regression. Yamamoto et al. [16] designed a multi-layer Back Propagation Neural Network (BPNN) for the prediction of the prepayment rate of a mortgage with the help of a correlation learning algorithm. Lowe et al. [17] developed an analogue Neural Network (NN) for the construction of portfolio under specified constraints. They also developed a feed forward NN for prediction of short-term equities in non-linear multi-channel time-series forecasting. Adedeji et al. [18] applied ANN for the analysis of risky economic projects. For the prediction of the potential returns on investment, an NN model could be used. On the basis of results obtained from the neural network, financial managers could select the financial project by comparing the results to those obtained from conventional models. The survey conducted in this paper for portfolio management concludes that ANN performs better in terms of accuracy. Without any time consuming and expensive simulation experiments, accuracy can be obtained by combining conventional simulation experiments with a neural network.

Research papers surveyed for portfolio management demonstrates that when compared to other traditional methods, ANN performs better particularly BPNN. Zimmermann et al. [19] demonstrated the application of the Back/Litterman portfolio optimization algorithm with the help of an error correction NN. Optimization of the portfolio includes (1) allocation that comply investors constraints and (2) controlled risk in the portfolio. The method was tested with internationally diversified portfolios across 21 financial markets from G7 countries. They stated that their approach surpassed conventional portfolio forecasts like Markowitz's mean-variance framework. Ellis et al. [20] performed a portfolio analysis by comparing BPNN with a randomly selected portfolio method and a general property method concluding that ANN performs better.

Stock Market Prediction

In recent years with the help of online trading, the stock market is one of the avenues where individual investors can earn sizeable profits. Hence there is a need to predict stock market behaviour accurately. With this prediction investors can take decisions about where and when to invest. Because of the volatility of financial market building a forecasting model is a challenging task.

ANN are a widely used soft computing method for stock market prediction and forecasting. White applied ANN on IBM daily stock returns and concluded that the NN outperformed other methods [21]. Kimoto et al. [22] reported the effectiveness of learning algorithms and prediction methods of Modular Neural Networks (MNN) for the Tokyo Stock Exchange price index prediction system. Kazuhiro et al. [23] investigated the application of prior knowledge and neural networks for the improvement of prediction ability. Prediction of daily stock prices was considered a real-world problem. They considered some non-numerical features such as political and international events, as well as a variety of prior knowledge that was difficult to incorporate into a network structure (the prior knowledge included stock prices and information about foreign and domestic events published in newspapers.) It was observed that event knowledge combined with an NN was more effective for prediction with a significance level of 5 %. Pai et al. [24] stated that ARIMA (autoregressive integrated moving average) along with SVM can be combined to deal with non-linear data. The unique strengths of ARIMA and SVM are used for more reliable stock-price forecasting. Thawornwong et al. [25] demonstrated that the NN model with feed-forward and probabilistic network for the prediction of stock generated high profits with low risk. Nakayama et al. [26] proposed a Fuzzy Neural Network (FNN) that contained a specific structure for realizing a fuzzy inference system. Every membership function consists of one or two sigmoid functions for inference rule. They concluded that their FNN performed better. Duke et al. [27] used Back Propagation Network (BPN) for the prediction of the performance of the German government's bonds

Risk Management

Financial risk management (FRM) is the process of managing economic value in a firm with the help of financial instruments to manage risk exposure especially market risk and credit risk. Financial Risk Management (FRM) is the process of identification of risk associated with the investments and possibly mitigating them. FRM can be qualitative or quantitative. FRM focuses on how and when hedging is to be done with the help of financial instruments to manage exposure to risk.

Treacy et al. [28] stated that the traditional approach of banks for credit risk assessment is to generate an internal rating that considers subjective as well as qualitative factors such as earning, leverage, reputation. Zhang et al. [29] compared Logistic Regression (LR), NN and five-fold cross validation procedures on the database of manufacturing firms. They employed Altman's five functional ratios along with the ratio current assets/current liabilities as an input to NN. They concluded that NN outperforms with accuracy 88.2 %. Tam et al. [30] introduced an NN approach to implement discriminant analysis in business research. Using bank data, linear classification is compared with a neural approach. Empirical results concluded that the neural model is more promising for the evaluation of bank condition in terms of adaptability, robustness and predictive accuracy. Huang et al. [31] introduced an SVM to build a model with a better explanatory ability. They used BPNN as a benchmark and obtained around 80 % prediction accuracy for both SVM and BPNN for Taiwan and United States markets.

Table 1.1 provides details of the literature that considers the application of ANN for portfolio management, stock market prediction and risk management.

2 Application of Expert Systems in Finance

An expert system is a computer system that is composed of a well-organized body of knowledge that emulates expert problem-solving abilities in a limited domain of expertise. Matsatsinis et al. [54] presented a methodology of acquisition of knowledge and representation of knowledge for the development of ES for financial analysis. Development of FINEVA (FINancial EVAluation), a multi-criterion knowledge base DSS (decision support software) for assessment of viability and corporate performance and the application of FINEVA was discussed. For a particular domain, a set of inference rules are provided by a human expert. The knowledge base is a collection of relevant facts, data, outcome and judgments [34]. Components of expert systems include the knowledge base, the user interface and the inference engine. Knowledge is represented through the techniques such as, predicate logic, frames and semantic nets but the most popular and widely used technique is the IF-THEN rule also referred as the production rule.

Liao et al. [55] carried out a review of the use of an ES in a variety of areas including finance during period 1995 to 2004. They observed that ES are flexible and provide a powerful method for solving a variety of problems, which can be used as and when required. Examples of the application of ES in finance domain follow.

Portfolio Management

It is a difficult and time-consuming task to explore and analyze a portfolio in relation to the requirements and objectives of the fund manager. Ellis et al. [34] examined the application of rule-based ES in the property market and portfolios randomly constructed from the market. They observed that rule-based outperform the random portfolio or market on risk adjusted return basis.

Bohanec et al. [56] developed a knowledge-based tool for portfolio analysis for evaluation of a project. This ES was developed for the Republic of Solvenia. The model is demonstrated with a tree structure supplemented by IF-THEN rules. Sanja Vraneš et al. [57] developed the Blackboard-based Expert Systems Toolkit (BEST) for combining knowledge from different sources, using different methodologies for knowledge acquisition. As far as investment decision making is concerned, information from proficient economist critical investment ranking might be combined with knowledge evolved from operational research methods. When decisions are made based on information combined from many sources, there is a probability of redundancy reduction and more promising results. Varnes et al. [58] suggested INVEX (investment advisory expert system) for investment management. This system assists investors and project analysts to select a project for investment. Mogharreban et al. [59] developed the PROSEL (PORtfolio SELection) system that uses a set of rules for stock selection. PROSEL consists of three parts (1) an information centre

Table 1.1 A Brief Review of ANNs Applied to Portfolio Management, Stock Market Prediction and Risk Management

Author	Objective	Data Set	Preprocessing	Approach used	Compared with	Evaluation metrics
Stoppiglia H, Idan Y, Dreyfus G [32]	To develop a neural network -aided model for portfolio management	The database comprises 398 companies, with 172 A companies, 172 B companies,& 54 C companies	Fifteen financial ratios such as, working capital/ fixedassets, profit after taxes and interest/net worth, per year	ANN	Statistical Method	Classification
Hans Georg Zimmermann, Ralph Neuneier and Ralph Grothmann, Siemens AG [33]	Portfolio Optimization	Financial markets of the G7 countries	Monthly data extracted from all databases	ANN	Mean-variance theory	Forecasting
Ellis C, Willson P [34]	To select portfolios	Australian property sector stocks	Not mentioned	BPNN	Random selection portfolio	Performance measure
Fernandez A, Gomez S [35]	Portfolio selection and portfolio management	Hang Seng in Hong Kong, DAX 100 in Germany, FTSE 100 in UK, S&P 100 in USA and Nikkei 225 in Japan	Weekly prices from data sets are extracted	ANN, GA and SA	Heuristic methods	Portfolio selection and optimization
Freitas FD, De Souza AF, De Almeida AR [36]	Portfolio selection and Portfolio optimization	IBOVESPA	Selected a subset of 52 stocks with long enough time-series for training the neural networks	BPNN	Mean-variance model	Prediction

(*continued*)

Table 1.1 (continued)

Author	Objective	Data Set	Preprocessing	Approach used	Compared with	Evaluation metrics
Po-Chang Ko, Ping-Chen Lin [37]	Portfolio selection and portfolio optimization	Taiwan stock exchange	Not mentioned	ANN	Traditional ANN model	Portfolio optimization
Chiang W-C, Urban TL, Baldridge GW [38]	Asset forecasting	US mutual fund	15 economic variables are identified	BPNN	Regression model	NAV prediction
Chen A-S, Leung MT, Daouk H [39]	Stock index forecasting	Taiwan StockExchange	Data is extracted on the basis of length of investment horizon	PNN	Random walk model and the parametric GMM models	Return on investment
O'Connor N, Madden MG [40]	To predict stock market movement forecasting	New York Stock Exchange and NASDAQ	Daily opening and closing values of DJIA index	BPNN	Simple benchmark functions	Accuracy
De Faria EL, Marcelo P, Albuquerque J, Gonzalez L, Cavalcante JTP, Marcio Albuquerque P [41]	To predict stock market movement forecasting	Brazilian stock market	Not mentioned	BPNN	Adaptive exponential smoothing method	Accuracy
Liao A, Wang J [42]	Stock index forecasting	SP500, SAI, SBI, DJI, HIS and IXIC	Data normalization and adjusted to remove the noise	BP stochastic time effective NN	Brownian motion	Forecasting

Hyun-jung Kim, Kyung-shik Shin [43]	To detect patterns in stock market	Korea StockPrice Index 200	Daily stock data is extracted	ATNN	TDNN	Accuracy
Kuo RJ, Chen CH, Hwang YC [44]	Stock market forecasting	Taiwan stock market	Not mentioned	ANN, GFNN	Qualitative and quantitative factors of NN	Performance evaluation
Md. Rafiul Hassan, Baikunth Nath, Michael Kirley [45]	Stock market forecasting	www.finance.yahoo.com	Daily data is extracted	ANN, GA, HMM	ARIMA model	Forecasting
Chye KH, Tan WC, Goh GP [46]	Credit risk assessment	Australian and German credit data sets	For each applicant 24 variables are selected	SVM classifier	Neural networks, genetic programming, and decision tree classifiers	Accuracy
Eliana Angelini, Giacomo di Tollo, Andrea Roli [47]	Credit risk evaluation	Bankin Italy	Sample group is categorized into two groups i.e. boins and default	ANN	classical feed forward neural network and special purpose feed forward architecture	Classification
Fanning, Cogger Shrivastava [48]	To develop a model using neural network to find managerial fraud	Management database	Not mentioned	ANN	Generalized adaptive neural network architectures (GANNA) and the Adaptive Logic Network (ALN)	Accuracy

(*continued*)

Table 1.1 (continued)

Author	Objective	Data Set	Preprocessing	Approach used	Compared with	Evaluation metrics
Feroz EH, Taek MK, Pastena VS, and Park K [49]	To test the ability of selected red flags for prediction of the targets of the investigations.	AAER	Not mentioned	ANN	Investigated versus non-investigated forms	Prediction
Kurt Fanning, Kenneth Cogger O [50]	To develop a model for detection of management fraud	FFS.	Not mentioned	ANN	Statistical methods	Efficiency
Brause R, Langsdorf T, Hepp M	To detect credit card fraud	GZS	38 field per transaction are extracted	ANN	Traditional systems	Prediction
Koskivaara [51]	To investigate the impact of various preprocessing models on the forecast capability of neural network for financial auditing.	Manufacturing firm	Monthly balances	ANN	Traditional systems	Prediction

Fen-May Liou [52]	To detect fraudulent financial reporting	Taiwan Economic Journal data bank and Taiwan Stock Exchange	Not mentioned	ANN	Logistic regression, neural networks, and classification trees.	Prediction
Chu, Jung [53]	To compare statistical methods and ANN in bankruptcy prediction.	International Stock Exchange Official Year Book from a Data stream database	Data distribution, group dispersion and orientation scheme	ANN with MDA	Logit, generalized Delta rule	Prediction
Yang et al. [29]	Bankruptcy prediction	COMPUSTAT database	Excluded from the sample are those that: (1) have operated in aregulated industry; (2) are foreign based and traded publicly in the USA; and (3) have previously been bankrupt.	ANN	MDA, ID3	Classification
Tam KY, Kiang MY [30]	To perform discriminant analysis in business research	Bank default data	Not mentioned	Neural Network	Linear classifier, logistic regression, k NN, and ID3	Predictive accuracy, adaptability

(*continued*)

Table 1.1 (continued)

Author	Objective	Data Set	Preprocessing	Approach used	Compared with	Evaluation metrics
Patrícia Xufre Casqueiro, António JL Rodrigues [112]	Portfolio Management	Trading Data	Not mentioned	NN, Q-Learning Algorithm	Various models of Neurocomputing	Forecasting, Risk optimization
Young-Woo Seo Joseph Giampapa Katia Sycara [113]	Portfolio Management	News articles	News articles were manually labeled into 5 classes	Classification and clustering	NaiveBayes Classification	Accuracy
Wang, Zhuowen [114]	Stock Market Prediction	Shanghai Stock Exchange Composite Index, and the Shenzhen Stock Exchange Component Index	Not mentioned	Revised Back Propagation	Correlation between the stock price	Prediction

(2) a fuzzy stock selector and (3) a portfolio constructor. A user-friendly interface is available in PROSEL to change rules at run time. Mogharreban et al. identified that PROSEL performed well.

Stock Market Prediction

One more promising area for ES is in stock market prediction. Many investment consultants use these types of systems to improve financial and trading activities. Midland Bank of London use an ES for interest rate swap, portfolios and currency management [34].

Grosan et al. [60] applied MEP (multi-expression programming), a genetic programming technique for prediction of the NASDAQ index of the NASDAQ stock market and the NIFTY stock index. The performance is compared with the help of an SVM, a Levenberg-Marquardt algorithm and a Takagi–Sugenoneuro-fuzzy inference system. They concluded that MEP performs outstandingly. Quek [61] applied neuro-fuzzy networks and ANFIS investor's measures forecasting to the US Stock Exchange where it proved best for stock price prediction. Trinkle [62] used ANFIS (adaptive network-based fuzzy inference system) and an NN for forecasting the annual excess return of three companies. The predictive ability of ANFIS and NN is compared with ARMA (autoregressive moving average). The result stated that ANFIS and NN are able to forecast significantly better. Afolabi et al. [63] used a neuro-fuzzy network, fuzzy logic and Kohonen's self-organizing plan for stock price forecasting. They concluded that, compared to other techniques, deviation of Kohonen's self-organizing plan is less. Yunos et al. [64] built a hybrid neuro-fuzzy model with the help of ANFIS to predict daily movements in the KLCI (Kuala Lumpur Composite Index). For data analysis four technical indicators were chosen. The conclusion showed that ANFIS performed better. Atsalakis et al. [65] developed a neuro-fuzzy adaptive control system for forecasting the price trends of stock for the following day of the NYSE and ASE index. The experimental analysis stated that the system performed well.

Risk Management

There is a vast potential in using ES in financial risk prediction and management. Matsatsinis et al. [54] presented a methodology for acquisition and representation of knowledge for the development of an ES. FINEVA is a multi-criteria knowledge-based ES for the assessment of viability and performance using an M4 ES shell. The interface uses forward and backward chaining method. Matsatsinis et al. concluded that the ranking of analyzed firms depends upon the class of risk.

Shue et al. [66] built an ES for financial rating of corporate companies. This ES was developed by integrating two knowledge bases (a) Portege—a domain knowledge base and (b) JES–an operational knowledge base. The model is tested and verified by inputting data from financial statements of various companies listed on the Taiwan stock market. Luke et al. [67] presented an ES, CEEES (credit evaluation and explanation ES) to take decisions about whether to allow credit lines to identified firms. CEEES used rule-based language for the decision-making process. They concluded that CEEES will recommend whether to consider or reject the application of credit.

Table 1.2 provides details of the literature that considers the application of ES for portfolio management, stock market prediction and risk management.

3 Applications of Hybrid Intelligence in Finance

HIS is a software system that is formed by combining methods and techniques of artificial intelligence, that is, a fuzzy expert system, a neuro-fuzzy system, a genetic-fuzzy system, for example. HIS systems are an effective learning system that combines the positive features and overcomes the weaknesses of the processing capabilities and representations of learning paradigms. HIS are used for problem solving in various domains [73]. Lertpalangsunti [74] proposed three reasons for creating HIS: (a) technique enhancement, (b) multiplicity of application task and (c) realizing multi-functionality. The degree of integration between the modules may vary from loosely coupled standalone modules to fully coupled. The application of HIS in the finance domain follows.

Portfolio Management

Portfolio management is a complex activity that involves a crucial decision-making process. It is an important activity of many financial institutes and organizations. In the past few years HIS has become widely applied in portfolio selection [75].

Kosaka et al. [76] applied NN and Fuzzy logic for stock portfolio selection. They concluded that the proposed model identified price tuning points with 65 % accuracy. Chen et al. [77] developed a portfolio-selection model.

Table 1.2 A Brief Review of ES Applications to Portfolio Management, Stock Market Prediction and Risk Management

Author	Objective	Data Set	Preprocessing	Approach used	Compared with	Evaluation metrics
Shaw M, Gentry J [68]	To manage business loan portfolios	Commercial Bank Data	Not mentioned	MARBLE	Conventional Method	Decision Rules, Classification
Bohanec M, Rajkovič V, Semolič B, Pogačnik A [56]	Portfolio selection and portfolio management	Ministry of Science and Technology of the Republic of Slovenia	Portfolio attribute selection and evaluation	Expert system shell: DEX	Expert reviews	Rule generation
Ellis C, Willson P [34]	To develop an expert system for portfolio management in property investment	Australian Stock Exchange (ASX), London Stock Exchange (LSE)	Not mentioned	Expert system	Conventional method	Rule generation
Kim SH, Chun SH [69]	Stock index forecasting		Not mentioned	BPNN	CBR, APN, PNN	Prediction
Wang Y-F [70]	Stock price prediction	Taiwan stock market	Every 5 min. data is recorded	Fuzzy grey prediction	GA	Prediction

(*continued*)

Table 1.2 (continued)

Author	Objective	Data Set	Preprocessing	Approach used	Compared with	Evaluation metrics
Quintana D et al. [71]	Bankruptcy prediction	552 US companies from COMPUSTAT database	Not mentioned	ANN with ENPC	Naive Bayes (NB), logistic regression (LR),C4.5, PART, SVM, multilayer perceptron (MLP) and ENPC	Classification
Lenard MJ, Alam P [72]	To develop a fuzzy logic model for development of clusters to represent red flags in detection of fraud	Not mentioned	Not mentioned	Fuzzy logic and expert reasoning	Statistical methods	Clustering

In the proposed model triangular fuzzy numbers are used to represent future return rates and risks of mutual funds. Quek et al. [78] developed a fuzzy-neural system for portfolio balancing with the help of GenSoFNN (Generic self-organizing fuzzy NN). They applied supervised learning methods in the network for detection of inflection points in the stock price cycle. Yu et al. [79] developed an NN-dependent mean-variance skewness model for portfolio section on the basis of the integration of an RBF (radial basis function) and a Lagrange multiplier theory of optimization. Li et al. [80] proposed a hybrid intelligent algorithm by assimilating NN, simulated annealing algorithm and fuzzy simulation techniques for solving portfolio selection problems. In the proposed model, NN is used for the approximation of expected value and variance of fuzzy returns. Fuzzy simulation generates the training data for NN. Their model and genetic algorithms are also compared. Quah et al. [81] compared the performance of ANFIS, MLP-NN and GGAP-RBF (general growing pruning radial basis function). Quah et al. also proposed the method of selection of equities through the use of a ROC (relative operating characteristics) curve.

Stock Market Prediction

The volatile nature of stock market requires a variety of computing techniques. As compared to other domains, hybrid AI systems are widely used for financial prediction because hybrid systems are able to combine the capabilities of various systems with their unique abilities.

Kuo et al. [82] developed a system for stock market forecasting. The proposed model deals with qualitative and quantitative factors simultaneously. The system was developed by integrating a fuzzy Delphi model with an NN for qualitative and quantitative factors respectively. The system was tested on the database of Taiwan Stock Market and found considerably better than. Romahi et al. [83] proposed a rule-based ES for financial forecasting. They combined rule induction and fuzzy logic and observed that their system performed better. Keles et al. [84] developed a model for forecasting domestic debt (MFDD). They applied ANFIS to few microeconomic variables of Turkish economy. They observed that the MFDD performed better in terms of forecasting. Huang et al. [85] combined an average autoregressive exogenous (ARX) model for prediction with grey system theory and a rough set to forecast the stock market automatically of the Taiwan stock exchange. They employed a GM (1,N) model for data reduction. After data reduction, clusters are formed by using K-means algorithm and later supplied to rough set classification module. Set of suitable stocks is selected by applying some decision rules. The results are then compared with GM(1,1). They observed that the hybrid method has greater forecasting ability for the selected stock.

Risk Management

Risk management is a decision-making activity that involves social, political, engineering and economic factors. Risk could arise in the form of fraud, bankruptcy and so forth. Elmer et al. [86] proposed a hybrid fuzzy logic and neural network algorithm for credit risk management. An HFNN (hybrid fuzzy logic-neural network) model is used for credit risk evaluation. Dadios and Solis conclude that the HFNN model is robust, accurate and reliable [87]. Lean et al. [88] proposed hybrid intelligence system for credit risk evaluation and analysis using a rough set (RS) and SVM. SVMs are used to extract features and for noise filtration. RS acted as a preprocessor for the SVM. Lean et al. concluded that the proposed model performed better.

Hyunchul et al. [89] focused on the important issue of corporate bankruptcy prediction. Various data driven approaches are applied to enhance prediction performance using statistical and AI techniques. Case based reasoning (CBR) is the most widely used data-driven approach. The model is developed by combining CBR with a genetic algorithm (Gas). It was observed that the model generates accurate results along with reasonable explanations. Zopounidis et al. [90] presented a review on the application of a knowledge base decision support system (KBDSS) in finance and management. KBDSS is developed by combining the features of an ES and DSS in many fields, for example, financial analysis, bankruptcy risk assessment and financial planning. Zopounidis et al. [89] described KBDSS for portfolio management, financial analysis and credit gaining problems. They observed that a KBDSS improvises the decision-making process by explaining the operations and the results generated by the system. Hua et al. [91] applied SVM for bankruptcy prediction and it proved competitive against neural network, logistic regression and linear multiple discriminant analysis. Hua et al. [90] developed an integrated binary discriminant rule (IBDR) for financial distress prediction. The experimental results proved that IDBR performs better when compared to the conventional SVM.

Table 1.3 provides details of the literature that considers the application of HIS for portfolio management, stock market prediction and risk management.

Table 1.3 A Brief Review of HIS Applications to Portfolio Management, Stock Market Prediction and Risk Management

Author	Objective	Data Set	Preprocessing	Approach used	Compared with	Evaluation metric
Li X, Zhang Y, Wong H-S, Qin Z [80]	To develop a model for portfolio selection	Data from B. Liu, *Theory and Practice of Uncertain Programming* Physica-Verlag, Heidelberg, 2002	Not mentioned	ANN, GA and Fuzzy simulation	GA	Robust and effective
Chen Y, Ohkawa E, Mabu S, Shimada K, Hirasaa K [92]	To develop a model for portfolio optimization	Japanese stock market	Not mentioned	GNP with controlled nodes, technical analysis rules	GA and B&H method, Conventional GNP-based methods	Effectiveness
Shian-Chang Huang [93]	Stock indices Forecasting	NASDAQ (US), NK225 Japan),TWSI (Taiwan) and KOSPI (South Korea).	Stock indices are transformedinto daily returns	Hybrid model	Traditional time domain models	Forecasting
Wang Y-F [70]	Stock price prediction	Taiwan stock market	Every 5 min. data is recorded	Fuzzy grey prediction	GA	Prediction
Lin S-H [94]	Credit risk assessment	Taiwan banks	Not mentioned	LR, ANN	ANN, logarithm LR and LR	Prediction
Arminger G, Enache D, Bonne T [95]	To analyze credit risk		Data is extracted on the basis some factors like age, sex, job etc.	BPNN,LDA	Discriminant Techniques	Classification

(*continued*)

Table 1.3 (continued)

Author	Objective	Data Set	Preprocessing	Approach used	Compared with	Evaluation metric
Kun Chang Lee, Ingoo Han, Youngsig Kwon [96]	To develop the hybrid neural network models for bankruptcy prediction	Korea Stock Exchange	Divided data set into two categories i.e. Failed firms and non-failed firms.	HNN	MDA, ID3, SOFM	Accuracy and adaptability
Garcia-Almanza, Alma Lilia, Biliana Alexandrova-Kabadjova, and Serafin Martinez-Jaramillo [97]	Bankruptcy prediction for banks	FDIC Data	Not mentioned	MP-EDR	Machine learning	Prediction
Kyung-Shik Shin, Yong-Joo Lee [98]	To predict corporate failure		528 externally audited mid-sized manufacturing firms.	GA	Traditional statistical methods	Rule extraction
Hsueh-Ju Chen, Shaio Yan Huang, Chin-Shien Lin [99]	To develop a model for bankruptcy prediction	Securities and Exchange Commission (SEC)	Not mentioned	Fuzzy logic	Traditional methods	Accuracy
Sung-Hwan Mina, Jumin Leeb, Ingoo Hanb [100]	Bankruptcy prediction	Commercial bank in Korea	Financial ratios categorized as stability, profitability, growth, activity and cash flow	SVM-based model	neural network (NN) and logistic regression	Feature subset selection and parameter optimization

Lenard, Watkins and Alam [101]	To detect financial statement frauds	Retail and manufacturing industries	Not mentioned	Fuzzy logic	External parties	Accuracy
Deshmukh, Romine and Siegel [102]	To provide a fuzzy sets model to assess the risk of managerial fraud	Business accounts	Not mentioned	Fuzzy logic	Traditional systems	Assessment
Juszczak et al. [103]	To detect fraud in financial statement	D1 and D2	Not mentioned	Supervised and semi-supervised classification	Supervised classification versus unsupervised classification	Classification
Lenard, Watkins and Alam [101]	To develop a fuzzy logic model for development of clusters to represent red flags in detection of fraud	Not mentioned	Not mentioned	Fuzzy logic & expert reasoning	Statistical methods	Clustering
Jerry W Lin, Mark I Hwang, Jack D Becker [115]	Risk Management	publicly traded firms that were charged by the US Securities and Exchange Commission (SEC) in the period from 1980 to 1995	Classification of firms.	Fuzzy Neural Network	ANN, Statistical Models	Prediction Accuracy by 97.5%

4 Conclusion

A comprehensive review is conducted on applications of AI in finance management. The review is organized by considering the type of techniques, their application domain, objectives and evaluation metrics. The review indicates that the traditional approach is not sufficient enough to tackle and analyze huge quantities of financial data. Hence, contemporary methods need to be applied.

AI techniques are an important dimension of contemporary research. An important conclusion that is drawn from this research is that researchers employ various AI techniques to solve, effectively, the problems associated with finance management. Comparative studies illustrate that ANN are successful in financial prediction. However, it is quite difficult to determine the required structure and size of a neural network to solve the given problem. The major difficulty with ANN is that they are trained using past data, which may not be repeated. An alternative method for this could be an ES as they generate predictions. The problem of ES is that they do not learn through experience and are unable to handle non-linear data. To overcome these problems hybrid intelligent systems, which are able to handle linear and non-linear data, could be implemented. HIS can combine the capabilities of various systems to overcome the limitations of individual techniques. It is observed that limited literature is available on ES and HIS in finance domain.

Computational finance is a blending of computational power and machine-learning techniques to cope with problems of practical interest in the financial domain. This is an application-oriented study that proposes innovative techniques to solve financial domain problems. Although we focus on portfolio management, stock market prediction and risk management related problems practically all types of financial problems can be addressed with computerized techniques and with AI techniques in particular. AI techniques are also used for building models of financial markets. Novel approaches in the field of research could have a fusion of different techniques [45]. A number of researchers in various research institutes are working in this area, for example, those based at the University of Essex, Illinois Institute of Technology, University of Washington.

The objective of this chapter was to provide introduction of the field of computational finance and how AI techniques are being used to deal with problems of practical interest in the financial domain. Some of the difficulties in the field of portfolio management, stock market prediction and risk management include resource balancing and making decisions in the absence of major information. These problems could be resolved using preliminary and detailed investigation, source capacity analysis, integrate portfolio management and many more [104, 105]. Furthermore, we wish to elaborate how researchers are exploring the strengths of certain AI techniques to overcome the problems that mere statistical techniques cannot deal with. We feel that this review could be a helpful guideline to study various AI techniques and we hope that the researchers will attempt to overcome the drawbacks of some techniques to try to develop an influential integrated system by utilizing strengths and complementary features of the different techniques.

5 Appendix 1

Frequently used data-mining tools are listed below:

Weka is a group of machine-learning algorithms for data- mining techniques. These algorithms can be applied directly or indirectly to a dataset. Weka comprises of tools for the preprocessing of data, regression, classification, association rule, clustering and visualization. It is open-source software [106].

Scilab is open source software that can be effectively used for scientific and engineering applications. Scilab contains a number of mathematical functions, optimization, 2D and 3D visualization, statistics and so forth. Scilab is available for the platforms like Mac OS, GNU/Linux Windows vista/XP/7/8 [107].

R is free software for graphics and statistical computing. It runs on a wide range of platforms like Windows, Unix and Mac OS [108].

SPMF is a specialized open-source software in pattern mining developed in Java. SPMF implements 109 data mining algorithms including sequential rule mining, association rule mining, sequence prediction, item-set mining and clustering [109].

A sample code for some data mining techniques follows.

Regression Analysis [7]

Regression is the process of fitting models to the data available. Following is an example of linear regression model using MATLAB.

```
% init params
N = 1000; % sample size
muX = 12; % params explanatory variable
sigmaX = 2.3;
coeff = 0.8; % regression coefficient
```

```
intcept = 4.3; % regression intercept
% simulate explanatory variable
X = normrnd(muX,sigmaX,N,1);
% simulate standard normally distributed innovations
epsilon = randn(N,1);
% calculate Y according to linear model
Y = intcept + coeff*X + epsilon; % do not use for loop
```
Parameters are estimated on the basis of values simulated.
```
% because of intercept, expand matrix of explanatory variables
X = [ones(N,1) X];
% OLS estimation, naive way
paramsHat = inv(X'*X)*X'*Y; % usual estimation formula
% avoiding single matrix inversion as mlint warning suggests
paramsHat2 = (X'*X)\(X'*Y); % faster way
paramsHat3 = X\Y; % best way
% calculate regression line
xLimits = [floor(min(X(:,2))) ceil(max(X(:,2)))]; % use nearest
% neighbouring integer numbers
grid = xLimits(1):0.1:xLimits(2);
vals = paramsHat(1)+paramsHat(2)*grid;
% plotting data
close
scatter(X(:,2),Y,'.'); % used for visualizing points cloud
% include regression line
hold on; % plot in same figure
plot(grid,vals,'LineWidth',2,'Color','r') % larger line width
set(gca,'xLim',xLimits)
xlabel('regressor variable')
ylabel('dependent variable')
title(['Linear model: estimated beta is ' num2str(paramsHat(2))])
```

Classification [7]

Classification can be completed with the help of naïve Bayes classifiers, discriminant analysis and decision trees. Data sets comprising observations along with measurements of various variables, which can also be referred to as predictors and their predetermined class labels. Predictors' class can be identified with the help of new observations obtained. Classification could be done with the help of following techniques [7, 110]:

- Fisher's Iris Data

  Sample Code for Fisher's Iris Data
  ```
  load fisheriris
  gscatter(meas(:,1), meas(:,2), species,'rgb','osd');
  xlabel('Sepal length');
  ylabel('Sepal width');
  N = size(meas,1);
  ```

- Linear and Quadratic Discriminant Analysis

  Sample Code for Linear and Quadratic Discriminant Analysis
  ```
  lda = fitcdiscr(meas(:,1:2),species);
  ldaClass = resubPredict(lda);
  ldaResubErr = resubLoss(lda)
  ```

- Naive Bayes Classifiers

  Sample Code for Naïve Bayes Classifiers
  ```
  nbGau = fitcnb(meas(:,1:2), species);
  nbGauResubErr = resubLoss(nbGau)
  nbGauCV = crossval(nbGau, 'CVPartition',cp);
  nbGauCVErr = kfoldLoss(nbGauCV)
  labels = predict(nbGau, [x y]);
  gscatter(x,y,labels,'grb','sod')
  ```

- Decision Tree

  Sample Code for Decision Tree
  ```
  t = fitctree(meas(:,1:2), species,'PredictorNames',{'SL' 'SW' });
  [grpname,node] = predict(t,[x y]);
  gscatter(x,y,grpname,'grb','sod')
  ```

Clustering [7]

Cluster Creation
```
close all, clear all, clc, format compact
% number of samples of each class
K = 100;
```

```
% define 4 clusters of input data
q = .6; % offset of classes
A = [rand(1,K)-q; rand(1,K)+q];
B = [rand(1,K)+q; rand(1,K)+q];
C = [rand(1,K)+q; rand(1,K)-q];
D = [rand(1,K)-q; rand(1,K)-q];
% plot clusters
figure(1)
plot(A(1,:),A(2,:),'k+')
hold on
grid on
plot(B(1,:),B(2,:),'bd')
plot(C(1,:),C(2,:),'k+')
plot(D(1,:),D(2,:),'bd')
% text labels for clusters
text(.5-q,.5+2*q,'Class A')
text(.5+q,.5+2*q,'Class B')
text(.5+q,.5-2*q,'Class A')
text(.5-q,.5-2*q,'Class B')
```

Fuzzy c-means clustering [7]

```
data = load('fcmdata.dat');  % load some sample data
n_clusters = 3;              % number of clusters
[center,U,obj_fcn] = fcm(data, n_clusters);
```

Back propagation Algorithm Code in MATLAB [111]

```
% BACKPROPAGATION ALGORITHM: ONLY FOR SINGLE HIDDEN
LAYER
pattern=[0.1 0.1 0.1
0.1 .95 .95
.95 0.1 .95
.95 .95 0.1];

eta = 1.0;                          % Learning rate
alpha = 0.7;            % Momentum
tol = 0.001;            % Error tolerance
Q = 4;                              % Total no. of the patterns to be input
n = 2; q = 2; p = 1;    % Architecture
```

```
Wih = 2 * rand(n+1,q) - 1;            % Input-hidden weight matrix
Whj = 2 * rand(q+1,p) - 1;            % Hidden-output weight matrix
DeltaWih = zeros(n+1,q);              % Weight change matrices
DeltaWhj = zeros(q+1,p);
DeltaWihOld = zeros(n+1,q);
DeltaWhjOld = zeros(q+1,p);
Si = [ones(Q,1) pattern(:,1:2)];      % Input signals
D = pattern(:,3);                     % Desired values
Sh = [1 zeros(1,q)];                  % Hidden neuron signals
Sy = zeros(1,p);                      % Output neuron signals
deltaO = zeros(1,p);                  % Error-slope product at output
deltaH = zeros(1,q+1);   % Error-slope product at hidden
sumerror = 2*tol;                     % To get in to the loop

while (sumerror > tol)   % Iterate
  sumerror = 0;
  for k = 1:Q
    Zh = Si(k,:) * Wih;   % Hidden activations
    Sh = [1 1./(1 + exp(-Zh))];       % Hidden signals
    Yj = Sh * Whj;                    % Output activations
    Sy = 1./(1 + exp(-Yj));           % Output signals
    Ek = D(k) - Sy;                   % Error vector
    deltaO = Ek .* Sy .* (1 - Sy);    % Delta output
    for h = 1:q+1
      DeltaWhj(h,:) = deltaO * Sh(h);  % Delta W: hidden-output
    end
    for h = 2:q+1                          % Delta hidden
      deltaH(h) = (deltaO * Whj(h,:)') * Sh(h) * (1 - Sh(h));
    end
    for i = 1:n+1                          % Delta W: input-hidden
      DeltaWih(i,:) = deltaH(2:q+1) * Si(k,i);
    end
    Wih = Wih + eta * DeltaWih + alpha * DeltaWihOld;
    Whj = Whj + eta * DeltaWhj + alpha * DeltaWhjOld;
    DeltaWihOld = DeltaWih;                    % Store changes
    DeltaWhjOld = DeltaWhj;
    sumerror = sumerror + sum(Ek.^2); % Compute error
  end
  sumerror   % Print epoch error
end
```

Sample Code of NN Using MATLAB for Finance Management

Required functions [6]

hist_stock_data
processData
LPPL
LPPLfit
constrFunc
LPPLinteractively

Load Historic DAX Prices

The code below is an example of the use of the function hist_stock_data which will be used to download historic data of stock prices by Yahoo Finance.

```
% specify ticker symbol as string variable
tickSym = '^GDAXI'; % specify stock data of interest
% specify beginning and ending as string variables
dateBeg = '01011990'; % day, month, year: ddmmyyyy
```

Dynamic date selection is also possible with the hep of MATLAb command "today"

```
% display date of today
fprintf(['\nToday is ' num2str(today) '.\n'])
% Note: fprintf is able to display a string to the command
% window, without having to assign it to a variable or
% MATLAB's placeholder for answer "ans" first. In order to
% get the input string, in this case we first have to
% concatenate smaller strings into one large string.
```

Plotting Financial Data [6]

At the time of plotting financial data, x-axis denoted date and can be done using the command "datetick"

```
figure('position',[50 50 1200 600]) % create gray window, left
% corner at latitude 50,
% height 50, with width 1200
```

```
% and height 600
subplot(1,2,1); % Include two different white windows within
% the gray figure window. 1,2 denotes
% arrangement (one row, two columns of white
% windows), while the last number (1) denotes
% the currently used window.
% use plot command without further adjustments
plot(DAX.prices) % since no x-values are specified, MATLAB
% automatically numbers observations from 1 to
% numel(DAX.dates).
subplot(1,2,2);
plot(DAX.dates,DAX.prices)
datetick 'x' % exact format of date labels can be chosen with
% additional input, e.g. try datetick('x',29) and
% datetick('x',10)
xlabel('dates')
ylabel('prices')
title('historic DAX values')
% crop x-axis to relevant size only
set(gca,'xLim',[DAX.dates(1) DAX.dates(end)])
```

CAPM [6]

The Capital Asset Pricing Model describes the prices of assets. It is based on the assumptions.

In order to find stock's position in the market, linear regression of the daily returns is used.

```
betas = zeros(1,29);
for ii=1:29
betas(ii) = regress(DAX_stocks.disRet(:,end),...
DAX_stocks.disRet(:,ii)); % no intercept involved
end
% plot betas with expected returns
close
scatter(betas,expRets(1:end-1),'.')
% estimate regression coefficients with intercept
betaHat = [ones(numel(betas),1) betas']\expRets(1:end-1)';
% include regression line
```

```
xLimits = get(gca,'XLim');
grid = linspace(xLimits(1),xLimits(end),200);
yVals = [ones(numel(grid),1) grid']*betaHat;
hold on;
plot(grid,yVals,'r')
xlabel('estimated beta coefficients')
ylabel('estimated mean returns')
title('CAPM disproved?')
```

Stock Price Prediction Based on Curve Fitting [6]

Following is the Sample Code for Stock Price Prediction

```
% get log prices
DAX.logPrices = log(DAX.prices);
% specify subperiod as strings
begT = '01-Jun-1993';
endT = '29-Jul-1998';
% find indices associated with considered period
indS = find(DAX.dates>datenum(begT,'dd-mmm-yyyy'),1);
indE = find(DAX.dates>datenum(endT,'dd-mmm-yyyy'),1);
% create figure window
close
figure('Position',[50 50 1200 600])
% plot DAX prices with subperiod highlighted ax(1) = subplot(2,1,1);
plot(DAX.dates,DAX.prices,'Color',[1 0.8 0.8]);
hold on;
plot(DAX.dates(indS:indE),DAX.prices(indS:indE));
datetick 'x'
title('linear scale')
% plot log DAX prices with subperiod highlighted
ax(2) = subplot(2,1,2);
plot(DAX.dates,DAX.logPrices,'Color',[1 0.8 0.8]);
hold on;
plot(DAX.dates(indS:indE),DAX.logPrices(indS:indE)); shg
datetick 'x'
title('logarithmic scale')
% connect axes of both graphs: zooming in applies to both plots
linkaxes([ax(1) ax(2)],'x');
```

References

1. Pal, S. K. and Srimani, P. K. (1996, March). Neurocomputing: Motivation, models, and hybridization. *Computer*, 29(3), 24–28. NY, USA: IEEE Computer Society. ISSN 0018-9162.
2. Buckland, M., & LaMothe, A. (2002). *AI Techniques for Game Programming*. Cincinnati: Premier Press.
3. https://en.wikipedia.org/wiki/Artificial_neural_network.
4. Lam, M. (2004). Neural network techniques for financial performance prediction, integrating fundamental and technical analysis. *Decision Support Systems, 37*, 567–581.
5. Rich, E. & Knight, K., (1991). *Artificial Intelligence* (2nd Ed., pp. 4–6), New York, NY: McGraw Hill.
6. Christian Groll. (2011, August). Working with financial data: Regression analysis and curve fitting.
7. MATLAB Academy. (2015). http://in.mathworks.com, date accessed 15 September 2015
8. Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: Fundamentals, computing, design, and application. *Journal of Microbiological Methods, 43*, 3–31.
9. Anderson, C. W., Devulapalli, S. V., & Stolz, E. A. (1995). Determining mental state from EEG signals using parallel implementations of neural networks. *Scientific Programming, special issue on applications analysis, 4*(3), 171–183.
10. Touzet, C. F. (1997). Neural reinforcement learning for behaviour synthesis. *Robotics and Autonomous Systems, 22*, 251–281.
11. Faller, W., & Schreck, S. (1995). Real-time prediction of unsteady aerodynamics: Application for aircraft control and maneuverability enhancement. *IEEE Transactions on Neural Networks, 6*(6), 1461–1468.
12. Sang, K. K., & Niyogi, P. (1995). Active learning for function approximation. In G. Tesauro (Ed.), *Neural information processing systems* (Vol. 7, pp. 497–504). Cambridge: The MIT Press.
13. Vellido, A., Lisboa, P. J. G., & Vaughan, J. (1999). Neural networks in business: A survey of applications (1992–1998). *Expert Systems With Applications, 17*, 51–70.
14. Atiya, A. F. (2001). Bankruptcy prediction for credit risk using neural networks: A survey and new results. *IEEE Transactions on Neural Networks, 12*(4), 929–935.
15. Motiwalla, L., & Wahab, M. (2000). Predictable variation and profitable trading of US equities: A trading simulation using neural networks. *Computers & Operations Research, 27*, 1111–1129.
16. Yamamoto, Y., & Zenios, S. A. (1993). Predicting prepayment rates for mortgages using the cascade correlation learning algorithm. *Journal of Fixed Income, 2*(4), 86–96.

17. Lowe, D. (1994). Novel exploitation of neural network methods in financial markets. *Neural Networks, 1994. IEEE World Congress on Computational Intelligence, 1994 IEEE International Conference on*. (Vol. 6. IEEE).
18. Badiru, A. B., & Sieger, D. B. (1998). Neural network as a simulation metamodel in economic analysis of risky projects. *European Journal of Operational Research, 105*(1), 130–142.
19. Zimmermann, H. J., Neuneier, R., & Grothmann, R. (2001). *Active portfolio-management based on error correction neural networks*, *Advances in neural information processing systems, NIPS*. Cambridge: The MIT Press.
20. Ellis, C., & Willson, P. (2005). Can a neural network property portfolio selection process outperform the property market? *Journal of Real Estate Portfolio Management, 11*(2), 105–121.
21. White, H. (1988). *Economic prediction using neural networks: The case of IBM daily stock returns*. In Proceedings of the second IEEE annual conference on neural networks, II (pp. 451–458).
22. Kimoto, T., Asakawa, K., Yoda, M., &Takeoka, M. (1990). *Stock market prediction system with modular neural networks*. In Proceedings of the international joint conference on neural networks (IJCNN) (Vol. 1, pp. 1–6.) San Diego.
23. Kohara, K., et al. (1997). Stock price prediction using prior knowledge and neural networks. *Intelligent systems in accounting, finance and management, 6*(1), 11–22.
24. Pai, P.-F., & Lin, C.-S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega, 33*(6), 497–505.
25. Thawornwong, S., & Enke, D. (2004). The adaptive selection of financial and economic variables for use with artificial neural networks. *Neurocomputing, 56*, 205–232.
26. Nakayama, S., Horikawa, S., Furuhashi, T., & Uchikawa, Y. (1992). Knowledge acquisition of strategy and tactics using fuzzyneural networks. *Proc. IJCNN'92*, II-751–II-756.
27. Duke, L., & Long, J. (1993). Neural network futures trading—a feasibility study. In Society for Worldwide Interbank Financial Telecommunications (Ed.), *Adaptive Intelligent Systems, Society for Worldwide Interbank Financial Telecommunications* (pp. 121–132). Amsterdam: Elsevier Science Publishers.
28. Treacy, W., & Carey, M. (2000). Credit risk rating at large US banks. *Journal of Banking & Finance, 24*, 167–201.
29. Zhang, G., Hu, M. Y., Patuwo, B. E., & Indro, D. C. (1999). Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operational Research, 116*(1), 16–32.
30. Tam, K. Y., & Kiang, M. Y. (1992). Managerial applications of neural networks: The case of bank failure predictions. *Management Science, 38*(7), 926–947.
31. Huang, Z., Chen, H., Hsu, C. J., Chen, W. H., & Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: A market comparative study. Decision Support System, 37(4), 543–558.

32. Stoppiglia, H., Idan, Y., & Dreyfus, G. (1996). Neural-network-aided portfolio management, Industrial applications of neural networks. In F. Fogelman & P. Gallinari (Eds.), Proceedings ofICNN95, conference of European Union Neural Network. Paris: World Scientific.
33. Zimmermann, H. G., Neuneier, R., and Grothmann, R. Active Portfolio-Management based on Error Correction Neural Networks. Munchen, Germany: Siemens AG Corporate Technology.
34. Ellis, C., & Willson, P. (2005). Expert system portfolios of Australian and UK securitized property investments. Pacific Rim Property Research Journal, 12(1), 107–127.
35. Fernandez, A., & Gomez, S. (2007). Portfolio selection using neural networks. Computers & Operations Research, 34, 1177–1191.
36. Freitas, F. D., De Souza, A. F., & de Almeida, A. R. (2009). Prediction based portfolio optimization model using neural networks. Neurocomputing, 72(10–12), 2155–2170.
37. Ko, P.-C., & Lin, P.-C. (2008). Resource allocation neural network in portfolio selection. Expert Systems with Applications, 35(1–2), 330–337.
38. Chiang, W.-C., Urban, T. L., & Baldridge, G. W. (1996). A neural network approach to mutual fund net asset value forecasting. International Journal of Management Science, 24(2), 205–215.
39. Chen, A.-S., Leung, M. T., & Daouk, H. (2003). Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. Computers & Operations Research, 30, 901–923.
40. O'Connor, N., & Madden, M. G. (2006). A neural network approach to predicting stock exchange movements using external factors. Knowledge-Based Systems, 19, 371–378.
41. Faria, E. L., Albuquerque, M. P., Gonzalez, J. L., & Cayalcante, J. T. P. (2009). Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods. Expert Systems with Applications, 36(10), 12506–12509.
42. Liao, A., & Wang, J. (2010). Forecasting model of global stock index by stochastic time effective neural network. Expert Systems with Applications, 37, 834–841.
43. Kim, H.-J., Shin, K.-S. A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets. Applied Soft Computing Volume 7, Issue 2, 2007, Pages 569–576
44. Kuo, R. J., Chen, C. H., & Hwang, Y. C. (2001). An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network. Fuzzy Sets and Systems, 118, 21–45.
45. Hassan, M. R., Nath, B., & Kirley, M. (2007). A fusion model of HMM, ANN and GA for stock market forecasting. Expert Systems with Applications, 33(1), 171–180.

46. Chye, K. H., Tan, W. C., & Goh, C. P. (2004). Credit scoring using data mining techniques. Singapore Management Review, 26(2), 25–47.
47. Angelini, E., di Tollo, G., & Roli, A. (2008). A neural network approach for credit risk evaluation. The Quarterly Review of Economics and Finance, 48(4), 733–755.
48. Fanning, K., Cogger, K. O., & Srivastava, R. (1995). Detection of management fraud: A neural network approach. Intelligent Systems in Accounting, Finance and Management, 4(2), 113–126.
49. Feroz, E. H., et al. (2000). The efficacy of red flags in predicting the SEC's targets: An artificial neural networks approach. International Journal of Intelligent Systems in Accounting, Finance & Management, 9.3, 145–157.
50. Fanning, K. M., & Cogger, K. O. (1998). Neural network detection of management fraud using published financial data. International Journal of Intelligent Systems in Accounting, Finance & Management, 7(1), 21–41.
51. Koskivaara, E., & Back, B. (2007). Artificial neural network assistant (ANNA) for continuous auditing and monitoring of financial data. Journal of Emerging Technologies in Accounting, 4.1, 29–45.
52. Liou, F.-M. (2008). Fraudulent financial reporting detection and business failure prediction models: A comparison. Managerial Auditing Journal, 23(7), 650–662.
53. Chu, J. (1997). Comparing statistical methods and artificial neural networks in bankruptcy prediction. Comparing statistical methods and artificial neural networks in bankruptcy prediction. PhD thesis, University of Warwick.
54. Matsatsinis, N. F., Doumpos, M., & Zopounidis, C. (1997). Knowledge acquisition and representation for expert systems in the field of financial analysis. Expert Systems with Applications, 12(2), 247–262.
55. Liao, S.-H. (2005). Expert system methodologies and applications a decade review from 1995 to 2004. Expert Systems With Applications, 28, 93–103.
56. Bohanec, M., Rajkovic, V., Semoil, B., & Pogacnik, A. (1995). Knowledge-based portfolio analysis for project evaluation. Information Management, 28(5), 293–302.
57. Vraneš, S., et al. (1996). INVEX: Investment advisory expert system. Expert Systems, 13(2), 105–119.
58. Vranes, S., Stanojevic, M., Stevanovic, V., & Lucin, M. (1996). INVEX: Investment advisory expert system. Expert Systems With Applications, 13(2), 105–119.
59. Mogharreban, N., & Zargham, R. (2005). PORSEL: An expert system for assisting in investment analysis and valuation, soft computing a fusion of foundations. Methodologies and Application, 9(10), 742–748.
60. Grosan, C., Abraham, A., Ramos, V., & Han, S. Y. (2005). Stock market prediction using multi expression programming. In Proceedings of Portuguese conference of artificial intelligence, workshop on artificial life and evolutionary algorithms (pp. 73–78). Portuguese: IEEE Press.

61. Quek, C. (2005). Predicting the impact of anticipator action on US stock market—An event study using ANFIS (a neural fuzzy model). Computational Intelligence, 23, 117–141.
62. Trinkle, B. S. (2006). Forecasting annual excess stock returns via an adaptive network-based fuzzy inference system. Intelligent Systems in Accounting, Finance and Management, 13(3), 165–177.
63. Afolabi, M., &Olatoyosi, O. (2007). Predicting stock prices using a hybrid Kohonen self-organizing map (SOM). In 40th Annual Hawaii international conference on system sciences (pp. 1–8).
64. Yunos, Z. M., Shamsuddin, S. M., & Sallehuddin, R. (2008). Data Modeling for Kuala Lumpur Composite Index with ANFIS. In Second Asia international conference on modeling and simulation, AICMS 08, Kuala Lumpur (pp. 609–614).
65. Atsalakis, G. S., & Valavanis, K. P. (2009). Forecasting stock market short-term trends using a neuro-fuzzy based methodology. Expert Systems with Applications, 36(7), 10696–10707.
66. Shue, L.-Y., Chen, C.-W., & Shiue, W. (2009). The development of anontology-based expert system for corporate financial rating. Expert Systems with Applications, 36(2), 2130–2142.
67. Hodgkinson, L., & Walker, E. (2003). An expert system for credit evaluation and explanation. Journal of Computing Sciences in Colleges, 19(1), 62–72.
68. Shaw, M., & Gentry, J. (1988). Using an expert system with inductive learning to evaluate business loans. Finance Management, 17(3), 45–56.
69. Kim, S. H., & Chun, S. H. (1998). Graded forecasting using an array of bipolar predictions: Application of probabilistic neural networks to a stock market index. International Journal of Forecasting, 14, 323–337.
70. Wang, Y.-F. (2002). Predicting stock price using fuzzy grey prediction system. Expert Systems with Applications, 22(1), 33–38.
71. Quintana, D., Saez, Y., Mochon, A., & Isasi, P. (2008). Early bankruptcy prediction using ENPC. Applied Intelligence, 29, 157–161.
72. Lenard, M. J., Alam, P. (2009). Application of fuzzy logic to fraud detection. Category: Networking & Telecommunication.
73. Taha, I. E. (1997). A hybrid intelligent architecture for revising domain knowledge. Ph.D. Report. University of Texas at Austin.
74. Lertpalangsunti, N. (1997). An implemented framework for the construction of hybrid intelligent forecasting systems. Ph.D. Report. University of Regina.
75. Lin, C., & Hsieh., P.-J. (2004). A fuzzy decision support system for strategic portfolio management. Decision Support Systems, 38, 383–398.
76. Kosaka, M., Mizuno, H., Sasaki, T., Someya, T., & Hamada, N. (1991). Applications of fuzzy logic/neural network to securities trading decision support system. Proc IEEE ConfSyst, Man, Cyber, 3, 1913–1918.
77. Chen, L.-H., & Huang, L. (2009). Portfolio optimization of equity mutual funds with fuzzy return rates and risks. Expert Systems with Applications, 36(2), 3720–3727.

78. Quek, C., Yow, K. C., Cheng, P. Y. K., & Tan, C. C. (2009). Investment portfolio balancing: Application of a generic self-organizing fuzzy neural network (GenSoFNN). Risk Anal Complex Syst:IntelligentSyst Finance, 16(1/2), 147–164.
79. Yu, L., Wang, S., & Lai, K. K. (2008). Neural network-based mean–variance–skewness model for portfolio selection. Computers & Operations Research, 35, 34–46.
80. Li, X., Zhang, Y., Wong, H.-S., & Qin, Z. (2009). A hybrid intelligent algorithm for portfolio selection problem with fuzzy returns. Journal of Computational and Applied Mathematics, 233(2), 264–278.
81. Quah, T.-S. (2008). DJIA stock selection assisted by neural network. Expert Systems with Applications, 35(1–2), 50–58.
82. Kuo, R. J., Lee, L. C., & Lee, C. F. (1996). Integration of artificial neural networks and fuzzy Delphi for stock market forecasting. IEEE International Conference on Systems, Man, and Cybernetics, 2, 1073–1078.
83. Romahi, Y., & Shen, Q. (2000). Dynamic financial forecasting with automatically induced fuzzy associations. IEEE Syst, 1, 493–498.
84. Keles, A., Kolcak, M., & Keles, A. (2008). The adaptive neuro-fuzzymodel for forecasting the domestic debt. Knowledge-Based System, 21(8), 951–957.
85. Huang, K. Y., & Jane, C.-J. (2009). A hybrid model for stock market forecasting and portfolio selection based on ARX, grey system and RS theories. Expert Systems with Applications, 36(3), 5387–5392.
86. Elmer, Peter, J., and David M. Borowski. (1988). An expert system approach to financial analysis: The case of S&L bankruptcy. Financial Management 66–76.
87. Dadios, E. P., & Solis, J. (2012). Fuzzy-neuro model for intelligent credit risk management. Intelligent Information Management, 4(05), 251.
88. Lean, Y. U., et al. (2008). Designing a hybrid intelligent mining system for credit risk evaluation. Journal of Systems Science and Complexity, 21(4), 527–539.
89. Ahn, H., & Kim, K.-j. (2009). Bankruptcy prediction modeling with hybrid case-based reasoning and genetic algorithms approach. Applied Soft Computing, 9(2), 599–607.
90. Zopounidis, C., Doumpos, M., & Matsatsinis, N. F. (1997). On the use of knowledge-based decision support systems in financial management: A survey. Decision Support Systems, 20(3), 259–277.
91. Hua, Z., et al. (2007). Predicting corporate financial distress based on integration of support vector machine and logistic regression. Expert Systems with Applications, 33(2), 434–440.
92. Chen, Y., Ohkawa, E., Mabu, S., Shimada, K., & Hirasaa, K. (2009). A portfolio optimization model using genetic network programmingwith control nodes. Expert Systems with Applications, 36(7), 10735–10745.
93. Huang, S.-C. (2011). Forecasting stock indices with wavelet domain kernel partial least square regressions. Applied Soft Computing, 11(8), 5433–5443.

94. Lin, S.-H. (2009). A new two-stage hybrid approach of credit risk in banking industry. Expert Systems with Applications, 36(4), 8333–8341.
95. Arminger, G., Enache, D., & Bonne, T. (1997). Analyzing credit risk data: A comparison of logistic discrimination classification tree analysis and feedforward networks. Computational Statistics, 12, 293–310.
96. Lee, K. C., Han, I., & Kwon, Y. (1996). Hybrid neural network models for bankruptcy predictions. Decision Support Systems: Artificial Intelligence and Expert Systems, 18(1), 63–72.
97. Garcia-Almanza, A. L., Alexandrova-Kabadjova, B., & Martinez-Jaramillo, S. (2013). Bankruptcy prediction for Banks: An artificial intelligence approach to improve understandability. In Artificial Intelligence, Evolutionary Computing and Metaheuristics (pp. 633–656). Berlin Heidelberg: Springer.
98. Shin, K.-S., & Lee, Y.-J. (2007). A genetic algorithm application in bankruptcy prediction modeling. European Journal of Operational Research, 180(1), 1–28.
99. Chen, H.-J., Huang, S. Y., & Lin, C.-S. (2009). Alternative diagnosis of corporate bankruptcy: A neuro fuzzy approach. Expert Systems with Applications, 36(4), 7710–7720.
100. Mina, S.-H., Leeb, J., & Hanb, I. (2006). Hybrid genetic algorithms and support vector machines for bankruptcy prediction. Expert Systems with Applications, 31(3), 652–660.
101. Lenard, M. J., Watkins, A. L., & Alam, P. (2007). Effective use of integrated decision making: An advanced technology model for evaluating fraud in service-based computer and technology firms. Journal of Emerging Technologies in Accounting, 4(1), 123–137.
102. Deshmukh, A., Romine, J., & Siegel, P. H. (1997). Measurement and combination of red flags to assess the risk of management fraud: A fuzzy set approach. Managerial Finance, 23(6), 35–48.
103. Juszczak, P., et al. (2008). Off-the-peg and bespoke classifiers for fraud detection. Computational Statistics & Data Analysis, 52(9), 4521–4532.
104. Cooper, R. G., Edgett, S. J., & Kleinschmidt, E. J. (2000). New problems, new solutions: Making portfolio management more effective. Research-Technology Management, 43(2), 18–33.
105. SAS Academy for Data Science. (2015a). www.sas.com/en_us/news/sascom/2012q1/risk-highlight.html, date accessed 15 September 2015
106. Weka. (2015a). www.cs.waikato.ac.nz/ml/weka/, date accessed 15 September 2015
107. Scilab. (2015a). Scilab Publisher, www.scilab.org/, date accessed 15 September 2015
108. R. (2015a). www.r-project.org/ date accessed 15 September 2015
109. SPMF. (2015a). www.philippe-fournier-viger.com/spmf/, date accessed 15 September 2015
110. LASIN—Laboratory of Synergetics. (2015a). http://lab.fs.uni-lj.si/lasin/wp/IMIT_files/neural/NN-examples.pdf, date accessed 15 September 2015

111. Kumar, S. (2004). Neural networks: A classroom approach. New Delhi: Tata McGraw-Hill Education, 184.
112. Casqueiro, P. X., & Rodrigues, A. J. L. (2006). Neuro-dynamic trading methods. European Journal of Operational Research, 175, 1400–1412.
113. Seo, Y.-W., Giampapa, J. A., Sycara, K. (2002). Financial news analysis for intelligent portfolio management. Carnegie Mellon University Research Showcase @ CMU Robotics Institute School of Computer Science.
114. Wang, Z. B. (2004). Prediction of stock market prices using neural network techniques. Ph.D. Thesis, University of Ottawa
115. Lin, J.W., Hwang, M. I., Becker, J. D., (2003) A fuzzy neural network for assessing the risk of fraudulent financial reporting. Managerial Auditing Journal, Vol. 18 Iss: 8, pp. 657–665.

Part II

Financial Forecasting and Trading

2

Trading the FTSE100 Index: 'Adaptive' Modelling and Optimization Techniques

Peter W. Middleton, Konstantinos Theofilatos, and Andreas Karathanasopoulos

1 Introduction

Modelling and trading financial indices remains a challenging and demanding task for market participants. Forecasting financial time-series can be extremely difficult because they are influenced by a large number of variables. Much of the analyzed data displays periods of erratic behaviour and as a result drastic declines and spikes in the data series are experienced. Existing linear methods are limited as they only focus on one time-series. Some of the older machine learning models also have trouble producing accurate and profitable forecasts due to their rigid architectures. In this chapter the proposed models improve on these inefficiencies to make the models more dynamic

P.W. Middleton (✉)
University of Liverpool, Liverpool, England
e-mail: peter.william.middleton@gmail.com

K. Theofilatos
Department of Computer Engineering and Informatics, University of Patras, Patras, Greece
e-mail: theofilk@ceid.upatras.gr

A. Karathanasopoulos
American University of Beirut (AUB), Beirut, Lebanon
e-mail: ak152@aub.edu.lb

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets*, New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0_2

similar to the time-series they are tasked with forecasting. This is particularly important in times of crises as the correlations between different asset classes and time-series increase. These inadequacies have been studied in great depth by the scientific community and many methodologies have been proposed to overcome the disadvantages of previous models [1].

The main disadvantage of existing non-linear financial forecasting and trading methodologies is that most of them search for global optimal estimators. The problem with this approach is that most of the time global estimators do not exist due to the dynamic nature of financial time-series. Moreover, the algorithms used for modelling financial time-series have a lot of parameters that need to be tuned and if this procedure is performed without careful consideration the accuracy of extracted prediction models will suffer and in some cases result in a data-snooping effect. In existing models the training of a prediction model is also generally performed separately from the generation of trading signals, which has been found to reduce overall performance. For example, most machine learning algorithms that are designed for forecasting financial time-series deploy only statistical metrics for the optimization steps of their training phase and do not apply an optimization step for improving their trading performances. Here a multi-objective algorithm is employed to optimize both statistical properties and trading performance.

The purpose of this chapter is to present a novel methodology that is capable of overcoming the aforementioned limitations. This methodology is based on a sliding window approach for the one day a head prediction of the FTSE100 returns. In order to provide a forecast at every time step the proposed model trains a machine learning model using a sliding window of explanatory variables. Thus the proposed method searches for the optimal predictor for each day. The machine learning model used was a hybrid combination of an adaptive version of the particle swarm optimization (PSO) algorithm [2]. Numerous existing papers utilize PSO RBF (particle swarm optimization radial basis function) neural networks to model financial time-series however many of these are limited in their application as they do not optimize the number of hidden neurons nor do they have a selection criteria for the input series [3]. The adaptive PSO algorithm (aPSO) applied by Ding et al. [3] was used for selecting the optimal feature subset and optimizing the structure of RBFNN. Moreover, a multi-objective approach was used to account for both statistical and trading performance. In particular two fitness functions are combined to minimize error and maximize annualized returns. This approach was first successfully applied to the modelling and trading of foreign exchange rates [4]. Another important limitation of existing methodologies for modelling and trading financial time-series is that only a small set of autoregressive (AR) inputs and technical indicators are used as explanatory variables. In this investigation, a FTSE100 specific superset of 50 inputs is evaluated.

The novelty of the proposed approach lies in the application of a sliding window machine learning approach for forecasting and trading the FTSE100 and at the superiority of the proposed machine learning technique. To the best of our knowledge this is the first time that this adaptive PSO algorithm has been combined with an RBFNN to model and forecast an equity index. Our proposed machine learning method also applies the PSO algorithm to select the more relevant inputs at each time step. This is different from many other existing non-linear models as most neural networks provide a prediction in the form of a weighted computation of all inputs that are fed into the network during the training process. Therefore, the proposed model has an ability to locate the optimal feature subset, which should be used as inputs. This enables the practitioner to introduce a more expansive universe of inputs without having to worry about a noticeable reduction in training times or a redundancy of features. Moreover, the feature selection is a dynamic procedure and not a static one with different feature subsets being selected in different time steps. This also helps remove the risk of survivorship bias when backtesting older data as all major equities can be included as inputs. During the backtest, and for trading, the algorithm records the number of times an input is selected, which indicates which variables were more influential than others over the examined time period.

The performance of the proposed methodology is compared with numerous linear and adaptive methodologies. To allow for a fair comparison, all non-linear methods included in the comparative analysis were trained with the same sliding window approach. Moreover, the deployed PSO algorithm was also deployed to optimize the AR and moving average terms in an ARMA (autoregressive moving average) model.

The rest of the chapter is organized as follows: Section 2 presents a review of literature focused on forecasting methodologies and in particular neural networks and the FTSE100. Section 3 describes the dataset used for the experiments and the descriptive statistics. Section 4 describes the proposed PSO RBF methodology. Section 5 is the penultimate chapter, which presents the empirical results and an overview of the benchmark models. The final chapter presents concluding remarks and future objectives and research.

2 Literature Review

The FTSE100 is an index that has been modelled and forecasted by many who focus their research on conventional, statistical and machine learning methods. Some of the earliest research was conducted by Weigend et al. [5], Lowe [6], Tamiz et al. [7], and Omran [8]. These earlier publications establish the value of neural networks (NN) when predicting the daily changes in closing prices for FTSE100. They conclude that NN have strong predictive abilities as they outperform linear methods. It is worth noting that for all of the above studies only AR returns of lagged closing prices for the FTSE100 are used in the input dataset to train the networks.

More recent research conducted by Lee and Ko [9] focuses on RBFNNs. Lee and Ko [9] proposed a NTVE-PSO (non-linear time-varying evolution particle swarm optimization) method that compares existing PSO methods, in terms of predicting the different practical load types of the Taiwan power system (Taipower) in terms of predicting one day ahead and five days ahead. Yan et al. [10] contribute to the applications of RBFNN by experiments with real-world data sets. Experimental results reveal that the prediction performance of RBFNN is significantly better than a traditional back propagation neural network models. Marcek et al. [11] estimate and apply ARCH-GARCH (autoregressive conditional heteroscedasticity-generalized autoregressive conditional heteroscedasticity) models for forecasting the bond price series provided by the VUB (Všeobecná úverová banka). Following the estimation of these models Marcek et al. [11] then forecast the price of the bond using an RBFNN. Cao and Tay [12] compare a support vector machine model with an RBF and a generic BPNN (back propagation neural network) model. In their methodology Cao and Tay [12] analyze five futures contracts that are traded on the CME. Empirical results from this analysis conclude that the RBFNN outperforms the BPNN while producing similar results to the SVRNN (support vector machine neural network). As an overall summary the predictive ability of an RBF is significantly stronger when compared to any of the aforementioned benchmark models. In some cases the performance is almost double that of other comparable models.

With the emergence of newer technology and faster processing power finance has seen numerous advancements in the area of artificial intelligence (AI). As a result, the accuracy and practicality of such models has led to AI being applied to different asset classes and trading strategies. Enke and Thawornwong [13] suggest that machine learning methodologies provide higher returns when compared to a buy-and-hold strategy. De Freitas et al. [14] propose a novel strategy for training NN using sequential Monte Carlo algorithms with a new hybrid gradient descent/sampling importance resampling algorithm (HySIR). The effectiveness of this model was validated following an application to forecasting FTSE100 closing prices. The HySIR model outperformed all the other benchmarks in terms of trading performance. Their novel technique was fixed from values with weights that generate a 200 input-output data test. The input test data was then used to train the model using the weights estimated at the 200th time step. Schittenkop et al. [15], Tino et al. [16], Sallans et al. [17], Jasic and Wood [18] and Bennel and Sutcliffe [19], show results that indicate that for all markets the improvement in the forecast by non-linear models is significant and highly accurate. Moreover, Eldeman [20] presented a hybrid Calman filter RBF model used in forecasting the FTSE100 and ISEQ one day ahead. This study used lagged returns from previous days as inputs. The results produced by Eldeman are favourable towards the RBF model as it outperformed the buy-and-hold strategy, a moving average model and even a traditional recurrent neural network.

The past few years of AI research has been continued with Ling Bing Tang et al. [21], which analyses the application and validity of wavelet support vector machine for volatility forecasting. Results from their computer simulations and experiments on stock data reveal that kernel functions in SVMs are unable to predict accurately the cluster feature of volatility. Miazhynskaia et al. [22] attempt to forecast volatility with numerous models. Their conclusion shows that statistical models account for non-normality and explain most of the fat tails in the conditional distribution. As a result, they believe that there is less of a need for complex non-linear models. In their empirical analysis, the return series of the Dow Jones Industrial Average index, FTSE100 and NIKKEI 225 indices over a period of 16 years are studied. The results are varied across each of the markets.

More recently Dunis et al. [23] have forecasted the volatility of FTSE100 with artificial intelligent models. In their analysis higher order neural networks (HONN) outperform all the others. Moreover, Nair et al. [24] propose a hybrid GA (genetic assisted) neural network which, when compared with benchmark models, outperforms displaying superior accuracy and overall performance. Nair et al. [25] forecasts one day ahead and uses closing prices from the FTSE100, BSE Sensex, Nikkei 225, NSE-Nifty and DJIA as inputs for their models. Lastly, Karathanasopoulos et al. [26] have used a sliding window approach that combines adaptive differential evolution and support vector regression for forecasting and trading the FTSE100.

3 Related Financial Data

A robust backtest was conducted taking the largest stocks by market capitalization to be included in the training of the networks as a representation of the FTSE100's most heavily weighted stocks over the examined time period. Over the five-year backtest although weightings have changed slightly

Table 2.1 Total dataset

Name of Period	Trading Days	Beginning	End
Total dataset	1629	1 January 2007	28 March 2013
In sample dataset	1119	1 January 2001	14 April 2011
Out of sample set	510	15 April 2011	28 March 2013

the overall picture is fairly consistent. Two notable 'dropouts' however include Barclays and Royal Bank of Scotland. As a result, financials and banks saw a reduction from 27.64 %[1] to 19.16 %.[2]

The FTSE100 index is weighted according to market capitalization, which currently comprises of 101 large cap constituents listed on the London Stock Exchange. For the purpose of the trading simulation the iShares FTSE100 (ISIN: IE0005042456) exchange traded fund is traded to capture daily movements of the FTSE100 index. Trading signals are generated based on the forecast produced by each of the models. When the model forecasts a negative return then a short position (sale) is assumed at the close of each day and when the model forecasts a positive return a long position (purchase) is executed. Profit/loss is determined by daily positions and in circumstances where consecutive negative or positive changes are forecasted the position is held as a trading decision for the following day.

Arithmetic returns are used to calculate daily returns and they are estimated using equation (2.1). Given the price level P_1, P_2,…,P_t, the arithmetic return at time t is formed by:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{2.1}$$

Table 2.1 presents the examined dataset.

As inputs to the algorithms, a combination of AR returns, moving averages, fixed income returns, commodity returns, equity returns, equity index returns and a volatility time-series were all included. The approach to selecting credible inputs was not optimized however, an element of logic was applied. For the individual stock time- series, only those with the largest market capitalizations were selected as these bear more weight on the FTSE100 index. The sum of the weightings in the FTSE100 index for the selected stocks amounted to more than 50 % of the FTSE100 index. Empirical results in Sect. 4 display the significance of each input during the sample period.

[1] As at 31 August 2007—source: FTSE Group (2007).

[2] As at 29 March 2013—source: FTSE Group (2013).

Each of the inputs was translated into a return series based on equation (2.1) using the closing prices for each.

Finally, the set of explanatory variables were all normalized in the interval of [−1,1] to avoid overrating inputs of higher absolute values.

4 Proposed Method

In this algorithm the adaptive PSO methodology was used to locate the parameters C_i of the RBFNN while in parallel locating the optimal number for the hidden layers of the network. This methodology is extended to the proposed algorithm to allow its application in a sliding window approach, to optimize the feature subset. The selected candidates are then used as inputs in the proposed model with the adaptive PSO methodology and to reduce the algorithms complexity by using a standard simple neural network topology that is able to improve the generalization properties of the model.

The PSO algorithm, proposed by Eberhart and Kennedy [26], is a population-based heuristic search algorithm based on the simulation of the social behaviour of birds within a flock. In PSO, individuals, which are referred to as particles, are placed initially randomly within the hyper-dimensional search space. Changes to the position of particles within the search space are based on the social-psychological tendency of individuals to emulate the success of other individuals. The outcome of modelling this social behaviour is that the search process is such that particles stochastically return towards previously successful regions in the search space.

The performance of an RBFNN highly depends on its structure and on the effective calculation of the RBF function's centres C_i and widths σ and the network's weights. If the centres of the RBF are properly estimated then their widths and the networks weights can be computed accurately with existing heuristic and analytical methodologies that are described below. In this approach the PSO searches only for optimal values of the parameters C_i and the optimal feature subset, which should be used as inputs. A sliding window approach is used and this enables for a prediction that is based on daily re-optimization of the model's parameters and input dataset. For the number of hidden neurons (the RBFNN structure) no further optimization procedure was followed but a simple ten-node architecture was selected. This simple topology enables us to alleviate the computational cost of the optimization procedure and to maintain the simplicity in the derived models to achieve better generalization performance.

Each particle i is initialized randomly to have ten hidden neurons (within a predefined interval starting from the number of inputs until 100 which is the maximum hidden layer size that we applied) and is represented as shown in equation (2.2):

$$C^i = \begin{array}{cccc|cccc} C^i_{1,1} & C^i_{1,2} & \dots & C^i_{1,d} & & & & \\ \cdot & \cdot & \cdot & \cdot & & & & \\ \cdot & \cdot & \cdot & \cdot & \text{Input 1} & \text{Input 2} & \dots & \text{Input } d \\ \cdot & \cdot & \cdot & \cdot & & & & \\ C^i_{10,1} & C^i_{10,2} & \dots & C^i_{10,d} & & & & \end{array} \tag{2.2}$$

Where: N is a large number to point that it does not represent an RBF centre. The variables Input1 to Input d take values from −1 to 1 with values larger than 0 indicating that this feature should be utilized as input.

In our PSO variation, initially we create a random population of particles, with candidate solutions represented as showed in equation (2.2), each one having an initially random velocity matrix to move within the search space. It is this velocity matrix that drives the optimization process, and reflects both the experiential knowledge of the particle and socially exchanged information from the particle's neighbourhood. The form of the velocity matrix for every particle is described in equation (2.3):

From the centres of the particle described in equation (2.2) using the Moody-Darken [27] approach we compute the RBF widths using equation (2.3).

$$\sigma^i_j = \left\| c^i_j - c^i_k \right\| \tag{2.3}$$

Where: c^i_k is the nearest neighbour of the centres c^i_j. For the estimation of the nearest neighbours we apply the Euclidean distance, which is computed for every pair of centres.

At this point of the algorithm the centres and the widths of the RBFNN have been computed. The computation of its optimal weights w^i is accomplished by solving equation (2.4).

$$w^i = \left(H^T_i \cdot H_i\right)^{-1} \cdot H^T_i \cdot Y \tag{2.4}$$

where

$$H_i = \begin{bmatrix} \phi_1^i(x_1) & \phi_2^i(x_1) & \dots & \phi_{10}^i(x_1) \\ \phi_1^i(x_2) & \phi_2^i(x_2) & \dots & \phi_{10}^i(x_2) \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ \phi_1^i(x_{n_1}) & \phi_2^i(x_{n_1}) & \dots & \phi_{10}^i(x_{n_1}) \end{bmatrix}$$

where n_1 is the number of training samples.

The calculation of $\left(H_i^T \cdot H_i\right)^{-1}$ is computationally intensive when the rows of H_i are highly dependent. In order to solve this problem the in-sample dataset is filtered and when the mean absolute distance of two training samples is less than 10^{-3} (from the mean values of their input values) then one of them is selected at random to be included in the final training set. As a result, the algorithm becomes faster while maintaining its accuracy. This analytical approach for the estimation of the RBFNN weights is superior in comparison with the application of meta-heuristic methods (PSO, genetic algorithms, swarm fish algorithm), which have been already presented in the literature, because it eradicates the risk of getting trapped into local optima and the final solution is assured to be optimal for a subset of the training set.

The algorithm is a multi-objective algorithm that addresses two main elements. The first is an error minimization algorithm as displayed in equation (2.5). The second is employed to optimize and improve the trading performance. Equation (2.6) optimizes annualized returns as first introduced by Sermpinis et al. [4].

$$E(C,\sigma,w_t) = \frac{1}{T}\sum_{t=1}^{T}\left(y_t - \tilde{y}_t\left(w_t,C,\sigma\right)\right)^2 \tag{2.5}$$

with y_t being the target value and T the number of trading days.

$$\mathrm{R}^{\mathrm{A}} - MSE - \left(n * 10^{-2}\right) \tag{2.6}$$

where:

R^A	annualized return
MSE	mean square error
n	number of inputs

Iteratively, the position of each particle is changed by adding in its velocity vector and the velocity matrix for each particle is changed using equation (2.7):

$$V^{i+1} = w * V^i + c1 * r1 * \left(C^i_{p\text{best}} - C^i\right) + c2 * r2 * \left(C^i_{g\text{best}} - C^i\right) \tag{2.7}$$

where w is a positive-valued parameter showing the ability of each particle to maintain its own velocity, $C^i_{p\text{best}}$ is the best solution found by this specific particle so far, $C^i_{g\text{best}}$ is the best solution found by every particle so far, $c1$ and $c2$ are used to balance the impact of the best solution found so far for a specific particle and the best solution found by every particle so far in the velocity of a particle. Finally, $r1$, $r2$ are random values in the range of [0,1] sampled from a uniform distribution.

Ideally, PSO should explore the search space thoroughly in the first iterations and so the values for the variables w and $c1$ should be kept high. For the final iterations the swarm should converge to an optimal solution and the area around the best solution should be explored thoroughly. Thus, $c2$ should be valued with a relatively high value and w, $c1$ with low values. In order to achieve the described behaviour for our PSO implementation and to avoid getting trapped in local optima when being in an early stage of the algorithm's execution we developed a PSO implementation using adaptive values for the parameters w, $c1$ and $c2$. Equations (2.8), (2.9) and (2.10) mathematically describe how the values for these parameters are changed through PSO's iterations helping us to endow the desired behaviour in our methodology.

$$w(t) = \left(0.4 / n^2\right) * \left(t - n\right)^2 + 0.4 \tag{2.8}$$

$$c1(t) = -2 * t / n + 2.5 \tag{2.9}$$

$$c2(t) = 2 * t / n + 0.5 \tag{2.10}$$

where t is the present iteration and n is the total number of iterations.

For the initial population of particles a small value of 30 particles (number of articles found with backtesting experiments) is used and the number of iterations used was 200 combined with a convergence criterion. Using this termination criterion the algorithm stops when the population of the particles is deemed as converged. The population of the particles is deemed as converged when the average fitness across the current population is less than 5 % away from the best fitness of the current population. Specifically, when the average fitness across the current population is less than 5 % away from the best fitness of the population, the diversity of the population is very low and evolving it for more generations is unlikely to produce different and better individuals than the existing ones or the ones already examined by the algorithm in previous generations.

In summary, the novelty of the algorithm lies in the following points. First, the feature selection optimizations step allows the utilization of a large number of candidate inputs and enables the final model to use only the most significant variables in order to model and trade on the FTSE100. The sliding window approach allows for the approximation of a dynamic time-series as is the case in live trading conditions. Moreover, the adaptive estimation of the models parameters with a single run helps traders to avoid over-fitting and data-snooping effects. Finally the problem specific fitness function allows for the extraction of models that present high statistical and trading performance

5 Empirical Results

Benchmark Models

Three linear benchmark models and five non-linear models were used to gauge the effectiveness of the proposed PSO RBF model. A naïve trading strategy, a MACD (moving average convergence divergence) strategy and a buy-and-hold strategy were all used to generate next day trading signals. The five non-linear models include a generic MLP model, an 'adaptive' ARMA model, a PSO ARMA model, a GA MLP model and a PSO RBF model. The latter is the same methodology as the proposed PSO RBF however only AR inputs are used as explanatory variables.

With the buy-and-hold strategy the practitioner simply buys the index at the beginning of the in-sample period and then sells it at the end of the sample period. Trading statistics and performance are then recorded. The naïve strategy however is a strategy that generates trades based on the previous day's returns. Therefore, the forecast for R_{t+1} is produced by using R. The MACD strategy is based on a long-term and short-term moving average. Trading signals are produced when these two moving averages intersect. A long position is triggered when the short-term moving average intersects the long-term moving average from below and a short position is triggered when the short-term moving average intersects the long-term moving average from above.

The non-linear models are all re-estimated daily using moving window parameters of '*x*' days. For instance the AR and MA terms in the ARMA model are re-estimated based on a window of '*x*' days to produce a forecast. The PSO ARMA is optimized by a PSO algorithm to find the optimal combination of AR and MA terms. The MLP model is estimated using a traditional back propagation algorithm to adjust the weights when forecasting next day returns. The GA MLP model uses a genetic algorithm to optimized weights while also imposing input selection criteria at each time step. Finally the PSO RBF neural network uses a PSO algorithm to select the optimal inputs from a set of AR returns of the FTSE100 index.

Neural networks exist in several forms in the literature. The most popular architecture is the MLP. Their most important problem is that they require a feature selection step and their parameters are hard to optimize. For the reasons outlined by Theofilatos et al. [28] GAs [29] were used to select suitable inputs. The Levenberg–Marquardt back propagation algorithm [30] is employed during the training procedure, which adapts the learning rate parameter during this procedure.

In the second benchmark machine learning model the authors proposed a hybrid GA model as it is designed to overcome some of the limitations of ANNs. More specifically in this methodology, a genetic algorithm is used to optimize the MLP parameters and on parallel to find the optimal feature subset. Moreover, this approach used a problem-specific fitness function, which is believed to produce more profitable prediction models.

Trading Performance

The statistical and trading performance for all the models is presented in Tables 2.2, 2.3, and 2.4. The trading strategy for all of the models is to trade based on the forecast produced by each of the models. If the model forecasts a positive return then the trader buys the iShares FTSE100 ETF and if the model predicts a negative return then the trader sells the iShares FTSE100 ETF. For consecutive positive or negative signals the trader holds the previous day's trade to minimize transaction costs. As the proposed model is trained using a multi-objective algorithm the second objective focuses on optimizing annualized returns. As a result, Table 2.4 displays results from a filtered trading simulation. These models only trade when the strength of each model's forecast is greater than the '*x*' basis points.

Table 2.2 Statistical results

Trading Performance	Naive	MACD	Adaptive ARMA	PSO ARMA	MLP	GA MLP	PSO RBF Autoregressive Inputs	PSO RBF Mixed Inputs
MAE	0.0155	0.0106	0.0108	0.0141	0.0177	0.0153	0.0125	0.0103
MAPE	464.88 %	178.37 %	175.93 %	432.66 %	720.34 %	531.21 %	378.37 %	242.46 %
RMSE	0.0218	0.0155	0.0156	0.0201	0.0240	0.0218	0.0172	0.0146
Theil's Inequality	0.7212	0.8406	0.7857	0.6664	0.6788	0.6701	0.6581	0.7319
Correct Directional Change	48.03 %	48.79 %	47.96 %	51.88 %	51.54 %	52.78 %	52.71 %	56.39 %

Table 2.3 Trading results (unfiltered)

Trading Performance	Naïve	MACD	Adaptive ARMA	PSO ARMA	MLP	GA MLP	PSO RBF autoregressive inputs	PSO RBF mixed inputs
Sliding training windows	Fixed	Fixed	252	252	252	252	252	252
Gross annualized return	−11.68 %	−3.70 %	−5.19 %	25.81 %	15.25 %	23.11 %	34.82 %	51.02 %
Annualized volatility	20.97 %	21.30 %	21.30 %	21.27 %	22.38 %	22.38 %	22.38 %	22.38 %
Maximum cumulative drawdown	−77.81 %	−60.97 %	−64.47 %	−31.99 %	−42.52 %	−25.58 %	−23.39 %	−16.71 %
Average max drawdown	−40.88 %	−27.74 %	−31.15 %	−6.02 %	−14.01 %	−6.27 %	−5.46 %	−2.86 %
Maximum drawdown duration (days)	1353	1160	1227	261	596	369	327	136
Average drawdown duration (days)	560	417	462	64	174	79	65	21
Calmar ratio	−0.15	−0.06	−0.08	0.81	0.36	0.90	1.49	3.05
Information ratio	−0.56	−0.17	−0.24	1.21	0.68	1.03	1.56	2.28
# Transactions (annualized)	137	32	140	131	97	124	109	131
Total trading days	1329	1329	1329	1329	1329	1329	1329	1329
Transaction costs	6.85 %	1.60 %	7.00 %	6.55 %	4.85 %	6.20 %	5.45 %	6.55 %
Annualized return (incl. costs)	−18.53 %	−5.30 %	−12.19 %	19.26 %	10.40 %	16.91 %	29.37 %	44.47 %
RANKING	8	6	7	3	5	4	2	1

Table 2.4 Trading results (filtered)

Trading Performance	Naïve	MACD	Adaptive ARMA	PSO ARMA	MLP	GA MLP	PSO RBF autoregressive inputs	PSO RBF mixed inputs
Sliding training windows	Fixed	Fixed	252	252	252	252	252	252
Threshold trading filter (basis points)	90 bps	15 bps	30 bps	45 bps	30 bps	40 bps	30 bps	30 bps
Gross annualized Return	−9.59 %	0.20 %	13.72 %	27.03 %	19.32 %	30.87 %	59.45 %	74.73 %
Annualized volatility	16.25 %	18.25 %	15.49 %	17.76 %	20.28 %	19.48 %	18.44 %	16.84 %
Maximum cumulative drawdown	−53.48 %	−46.35 %	−24.84 %	−24.40 %	−26.69 %	−20.54 %	−17.08 %	−14.73 %
Average max drawdown	−23.42 %	−18.19 %	−6.26 %	−5.59 %	−8.51 %	−4.82 %	−3.48 %	−1.60 %
Maximum drawdown duration (days)	1353	1160	987	470	579	369	342	176
Average drawdown duration (days)	560	415	307	117	170	78	48	22
Calmar ratio	−0.18	0.00	0.55	1.11	0.72	1.50	3.48	5.08
Information ratio	−0.59	0.01	0.89	1.52	0.95	1.59	3.22	4.44
# Transactions (annualized)	201	48	211	179	110	162	149	199
Total trading days	1329	1329	1329	1329	1329	1329	1329	1329
Transaction costs	10.05 %	2.40 %	10.55 %	8.95 %	5.50 %	8.10 %	7.45 %	9.95 %
Annualized return (incl. costs)	−19.64 %	−2.20 %	3.17 %	18.08 %	13.82 %	22.77 %	52.00 %	64.78 %
RANKING	8	7	6	4	5	3	2	1

This enables the trader to capitalize on more significant moves in the index while avoiding trading during less significant periods. The confirmation filter restricts the model for trading when the forecasted value is less than the optimal confirmation threshold for its sliding window period. Finally, as the non-linear methodologies are stochastic by nature a single forecast is not sufficient to represent a credible forecast. For this reason, an average of ten estimations were executed to minimize variance.

By observation, it can be seen that the proposed mixed input PSO RBF model is the strongest statistically. It also predicts the highest number of correct directional changes.

As it was expected the proposed methodology clearly outperformed the existing models with leading results across all the examined metrics. Another interesting observation is made when comparing the proposed PSO RBF model with the PSO RBF model, which was trained using only AR inputs. It is clearly beneficial for the trader to examine a larger and more expansive universe of explanatory variables as it reduces volatility, maximum drawdowns and improved annualized returns.

In the filtered trading simulation the PSO RBF maintains its ranking as the best model. This threshold filter is optimized during the in sample and applied to the examined dataset. Results are improved under the supervision of a trading filter as overall annualized returns are increased. Furthermore, overall volatility and maximum drawdowns are also improved. Overall rankings are slightly different as the PSO ARMA model is now ranked fourth rather than third during the unfiltered trading scenario.

The input selection PSO and GA algorithms can also be compared and it can be concluded that the PSO algorithm was able to discriminate the significance of each input over the sample period with greater precision as a wider range of percentages is observed. The GA algorithm tended to select more inputs at each time step, which was more computationally intensive. Therefore, it was not as precise when selecting the 'more significant' variables.

By observations of Table 2.5 the more significant inputs are clearer when reviewing the PSO output. The GA results aren't differentiated as much however it does highlight a few 'more significant' inputs. Notably, the GA algorithm selects the FTSE100 returns lagged by three days, Vodafone Plc, AstraZeneca Plc, British Pound (CME) Continuous, and the S&P 500 returns as the variables which most explain the daily change in the FTSE100 index. The PSO algorithm however highlights a few more explanatory variables over the examined sliding windows, which include the SNP500, BHP Billiton PLC, AstraZeneca PLC, the MSCI World Index, the MSCI All Country World index and the CBOE VIX index.

Table 2.5 Input selection GA vs. PSO

Explanatory variable	Lag(s)	Percentage selected during backtest GA MLP	Percentage selected during backtest PSO RBF
Autoregressive Returns	1	55.12	50.54
Autoregressive Returns	2	54.76	46.71
Autoregressive Returns	3	57.45	45.55
Autoregressive Returns	4	56.27	48.39
Autoregressive Returns	5	56.59	44.97
Autoregressive Returns	6	57.40	45.68
HSBC Holdings Plc	1	55.38	47.04
Vodafone Group Plc	1	57.10	48.40
BP Plc.	1	55.89	50.57
Royal Dutch Shell Plc Class A	1	54.02	47.41
GlaxoSmithKline Plc	1	55.18	49.81
British American Tobacco Plc	1	56.23	47.30
Royal Dutch Shell Plc Class B	1	54.88	48.72
BG Group Plc	1	56.26	44.45
Diageo Plc	1	55.73	47.54
BHP Billiton Plc	1	56.63	53.22
Rio Tinto Plc	1	54.85	49.93
AstraZeneca Plc	1	57.01	53.35
Gold (NYM $/ozt) Continuous	1	55.14	51.13
Silver (NYM $/ozt) Continuous	1	55.81	48.79
British Pound (CME) Continuous	1	57.19	49.09
British Pounds per Euro	1	53.82	48.43
Euro per British Pounds	1	53.40	48.65
British Pounds per Swiss Franc	1	53.50	49.13
Swiss Franc per British Pounds	1	55.60	49.58
Japanese Yen per British Pounds	1	53.73	50.00
British Pounds per Japanese Yen	1	54.60	49.63
US Dollar per British Pounds	1	55.38	48.15
British Pounds per US Dollar	1	55.04	48.77
Euro STOXX 50	1	56.30	50.40
S&P 500	1	57.07	71.01
MSCI EAFE	1	55.40	50.23
MSCI The World Index	1	55.48	56.92
MSCI AC World	1	56.10	55.98
CBOE Market Volatility Index	1	56.78	54.10
Crude Oil (NYM $/bbl) Continuous	1	56.29	46.89
Brent Crude (ICE $/bbl) Continuous	1	55.21	37.93
US Benchmark Bond—6 Month	1	55.02	48.53
US Benchmark Bond—5 Year	1	55.76	48.62
US Benchmark Bond—30 Year	1	55.35	50.40
US Benchmark Bond—3 Month	1	53.85	48.70
US Benchmark Bond—2 Year	1	56.42	43.14
US Benchmark Bond—10 Year	1	53.76	47.03

(*continued*)

Table 2.5 (continued)

Explanatory variable	Lag(s)	Percentage selected during backtest GA MLP	Percentage selected during backtest PSO RBF
US Benchmark Bond—1 Month	1	54.05	47.70
21 Day MA	21	54.04	43.90
50 Day MA	50	49.93	41.36
100 Day MA	100	49.61	50.19
150 Day MA	150	48.46	49.03
200 Day MA	200	49.36	50.40
250 Day MA	250	48.17	51.23

6 Conclusions and Future Work

This chapter introduced a novel methodology for acquiring profitable and accurate trading results when modelling and trading the FTSE100 index. The proposed PSO RBF methodology is a sliding window combination of an adaptive PSO with a RBF neural network. It not only addresses the limitations of existing non-linear models but it also displays the benefits of using an adaptive hybrid approach to utilizing two algorithms. Furthermore, this investigation also fills a gap in current financial forecasting and trading literature by imposing input selection criteria as a pre-selection system before training each of the neural networks. Furthermore, a GA and PSO algorithm are both tasked with optimizing inputs which, based on an extensive literature review, has never been done before. The application of a PSO algorithm to a traditional ARMA model is also an innovation. Lastly, as first employed by Sermpinis et al. [4] the multi-objective approach to optimizing statistical and trading performance is applied to an equity index for the first time.

Experimental results proved that the proposed technique clearly outperformed the examined linear and machine learning techniques in terms of an information ratio and net annualized return. This technique is now a proven and profitable technique when applied to forecasting a major equity index. Future applications will focus on other equity indices to test the robustness of the PSO RBF model as well as other asset classes. In addition, the lag structure of the inputs will be of more focus in future applications as traders could also benefit from the 'optimization' of such parameters. The universe of explanatory variables could be enriched further to include more technical time-series such as the VWAP (volume weighted average price), high, low and opening prices. Other models outputs could also be included in the input dataset to benefit from the informational content of both existing conventional models and other non-linear methodologies.

References

1. Li, Y., & Ma, W. (2010). Applications of artificial neural networks in financial economics: A survey. *Proceedings of International Symposium on Computational Intelligence and Design (ISCID), 1*, 211–214.
2. Kennedy, J., & Eberhart, R. C. (1997). *Particle Swarm Optimization* (pp. 1942–1948). Orlando: Proceedings of International Conference on Neural Networks.
3. Ding, H., Xiao, Y., & Yue, J. (2005). Adaptive training of radial basis function networks using particle swarm optimization algorithm. *Lecture Notes in Computer Science, 3610*, 119–128.
4. Sermpinis, G., Theofilatos, K., Karathanasopoulos, A., Georgopoulos, E., & Dunis, C. (2013). Forecasting foreign exchange rates with adaptive neural networks using radial-basis functions and particle swarm optimization. *European Journal of Operational Research, 225*(3), 528–540.
5. Weigend, A. S., Huberman, B. A., & Rummelhart, E. D. (1990). Predicting the future: A connectionist approach. *International Journal of Neural Systems, 1*(3), 345–346.
6. Lowe, D. (1994). Novel Exploration of Neural Network Methods in Financial Markets, Neural Networks, 1994. *IEEE World Congresson Computational Intelligence, 6*, 3623–3428.
7. Tamiz, M., Hasham, R., Jones, D. F., Hesni, B., & Fargher E., K. (1996). A Two Staged Goal Programming Model for Portfolio Selection. *Lecture Notes in Economics and Mathematical Systems, 432*, 286–299.
8. Omran, M. F. (1996). Non Linear Dependence and Conditional Heteroscedacity in Stock Returns : UK evidence. *Applied Economics Letters, 4*(10), 647–650.
9. Lee, C. M., & Ko, C. N. (2009). Time prediction using RBF Neural Netwoks with nonlinear time-varying evolution PSO algorithm. *Neurocomputing, 73*(1-3), 449–460.
10. Yan, X. B., Wang, Z., Yu, S. H., & Li, Y. J. (2005). Time Series Forecasting with RBF Neural Network. *Machine Learning and Cybernetics, Proceedings, 8*, 4680–4683.
11. Marcek, D., Marcek, M., & Babel, J. (2009). Granular RBF NN Approach and statistical Methods Applied to Modelling and Forecasting High Frequency Data. *International Journal of Computational Intelligence systems, 2*(4), 353–364.
12. Cao, L. J., & Tay, F. E. H. (2003). Support vector machine with adaptive parameters in financial time series forecasting. *IEEE transactions on Neural Networks, 14*(6), 1506–1518.
13. Enke, D., & Thawornwong, S. (2005). The use old data mining and neural networks for forecasting stock markets returns. *Expert Systems with Applications, 29*(4), 927–940.
14. De Freitas, J. F. G., Niranjan, M., Gee, A. H., & Doucet, A. (2000). Sequential Monte Carlo Methods to Train Neural Networks Models. *Neural Computation, 12*(4), 995–993.

15. Schittenkopf, C., P. Tino, and G. Dorffner. The profitability of trading volatility using real-valued and symbolic models. InIEEE/IAFE/INFORMS 2000 Conference on Compuational Intelligencefor Financial Engineering (CIFEr), pages 8–11, 2000.
16. Tino, P., Schittenkopf, C., & Dorffner, G. (2001). Financial volatility trading using recurrent neural networks. *IEEE Transactions on Neural Networks, 12*(4), 865–874.
17. Sallans, B., Pfister, A.,Karatzoglou, A., & Dorffner, G. (2003). Simulation and validation of an integrated markets model. Journal of Artificial Societies and Social Simulation, 6(4).
18. Jasic, T., & Wood, D. (2004). The profitability of daily stock indices trades based on neural network predictions: Case study for the S&P 500, the Dax, the Topix, and the FTSE100 in the Period 1965–1999. *Applied Financial Economics, 14*(4), 285–297.
19. Bennell, J., & Sutcliffe, C. (2005). Black and scholes versus artificial neural networks in pricing FTSE100 options. *Intelligent Systems in Accounting, Finance and Management: An International Journal, 12*(4), 243–260.
20. Edelman, D. (2008). Using Kalman-Filtered Radial Basis Function Networks for index Arbitrage in the financial markets. *Natural Computing in Computing in Computational Finance Studies in Computational Intelligence, 100*, 187–195.
21. Tang, L.-B., Tang, L.-X., & Sheng, H.-Y. (2009). Forecasting Volatility based on Wavelet Support Vector Machines. *Expert Systems with Applications, 36*(2), 2901–2909.
22. Miazhynskaia, T., Fruhwirth-Schnatter, S., & Dorffner, G. (2006). Bayesian testing for non-linearity in volatility modeling. *Computational Statistics and Data Analysis, 51*(3), 2029–2042.
23. Dunis, C., Likothanassis, S., Karathanasopoulos, A., Sermpinis, G., & Theofilatos, K. (2013). A hybrid genetic algorithm-support vector machine approach in the task of forecasting and trading. *Journal of Asset Management, 14*, 52–71.
24. Nair, B., Gnana Sai, S., Naveen, A. N., Lakshimi, A., Venkatesh, G. S., & P., M. V. (2011). A GA-Artificial neural network hybrid system for financial time series forecasting. *Information Technology and Mobile Communication, 147*, 499–506.
25. Karathanasopoulos, A., Theofilatos, K., Leloudas, P., and Likothanassis, S. (2010). Modeling the ASE 20 Greek index using Artificial Neural Networks Combined with Genetic Algorithms. *Lecture notes in computer science LNCS* (*including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics*) (part 1, vol. 6352, pp. 428–435).
26. Eberhart, R.C., and Kennedy, J. (1995). A new optimizer using particle swarm theory. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 39–43. Piscataway, NJ: IEEE Service Center.
27. Moody, J. and Darken, C. Fast learning in networks of locally-tuned processing units. Neural Computation, *1*, 281–294, 1989.

28. Theofilatos, K., et al. (2012). Modelling and Trading the DJIA Financial Index using neural networks optimized with adaptive evolutionary algorithms. *Engineering Applications of Neural Networks, 311*, 462–453.
29. Holland, J. (1995). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control and artificial intelligence*. Cambridge, MA: MIT Press.
30. Roweis, S. (2013). *Levenberg-arquardt optimization*. http://www.cs.toronto.edu/roweis/notes/lm.pdf retrieved on 30.9.2013.

3

Modelling, Forecasting and Trading the Crack: A Sliding Window Approach to Training Neural Networks

Christian L. Dunis, Peter W. Middleton, Konstantinos Theofilatos, and Andreas Karathanasopoulos

1 Introduction

Petroleum refiners are exposed to price fluctuations on both sides of the refining process, which may reduce profit margins. Refiner's primary risk is that posed by an increase in input (raw materials) prices while output prices such as RBOB (Reformulated Gasoline Blendstock for Oxygen Blending) gas and heating oil remain static or simultaneously decrease. This would result in a narrowing of the spread and perhaps momentarily result in a negative spread as the price of crude becomes greater than the sum of output prices. In order to hedge this

C.L. Dunis (✉)
ACANTO Holding, Hannover, Germany
e-mail: christian.dunis@acantoholding.com

P.W. Middleton
University of Liverpool, Liverpool, England
e-mail: peter.william.middleton@gmail.com

K. Theofilatos
Pattern Recognition Laboratory, Department of Computer Engineering & Informatics, University of Patras, 26500 Patras, Greece
e-mail: theofilk@ceid.upatras.gr

A. Karathanasopoulos
American University of Beirut (AUB), Beirut, Lebanon
e-mail: ak152@aub.edu.lb

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets,*
New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0_3

risk the 'Crack Spread' is traded to safeguard profit margins. The process of converting crude oil into 'refined' outputs, which include petroleum gas, gasoline, kerosene, diesel, industrial fuel oil (heating oil), lubricating oil, paraffin wax and asphalt, is known as 'cracking'. Crude oil is cracked to produce each by-product. In the refining industry there are two widely used crack ratios as the hedge traded by each refiner varies based on variables such as capacity and operational configuration. Furthermore, both the inputs (grades of crude oil) and outputs vary from region to region depending on requirements for delivery and demand for finished products. The RBOB unleaded gasoline contract traded here is relatively new to the NYMEX exchange as the grade of gasoline changed in 2005 to include ethanol in the mix.

The first hedge is based on the 3:2:1 ratio, which means that three barrels of crude oil are required to 'crack' two barrels of gasoline and one barrel of distillate heating oil fuel. The other ratio that refiners also trade is known as the 5:3:2 ratio. In this case, five barrels of crude are 'cracked' into three barrels of gasoline and two barrels of heating oil. Refiners that crack crude with a lower yield of gasoline relative to distillate are more likely to trade using the latter of the two combinations.

The spread is positive and hence profitable when the sum of the by-products is greater than the cost to procure crude oil. As the hedge is executed based on the output side of the spread refiners generally purchase crude oil futures to hedge rising crude prices and sell both the gasoline and heating oil futures to hedge decreasing output prices. This would be considered 'shorting' or selling the spread. Furthermore, these counteracting positions allow the market participant to 'lock into' a predetermined margin. For the purpose of this investigation, a spread between crude oil, gasoline and heating oil is formed by trading three futures contracts of crude oil, two futures contracts of RBOB unleaded gasoline and one futures contract of heating oil. This spread most closely represents the WTI Cushing/NYH RBOB 3:2:1 Crack as displayed in Fig. 3.1.

Motivation for this investigation derives from the initial analysis carried out by Dunis et al. [1] who model the Crack Spread between NYMEX West Texas Intermediate (WTI) for crude oil and NYMEX Unleaded Gasoline (GAS). Conclusions reveal that NNs offer interesting results and the aim here is to offer more insight into the benefits of using non-linear modelling, by expanding the universe of explanatory variables, to train the network over different sliding windows using both a PSO algorithm and a traditional back propagation algorithm. In addition, each model is filtered using a threshold confirmation filter to enhance performance. Furthermore, the Spread, which is investigated here, also includes heating oil as an output whereas previous research does not. Therefore, as a result of heating oil not being traded until 2008, the Crack Spread is calculated using three variables and not just crude

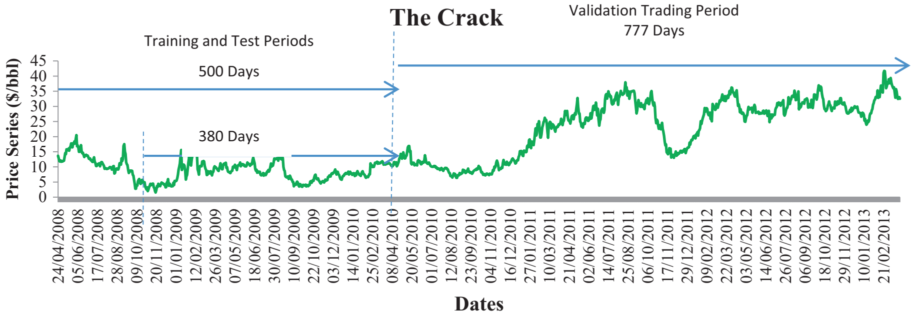


Fig. 3.1 Full sample price series of the 3:2:1 Crack Spread

oil and gasoline as traded in Dunis et al. [1]. Here, a more in-depth application of NNs is examined in order to more accurately predict daily changes in the Crack Spread.

The Crack Spread is calculated using three variables. The input variable is crude oil (CL), which is denominated in US dollars per barrel, while the outputs consist of gasoline (RBOB) and heating oil (HO) of which both are denominated in US cents per gallon. In order to create the spread a conversion of units is required. As the quantity of a crude contract is 1,000 barrels per contract and both the gasoline and heating oil amount to 42,000 gallons per contract then the latter two are multiplied by 0.42. This is based on the calculation that there are 42 gallons of oil per barrel. Using this conversion of units the outputs are converted into US dollars per barrel as mathematically depicted in equation 3.1.

$$3:2:1 \text{ CRACK SPREAD } St = \frac{\left(\left(\left(2\times \text{RBOB}\times 0.42\right)+\left(1\times \text{HO}\times 0.42\right)\right)-\left(3\times \text{CL}\right)\right)}{3} \quad (3.1)$$

The methodology applied throughout this investigation in order to calculate the returns of the Crack Spread can be seen below as provided by Butterworth and Holmes [2] and more recently by Dunis et al. [3] and Dunis et al. [4]:

$$\Delta S_t = \left[\left(\frac{\left(P_{RBOB(t)} - P_{RBOB(t-1)}\right)}{\left(P_{RBOB(t-1)}\right)} + \frac{\left(P_{HO(t)} - P_{HO(t-1)}\right)}{\left(P_{HO(t-1)}\right)}\right) - \left(\frac{\left(P_{CL(t)} - P_{CL(t-1)}\right)}{\left(P_{CL(t-1)}\right)}\right)\right] \quad (3.2)$$

Where

ΔS_t Percentage change in returns of the Crack Spread at time $_t$
$P_{RBOB(t)}$ is the price of RBOB at time $_t$ (in dollar per barrel)

$P_{RBOB(t\text{-}1)}$ is the price of RBOB at time $_{t\text{-}1}$ (in dollar per barrel)
$P_{HO(t)}$ is the price of Heating Oil at time $_{t}$ (in dollar per barrel)
$P_{HO(t\text{-}1)}$ is the price of Heating Oil at time $_{t\text{-}1}$ (in dollar per barrel)
$P_{CL(t)}$ is the price of Crude Oil at time $_{t}$ (in dollar per barrel)
$P_{CL(t\text{-}1)}$ is the price of Crude Oil at time $_{t\text{-}1}$ (in dollar per barrel)

The larger cap refiners include Exxon Mobil Corp., Total S.A., Royal Dutch Shell Plc., Chevron Corp., ConocoPhillips and BP Plc., as displayed in Fig. 3.2. Figure 3.3 on the other hand, focuses on small- to medium-sized refiners such as Western Refining Inc., Alon USA Energy Inc., Hess Corp., Tesoro Corp., and Valero Energy Corp. Both Figs. 3.2 and 3.3 display price performance (rebased to 100) of each company compared to the Crack Spread traded over the period 9 April 2010 to 28 March 2013.

By observation, Fig. 3.2 displays a clear and strong relationship with each of the refining company's equity. Refiner's equity increases as the spread widens and decreases as it narrows. The one exception or break in this relationship is in the summer of 2010 when BP Plc's stock price declined as a result of the oil spill in the Gulf of Mexico. This however, shows how many other additional factors besides endogenous factors such as operational efficiency also affect profit margins. Refining margins are also eroded by fixed costs and in general

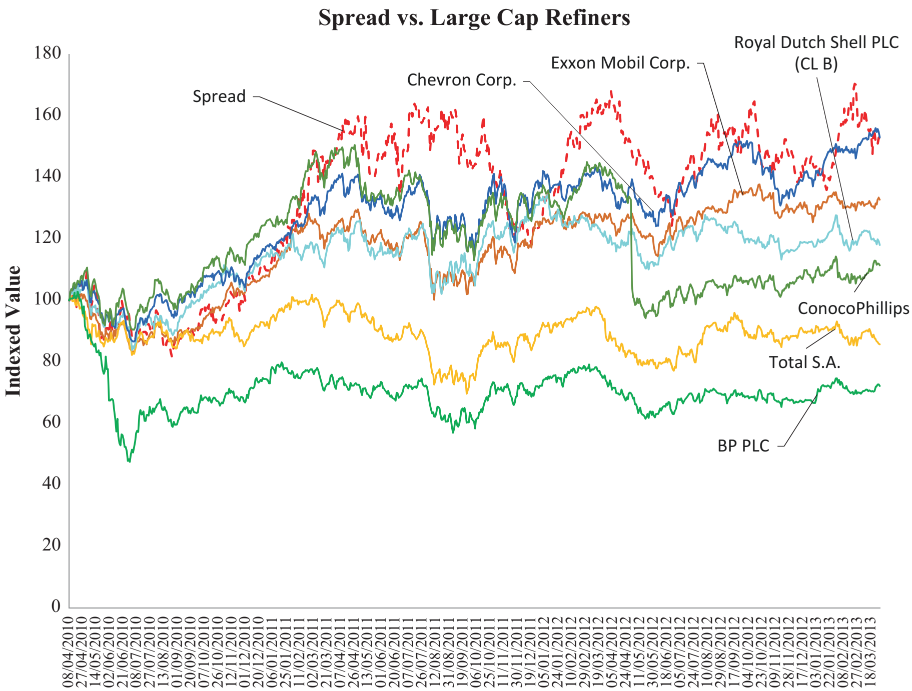


Fig. 3.2 Trading dataset price performance (rebased/indexed to 100). The 'Crack' vs. Large Cap Refiners Equity

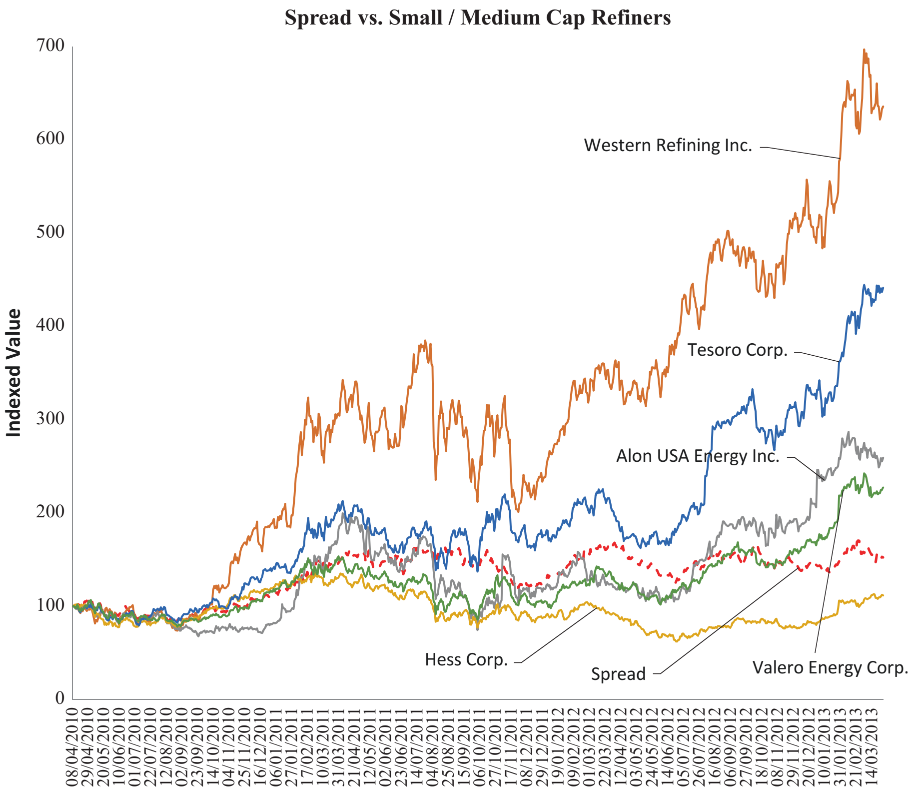


Fig. 3.3 Trading dataset price performance (rebased/indexed to 100). The 'Crack' vs. Small/Medium Cap Refiners Equity

refiners aim to operate at their determined break-even points to avoid inefficiency associated with excess capacity. Furthermore, as explained by Dunis et al. [1] the magnitude of adjustments on the upside tend to be greater and more favourable than the losses endured on the downside. This could indicate that larger refiners have more influence on the Crack Spread and some may even manipulate margins to enhance their earnings.

Instead of benchmarking the proposed PSO RBF (particle swarm optimization radial basis function) and MLP (multi-layer perceptron) models against linear models, which is frequently criticized, this investigation utilizes informational content from traditional models. In the literature there are many examples that compare traditional methodologies with NN and other AI models however here the focus is on comparing two competing NNs using different configurations for each. Traditional models are included in the universe of inputs to produce a mixed-model approach in an attempt to improve the accuracy and trading performance of each neural network. In particular, the inclusion of a GARCH (generalized autoregressive conditional heteroscedasticity) volatility time-series was justified as it enhanced performance by reducing volatility and maximum drawdowns.

Preliminary research has led to a number of unanswered questions when using neural networks as a methodology for forecasting commodity spread time-series. For instance, how large should the training window be? Should the inputs be preprocessed (i.e. normalization of inputs or the removal of outliers)? How should he network be configured (e.g. number of hidden neurons, number of layers, etc.)? What algorithm should be used to train the data (back propagation, PSO, genetic algorithms, to mention a few)? In an attempt to answer these questions the remaining structure of this chapter is presented as follows. Section 2 provides a review of all current literature relevant to modelling the Crack Spread and other gasoline spreads. A review of literature, which draws on previous research on sliding windows as a technique to train networks is also included in Sect. 2.2. Section 3 presents descriptive statistics of the data used to model the spread. Section 4 presents the methodologies and estimation parameters for the PSO, RBF, NN and MLP NN. Section 5 offers an evaluation of empirical results and trading performance. This is then followed by concluding remarks and research limitations.

2 Literature Review

Modelling the Crack

Numerous linear methodologies have been applied to the task of modelling and trading various combinations of gasoline spreads as well as the Crack Spread investigated here. For instance, Al-Gudhea et al. (2006) use threshold co-integration models to capture the relationships between crude, spot wholesale and retail gasoline price adjustments during the period of December 1998 to January 2004. In total four spreads are analyzed. The first is a spread between crude oil prices and retail gasoline prices, the second is between crude oil prices and spot gasoline prices, the third spot gasoline prices and wholesale gasoline prices, and the fourth spread is that of wholesale gasoline prices and retail gasoline prices. Test statistics from each of these spreads confirm that they are all co-integrated with evidence of asymmetric adjustments toward long-run equilibrium.

Chen et al. [5] also utilize threshold co-integration models when examining price adjustments for the spread between crude oil and gasoline prices. They find evidence of asymmetry in both the short- and long-run adjustments using both futures data and spot prices. In particular, conclusions reveal that retail gasoline prices respond asymmetrically to crude oil price changes.

In a similar approach to modelling the Crack Spread, Dunis et al. [1] use both the aforementioned Enders and Granger [6] threshold co-integration technique and numerous neural network architectures. They apply a Higher

Order Neural Network (HONN), a Recurrent Neural Network (RNN) and a Multilayer Perceptron Neural Network (MLP NN) to the task of predicting next day spread returns. All of which are trained using the most common BP algorithm. A fixed training period is used to train each of the networks with the training set being divided into training and test datasets in order to avoid 'over-fitting'. Over-fitting in this application is largely avoided due to the fact that the training window slides 'x' amount of days splitting each period into training and test datasets each time a forecast is produced. This is however discussed in more detail in Sect. 3. Results from Dunis et al. [1] reveal that the spread does in fact exhibit asymmetric adjustment. It is also observed that movements away from fair value are almost three times larger on the downside than on the upside. Overall the fair value co-integration model produces the most profitable trading performance. Out of three neural networks the HONN produces profits in excess of those achieved by the RNN and MLP neural networks.

Training of Neural Networks

Different approaches to training neural networks have been explored by many over the years and even more so in recent years. Kaastra and Boyd [7] discuss various techniques used to train neural networks. The most popular and widely used approach is one where the practitioner elects fixed windows for training and validation datasets. For example, this training approach was adopted by Dunis et al. [1] who also model the Crack Spread. Using a fixed training and test dataset they train the network using 80 % of the data and then validate the neural parameters over the remaining 20 % of the dataset (out-of-sample trading). Training datasets usually account for 70–90 % of the in-sample period while the validation dataset covers anywhere from 30 % to as little as 10 %. Another approach is one where the practitioner randomly selects the validation data set that is usually within the training dataset. This however may bias the test and reduce the accuracy when validating the training using larger out-of-sample datasets. For this reason the first approach is usually favoured by practitioners. In addition, the first approach of selecting simultaneous in-sample and out-of-sample datasets allows practitioners to test the parameters of the 'trained' neural network on more recent data, which is usually more relevant than older data. The final approach Kaastra and Boyd [7] propose is a 'sliding window' approach as used in this investigation when training both the PSO RBF and MLP neural networks. Kaastra and Boyd [7] call this a 'walk-forward' testing routine, which is commonly adopted by commodity trading systems to model and trade data in dynamic and changing market conditions. In order to adapt to these changing conditions a sliding window is utilized to provide a more robust and time varying approach. Furthermore, NNs are trained more frequently using more recent data. Hence, this technique continuously updates the training dataset and as a result it also provides a more practical and realistic approach to trading financial assets.

More recently, Tsai and Wang [8] use an average of different sliding windows to obtain an ensemble forecast when predicting next day returns for Taiwanese electronic stocks. They run four different sliding windows and take an average of these four training sets to produce a forecast. Chang et al. [9] find that the performance of neural networks is enhanced when using ensemble and hybrid techniques such as combining multiple forecasts of varying sliding windows. Thawornwong and Enke [10] use a sliding window training technique to forecast an S&P500 monthly time-series using a total of 31 inputs from 24 years of data. In particular they use four different sliding windows to capture different trends while also registering the significance of inputs during each of these windows. Over the four training periods, Thanwornwong and Enke [10] find that six inputs were consistently selected. These include the consumer price index for all urban consumers (CP), the money supply (M1), the 3-month T-bill rate (T3), the one-month certificate of deposit rate (CD1), the default spread between BAA and AAA credit ratings, the default spread between BAA and T120 (DE2) and the default spread between BAA and T3 (DE5). Therefore, it can be assumed that these variables were 'reasonably' significant as explanatory variables for the prediction of a monthly S&P500 time-series.

In a comparative analysis of ANNS and Genetic Evolutionary Algorithms (GEAs) Cortez et al. [11] discuss the implications that may arise when selecting the duration of sliding windows. For instance, a large sliding window may increase the complexity of the neural network that could ultimately reduce the learning capabilities of the model. On the other hand, smaller windows may not contain a sufficient amount of information for the neural network to be able to train the data and produce 'informationally' significant forecasts.

3 Descriptive Statistics

All data was sourced from Bloomberg for the period of 24 April 2008 to 28 March 2013 for WTI Crude, RBOB Unleaded Gasoline and Heating Oil futures contracts. The RBOB Unleaded Gasoline contract is fairly new to the exchange as it replaces the old Unleaded Gasoline contract as Methyl Tertiary-Butyl Ether (MTBE) was phased out in 2005.[1] Segregation of the dataset is displayed in Table 3.1.

[1] This was seen to be less environmentally friendly than its alternative ethanol. As a result this new blend now comprises of 10 % ethanol.

Table 3.1 Dataset

Period	In sample training/ test days	Trading days	Beginning	End
Total Dataset	510[a] 390[b]	777	24/04/2008 10/10/2008	28/03/2013
380-Day Training (Initial Window)	380	380	10/10/2008	08/04/2010
500-Day Training (Initial Window)	500	500	24/04/2008	08/04/2010
Validation set (out-of-sample)	0	777	09/04/2010	28/03/2013

[a]500 days for all inputs except the autoregressive inputs and 10 days for autoregressive inputs which regress 10 lags.
[b]300 days for all inputs except the autoregressive inputs and 10 days for autoregressive inputs which regress 10 lags.

As presented in Table 3.1 the modelling and trading of the PSO RBF and the MLP NN is based on two sliding training windows using 380 for the shortest period and 500 days of data points for the longer period. The first represents 1.5 years of working days and the second covers two full years of working days. Anything less than 380 days was found to produce unsatisfactory results therefore it is assumed that the training period did not include sufficient data points to accurately capture patterns within the data.

For the proposed PSO RBF and MLP models, over-fitting is dealt with using a two-pronged approach. Firstly, each of the sliding windows is separated into training and test datasets. Training sets account for 66.66 % of the sliding window while the remaining 33.33 % is allocated for testing. The second control that has been tested during the in-sample backtest and implemented for the validation period uses a fixed and constant amount of neurons in the hidden layer. For instance, for the PSO RBF model a total of ten neurons were found to produce adequate results during the in-sample backtest while avoiding over-fitting. In the absence of a ‘feature selection method’ all inputs are selected during the training process for the MLP model. The complexity of the network is calculated based on the number of inputs as displayed in equation 3.3.[2]

$$h = (n+1)/2 \tag{3.3}$$

Where:

h number of hidden neurons
n number of inputs

[2] For this application a total of 30 hidden neurons were used during the MLP training process.

A PSO algorithm is used to calculate the number of hidden neurons for the RBF NN. This algorithm is programmed to adapt, search for and identify the 'optimal' number of neurons. Results from these experiments produced an average of 25–30 neurons in the hidden layer. In this case, the complexity of a network with as many as 30 neurons was found to 'over fit' the dataset. Therefore, it was decided to use fewer (ten neurons) neurons in order to reduce the risk of over fitting with a less complex network topology.

Each of these training periods produces one-day-ahead forecasts. In a similar approach, Von Mettenheim and Breitner [12] use a sliding window of 128 days to produce forecasts for 10 days ($t, t_{+1}, t_{+2} \ldots t_{+10}$) ahead when modelling various stocks and ETFs. As the training process is rolling so too are the forecasts. For instance, the model, which is trained over 380 days, uses 380 data points in addition to the ten autoregressive spread return series commencing on 10 October 2008, which is 390 days before the beginning of the validation period 9 April 2010. This would provide a forecast for $_{t+1}$. In order to obtain the predicted $_{t+2}$ output the window moves forward by one day to include the forecasted $_{t+1}$ in the training period, which is used to estimate $_{t+2}$. Then $_{t+1}$ and $_{t+2}$ are used in the training window to produce $_{t+3}$ and so on. Therefore, the sliding window approach is where the PSO RBF and MLP networks are trained to use the last k values of a series ($t_{n-k} \ldots t_n$) to predict the next value at t_{n+1}. In practice, this means that the model only needs to be trained every 'x' day(s) depending on how far into the future one wishes to forecast. In this case, the neural network is retrained every day to produce a forecast as traded during the out-of-sample validation period. More frequent retraining of a sliding window is not problematic as it takes a matter of minutes to retrain and generate forecasts. This can be done over a weekend or outside trading hours such as in the morning before the market opens.

The two models used here are trained to forecast the next day change in the Crack Spread (S_{t+5}) using historical returns from 59 different explanatory variables. $S_{t,}$ is essentially the daily change in the spread as calculated in equation 3.2. Simple returns are used as inputs due to the fact that they enable neural networks to converge much more quickly than price series data. Furthermore, many simple return time-series are also found to be stationary, which is the main reason for quicker convergence. This is however not always the case as some time-series display unit roots.

The selection of input variables is a modelling decision that can greatly affect a model's performance. Dunis et al. [1], who also model and trade the Crack Spread, only use autoregressive returns of the spread to produce non-linear forecasts however for the purpose of this application a more comprehensive and significant set of inputs are considered. The aim here is to accurately capture and forecast the directional change of the spread by enriching the input dataset with more explanatory variables. A larger universe of inputs was initially evaluated over the duration of each training window. Following numerous backtests a total of 59 inputs were retained for out-of-sample trading. Included in these 59 inputs are various moving average time-series based on 21, 50, 100, 150, 200 and 250 days, changes in daily implied volatility was also included by using the CBOE VIX index, five ARMA (autoregressive moving average) and two GARCH models are also incorporated into the training process. Research conducted by Dunis et al. [13] found that the inclusion of the ARMA models as inputs to a 'mixed neural network' improves both statistical accuracy and trading performance as the training of the neural network is enhanced. Therefore, the inclusion of linear models as inputs for neural network training is justified. For the most part, in this application the inclusion of volatility models is found to effectively reduce overall volatility while also improving maximum drawdowns during the training period.

The majority of existing neural network literature uses fixed training windows during in- sample datasets, which is not realistic especially during times when the dataset is continuous or when it experiences various regime changes. Furthermore, for the proposed RBF neural network the application of a PSO algorithm in input selection also provides more insight into the significance of each input as the percentage of time each is selected during the sliding windows is also recorded as displayed in Table 3.2. This enables practitioners to see which explanatory variables are more influential during the period 9 April 2010 to 28 March 2013 (777 trading days). The difference of results between the 380- and 500-day sliding windows may indicate that each sliding window identifies different trends in the data with different inputs becoming more significant at times than others.

The percentage of time each input is selected over all of the training periods is estimated based on:

$$\text{Input Selection Percentage} = N / R^{*} \qquad {}^{*}\text{with } R = \left(S - X\right) / S_{t+n} \quad (3.4)$$

Where:

N number of sliding window repetitions an input was selected
R repetitions
S total sample dataset
X days of sliding window
S_{t+n} spread forecast horizon

Table 3.2 PSO RBF input selection during the training windows

		SELECTION AS A % OF THE TRAINING WINDOW	
		PSO RBF 380-Day Sliding Window	PSO RBF 500-Day Sliding Window
NEURAL INPUTS = 59 Total	Lags	1-Day Forecast	
Autoregressive Returns	1	46.26	47.01
Autoregressive Returns	2	68.06	65.92
Autoregressive Returns	3	47.91	49.15
Autoregressive Returns	4	45.19	47.01
Autoregressive Returns	5	46.37	48.83
Autoregressive Returns	6	44.45	45.19
Autoregressive Returns	7	48.82	48.72
Autoregressive Returns	8	50.54	46.26
Autoregressive Returns	9	50.43	53.10
Autoregressive Returns	10	46.69	39.74
Exxon Mobil Corp. Stock Price Returns	1	43.70	43.91
Total S.A. Stock Price Returns	1	47.65	51.50
Royal Dutch Shell Stock Price Returns	1	52.35	50.85
Chevron Corp. Stock Price Returns	1	46.15	52.03
ConocoPhillips Stock Price Returns	1	51.39	44.44
BP PLC Stock Price Returns	1	52.46	52.78
Western Refining Inc. Stock Price Returns	1	48.61	54.28
Alon USA Energy Inc. Stock Price Returns	1	53.10	52.89
Hess Corp. Stock Price Returns	1	50.43	51.28
Tesoro Corp. Stock Price Returns	1	49.57	45.94
Valero Energy Corp. Stock Price Returns	1	50.43	55.02
Crude Oil (NYM $/bbl) Returns	1	50.97	50.96
Brent Crude (ICE $/bbl) Returns	1	34.08	26.71
NY Harb RBOB (NYM $/gal) Returns	1	47.97	49.25
Heating Oil (NYM $/gal) Returns	1	47.12	45.41
Natural Gas (NYM $/btu) Returns	1	48.82	51.28
CBOE Market VIX Return Series	1	48.18	51.71

(*continued*)

Table 3.2 (continued)

		SELECTION AS A % OF THE TRAINING WINDOW	
		PSO RBF 380-Day Sliding Window	PSO RBF 500-Day Sliding Window
NEURAL INPUTS = 59 Total	Lags	1-Day Forecast	
Gold (NYM $/ozt) Return Series	1	47.76	47.54
Silver (NYM $/ozt) Return Series	1	48.51	52.67
British Pound (CME) Return Series	1	46.15	46.90
U.S. Dollar per Euro Return Series	1	48.18	45.09
USD / JPY Return Series	1	48.83	46.15
USD / CHF Return Series	1	47.65	47.33
USD / CAD Return Series	1	47.86	46.90
USD / AU D Return Series	1	46.69	46.15
USD / GBP Return Series	1	50.00	47.22
Euro STOXX 50 Return Series	1	50.75	48.08
S&P 500 Return Series	1	48.29	50.64
FTSE 100 Return Series	1	49.79	47.44
MSCI EAFE Return Series	1	48.83	49.47
MSCI The World Index Return Series	1	49.36	49.47
MSCI AC World Return Series	1	50.54	50.43
US TREASURY Bond 2 yr. Return Series	1	51.39	48.93
US TREASURY Bond 5 yr. Return Series	1	51.82	51.07
US TREASURY Bond 10 yr. Return Series	1	54.81	49.89
US TREASURY Bond 30 yr. Return Series	1	53.31	48.18
21 Day MA Return Series	21	53.74	49.57
50 Day MA Return Series	50	51.39	48.72
100 Day MA Return Series	100	53.95	48.61
150 Day MA Return Series	150	50.32	47.65
200 Day MA Return Series	200	50.11	51.49
250 Day MA Return Series	250	54.38	54.06
ARMA 1 Returns	(10,10)	43.80	48.29
ARMA 2 Returns	(8,8)	45.62	44.55
ARMA 3 Returns	(13,13)	55.99	63.57
ARMA 4 Returns	(4,4)	47.65	47.76
ARMA 5 Returns	(12,12)	55.63	52.99
GARCH 1 Returns	(16,16)	59.61	59.40
GARCH 2 Returns	(15,15)	59.08	68.27

Table 3.3 Most significant explanatory variables

Explanatory Variable	Lags (days)	380-Day Sliding Window (%)	500-Day Sliding Window (%)
Spread Return Series	2	68.06	65.92
BP PLC Stock Price Returns	1	52.46	52.78
Western Refining Inc. Stock Price Returns	1	48.61	54.28
Alon USA Energy Inc. Stock Price Returns	1	53.10	52.89
Valero Energy Corp. Stock Price Returns	1	50.43	55.02
US TREASURY Bond 10 yr. Return Series	1	54.81	49.89
US TREASURY Bond 30 yr. Return Series	1	53.31	48.18
250 Day MA Return Series	250	54.38	54.06
ARMA (13,13)	13	55.99	63.57
ARMA (12,12)	12	55.63	52.99
GARCH (16,16)	16	59.61	59.40
GARCH (15,15)	15	59.08	68.27

Table 3.3 provides a summary of the most significant PSO RBF neural inputs. In particular, the ARMA and GARCH inputs prove to be among the most valuable as explanatory variables.

By including ARMA and GARCH time-series the trading performance and statistical accuracy of the models was increased substantially. In addition, autoregressive time-series of spread returns were also included in the modelling of the Crack Spread. The most significant input of the lagged spread returns from lags of one to ten days was the two-day lag with this being selected as often as 68.06 % of the time during the 380 days sliding window period and 65.92 % during the 500-day sliding window. Other more influential inputs included the daily changes in some of the refiners' share prices. For instance, BP Plc., Valero Energy Corp., Alon Energy Inc. and Western Refining Inc. were all seen as more significant relative to the other refiners. Each of these inputs is lagged by one day. Furthermore, of the daily changes in US treasury rates the 10 and 30 year rates were selected as much 54.81 % and 53.31 % respectively. Interestingly, Brent was the least selected input as it was only included 26.71 % of the time using the 500-day sliding window.

A histogram of the spread's return series over the entire sample period is displayed in Fig. 3.4. This is found to display a leptokurtic distribution with positively high kurtosis. This however is quite common when observing normal distributions of return series as data points tend to be highly concentrated around the mean. Furthermore, all of the spreads are confirmed to be non-normal (confirmed at a 99 % confidence level by the Jarque-Bera test statistic).

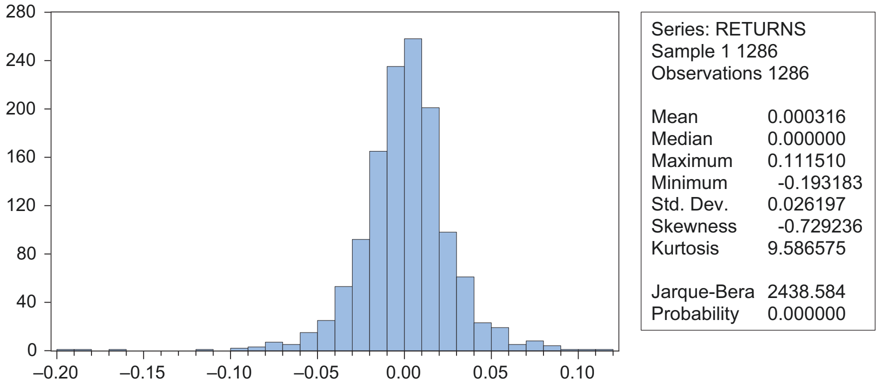


Fig. 3.4 Spread returns (out of sample)

Equations and estimation output for each of the ARMA and GARCH models have been included in Appendices A3 and A4 respectively. All ARMA models were found to be significant at a 95 % confidence level as their p-values were less than 0.05 for each of the estimated (p,q) terms. The GARCH models were deemed stable and terms for both models were also significant at 95 % confidence level. Residuals were tested for serial correlation using the squared residual test revealing that serial correlation is not present in either of the models. Therefore, the estimated models are deemed adequate and have been used to estimate the change in spread as two of the explanatory variables that are included during the training sliding window process of the neural network.

4 Methodology

The MLP Model

The multi-layer perceptron allows the user to select a set of activation functions to explore including identity, logistic, hyperbolic tangent, negative exponential and sine.[3] These activation functions can be used for both hidden and output neurons. MLP also trains networks using a variety of algorithms such as gradient descent, conjugate gradient and BFGS (Broyden, Fletcher, Goldfarb and Shanno). Here, the logistic activation function and gradient descent algorithm are used.

[3] This activation function is considered to be non-monotonic in that it is difficult to make weights vary sufficiently from their initial position. This can result in much larger numbers of local minima in the error surface [14].

The network architecture of a conventional MLP network can best be illustrated in Fig. 3.5.

Where:

$x_t^{[n]}$ $(n = 1, 2, \ldots, k+1)$	model inputs (including the input bias node) at time t
$h_t^{[m]}$ $(m = 1, 2, \ldots, j+1)$	hidden node outputs (including the hidden bias node)
$\tilde{y}_t$	MLP model's output
u_{jk} and w_j	network weights

 sigmoid transfer function

$$S(x) = \frac{1}{1+e^{-x}}, \tag{3.5}$$

 linear output function

$$F(x) = \sum_i x_i \tag{3.6}$$

The error function to be minimized is

$$E(u_{jk}, w_j) = \frac{1}{T}\sum_{t=1}^{T}\left(y_t - \tilde{y}_t(u_{jk}, w_j)\right)^2 \tag{3.7}$$

with y_t being the target value and T the number of trading days.

Training and selection of a network is halted once profit (in the form of an annualized return) is at its greatest during the in-sample test period.

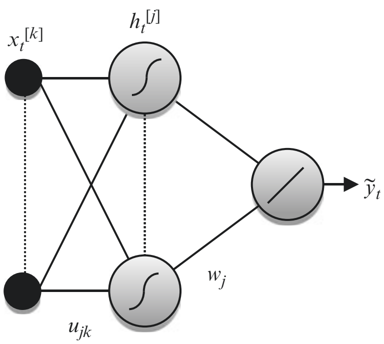


Fig. 3.5 A single output, inter-connected MLP model

The PSO Radial Basis Function Model

An RBF NN is a feed-forward neural network where hidden layers do not implement an activation function, but instead a radial basis function. As discussed by Park et al. [15], input values in an RBF network are each assigned to a node in the input layer and then passed directly through to the hidden layer without weights. On the other hand, traditional neural networks such as the MLP pass inputs through to the hidden layer as weighted computations.

The PSO aspect introduces a hybrid approach to the training of a network and hence the refinement of its forecasting accuracy has been compared to that achieved by Genetic Programming Algorithms (GPA). PSO was first introduced by Kennedy and Eberhart [16] as a stochastic optimizer during the neural network training process. Kennedy and Eberhart [16] developed the PSO algorithm based on observations found within nature such as the social behaviour found within a flock of birds or a school of fish. With these observations as a basis, the algorithm is developed to search a fixed space in an attempt to identify optimal positions within this space to best solve a predefined problem. In particular, PSO optimization reduces the time it takes to train neural networks by simplifying the complex calculations found within traditional neural networks and determining the optimal number of hidden layers.[4] Many academics have previously researched standard Radial Basis Function Neural Networks however the combination of PSO and NNs is relatively new to time-series analysis. As explained by Chen and Qian [17], PSO optimizes parameters within a traditional RBF. In particular, this optimization helps overcome inefficiencies associated within the standard back propagation algorithm.

The RBF neural network approximates a desired function by the superposition of non-orthogonal, radially symmetric functions as discussed in more detail by Konstantinos et al. [18]. The networks architecture is depicted below in Fig. 3.6 [19].

Here, the Gaussian radial basis function is used in the hidden layer (as seen in equation 3.8) as this is the most common found in existing financial time-series literature.

[4] For the purpose of forecasting, the proposed PSO RBF model utilizes a constant layer of ten neurons. Tests were conducted using the algorithm to search for the 'optimal' number of hidden neurons. Results from these tests produce a lot more than ten neurons and as a result the PSO RBF was found to 'over-fit' the data in most cases. This can be checked by observing the best weights output and comparing training using fewer fixed neurons with what the algorithm would use if it was tasked with identifying the 'optimal' number of neurons. With this in mind, a number of experiments were run using varying numbers of hidden neurons. All of the PSO RBF parameters are provided in Appendix A4. The best weights for each of the models are included in Appendix A5.

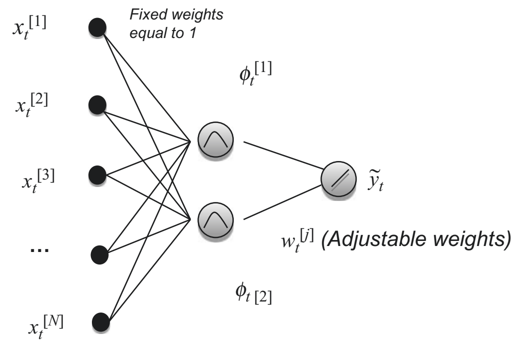


Fig. 3.6 Radial basis function neural network (with two hidden nodes)

$x_t(n = 1, 2, \ldots, N + 1)$	are the model inputs (including the input bias node)
$\tilde{y}_t$	RBF are the model's output
$x_t^{[N]}(j = 1, 2)$	are the adjustable weights
	is the Gaussian function

$$\phi_t^i(x) = e^{-\frac{\|x - C_i\|^2}{2\sigma_i^2}} \tag{3.8}$$

where:

C_i is a vector indicating the centre of the Gaussian Function and σ_i is a value indicating its width. C_i, σ_i and the weights w_i are parameters that are optimized by the PSO algorithm during a learning phase while training the RBF neural network.

	is the linear output function

$$U(x) = \sum_i x_i \tag{3.9}$$

The error function to be minimized is:

$$E(C, \sigma, w_t) = \frac{1}{T}\sum_{t=1}^{T}\left(y_t - \tilde{y}_t(w_t, C, \sigma)\right)^2 \tag{3.10}$$

with y_t being the target value and T the number of trading days.

In order to maximize annualized returns an additional fitness function is employed as defined in equation 3.11. This approach was first introduced by Sermpinis et al. [19].

The annualized return function to be maximized is:

$$R^A - MSE - \left(n * 10^{-2}\right) \tag{3.11}$$

Where:

R^A is the annualized return
MSE mean square error defined in equation 3.10.[5]
n number of inputs.

The R^A terms range from −0.4 to 0.5 while experimental results indicated that the maximum value for the MSE term is 0.01. These parameters are established so that the algorithm can primarily search for profitable forecasts with statistical performance becoming of secondary importance.

The hybrid methodology of combining a PSO with an RBF NN was first inspired by Li and Xios [20] and is also an extension of the PSO RBF methodology proposed by Sermpinis et al. [19]. The PSO methodology is used to locate the parameters C_i, of the RBF NN, while at the same time locating the optimal feature subset, which should be used as inputs to the RBF network.

The complexity of a traditional neural network is reduced by applying the PSO algorithm to refine the training process. As applied by Konstantinos et al. [18], the PSO algorithm encodes network weights as particle components with each particle evaluating inputs based on minimizing the error function in equation 3.10. PSO parameters are also 'adaptive' as depicted in equations 3.12–3.14. This proves beneficial to a wider range of users. Therefore 'velocity' as described originally by Kennedy and Eberhart [14] is adaptable with the algorithm retaining knowledge of an input's (particle) best position within the population (swarm).

With the PSO algorithm the traditional neural network weight matrix is reorganized as an array of randomly initialized particles to commence the optimization procedure. During this search the PSO algorithm is assessing 'global' and 'local' variants. A local variant is an individual particle's best solution achieved thus far while the global variant is the best solution achieved

[5] The number of hidden neurons is multiplied with 10^{-2} because the simplicity of the derived neural network is of secondary importance compared with the other two objectives (maximize the annualized return and minimizing the MSE).

in the entire population of particles. Furthermore, Mohaghegi et al. [21] note that particles have a tendency to repeat their past behaviour (cognitive) as well as follow the behaviour of those particles deemed 'fit' (socialization). The eventuality of this behaviour is that the population of particles converges to create an optimal solution. Upon the completion of iterations the particles return to their best position, which is identified during the search/training process. Predefined parameters for the PSO algorithm can be found in Appendix A4 (Tables 3.12 and 3.13). For a more detailed explanation please refer to Eberhart et al. [22] and Konstantinos et al. [19].

$$W_{(\mathrm{T})} = \left(0.4 / N^2\right) * \left(T - N\right)^2 + 0.4 \tag{3.12}$$

$$\mathrm{c1}_{(\mathrm{T})} = \left(-2\right) * T / N + 2.5 \tag{3.13}$$

$$\mathrm{c2}_{(\mathrm{t})} = \left(2\right) * T / N + 0.5 \tag{3.14}$$

where:

T is the current iteration
N is the total number of iterations.

Weights are decreased from 1.0 to 0.4 during the training phase in search of a candid solution to the proposed problem. In selecting the appropriate training set the termination criterion applied to the PSO algorithm is 10^{-3}. Ultimately, training is stopped once the number of iterations reaches 100 or the profit in the form of annualized returns is at its maximum.

5 Empirical Results

The general trading rule is to long the spread on a positive forecast and short the spread when a negative forecast is indicated. When consecutive positive or negative signals are generated then the position is held from the previous signal. Longing the spread or buying the spread is when WTI Crude Oil is sold and both Heating Oil and RBOB Gasoline are bought. Shorting the spread or selling the spread occurs when WTI Crude is bought and both Heating Oil and RBOB Gasoline are sold.

Statistical Accuracy

Statistics are computed by taking the average of ten executions in order to reduce the variance of each forecast. As neural networks are stochastic by nature it is in the best interest of a practitioner to use an average derived from numerous models. Computationally this is not too time consuming as forecasts are generated by numerous computers (Table 3.4).[6]

From a statistical perspective the PSO RBF model that is trained over 380 days is the most accurate when predicting $_{t+5}$ returns. In particular the Correct Directional Change (CDC) statistic is more than 50 %. A CDC of greater than 50 % is more desirable. Both of the MLP sliding window models are also found to be less accurate in comparison to the PSO RBF models. For all other statistics the lower they are the more accurate a model is considered to be. As explained by Dunis et al. [23] the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) statistics are 'scale-dependent' measures. These provide a modeller with statistics to compare each of the models with actual Crack Spread returns. The Theil-U statistic is one that falls between 0 and 1 with a model producing 0 being considered a 'perfect' model. Despite the significance of statistical accuracy the ultimate test is for a model to produce profit at acceptable levels of risk. Therefore, many traders will be more interested in how a model trades (discussed in Sect. 5.2).

Trading Performance

During the training process the best weights for each of the PSO RBF models were registered. These have been included in Appendix A5 (Tables 3.14 and 3.15). In total there are ten sets of best weights as each model is based on the average of five underlying models.

Numerous sliding windows were backtested and then traded for the purpose of forecasting the Crack Spread. As mentioned previously, any windows with less than 380 days of observations were found to produce unsatisfactory results.

Unfiltered results from both 380 and 500-day sliding training windows are presented in Table 3.5. Table 3.5 shows that the PSO RBF that was trained using a 380-day sliding window achieved the highest annualized returns and the best risk return trade off. This is challenged closely by the PSO RBF trained over 500 days. The MLP models, which were also trained over 380 and 500 days ranked, third and fourth consecutively. Another interesting observation is that both the PSO RBF and the MLP models that were trained using a sliding window of 380 days had much worse maximum drawdowns in comparison to their respective 500-day models. However, in order to minimize maximum drawdowns a threshold filter, which was optimized during the training period, is used to filter each model.

[6] Intel core i5 processors were used during both the backtesting and forecasting phases. Furthermore, in order to reduce the estimation time four out of the five cores are utilized by executing the Parallel Toolbox function in Matlab 2011.

Table 3.4 Out-of-sample trading statistics

Statistical Performance	PSO RBF Model		MLP Model	
Sliding Training Windows	380	500	380	500
Forecast	5 days ahead	5 days ahead	5 days ahead	5 days ahead
MAE	0.0147	0.0148	0.205	0.0203
MAPE	166.87 %	158.46 %	420.05 %	442.12 %
RMSE	0.0194	0.0196	0.0263	0.0260
THEIL-U	0.8369	0.8349	0.6877	0.6974
Correct Directional Change (CDC)	52.38 %	51.87 %	50.32 %	50.84 %

Results from a filtered trading simulation are presented in Table 3.6. With this threshold filter the model only trades when the PSO RBF and MLP NN models produce forecasts greater than '*x*' or less than '*x*'. These '*x*' parameters are optimized during the in sample period as a threshold for trading each of the models.[7] When comparing each of the forecasted return series it is clear that the MLP models are more erratic as they did not include the additional fitness function defined in equation 3.11, which maximizes annualized returns. Using this filter, only larger more significant forecasts are traded while smaller less significant changes in the spread are filtered out. This minimizes maximum drawdowns and reduces volatility while also increasing annualized returns. Model rankings remain constant with the PSO RBF model, which is trained over a 380-day sliding window, producing the highest annualized returns and best risk/return profile.

When trading futures contracts a trader has to be aware of margins. At present, margins are around 9 % for each of the contracts however most brokers calculate margins on an aggregate level and in this case margins would be calculated based on the spread performance of WTI crude, RBOB gasoline and heating oil. With this in mind traders could consider Calmar ratios as an indicator of how much return a model produces for one unit of drawdown as part of their criteria for selecting suitable models. Therefore, similar to the information ratio, a model that produces more than one would be considered a 'good' model. The formula used to calculate the Calmar ratio is displayed in the Appendix A1. In this case, a trader would trade a filtered PSO RBF model using 380 days to train the network in order to forecast five days ahead. The Calmar ratio provides an indication of risk-adjusted performance and for the proposed PSO RBF filtered (380-day sliding window) model it is 1.69. Therefore it can be assumed that for one unit of drawdown a 1.69 unit of return is produced. This is more than double the unfiltered performance, which only produces 0.73 as maximum drawdowns are substantially higher with the filter. In terms of volatility the PSO RBF (380-day sliding window) also produces the most attractive risk/return profile as it trades with a 1.83 information ratio. As a result of the filter the model trades

[7] For the RBF 380- and 500-day models the '*x*' parameter = 0.20 %. For the MLP 380-day model the '*x*' parameter = 1.90 % and for the MLP 500-day model '*x*' = 1.45 %.

Table 3.5 Out-of-sample unfiltered trading performance

Trading Performance	PSO RBF MODEL		MLP MODEL	
Sliding Training Windows	380	500	380	500
Forecast Horizon	5 days ahead	5 days ahead	5 days ahead	5 days ahead
Gross Annualized Return	32.99 %	28.84 %	22.02 %	18.86 %
Annualized Volatility	23.92 %	23.94 %	23.94 %	23.95 %
Maximum Cumulative Drawdown	−44.90 %	−29.10 %	−46.31 %	−30.09 %
Average Daily Drawdown	−6.13 %	−6.50 %	−7.07 %	−5.22 %
Maximum Drawdown Duration (days)	248	234	318	191
Average Drawdown Duration (days)	29	44	50	30
Calmar Ratio	0.73	0.99	0.48	0.63
Information Ratio	1.38	1.21	0.92	0.79
# Transactions (annualized)	109	94	104	113
Total Trading Days	777	777	777	777
Transaction costs (annualized)	10.83 %	9.37 %	10.38 %	11.29 %
Net Annualized Return	22.16 %	19.47 %	11.64 %	7.31 %
RANKING	1	2	3	4

Using a ten basis point (bps) round trip transaction cost as offered by interactive brokers.

Table 3.6 Out-of-sample filtered trading performance

Trading Performance	PSO RBF MODELS		MLP MODELS	
Sliding Training Windows	380	500	380	500
Forecast Horizon	5 days ahead	5 days ahead	5 days ahead	5 days ahead
Gross Annualized Return	33.26 %	28.60 %	19.84 %	21.08 %
Annualized Volatility	18.19 %	17.77 %	13.28 %	14.72 %
Maximum Cumulative Drawdown	−19.70 %	−19.35 %	−16.68 %	−16.72 %
Average Daily Drawdown	−2.75 %	−3.16 %	2.63 %	−2.78 %
Maximum Drawdown Duration (days)	102	162	194	148
Average Drawdown Duration (days)	19	22	32	22
Calmar Ratio	1.69	1.48	1.19	1.26
Information Ratio	1.83	1.61	1.49	1.43
# Transactions (annualized)	90	88	55	71
Total Trading Days	777	777	777	777
Transaction costs (annualized)	8.98 %	8.79 %	5.48 %	7.04 %
Net Annualized Return (incl. costs)	24.28 %	19.81 %	14.36 %	14.05 %
Annualized Returns Filter Effect	2.12 %	0.34 %	2.72 %	6.74 %
Volatility Reduction	5.73 %	6.17 %	10.66 %	9.23 %
Drawdown Reduction	25.20 %	9.75 %	29.63 %	13.37 %
RANKING	1	2	3	4

less frequently, which reduces the impact of transaction costs. High transaction costs is one of the main drawbacks highlighted by Dunis et al. [1] with an annualized average of 17.03 % in transaction costs and between 93 and 106 trades per year being triggered for an MLP, an RNN and a HONN model.

Figures 3.7 and 3.8 display the best two unfiltered trading performances over the out-of-sample trading periods for each of the sliding windows. By observation, both PSO RBF models experience periods of long drawdowns particularly from 18 April 2012 to 28 March 2013. However, the sliding window of 500 days recovers slightly and hits a new high watermark on 27 February 2013.

Figures 3.9 and 3.10 display the best two filtered trading performances. Notably, the period of prolonged drawdowns mentioned during the unfiltered simulation is reduced as new high watermarks are more frequently achieved. Each of these models is also less erratic, which reduces volatility by between 1.51 % and 5.82 %.

A threshold filter is applied to reduce the frequency of trading while lessening volatility and maximum drawdowns. In Dunis et al. [1] high transaction costs were found to significantly reduce profitability of each neural network.

6 Concluding Remarks and Research Limitations

Results from empirical analysis clearly show that the sliding window technique for training the proposed PSO RBF neural network offers a mixture of positive results. The same is also true for the MLP neural network. Furthermore the

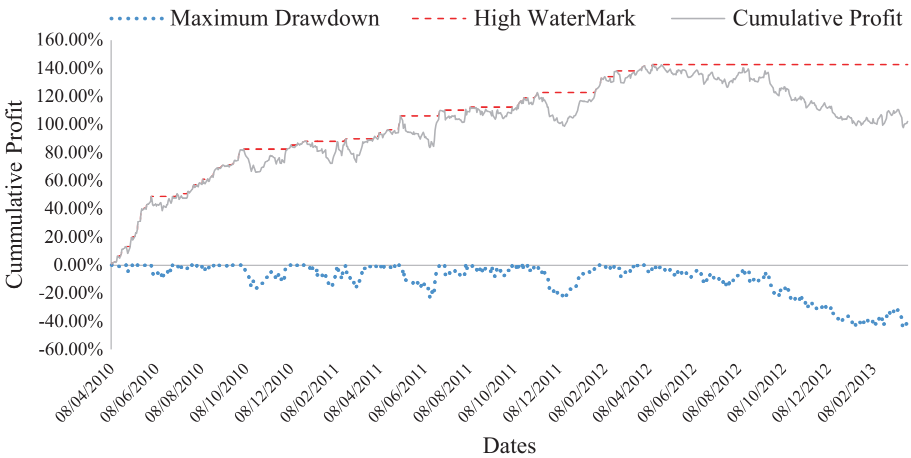


Fig. 3.7 PSO RBF unfiltered trading performance (380 days sliding window)

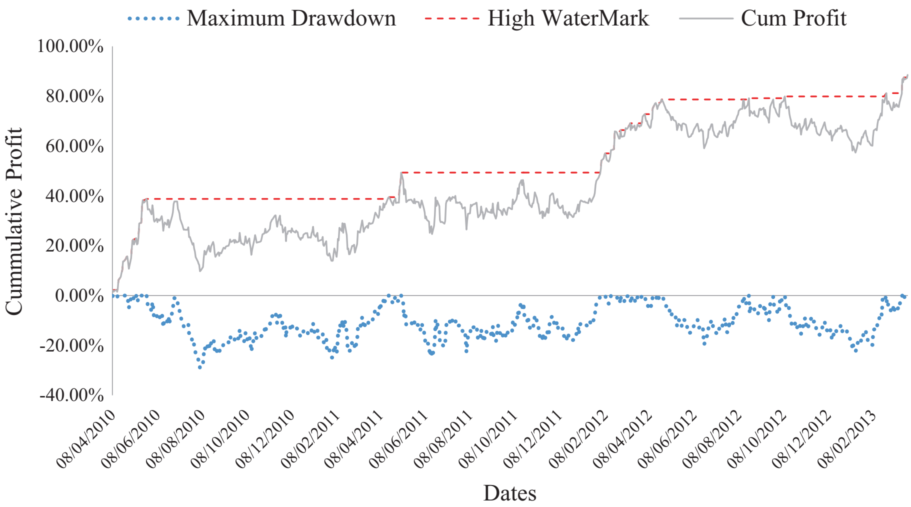


Fig. 3.8 PSO RBF unfiltered trading performance (500 days sliding window)

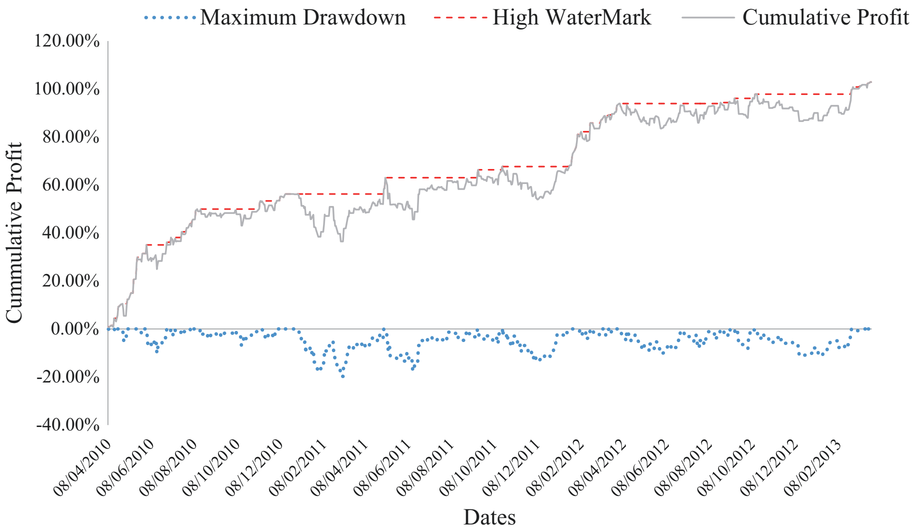


Fig. 3.9 PSO RBF filtered trading performance (380 days sliding window)

inclusion of linear models as inputs also assists in enhancing the performance of both PSO RBF and MLP models. This is corroborated by Makridakis [24], Clemen [25], Newbold and Granger [26], and Palm and Zellner [27] who all establish that forecasts are improved by combining different linear forecasting methodologies when compared to individual forecasts. For the PSO RBF models a feature selection method is explored by using the PSO algorithm to

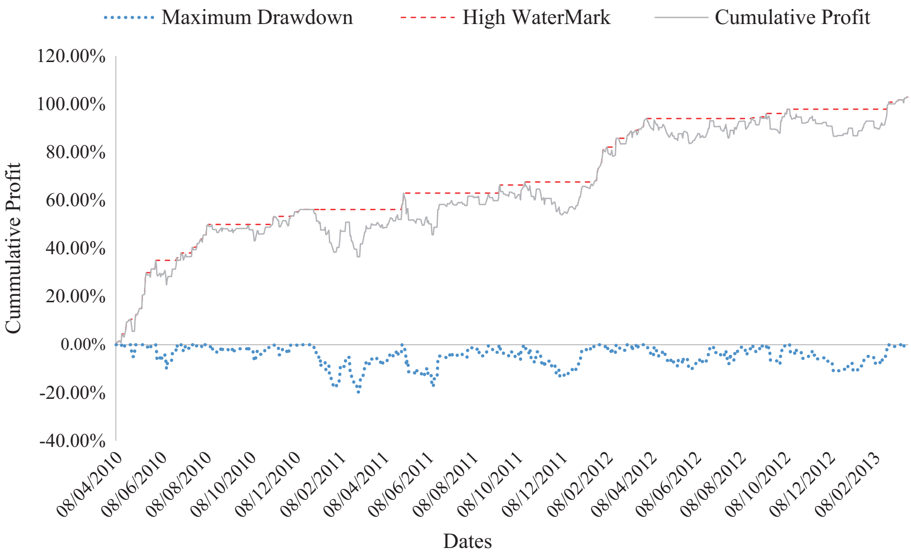


Fig. 3.10 PSO RBF filtered trading performance (500 days sliding window)

optimize the inputs. During both sliding windows only the more significant inputs are selected to train the PSO RBF NN. Each time an input is selected the algorithm produces a '1' and when an input is not selected then a '0' is generated. At the end of the trading period the algorithm then calculates a total for each input as a percentage of time each were selected. Over the 380 and 500-day sliding windows a handful of more significant explanatory variables emerged. Table 3.3 in Sect. 3 summarizes the more significant inputs over these periods. In summary, the longer term moving average inputs along with the ARMA and GARCH inputs ranked among the most significant. On the other hand, the MLP used all of the inputs for its training as no optimization algorithms were employed during the input selection phase.

Unfiltered trading simulations are generated from sliding training windows of 380 and 500 day. Each of these models forecast five days ahead using a total of 59 explanatory variables to train both the PSO RBF and MLP models. Empirical results for the RBF NN produced 22.16 % and 19.47 % in annualized returns respectively. Information ratios for the RBF models were 1.38 for the 380-day window and 1.21 for the 500-day sliding window. Calmar ratios were slightly lower with 0.73 and 0.99 respectively. MLP models generated 11.64 % and 19.47 % in annualized returns, 0.92 and 0.79 as information ratios, 0.48 and 0.63 as Calmar ratios. Transaction costs for each scenario were extremely high as the models were frequently trading even during times of little change. This was also found to be the case by Dunis et al. [1] who initially model the Crack Spread. For this reason a threshold confirmation filter was imposed.

The threshold confirmation filter only generates a trading signal once each of the forecasts is greater than x % or less than –x %. This way each of the models only trades when the forecasts indicate more significant movements in the spread. As a result Information and Calmar ratios are significantly increased. The RBF model that is trained by a 380-day sliding window now trades with an information ratio of 1.83 and a Calmar ratio of 1.69. The other RBF model that is trained using a 500-day sliding window produces 1.61 as an Information ratio and 1.48 for its Calmar ratio. Filtered returns for the 380- and 500-day MLP sliding window models were also improved with 14.36 % and 14.05 % respectively. Similar to the RBF models both the Information and Calmar ratios are also enhanced considerably. In summary, the risk/return and maximum drawdown/return profiles for each of the simulations are improved. All models return more than one unit of return (annualized return) for every one unit of risk (annualized volatility). As spread trading of futures contracts routinely requires market participants to meet margin calls a trader has to be aware of adverse movements in the spread. With this in mind, a trader would aim to select a model that produces the highest return relative to drawdowns. In this case, a trader would select the PSO RBF 380-day sliding window model as it trades with a superior Calmar ratio of 1.69.

There are a few limitations found within this research. For one, only a few sliding windows are analyzed and traded with two of the most suitable periods being displayed in the empirical findings. Results taken from combined sliding windows may enhance performance and will be researched in future applications. Further research could also be conducted to produce forecasts from an ensemble of many models as proposed by Mettenheim and Breitner [12] who use a Historically Consistent Neural Network (HCNN) to provide forecasts. Finally the proposed PSO RBF could also be applied to other asset classes such as equities, foreign exchange, derivatives and fixed income in order to test its robustness.

7 Appendix

Performance Measures

See Table 3.7

Supplementary Information

See Table 3.8

Table 3.7 Statistical and trading performance measures

Root Mean Squared Error (RMSE)	$RMSE = \sqrt{(1/N) * \sum_{t=1}^{t+N} (\bar{\sigma}_t - \sigma_t)^2}$
Mean Absolute Error (MAE)	$MAE = (1/N) * \sum_{t=1}^{t+N} \lvert \bar{\sigma}_t - \sigma_t \rvert$
Mean Absolute Percentage Error (MAPE)	$\mathrm{MAPE} = (1/N) * \sum_{t=1}^{t+N} \left\lvert \frac{\bar{\sigma}_t - \sigma_t}{\sigma_t} \right\rvert$
Theils-U Statistic	$\mathrm{THEIL} - U = \sqrt{(1/N) * \sum_{t=1}^{t+N} (\bar{\sigma}_t - \sigma_t)^2} \Big/ \left[\sqrt{(1/N) * \sum_{t=1}^{t+N} (\bar{\sigma}_t)^2} + \sqrt{(1/N) * \sum_{t=1}^{t+N} (\sigma_t)^2} \right]$
Correct Directional Change (CDC)	$CDC = (100/N) * \sum_{t=1}^{t+N} Dt$ Where $Dt = 1$ if $(\sigma_t - \sigma_{t-1}) * (\bar{\sigma}_t - \sigma_{t-1}) > 0$, Else $D\,t = 0$.
Annualised Return	$R^A = 252 * \frac{1}{N} \sum_{t=1}^{N} R_t$ *with R_t being the daily return*
Cumulative Return	$R^C = \sum_{t=1}^{N} R_t$ *with R_t being the daily return*
Annualised Volatility	$\sigma^A = \sqrt{252} * \sqrt{\frac{1}{N-1} * \sum_{t=1}^{N} (R_t - \bar{R})^2}$
Information Ratio	$IR = \frac{R^A}{\sigma^A}$
Maximum Drawdown	*Maximum negative value of* $\sum (R^c{}_t)$ *over the period* $\mathrm{Max}\,DD = \mathrm{Min}\left[R_t - \mathrm{Max}\left(\sum_{t=1}^{N} R_t \right) \right]$
Calmar Ratio	$CR = \frac{R^A}{\lvert \mathrm{Max}\,DD \rvert}$

Table 3.8 The refiner's market capitalization

Refiner	Market capitalization (m$)	As at:	Source
Exxon Mobil Corp.	384,819	03/05/2013	FactSet (2013)
Chevron Corp.	222,559	03/05/2013	FactSet (2013)
Royal Dutch Shell Plc (CL B)	134,994	03/05/2013	FactSet (2013)
BP PLC	84,283	03/05/2013	FactSet (2013)
Total S.A.	83,399	03/05/2013	FactSet (2013)
ConocoPhillips	69,458	03/05/2013	FactSet (2013)
Hess Corp.	22,959	03/05/2013	FactSet (2013)
Valero Energy Corp.	21,190	03/05/2013	FactSet (2013)
Tesoro Corp.	6,737	03/05/2013	FactSet (2013)
Western Refining Inc.	2,629	03/05/2013	FactSet (2013)
Alon USA Energy Inc.	1,056	03/05/2013	FactSet (2013)

Table 3.9 ARMA equations

ARMA Models	Equations	
(10,10)	$Y_t = 5.22*10^{-4} - 0.583Y_{t-1} + 0.458Y_{t-6} - 0.481Y_{t-9} - 0.568Y_{t-10} - 0.575\varepsilon_{t-1} + 0.450\varepsilon_{t-6} - 0.516\varepsilon_{t-9} - 0.595\varepsilon_{t-10}$	(3.16)
(8,8)	$Y_t = 4.79.10^{-4} - 1.161Y_{t-1} - 0.208Y_{t-4} + 0.122Y_{t-8} - 1.067\varepsilon_{t-1} - 0.257\varepsilon_{t-4} + 0.159\varepsilon_{t-8}$	(3.17)
(13,13)	$Y_t = 3.91.10^{-4} - 0.605Y_{t-1} + 0.503Y_{t-5} + 0.220Y_{t-13} - 0.603\varepsilon_{t-1} + 0.551\varepsilon_{t-5} + 0.222\varepsilon_{t-13}$	(3.18)
(4,4)	$Y_t = 4.33.10^{-4} - 0.510Y_{t-1} + 0.109Y_{t-2} - 0.558Y_{t-3} - 0.891Y_{t-4} - 0.519\varepsilon_{t-1} + 0.113\varepsilon_{t-2} - 0.569\varepsilon_{t-3} - 0.953\varepsilon_{t-4}$	(3.19)
(12,12)	$Y_t = 4.80.10^{-4} + 0.551Y_{t-1} - 0.699Y_{t-3} - 0.345Y_{t-7} - 0.177Y_{t-12} + 0.554\varepsilon_{t-1} - 0.709\varepsilon_{t-3} - 0.283\varepsilon_{t-7} - 0.171\varepsilon_{t-12}$	(3.20)

ARMA Equations and Estimations

Autoregressive moving average (ARMA) models assume that the future value of a time-series is governed by its historical values (the autoregressive component) and on previous residual values (the moving average component). A typical ARMA model takes the form of equation 3.15.

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t - w_1\varepsilon_{t-1} - w_2\varepsilon_{t-2} - \ldots - w_q\varepsilon_{t-q} \qquad (3.15)$$

Where:

Y_t	is the dependent variable at time t
Y_{t-1}, Y_{t-2}, and Y_{t-p}	are the lagged dependent variables
ϕ_0, ϕ_1, ϕ_2, and ϕ_p	are regression coefficients
ε_t	is the residual term
ε_{t-1}, ε_{t-2}, and ε_{t-p}	are previous values of the residual
w_1, w_2, and w_q	are weights.

Table 3.10 GARCH model # 1

GARCH MODEL # 1 (16,16)				
Dependent Variable: RETURNS				
Method: ML—ARCH				
Sample (adjusted): 4/22/2008 3/29/2012				
Included observations: 1028 after adjustments				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.001335	0.000660	2.023991	0.0430
AR(1)	–0.114106	0.016479	–6.924442	0.0000
AR(2)	–0.069719	0.013631	–5.114936	0.0000
AR(10)	–0.019143	0.007719	–2.480175	0.0131
AR(16)	–0.879046	0.013540	–64.92089	0.0000
MA(1)	0.120446	0.016908	7.123546	0.0000
MA(2)	0.065817	0.012455	5.284208	0.0000
MA(10)	–0.018657	0.008031	–2.323231	0.0202
MA(16)	0.897817	0.012306	72.95477	0.0000
	Variance Equation			
C	1.42E-05	4.07E-06	3.478155	0.0005
RESID(-1)^2	0.083038	0.012924	6.425038	0.0000
GARCH(-1)	0.896013	0.014846	60.35321	0.0000
R-squared	0.046906	Mean dependent var		0.000433
Adjusted R-squared	0.039424	S.D. dependent var		0.027925
S.E. of regression	0.027369	Akaike info criterion		–4.622048
Sum squared resid	0.763277	Schwarz criterion		–4.564436
Log likelihood	2387.733	Hannan-Quinn criter.		–4.600181
Durbin-Watson stat	1.987245			

Table 3.11 GARCH model # 2

GARCH MODEL # 2 (15,15)				
Dependent Variable: RETURNS				
Method: ML—ARCH				
Sample (adjusted): 4/21/2008 3/29/2012				
Included observations: 1029 after adjustments				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.001178	0.000697	1.691729	0.0907
AR(1)	–0.994472	0.026416	–37.64675	0.0000
AR(4)	–0.289867	0.024123	–12.01638	0.0000
AR(15)	0.153101	0.020739	7.382238	0.0000
MA(1)	1.016204	0.017594	57.75922	0.0000
MA(4)	0.322259	0.016612	19.39862	0.0000
MA(15)	–0.170083	0.014273	–11.91625	0.0000
	Variance Equation			
C	1.75E-05	5.09E-06	3.437281	0.0006
RESID(-1)^2	0.091396	0.016350	5.589899	0.0000
GARCH(-1)	0.883565	0.020421	43.26752	0.0000
R-squared	0.005203	Mean dependent var		0.000431
Adjusted R-squared	–0.000638	S.D. dependent var		0.027911
S.E. of regression	0.027920	Akaike info criterion		–4.594528
Sum squared resid	0.796679	Schwarz criterion		–4.546556
Log likelihood	2373.885	Hannan-Quinn criter.		–4.576321
Durbin-Watson stat	2.051188			

Using a correlogram as a guide in the training and the test sub-periods the below restricted ARMA models were selected to trade each spread. All coefficients were found to be significant at a 95 % confidence interval. Therefore, the null hypothesis that all coefficients (except the constant) are not significantly different from zero is rejected at the 95 % confidence interval (Table 3.9).

GARCH Equations and Estimations

Each of the GARCH models (16,16) and (15,15) are deemed stable and significant at a 95 % confidence level. Following the initial estimation of significant terms a squared residuals test, Jarque-Bera test and an ARCH test are all conducted to test the reliability of the residuals. For the sake of brevity, outputs from these tests are not included. These can be obtained on request from the corresponding author. Autocorrelation is absent from both models and as a result returns derived from each model were used as inputs during the training of the proposed PSO RBF Neural Network (Table 3.10).

Table 3.12 PSO RBF parameters

Characteristics	380 day Sliding Window	500 day Sliding Window
Iterations	100	100
Number of Particles	30	30
Inertia Constant (*w*)	Adaptive	Adaptive
Cognitive Acceleration Constant (C_1)	Adaptive	Adaptive
Social Acceleration Constant (C_2)	Adaptive	Adaptive
Maximum Velocity	2/Number of particles	2/Number of particles
Number of Neurons (1 hidden layer)	10	10
Constant Size of Hidden Layer	Yes	Yes
Input Nodes	59	59
Output Nodes	1	1

Table 3.13 Neural characteristics

Parameters	RBF	MLP
Learning algorithm	PSO	Gradient descent (Levenberg Marquardt variation)
Learning rate	Not Applicable	0.001
Momentum	Not Applicable	0.003
Iteration steps	100	5000
Initialization of weights	No Initialization Required. Deterministic method for finding them	$N(0,1)$
Input nodes	59	59
Hidden nodes (1 layer)	10	30
Output node	1	1

Table 3.14 Best weights obtained from the 380 day training window

MODEL 1		MODEL 2		MODEL 3		MODEL 4		MODEL 5	
Input Weights (59)	Neuron Weights (10)	Input Weights (59)	Neuron Weights (10)	Input Weights (59)	Neuron Weights (10)	Input Weights (59)	Neuron Weights (10)	Input Weights (59)	Neuron Weights (10)
1.9767	−0.0080	1.4888	0.8273	1.0584	−0.9852	0.8129	−0.0870	1.3663	0.0686
1.2275	−0.5489	1.5138	0.0075	1.2021	−0.1254	1.2912	0.1183	1.4555	0.0094
1.2147	0.2786	0.5293	0.2896	1.0792	−0.3226	0.8654	−0.0422	0.9676	−0.0908
1.2037	−0.9810	1.4406	−0.1087	1.4536	−0.1381	1.0240	−0.2030	1.4662	0.4919
0.5465	0.7071	1.5420	−0.4135	0.6496	1.1087	1.1924	−0.1213	1.1912	−0.4026
1.7877	−0.6872	1.4698	−1.0310	1.6624	0.3610	0.3727	−0.1485	0.8524	0.2388
1.2847	0.1213	0.2456	−0.5791	0.0439	0.6602	1.1430	−0.3494	0.6324	0.2551
0.9546	0.8774	1.1454	0.5787	1.7873	0.1058	1.2559	0.1910	0.4705	0.1141
1.3211	0.8775	0.5560	0.6667	0.9773	0.0443	0.9835	−0.1413	1.3414	−0.4315
1.6836	−0.7672	0.9511	−0.1383	0.7939	−0.5638	0.7293	0.8557	1.4916	−0.2457
0.7538	#N/A	1.2515	#N/A	1.2831	#N/A	0.9263	#N/A	0.2890	#N/A
1.1182	#N/A	0.8281	#N/A	0.3945	#N/A	1.4340	#N/A	0.8577	#N/A
1.3411	#N/A	0.6596	#N/A	1.6227	#N/A	0.6905	#N/A	0.6781	#N/A
1.0868	#N/A	1.1913	#N/A	1.6811	#N/A	1.0471	#N/A	0.8857	#N/A
0.1301	#N/A	0.8083	#N/A	0.3740	#N/A	1.7155	#N/A	0.3531	#N/A
1.3963	#N/A	1.5124	#N/A	0.4884	#N/A	1.2705	#N/A	0.9914	#N/A
1.5567	#N/A	1.6089	#N/A	0.7847	#N/A	0.6418	#N/A	0.7139	#N/A
1.5600	#N/A	1.1842	#N/A	1.3060	#N/A	0.6828	#N/A	1.0258	#N/A
1.4989	#N/A	1.3692	#N/A	0.6556	#N/A	0.9276	#N/A	0.9841	#N/A
0.6817	#N/A	0.3603	#N/A	1.3789	#N/A	0.9137	#N/A	0.5429	#N/A
1.1011	#N/A	1.6180	#N/A	1.0049	#N/A	0.9241	#N/A	1.5708	#N/A
0.2932	#N/A	0.4895	#N/A	1.4075	#N/A	1.0258	#N/A	0.7604	#N/A
1.2537	#N/A	0.6291	#N/A	0.0129	#N/A	1.0547	#N/A	0.5180	#N/A
1.4219	#N/A	1.4403	#N/A	0.2804	#N/A	1.4634	#N/A	1.2458	#N/A
1.4897	#N/A	0.8672	#N/A	1.3861	#N/A	0.9816	#N/A	1.0990	#N/A
1.6812	#N/A	0.6399	#N/A	1.6276	#N/A	1.1977	#N/A	0.9839	#N/A
1.4353	#N/A	1.2362	#N/A	0.9870	#N/A	0.8731	#N/A	1.4542	#N/A
1.9373	#N/A	1.2822	#N/A	0.4746	#N/A	1.2274	#N/A	0.5208	#N/A

0.7310	#N/A	0.9915	#N/A	1.6292	#N/A	0.9713	#N/A	1.2539	#N/A
1.0500	#N/A	1.7610	#N/A	0.9946	#N/A	0.7712	#N/A	0.4589	#N/A
0.5196	#N/A	1.1448	#N/A	0.8270	#N/A	1.1352	#N/A	0.8601	#N/A
1.1944	#N/A	0.3551	#N/A	1.0830	#N/A	1.6924	#N/A	0.4033	#N/A
1.7218	#N/A	1.5575	#N/A	1.2248	#N/A	0.5942	#N/A	1.5243	#N/A
1.4142	#N/A	0.3726	#N/A	0.4434	#N/A	0.9198	#N/A	1.1583	#N/A
1.0300	#N/A	0.4051	#N/A	0.3318	#N/A	1.7501	#N/A	0.5996	#N/A
1.3445	#N/A	0.4431	#N/A	0.4533	#N/A	0.9295	#N/A	0.9293	#N/A
1.2052	#N/A	0.5780	#N/A	0.8925	#N/A	1.3986	#N/A	0.9627	#N/A
1.5575	#N/A	0.4983	#N/A	0.8876	#N/A	1.2384	#N/A	0.5914	#N/A
1.5098	#N/A	0.6500	#N/A	1.4789	#N/A	1.0788	#N/A	0.3870	#N/A
0.5978	#N/A	1.2786	#N/A	0.5971	#N/A	0.7477	#N/A	1.0267	#N/A
0.6767	#N/A	1.5096	#N/A	1.5781	#N/A	1.1139	#N/A	1.0592	#N/A
1.3888	#N/A	0.8247	#N/A	0.8714	#N/A	1.3540	#N/A	0.9540	#N/A
0.9817	#N/A	1.3836	#N/A	0.9597	#N/A	1.8071	#N/A	0.9369	#N/A
0.8911	#N/A	0.4707	#N/A	1.4192	#N/A	0.4874	#N/A	1.1501	#N/A
1.2582	#N/A	1.0670	#N/A	1.3436	#N/A	0.5496	#N/A	1.1824	#N/A
1.2932	#N/A	1.3856	#N/A	0.6297	#N/A	1.1938	#N/A	1.1921	#N/A
0.2583	#N/A	0.7562	#N/A	0.6838	#N/A	1.5778	#N/A	1.1152	#N/A
1.0916	#N/A	0.5513	#N/A	0.8909	#N/A	1.0532	#N/A	0.9206	#N/A
2.0000	#N/A	1.3917	#N/A	0.2423	#N/A	1.2421	#N/A	0.5978	#N/A
1.3646	#N/A	1.0909	#N/A	1.1447	#N/A	0.9372	#N/A	0.9002	#N/A
0.5081	#N/A	1.2711	#N/A	0.6872	#N/A	1.2271	#N/A	0.7934	#N/A
0.3709	#N/A	1.2379	#N/A	1.1902	#N/A	0.8882	#N/A	1.1771	#N/A
0.5136	#N/A	1.5178	#N/A	1.5556	#N/A	1.7429	#N/A	0.6308	#N/A
1.0695	#N/A	0.5873	#N/A	1.0279	#N/A	0.9654	#N/A	1.0861	#N/A
1.1609	#N/A	0.5277	#N/A	0.6760	#N/A	0.7039	#N/A	1.0185	#N/A
0.6120	#N/A	1.5059	#N/A	1.6096	#N/A	1.1453	#N/A	1.4657	#N/A
1.5132	#N/A	1.3238	#N/A	0.9222	#N/A	0.7755	#N/A	0.6311	#N/A
1.3349	#N/A	1.3526	#N/A	1.4649	#N/A	1.2904	#N/A	0.1833	#N/A
0.6911	#N/A	0.4935	#N/A	0.3462	#N/A	0.7502	#N/A	0.6867	#N/A
1.8144	#N/A	0.6191	#N/A	1.0211	#N/A	0.9639	#N/A	1.5926	#N/A

Table 3.15 Best weights obtained from the 500-day training window

MODEL 1		MODEL 2		MODEL 3		MODEL 4		MODEL 5	
Input Weights (59)	Neuron Weights (10)	Input Weights (59)	Neuron Weights (10)	Input Weights (59)	Neuron Weights (10)	Input Weights (59)	Neuron Weights (10)	Input Weights (59)	Neuron Weights (10)
0.9070	0.4015	0.6819	–0.0703	0.9070	0.4015	1.7253	0.4609	0.8720	–0.1244
1.4054	–0.1953	1.6728	0.2570	1.4054	–0.1953	1.3730	0.1288	1.2579	–0.0285
0.7730	–0.1110	0.7241	–0.2934	0.7730	–0.1110	0.7039	0.1699	0.8418	1.0970
1.3742	–0.4231	1.2879	–0.1102	1.3742	–0.4231	0.5764	0.2313	0.8797	0.1393
0.9698	–0.0118	1.3060	–0.3200	0.9698	–0.0118	1.0941	–0.1848	1.4859	–0.7727
1.1820	–0.4267	0.7503	0.6308	1.1820	–0.4267	0.9311	–0.5983	0.6381	0.2812
0.9609	0.0749	0.8368	0.8377	0.9609	0.0749	1.0947	–0.1012	1.0521	–0.3009
1.1526	–0.1345	1.5200	–0.8396	1.1526	–0.1345	1.2402	–0.5187	0.7102	–0.3916
0.8540	–0.0502	0.2606	–0.0575	0.8540	–0.0502	0.9163	0.3221	0.9099	–0.6510
0.7939	0.8447	1.2288	0.0940	0.7939	0.8447	1.0475	0.0367	0.4763	0.5374
0.7716	#N/A	0.3574	#N/A	0.7716	#N/A	1.2223	#N/A	1.0691	#N/A
0.6642	#N/A	0.9626	#N/A	0.6642	#N/A	1.4063	#N/A	0.5522	#N/A
0.9845	#N/A	0.9560	#N/A	0.9845	#N/A	0.8503	#N/A	0.9638	#N/A
0.8801	#N/A	0.6218	#N/A	0.8801	#N/A	0.5076	#N/A	1.0708	#N/A
1.3758	#N/A	0.5021	#N/A	1.3758	#N/A	1.3177	#N/A	0.7855	#N/A
0.9750	#N/A	0.5827	#N/A	0.9750	#N/A	0.7341	#N/A	0.9721	#N/A
0.5542	#N/A	1.5309	#N/A	0.5542	#N/A	0.6713	#N/A	0.5005	#N/A
0.8604	#N/A	1.6899	#N/A	0.8604	#N/A	0.9526	#N/A	1.2341	#N/A
1.3104	#N/A	0.8265	#N/A	1.3104	#N/A	1.4329	#N/A	0.8003	#N/A
1.3563	#N/A	1.2091	#N/A	1.3563	#N/A	0.3215	#N/A	0.4206	#N/A
0.8165	#N/A	0.2437	#N/A	0.8165	#N/A	1.3214	#N/A	1.3452	#N/A
1.7920	#N/A	1.2195	#N/A	1.7920	#N/A	1.2343	#N/A	1.0293	#N/A
1.3503	#N/A	0.6166	#N/A	1.3503	#N/A	0.7596	#N/A	0.7727	#N/A
1.3684	#N/A	1.0831	#N/A	1.3684	#N/A	0.8708	#N/A	0.6222	#N/A
1.0686	#N/A	1.4177	#N/A	1.0686	#N/A	0.7813	#N/A	0.7203	#N/A
1.6583	#N/A	0.9189	#N/A	1.6583	#N/A	1.5476	#N/A	1.3443	#N/A
0.7123	#N/A	1.5404	#N/A	0.7123	#N/A	1.4377	#N/A	1.3139	#N/A
0.6076	#N/A	0.9581	#N/A	0.6076	#N/A	0.8825	#N/A	1.5451	#N/A

0.8656	#N/A	1.5559	#N/A	0.8656	#N/A	1.9318	#N/A	0.9862	#N/A
0.2273	#N/A	0.6753	#N/A	0.2273	#N/A	1.0459	#N/A	1.1460	#N/A
1.0345	#N/A	0.2068	#N/A	1.0345	#N/A	0.7238	#N/A	1.2063	#N/A
0.7623	#N/A	0.4673	#N/A	0.7623	#N/A	0.9009	#N/A	0.9245	#N/A
0.7128	#N/A	0.6491	#N/A	0.7128	#N/A	0.8554	#N/A	1.7615	#N/A
1.2197	#N/A	1.1203	#N/A	1.2197	#N/A	0.8877	#N/A	0.8165	#N/A
0.7839	#N/A	1.5416	#N/A	0.7839	#N/A	0.9554	#N/A	0.6599	#N/A
0.5538	#N/A	1.3938	#N/A	0.5538	#N/A	1.0397	#N/A	1.3623	#N/A
1.5902	#N/A	0.6421	#N/A	1.5902	#N/A	1.1515	#N/A	1.3136	#N/A
0.9899	#N/A	0.6421	#N/A	0.9899	#N/A	0.0879	#N/A	0.9088	#N/A
0.8756	#N/A	1.7203	#N/A	0.8756	#N/A	0.4917	#N/A	0.5814	#N/A
0.3445	#N/A	0.5204	#N/A	0.3445	#N/A	1.3025	#N/A	1.1367	#N/A
0.9331	#N/A	1.7534	#N/A	0.9331	#N/A	0.5988	#N/A	0.7868	#N/A
0.9731	#N/A	0.6174	#N/A	0.9731	#N/A	0.9355	#N/A	1.2207	#N/A
0.9116	#N/A	0.8339	#N/A	0.9116	#N/A	0.2439	#N/A	1.1726	#N/A
0.7557	#N/A	0.9006	#N/A	0.7557	#N/A	1.3068	#N/A	1.0262	#N/A
1.5539	#N/A	0.8249	#N/A	1.5539	#N/A	1.5354	#N/A	1.0560	#N/A
0.3683	#N/A	0.3886	#N/A	0.3683	#N/A	1.0518	#N/A	0.8511	#N/A
0.5836	#N/A	1.0118	#N/A	0.5836	#N/A	0.8430	#N/A	1.1025	#N/A
1.6273	#N/A	0.4865	#N/A	1.6273	#N/A	0.8351	#N/A	0.9436	#N/A
1.4291	#N/A	0.6186	#N/A	1.4291	#N/A	0.7076	#N/A	0.3921	#N/A
1.1770	#N/A	0.2630	#N/A	1.1770	#N/A	0.5120	#N/A	1.3862	#N/A
0.6836	#N/A	1.2072	#N/A	0.6836	#N/A	0.7497	#N/A	1.0211	#N/A
1.4776	#N/A	1.4199	#N/A	1.4776	#N/A	1.4651	#N/A	1.2946	#N/A
0.3409	#N/A	1.0153	#N/A	0.3409	#N/A	1.1771	#N/A	1.2813	#N/A
1.1485	#N/A	1.2068	#N/A	1.1485	#N/A	0.3835	#N/A	1.5403	#N/A
0.9180	#N/A	0.3319	#N/A	0.9180	#N/A	0.5203	#N/A	0.8887	#N/A
1.2806	#N/A	0.6448	#N/A	1.2806	#N/A	0.8930	#N/A	0.5769	#N/A
0.8480	#N/A	1.5232	#N/A	0.8480	#N/A	0.6923	#N/A	1.4762	#N/A
1.3163	#N/A	1.0349	#N/A	1.3163	#N/A	1.2372	#N/A	0.2311	#N/A
0.6303	#N/A	1.2965	#N/A	0.6303	#N/A	1.2530	#N/A	1.3228	#N/A
0.7381	#N/A	0.5941	#N/A	0.7381	#N/A	1.0344	#N/A	1.3092	#N/A

Observation
The AR(1), AR(2), AR(10), AR(16), MA(1), MA(2), MA(10) and MA(16) terms are all deemed significant at a 95 % confidence level. The model is also deemed stable due to the fact that the sum of GARCH(-1) and RESID(-1)^2 is less than 1. In this case it is, 0.896013 + 0.083038 = 0.979 (Table 3.11).

Observation
The AR(1), AR(4), AR(15), MA(1), MA(4), and MA(15) terms are all deemed significant at a 95 % confidence level. The model is also deemed stationary due to the fact that the sum of GARCH(-1) and RESID(-1)^2 is less than 1. In this case it is, 0.883565 + 0.091396 = 0.974961.

PSO Parameters

See Tables 3.12 and 3.13

Best Weights over the Training Windows

See Tables 3.14 and 3.15

References

1. Dunis, C. L., Laws, J., & Evans, B. (2005). Modelling and trading the gasoline crack spread: A non-linear story. *Derivatives Use, Trading & Regulation, 12*(1-2), 126–145.
2. Butterworth, D., & Holmes, P. (2002). Inter-market spread trading: Evidence from UK Index Futures Markets. *Applied Financial Economics., 12*(11), 783–791.
3. Dunis, C. L., Laws, J., & Evans, B. (2006). Modelling and trading the soybean-oil crush spread with recurrent and higher order networks: A comparative analysis. *Neural Network World, 13*(3/6), 193–213.
4. Dunis, C. L, Laws, J; Middleton, P.W. (2011). Modelling and Trading the Corn/Ethanol Crush Spread with Neural Networks. *CIBEF Working Paper*, Liverpool Business School, Available at www.cibef.com.
5. Chen, L. H. C., Finney, M., & Lai, K. S. (2005). A threshold cointegration analysis of asymmetric price transmission from crude oil to gasoline prices. *Economics Letters, 89*, 233–239.
6. Enders, W., & Granger, C. W. J. (1998). Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates. *Journal of Business and Economic Statistics, 16*(3), 304–311.

7. Kaastra, I., & Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing, 10*(3), 215.
8. Tsai, C. F., & Wang, S. P. (2009). Stock price forecasting by hybrid machine learning techniques. *Proceedings of the International Multi-Conference of Engineers and Computer Scientists, 1*, 755–760.
9. Chang, P. C., Wang, Y. W., & Yang, W. N. (2004). An investigation of the hybrid forecasting models for stock price variation in Taiwan. *Journal of the Chinese Institute of Industrial Engineering., 21*(4), 358–368.
10. Thawornwong, S., & Enke, D. (2004). The adaptive selection of financial and economic variables for use with Artificial Neural Networks. *Neurocomputing., 56*, 205–232.
11. Cortez, P., Rocha, M; Neves, J. (2001) Evolving Time Series Forecasting Neural Network Models. *In Proc. of the 3rd Int. Symposium on Adaptive Systems*: *Evolutionary Computation and Probabilistic Graphical Models* (*ISAS 2001*), (pp. 84–91).
12. Mettenheim, V. H. J; Breitner, M. H. (2012). Forecasting daily highs and lows of liquid assets with neural networks. *Operations Research Proceedings*, Selected Papers of the Annual International Conference of the German Operations Research Society, Hannover.
13. Dunis, C. L., Laws, J., & Karathanassopoulos, A. (2011). Modelling and trading the Greek Stock Market with Mixed Neural Network Models. *Applied Financial Economics*., *21(23)*, 1793–1808
14. Sopena, J. M., Romero, E., and Alquezar, R. (1999). Neural networks with periodic and monotonic activation functions: a comparative study in classification problems. In Proceedings of the 9th International Conference on Artificial Neural Networks.
15. Park, J. W., Harley, R. G., & Venayagamoorthy, G. K. (2002). Comparison of MLP and RBF Neural Networks using deviation signals for on-line identification of a synchronous generator. *Proceedings of Power Engineering Society Winter Meeting, IEEE, 1*(27–31), 274–279.
16. Kennedy, J., & Eberhart, R. (1995). Particle Swarm Optimization. *Proceedings of the IEEE International Conference on Neural Networks, 4*, 1942–1948.
17. Chen, Z., & Qian, P. (2009). Application of PSO-RBF Neural Network in Network Intrusion Detection. *Intelligent Information Technology Application, 1*, 362–364.
18. Konstantinos, T.; Parsopolous, E.; Vrahatis, M. N. (2010). Particle swarm optimization and intelligence: Advances and applications. *IGI Global*, 149–164.
19. Sermpinis, G., Theofilatos, K., Karathanasopoulos, A., Georgopoulos, E., & Dunis, C. (2013). Forecasting foreign exchange rates with Adaptive Neural Networks using radial-basis functions and Particle Swarm Optimization. *European Journal of Operational Research, 225*(3), 528–540.
20. Li, J.; Xiao, X (2008). Multi-swarm and multi-best particle swarm optimization algorithm. In *Proceedings of the 7th World Congress on Intelligent Control and Automation*, China. pp. 6281–6286.

21. Mohaghegi, S., Valle, Y., Venayagamoorthy, G., Harley, R. (2005). A comparison of PSO and backpropagation for training RBF Neural Networks for identification of a power system with Statcom. *In Proc. IEEE Swarm Intelligence Symposium*, pp. 381–384.
22. Eberhart, R.C., Simpson, P.K., Dobbins, R.W., 1996. Computational Intelligence PC Tools. Academic Press Professional. Boston, MA.
23. Dunis, C. L. (1996). The economic value of neural network systems for exchange rate forecasting. *Neural Network World, 1*, 43–55.
24. Makridakis, S. (1989). Why combining works? *International Journal of Forecasting, 5*, 601–603.
25. Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting, 5*, 559–583.
26. Newbold, P., & Granger, C. W. J. (1974). Experience with forecasting univariate time series and the combination of forecasts (with discussion). *Journal of Statistics, 137*, 131–164.
27. Palm, F. C., & Zellner, A. (1992). To combine or not to combine? Issues of combining forecasts. *Journal of Forecasting., 11*, 687–701.

4

GEPTrader: A New Standalone Tool for Constructing Trading Strategies with Gene Expression Programming

Andreas Karathanasopoulos, Peter W. Middleton, Konstantinos Theofilatos, and Efstratios Georgopoulos

1 Introduction

Many technical trading algorithms have been introduced in the past decade with Artificial Neural Networks (ANNs) being the dominant machine learning technique for financial forecasting applications [1]. However, these techniques present certain drawbacks such as, overfitting issues, a data snooping bias, and a curse of dimensionality [2]. Moreover, despite the vast number of different financial forecasting algorithms that are continuously being published, only a

A. Karathanasopoulos (✉)
American University of Beirut, Beirut, Lebanon
e-mail: ak152@aub.edu.lb

P.W. Middleton
University of Liverpool, Liverpool, England
e-mail: peter.william.middleton@gmail.com

K. Theofilatos
Pattern Recognition Laboratory, Department of Computer Engineering and Informatics, University of Patras, 26500 Patras, Greece
e-mail: theofilk@ceid.upatras.gr

E. Georgopoulos
Technological Educational Institute of Peloponnese, Kalamata, Greece
e-mail: sfg@teikal.gr

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets,*
New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0_4

few of them are being combined to construct efficient trading strategies and even fewer are provided to the masses through user-friendly graphical user interfaces.

This chapter introduces GEPTrader as a new standalone free tool for constructing financial forecasting models using a variation of the Gene Expression Programming algorithm (GEP) [3]. GEP algorithm belongs to the wider category of evolutionary and genetic programming algorithms. Evolutionary and genetic programming tools are becoming popular forecasting tools with an increasing number of financial applications [4, 5].

To prove the superiority of GEPTrader in financial forecasting and trading tasks we applied it to the task of forecasting and trading the SPDR Dow Jones Industrial Average ETF Trust (DIA-US), the SPDR S&P 500 ETF Trust (SPY-US) and the PowerShares QQQ Trust (QQQ-US) EFTs. Its performance is benchmarked with a simple random walk (RW) model, a Moving Average Convergence Divergence (MACD) model, a Genetic Programming (GP) algorithm, a Multi-Layer Perceptron (MLP), a Recurrent Neural Network (RNN) and a Higher Order Neural Network (HONN). All models are evaluated in terms of statistical accuracy and trading profitability.

The models studied were deployed to forecast the one-day-ahead return of each of the aforementioned ETF time-series. Based on the sign of the forecasted return, a trading signal was generated. In order to further improve the trading performance of the proposed approach, three trading strategies based on one-day-ahead volatility forecasts of the series under study and the strength of the forecasts were introduced. These two derived trading strategies were combined to improve the trading performance of the best performing models and to reduce the risk from the trading signals generated. This resulted in a more elaborate trading simulation that displayed superior performance.

The rest of the chapter is organized as follows: Section 2 reviews existing literature concerning GP and GEP and their applications to financial forecasting and trading. Section 3 presents the datasets analyzed while Section 4 describes the methodology of the proposed algorithm and the GEPTrader standalone tool. Section 5 introduces the benchmark models and discusses experimental results of the proposed model relative to the benchmark models. Finally, Section 6 provides some concluding remarks, research limitations and future areas of research.

2 Literature Review

Genetic Programming and Its Applications to Financial Forecasting

The first form of GEP algorithms is the GP, which has had numerous empirical applications over the years. GP is an evolutionary algorithm based on the Darwinian principle of the survival of the fittest. It models regression and classification problems' solutions using a tree-based typology. These tree-based structures represent models of input–output relations and are evolved in order to produce new solutions until a solution of acceptable performance is found or other more elaborate termination criteria are reached.

Kaboudan [6] successfully applied GP in predicting the daily highest and lowest stock prices of six US stocks while Alvarez-Diaz et al. [7] employed the same algorithm to the task of forecasting two exchange rates. Esfahanipour and Mousavi [8] deployed a GP algorithm to extract trading rules using three signals of buy, sell and no trade. Their model was applied to the Tehran Stock Exchange indices taking transaction costs into consideration. Experimental results from this study indicate the superiority of GP techniques over fundamental trading strategies such as the buy-and-hold technique. Vasilakis et al. [9] present a novel GP variation specific for financial time-series prediction and incorporated it to successfully model and trade foreign exchange rate time-series.

Manahov et al. [10] introduced a new trading method called the Strongly Typed Genetic Programming (STGP)-based trading algorithm. The STGP-based trading algorithm produces one-day-ahead return forecasts for groups of artificial traders with different levels of intelligence and different group sizes. This method was applied to short-term trading tasks presenting improved empirical metrics. The authors also deployed this method to test some theoretical financial hypothesis presenting some evidence for supporting the Hayek Hypothesis [11].

Gene Expression Programming and Previous Applications

The GEP algorithm is a domain-independent technique that runs in various environments. These environments are structured in a manner that approximates problems in order to produce accurate forecasts. GEP is based on the Darwinian principle of reproduction and survival of the fittest. It applies biological genetic operations such as crossover recombination and mutation to identify complex non-linear and non-stationary patterns. Koza [12] and [13] was one of the first to establish that Evolutionary Algorithms address and quantify complex issues as an automated process via programming. This enables computers to process and solve problems using a stochastic process based on evolution.

In financial forecasting although there are several applications of Neural Networks (NNs), the empirical evidence of GEP is quite limited with the notable exceptions of Divsalar et al. [14] and Sermpinis et al. [15]. This can be explained by the complexity of the algorithm compared to NNs (see [16]). However, in the past three years several applications of GEP methods in financial forecasting and trading have been proposed. For instance, Chen et al. [17] use GEP to extract trading rules that are responsible for locating the optimal buying and selling timings. These trading rules from different time-series are then combined to dynamically form a portfolio using Sortino ratios. The overall model was successfully evaluated when applied on constructing portfolios from mutual funds. A similar approach was followed by Lee et al. [18] and when it was applied in Taiwan stock market the empirical results produced highly profitable portfolios of stocks. Algieth et al. [19] performed an empirical comparison of GEP, GP and genetic algorithms over different horizons focusing on short-term periods such as five-days and medium-term 56-days. This analysis used several equity time-series indicating the superiority of GEP technique in terms of predictive accuracy.

Despite its limited applications on financial forecasting and trading, GEP has been applied successfully in other fields of science, such as mining and computing [20, 21]. Here, we intend to expand the financial literature and provide empirical evidence of its benefits when applied to the forecasting of financial time-series.

3 Dataset

In this study, seven forecasting models are employed to the task of forecasting and trading the one-day-ahead logarithmic returns of the DIA, SPY and QQQ ETFs.[1] The DIA, the SPY and the QQQ ETF are designed to mimic the Dow Jones Industrial Average, the S&P 500 and NASDAQ 30 stock market indices respectively. Exchange-Traded Funds (ETFs) offer investors the opportunity to be exposed to an index without having to trade and own every stock in the index. This helps avoid transaction costs associated with trading individual stocks as well as the costs associated with frequent rebalances of the index.

Each model is trained using a rolling window process with the initial in sample period running from 2 September 2002 to 31 December 2008. The models are validated from 2 January 2009 to 31 December 2012. During the training period each of the parameters and weights are optimized and carried forward to be tested on 'unseen data', in a validation period, in order to test each model's performance. This estimation period is then rolled forward each year. For example, initially the models were be trained from 2

[1] ETFs are investment funds that are designed to replicate stock market indices.

September 2002 to 31 December 2008 and validated from 2 January 2009 to 31 December 2012. Then the in-sample period is rolled forward one year (2 January 2003 to 31 December 2010) and the forecasting models are revalidated from 3 January 2011 to 30 December 2011. This rolling forward procedure is conducted three times to complete the modelling and trading process on 31 December 2012.

4 GEPTrader

Proposed Algorithm

GEP models are symbolic strings of a fixed length resembling the biological definitions of chromosome/genotype (genome) of an organism. They are encoded as non-linear entities of different sizes and shapes determining an organism's fitness. GEP chromosomes consist of multiple genes with equal lengths across the structure of the chromosome. Each gene is comprised of a head (detailing symbols specific to functions and terminals) and a tail (includes only terminals). These can be represented mathematically by the following equation:

$$t = (n-1)h + 1 \tag{4.1}$$

Where:

- h the head length of the gene.
- t the tail length of the gene.
- n total number of arguments within the function (maximum arity).

The set of terminals included within both the heads and tails of the chromosomes consist of constants and specific variables. Each gene is equal and fixed in size and they hold the capacity to code for multiple and varied expression trees (ETs). The structure of a GEP is able to cope in circumstances when the first element of a gene is terminal producing a single node as well as when multiple nodes ('sub-trees' reproduced by functions) are produced in search for eventual terminality. In GEP, valid ETs are always generated while in GP this is not guaranteed. Each gene encodes an ET and in situations where multiple generations arise, GEP codes for sub-ETs with interlinking functions enable reproduction of generations. The parameters of our GEP algorithm

Table 4.1 GEPTrader default parameters

Parameter	Value
Head length (h)	30
Initial population size	1000
Tournament size	20
Maximum number of generations	100,000
Initial mutation probability	10 %
Final mutation probability	1 %
Crossover probability	90 %
Local optimization of the constants Probability	5 %

are based on the guidelines of Sermpinis et al. [15] and Ferreira [16] and the most significant of them are described in Table 4.1.

In the beginning of the GEP's evolutionary process, it randomly generates an initial population from populations of individuals and all succeeding populations are spawned from this initial population. In the spawning of new generations genetic operators evolve each of the individuals by 'mating' them with other individuals in the population using a two-point crossover operator. With a reduced probability the mutation operator is also applied to locally modify a gene of a given solution. The mutation probability starts with an increased value to perform a global search in initial iterations, and it is linearly decreased to a small value to perform a local search in the final steps of the algorithm. Another variability operator, which is introduced in the proposed approach, is the local optimization of the constants operator. This operator searches for constant genes and perform local optimization for their values to improve the overall performance of the model.

In the following step of GEP expression, trees are created from the chromosomes of the population. The structure of each ET is in such a way that the root or the first node corresponds with beginning of each gene. The resulting offspring evolved from the first node is dependent on the number of arguments. In this process of evolution the functions may have numerous arguments however the terminals have a parity of zero. Each of the resulting offspring's characteristics is populated in nodes ordered from left to right. This process is concluded once terminal nodes are reached. Later the initial population is generated and the resulting ETs are developed. This is explained in detail by Ferreira [16]. In order to create an accurate model suited to our forecasting requirements it is imperative that a function that minimizes error and improves accuracy is used. However, it is even more important to optimize the annualized return of the overall model and minimize the complexity of the model. Thus for our approach we utilized the following fitness function:

$$\text{Fitness Function} = \text{Annualized\_Return} + 1/MSE - \text{Normalized\_Solution\_Size} \quad (4.2)$$

Where:

Normalized_Solution_Size is the number of nodes in the ET form of the solution/the maximum number of nodes, which is the sum of head and tail lengths as described in equation 4.1.

The main principal during the process of evolution is the generation of offspring from two superior individuals to achieve 'elitism'. As a consequence, the best individuals from the parent generation produce offspring with the most desirable traits whilst the individuals with less desirable traits are removed. On this basis our model minimizes error and maintains superior forecasting abilities. As explained in greater detail by Ferreira [16], elitism is the cloning of the best chromosome(s)/individual(s) to the next population (also called generation). The role of 'elitism' (via suited genetic operators) enables the selection of fitter individuals without eliminating the entire population. The selection of individuals based on their 'fitness' is carried out during the 'tournament' selection for reproduction and modification. This process selects the individuals at random with the superior ones being chosen for genetic modification in order to create new generations. In the reproduction of future generations, we apply the mutation and recombination genetic operators. Then the tournament losers are replaced with the new individuals created by the reproduction in the population. A check is made to determine whether the termination criterion is fulfilled, if it is not we return to the second step. As a termination criterion we used a maximum number of generations during which the GEP was left to run. As a result the best individual found during the evolution process is provided.

GEPTrader Graphical User Interface

The GEPTrader tool is a standalone tool implemented in Java programming language that aims to provide the proposed GEP variation through a user-friendly interface (see Fig. 4.1) to analysts and forecasters. Historical data can be imported in GEPTrader's GUI using comma separated files, and the parameters of the GEP variation can be further tuned. Moreover, GEPTrader supports multi-threading parallel executions enabling the users to select the number of threads that will be used during the training phase. When training is finished the users are provided with the extracted models in terms of a mathematical input output equations as well as their performance metrics and some graphical representations of the real versus the estimated values.

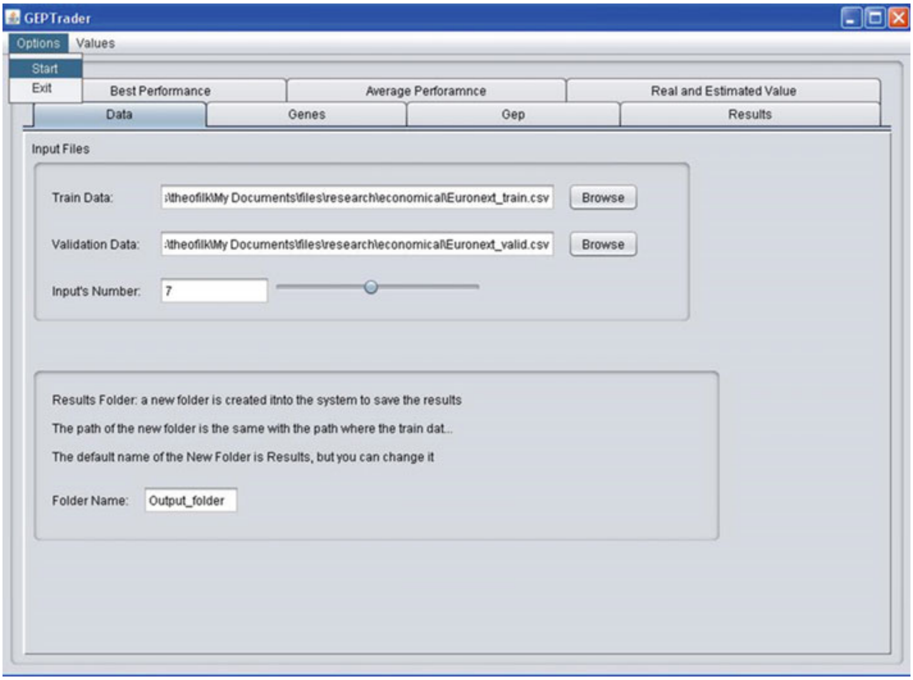


Fig. 4.1 GEPTrader Graphical User Interface

5 Empirical Results

Benchmark Models

A RW and a MACD model were used as linear benchmarks to the GEP algorithm. Three neural network models and a GP algorithm acted as a non-linear benchmark with all deployed models being well documented in the literature. NNs usually consist of three or more layers. The first layer is called the input layer (the number of its nodes corresponds to the number of independent variables). In this study the inputs were selected among the first 12 autoregressive lags of the forecasting series. The specific choice of each set of inputs was based on a sensitivity analysis in the in-sample period. The last layer is called the output layer (the number of its nodes corresponds to the number of response variables). An intermediary layer of nodes, the hidden layer, separates the input from the output layer. Its number of nodes defines the level of complexity the model is capable of fitting. Each node generally connects one layer to all the other nodes of the next layer. The training of the network (which is the adjustment of its weights in the way that the network maps the input value of the training data to the corresponding output value) starts with randomly chosen weights and proceeds by applying a learning algorithm called backpropagation of errors [22].The iteration length is optimized by minimizing the MSE in a subset of in-sample dataset (the last year of the in-sample period each time). The most popular architecture NN model is the MLP. RNNs have the ability to embody short-term memory and past errors while HONNs are able to capture higher order correlations (up to the order three or four) within the dataset.

GP are algorithms that evolve algebraic expressions that enable the analysis/optimization of results in a 'tree like structure'. A complete description of GP is provided by Koza [13].

Statistical Performance

In Table 4.2, the statistical performance in the out-of-sample period of all models is presented for the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE) and Theil-U statistics. Interpretation of results is such that, the lower the output, the better the forecasting accuracy of the model concerned. The Pesaran and Timmermann [23] (PT) test examines whether the directional movements of the real and forecast values are in step with one another. Furthermore, it checks how well rises and falls in the forecasted value follow the actual rises and falls of the time-series. The null hypothesis is such that the model under study has 'no predictive power' when forecasting the ETF return series. The Diebold and Mariano [24] (DM) statistic for predictive accuracy statistic tests the null hypothesis of equal predictive accuracy. Both the DM and the PT tests follow the standard normal distribution.

From Table 4.2 we note that GEP outperforms all benchmarks for all of the statistical measures retained. The HONN model presents the second best performance while the GP algorithm produces the third most statistically accurate forecasts. The PT statistics indicate that all non-linear models under study are able to forecast accurately the directional movements of the three ETF return series while the DM statistics confirm the statistical superiority of the GEP forecasts. On the other hand, the two statistical benchmarks seem unable to provide statistically accurate forecasts for the series and period under study.

Trading Performance

The trading performance of our models in the out-of-sample period is presented in Table 4.3.

Table 4.2 Out-of-sample statistical performance

		RW	MACD	GP	MLP	RNN	HONN	GEP
DIA	RMSE	1.027	0.239	0.073	0.136	0.085	0.069	0.057
	MAE	0.833	0.162	0.059	0.097	0.077	0.054	0.046
	Theil-U	0.989	0.756	0.634	0.673	0.641	0.620	0.601
	PT	0.02	0.35	5.89	5.02	5.68	6.44	6.99
	DM	−14.85	−10.47	−4.62	−5.91	−5.73	−4.45	–
SPY	RMSE	1.021	0.287	0.068	0.143	0.072	0.064	0.053
	MAE	0.831	0.195	0.055	0.103	0.061	0.052	0.042
	Theil-U	0.977	0.792	0.621	0.686	0.672	0.613	0.597
	PT	0.05	0.30	6.48	4.85	5.79	6.52	7.11
	DM	−14.56	−11.41	−5.59	−6.65	−5.97	−5.08	–
QQQ	RMSE	1.022	0.295	0.069	0.127	0.071	0.067	0.055
	MAE	0.834	0.204	0.058	0.088	0.059	0.055	0.044
	Theil-U	0.987	0.799	0.632	0.658	0.644	0.616	0.599
	PT	0.08	0.25	6.14	5.19	5.73	6.50	7.08
	DM	−15.31	−12.04	−5.42	−5.81	5.70	−4.77	–

By observation of Table 4.3, we note that our GEP variation clearly outperforms its benchmarks. In the next section, two trading strategies are introduced to further increase the trading performance of GEP.

In order to further improve the trading performance of our models we introduce leverage based on one-day-ahead volatility forecasts. The intuition for the strategy is to exploit the changes in the volatility. As a first step we forecast the one-day-ahead exchange rate volatility with a GARCH [25], GJR, RiskMetrics [26] and EGARCH model [27] in the test and validation subperiods. Each of these periods is then split into six sub-periods called regimes, ranging from regimes of extremely low volatility to regimes of extremely high volatility. Periods with different volatility levels are classified in the following way: first the average (μ) difference between the actual volatility in day t and the forecast for day $t + 1$ and its 'volatility' (measured in terms of standard deviation (σ)) are calculated; those periods where the difference is between μ plus one σ are classified as 'Lower High Vol. periods'. Similarly, 'Medium High Vol.' (between $\mu + \sigma$ and $\mu + 2\ \sigma$) and 'Extremely High Vol.' (above $\mu + 2\ \sigma$ periods). Periods with low volatility are also defined following the same approach, but with a minus sign. For each sub-period a leverage is assigned starting with 0 for periods of extremely high volatility to a leverage of 2.5 for periods of extremely low volatility.

The second trading strategy is based on the absolute values of our forecasts. The GEP algorithm is forecasting the one-day-ahead returns for the three indices. For this strategy, a directional decision is taken based on the strength of the forecast. Only forecasts with stronger conviction are used to make trading decisions. Furthermore, these trades are held for that index until the sign

Table 4.3 Out-of-sample trading performance

		MACD	GP	MLP	RNN	HONN	GEP
DJIA	Information ratio	–0.17	0.76	0.63	0.67	0.79	0.94
	Annualized return (including costs)	2.16 %	15.22 %	10.01 %	12.94 %	15.92 %	23.55 %
	Maximum drawdown	–23.24 %	–22.04 %	–20.91 %	–19.95 %	–21.42 %	–17.87 %
S&P 500	Information ratio	–0.24	0.85	0.51	0.74	0.92	1.28
	Annualized return (including costs)	–2.12 %	19.57 %	7.65 %	16.79 %	19.67 %	26.68 %
	Maximum drawdown	–36.16 %	–30.12 %	–35.92 %	–34.16 %	–32.54 %	–26.56 %
NASDAQ 100	Information ratio	–0.38	0.78	0.70	0.76	0.86	1.21
	Annualized return (including costs)	–3.65 %	17.57 %	14.51 %	16.42 %	19.54 %	24.02 %
	Maximum drawdown	–32.85 %	–28.19 %	–30.81 %	–31.91 %	–27.91 %	–24.41 %

Table 4.4 Trading Strategies

	GEP-DJIA with Time-Varying Leverage Trading Strategy	GEP-S&P500 with Time-Varying Leverage Trading Strategy	GEP-NASDAQ 100 with Time-Varying Leverage Trading Strategy	GEP-The Strongest Signal Trading Strategy
Information Ratio	1.03	1.32	1.29	1.96
Annualised Return (including costs)	26.85 %	28.14 %	26.98 %	34.54 %
Maximum Drawdown	–18.95 %	–28.95 %	–22.95 %	–25.81 %

of the forecast changes, which may mean assuming the same position for a number of days limiting the number of transactions.

In Table 4.4, we report the performance for our trading strategies.

We note that all trading strategies were successful but the combined trading strategy using the level of confidence approach was significantly more profitable.

6 Conclusions

In this study, a GEP variation was applied to the task of forecasting and trading the DJIA, S&P 500 and NASDAQ 100 indices. It was benchmarked against several non-linear models. The GEP forecasts outperformed its benchmarks in terms of annualized return and information ratio.

Although empirical research is limited in the area of forecasting of time-series using GEP compared to NNs and GP, here we find that it is superior. GP classifies its individuals as non-linear comprising of different shapes and sizes (tree like structures). On the other hand, GEP also classifies individuals as symbolic strings of fixed size (i.e. chromosomes). Furthermore, it clearly distinguishes the differences between the genotype and the phenotype of individuals within a population. Ferreira [16] argues that GEP represents not only an individual's genotype, in the form of chromosomes, but also its phenotype as a tree like structure of expressions in order to establish fitness. Compared with NNs, GEP does not have a risk of getting trapped in local optima and it is able to reach the optimal solution quicker. The findings of this chapter support these arguments as the results show that the modified GEP algorithm outperforms all other models in terms of trading efficiency.

This trading performance was further enhanced by the introduction of two new trading strategies. The present study indicates that even the more complicated modelling techniques can be a profitable tool for investors as it enables forecasters to capture non-linear relationships that are otherwise ignored using more traditional methodologies. Furthermore, forecasts from GEP models, in particular, can be used to deduce additional trading rules in order to extract more value from the outputs.

Ultimately, this empirical application should go a step further in proving that stochastic models such as the proposed GEP model presented here is beneficial to the investment process. Furthermore, these models can be used to by the growing number of quantitative fund managers and academics. They enable users to experiment beyond the bounds of traditional statistical and neural network models.

The main limitation of the present approach is the fact that this approach and tool does not take into account the dynamic nature of financial time-series. Specifically this approach assumes that a single model of good performance exists for all the historical data of every time-series. However, this is not the case and the trading results could be significantly improved if a sliding window approach is incorporated into the GEPTrader tool. Moreover, even if GEP algorithm performs feature selection by nature, its performance will be further improved if it is combined with other feature selection mechanisms.

Finally, we strongly believe that these newly introduced algorithms should be provided as tools in the form of a graphical user interface to reach a wider audience. This will help remove the current barriers and knowledge gap that is present in the area of AI and in particular evolutionary algorithms.

References

1. Zhang, G. P. (2012). *Neural networks for time-series forecasting*, *Handbook of Natural Computing* (pp. 461–477). Berlin, Heidelberg: Springer.
2. Dunis, C., Likothanassis, S., Karathanasopoulos, A., Sermpinis, G., & Theofilatos, K. (2014). *Computational intelligence techniques for trading and investment*. London: Routledge.
3. Sermpinis, G., Fountouli, A., Theofilatos, K., & Karathanasopoulos, A. (2013). *Gene expression programming and trading strategies*, *Artificial intelligence applications and innovations* (pp. 497–505). Berlin, Heidelberg: Springer.
4. Chen, S. (2002). *Genetic algorithms and genetic programming in computational finance*. Amsterdam: Kluwer Academic Publishers.
5. Lisboa, P., Edisbury, B., & Vellido, A. (2000). *Business applications of neural Networks:The State-of-the-Art of Real-World Applications*. Singapore: World Scientific.
6. Kaboudan, M. A. (2000). Genetic programming prediction of stock prices. *Computational Economics, 16*, 207–236.

7. Alvarez-Díaz, M., & Alvarez, A. (2005). Genetic multi-model composite forecast for non-linear prediction of exchange rates. *Empirical Economics, 30*(3), 643–663.
8. Esfahanipour, A., & Mousavi, S. (2011). A genetic programming model to generate risk-adjusted technical trading rules in stock markets. *Expert Systems with Applications, 38*(7), 8438–8445.
9. Vasilakis, G. A., Theofilatos, K. A., Georgopoulos, E. F., Karathanasopoulos, A., & Likothanassis, S. D. (2013). A genetic programming approach for EUR/USD exchange rate forecasting and trading. *Computational Economics, 42*(4), 415–431.
10. Manahov, V., Hudson, R., & Hoque, H. (2015). Return predictability and the 'wisdom of crowds': Genetic Programming trading algorithms, the Marginal Trader Hypothesis and the Hayek Hypothesis. *Journal of International Financial Markets, Institutions and Money, 37*, 85–98.
11. Smith, V. L. (1982). Markets as economizers of information: Experimental examination of the "Hayek hypothesis". *Economic Inquiry, 20*(2), 165–179.
12. Koza, J. R. (1992). *Genetic programming: On the programming of computers by means of natural selection*. Cambridge: MIT Press.
13. Koza, J. R. (1998). Genetic programming. In J. G. Williams & A. Kent (Eds.), *Encyclopedia of computer science and technology* (Vol. 39, Supplement 24, pp. 29–43). New York, NY: Marcel-Dekker.
14. Divsalar, M., et al. (2012). A robust data-mining approach to Bankruptcy prediction. *Journal of Forecasting, 31*, 504–523.
15. Sermpinis, G., Laws, J., Karathanasopoulos, A., & Dunis, C. (2012). Forecasting and Trading the EUR/USD Exchange Rate with Gene Expression and Psi Sigma Neural Networks. *Expert Systems with Applications, 36*, 8865–8877.
16. Ferreira, C. (2001). Gene Expression Programming: A new adaptive algorithm for solving problems. *Complex Systems, 13*, 87–129.
17. Chen, H. H., Yang, C. B., & Peng, Y. H. (2014). The trading on the mutual funds by gene expression programming with Sortino ratio. *Applied Soft Computing, 15*, 219–230.
18. Lee, C. H., Yang, C. B., & Chen, H. H. (2014). Taiwan stock investment with gene expression programming. *Procedia Computer Science, 35*, 137–146.
19. Alghieth, M., Yang, Y., & Chiclana, F. (2015, September). Development of 2D curve-fitting genetic/gene-expression programming technique for efficient time-series financial forecasting. In *Innovations in Intelligent SysTems and Applications (INISTA), 2015 International Symposium on* (pp. 1–8). IEEE.
20. Dehuri, S. and Cho S. B., (2008), Multi-objective classification rule mining using gene expression. Third International Conference on Convergence and Hybrid Information.
21. Lopez, H. S., & Weinert, W. R. (2004). An enhanced gene expression programming approach for symbolic regression problems. *International Journal of Applied Mathematics in Computer Science, 14*, 375–384.

22. Shapiro, A. F. (2000). A Hitchhiker's guide to the techniques of adaptive nonlinear models. *Insurance, Mathematics and Economics, 26*(2), 119–132.
23. Pesaran, M. H., & Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *The Journal of Finance, 50*(4), 1201–1228.
24. Diebold, F., & Mariano, R. (1995). *Comparing predictive accuracy*, 3HMJG; EH@.
25. Duan, J. C. (1995). The GARCH option pricing model. *Mathematical finance, 5*(1), 13–32.
26. Mina, J., & Xiao, J. Y. (2001). *Return to RiskMetrics: The evolution of a standard*. New York: RiskMetrics Group.
27. Brandt, M. W., & Jones, C. S. (2006). Volatility forecasting with range-based EGARCH models. *Journal of Business & Economic Statistics, 24*(4), 470–486.

Part III

Economics

5

Business Intelligence for Decision Making in Economics

Bodislav Dumitru-Alexandru

1 Introduction

The real meaning of economic growth was defined during the economic crisis of 2007/2008. In particular, economic growth emerged as a topic of much discussion due to the lack of growth in both developed and emerging economies.

What could be done to create a transparent economic model that simulates an economy and also presents sustainability into the future? The twenty-first century has experienced rapid informational and technological advancement, which is now at a stage where it can be applied to process large amounts of data in order to make decisions. In many economies these advancements have been the main contributors to growth. An economy that is developed through the creation of a hybrid information and a technological model driven by Business Intelligence (BI) strategies with the support of Big Data analysis could revolutionize the global economy. The revolution could be based on developing scalable models that are used to predict crises, create growth models and re-ignite the financial markets in areas that were perceived as weak such as the development of derivative funds [1].

B. Dumitru-Alexandru, PhD (✉)
Bucharest University of Economic Studies, București 010374, Romania
e-mail: alexandru.bodislav@infinitumgroup.com

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets*,
New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0_5

For the past three decades the growth of many global economies has been due to economic development through modern capitalism, which was based on the partial annulment of the Bretton Woods agreement, regarded as the gold standard. In this system individuals and corporations are incentivized by creating wealth while at the same time improving the standard of living for future generations. Currently the traditional social structure is less defined as a result of this social pressure on those within the 'system' to create 'change for a better future'. Therefore, the modern capitalist system allows individuals of all classes to capitalize on opportunities to create wealth and improve their standard of living. In order for this system to work efficiently it is important for both public and private sectors to collaborate and share strategies and political development by re-defining the capitalist system.

This research is an add-on to existing research that covers a similar area of study. The Obama Administration underlined the fact that the potential is immense for automated mechanisms [2] that filter large volumes of data (Big Data) generated by the actual economic system. Its implementation in the long run represents a reliability boost for proactive measures [3]. On the subject there are different views, from those that encourage the use of data for defining patterns and behaviours in individuals or economic entities that help in mapping needs before they appear [4], to those that use the data and its processing for mapping brain logic [5, 6]. The use of data, processing and BI leads to secondary developments like creating the worldwide map of DNA [7] and have an anticipatory behaviour towards developments in social, economic or political aspects [8]. Although there is a need for limiting global coverage of interconnected issues, those could be anticipated with the use of BI and Big Data. This could lead to a control of human life because the needs of future generations could be identified prior to birth. This has implications for morality and security [9].

The support for automated decision making is given by the complexity of the decision needed to be made, which depends on the type of work analyzed. Reports, alerts, visualization, business analytics, simulation tools, collaboration tools and knowledge management [10] need to be aggregated to create a super-tool that could offer an automated facility for decision making. If decision making is seen as technologically achievable then sensors, actuators, data architecture, data mining, rule engines, web services, workflow applications and enterprise systems are needed in the process [11, 12].

The literature review discusses the impact of the research by addressing all aspects of the proposed research. Reaching automated decision making by using quantitative and qualitative analysis relies on observing how an economy works when it comes to creating and implementing policies. The literature review also analyzes the technological influence on developing GDP and how it could be moulded into a model by selecting similar sectors as those that form a modern economic system. The methodology presents the processes through which the model was developed and how the companies were selected. This section also includes the technical perspective on how an economy is partitioned into its main economic sectors by using the GICS (Global Industry Classification System). Altering the GICS sector classification helps our model to understand better the influences of different sectors and constituents, in order to clarify how the model works. In addition, the methodology also highlights how advanced and emerging economies performed over the same sample period and how main stock market indicators reacted to the macroeconomic environment through the lifespan of the Business-Automated Data Economy Model (BDM). In the empirical results section, we evaluate how the model performed during its lifespan. In particular, the analysis is evaluated from the perspective of the end user who observes the final result. It also focuses on the sectorial evolution of the model. The conclusion opens the door for future research and contemplates technological evolution and the macroeconomic reordering created by the current information era.

2 Literature Review

There are questions that can be asked about modern capitalism by bringing the idea of social inequality into the spotlight. For instance, the earnings of a company's CEO is estimated to be 500 times that of newly hired personnel [11].This disparity creates what is perceived by many to be an unethical situation that underlines social differences. In addition to this issue we have the inefficient situation of the re-distribution of tax revenues that instead of being distributed for investment in the creation of better social welfare, by building schools and improving the healthcare system are distributed for the payment of the budgetary deficit [13].

All associated issues with modern capitalism are founded on the development of the network economy[1] [14]. What is required is an economic model that through the influence of globalization helps the capitalist system to reform itself and maintain a long-term focus. This starts with healthcare, education and pension/social insurance systems [13]. From this point on we

[1] A network economy is a new organizational form of economies that is not situated in a hierarchical system but in a horizontal distribution of power and relations. As a result, so there are no clear leaders that impose themselves on smaller economies.

need to encourage cooperation and forge a bilateral partnership between the public sector (execution availability) and the private sector (implementation availability) and develop them together.

Meanwhile, the global trend is to create a corporate and social capitalist model (corporate, governance and state integration). This model is based on the solutions developed on a small scale and moulded for national, regional, union and global implementation. In this context the horizon for implementation must be discussed. From the public sector's perspective, in developed countries, the notion of economic vision is conceived as having a short life span similar to speculative investments carried out in the private sector [15]. This system of having different event horizons is based on the shift of logic created by political cycles, which introduces the idea that the funds spent by politicians in their countries depend on the short timeframe they are in power [16].

The sustainable evolution of an emergent economy can be framed into a corporate model, bringing in here Schumpeter's idea of an underlined benefit. Schumpeter's [17] corporate model gives better protection to a nation, region or state union by lowering the necessity of an intervention in the economy through government action. This model however emphasizes the importance of maintaining the connection between market needs and the private sector. This is particularly significant because the market is the place where demand meets supply and if the private sector is directly connected to the market then it can better satisfy the needs of the consumers that form the demand. The result of this Schumpeterian logic is a linear cyclicality of demand, this way the chances of technological under-utilization of production factors and its results in the economy decrease [13]. We also should consider corporate welfare as well as the social welfare of the entire nation but within limits that do not harm total social welfare.

Links are created between the state and corporations if there are logical transfers between decision paths used by the private sector that cross into public administration. Connections defined by their capitalist structure that result in growing economic performance must be maintained at a sustainable level in order to prevent over-heating the economy [20].[2] Another idea from Schumpeter could be considered here, the idea of '*healthy*' economic growth

[2] Overheating the economy leads to faster pace production. This leads to an increased consumption of resources in order to maintain the production cycle at a microeconomic level. At a macroeconomic level the unemployment rate decreases and the overall standard of the economy improves. Contributing factors may include the overall output of an economy and a reduction in the rate of unemployment to lower than the natural rate of unemployment. As a result, revenues to the government in the form of taxes are increased in a way similar to the evolution of the Okun's law [18]. Furthermore, if the economic output is higher than the potential output of the economy then the economy enters a boom cycle and in the long run this increases the chance of a 'crash' or a severe adjustment/correction to normal levels.

of 2 % per year (according to the analysis done on the 1890–1940 period, including here the unsustainable growth of 3.8 % per year pre-crisis—the 1890–1929 period[3]) [17].

Capitalism without democracy is not possible, so we need to bring to the discussion the theory of the functionality of political democracy in ideal conditions [19]. This theory underlines the idea that an ideal democracy is similar to the system of free enterprise in a given market, stressing the ideological convergence between democracy and capitalism. This vision is designed through the idea of efficiency in '*assuring functionality*' (corporate governance) of an economic sector at a higher level than that which the state can offer through government action or the political party through its doctrinal view [13]. The exemplification refers to one sector and not to an entire economy because according to the approach on competition in a democracy [19] we can draw similarities between the connections with an economy and its sectors, in parallel with a corporation and the divisions that create it. This way the divided purpose creates a better way to obtain the output based on specialized governance, which is done by individuals from the corporation (division directors, department managers) and by individuals or work groups from the economic sectors (ministry and state secretaries).

This perspective can be treated and moulded as an innovation brought to the debate by the thesis of Albert Hirschmann in *Rhetoric of Reaction* [20], in which he shows the problem of wealth creation in the 20th century: the perversity thesis—the desired political, social and economic action creates a backwards reaction against what was intended; the futility thesis—the social transformation doesn't have any results (through political and economic implementation); and the risk thesis—the proposed solution can cancel other results and then the final cost would be too large for the implementation of the solution [13]. These theses are taken into consideration because the deliverable must be a realistic solution, feasible and also have the ability to be implemented in the short term but with a long-term perspective. Innovation is the core of economic vitality, which drives growth, creates jobs, builds healthcare, provides employees with a purpose, regenerates organizations and consolidates the consumer's life span through the creating new products and revolutionizing services [21].

Returning to the idea of economic growth and its connection with investment we could see general and abstract views. The abstract view is connected

[3] The sample created by Schumpeter was developed by observing the data offered by Western and Central European countries, that is, countries that were economically affected by World War I and II, but were booming in pre-war.

to the microeconomic perspective and scaled for the general view or the macroeconomic perspective. If we blend the microeconomic and macroeconomic perspective we then observe companies competing against countries in the field of economic growth, in their ability to create or absorb innovation and in generating revenues. For example, in 2012 Switzerland was in competing with Apple as the nineteenth world's largest economic entity. The Swiss economy was producing a GDP of 545 billion USD and Apple had an average market valuation of 560 billion USD [22]. At the end of February 2015, Apple was the biggest corporate financial entity in the world, with a market capitalization of more than 755 billion USD (Bloomberg Professional Service). This rapid expansion and growth is now competing with the likes of Saudi Arabia's GDP, estimated to be around 778 billion USD [23]. The wage level and prices in developed and technologically advanced countries are at a stage of accelerated growth. As a result, a transfer of jobs from developed countries towards emerging markets has been occurring on a global scale. Similar deployments are seen in the de-industrialization of developed countries versus the industrialization of emerging economies. Emerging economies could create an advantage through the surplus in their commercial balance, however, they are unable to evolve into a real global players until they fulfil all the macro-positions (producer, exporter and consumer)[4]. This bidirectional issue pressurizes the process of de-industrialization in developed countries, especially in those that are members of OECD.

John Maynard Keynes developed the well-known macroeconomic formula for macroeconomic output [24], this is known more commonly as GDP [22, 25]:

General Equations for Macroeconomic Output

$$\mathrm{Y} = \mathrm{C} + \mathrm{I} + \mathrm{G} + \mathrm{X} \tag{5.1}$$

$$\mathrm{X} = \exp - Imp \tag{5.2}$$

$$\mathrm{C} + Imp + \mathrm{S} = \mathrm{I} + \exp + \mathrm{C} \Leftrightarrow \mathrm{C} + \mathrm{S}\text{-}\mathrm{I} = \exp - Im \tag{5.3}$$

where:
Y—Gross Domestic Product
C—Consumption

[4] China represents a paradigm, because although it represents a producer, consumer, exporter and importer, it has struggled to become a global leader when it comes to Research and Development.

G—Government Spending
X—Net Exports: (Exp—Exports; Imp— Imports)
S—Savings
I—Investment

The national surplus (S–I) must equal Net Exports if a developed country imports from emerging countries whereas a commercial deficit will occur when Exp–Imp is negative [22]. This import of goods is balanced by investment in emerging countries that occurs due to favourable savings rates in developed countries. This allows investments to be made outside the country of origin. S–I is positive because we have a positive output from S–I and a negative Net Exports and we also know that these two macrocomponents are equal [22]. Although reducing a deficit may seem like a way to promote import constraints, the Aggregate Demand of the population stands as a reactive force in using the deficit as a measure to dilute investments. After analyzing the macroeconomic technical framework[5]we can see that economic growth represents an increase in capacity to produce goods and services from one period to the next. A macrocomponent that accelerates growth is technology and information technology and this accelerator is used as a principle in the model we will discuss later in the chapter.

The societal values created in a company or in a ministry could lead to the rejection of objective BI solutions and thus annul any government initiatives in the macroeconomic sector [26]. Macroeconomic policies are seen as projects that are implemented with the help of BI supervision and evolve to be successful through the convergence of many factors. The most important factors are people, technical strategies and technology.

The created advantages of BI software, with a bias towards strategy and executive implementation, are based on offering added value to companies situated in a changing economic environment.

Creating BI completely bespoke software for a company or a governmental body could take years, political cycles or economic cycles, and the technological evolution must be also considered. The main driver of BI software is the immense volume of data that is analyzed to obtain efficient and direct solutions, which consume minimum amounts of resources, effort and time. The effort is linear during time and it is also the factor that increases the success rate for implementing BI. This way it is also able to deliver in line with the needs created in the twenty-first century, be it for a corporation or for any form of government or governance.

[5] The technical framework represents how the economic aggregates function, how they influence global economic growth and their relationship with consumption and investment/savings.

BI software that is biased on variable forecasting like the IBM Cognos is based not only on the quantitative result, but also on reaching the qualitative side by implementing it in the community. Through these connections, nationwide goals and corporate goals are reached, and a new complex BI system is developed that also sustains excellence in its domain.

Furthermore, in shaping the idea of automation, excellence should be discussed as a standard, which could be developed in the long run as a macro-picture that highlights the automation of responses. The responses received from the system are given from the relation between technology and individuals that are filtered through institutional practices seen in both the public and private sectors [26, 27].

Automation as a proposed solution represents the beginning when it comes to automated corporate governance because it is based on principles for reporting, analyzing, integrating work flow, visualization, and metric yielding (scorecard). These principles work based on standardized processes for performance management (financial analysis, planning, budgeting, forecasting) [27], standardized processes for advanced analysis (predictive analysis, data mining) and standardized processes for information management (data profiling, data quality testing, data warehousing and data integration).

An automated transformational process along with technological development could change not only the system's evolution but also cohesion in the system's structure, which needs to be developed as a multi-dimensional strategy [28].

Automation represents a new '*social*' process (i.e. *social* understood as a process that the state follows to create social welfare and to raise the standard of living for its inhabitants), which reduces the chance of creating returns for the state. It could be seen as a function for minimizing profits and for giving back to society, and for its members to increase their long-term living standards. Automation can also be used for proactive crisis management where the results from a range of business models are filtered through BI software.

Successive recessions lead towards the replacement of labour with capital, but the issue arises from the fact that economic growth is re-ignited each time and economic output returns to its normal trend. However, the workflow or labour quantity decreases, as a result certain jobs will become redundant.

Economists, starting from Joseph Schumpeter, followed the scientific purpose of creating sustainable economic growth that could also be shifted towards emergent markets. Furthermore, Robert Solow followed the idea and proposed the theory of long-run growth [29] that gained momentum in the early 1980s.[6] His theory was rapidly replaced by financial growth theories,

[6] Robert Solow's theory of economic growth includes the idea that growth is accelerated when capital is preponderant in the economy. This is because the scalability of employees is lower and they can more

which maintain the idea that our economy could be developed by following the intangible side of the economy, such as financial derivatives or the service side of the demand. The intangible evolutionary component represents the difference between traditional economics and modern economics. Meanwhile, it also represents the acceleration towards the future, based on the need to ride the innovation wave for positioning mankind on the enlightened side of evolution. This side of evolution allows lowering losses and decreasing the number of paths to follow that are developed at state, regional or union level (e.g. infrastructure, recreating comparative advantages and creating competitive poles).

3 Methodology for Creating the Business-Automated Data Economy Model

The proposed Business-Automated Data Economy Model (BDM) is designed to improve the efficiency of closed funds by developing an algorithm that uses data from the US stock market. The secondary output is to use the same algorithm as a model that is scaled to fit to solve issues regarding automated decision making at a government level. All these solutions are filtered from the BI perspective of software algorithms similar to those found in solutions like the IBM Cognos.

Our proposed model helps capture large amounts of data in order to provide viable solutions for decision making particularly when simulating an entire nation's economy. It represents a new vantage point on the traditional view of economic growth and the idea of developing a pseudo-economy that replicates how a country functions. To achieve this, many companies that capture and represent the broad economy are included. These companies operate in many different sectors as classified by GICS sector classification methodology.[7] In particular, companies used to create the BDM are listed on NASDAQ OMX exchange.[8]

easily be replaced by technologies that evolve exponentially over time. This way the momentum increases exponentially at the macroeconomic level as well. Solow was the pioneer behind the idea of growth through capital accumulation by correlating the savings function with the investment behaviour of individuals. Reference to the GDP formula can be made from equation 5.1 in this chapter.

[7] The Global Industry Classification Standard has classified the sectors as follows [30]: energy, materials, industrials, consumer discretionary, consumer staples, health care, financials, information technology, telecommunication services and utilities.

[8] NASDAQ OMX is the largest electronic equities exchange in the United States. The company owns several stock exchange platforms worldwide. NASDAQ OMX Group also offers services that help com-

BDM is developed by analyzing how the economy works from the entities that form it, from the sectors that, through its business processes, develop to form a nationwide economy. The companies that form the BDM also underline the difference between profit and welfare (the purpose of a company versus the purpose of a government or state). Practically, the efficiency of companies represents the backbone of the developed model and validates this research.

To develop a country in a successful and functional manner you need to surpass the unification processes, the national consciousness[9] and the historical dependence that usually form a nation, and go further than the event horizon. Macroeconomic performance shows in a relative way how economic evolution is ignited in a country. However the quality and living standards are also important details in the path of development.

When developing the proposed economic model, we initially evaluated companies that were most appropriate for the economic model that were most influential and representative of a modern nationwide economy. For this to be possible we need to use a transparent and open market which also has minimal transaction costs, both requirements being achieved by NASDAQ OMX[10]. We used the advantages offered by NASDAQ OMX (as an operational platform) and Bloomberg Professional (as the service that provides analytical data relevant to our research to help us retrieve data in order to build our model). Before all these could be used we needed to identify the main sectors that form a functional economy. This had to be in line with the latest trends of the twenty-first century, so we sliced the main macroeconomic sectors into ten. Using the GICS classification [30] as a basis we modified some sectors accordingly to provide a better fit for our model and because we needed to reach the actual volatile status of a modern economy,[11] for instance, the Financials sector was divided into two sectors: financial and derivatives. The Consumer Discretionary sector was renamed as the Cyclical Consumption sector because of the faster rate of consumption—the goods and services that

panies with investor relations, market intelligence, board relationships and news dissemination. All of which create transparency when it comes to investing in the companies listed on NASDAQ [31].

[9] We understand national consciousness to be the desire to unify a geo-political territory that has a population that wants to speak the same language and have the same governing laws.

[10] The theory of transaction costs underlines the fact that when developing an economic model, a company or a business venture, the development or maintenance of the enterprise will have some operating costs, also known as transaction costs [32]. When it comes to our model we needed those costs to be near zero and the only transparent and open market with near zero costs is the NASDAQ OMX platform.

[11] 'Volatile status of a modern economy' as we understand it means the actual development of modern economies, which have as the most important sector the Financial and the Derivatives sectors, because they fuel all other sectors. These two sectors represent the blood stream for all other sectors and they need to be highlighted accordingly.

form this sector have a higher replacement rate. The Consumer Staples sector was also modified to be the Non-Cyclical Consumption sector because of a lower rate of consumption of those types of goods and services. The modified sectors are: Materials, Derivatives, Telecommunication Services, Cyclical Consumption, Non-Cyclical Consumption, Energy, Industrials, Financial, Information Technology and Utilities. The Financial and Derivatives form the tertiary sector that is seen as the money flow side of an economy. The aggregate demand is based on the evolution of the Telecommunication Services, Cyclical Consumption, Non-Cyclical Consumption, Information Technology, Energy and Utilities sectors. These sectors offer consumers end products and services. Furthermore, they also enter the creation of aggregate supply when they enter the production process providing intermediary goods or services as is the case in the Materials and Industrials sectors. The companies used to form the model are traded on NASDAQ OMX and were selected by using the investment service offered by Zacks Investment Management (the Value Investor and Insider Trader solutions) [33, 34] thus, simulating the services offered by all investment companies to corporate clients. The 58 companies that form the economic model are representative to emulate and simulate a nationwide economy and its sectors, but are not representative of the NASDAQ OMX platform, which includes more than 3000 companies. The 58 companies were selected from a review of publicly available financial performance data,[12] by matching them with the previously mentioned sectors and by analyzing their potential for growth.

One of the most important characteristics of the model is that the 58 companies used in developing the model could be replaced with other companies that have a similar market niche or sector and also have the potential to reach the same value of the portfolio so that in the end the model will perform in a similar manner.[13] This exchangeability could be seen as a new path in validating private sector practices as a better approach in implementing projects by the public sector or the government. In the long run there could be developed a set of rules that could help policy makers in developing proactive macroeconomic solutions that have as a final goal sustainable economic growth.

[12] This was the initial path followed when selecting the companies. The fact that these companies are from different sectors and have different markets in which they trade creates diversification. This method of selecting the companies also offers insight into the idea of autonomous systems that are created by using a simple selection process. The idea of simulating an economy offers an optimal path for creating economic growth by emulating real economies.

[13] The process of creating the model is transparent, simple, and offers a new tool to explain financial growth instruments without the need to have components that are tested beforehand for them to be used in the model; it applies to all similar components (replacing companies with others which have the same activity domain and that the replaced financial value in the model is of a comparable value).

The selection of the companies that form the model was based on a previous research conducted in 2011 [35]. This selection process is explained and justified in more detail. The companies were selected by using the solution utilized by investment companies like Zacks Investment Research [33, 34] and Goldman Sachs [36]. The former uses both the Insider Trader solution [34] and Value Investor [33] solution.[14] These are tailored to the needs of investors who have theoretical knowledge of the domain but require further assistance with the practical aspects. Goldman Sachs Research [37] use the Market Themes 2015 [36, 38].The selection was initially made in the August–September 2011 timeframe using the solutions specified previously. The execution interval for the model was chosen based on the concept that the global economy is in a continuous muddling through process[15] and on considering that the model's execution is based on stock flow on a competitive market especially created. The stock prices also include investors' optimism and the fact that almost all chosen companies are listed in the USA on the NASDAQ OMX platform. To underline the resulted growth obtained in two stagnating markets (USA and European stock markets) with premises for a new recession, a growth of 20 % per year was needed for the model to be validated and represent attractiveness as an investment fund with low to average risk [39], and as a validation model for the intervention of corporate governance in state governance for creating policies and implementing and executing economic models. The rate of 20 % per year was established between the intersection of investor's attractiveness towards the model, the risk implied by the investment and the volume of the investment.

The companies selected for building the pseudo-macroeconomic BDM are shown in Table 5.1.

To rollover the basic model for validating corporate efficiency the timeframe was chosen between 3 October 2011 and 2 April 2015. This period was selected in order for the financial exercise because it was started from the speculative component found on the stock market each October. October represents the '*Earnings Season*' for the third quarter of the year and also the part where dividends are created and declared to be paid in the following quarter, so this allows market players to make a fast return on their investment by hunting annual dividends offered to shareholders. During the model's lifespan the global econ-

[14] Zacks Investment offers solutions like Insider Trader and Value Investor to help investors specify their investments based on preferences, individual circumstances and tolerance for risk. This may be for long-term investing, hedging or pure speculation.

[15] '*Muddling through*' is a concept used by the Obama Administration in the 2008–2012 mandate for explaining the inefficiency in obtaining sustained economic growth, although economic measures were implemented.

Table 5.1 Companies that make up the BDM

No.	Company Name	Bloomberg ticker	Profile[17]
1.	Schweitzer-Mauduit International	SWM	Premium paper
2.	Paramount Gold and Silver Corp.	PZG	Rare metals mining
3.	Goldcorp Inc	GG	Rare metals mining
4.	First Majestic Silver Corp	AG	Silver mining
5.	SPDR Gold Trust	GLD	ETF
6.	Telefonica SA*	TEF	Telecom
7.	Stamps.com	STMP	Delivery services
	OpenTable	OPEN	Online Reservations
9.	Google Inc.	GOOG	Internet Search & SAAS
10.	Watsco	WSO	AC technologies
11.	Town Sports International Holdings	CLUB	Fitness
12.	Steven Madden	SHOO	Footwear
13.	Ross Stores	ROST	Discount Stores
14.	Nordstrom	JWN	Fashion Retailer
15.	Men's Wearhouse	MW	Men's suits
	Maidenform Brands	MFB	Fashion & Retail
17.	LuLulemonAthletica	LULU	Technical Fashion
	Watson Pharmaceuticals	WPI	Pharmaceuticals
19.	Techne Corporation	TECH	Biotechnology
20.	On Assignment	ASGN	HR Specialized Solutions
21.	Jazz Pharmaceuticals	JAZZ	Pharmaceuticals
	Cross (A.T.) Company	ATX	Writing instruments
23.	Western Refining	WNR	Refining
24.	Sunoco Logistics Partners	SXL	Oil 360
25.	Patterson-UTI Energy	PTEN	Energy production
26.	Exxon Mobil Corporation	XOM	Energy (exploration & production)
27.	Alon USA Energy	ALJ	Petrol (production & distribution)
28.	Templeton Russia Fund	TRF	Financial derivate
29.	Tanger Factory Outlet Centers	SKT	REIT
30.	Rayonier Inc.	RYN	Building materials
31.	Plum Creek Timber Company	PCL	Wood industry
32.	Medallion Financial Corp.	TAXI	Financial taxi business
33.	CME Group Inc	CME	Trading & Investment Banking
34.	Berkshire Hathaway (1/100)	BRK/A	Business Solutions
35.	Bank of America Corporation	BAC	Banking & financial services
36.	American Campus Communities Inc	ACC	REIT
37.	Westinghouse Air Brake Technologies Corporation	WAB	AC technologies
38.	MasTec	MTZ	Building grounds
39.	GSI Group	GSIG	Lasers & Electro-optical solutions
40.	GrupoAeroportuario del Sureste**	ASR	Aeroports
41.	General Electric Company	GE	Technology & Financial Services
42.	Embraer-EmpresaBrasileira de Aeronautica**	ERJ	Aviation 360
43.	Colfax Corporation	CFX	Pipelines
44.	Boeing Company	BA	Aviation development
45.	Bae Systems Plc.*	BAESY	Aerospace &Defense Systems
46.	3M Company	MMM	Technology Company
	Goldman Sachs Income Builder Fund A Shares	GSBX	Mutual funds
48.	Ultratech	UTEK	Innovation
49.	Stratasys	SSYS	3D printers
50.	Microsoft Corporation	MSFT	IT
51.	j2 Global	JCOM	Cloud Computing & SAAS
52.	International Business Machines Corporation	IBM	BI & Technical Innovation
53.	Intel Corporation	INTC	Microprocessors & chipsets
54.	CACI International	CACI	Enterprise Information Technology
55.	Apple Inc.	AAPL	PCs & multimedia devices
56.	ACI Worldwide	ACIW	Electronic Payments
57.	3D Systems Corporation	DDD	3D printers
58.	American Water Works	AWK	Utilities - water works

1. The companies/funds that are highlighted were initially listed on the stock market, however, they were bought and delisted to become private (for example: Open Table, Maidenform Brands, Watson Pharmaceuticals and Cross (A.T.) Company, the Goldman Sachs Income Builder Fund was also initially part of the model, but was removed once the closed fund reached its maturity and was subsequently closed by Goldman Sachs.)
2. Most companies are primarily listed on NASDAQ, the two exceptions from Latin American are GrupoAeroportuario del Sureste and Embraer-EmpresaBrasileira de Aeronautica, which are Mexican and Brazilian companies respectively. They do however trade on the NASDAQ as ADRs. As a result, both companies were

(*continued*)

omy went through multiple major events, which reached the social, economic and humanitarian side of the global population, from the Arab Spring, Japan's earthquake and tsunami to reaching the upper limit of the national debt of the USA. All these events destabilized the global economy and lowered expectations and market forecasts, this way turning south the evolution of all companies that are listed on any stock exchanges worldwide. The evolution of the Greek situation starting with the beginning of 2015 also created external pressure on European and US stock exchanges. To emphasize the model's robustness, we tested the model for a couple of years earlier than the aforementioned timeframe and observed that BDM performs better than the stock market each time, without any direct influence experienced by economic crisis.

To validate the model's performance and not to consider it a '*black swan*' exception [40], the model was circulated for three-and-a-half years and this way it included expectations that were decreased all over the world [1]. Obtaining economic growth that gravitates around zero value sends a disturbing message to all advanced economies. Reaching the debt limit for the USA in 2013 and validating as a solution eliminating the debt ceiling on an undefined period of time sent the world economy into an unrealistic booming period. The FED (Federal Reserve) through its Quantitative Easing (QE) (1–3, and) and Operation Swap programmes and seconded by the European

Table 5.1 (continued)

included in our model because they are relevant for capturing all sides/sectors of a simulated economy. These two companies have integrated the entire airline services, including airports, personnel, luggage management, catering and maintenance for airplanes and airports. The North American equivalent is not available because US airlines are separated from airport management and all other integrated components in airport services.

3. Other companies that are not primarily listed on the NASDAQ but are included in our model include Telefonica and BAE Systems Plc. These two European companies are situated in Madrid, Spain and London, England respectively. They were selected as part of the model in place of their US equivalents, AT&T, Verizon and T-Mobile. Although these two were possible candidates to be included in the model they were deemed US centric and too interconnected with the evolution of the US economy. A similar situation is seen in the selection of BAE Systems Plc rather than Lockheed Martin Corp. or Halliburton Co. BAE Systems has contracts with more than 50 countries worldwide while Lockheed Martin and Halliburton mainly contract with the US Government.
4. With regards to Berkshire Hathaway, we diluted the value of the stock at 1 % of its value because its entire value would tip the evolution of the model towards the evolution of the Berkshire stock.
5. During the model's life span Apple Inc. decreased the volume of shares available on the stock market through a buyout operation (during April and May 2015), so the actual value of the shares was diluted seven fold.

Central Bank through the QE programme injected cheap money that ignited" should be changed in "(1, 2 and 3) and Operation Swap programmes that was seconded by the European Central Bank through the QE programme, they injected cheap money that ignited consumption in developed economies. To observe the efficiency of the QE programmes we can see the evolution in the Tables 5.2 and 5.3 of some of the main stock markets worldwide and the evolution of the output in advanced economies and emerging ones.

Compared with the evolution of economic growth from countries that are in direct economic relation and represented through decreasing or increasing economic growth, we also considered the evolution other main players in the global economy, that is, the countries that represent the BRICK[16] (Brazil, Russian Federation, India, China and South Korea) group and other emerging countries like Mexico, Turkey and Romania [1]. The social implications experienced in each powerful country as a result of globalization influences

Table 5.2 The evolution of strategic stock markets worldwide for the 2011–2015 timeframe

No. crt.	Stock Market	Returns for 1st period (%)	Returns for 2nd period (%)	Returns for 3rd period (%)	Returns for 4th period (%)
1	NASDAQ	10.31	28.10	56.63	109.22
2	S&P 500	6.21	22.03	42.97	88.04
3	DJIA	5.73	17.38	31.75	66.71
4	FTSE100	6.00	6.15	17.63	38.20
5	CAC40	19.60	22.03	48.62	73.37
6	NIKKEI225	23.31	36.29	119.81	126.00
7	DAX	30.14	38.68	63.81	122.58
8	BE500	19.36	27.18	45.19	74.32
9	Hang Seng	9.59	9.97	21.00	55.54
10	STOXX50	19.60	27.31	49.06	73.75
11	Shanghai Shenzen CSI300	68.29	71.22	77.78	61.31
12	BOVESPA Brasil	-0.74	0.04	-10.30	4.59
13	BOLSA Mexico	-0.12	7.03	7.29	34.09
14	KOSPI Korea	2.68	1.48	1.66	18.92
15	Bucharest SE	-1.76	17.78	48.94	67.48

The calculations are made highlighting the evolution for the actual situation (from 3 October 2011 to 2 April 2015).
The returns are segmented into 4 periods:
1st period: 2 April 2015–2 October 2014
2nd period: 2 April 2015–2 October 2013
3rd period: 2 April 2015–2 October 2012
4th period: 2 April 2015–2 October 2011
Source: Derived from Bloomberg Professional services data

[16] Although the acronym is BRIC (Brazil, Russian Federation, India and China), because of the advancement of South Korea it is considered that this country could enter in the select club of powerful emerging economies

Table 5.3 The evolution of developed countries, the BRICK group and a selection of emergent countries that are in full process of catching up

	GDP Growth	(%)				
Country (GDP hierarchy)	2011	2012	2013	2014	Evolution 2011 - 2014	2010 - 2014
Global	**3.90**	**3.10**	**3.00**	3.70	10.12	14.42
OECD	**2.00**	**1.50**	**1.30**	2.20	5.08	7.18
EU28	**1.60**	**-0.60**	**-0.40**	0.80	-0.21	1.39
USA	1.80	2.80	1.90	2.60	7.48	9.41
China	9.30	7.70	7.70	7.40	24.58	36.16
Japan	-0.50	1.40	1.50	1.20	4.16	3.64
Germany	3.40	0.90	0.50	1.90	3.33	6.84
France	2.00	0.00	0.30	0.90	1.20	3.23
Brazil	2.70	1.00	2.30	1.80	5.18	8.02
United Kingdom	1.10	0.30	1.70	3.20	5.27	6.43
Italy	0.60	-2.40	-1.80	0.50	-3.68	-3.10
Russia	4.30	3.40	1.50	2.00	7.05	11.65
India	6.80	3.20	4.40	5.40	13.56	21.28
Mexico	3.90	3.70	1.20	3.00	8.09	12.31
South Korea	3.70	2.30	3.00	4.00	9.58	13.64
Indonesia	6.50	6.30	5.80	5.70	18.88	26.60
Turkey	8.80	2.10	4.00	2.80	9.16	18.76
Romania	2.50	0.60	3.50	2.50	6.72	9.39

Source: Derived from data from the World Bank database [41] and the International Monetary Fund's database [23], June 2015

other remotely linked countries, a phenomenon that could likened to the butterfly effect.

In Table 5.3 it can be seen that between 2011 and 2013 the global output's pace decreased from 3.90 % growth year on year to 3.00 % growth year on year and increased for 2014 to 3.70 %, a similar pace being seen in the OECD's level, the EU28's level and in the USA. For 2015 the increasing trend of the economy is given by the results of the QE programmes ignited in the USA by the FED. The FED, by creating money and injecting it into the economy, artificially grew the economy by stimulating demand for goods, services and stocks, and this way created more jobs in the economy. This came with a long-term cost that is going to be experienced through increasing inflation, a boom in the housing market and an increase in the volume of stocks traded on stock exchanges worldwide.

4 Empirical Results of the Model

Here we present the performance of each of our constituent companies over four specific time periods.

The principles resulting from testing the BDM also form the basic algorithm used to simulate or to emulate a nationwide economy (especially in the case of emerging ones). Through the use of BI analysis and corporate governance practices the following are derived:

1. *A financial biased output*—if you develop a closed investment fund by using the developed model you'll have a fund that will be rated by S&P or Moody's as Investment Grade because the companies that form the model are mature on the US stock exchange and transparently traded on NASDAQ OMX. They have also been analyzed by companies like PriceWaterhouseCoopers, KPMG, Deloitte or Ernst & Young and have been evaluated by one of the earlier mentioned rating agencies. They are considered safe so are included in the Investment Grade. The mix of companies presented in the model has large returns in this model so this way they reach returns similar to those experienced from high-risk investment funds (returns of more than 20 % a year) that are used by risk-taking clients for large returns, but with a high chances of loss.
2. *A macroeconomic biased output*—in previous research it was discovered that there is correlation between this created model and the procedures used by a government to take its next steps toward their goal [27, 35]. Simulating a national economy, as in the conception and development of this model, shows there are certain governmental performance improvements that occur between three and five years.

Table 5.4 shows the companies that made up the model and the model's evolution during the three-and-a-half years timeframe.

Table 5.4 shows that the results of the BDM as an arithmetic return is 81.44 %, which outperforms the average returns for the sample stock markets listed in Table 5.5. The results correlate better with consumer expectations for the year 2012 and they are in tune with the moments when economic inefficiency surpassed economic logic (the year 2013 was a year full of counter economic measures). For 2013 the model performed above average when compared to the global stock markets, which yielded a rate of 56.21 %, while for 2014 the momentum of the model slowed and the global economy changed course. This was in part due to Greece's debt negotiations with European lenders and the political instability that became apparent during their local elections. Political differences also arose when the European Union held elections in 2014. The model does not include fees for trading on the NASDAQ OMX or stamp duty for stocks that are primarily listed on the London Stock Exchange (e.g. BAE Systems is also listed on NASDAQ). These charges were not included because the main intention of the model was to simulate an

Table 5.4 BDM and its evolution

No.	Company Name	Ticker	End Value	Interm. 3	R.4. (%)	Interm. 2	R.3. (%)	Interm. 1	R.2. (%)	Starting Value	R.1. (%)
1.	Schweitzer-Mauduit International	SWM	45.77	40.14	14.03	59.47	-23.04	32.45	41.05	27.57	66.01
2.	Paramount Gold and Silver Corp.	PZG	1.40	0.88	59.09	1.30	7.69	2.59	-45.95	2.17	-35.48
3.	Goldcorp Inc	GG	18.85	23.60	-20.13	25.47	-25.99	45.62	-58.68	45.36	-58.44
4.	First Majestic Silver Corp	AG	5.49	7.84	-29.97	11.58	-52.59	22.86	-75.98	14.90	-63.15
5.	SPDR Gold Trust	GLD	115.28	116.74	-1.25	127.06	-9.27	172.10	-33.02	160.96	-28.38
6.	Telefonica SA	TEF	14.62	15.10	-3.18	16.03	-8.80	13.53	8.06	18.38	-20.46
7.	Stamps.com	STMP	66.04	32.11	105.67	45.83	44.10	22.86	188.89	19.36	241.12
	OpenTable	OPEN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9.	Google Inc.	GOOG	534.06	570.08	-6.32	887.99	-39.86	756.99	-29.45	495.52	7.78
10.	Watsco	WSO	125.23	87.35	43.37	94.01	33.21	76.92	62.81	51.31	144.07
11.	Town Sports International Holdings	CLUB	6.75	6.25	8.00	12.46	-45.83	13.22	-48.94	6.87	-1.75
12.	Steven Madden	SHOO	38.20	32.21	18.60	35.93	6.32	43.87	-12.92	28.36	34.70
13.	Ross Stores	ROST	51.96	75.81	-31.46	74.01	-29.79	65.95	-21.21	38.21	35.99
14.	Nordstrom	JWN	79.96	68.83	16.17	56.80	40.77	55.63	43.74	45.37	76.24
15.	Men's Wearhouse	MW	52.68	46.67	12.88	34.30	53.59	34.31	53.54	25.01	110.64
	Maidenform Brands	MFB	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17.	LuLulemonAthletica	LULU	63.35	42.36	49.55	74.46	-14.92	76.35	-17.03	44.80	41.41
	Watson Pharmaceuticals	WPI	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19.	Techne Corporation	TECH	98.52	93.64	5.21	80.57	22.28	73.38	34.26	65.29	50.90
20.	On Assignment	ASGN	38.69	26.58	45.56	33.01	17.21	19.95	93.93	6.75	473.19
21.	Jazz Pharmaceuticals	JAZZ	168.86	154.55	9.26	90.28	87.04	58.76	187.37	37.35	352.10
22.	Cross (A.T.) Company	ATX	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23.	Western Refining	WNR	46.32	42.04	10.18	30.50	51.87	26.58	74.27	11.80	292.54
24.	Sunoco Logistics Partners	SXL	41.53	47.88	-13.26	66.46	-37.51	49.00	-15.24	29.46	40.97
25.	Patterson-UTI Energy	PTEN	19.37	30.27	-36.01	22.39	-13.49	15.80	22.59	15.83	22.36
26.	Exxon Mobil Corporation	XOM	84.30	93.30	-9.65	86.08	-2.07	91.72	-8.09	71.15	18.48
27.	Alon USA Energy	ALJ	16.69	14.49	15.18	10.26	62.67	13.95	19.64	5.47	205.12
28.	Templeton Russia Fund	TRF	10.33	12.39	-16.63	14.79	-30.16	15.55	-33.57	14.06	-26.53
29.	Tanger Factory Outlet Centers	SKT	35.85	32.95	8.80	33.08	8.37	32.02	11.96	24.73	44.97
30.	Rayonier Inc.	RYN	26.45	31.14	-15.06	55.78	-52.58	48.65	-45.63	35.65	-25.81
31.	Plum Creek Timber Company	PCL	43.41	39.04	11.19	47.45	-8.51	42.94	1.09	33.75	28.62
32.	Medallion Financial Corp.	TAXI	10.34	11.75	-12.00	14.92	-30.70	11.97	-13.62	8.91	16.05
33.	CME Group Inc	CME	92.91	80.42	15.53	73.97	25.60	57.11	62.69	49.70	86.94
34.	Berkshire Hathaway (1/100)	BRK/A	2165.00	2062.50	4.97	1715.00	26.24	1332.27	62.50	1047.01	106.78
35.	Bank of America Corporation	BAC	15.54	16.88	-7.94	14.06	10.53	8.93	74.02	5.53	181.01
36.	American Campus Communities Inc	ACC	43.36	36.29	19.48	34.31	26.38	43.49	-0.30	36.15	19.94
37.	Westinghouse Air Brake Technologies Corporation	WAB	94.29	77.91	21.02	63.43	48.65	81.38	15.86	50.15	88.02
38.	MasTec	MTZ	19.49	29.60	-34.16	30.54	-36.18	20.55	-5.16	16.93	15.12
39.	GSI Group	GSIG	13.43	11.60	15.78	9.47	41.82	8.79	52.79	7.27	84.73
40.	GrupoAeroportuario del Sureste	ASR	141.11	125.21	12.70	115.43	22.25	89.90	56.96	48.94	188.33
41.	General Electric Company	GE	24.94	25.12	-0.72	24.33	2.51	22.79	9.43	14.69	69.78
42.	Embraer-EmpresaBrasileira de Aeronautica	ERJ	31.29	37.95	-17.55	33.73	-7.23	26.22	19.34	24.67	26.83
43.	Colfax Corporation	CFX	48.48	57.10	-15.10	57.66	-15.92	36.12	34.22	19.07	154.22
44.	Boeing Company	BA	149.28	124.17	20.22	117.84	26.68	69.53	114.70	58.25	156.27
45.	Bae Systems Plc	BAESY	31.09	29.47	5.50	29.13	6.73	21.60	43.94	15.78	97.02
46.	3M Company	MMM	162.80	138.67	17.40	119.20	36.58	93.54	74.04	70.93	129.52
	Builder Fund A Shares	GSBX	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
48.	Ultratech	UTEK	17.14	22.80	-24.82	29.97	-42.81	31.99	-46.42	16.12	6.33
49.	Stratasys	SSYS	52.89	115.05	-54.03	103.50	-48.90	56.55	-6.47	18.00	193.83
50.	Microsoft Corporation	MSFT	40.29	45.76	-11.95	33.92	18.78	29.66	35.84	24.53	64.25
51.	j2 Global	JCOM	67.20	50.18	33.92	50.25	33.73	32.78	105.00	25.50	163.53
52.	International Business Machines Corporation	IBM	160.45	186.91	-14.16	184.96	-13.25	209.84	-23.54	173.29	-7.41
53.	Intel Corporation	INTC	30.81	33.52	-8.08	22.89	34.63	22.84	34.89	20.62	49.42
54.	CACI International	CACI	88.33	70.88	24.62	68.90	28.20	52.31	68.86	46.81	88.70
55.	Apple Inc *	AAPL	877.24	699.30	25.45	487.97	79.77	661.29	32.66	374.57	134.20
56.	ACI Worldwide	ACIW	21.48	18.79	14.32	55.19	-61.08	42.20	-49.10	25.26	-14.96
57.	3D Systems Corporation	DDD	27.45	43.71	-37.20	55.22	-50.29	34.21	-19.76	13.50	103.33
58.	American Water Works	AWK	54.89	48.01	14.33	41.53	32.17	36.82	49.08	29.27	87.53
	TOTAL/ WEIGHTED AVERAGE		6331.48	5983.79	5.81	5714.67	10.79	5058.23	25.17	3616.89	75.05
	Normal Average				4.86		4.45		22.26		81.44

The periods coincide with those highlighted in the previous table and offer support for analyzing the evolution of the model in the same periods as main advanced and emerging indicators at stock market level.

The starting value date is 3 October 2015 and the End value date is 2 April 2015, with the following intermediates: 1. 2 October 2012; 2. 2 October 2013; 3. 2 October 2014.

The returns are segmented into four periods:

R.4.—4th period: 2 April 2015–2 October 2014

R.3.—3rd period: 2 April 2015–2 October 2013

R.2.—2nd period: 2 April 2015–2 October 2012

R.1.—1st period: 2 April 2015–2 October 2011

(*continued*)

Table 5.5 Main stock market indicators on developed and emergent markets

No.	Stock Market	Ticker	End Value	Interm. 3	Interm. 2	Interm. 1	Starting Value
1	NASDAQ	CCMP	4886.937	4430.19	3815.02	3120.04	2335.83
2	S&P 500	SPX	2066.96	1946.17	1693.87	1445.75	1099.23
3	DJIA	INDU	17763.24	16801.05	15133.14	13482.36	10655.3
4	FTSE100	UKX	6833.46	6446.39	6437.5	5809.45	4944.44
5	CAC40	CAC	5074.14	4242.67	4158.16	3414.23	2926.83
6	NIKKEI225	NKY	19312.79	15661.99	14170.49	8786.05	8545.48
7	DAX	DAX	11967.39	9195.68	8629.42	7305.86	5376.7
8	BE500	BE500	270.67	226.76	212.82	186.42	155.27
9	Hang Seng	HSCI	25275.64	23064.56	22984.48	20888.28	16250.27
10	STOXX50	SX5E	3715.27	3106.42	2918.31	2492.48	2138.24
11	Shanghai Shenzen CSI300	SHSZ300	4124.776	2450.99	2409.04	2320.16	2557.08
12	BOVESPA Brasil	IBOV	53123.02	53518.57	53100.18	59222.08	50791.53
13	BOLSA Mexico	MEXBOL	44202.94	44254.43	41300.66	41199.29	32966.23
14	KOSPI Korea	KOSPI	2029.07	1976.16	1999.47	1996.03	1706.19
15	Bucharest SE	BET	7101.95	7228.82	6029.94	4768.31	4240.47
	Total (in market units)		207748.253	194550.85	184992.5	176436.79	146689.09

The starting value date is 3 October 2015 and the End value date is 2 April 2015, with the following intermediates: 1. 2 October 2012; 2. 2 October 2013; 3. 2 October 2014.
Source: Data derived from the NASDAQ OMX [45] stock exchange and trading platform and Bloomberg Professional Service [46].

economy and we apply the theory of transaction costs [32] that helps us to ignore scalable models' fees from an analytical perspective.

Global stock market indices, which represent the major developed and emergent economies, were analyzed over the same intervals, and the results are shown in Table 5.5.

It can be observed from Table 5.5 that the calculated yields from the beginning of the sample period to the end for each major global stock market index (including emerging markets) produced an arithmetic return of 20.28 % at the end of 2012, 26.11 % at the end of year 2 (2013), 32.63 % at the end of the third year (2014) and 41.62 % at the beginning of April 2015.

Table 5.4 (continued)

The weighted average and normal average represent the returns for the entire model and are in line with the returns for each company as R.1., R.2., R.3. and R.4.
Source: Data derived from the NASDAQ OMX stock exchange and trading platform and Bloomberg Professional Service.

Figure 5.1 shows the graphical designed metadata of all companies that are components of the BDM, to which is also added to their average return (weighted and arithmetic) that underlines the resulted performance by executive management from the private sector.

Figure 5.2 shows the graphical designed metadata of all stock market indices relevant for comparison with the model's results (the average at the end). A comparison of Figs. 5.1 and 5.2 shows that the private sector is better than governments at creating economic growth (because BDM's growth is

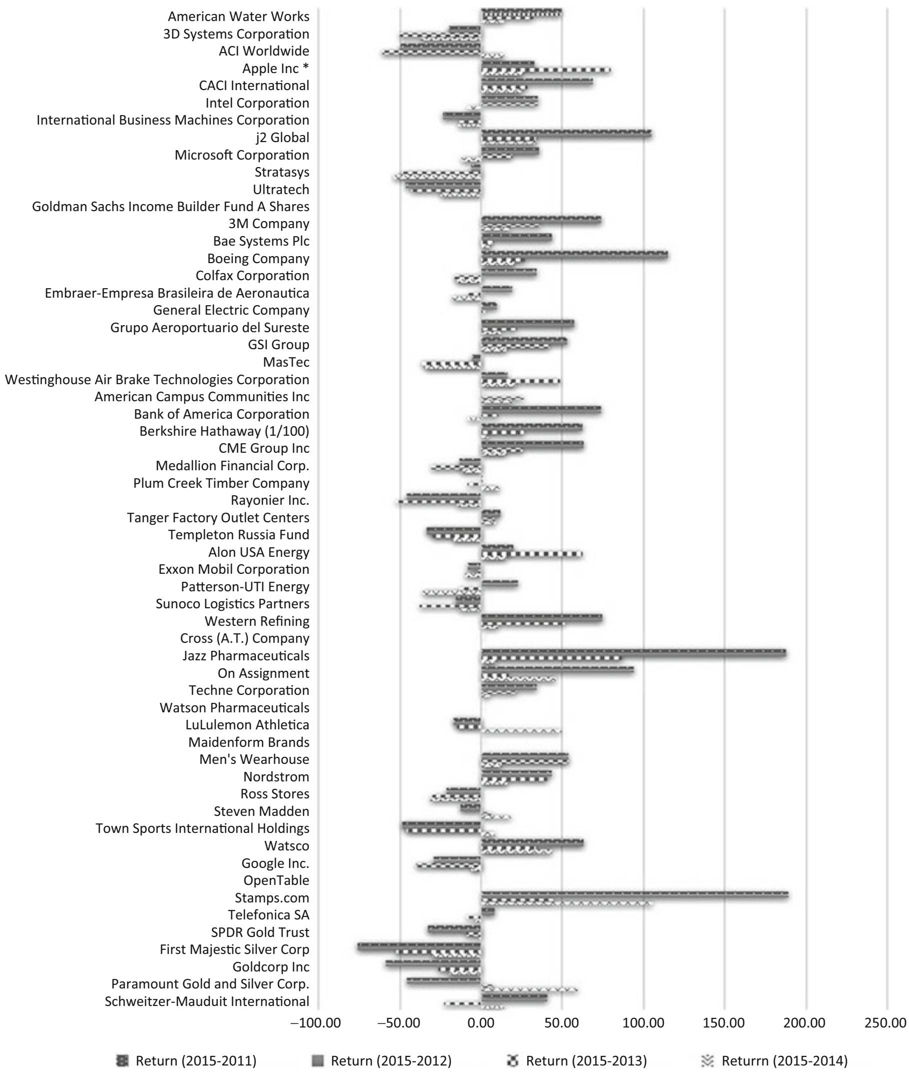


Fig. 5.1 The evolution of the components of the BDM. The scale is presented in percentage points. *Source*: Data derived from Bloomberg Professional and NASDAQ OMX—19 June 2015 [39].

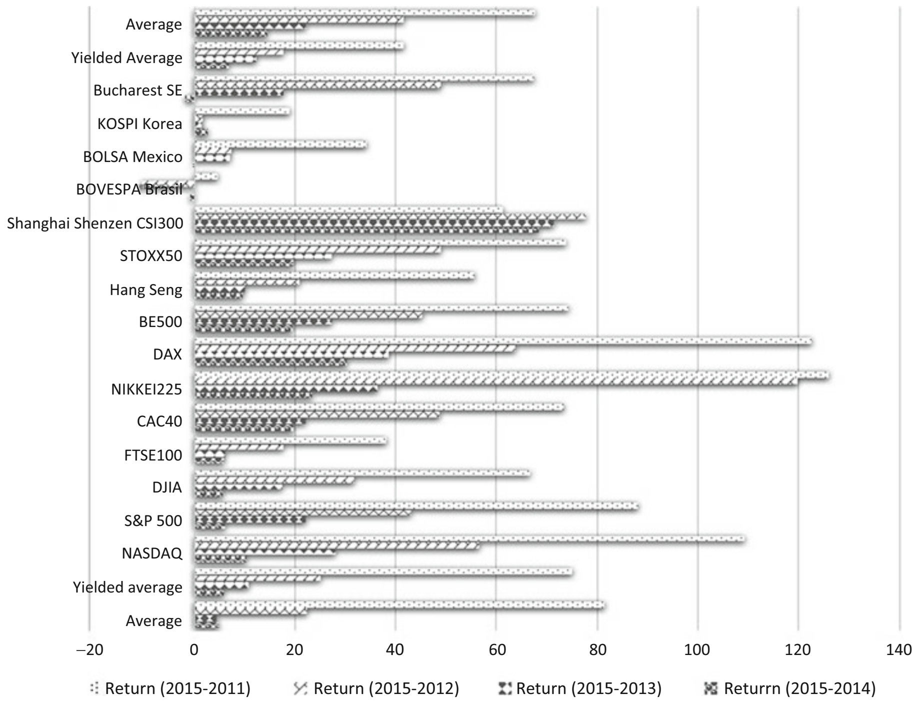


Fig. 5.2 The evolution of the main stock market indices that are relevant to be compared with BDM. The scale is presented in percentage points. *Source*: Data derived from Bloomberg Professional and NASDAQ OMX—19 June 2015.

higher than the average growth of the stock market indices used). The main role of the model that is also the central concept of this research is: *creating and underlining the feasibility of automated decision making by using Big Data analysis and qualitative validation through BI principles* [42]. This follows the sustainable feature presented through the production economy, the economy based on information technology and the banking sector.

To better observe the sectors of the economy we return to the changes made to the Global Industry Classification Standard developed by S&P in 1999 [30]. The Financials sector was divided into two sectors: Financial and Derivatives. The Consumer Discretionary sector was renamed the Cyclical Consumption sector because of the faster rate of consumption, the goods and services that form this sector have a higher replacement rate. The Consumer Staples sector was modified as the Non-Cyclical Consumption sector because of a lower rate of consumption of those types of goods and services. When selecting the companies that form the BDM we also tried to make a similar distribution of revenues according to the performance of an actual nationwide economy (the graphical distribution is available in Figs. 5.3, 5.4 and 5.6). The distribution we followed for the model's sectors was briefly described by the European Central Bank in its Economic Bulletins [43].

Table 5.6 shows the evolution of the absolute values of the model, which are better explained as returns in Table 5.7. Cyclical evolutions for the model's maturity and for validating the model's performance are underlined in Table 5.7 that follows:

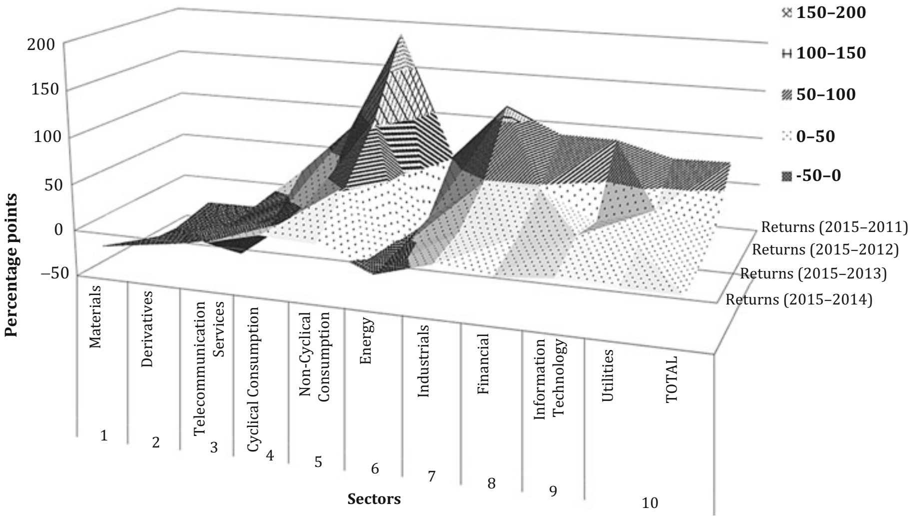


Fig. 5.3 BDM—sectorial evolution during 2011–2015

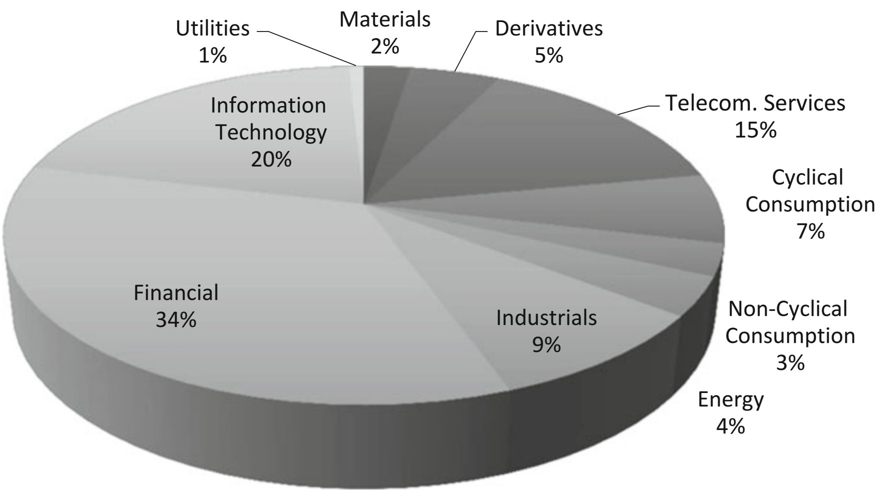


Fig. 5.4 The BDM—starting value distribution

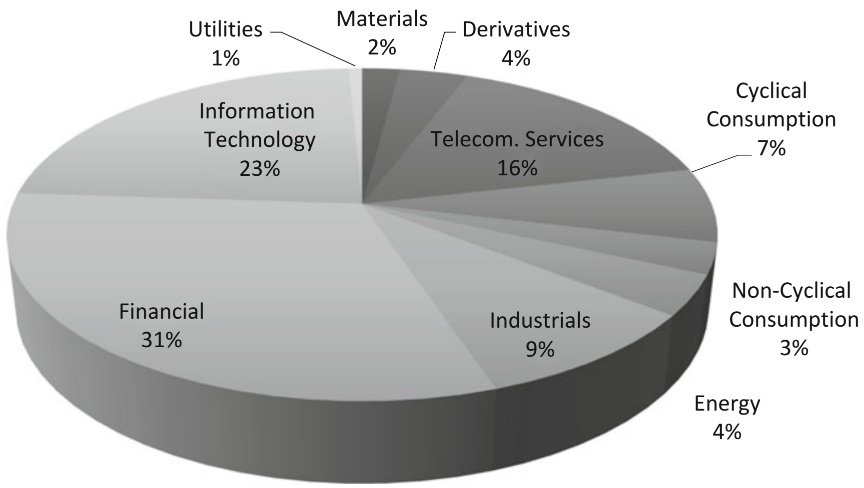


Fig. 5.5 The BDM—annual distribution according to its performance (as 2 October 2012)

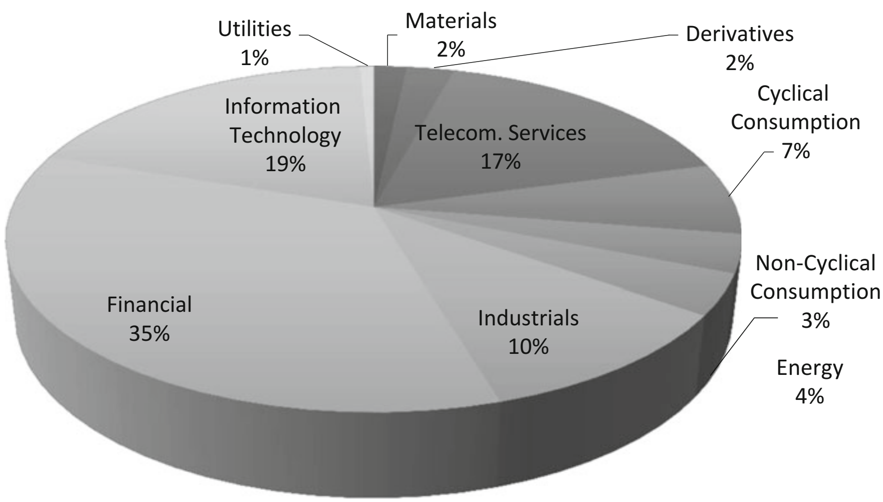


Fig. 5.6 BDM—annual distribution according to its performance (as 2 October 2012)

Table 5.7 shows that from the beginning of the sample period some sectors were negative. This highlights the fact that there was a poorly developed global Aggregate Demand for creating or repairing infrastructure and real estate between 2011 and 2014 with particular focus on the Materials sector (it went from −20.54 % in 2012 to −30.92 % and −26.90 % in 2013 and 2014, before recovering at the beginning of 2015). The entire table provides a better overview of the economy and an optimistic aggregate feeling because

Table 5.6 The value distribution of BDM's sectors

No. Crt.	Sector	End Value (02.04.2015)	Intermediate Value (02.10.2014)	Intermediate Value (02.10.2013)	Intermediate Value (02.10.2012)	Starting Value (03.10.2011)
1	Materials	71.51	72.46	97.82	103.52	90.00
2	Derivatives	125.61	129.13	141.85	187.65	175.02
3	Telecommunication services	614.72	617.29	949.85	793.38	533.26
4	Cyclical Consumption	418.13	359.48	381.97	366.25	239.93
5	Non-Cyclical Consumption	306.07	274.77	203.86	152.09	109.39
6	Energy	208.21	227.98	215.69	197.05	133.71
7	Industrials	716.20	656.80	600.76	470.42	326.68
8	Financial	2432.86	2310.97	1988.57	1577.38	1241.43
9	Information Technology	1383.28	1286.90	1092.77	1173.67	738.20
10	Utilities	54.89	48.01	41.53	36.82	29.27
	TOTAL	6331.48	5983.79	5714.67	5058.23	3616.89

The values are denominated by thousands USD for between the period expressed in each column.
Source: the author, according to the BDM.

Table 5.7 The evolution of the sectors that make up an emerging economy

No. Crt.	Sector	Returns 4th period	Returns 3rd period	Returns 2nd period	Returns 1st period
1	Materials	-1.31	-26.90	-30.92	-20.54
2	Derivatives	-2.73	-11.45	-33.06	-28.23
3	Telecommunication services	-0.42	-35.28	-22.52	15.28
4	Cyclical Consumption	16.32	9.47	14.17	74.27
5	Non-Cyclical Consumption	11.39	50.14	101.24	179.80
6	Energy	-8.67	-3.47	5.66	55.72
7	Industrials	9.04	19.22	52.25	119.24
8	Financial	5.27	22.34	54.23	95.97
9	Information Technology	7.49	26.59	17.86	87.39
10	Utilities	14.33	32.17	49.08	87.53
	TOTAL	5.81	10.79	25.17	75.05

Notes to accompany Table 7:

The returns are segmented into four periods:

R.4—4th period: 2 April 2015–2 October 2014
R.3—3rd period: 2 April 2015–2 October 2013
R.2—2nd period: 2 April 2015–2 October 2012
R.1—1st period: 2 April 2015–2 October 2011
The values are calculated as percentage points (%) for between the periods expressed in each column.
Source: the author, according to the BDM

during the final stage of the model (2 October 2014–2 April 2015) all negative returns were lowered (the Derivatives sector went from –28.23 % with a maximum of –33.06 % to –2.73 % in April 2015). An issue is that all of the sectors that seemed to be booming[17] (Non-Cyclical Consumption, from +179.80 % in 2012 to 11.39 % in April 2015, Industrials, from +119.24 % in 2012 to 9.04 % in April 2015) were deflated to a relative growth level, which means that the economy is poor on stimulating Aggregate Supply because the anticipation is poor on meeting the forecasted outputs of Aggregate Demand.

The evolution of all sectors is presented in Fig. 5.3.

Figure 5.3 shows that some sectors are still in negative forecasted environments (Materials, Derivatives and Telecommunication Services) while others are deflated from their booming period to regular values (Financial, Industrials and Non-Cyclical Consumption).

The initial framework that was developed into the model has the following distribution:

By observation of Fig. 5.4 we can see that the value distribution is biased towards the financial sector as primary stream. The model needs to highlight the Information Technology sector as a model for building a strong path when developing an economy.

Figure 5.5 shows that after the first year (2 October 2011–2 October 2012) the model performs in perfect harmony, with only the Technology sector developing at a faster pace.

Figure 5.6 shows that after the second year (2 October 2012–2 October 2013) the model performs in perfect harmony, with only the Financial sector developing at a faster pace and the Information Technology sector reducing its accelerated growth.

Figure 5.7 shows that the evolution of the model's sectors is similar to the momentum of the model and it develops growth harmonization when it comes to maintaining the output of the BDM with a positive trend (no sectors are considered to be peaking the model is designed to maintain an equilibrium).

Figure 5.8 shows the distribution structure of the sector's performance. It can be seen that growth is harmonized through the entire BDM and it is not biased towards a sector boom, an element that is frequently seen in speculative models. During the sample period (3 October 2011 to 2 April 2015) the Materials and Derivatives sectors were hit by uncertainty with little interest in investing in the long term. The Materials and Derivative sectors had negative outputs during the model's deployment. Meanwhile, sectors like Non-Cyclical Consumption and Industrials grew in a range of between 119 % and 180 % over 42 months

[17] In calculating the growth of an economy or of one of its sectors there are no theories or mathematical models that underline a certain value above which identifies a booming period, however it is acknowledged that when a certain sector or an entire economy is running at a pace well above the average or a trend for more than one year then it can be considered to be booming.

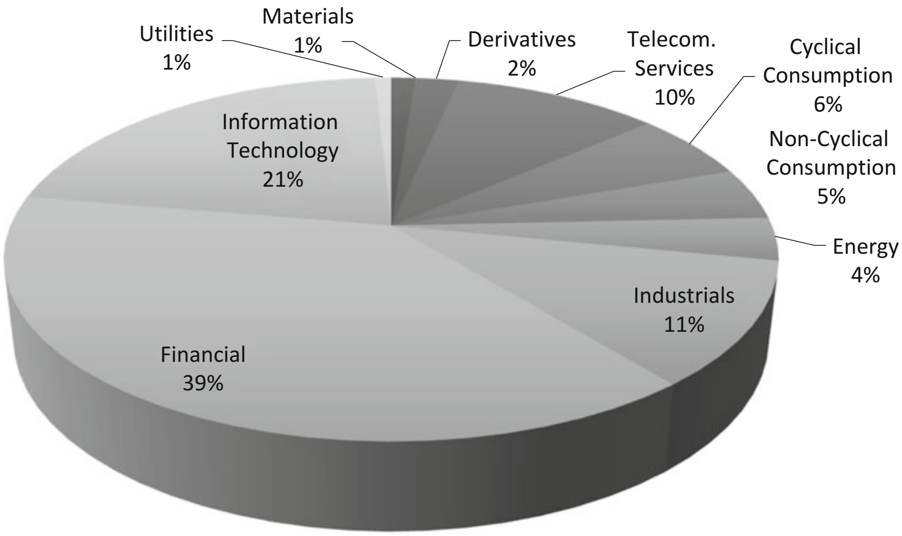


Fig. 5.7 BDM—annual distribution according to its performance (as 2 October 2014)

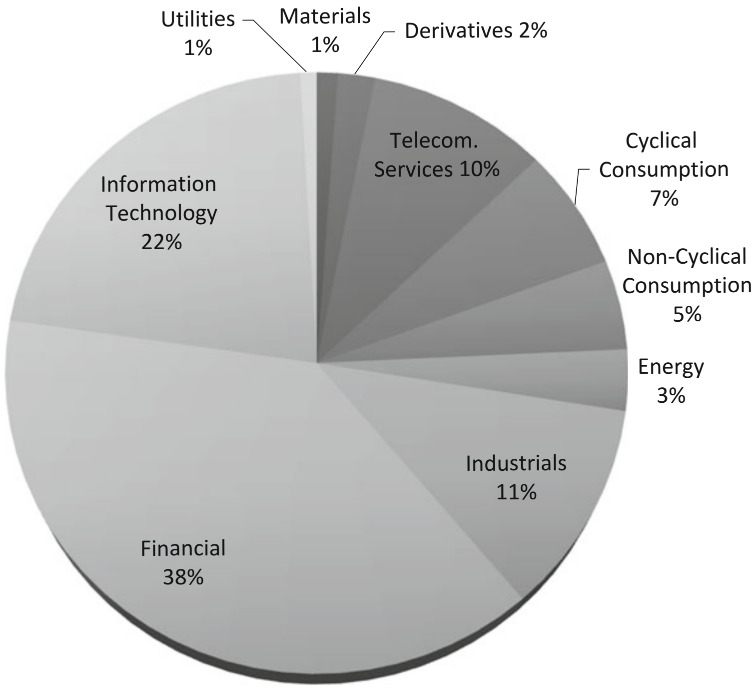


Fig. 5.8 BDM—maturity status (2011–2015). Data weighted for the End value date of the model—2 April 2015.

(three-and-a-half years). This fact shows that real estate development and infrastructure are the stars of the global economy and could result in increased revenues for those who are already big players in those sectors.

Figure 5.9 highlights sector growth and the anomalies associated with the hyper-acceleration of the Non-Cyclical Consumption and Industrials sectors. Furthermore, from Fig. 5.6 and the more detailed figures presented earlier we can see that in three-and-a-half years that the functionally developed economy (country) evolved better than all economies worldwide (almost doubled in volume). In its evolution of sectors, the model was situated in steady state when it came to endogenous development.

A question is raised by the fact that the Information Technology sector grew by 40 % (from 20 % to 28 %), while the economy doubled in volume. This fact technically means it actually grew 2.8 times from the initial value of the sector in just three years. The pace of the Information Technology sector represents an anomaly because all other sectors grew around 10 % in three years when it comes to volume (2.1 times in financial value) and this represents an ordinary pace. Another anomaly is the decrease of the weighted value of the Telecommunication Services sector as it decreased from 15 % to 9 % in total volume. Although the Telecommunication Services sector grew by 20 % in size, it decreases compared with the entire model, which doubled in size from its initial value.

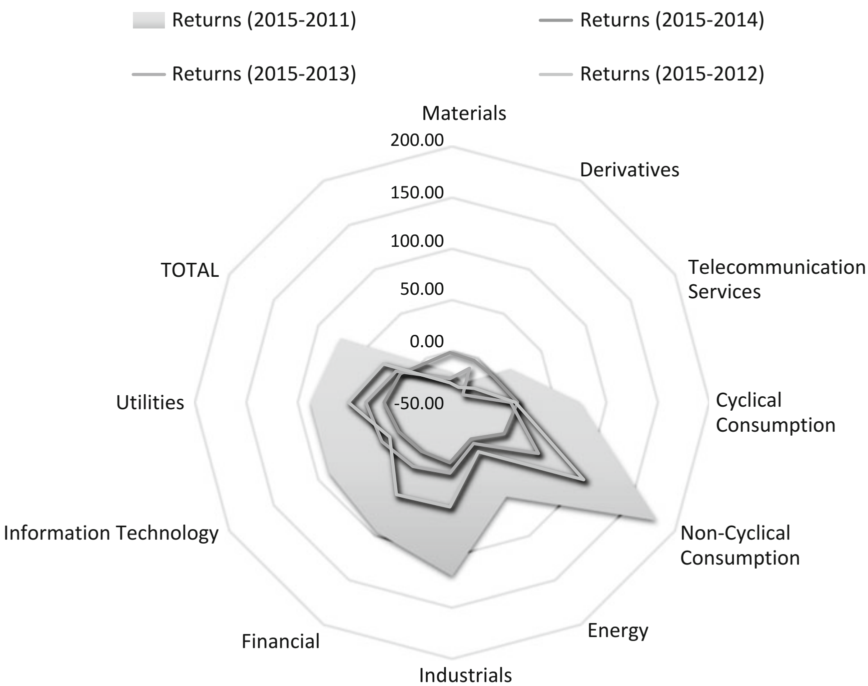


Fig. 5.9 The harmonized distribution of economic sector growth in the BDM

After reviewing the results of this model for a country economy it could be observed that replacing state governance with corporate governance leads to economic growth, which is almost 100 % higher [1]. This way, BDM is validating two new concepts:

- From the **State Governance** perspective we have the following fact: the public sector–private sector relationship represents synergy if public institutions are in a state of continuous optimization and re-shaping through corporate governance. Furthermore, public institutions must make transparent decisions and certainty, more accurate forecasting process, because it also involves the private sector versus the public sector.
- From the **Corporate Governance** perspective we have the following scenario: the public sector–private sector relationship represents the starting point and by filtering it with the law of large numbers synergy occurs. This accelerates performance towards its optimal level, which is an efficiency experienced in large corporations.

If these ideas are transferred into the economic environment it will ease the search for the *true North* and lead to linearity in economic foresight for developed and emerging economies [1]. Microeconomic constraints that reduce seesaw-like evolution between Aggregate Demand and Aggregate Supply could be encountered. As a result, an unsustainable overcharge of production facilities could be faced and surpassing the maximum production level could lead to inefficient technical yields, lower profits and the market equilibrium shifted towards a loss-loss situation for consumers and producers.

5 Conclusions

The process of building the model and highlighting its feasibility over time offers a new analytical approach towards the field of investment banking or for sketching the broader image of state governance linked with policy making for creating sustainable economic growth. The big picture represents the idea of the information economy in the digital era and going from analogous consumers (based on traditional experiences when it comes to satisfying their needs through demand that is met by supply in a certain given space, be it a store, bank, generally known as the market place) towards digitally filtered consumers (they are satisfied by solutions created in the virtual market place, online or in augmented or virtual reality). The digital advance and bias towards digitalization has as its main features: decreasing production costs that imply policies similar to the ‘just-in-time’ stock management created by Toyota [47] in the 1970s; and in the long run representing upscale lowering prices. As a result, the demand for Toyota's goods is targeted to match perfectly supply and demand and estimated in line with consumer's needs, the counterpart, the producer's supply will have an immediate response and this way the stocks of goods will be very close to zero. All these results seen as the bigger picture form a great competitive advantage: the optimization of consumer anticipation. At the *microeconomic* level we have individual clients, such as large corporations for Business to Business relationships and at *macroeconomic* level we deal with countries that function like hyper-corporations. These must attract new investors, such as other corporations and other states. The attractiveness is created by an optimized labour force which involves individuals who accept offers that others do not. For instance, advanced economies have jobs/occupations which aren't attractive to its residents. For instance, in agriculture many European countries have jobs such as orange picking that aren't well paid so they employ people from Eastern Europe in order to do those jobs. The same is happening with the utilities sector in the USA where gardening and cleaning services, sewerage maintenance and other low-skilled jobs are occupied by cheaper labour. When there are countries that could outsource or export this type of low-skilled labour it means they have a modern competitive advantage resulting from their level of development (i.e. Mexico, Romania and Bulgaria are emerging markets). Although Spain, Italy and Greece are going through extreme economic conditions because of their national debt and unemployment rate, they still outsource jobs to poorer countries. This path is also followed by corporations by outsourcing some of their low-skilled jobs.

The BDM highlights some irregularities or *black swan* events especially when it comes to the evaluation of its performance reviewed at a sectorial level. At a microeconomic level, the irregularity is caused by the Financial and Information Technology sectors due to overheating, leading to a faster momentum in their development cycles and an increased need for more resources to be consumed in the production cycle. At a macroeconomic level the output of the economy increases with a rate above its potential, which could backfire in the long run through the accumulation of the overproduction of goods and services. This process shifts the Aggregate Demand that is also fuelled by printed money created through QE programmes and operations of swapping short-run maturity debts with long-run maturity debts. In the long run the overheating of the Financial and Information Technology sector leads to their failure to the point where demand isn't fuelled by macroeconomic financing schemes like QEs. Another sector that shows overheating is the Utilities sector, not because they are more expensive, but because social disparities, global warming and information asymmetry works against the demand for utilities. The overheating of the Utilities sector is based on the malfunction of the global economy by uneven development of social classes and of countries worldwide. This issue leads to the need, by those who need utilities in underdeveloped economies, to pay a premium. All these issues represent a negative externality that increases the costs of living, especially because those costs are formed largely from the prices for goods and services used to satisfy basic needs.

The results show that the model (BDM) reacted and evolved according to the initial suppositions and hence offers insight into the possibility of shifting microeconomic solutions into the macroeconomic environment. This is similar to shifting corporate governance practices into public administration and policy making.

From the research we also observed issues that have a non-technical nature and could be filtered out with the help of BI factors [27]. The research also offers a solution to surpass the results developed by well-established political cycles. Some of the possible issues that could arise are:

- Lack of a long-run strategy;
- Metrics for measuring the success of a created policy applied at macroeconomic level are not clearly defined;
- Both political class and culture are interfering in setting a sustainable executive vision;
- Direct implementation and functional connections between properties of SAAS (software as a service) or BI technologies are non-existent;
- Solutions used at macroeconomic level are not connected and have low coverage at extremely high costs, thus the efficiency of implemented solution cannot be identified.

The Digital Information economy aligns expected demand with supply, which is created by small, medium or large companies, or by a state seen as a company. The alignment of business strategy with economic or national strategy results in the efficient execution of proposed goals. By testing different models, we observed that all hypotheses were proven and the results obtained converge towards the idea of decision automation (not understood as taking a decision in an automatic manner through predetermined algorithms, but by stimulating the idea of accessing only some paths to follow and flow proactivity in the work process).

Decision automation represents a normal step in the circuit of knowledge flow obtained from the quantic revolution and by treating information not as a final good, but as an intermediary flow that could have a multitude of endings. All of which are correlated with the user's needs and/or wants.

References

1. Bodislav, D. A., Pascu, L., Pascu, B., Ratezanu, I., Badea, G. (2015). Emulating Business Intelligence filtered with Big Data Analysis. The Private—Public Sector Transfer. *Proceedings of IBIMA 25th International Conference 2015*, 8th of May, Netherlands, 1471–1481
2. Podesta, J., Pritzker, P., Moniz, E., Holdren, J., & Zientis, J. (2014). *Big data: Seizing opportunities, preserving values*. Washington, DC: Executive Office of the President.
3. Armah, N. A. (2013). Big Data Analysis: The Next Frontier. *Bank of Canada Review*, Summer edition, 32–39.
4. Pitt, J. (2012). Design contractualism for pervasive/affective computing? *IEEE Technology and Society Magazine, 31*(4), 25–28.
5. Michael, L., & Miller, K. (2013). Big Data: New opportunities and new challenges. *IEEE Computer Society, 6*, 22–24.
6. Strickland, E. (2013). The gene machine and me. *IEEE Spectrum, 3*, 26–32.
7. Michael, K., & Michael, M. (2011). The social and behavioural implications of location-based services. *Journal Location-Based Services, 5*(3–4), 121–137.
8. Sestini, F. (2012). Collective awareness platforms: Engines for sustainability and ethics. *IEEE Technology and Society Magazine, 31*(4), 54–62.
9. Orwell, G. (1950). *1984* (1 ed.). London: Signet.
10. Harris, J., Davenport, T. (2005). Automated decision making comes of age. *MIT Sloan Management Review*, January, 2–10
11. Bodislav, D. A. (2015). *The Big Data—Business Intelligence Synergy. The Solution for Building Relevant Indicators for an Emerging Economy* (1 ed.). Bucharest: ASE Publishing.
12. Stjerne, D. (2011). *Introduction to Business Intelligence. What to consider implementing a Data Warehouse and Business Intelligence*. http://www.stjerne.biz/wp-content/uploads/2011/11/Business-Intelligence-An-introduction.pdf
13. Bran, F., Bodislav, D. A., Radulescu, C. V., & Ioan, I. (2014). Corporate Governance Intervention for a Sustainable Socio-Economic Model. *Revista de Cercetare si Interventie Sociala, 46*, 216–226.
14. Cairncross, F. (2001). *The death of distance: How the communications revolution is changing our lives* (revised ed.). Cambridge: Harvard Business School Press.
15. Bodislav, D. A. (2011). The Birth of the Regulated Company in the Macroeconomic Environment. *Theoretical and Applied Economics, 18*(6), 129–140.

16. Alesina, A., Roubini, N., & Cohen, G. (1997). *Political cycles and the Macroeconomy* (1 ed.). Cambridge: MIT Press.
17. Schumpeter, J. (2011). *Poate supraviețui Capitalismul? Distrugerea creatoare și viitorul economiei globale (Translation: Can Capitalism survive? Creative destruction and the future of the global economy)*. Bucharest: Publica Publishing.
18. Knotek, E. (2007). How useful is Okun's Law. Kansas FED. https://www.kansascityfed.org/publicat/econrev/pdf/4q07knotek.pdf. Accessed on 7 September 2015
19. Becker, G. (2003). *The Human Capital* (2 ed.). Bucharest: C.H. Beck Publishing.
20. Hirschmann, A. (1991). *The Rhetoric of reaction: Perversity, Futility, Jeopardy*. Cambridge: Harvard University Press.
21. Branson, R., Bodislav, D. A., Stoynova, P. (2010, August 8). If I could do it all over again. *The Wall-Street Journal*, A4
22. Pascu, L., Pascu, B., Bodislav, D. A., Rătezanu, I., Badea, G. (2015). Micro-accounting and macro-governance. A comparative perspective and the oversight of collaboration. In *Proceedings of IBIMA 25th International Conference 2015*, 8th of May, Netherlands, 1482–1489
23. IMF. (2015). International Monetary Fund. *World Economic Outlook*. http://www.imf.org/external/pubs/ft/weo/2015/update/02/index.htm. Accessed on 15 June 2015
24. Keynes, J. M. (1936). *The general theory of employment, interest and money*. London: Palgrave Macmillan.
25. Brakman, S., Garretsen, V., Marrewijk, V., & Witteloostujin, A. (2009). *Nations and Firms in the Global Economy*. New York: Cambridge University Press.
26. Bodislav, D. A. (2015). Transferring business intelligence and big data analysis from corporation to governments as a hybrid leading indicator. *Theoretical and Applied Economics*, *22*, no. 1(602), 257264
27. Boyer, J., Frank, B., Green, B., Harris, T., & Van De Vanter, K. (2011). *Business Intelligence Strategy—a Practical Guide for Achieving BI Excellence*. Ketchum: MC Press.
28. Bodislav, D. A. (2015). Decision Automation and the Feasibility of the Resulted Solution at Macroeconomic Level. *National Conference of the Doctoral/Postdoctoral Programme*, 29th of May 2015, Romania.
29. Solow, R. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics, 70*(1), 65–94.
30. GICS. (2015). *Global Industry Classification Standard*. https://en.wikipedia.org/wiki/Global_Industry_Classification_Standard. Accessed on 8 September 2015.
31. Amadeo, K. (2015). *What does NASDAQ stand for? Then and Now*. http://useconomy.about.com/od/themarkets/f/NASDAQ.htm. Accessed on 10 September 2015
32. Coase, R. (1937). The nature of the firm. *Economica, 4*(16), 386–405.
33. Zacks. (2015). *Zacks Investment Management Value Investor*. http://www.zacks.com/valueinvestor/?adid=TOP_ONLINE_NAV. Date accessed 1 May 2015.
34. Zacks. (2015). *Zacks Investment Management Insider Trader*. http://www.zacks.com/insidertrader/. Accessed on 1 May 2015.
35. Bodislav, D. A. (2013). *The optimal corporative model for sustainable economic growth for an emergent country* (1 ed.). Bucharest: ASE Publishing.

36. GSO. (2015). *Goldman Sachs Outlook*. http://www.goldmansachs.com/our-thinking/archive/2015/index.html#infographic. Accessed on 1 May 2015
37. GSR. (2015). *Goldman Sachs Research*. http://www.goldmansachs.com/what-we-do/research/index.html. Accessed on 1 May 2015
38. GSIB. (2015). *Goldman Sachs Investment Banking*. http://www.goldmansachs.com/what-we-do/investment-banking/services/financing.html. Accessed on 1 June 2015
39. Baker, H. K., & Filbeck, G. (2015). *Investment Risk Management* (1 ed.). New York: Oxford University Press.
40. Taleb, N. N. (2010). *Lebăda Neagră—Impactul foarte puțin probabilului, (Translation: The Black Swan—the impact of the least probable)* (1 ed.). Bucharest: Curtea Veche Publishing.
41. WB. (2015). World Bank database. http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators, Date accessed 15 June 2015
42. ECB. (2015). European Central Bank, *Economic Bulletin.* http://bnro.ro/Buletinul-economic-al-BCE-3222.aspx. Accessed on 5 September 2015
43. IMF. (2015). International Monetary Fund. *Financial Derivatives*. http://www.imf.org/external/np/sta/fd/. Accessed on 15 September 2015.
44. NASDAQ OMX. (2015). http://business.nasdaq.com. Accessed on 1 September 2015
45. Bloomberg Professional service. (2015). http://www.bloomberg.com/markets
46. Bodislav, D.A. (2015). The Synergy between Business Intelligence and Big Data—lowering the gap between the Future and Present Day. *Proceedings of 8th Edition International Conference Financial and Actuarial Mathematics*, 26th of June 2015, Bulgaria
47. TOYOTA. (2015). *Just in time manufacturing*. http://www.toyota-global.com/company/vision_philosophy/toyota_production_system/just-in-time.html

Part IV

Credit Risk and Analysis

6

An Automated Literature Analysis on Data Mining Applications to Credit Risk Assessment

Sérgio Moro, Paulo Cortez, and Paulo Rita

1 Introduction

Financial crises are catalysts to an outstanding increase in research that develops innovative techniques which attempt to anticipate financial crises. This enables market participants to take preventative action and learn from the past [1]. Therefore, research in financial credit risk has become one of the most prominent and prolific subjects in recent years. In fact, the 2008 global financial crisis arose due to a poor assessment of the risk associated with the various ways through which banks have transferred credit risk in the financial system [2]. This then spread worldwide highlighting global systemic risk.

Several studies have shown that there was a large increase in publications related to credit risk after 2009 [3, 4]. There are numerous reasons for such an event. First, the financial crisis sounded massive alarms that triggered research

S. Moro (✉)
Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR-IUL, Lisboa, Portugal

ALGORITMI Research Centre, University of Minho, Minho, Portugal

P. Cortez
ALGORITMI Research Centre, University of Minho, Minho, Portugal

P. Rita
Instituto Universitário de Lisboa (ISCTE-IUL), Business Research Unit (BRU-IUL), Lisboa, Portugal

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets*,
New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0_6

in two domains: bankruptcy detection for anticipating the impact of defaulting and regulatory reporting for providing a tighter control of financial institutions. Large-scale regulatory projects potentially benefiting from advanced machine learning techniques gained relevance in the post-crisis financial market, such as the IRB (Internal Ratings-Based Approach) in Europe [5]. The more rigid control on banks allowed for the detection of some severe fraud situations, with the institutions being affected by a loss of trust [6]. Other credit-risk related domains are credit scoring and research specifically on credit cards. While these were widely studied subjects prior to the crisis the global crisis has increased research on these topics. It also resulted in credit institutions requiring a tighter control on individual loans.

Data mining (DM) is a concept that encompasses techniques and methodologies for unveiling patterns of knowledge from raw data [7]. Typical DM projects include data understanding and preparation followed by the application of machine learning algorithms for finding interrelations between data that can be translated into valuable knowledge. The previous steps may involve data sampling and feature selection, depending on the data, and also data quality operations to improve the value of information before it can be used as input for subsequent steps. The later steps may include one of several widely studied algorithms, such as decision trees, artificial neural networks and support vector machines, or even an ensemble of a few different algorithms.

Real-world problems are often based on collected data from which a new insight is needed for leveraging business. Such problems may be addressed through a data-driven approach including DM. The credit risk domain typically involves an analysis of past history to understand which variables influence behaviour of credit holders, making it an excellent subject for the application of DM techniques. Being such an interesting applied field of research, a few literature analysis and reviews were published recently on DM applications to credit risk. Marques et al. [8] conducted a literature review by collecting 56 papers published from 2002 to 2012 on the application of evolutionary computing to credit scoring. Their methodology consisted in a manual analysis by dividing into sub-problems to which evolutionary computing techniques have been applied: classification, variable selection, parameter optimization, and other miscellaneous problems. The conclusions show that variable selection gets most of the attention regarding evolutionary techniques. Guerrero-Baena et al. [9] evaluated literature in terms of the application of multi-criteria decision-making techniques to corporate finance issues during the period 1980–2012. They used a total of 347 publications from the Scopus database. The method used was a manual classification of every article and a descriptive statistical analysis over that classification. The results presented by their study show bankruptcy prediction and credit risk assessment as receiving 4.6 % and

3.7 % of the attention. Nevertheless, that paper focused specifically on corporate finance, with capital budgeting receiving the most of the attention (64 %) however, there exists a high number of publications on DM applications relating to individual credit risk and default (e.g. [10]).

Literature analyses can be conducted through automated methods that are able to parse the relevant terms from each publication and then build logical clusters of articles, providing a meaningful structure from which new insights can be obtained. One of the most recent used methods includes Text Mining (TM), as the work of Delen and Crossland [11] demonstrates. Furthermore, the Latent Dirichlet allocation (LDA) algorithm may be used for organizing the articles in logical topics [4]. It should be noted that such a procedure has not yet been applied to credit risk assessment applications; it would be interesting to understand if the insights achieved from this procedure identifying potential research areas.

This study presents an automated literature analysis over a significant set of articles that focus on DM applications to credit risk assessment. The main highlights are:

- Collecting the 100 most relevant articles on credit risk using DM, according to Google Scholar relevance criteria.
- Using articles' keywords for building a lexical dictionary of relevant terms, followed by the application of TM for understanding the subjects that are deserving of the most attention.
- Applying an LDA algorithm for building a logical classification of articles, identifiable by the terms that characterize each of the topics.

The next section elucidates the materials and methods used for the experiments, whereas Sect. 3 analyzes the results achieved. Finally, Sect. 4 completes the chapter with conclusions and final remarks.

2 Materials and Methods

Search Criteria

Several academic search engines currently available are widely known to the scientific community, such as Google Scholar (GS), Web of Science (WoS) and Scopus [12]. GS is a free service that uses web crawlers for retrieving scholarly publications from several sources available on the Internet, while WoS and Scopus are indexing systems for specifically selected sources. Recent articles on this theme have shown that GS has progressively improved and is at a mature stage that enables it to compete with the more traditional source-enclosed indexing services [12, 13]. Furthermore, GS, according to its website, provides by default search results ordered by relevance, measuring it not only by the number of citations but also considering the full text of each source as well as the source's author and the publication in which the source appeared. The use of the GS ranking to obtain a certain number of the most relevant articles has been used in several relevant reviews, such as Hall [14] and Tabuenca et al. [15]. Therefore, for the present analysis, GS was the chosen search engine for selecting the most relevant articles on DM applications to credit risk.

The search query included every publication that contained both 'credit risk' and 'data mining' within the 2010–2014 timeframe, with both the 'include patents' and 'include citations' unchecked, and leaving the default option of 'sort by relevance' for allowing the most relevant hits to be presented as the top pages. Also, it should be stressed that only English language journals were considered because English is the major research dissemination language, although this left a large number of publications out of the study [16]. Additionally, the proposed automated approach would not be viable considering the need to fully translate each article to a common language and the fact that most human languages have an intrinsic subjectivity; a direct word to word translation does not usually encompass this subjectivity [17]. The search was executed on 16 May 2015, resulting in 2970 hits. Then, articles began to be collected starting at the top of the first page, benefiting from the sort by relevance of GS. It should be noted that only journal and conference articles were included—books, book chapters, presentations and other sorts of material were excluded. From 16–18 May 2015 articles were collected one by one, and each one of them was evaluated to validate if it matched the subjects in analysis— only those that matched were included. This process stopped when the number of valid articles required for the study reached 100. To achieve that figure 116 articles—16 of those, most of them in the latter pages of the search results list were not relevant for the present analysis and were excluded. Hence, above a certain threshold of articles it became less probable that relevant articles would be found. Therefore, the subsequent steps of this literature analysis were performed on the 100 most relevant articles in DM applications to credit risk, according to GS.

Text Mining

TM allows extracting knowledge from unstructured data, for example, a collection of texts [18]. Therefore, it can be useful in the process of analyzing a large number of literature publications, providing an automated mean of summarizing the article contents. Previous works followed two distinct approaches: by extracting all the lexical words, excluding only the more common words such as pronouns [11]; or by using specific dictionaries of terms composed of one or more words [19] defined by experts in the studied subjects [4]. The current work followed the latter approach, with a significant enhancement: instead of asking the assistance of experts to define an unguided dictionary, all the keywords for the 100 articles were collected, resulting in a list with 466 words. Then all duplicates were removed, and similar words in different formats (i.e. singular versus plural) were also reduced to a single word. Also, common or too generic terms such as 'banking' or 'data mining' were removed. This literature analysis focussed on specific DM methods and techniques applied to a subset of problems within credit risk and as such these words were considered to be too general. Finally, the dictionary of equivalent terms was built with the remaining 111 terms, as displayed in Table 6.1 (similar terms in different formats such as singular versus plural are not shown, for simplification purposes).

Table 6.1 is divided in two areas: the first for the specific DM methods, and the second for sub-problems within credit risk. Also, abbreviations were included when these do not have a meaning in the English language (to avoid a mismatch) and because they were frequently used in the articles. It should be emphasized that the procedure of using articles' keywords is less prone to the subjectivity associated with human experts' definition of relevant terms.

For the experiments, the R statistical environment was adopted because it has the advantage of being open source with the support of a large community, which provides a vast number of packages in a wide range of applications. Moreover, the 'tm' package was chosen for the TM functions, and the 'word cloud' package for generating visually appealing word clouds, with a few other packages also included for supporting auxiliary functions.

Using the dictionary and the R packages, the procedure adopted can be summarized as follows:

1. Create a corpus of documents (i.e. articles).
2. Remove extra spaces and convert all words into lower case for simplifying term matching.
3. Apply a transformation to convert every equivalent term in the dictionary into a unique term.
4. Build the document term matrix.

The document term matrix is a bi-dimensional matrix that counts the frequency that each term (in columns) occurs in each of the documents (in rows). Such structure is of paramount relevance for TM because it is the basic input for constructing easy to interpret structures such as a table of frequencies and a word cloud.

Table 6.1 Dictionary of terms (in lower case)

Term	Equivalents
Data mining methods and techniques	
ANFIS	Adaptive neuro-fuzzy inference system, adaptive network-based fuzzy inference system
Bagging	bootstrap aggregating
Bayesian	
Case-based reasoning	CBR
Clustering	Self-organizing map, k-nearest neighbour
Data quality	Information quality
Decision support system	DSS, expert systems
Decision tree	DT, random forest, rotation forest, chaid
Discriminant analysis	CPDA, LPDA
Ensemble	
Feature selection	Filtering, variable selection
Genetic algorithm	GA
Hybrid	bayesian ensemble hybrid particle swarm credit card
Logistic regression	LR
Multiple criteria	MCDM
Neural network	NN, ANN, multilayer perceptrons
Particle swarm	
Rough set	Set theory, fuzzy sets
Sampling	Sample selection, random subspace
Support vector machine	SVM
Credit risk sub-problems	
Bankruptcy	Insolvency, default detection, early warning, financial distress
Credit card	
Fraud	Money laundering
Regulatory	Loss given default, probability of default, IRB, basel
Scoring	Credit risk classification, rating

Topics of Articles

A simple TM consisting in word counting and displaying summarized information about the contents of documents may be interesting on its own for providing some insights. However, it is more informative to build a body of knowledge using clustering DM techniques for unveiling previously unknown trends that may enrich understanding on a given subject. The main goal is to group articles in logical clusters, which may be characterized by some common denominators. For this task, the LDA algorithm was adopted for both its simplicity and effectiveness given the large number of publications. This technique provided interesting results [20].

LDA is the most popular topic-analysis method, with applications in a wide range of domains. It produces a weighted list of topics for every document in a collection dependent on the properties of the whole [20]. The number of topics is a needed input for LDA. Following similar approaches [4, 11], this value was set to half of the terms considered. Each word w_j has a probability of matching a given topic z_i given by Equation 1.

$$\beta_{ij} = p\,(w_j = 1 | z_i = 1); \qquad (1\text{—}\beta \text{ distribution}) \quad (6.1)$$

Thus the LDA computes a bi-dimensional probability matrix β of k topics versus V different words. Therefore β provides a simple metric for measuring the relation of each term to a given topic. A value closer to zero indicates a stronger relation to that topic [21]. For the experiments presented next section, the R package 'topicmodels' was adopted, since it can be fed directly with the document term matrix produced from the 'tm' package, thus facilitating the procedure.

The LDA output is a tri-dimensional matrix encompassing terms, documents and topics built by the algorithm. Thus, for every topic it is possible to obtain a measure of its relationship to one of the dictionary terms through the β distribution. Also, for every document it is possible to identify which topic it best fits. Considering the 25 reduced terms defined in Table 6.1, the 100 articles, and the 13 topics, gives a structure containing 32,500 values. Since the goal is to analyze the groups of articles represented by the topics and its characterization and more specifically different DM approaches to credit risk problems, for each topic only the most relevant credit risk problem and the most relevant DM method (as measured by the β distribution) are scrutinized.

Proposed Approach

This section describes the whole approach undertaken and draws on the procedures described in Sects. 2.2 and 2.3, this approach is illustrated in Fig. 6.1.

The input is the set of articles collected through the methods explained in Sect. 2.1. From each article, the title, abstract, keywords and body text are retained, discarding images and also the references section, the latter to avoid the inclusion of terms that appear only in the titles of each reference. These articles constitute the corpus of documents used for the TM procedures. The TM procedure then takes place over the corpus by preparing and analyzing the contents of each document. A code excerpt presented next illustrates such steps:

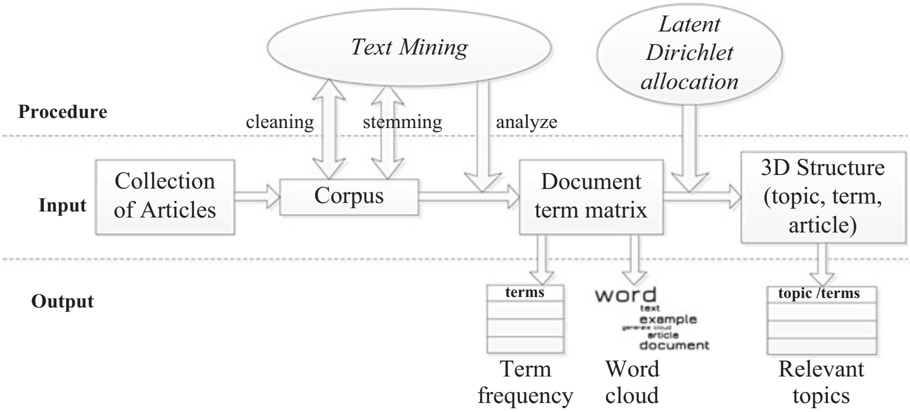


Fig. 6.1 Proposed approach

```
articles <- Corpus(DirSource(path), readerControl = list(language
= 'en'))

articles <- tm_map(articles, content_transformer(stripWhitespace))
# remove extra space

articles <- tm_map(articles, content_transformer(tolower)) # lower
case

equivTerms <- stemFromFileLoad('equivalent.txt')

reducedDictionary <- as.vector(intersect(unique(equivTerms[[1]]),
dictionary))

articles <- tm_map(articles, content_transformer(function(x)
stemFromFile(doc=x, equivTerms=equivTerms)))

phraseTokenizer <- function(x) RWeka::NGramTokenizer(x, Weka_
control(min = 1, max = 6))

dtm <- DocumentTermMatrix(articles, control = list(

  tokenize = phraseTokenizer,

  dictionary = reducedDictionary))

dtmMatrix <- as.matrix(dtm)

v <- sort(colSums(dtmMatrix),decreasing = TRUE)

d <- data.frame(word = names(v), freq = v) # term frequency list

wordcloud(d$word,d$freq) # generate the word cloud
```

First, the corpus of documents is read from the file system. Then, extra spaces are removed and all contents are converted to lower case, to allow for a direct comparison. Next, stemming occurs for finding all relevant keywords (the right-hand column of Table 6.1) and transforming them to the corresponding reduced terms (the left-hand column from Table 6.1), in replacement of their equivalents. The document term matrix is built on the analysis of the documents and retaining the ones existing in the reduced dictionary. Furthermore, it should be noted that a tokenizer is used for finding the relevant terms considering each of them may be constituted of one to six words (e.g. 'adaptive network-based fuzzy inference system'). Finally, besides the document term matrix, used as an input for the LDA, the other two direct outputs are the term frequency list and the word cloud.

The execution of the LDA model is very simple using the 'topicmodels' package, by simply invoking the LDA function with both the document term matrix (dtm) and the number of topics to be modelled in the two parameters. Then the relation to each term is collected, as well as the most relevant topic in which each article is fitted initially (see code below).

```
lda <- LDA(dtm, 13)
terms <- terms(lda, length(reducedDictionary))
topics <- topics(lda,1)
```

The resulting approach is relatively straightforward, allowing it to be applied to other numerous contexts involving text analysis (e.g. analysis of online news).

3 Results and Analysis

Articles

In this section, the articles selected are summarized in three categories: publication names (Table 6.2), publication types (Table 6.3) and publishers (Table 6.4). For Tables 6.2 and 6.4 the publication names/publishers that contribute with just one article are not presented, for page space optimization purposes only. Notably, *Elsevier's Expert Systems with Applications* journal contributed with 33 articles, helping to consolidate Elsevier's dominant position, with 62 articles. In fact, from the publications contributing more than one article (Table 6.2), only the *International Journal of Neural*

Table 6.2 Journals containing the articles

Publication name	Number of articles
Expert Systems with Applications	33
Applied Soft Computing	6
Decision Support Systems	4
European Journal of Operational Research	3
Knowledge-Based Systems	2
Procedia Computer Science	2
International Journal of Neural Systems	2
International Journal of Forecasting	2
Information Sciences	2

Table 6.3 Publication types for the articles

Publication type	Number of articles
Journal	92
Conference	8

Table 6.4 Publisher names for the articles

Publisher name	Number of articles
Elsevier	62
IEEE	7
Springer	5
Wiley	3
AIRCC	2
World Scientific	2

Systems is not published by Elsevier. It should also be noted that from Table 6.2, most of the journals are strongly technology related, with the exception of the *International Journal of Forecasting* and the *European Journal tof Operational Research* that are more management related, although also both encourage contributions benefiting from technology approaches. Finally, Table 6.2 includes only journals, emphasizing the result from Table 6.3, which shows that 92 articles from the 100 are published in journals. Hence, GS relevance order appears to favour journals.

Text Mining

Following the experiments on the articles' contents with TM, the frequency of the relevant terms defined in the dictionary is show on Tables 6.5 and 6.6. On the left, the results for credit risk problems are presented. It is shown that credit scoring accounts for more than half of the credit risk problems being

Table 6.5 Frequency of terms for credit risk

Term	#	%
Scoring	2082	57.0
Bankruptcy	710	19.4
Credit card	394	10.8
Regulatory	339	9.3
Fraud	129	3.5
Total	3654	100.0

Table 6.6 Frequency of terms for DM

Term	#	%
Neural network	1834	18.7
Support vector machine	1397	14.2
Decision tree	1182	12.1
Ensemble	920	9.4
Logistic regression	675	6.9
Hybrid	616	6.3
Clustering	604	6.2
Sampling	522	5.3
Genetic algorithm	385	3.9
Feature selection	371	3.8
Bagging	302	3.1
Discriminant analysis	252	2.6
Multiple criteria	206	2.1
Rough set	130	1.3
Bayesian	108	1.1
Anfis	80	0.8
Decision support system	76	0.8
Data quality	72	0.7
Case-based reasoning	55	0.6
Particle swarm	17	0.2
Total	9804	100.0

All the values shown were rounded to the first decimal.

addressed by DM methods for the selected set of 100 most relevant articles. Next appears bankruptcy, a subject that has been largely debated due to the impact of the financial crisis. Credit cards still receive a lot of attention, even though it is a subject that has also been widely studied prior to the crisis [22]. Regulatory projects including DM get almost 10 % of the attention, being one of the subjects highly boosted by the crisis [23]. Finally, fraud is the least studied from the subjects in analysis, with 3.5 %. Nevertheless, it can be argued that the improvement in credit scoring evaluation tends to reduce fraud [24].

In Table 6.6, results are shown for the frequency of DM terms. Neural networks are advanced machine learning techniques that try to mimic human brain action using artificial neurons for apprehending non-linear relations between input variables [25]. This technique has been largely studied in the

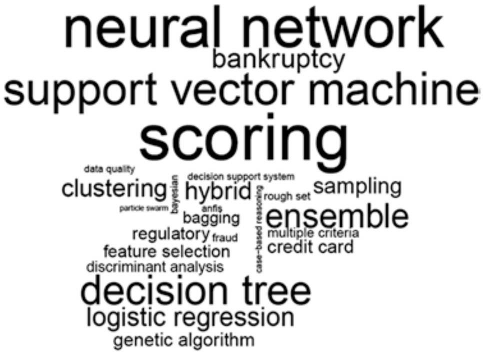


Fig. 6.2 Word cloud

literature [26], with good modelling results, therefore, it is a good candidate for enhancing solutions for credit risk problems, appearing at the top of our list, with almost 19 % of the total occurrences. A support vector machine is the second most mentioned DM method. This modelling technique emerged in the 1990s [27] and became one of the most complex and successful among those in the machine learning domain. The traditional decision trees stand the test of time, in third place, when considering the 2010–2014 timeframe. In fourth comes ensemble modelling, in which a few different techniques are combined for obtaining a better result than any of the isolated methods. Several other modelling techniques are included in the list. Also methods for selecting the appropriate records (e.g., sampling and feature selection) have been studied for improving credit risk assessment. Surprisingly, data quality, a key issue for large DM projects, particularly in the regulatory domain [28], is still weakly associated with credit risk. This is an interesting gap for researchers to fill. Figure 6.2 complements Tables 6.5 and 6.6 by visually displaying the differences between terms in the dictionary thus enhancing understanding.

Topics of Articles

As stated in Sect. 2.3, the result of LDA is a set of topics grouping articles logically according to the frequency of each term in the dictionary. Table 6.7 summarizes the findings under 13 topics, showing the number of articles as well as the most relevant credit risk domain and DM method in each topic. Also for exemplification purposes, one article was selected from each topic.

By looking at Table 6.7 it is easy to see that scoring is receiving the most attention, with more than half of the articles (53 out of 100) and six topics. Nevertheless, according to the β values, for three of the topics (totalling 22 articles) with β greater than four, it is a weak relation, even though scoring is

Table 6.7 Topics of articles

	Credit risk		DM method		
#	Term	β	Term	β	Example
19	Scoring	0.23	Logistic regression	2.25	Yap et al. [29]
12	Scoring	4.64	Neural network	0.23	Khashman [26]
7	Scoring	5.44	Decision tree	0.08	Mandala et al. [30]
6	Scoring	3.75	Hybrid	0.59	Tsai and Chen [31]
6	Scoring	2.77	Feature selection	0.59	Marinaki et al. [32]
3	Scoring	4.55	Sampling	0.42	Zhou et al. [33]
53	Total scoring articles				
10	Bankruptcy	5.81	Ensemble	0.18	Verikas et al. [34]
5	Bankruptcy	0.28	Clustering	2.19	De Andrés et al. [35]
4	Bankruptcy	4.88	Genetic algorithm	0.30	Oreski and Oreski [10]
19	Total bankruptcy articles				
9	Fraud	1.94	Clustering	1.00	Wu et al. [36]
8	Fraud	5.79	Support vector machine	0.03	Hens and Tiwari [37]
17	Total fraud articles				
6	Credit card	0.26	Neural network	3.37	Chen and Huang [38]
5	Regulatory	0.26	Data quality	2.22	Moges et al. [28]

the credit risk problem more closely related to those topics. Such findings reveal that those three topics are more closely related with DM techniques exploration than with benefiting credit risk scoring. In fact, both the examples selected for two of those topics confirm this hypothesis: Khashman's [26] work focused on different neural models and learning schemes, while Zhou et al. [33] used the nearest sub-space method for improving classification, which is a technique based on training samples' selection. The two papers are highly technology related. The exception is the example chosen for Mandala et al. [21] which tries to improve credit risk scoring in a local context through decision trees. Still analyzing scoring topics, it is also interesting to note that the topic that includes most of the articles (19) is related to logistic regression, which is one of the most basic techniques. Perhaps this is another indicator that there is still room for research in advanced DM techniques that can translate to a direct improvement in credit scoring.

Bankruptcy is the second most mentioned credit risk problem, including 19 articles and three topics. Again in an almost repeat of what was observed for scoring, the two topics mostly related with the advanced DM methods are ensemble modelling (combination of a few techniques for improving the isolated techniques' results) and genetic algorithms are barely related to bankruptcy, with β values of 5.81 and 4.88, respectively. Such findings reinforce the suspicion that DM research is still failing to explicitly focus on the benefits for bankruptcy, as occurred with scoring. On the other hand, bankruptcy is strongly associated with clustering with five articles, which provides a mean for grouping enterprises in terms of bankruptcy risk.

Fraud also receives some attention, with 17 articles, being more related to clustering for the same reasons as bankruptcy, than with a support vector machine, which is very weakly associated with fraud. Such result contrasts with Table 6.5 and may be justified by a higher concentration of credit fraud issues in a smaller number of articles. The remaining two topics show two median relations: between credit cards and neural networks; and between regulatory issues and data quality. The former topic shows through its example [38] a more mature trend of research where advanced neural networks are applied toward a solution for a real credit risk problem. As stated previously, such a trend has been widely studied prior to the crisis [22]. The latter topic emphasizes a real problem that emerged particularly after the crisis, given governmental pressures to audit financial assets and liabilities: financial institutions are posed with a huge problem of reorganizing management information systems to improve data quality in order to respond to an increasingly large number of regulatory reports. In fact, the demand for highly detailed reports has emphasized a growing pressure over financial institutions and justifies the significant relation found between data quality and regulatory issues (β values of 2.22 and 0.26, respectively).

4 Conclusions

Credit risk poses several interesting problems where solutions can benefit directly from DM approaches. Some of the most widely studied problems include credit scoring, bankruptcy, credit fraud, credit cards and regulatory issues. The 2008 global financial crisis proved that previous solutions were not adequate to predict credit risk on a global scale, although specialized DM approaches to problems such as credit cards provided an already effective method. Particularly, bankruptcy and regulatory issues have received significant analytic attention in the 2010–2014 post-crisis period.

This chapter presents an automated literature analysis approach to credit risk problems being addressed by DM methods. The automation included the usage of TM for analyzing contents and the LDA for organizing the identified articles into topics. Of the relevant articles 100, including both 'credit risk' and 'data mining', were selected for analysis, according to Google Scholar's relevance criterion.

Credit scoring is by far the most mentioned credit risk problem, followed by bankruptcy and fraud. The most cited DM techniques include neural networks and support vector machines, which are two advanced methods, showing that these can be directly applied to credit risk problems. Ensembles that try to bring the best features of a few established techniques by combining their results are also figure largely in credit risk problems.

By analyzing the topics built on the LDA algorithm, one of the major conclusions is that research on the most advanced and recent DM methods and techniques such as support vector machines and ensembles is more focused on a fine tuning of those techniques than in assessing real benefits for credit risk. More work needs to be done to take advantage of those techniques in relation to real-world credit risk applications, thus providing an interesting research opportunity. Another finding is that regulatory issues demand research in data quality. Such a trend is directly related to the huge increase, in the post-crisis period, of highly detailed regulatory reports that sustain more frequent auditing processes for the financial institutions.

The full approach undertaken can, potentially, be applied to any kind of literature analysis. In fact, it can also be used for analyzing other collections of texts such as comments within a website. Furthermore, the approach is both flexible and extensible: a full English word analysis may be used instead of a specific dictionary, and other clustering or topic analysis may also be applied.

References

1. Claessens, S., Kose, M. A., Laeven, M. L., and Valencia, F. (2014). Financial Crises: Causes, Consequences, and Policy Responses. *International Monetary Fund*.
2. Nijskens, R., & Wagner, W. (2011). Credit risk transfer activities and systemic risk: How banks became less risky individually but posed greater risks to the financial system at the same time. *Journal of Banking & Finance, 35*(6), 1391–1398.
3. Galati, G., & Moessner, R. (2013). Macroprudential policy—A literature review. *Journal of Economic Surveys, 27*(5), 846–878.
4. Moro, S., Cortez, P., & Rita, P. (2015). Business intelligence in banking: A literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation. *Expert Systems with Applications, 42*(3), 1314–1324.
5. Tobback, E., Martens, D., Van Gestel, T., & Baesens, B. (2014). Forecasting Loss Given Default models: Impact of account characteristics and the macroeconomic state. *Journal of the Operational Research Society, 65*(3), 376–392.
6. Macey, J. R. (2012). Regulator effect in Financial regulation. *The Cornell Law Review, 98*, 591–636.
7. Turban, E., Sharda, R., Delen, D. (2011). Decision support and business intelligence systems, (9th edn), New Delhi, Pearson.

8. Marques, A. I., García, V., & Sanchez, J. S. (2013). A literature review on the application of evolutionary computing to credit scoring. *Journal of the Operational Research Society, 64*(9), 1384–1399.
9. Guerrero-Baena, M. D., Gómez-Limón, J. A., & Fruet Cardozo, J. V. (2014). Are multi-criteria decision making techniques useful for solving corporate finance problems? A bibliometric analysis. *Revista de Metodos Cuantitativos para la Economia y la Empresa, 17*, 60–79.
10. Oreski, S., & Oreski, G. (2014). Genetic algorithm-based heuristic for feature selection in credit risk assessment. *Expert Systems with Applications, 41*(4), 2052–2064.
11. Delen, D., & Crossland, M. D. (2008). Seeding the survey and analysis of research literature with text mining. *Expert Systems with Applications, 34*(3), 1707–1720.
12. Harzing, A. W. (2013). A preliminary test of Google Scholar as a source for citation data: A longitudinal study of Nobel prize winners. *Scientometrics, 94*(3), 1057–1075.
13. De Winter, J. C., Zadpoor, A. A., & Dodou, D. (2014). The expansion of Google Scholar versus Web of Science: A longitudinal study. *Scientometrics, 98*(2), 1547–1565.
14. Hall, C. M. (2006). The impact of tourism knowledge: Google scholar, citations and the opening up of academic space. *E-Review of Tourism Research, 4*(5), 119–136.
15. Tabuenca, B., Kalz, M., Ternier, S., and Specht, M. (2014). Mobile authoring of open educational resources for authentic learning scenarios. Universal Access in the Information Society, Special Issue: The Use of Mobile Technology and Ubiquitous Computing for Universal Access in Online Education, pp. 1–15.
16. Crystal, D. (2012). *English as a global language*. Cambridge: Cambridge University Press.
17. Hatim, B., & Mason, I. (2014). *Discourse and the Translator*. London: Routledge.
18. Fan, W., Wallace, L., Rich, S., & Zhang, Z. (2006). Tapping the power of text mining. *Communications of the ACM, 49*(9), 76–82.
19. Soper, D. S., & Turel, O. (2012). An n-gram analysis of communications 2000–2010. *Communications of the ACM, 55*, 81–87.
20. Campbell, J. C., Hindle, A., & Stroulia, E. (2014). Latent Dirichlet allocation: Extracting Topics from Software Engineering Data. In M. Bird & T. Zimmermann (Eds.), *The art and science of analyzing software data* (1 ed.). Burlington, MA: Morgan Kaufmann.
21. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research, 3*, 993–1022.
22. Huang, C. L., Chen, M. C., & Wang, C. J. (2007). Credit scoring with a data mining approach based on support vector machines. *Expert Systems with Applications, 33*(4), 847–856.

23. Gerding, E. F. (2009). Code, crash, and open source: The Outsourcing of Financial Regulation to Risk Models and the Global Financial Crisis. *Washington Law Review, 84*, 127–198.
24. Siddiqi, N. (2012). *Credit risk scorecards: Developing and implementing intelligent credit scoring* (Vol. 3). New York: Wiley.
25. Moro, S., Cortez, P., & Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems, 62*, 22–31.
26. Khashman, A. (2010). Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes. *Expert Systems with Applications, 37*(9), 6233–6239.
27. Tian, Y., Shi, Y., & Liu, X. (2012). Recent advances on support vector machines research. *Technological and Economic Development of Economy, 18*(1), 5–33.
28. Moges, H. T., Dejaeger, K., Lemahieu, W., & Baesens, B. (2013). A multidimensional analysis of data quality for credit risk management: New insights and challenges. *Information & Management, 50*(1), 43–58.
29. Yap, B. W., Ong, S. H., & Husain, N. H. M. (2011). Using data mining to improve assessment of credit worthiness via credit scoring models. *Expert Systems with Applications, 38*(10), 13274–13283.
30. Mandala, I. G. N. N., Nawangpalupi, C. B., & Praktikto, F. R. (2012). Assessing Credit Risk: An application of data mining in a Rural Bank. *Procedia Economics and Finance, 4*, 406–412.
31. Tsai, C. F., & Chen, M. L. (2010). Credit rating by hybrid machine learning techniques. *Applied Soft Computing, 10*(2), 374–380.
32. Marinaki, M., Marinakis, Y., & Zopounidis, C. (2010). Honey bees mating optimization algorithm for financial classification problems. *Applied Soft Computing, 10*(3), 806–812.
33. Zhou, X., Jiang, W., & Shi, Y. (2010). Credit risk evaluation by using nearest subspace method. *Procedia Computer Science, 1*(1), 2449–2455.
34. Verikas, A., Kalsyte, Z., Bacauskiene, M., & Gelzinis, A. (2010). Hybrid and ensemble-based soft computing techniques in bankruptcy prediction: A survey. *Soft Computing, 14*(9), 995–1010.
35. De Andrés, J., Lorca, P., de Cos Juez, F. J., & Sánchez-Lasheras, F. (2011). Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS). *Expert Systems with Applications, 38*(3), 1866–1875.
36. Wu, D. D., Chen, S. H., & Olson, D. L. (2014). Business intelligence in risk management: Some recent progresses. *Information Sciences, 256*, 1–7.
37. Hens, A. B., & Tiwari, M. K. (2012). Computational time reduction for credit scoring: An integrated approach based on support vector machine and stratified sampling method. *Expert Systems with Applications, 39*(8), 6774–6781.
38. Chen, S. C., & Huang, M. Y. (2011). Constructing credit auditing and control & management model with data mining technique. *Expert Systems with Applications, 38*(5), 5359–5365.

7

Intelligent Credit Risk Decision Support: Architecture and Implementations

Paulius Danenas and Gintautas Garsva

1 Introduction

The expansion of loan portfolios and the potential to offer new financial products to customers are among the main objectives of financial institutions. Precise quantitative analysis and scoring are some of the approaches they take to achieve these goals. In particular, the accuracy in credit scoring must be ensured in order to minimize losses, while at the same time optimizing loan portfolio performance by quantifying risk of different loans and instruments. For such goals to be achieved, a considerable amount of research has been performed to develop new techniques and enhance existing ones while predominately focusing on classification and forecasting as part of statistical learning. The financial sector is specifically interested in recent developments and applications of novel techniques in the area of artificial intelligence (AI) and in particular machine learning (ML). Various applications of AI have provided examples of their superiority over traditional statistical and mathematical techniques. For the most part, AI overcomes many of the limitations associated with these older and more rigid methodologies while presenting analysts with new challenges. Recent research

P. Danenas (✉) • G. Garsva
Center of Information Systems Design Technologies, Kaunas University of Technology, K. Baršausko Str. 59, Kaunas, Lithuania

Department of Informatics, Kaunas Faculty, Vilnius University, Muitines Str. 8, Kaunas, Lithuania

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets*,
New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0_7

has also shown that AI is capable of driving the most complex and computationally demanding decision processes. This has led to the development of numerous models to evaluate credit risk and has also attracted a lot of attention from both academic researchers and financial analysts in financial institutions.

The rest of the chapter is structured to provide an informative analysis of AI and its application to the area of credit risk. Section 2 describes the main AI techniques applied to decision support in credit risk evaluation. Due to the large number of existing techniques, it is limited to the description of three core techniques that we consider to be current and relevant. Section 3 covers recent developments in decision support systems (DSS) applications relevant to credit risk. DSS provides high-level sophisticated, automated and uniform decision making at different levels, which simplifies and improves management tasks, as well as reduces costs of providing existing services. This section reviews the taxonomies of DSS structures at both architectural and decision support levels, identifies their future trends and components, which must be developed to support their functionality. A considerable amount of attention is also given to the requirements for the development of such systems. In addition, Section 3 describes a recent framework for developing a novel DSS based on AI techniques and financial data exchange standards that support real time development and an update of credit risk models.

2 Literature Review

In this section three relevant AI (particularly ML) techniques, often applied to develop credit risk evaluation models, mainly artificial neural networks (ANN), support vector machines (SVM) and decision trees, are reviewed, with particular focus on classification. While Sections 2.1 and 2.2 present core concepts of these techniques and ML techniques overall, Section 2.3 describes credit risk problems, which can be solved using these techniques, together with examples of related research.

Machine Learning Techniques

The *machine learning* paradigm is one of the sub-paradigms in the field of AI targeted at computational imitation of human reasoning and behaviour. Computational intelligence principles can also be carried out to explain and simulate the behaviour of different living species or their groups, applying their rationale to produce optimal solutions. Generically it can be described as a process with an objective of learning from examples and obtaining the most generalizing structure, using various operators to produce the best predicted results for the output of any given generator. This is performed by mapping input vectors x to output vectors y, using a probability distribution function $f(x)$, and minimizing a loss function. *Data mining* (also often referred to as or used in context of *knowledge discovery*) can be considered as an individual field that complements statistical learning with ML algorithms and methodology for their development and application. One of its main concerns is data acquisition— its preprocessing, imputation, analysis and storage; CRISP-DM (Cross Industry Standard Process for Data Mining) methodology [1] provides guidelines for applying data mining techniques on actual data.

Pattern recognition (identification of a connection for monitored instances to one of k classes) and *classification* (prediction of the class for an unseen input vector) are two of the most popular and widely applied techniques in credit risk, which are solved by ML. The latter is often referred to as *supervised learning*, due to its use of given identified observations. This chapter is limited to classification problems in credit risk motivated by the number of problems solved using these sorts of techniques. We will use the formal classification definition similar to the given by Dunham [2]: given a database $D = \{t_1, t_2, \ldots, t_n\}$ and a set of classes $C = \{C_1, C_2, \ldots, C_n\}$ the classification problem is a mapping $f: D \rightarrow C$, where each t_i is assigned to one of the classes from set C. $C_j = \{t_i | f(t_i) = C_j, 1 \leq i \leq n, t_i \in D\}$, that is, class C_j contains only entries assigned with it. The class is usually represented as nominal value. Another type of learning—*ensemble learning*—deals with combining outputs from multiple classifiers, using various strategies, such as majority voting. Many authors have shown that ensemble classifiers may produce more accurate results than single classifiers, when they are applied to problems relevant to credit risk [3–6].

Classification techniques can be used to solve tasks, providing multidimensional feature inputs, represented as numerical or nominal (categorical) feature vectors often returning robust-to-noise solutions. The training time may vary depending on the technique, its implementation, the optimization algorithm and its configuration, which are all selected during the training phase. The training of ANN- and SVM-based models is more computationally demanding, compared with other ML techniques, such as decision trees or rule classifiers. However, applying trained models to predict classes for new instances would typically be very quick. ANN and SVM frequently show better results in terms of accuracy, compared to other techniques such as single decision trees. The level of interpretability of the generated model may also vary as techniques such as ANN and SVM produce so-called 'black-box' models. These are difficult to interpret and explain although techniques for rule extraction help to eliminate this problem, producing a set of logical rules. Diederich [7] provides an extensive review of such techniques, applied for SVM classifiers.

Techniques for Classification

This section provides a concise introduction to three relevant classification techniques for credit risk related research. For comprehensive description and links on other techniques, along with more information on techniques in this chapter, the reader is referred to relevant ML literature [8, 9], as well as extensive surveys of relevant work, focused on intelligent techniques in the credit risk domain [3, 5, 10–13].

ANNs are the most widely known ML technique that offers a large level of flexibility to estimate or approximate internal relationships that can depend on large number of inputs. The idea of an ANN is based on imitation of the capabilities of the interconnected neurons in the human brain to recognize, identify, adapt and learn from patterns and ultimately to process this information. Multiple ANN definitions can be found in the relevant literature; it can be defined as a constrained oriented graph [2] or as a non-linear mapping [14]. We will define an ANN as 3-tuple $\Sigma = <G, A, MT>$, where G is a graph, describing an ANN structure, A is an ANN learning algorithm, and MT is a set of information processing techniques. The ANN may also be described in terms of neurodynamics and architecture. Neurodynamics describes the features of each neuron, and the architecture is defined by an overall ANN structure, such as the number of hidden layers, the number of hidden neurons in each layer, the transfer (transformation, activation) functions and the starting weights initialization. The choice of architecture is sophisticated and often performed manually, although heuristic techniques, such as genetic algorithms or swarm intelligence, may be employed to automate this task.

Another criticism of ANN concerns its tendency to get trapped in local minimums, if such exist. Nevertheless, this is a very flexible technique, with numerous algorithms and modifications available. ANN was one of the first AI techniques that was intensively developed, improved and applied in the business and financial sectors [15] and still remains one of the most relevant techniques in the credit risk domain. It often outperforms other similar techniques [16–19].

SVM [20, 21] are learning machines, capable of performing binary classification or regression tasks by mapping n-dimensional input data space to another feature space, which can be used in linear classification. At the same time the empirical classification error is minimized and the geometrical margin is maximized. Due to such characteristics SVM is also called a maximum margin classifier. The classifier may utilize similarity measures between data instances in form of a kernel *function* $K:\left(x,x'\right)\rightarrow k\left(x,x'\right);x,x'\in\mathbb{R}$, which returns a scalar characterizing the similarity between x and x'. Such representation gives an opportunity to develop a large set of various learning algorithms, as well as to use some form of expert knowledge, resulting in more precise and appropriate non-linear mapping with respect to the specifics of a particular problem. Formally, the SVM classifier is described as a separating hyperplane with binary solutions on both of its sides (i.e. solutions equal to +1 or −1), minimizing margin error. This hyperplane is described as a set of support vectors—instances for which and only for which Lagrangian is not equal to zero. Finding these vectors from training data can be formulated as a solution of the following convex optimization problem [21, 22]:

$$\min_{w,b,\zeta}\frac{1}{2}\|\mathbf{w}\|^2+C\sum_{i=1}^{n}\zeta_i^k \tag{7.1}$$

$$s.t.\quad y_i\left(w^T\phi\left(x_i\right)+b\right)\geq 1-\zeta_i,\ \zeta_i\geq 0,\ i=1,..,n$$

where C is a regularization parameter that determines the trade-off between the maximum margin and the minimum classification error, which comes from data left on the 'wrong' side of a decision boundary (data inside so-called soft margin, i.e. margin that allows mislabelled examples, considering the degree of misclassification, defined by the ζ variable). C is also referred to as a penalty parameter that determines the trade-off between the training error and Vapnik-Chervonenkis dimension of the model.[1] This is usually solved for $k = 1$ or $k = 2$, and such soft margin SVMs are referred as L1-SVM and L2-SVM [98]. The decision function for SVM is defined as

$$\langle\phi(\mathbf{x})\cdot\mathbf{w}\rangle+b=0 \tag{7.2}$$

Several SVM implementations have been developed, such as LibSVM [22], including canonical C-SVM implementation and υ-SVM [23], least squares SVM or LS-SVM [24], efficiently simplifying the SVM problem into a solution of a set of linear equations, SVM$^{\text{Light}}$ [25], linear SVM implementation

[1] The Vapnik-Chervonenkis dimension, or VC dimension, measures the flexibility of the classification algorithm [97].

LIBLINEAR [26]. These are frequently used as a basis for research and the development of new algorithms.

Decision trees (DT) are one of the oldest and most widely applied ML techniques. They form a tree type pattern consisting of a set of *if ... then* rules [27]. The classification is performed by recursively partitioning each sub-set of the source set, obtained at the previous iteration, using an attribute value test, based on measures, such as information gain,[2] to determine a variable that would best split this sub-set. The search is terminated if partitioning can no longer influence predictability. This algorithm can be applied when the attributes are either nominal or numerical and it can also perform well when the dataset contains missing or erroneous values. However, it may suffer from overtraining and unsatisfactory generalization and as a result the performance of the decision tree may be relatively poor due to it being susceptible to noisy redundant attributes [28]. One of the most widely used DT extensions is the Random Forest algorithm [29], which performs classification using randomly generated tree ensembles and majority voting. It exploits the potential to develop a stable and well-performing classifier by combining several less stable ('weak') classifiers. This technique also incorporates several useful characteristics such as ability to rank variables by their importance and has been used by many practitioners and researchers [28], [30], [31].

Credit Risk Problems, Solved by Artificial Intelligence

As discussed in earlier sections, credit risk problems are often formulated as classification techniques. For example, the main goal is the development of mapping $F: X \rightarrow Y$, where $y \in Y$ is a representation of particular problem-dependent class (whether it is insolvency, loan classification, etc.); $\mathbf{x} \in X$ is the set of attributes that represent the financial conditions of a particular company that influenced y at some particular moment. Depending on the problem and data availability, different attributes can be used to form X — financial and managerial attributes are among the most selected types of data, although other types of data are also applied. Wang et al. [32] also attest that other types of data may be involved in scoring systems, depending on the type of credit. As their work is based on analysis of green loans (for funding solar, agricultural or other projects) it proposes the inclusion of inputs that reflect environmental and energy factors (e.g. environment quality situation, emission quantity, project green level), together with financial, managerial

[2] A measure based on entropy from the area of information theory.

qualitative factors. Bellovary et al. [33] surveyed more than 150 models, from 1930 to 2007, and identified 752 unique financial, economic, social and other variables, although 674 of them were utilized in just one or two studies. It turned out that financial ratios, representing profitability, financial leverage, debt situation, liquidity, capital structure, were dominating in such research. Another interesting finding was that a large number of factors does not necessarily result in an increased accuracy rate and predictive ability of the model. Our analysis, performed on relevant patents in this field to show the current situation in the practical application of intelligent techniques for financial risk evaluation, also confirmed that, besides financial data, other data might be used in real-world systems, for example, public, macroeconomic and/or mortgage data [11].

One of the most popular credit risk problems, solved using credit scoring techniques, concerns the applicability of the customer to deal with future obligations. By classifying customers into two categories, as ‘good’ and ‘bad’, the goal is to make a distinction between the candidates in order to determine suitability for a loan or an extension of an existing loan. This enables financial institutions, such as banks, to determine appropriate candidates in order to avoid and/or minimize losses. This type of problem is one of the main objectives for most credit risk related research. Most researchers choose to focus on binary classification [18, 31, 34], although multi-class classification, which may be more relevant to real-world applications, is also employed to produce solutions [97].

Financial distress and bankruptcy prediction are other types of problems solved by hybrid classification techniques. Given a set of certain attributes describing the financial situation of a company, the main goal is to obtain a likelihood that a company will fail or face financial difficulties. Depending on the availability of data for research, most researchers worked with datasets including data from different sectors. Yet, one can also find examples of differentiated research, such as banks [35], [36], construction companies [37], manufacturing industries [33], insurance companies [38], dotcom companies [39], restaurants [40] and social lending [31]. The country factor is another interesting dimension—intelligent techniques were applied on actual data from South Korea [34], Taiwan [41], Belgium and The Netherlands [42], China [43], Croatia [44], Portugal [37], Lithuania [45], [46], Turkey [35], Ireland [47] and various other countries. Financial data can also be studied using techniques that allow the mapping of numerous attributes to another space in order to extract significant features that can be used later for visualization or further modelling steps, such as regression analysis [48]. Self-organizing maps, often referred as Kohonen maps as initially applied by

Kohonen [49], are one of the most flexible clustering techniques used for failure prediction [50]–[52] and visualization [53].

Another type of relevant problem is behavioural scoring or performance scoring, which is applied after granting credit and used to monitor the financial condition of the obligor. The main goal is an estimation of the probability of default during a given period (for instance, 1 month and 3 month periods). The dataset is formed using historical loan repayment data together with factors, which describe the credibility of the obligors. The factors can be drawn both from loan application data (loan amount, term of loan, payment frequency and customer information), as well as behavioural data (loan balance and repayment factors). Kennedy et al. [47] used both application data and behavioural data, together with arrears repayment information. Their experiments were performed using arrears repayment information for 3-month, 6-month, 12-month, 18-month and 24-month outcome windows, showing that the 12-month window was the most optimal in terms of accuracy. Nevertheless, performance scoring faces the same challenges as credit scoring, also relevant to data mining, such as segmentation alignment, optimal feature selection and imbalanced learning. For an extensive coverage and review of such problems refer to Sun et al. and Tsai [54], [55].

The feature selection (FS) problem is relevant to almost every modelling problem, as the reduced number of properly selected modelling attributes may influence the quality of the final model as well as reduce the amount of information necessary for this model to be applied in practice. Two types of FS techniques can be identified: filter-based (using particular measure to select attributes by its value, such as correlation, information gain, similarity measure) and wrapper-based (performing selection iteratively according to the classifier performance). Attribute ranking can be viewed as one of the filtering-based FS forms. Several specific techniques were also proposed—Lin et al. [56] propose novel expert knowledge-based FS technique for financial distress prediction which, according to their research, outperformed generic techniques. The overall effect of different FS techniques has also been carefully studied in different sources [55], [57], [58].

Rating analysis is another type of credit risk problem that can benefit from a classification approach; it can also be referred to risk ratings reverse engineering [36]. Such approaches can lead to improved ratings by questioning and reinforcing their validity, as well as providing extracted patterns to distinguish companies or financial institutions rated by their financial abilities and probability of insolvency. Prediction of financial ratings' changes is another relevant problem that can make use of AI techniques. Jones et al. [5] performed empirical evaluation of different binary classifiers, particularly

ANN, logistic regression, random forests and SVM, using data for changes in international corporate credit ratings, occurring between 1983 and 2013. The dataset consisted of 5053 ratings changes (2,660 ratings downgrades and 2393 upgrades), where 2891 ratings relate to BBB- or above on S&P ratings scale (investment grade), and 2162 ratings changes are BB+ and below (speculative grade). Financial, corporate government, market, expert knowledge-based and macroeconomic variables, together with attributes reflecting changes from previous ratings, were used as inputs for their developed models. Their findings concluded that, besides their objective to obtain the relevant models, ensemble classifiers such as AdaBoost [59] or Random forest were the most successful choice, also resulting in robustness to outliers and missing values.

Ye et al. [60] performed research on supply chain disruption, which is a more specific sub-problem of financial distress. Although this is not directly related to credit decisions, economic loss might have a huge impact on the clients of a particular company leading to financial loss and possible bankruptcy. Financial institutions, which have issued credit to such companies, would also experience severe loss. Multi-class SVM classification was applied for modelling; the class marks the cause of disruption (demand disruption, supply disruption, product disruption or external disruption). Sample data consisted of data from 926 announcements (financial statements and economic performance) of 200 companies from 2007 to 2011 divided by sector (manufacturing, agriculture or financial industry). The results indicated a comparatively high performance of SVM for supply chain disruption identification problem.

3 Decision Support and Expert Systems for Credit Risk Domain

DSS and expert systems (ES) play a critical role in solving various financial and business problems, where data processing for deriving new information, yielding possible solutions or their alternatives is a significant part of the relevant computations. Section 3.1 gives a brief introduction to DSS and ES, discusses their goals and main differences from standard information systems (IS). Section 3.2 reviews the main types and taxonomies of DSS, while relating them to financial risk-oriented problems. Section 3.3 discusses recent developments in DSS for financial problems, related to credit risk, while Sect. 3.4 gives a number of requirements for modern DSS dedicated to banking decisions.

Furthermore, we discuss the development of the novel DSS based on AI techniques, described in Sect. 2, and modern financial Extensible Markup Language-(XML) based standards, by describing one such standards in Sect. 3.5. We introduce the developed framework in Sect. 3.6 with a discussion of the main topics related to its implementation.

Decision Support Systems: Definitions, Goals, Premises

The evolution and structure of DSS and ES are described in numerous sources [61]. Although such systems may share similar architectural patterns and goals, they also possess some differences. While DSS definition may be more relevant for modern decision support and automation, it is reasonable to begin with the review of ES, which pioneered decision support at the beginning of artificial-intelligence-based analysis.

ES can be defined as an AI-driven computational system that uses a knowledge base of human expertise to aid in problem solving. Several types of expert systems can be identified according to their types, for instance, Liao [62] identified several types of expert systems according to techniques used for their development such as, rule-based systems, knowledge-based systems, neural networks, case-based reasoning, intelligent agents and modelling-based. According to different or various sources, features of ES can be identified:

1. It applies human knowledge and expertise, collected in a knowledge base as different forms (unstructured information, rules, formulas, models, restrictions or constraints).
2. It may use different heuristics for autonomous reasoning to obtain optimal solutions or items to store in the knowledge base.
3. It aims to give answers at least with the same level of accuracy and validity as a human expert while attempting to eliminate or minimize human error.
4. It contains at least a knowledge base, an inference engine and a set of models (rules, etc.) that form basis of the ES.

From a technical standpoint it can be observed that ES apply some form of inference, mostly rule-based and outdated, to generalize for overall decision support.

Initial use of ES showed that it had promising automated reasoning abilities. However, actual realizations were found to be inflexible, lacking insight, decision-support and logic for explaining solutions, and verification of how the solution functions. Implementation of thousands or tens of thousands of rules also resulted in increased complexity of design, implementation, intercommunication and support. These problems were fully or partially eliminated after applying ML techniques, such as neural networks, which enabled learning from trends found in historical data, whereas new technologies, such as Web Services, integrated tools and environments to support distributed expert system development were later introduced.

DSS are defined in various sources, for instance, Raynor [63] defines DSS as 'data modelling and reporting system that has been structured to answer specific ongoing business questions or issues'. Real-time and interactive analysis, which support real-time decision making are mainly key factors that actually distinguish DSS from traditional IS. Such definitions also extend ES definitions, which deal mainly with problem-solving capabilities, excluding constraints on technology or computational implementation.

Both DSS and ES have a subset of similar components (knowledge base, inference engine, and user interface), similar roles (experts, knowledge engineer and user) and artefacts (documented knowledge). According to Beemer and Gregg [64], DSS usually includes a monitoring agent to help to identify unstructured decisions that need to be addressed, and to identify the domain variables that are given to the inference engine. The common processes may also differ in the level of complexity of interfacing with the user guiding the overall decision-making process and capturing, formalizing and organizing knowledge. In addition to architectural differences, it is stated that main differences between canonical ES and intelligent advisory DSS lie in the overall decision structure, which is more structured in ES, AI methodology (rule-based approach in ES versus case-based, ML and hybrid approaches in DSS), as well as their role in the decision process [64], [65]. ES frequently perform decisions and automatically take predefined actions, while DSS are meant to support human decisions with the final decision made through human interaction. Unlike ES, intelligent DSS are capable of solving structured or semi-structured problems and model various alternatives or strategies for solving them. Furthermore, they are quicker, less subjective in obtaining their results and use concentrated or derived information, as well as provide sophisticated tools for graphical analysis.

Main Types of Decision Support Systems

Different sources give their own taxonomies of such systems, with respect to their nature and main goals. The early taxonomy given by Alter [66] included systems based on consequence analysis of possible actions (using particular financial or mathematical models or simulations); optimization analysis, which suggests an action according to an optimal solution consistent with provided constraints; and recommendations and suggestions that infer a decision for specific structured tasks. Alter [66] attests that in 1980s DSS were already considered more an intelligent tool to guide decision support than a typical IS. The Holsapple and Whinston taxonomy [67] identified five types of DSS, according to the knowledge management techniques on which they are based: text-oriented DSS, database-oriented DSS, spreadsheet-oriented DSS, solver-oriented DSS and rule-oriented DSS. The analysis of intelligent structures for credit risk evaluation and management [11], together with analysis of recent commercial solutions, show that modern DSS for the banking domain may be classified either as one of such types, or may include characteristics that belong to two or more of these types (so-called integrated DSS). For example, one such system may use text-based analysis to derive information or attributes that can be utilized in optimization driven tasks or to optimize rule-based systems. Text-oriented risk assessment systems extract relevant information (items, patterns or characteristics) about particular entities from documents or corpus, using advanced text-mining techniques and perform further steps, such as train models, develop profiles or perform clustering/comparison operations. Therefore, such systems may involve advanced storage, search and processing components. Database-oriented DSS (also referred as data-driven DSS) apply more structured information in order to infer decisions. Data-driven systems implement retrieval and manipulation of internal and external historical data, such as data management, time-series, and real-time data use data warehouse facilities to store and retrieve data as well as advanced tools such as online-analytical processing, data cubes and data mining. Spreadsheet-based DSS apply spreadsheet technique for knowledge management, enabling advanced access and modification of knowledge in the knowledge base, together with processing instructions in spreadsheets. Microsoft Excel-based scorecards may be a good illustration of such systems. Solver-based DSS often apply the most advanced type of techniques (solvers) to infer solutions to different problems, according to their domain (finance, economics, investment or insolvency analysis) or problem type (optimization, forecasting, planning or statistical analysis). Such DSS often apply advanced and novel techniques, such as mathematical optimization, data-driven analytics, AI and/or ML. The structure of such problems, as well the structure of the DSS themselves, may be rather specific. The solvers can be chained to produce solutions that can become an input to another solver; hybrid techniques may be a good illustration of such decision-making pattern.

The last type of DSS, defined in Holsapple and Whinton's taxonomy, is a rule-oriented DSS, which process a set of user defined rules and produce a recommendation together with an explanation of performance during the reasoning process, providing relevant information for a more detail analysis. The rules are frequently expressed in form of *if…then* rules, although other types of rules may be applied as well. Several recent standards enable standardized rule definition and formalization, mainly Production Rule Representation [68] or Semantics of Business Vocabulary and Rules [69], issued and maintained by the Object Management Group. One of the intents of its recent Decision Model Notation is to provide decision formalization [70]. Holsapple and Whinton taxonomy was updated in 2008 [96] by adding hypertext-oriented DSS and multi-participant DSS. The latter type of DSS is typically used by multiple participants which do not have authorities to make the whole decision but who can influence this decision with their contribution. Several types of AI techniques, relevant to financial risk problems, can be used to simulate such decision-making. In the ensemble techniques, such as previously mentioned random forests, each classifier contributes to the whole decision with its own "vote". Swarm intelligence [71] is itself based on the collective behaviour of self-organized systems, where each individual contributes to the decision of the whole system by some compound, depending on the solution that it is capable to obtain itself. Such local solution may be later used as a guideline, which could lead to improved global solution, obtained by the whole system.

Arnott, Pervan [72] and Power [73] also describe their own DSS taxonomies. Arnott and Pervan [72] distinguish between data-oriented DSS and model-oriented (or model-driven) DSS; the latter typically uses limited data and models with parameters obtained automatically or provided by decision makers. Its performance is less impacted by the real-time data processing, compared to the data-driven DSS. According to the definition by Holsapple and Whinston, it can be viewed as corresponding to solver-oriented and rule-oriented DSS. The present wave of Big-Data-enabled decision-support solutions is a combination of both of these types, relying on the unlimited use of data with the intention to extract relevant patterns and turn it into useful models. Such classification is also based on several factors, particularly the number of agents that participate in decision making using the DSS (personal DSS, group DSS, negotiation DSS), number of intelligent techniques to infer decisions (intelligent DSS, knowledge-management-based DSS) and data-driven nature, obligatory to obtain insightful business decisions (executive information systems/business intelligence and data warehousing). The classification by Power [73] also includes five groups of DSS, is more based on the problems that they are targeted at, together with dominant architectural components, and is more of a combination of Holsapple and Whinston and Arnott [72] taxonomies.

Recent Developments in Decision Support Systems for Banking Problems

Although the importance of the evaluation problem is significantly highlighted in relevant literature, DSS structure and development topics are not widely covered. Two key factors may influence this situation:

- The availability of developed systems—such systems are frequently developed for financial institutions, which either disclose a scarce amount of information on such systems, or do not provide any. In addition, at the moment of writing, no open frameworks or architectures for banking DSS were available. Most solutions are commercial, developed by leading vendors and usually are not oriented towards flexibility or applications of novel techniques.
- The availability of data required for research—the data that is necessary for data-driven analysis, is usually not available. Several freely available datasets (e.g. German and Australian credit datasets in UCI repository[3]) are frequently used for credit risk research, however, it may be difficult to adopt the models, produced using such data sources, for real-world decision-support, as the actual data may be skewed, include a large amount of data points or have different characteristics. A lot of the researchers, which use real-world datasets, not to disclose the data for benchmarking or testing, which complicates actual comparison of results obtained by different techniques or by other researchers.

Overall, descriptions of DSS architectures for banking problems are rather scarce, including only high-level structure and the main components [74], [75], [76]. This chapter is limited to DSS described in scientific literature; the description of the functionality of commercial tools can be found in relevant whitepapers or documentation provided by vendors of the software. Early research applied model-based [77] or rule-based principles, resulting in knowledge-driven expert systems [78]. Credit risk DSS for credit card assessment is described in Matsasinis [79]. Zhang et al. [80] present their framework of DSS, structured as multi-layer system, consisting

[3] Machine learning repository, hosted by Center for Machine Learning and Intelligent Systems at the University of California, Irvine, which provides many of datasets that are frequently used in ML-related research.

of information integrated platform layer, utilization layer and information representation layer. Mahmoud et al. [81] describe a banking credit ES, developed using data from periodicals, references and books, banking reports and publications, research, working papers and banking studies that are routinely updated by domain experts. The knowledge base for their tool consists of five main components, each representing their purpose (economic feasibility study, financial feasibility study, marketing feasibility study, technical feasibility study and collaterals). Tsaih et al. [82] proposed N-tier architecture with internal credit scoring model transformation into an XML document. Middle tiers include the web server, a management application server, a loan-processing sub-system with case processing application server, an evaluation module with XML parser and model engine sub-modules, and a model installing sub-system, which consisted of model defining application server and model recording module. Kotsiantis et al. [83] developed a distributed credit evaluation system with an application of a decision tree (particularly C4.5) algorithm for scoring, which uses data sources associated with ontologies, together with data exchange and reasoning facilities powered by semantic web technologies. Although they refer to XBRL (Extensible Business Language) as one of the options, they chose their own developed ontology to represent financial statements.

Hence, none of the previously discussed system architectures defined modules for functionality of automated data collection, processing and updating models. Although this subject is explored further in stock trading systems (especially high-frequency trading), credit scoring systems may also benefit from such an approach by continuously updating models in the knowledge base using new data, obtaining new patterns or producing models for analysis in different dimensions. Additional attributes may be automatically integrated and explored in order to produce more accurate and insightful decisions. Finally, an additional incentive to develop such systems is also inspired by Open Data [84] and similar initiatives that promote sharing data online; this leads to an increased quantity of data that can be used to support and improve DSS performance, functionality, provide new insights or improve existing systems.

Requirements for Credit Risk DSS

Different internationally renowned regulations, such as Basel, Sarbanes-Oxley Act, International Financial Reporting Standards (IFRS), United

States Generally Accepted Accounting Principles (US GAAP) and so forth provide conformance, restrictions and guidelines for worldwide banking and financial systems. Basel III [85] is an internationally recognized framework for risk management and governance in the banking sector and provides guidelines to support financial stability, transparency and integrate better risk management practices. Three main pillars form the core of this standard [85]: Pillar 1 defines requirements for capital, risk management and financial leverage, Pillar 2 is focused on risk management and supervision, while Pillar 3 discloses requirements for market discipline. To meet such regulations, an integrated approach for intelligent decision support and decision making is essential, including the design and development of centralized information architecture, data warehouse, advanced analytics and reporting tools. Inconsistency and differences in internal financial reporting systems complicate their alignment and analytical tasks, such as, consolidation, integration and reporting. Alignment and conformance with coexisting regulations and frameworks must also be ensured, together with minimization of changes in processes, systems and data. Thus, financial standardization frameworks and initiatives, such as XBRL, are embraced (although not so easily adopted) by the banking sector, as they provide flexibility, improve the general reporting process and ensure standardized financial information flows.

One of the core components of decision support in the banking sector is rating systems; Balthazar [86] provides an extensive reference of their development requirements. Here we list the requirements that are considered the most important:

- Calculation of the main risk measures, such as probability of default (PD) and loss given default (LGD). PD is quantitative estimate of the likelihood that a debtor will not meet its obligations, while LGD measures the amount that would be lost according to this likelihood.
- At least seven rating grades for non-defaulted companies (and one for defaulted). Such a scale is a common practice for rating agencies, such as Fitch, Standard & Poor's and Moody's.
- Consistency across subsidiaries, locations, businesses.
- Transparency to auditors and external parties (common reference and metadata vocabulary, risk description and classification).
- Proven and documented accuracy, effectiveness, supervision, auditing and proper application of the scoring model used and the model itself, realized by proper manual and automated logging facilities.
- Integration of all available data, including external rating information. Sample warehouse structure is presented in Fig. 7.1.
- Integration of the debtor's solvency despite adverse economic conditions.

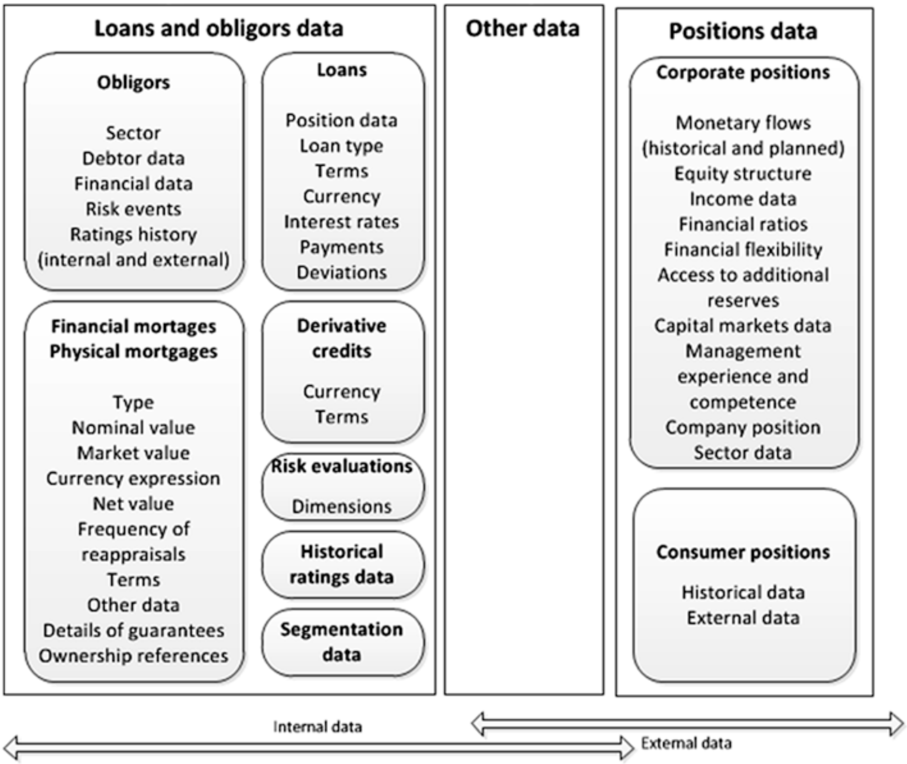


Fig. 7.1 Sample warehouse for credit risk DSS

- Regular model development validation and performance.
- Documented and justified scoring model overrides (for circumstances when human decision was preferred over the output of the scoring model).
- Persistent recording of the data used for rating and default history.

Basel III compatible DSS must satisfy the following requirements [86] [87]:

- Support an integrated approach for evaluation, management and the reporting of market, credit and liquidity risks, including risk-weighted assets, regulatory and economic capital.
- Integrate all available information (sample structure of a warehouse is given in Fig. 7.1) to support evaluation and management of risks, listed above.
- Be both data-driven and model-driven. It can combine features from other types of DSS, such as rule-based, solver-based or spreadsheet-based, depending on the implementation.
- Satisfy functional and non-functional requirements for system integrity, flexibility, scalability, supervisory access and usability.

- Calculate various risk components, such as PD, LGD, perform 'what-if' analysis and stress testing on different stress-testing scenarios, such as margin requirements, cash-flow generation or backtesting.
- Evaluate market value of credit risk, or counterparty valuation adjustments (CVA).
- Contain monitoring and multi-level logging functionality.

Financial Standards Based Decision Support

Data-driven decision support cannot be successfully implemented in model-oriented DSS without the necessary data for model development. Moreover, the model can be trained using sample data at the initial development stage, while updating it to reflect the current situation presents further challenges, requiring additional data in the same structure. This is especially influential in intelligent learning driven tasks where a model must be generated by continuously learning relationships between a dependent variable and a set of independent variables. To support the exchange of data, different formats may be used, such as XML, which is one of the core standards in data exchange between systems and is also the core standard in web services technology. XML enables definition of other domain-specific formats or languages by defining relevant internal representations and combining them to form descriptions representing concepts in different domains, together with their relations. Moreover, technologies, such as Extensible Stylesheet Language (XSL), XQuery and XPath enable advanced transformation and querying operations with XML documents. Development of such descriptions usually requires a high level of expertise in particular domains. Nevertheless, such approaches have already resulted in a significant number of relevant standards, including those that are relevant for the financial sector. A survey of such standards can be found in one of our previous papers [88], with XBRL and Statistical Data and Metadata Exchange (SDMX) among them. This section is focused on the XBRL standard, which, in addition to improved reporting process and accuracy, also offers advanced features that could be used to develop integrated DSS for financial risk analysis and is mature enough to be applied in real-world situations.

XBRL is defined by two primary concepts: *taxonomy* and *instance* [89]. *Taxonomy* defines all financial concepts that are used by a particular entity, together with their inner relationships and internal or external resources; *instance* is defined as the list of facts that has defined the structure in the taxonomy. XBRL taxonomy consists of XBRL schema, which stores information about taxonomy elements such as an unstructured list of elements and references to *linkbase* files, which link to specified external resources. Five types

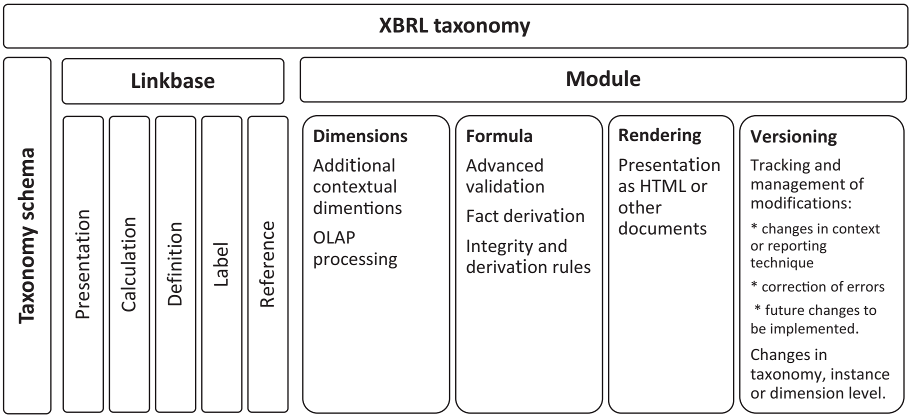


Fig. 7.2 XBRL taxonomy structure

of linkbases are defined in the XBRL specification: presentation, calculation, definition, label and reference [89].

Several complementary modules are also defined as extensions, such as XBRL Dimensions 1.0 [90], XBRL Formula 1.0 [91], XBRL Rendering specification [92] and XBRL Versioning [93] (Fig. 7.2). XBRL Dimensions complement XBRL documents with additional dimensions, enabling multi-dimensional processing. XBRL Formula is of particular interest to us as it provides advanced validation and fact derivation using integrity and derivation business rules, created according to accounting and financial assumptions. It applies such concepts, as formula, variable, completeness, accuracy assertion, filter and precondition. Using this standard, one can calculate, validate and store dimensional rollups, beginning and ending balances, concept equivalence and so forth. From the perspective of data analysis, it can be applied to derive new attributes or their transformations, which could be used in statistical modelling. XBRL versioning functionality could also be of interest, especially if modelling were to be based on a single taxonomy level.

For more technical information about XBRL and its possible integration to DSS for credit risk refer to Garsva and Danenas [88]. This paper introduces a mapping model for XBRL integration into an intelligent decision support model for credit risk evaluation. The dataset, which is used in such research, is typically defined as a two-dimensional m x n array of n instances of data having m attributes, or, as a more appropriate form for classification problem, as a set of n vectors $\{x_1, x_2, \ldots x_m, y\}$. This process is quite straightforward, if only instances from single taxonomies are used in the modelling process, as the set of financial attributes is finite. However, in case of multiple open XBRL taxonomies (such as the US SEC XBRL (US Securities XBRL) taxonomy) the application of the full XBRL knowledge base is not so trivial. It is obvious that only financial ratios defined in all used taxonomies can be used to form a dataset for a ML task. Note, that different adjustments should be taken into account in order to make comparisons between financial ratios of companies that use different systems for example, a US company and a European company. This is not explicitly addressed in this framework yet; currently only conversions according to the actual currency rates are performed to present financial attributes using a single currency. More formally, given a set of taxonomies $T = \{T_i | 0 \le i \le |T|\}$, where $|T|$ is the number of taxonomies in T, it is possible to have $\left\{F_{T_i}^k | N^{\min} \le k \le N^{\max}\right\}$ financial attributes, which are a common (overlap) in every taxonomy. Here the number of common financial ratios, represented by mappings in a taxonomy mappings data store, is defined as the minimal number of intercepting concepts F

$$N^{\min} = \min\left(k\right) | F_{T_i}^k \cap F_{T_j}^k, T_i \in T, 0 \le i \le |T|, 0 \le j \le |T| \tag{7.3}$$

and the maximum number of common financial concepts in the taxonomy

$$N^{\max} = \max\left(k\right) | F_{T_i}^k, T_i \in T, 0 \le i \le |T| \tag{7.4}$$

Suppose that the dataset represented by m instances from the taxonomy set $T = \bigcup_{i=1}^{|M|} T_i$ needs to be defined for classification task. We identify three cases for dataset formation:

- Dataset represented as $m \times N^{\min}$ matrix is formed, consisting of financial ratios common to each taxonomy. Such a strategy results in maximum loss of possible information available for model development, but in a least sparse dataset (i.e. it has the least number of attributes with missing values);
- Dataset represented as matrix $m \times k, N^{\min} < k < N^{\max}$; here, together with common financial ratios, additional ratios are partially integrated. The higher k is selected, the sparser dataset is obtained.
- Dataset represented as matrix $m \times N^{\max}$, which incorporates the maximum number of available data but results in a dataset with the largest sparseness.

To test this approach, SEC (Securities and Exchange Commission) taxonomies were selected as a base because they are well developed, applied and include the RATINGS taxonomy, which provides the means to define, structure and retrieve ratings data; another reason to select SEC (securities) taxonomies comes from the nature of the data from the SEC EDGAR database used in research. Ratings information may be used to perform a reverse-engineering task for the ratings (as discussed in Sect. 2.3) or to explore the inner relationships between financial, management or corporate data and these ratings by mapping financial and ratings data to produce a single dataset that can be used for analysis.

Developed Architecture for XBRL-Integrated DSS

A DSS for credit risk evaluation was designed using principles of XBRL integration described in Sect. 3.5 and ML techniques (Sect. 2). Only fundamental concepts are given here; for further reading please refer to the authors' original research describing this system [94], [95].

The system supports the full lifecycle of the data-driven model development; its framework is given in Fig. 7.3. Model development process itself is a composite task with sub-tasks such as data preprocessing, feature selection, training and parameter selection using manual selection or soft-computing technique-based heuristics. Analytical tasks are also supported, using statistical, financial and visual analysis. Credit risk analysis is viewed as an aggregation of these three. The system architecture is represented as a modular multi-layer system, aggregating functionality from such domains: ML, which implements relevant statistics and ML functionality; data and metadata management, including XBRL metadata and model repository together with metadata of these models, such as execution log and evaluation results; and credit risk evaluation. The credit risk evaluation layer (CRE layer) defines whole analytics, modelling, forecasting and evaluation functionality. Financial analysis, modelling and forecasting modules are defined particularly for the CRE layer as they present analytics, simulations and forecasting of a particular domain. Notice that the data management layer functionality intersects both AI/ML (model repository) and financial/credit risk domains (XBRL metadata, relevant business rules). The functionality of various supporting services is embedded in additional data source interaction, information processing and representation layers.

The proposed architecture extends the combined database and solver-oriented conceptual DSS model proposed by Holsapple [96]. Such framework is popular for development of a DSS capable of financial forecasting, analysis and optimization tasks. DSS that are based on this framework and combine data warehouse facilities with analytical (such as OLAP; Online Analytical Processing) or data-mining solvers are currently frequently used in large organizations. While the

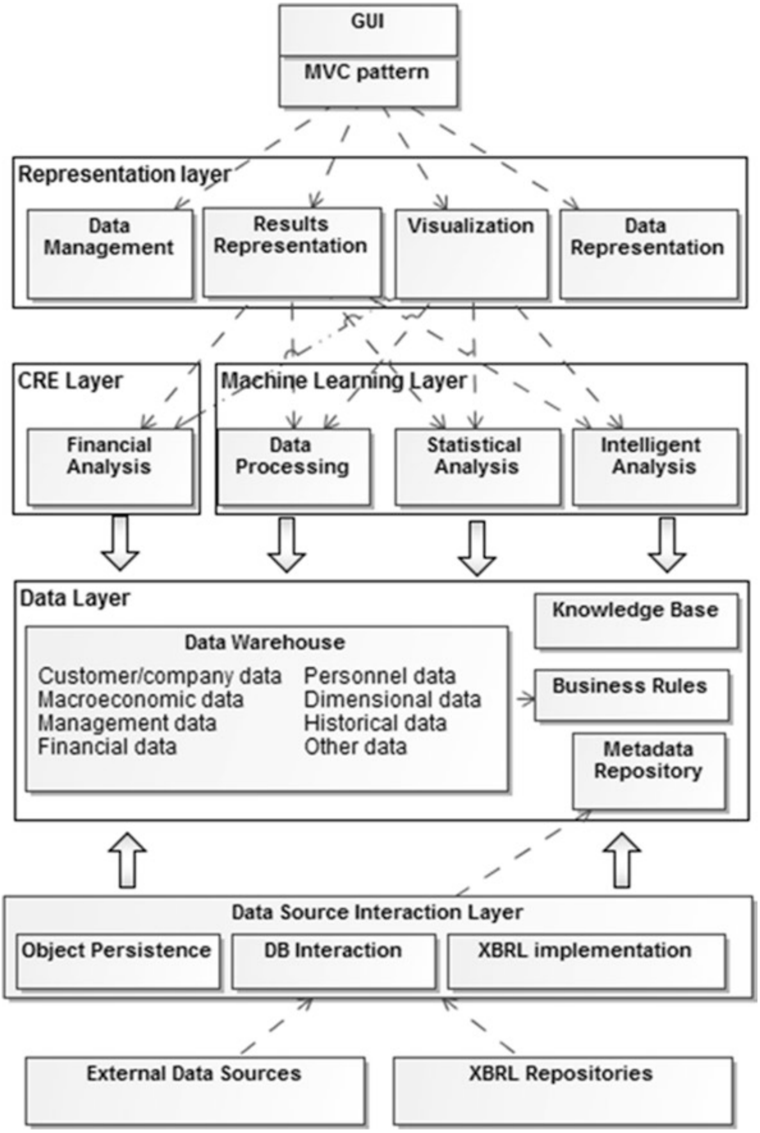


Fig. 7.3 DSS conceptual structure

original architecture included knowledge and problem processing sub-systems, to support automated data import and ML driven model update functionality together as management of metadata (taxonomy data, mappings) additional data and model processing systems are added (Fig. 7.4). A further extension is metadata management (taxonomy data, mappings), which is defined in the problem-processing system.

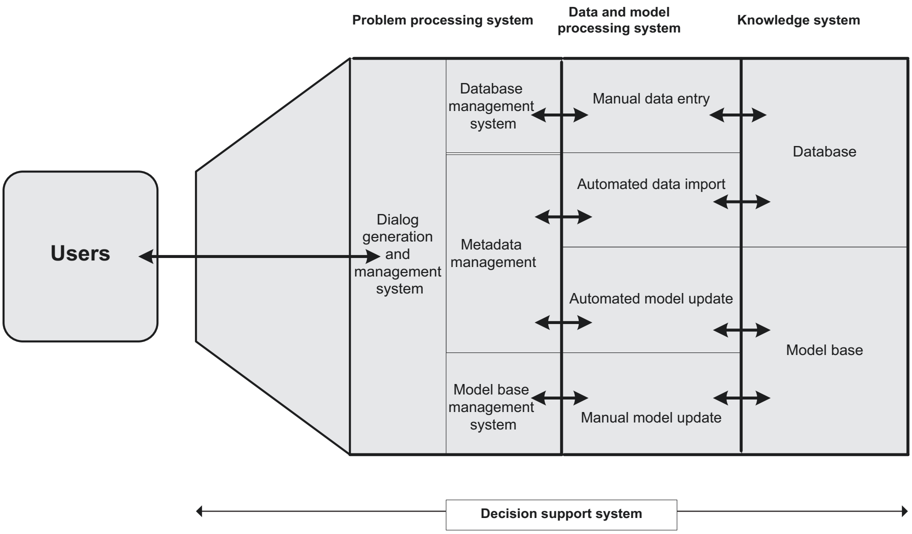


Fig. 7.4 Extension of Holsapple's combined database and solver-driven DSS architecture using data and model processing layers

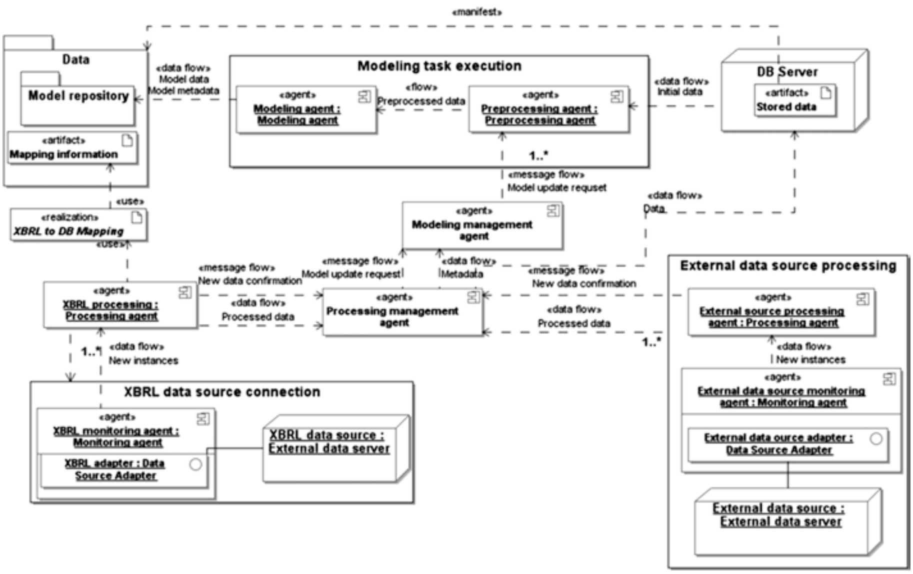


Fig. 7.5 DSS architecture as a set of agents

To illustrate exchange with automated self-learning functionality, the system is modelled as a set of agents with particular goals. Such model is given in Fig. 7.5, with messages between different agents (notifications) and data flows modelled using «*message flow*» and «*data flow*» stereotypes, respectively. Such agents are modelled:

- *Monitoring agent*: two specializations—XBRL monitoring agent and external source processing agent—are identified. These agents act similarly and have the primary goal of monitoring the data source for new XBRL document instances or changes in existing ones.

- *Processing agent*—two sub-classes are also identified, but, conversely from monitoring agents, they have slightly different goals. The XBRL processing agent processes XBRL instances using taxonomic information that is stored in a data-storage facility. If this instance comes with an unidentified taxonomy that is included with the instance it is also processed and included in the mappings information database. An external data source processing agent processes available data using rules that are implemented in its adapter. This adapter, included with each processing agent, implements the data access interface enabling specific functionality and rules necessary for data processing. Note that there can be multiple external XBRL or other external data sources with a single agent assigned to monitor each of them therefore, *1-to-many* is used to define such tasks.
- *Processing management agent*—acts as a manager and coordinator of processing agents. As there can be multiple processing agents this agent gets notifications of new data availability as well as the new data ready to be sent to the data warehouse. After the data is successfully sent to the data-storage facility this agent emits a notification to the *modelling management agent* together with metadata of the new instances that identifies the models to be updated.
- *Modelling management agent* also acts as a coordinator that triggers model update procedures by notifying corresponding *preprocessing* agents of new data available in the data store. At least one preprocessing and at least one corresponding modelling agent must be present to perform such tasks. Notice that cardinality between preprocessing and corresponding agents in this extension is modelled as *1-to-1*, although it can be extended to *1-to-many* if ensemble techniques not to be restricted to classification, with each agent acting as an individual in this ensemble, are also considered.
- *Preprocessing agent* performs data preprocessing tasks for retrieved data such as data cleansing, imputation and transformation. This data is further passed to the *modelling* agent (or multiple agents, in case of ensemble learning) to perform modelling tasks, such as model (re)training, testing and/or validation, and update the model in the repository. If a new model is developed then it is also added to the repository together with its metadata.

Such DSS extensions can be viewed as a self-organizing and learning system, able to adopt its decisions and recommendations on historical knowledge that it is configured to obtain. Moreover, it can be transformed to *Analytics-as-a-service*

(AaaS) by providing its decision through web services. However, several topics should be addressed during implementation, including the following:

- *Data compatibility*. It is possible that the client would not have all the data required for a particular model in DSS. Thus, it might be necessary to generate multiple models using different subsets of available data to satisfy different requirements.
- *Variable compatibility*. The client and the DSS may use different types of variables to describe the same object (e.g. variable 'job length' may be defined as nominal variable with value '8 years' and as a numerical variable with integer value '8').
- *Dynamic nature of developed models*, *defined by feature selection*. This leads to permanently changing models, which would also require dynamic generation of an interface to enter data for prediction for the client. Therefore, it requires that model metadata should be sent to the client. and that the client would retain the properly formatted data for evaluation each time they used the evaluation service. It might also require additional strategies for model development (e.g. some models might use only variables, which are most typical and accessible to all financial institutions, while others would incorporate more attributes yet would be usable only by a sub-set of institutions having the relevant data).

4 Conclusions

Credit risk evaluation has been one of the core problems for the banking domain since the time they started issuing loans to customers. State-of-the-art developments in modern AI with a focus on the imitation of human learning and behaviour offer many new challenges for their adoption to modern complex decision support in this area. Recent research has shown that ML techniques, such as ANN, SVM or random forests, frequently outperform typical statistical techniques and show promising results in their practical applications. Their combination with other techniques, such as evolutionary computation, ensemble learning, swarm intelligence or fuzzy logic, often outperforms standalone implementations of these techniques [4], therefore, research in this area is still active. The multitude of proposed techniques, the 'black-box' nature of most relevant techniques, the complexity of their implementation and the determination of the best approach are also factors that complicate their practical applications.

Highly configurable data-driven and model-driven systems for decision support are a prerequisite for domain experts to perform relevant analytical and modelling tasks. Proper architecture, design and implementation methodology must be applied to develop such systems, including the development, improvement and evaluation of techniques, which are used to create decision-making functionality. It is also reasonable to adapt them to existing standards, reuse and exploit their advantages to ensure data integrity or perform updates of the model. This is particularly advantageous in the financial domain as different financial systems contain a large number of constraints in the form of business rules. This chapter described an attempt to design a framework for such a system, by complementing machine-learning-based credit risk DSS functionality with additional components for financial metadata, which come from XBRL-based taxonomies. Such an approach may be applied to enforce particular constraints in data management, improve the quality of querying and the formation of datasets for continuous model re-training. Finally, an increase in the level of automation may lead to a self-learning intelligent decision support that may reduce the burden of the model repository update, as well as minimize the 'human error' factor in this process.

Several future research directions emerge. The developed framework incorporates a single financial information exchange standard (XBRL), whereas it may also benefit from other open financial exchange standards, such as FIX (Financial Information Exchange), SDMX (Statistical Data and Metadata eXchange), which would provide additional variables for the final datasets. Research of such variables in the context of credit risk problems and the need for their integration is another topic for discussion, as the significance of macroeconomic, statistical or market variables for credit risk is also described in numerous sources. However, no consensus on their usage has been achieved although most studies refer to ratios, based on financial statements, some of them incorporate other kinds of variables into their models. The automated management of differences in accounting principles between different regulations (e.g. IFRS and US GAAP), as well as the adjustment of data from different financial statements for comparison is another challenging area that is not addressed by this framework. Yet, some techniques (e.g. in the form of transformation or derivation rules) could be employed to fully or partially automate this task. Finally, such a system would also benefit from novel improvements in ML algorithms and their adaptability to changes in its environment, leading to an increased level of decision process automation and system autonomy.

References

1. Shearer, C. (2000). The CRISP-DM model: The new blueprint for data mining. *Journal of Data Warehousing, 5*(4), 13–22.
2. Dunham, M. H. (2003). *Data Mining: Introductory and advanced topics*. Upper Saddle River: Prentice-Hall.
3. Danenas, P., & Garsva, G. (2009). Support vector machines and their application in credit risk evaluation process. *Transformations in Business & Economics, 8, No. 3*(18), 46–58.
4. Jones, S., Johnstone, D., & Wilson, R. (2015). An empirical evaluation of the performance of binary classifiers in the prediction of credit ratings changes. *Journal of Banking & Finance, 56*, 72–85.
5. Lin, W.-Y., Hu, Y.-H., & Tsai, C.-F. (2011). Machine Learning in Financial Crisis Prediction: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42*(4), 421–436.
6. Wang, G., Ma, J., & Yang, S. (2014). An improved boosting based on feature selection for corporate bankruptcy prediction. *Expert Systems with Applications, 41*(5), 2353–2361.
7. Diederich, J. (2008). *Rule extraction from support vector machines*. Berlin Heidelberg: Springer.
8. Engelbrecht, A. P. (2007). *Computational intelligence: An introduction* (2 ed.). Chichester, UK: John Wiley & Sons Ltd.
9. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). New York: Springer.
10. Danenas, P., & Garsva, G. (2014). Intelligent techniques and systems in credit risk analysis and forecasting : A review of patents. *Recent Patents on Computer Science, 7*, 12–23.
11. Marques, A. I., Garcia, V., & Sanchez, J. S. (2013). A literature review on the application of evolutionary computing to credit scoring. *Journal of the Operational Research Society, 64*(9), 1384–1399.
12. Paliwal, M., & Kumar, U. A. (2009). Neural networks and statistical techniques: A review of applications. *Expert Systems with Applications, 36*(1), 2–17.
13. Verikas, A., Kalsyte, Z., Bacauskiene, M., & Gelzinis, A. (2010). Hybrid and ensemble-based soft computing techniques in bankruptcy prediction: a survey. *Soft Computing, 14*(9), 995–1010.
14. Engelbrecht A.P., Computational Intelligence: An Introduction, 2nd ed., John Wiley & Sons, Ltd., 2007.
15. Vellido, A., Lisboa, P. J. G., & Vaughan, J. (1999). Neural networks in business: A survey of applications (1992–1998). *Expert Systems with Applications, 17*(1), 51–70.

16. Yu, Q., Miche, Y., Séverin, E., & Lendasse, A. (2014). Bankruptcy prediction using Extreme Learning Machine and financial expertise. *Neurocomputing, 128*, 296–302.
17. Jeong, C., Min, J. H., & Kim, M. S. (2012). A tuning method for the architecture of neural network models incorporating GAM and GA as applied to bankruptcy prediction. *Expert Systems with Applications, 39*(3), 3650–3658.
18. Khashman, A. (2011). Credit risk evaluation using neural networks: Emotional versus conventional models. *Applied Soft Computing, 11*(8), 5477–5484.
19. Tseng, F.-M., & Hu, Y.-C. (2010). Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks. *Expert Systems with Applications, 37*(3), 1846–1853.
20. Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine learning, 20*, 273–297.
21. Vapnik, V. N. (1998). *Statistical learning theory* (p. 1998). New York: John Wiley & Sons Inc..
22. Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology, 2*(3), 27.
23. Huang, T.-M., Kecman, V., & Kopriva, I. (2006). *Kernel based algorithms for mining huge data sets: supervised, semi-supervised and unsupervised learning*. Berlin, Heidelberg: Springer.
24. Schölkopf, B. (2001). The kernel trick for distances. *Advances in Neural Information Processing Systems, 13*, 301–307.
25. Suykens, J., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Processing Letters, 9*(3), 293–300.
26. Joachims, T. (1999). Making large-scale SVM learning practical. In B. Schölkopf, C. Burges, & A. Smola (Eds.), *Advances in Kernel Methods—Support vector learning*. Cambridge, MA: MIT-Press.
27. Fan, R., Chang, K., Hsieh, C., Wang, X., & Lin, C. (2008). LIBLINEAR: A library for large linear classification. *The Journal of Machine Learning Research, 9*, 1871–1874.
28. Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning, 1*, 81–106.
29. Wang, G., Ma, J., Huang, L., & Xu, K. (2012). Two credit scoring models based on dual strategy ensemble trees. *Knowledge-Based Systems, 26*, 61–68.
30. Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5–32.
31. Kruppa, J., Schwarz, A., Arminger, G., & Ziegler, A. (2013). Consumer credit risk: Individual probability estimates using machine learning. *Expert Systems with Applications, 40*(13), 5125–5131.
32. Malekipirbazari, M., & Aksakalli, V. (2015). Risk assessment in social lending via random forests. *Expert Systems with Applications, 42*(10), 4621–4631.
33. Wang Q., Lai K. K., and Niu D. (2011). Green credit scoring system and its risk assessment model with support vector machine. *2011 Fourth International Joint Conference on Computational Sciences and Optimization*, pp. 284–287.

34. Bellovary, J., Giacomino, D., & Akers, M. (2007). A review of Bankruptcy prediction studies : 1930 to present. *Journal of Financial Education, 33*, 1–43.
35. Bae, J. K. (2012). Predicting financial distress of the South Korean manufacturing industries. *Expert Systems with Applications, 39*(10), 9159–9165.
36. Erdogan, B. E. (2013). Prediction of bankruptcy using support vector machines: an application to bank bankruptcy. *Journal of Statistical Computation and Simulation, 83*(8), 1543–1555.
37. Hammer, P. L., Kogan, A., & Lejeune, M. A. (2012). A logical analysis of banks' financial strength ratings. *Expert Systems with Applications, 39*(9), 7808–7821.
38. Horta, I. M., & Camanho, A. S. (2013). Company failure prediction in the construction industry. *Expert Systems with Applications, 40*(16), 6253–6257.
39. van Gestel, T., Martens, D., Baesens, B., Feremans, D., Huysmans, J., & Vanthienen, J. (2007). Forecasting and analyzing insurance companies' ratings. *International Journal of Forecasting, 23*(3), 513–529.
40. Chandra, D. K., Ravi, V., & Bose, I. (2009). Failure prediction of dotcom companies using hybrid intelligent techniques. *Expert Systems with Applications, 36*(3), 4830–4837.
41. Kim, S. Y., & Upneja, A. (2014). Predicting restaurant financial distress using decision tree and AdaBoosted decision tree models. *Economic Modelling, 36*, 354–362.
42. Chen, M.-Y. (2011). Using a hybrid evolution approach to forecast financial failures for Taiwan-listed companies. *Quantitative Finance, 14*(6), 1–12.
43. van Gestel, T., Baesens, B., Suykens, J. A. K., van den Poel, D., Baestaens, D. E., & Willekens, M. (2006). Bayesian kernel based classification for financial distress detection. *European Journal of Operational Research, 172*(3), 979–1003.
44. Geng, R., Bose, I., & Chen, X. (2015). Prediction of financial distress: An empirical study of listed Chinese companies using data mining. *European Journal of Operational Research, 241*(1), 236–247.
45. Oreski, S., & Oreski, G. (2014). Genetic algorithm-based heuristic for feature selection in credit risk assessment. *Expert Systems with Applications, 41*(4), 2052–2064.
46. Boguslauskas, V., & Mileris, R. (2009). Estimation of credit risk by artificial neural networks models. *Engineering Economics, 4*, 7–14.
47. Purvinis, O., Virbickaite, R., & Šukys, P. (2008). Interpretable nonlinear model for enterprise bankruptcy prediction. *Nonlinear Analysis: Modelling and Control, 13*(1), 61–70.
48. Kennedy, K., Mac, N. B., Delany, S. J., O'Sullivan, M., & Watson, N. (2013). A window of opportunity: Assessing behavioural scoring. *Expert Systems with Applications, 40*(4), 1372–1380.
49. Chen, M.-Y. (2012). Visualization and dynamic evaluation model of corporate financial structure with self-organizing map and support vector regression. *Applied Soft Computing, 12*(8), 2274–2288.
50. Kohonen, T. (1995). *Self-organizing maps*. Berlin, Heidelberg: Springer.

51. Huysmans, J., Baesens, B., Vanthienen, J., & van Gestel, T. (2006). Failure prediction with self organizing maps. *Expert Systems with Applications, 30*(3), 479–487.
52. Kiviluoto, K. (1998). Predicting bankruptcies with the self-organizing map. *Neurocomputing, 21*, 191–201.
53. Merkevicius, E., Garsva, G., & Girdzijauskas, S. (2006). A hybrid SOM-Altman model for bankruptcy prediction. *Lecture Notes in Computer Science, 3994*, 364–371.
54. Chen, N., Ribeiro, B., Vieira, A., & Chen, A. (2013). Clustering and visualization of bankruptcy trajectory using self-organizing map. *Expert Systems with Applications, 40*(1), 385–393.
55. Sun, J., Li, H., Huang, Q.-H., & He, K.-Y. (2014). Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches. *Knowledge-Based Systems, 57*, 41–56.
56. Tsai, C.-F. (2009). Feature selection in bankruptcy prediction. *Knowledge-Based Systems, 22*(2), 120–127.
57. Lin, F., Liang, D., Yeh, C.-C., & Huang, J.-C. (2014). Novel feature selection methods to financial distress prediction. *Expert Systems with Applications, 41*(5), 2472–2483.
58. Chen, F.-L., & Li, F.-C. (2010). Combination of feature selection approaches with SVM in credit scoring. *Expert Systems with Applications, 37*(7), 4902–4909.
59. Liang, D., Tsai, C.-F., & Wu, H.-T. (2015). The effect of feature selection on financial distress prediction. *Knowledge-Based Systems, 73*, 289–297.
60. Freund, Y. (1995). Boosting a weak learning algorithm by majority. *Information and Computation, 121*(2), 256–285.
61. Ye, S., Xiao, Z., & Zhu, G. (2014). Identification of supply chain disruptions with economic performance of firms using multi-category support vector machines. *International Journal of Production Research, 53*(10), 3086–3103.
62. Turban, E., Aronson, J. E., Liang, T.-P., & Sharda, R. (2006). *Decision support and business intelligence systems* (8th ed.). Upper Saddle River: Prentice Hall.
63. Liao, S. H. (2005). Expert system methodologies and applications-a decade review from 1995 to 2004. *Expert Systems with Applications, 28*(1), 93–103.
64. Raynor, W. J. (1999). *The international dictionary of artificial intelligence*. London: The Glenlake Publishing Company, Ltd.
65. Beemer, B. A., & Gregg, D. G. (2008). Advisory systems to support decision making. In F. Burstein & C. W. Holsapple (Eds.), *Handbook on Decision Support Systems 1: Basic Themes* (pp. 511–527). Berlin Heidelberg: Springer.
66. Turban, E., & Watkins, P. R. (1986). Integrating expert systems and decision support systems. *MIS Quarterly, 10*(2), 121–136.
67. Alter, S. L. (1980). *Decision support systems: Current practice and continuing challenge*. Reading MA: Addison-Wesley.
68. Holsapple, C., & Whinston, A. (1996). *Decision support systems: A knowledge-based approach*. Eagan, MN: West Publishing Company.

69. OMG. (2009). Production Rule Representation (PRR) v.1.0 (December 2009).
70. OMG. (2013). Semantics of Business Vocabulary and Business Rules (SBVR) v.1.2 (November 2013).
71. Holsapple, C. W. (2008). DSS architecture and types. In F. Burstein & C. W. Holsapple (Eds.), *Handbook on decision support systems 1: Basic themes* (pp. 163–189). Berlin, Heidelberg: Springer.
72. OMG. (2015). Decision Model and Notation (DMN) v.1.0 (March 2015).
73. Kennedy, J., Eberhart, R. C., & Shi, Y. (2001). *Swarm intelligence*. San Francisco, CA: Morgan Kaufmann Publishers.
74. Arnott, D., & Pervan, G. (2005). A critical analysis of decision support systems research. *Journal of Information Technology, 20*(2), 67–87.
75. Power, D. J. (2008). Decision support systems: A historical overview. In F. Burstein & C. W. Holsapple (Eds.), *Handbook on decision support systems 1: Basic themes*. Berlin Heidelberg: Springer.
76. Cheng, H., Lu, Y.-C., & Sheu, C. (2009). An ontology-based business intelligence application in a financial knowledge management system. *Expert Systems with Applications, 36*(2), 3614–3622.
77. Guo-an Y., Hong-bing X., and Chao W. (2003) Design and implementation of an agent-oriented expert system of loan risk evaluation. *Proc. Of International Conference on Integration of Knowledge Intensive Multi-Agent Systems* (pp. 41–45).
78. Huai, W. (2010) The framework design and research on enterprises group financial decision support system. In: *Proceedings of Management and Service Science (MASS)*, Wuhan, China, pp.1–4.
79. Kuo, H.-C. (1997). Cognitive management system based support for bank credit granting decision: An integrated and practical design for Taiwan. *Journal of Systems Integration, 7*, 77–91.
80. Nedović, L., & Devedžić, V. (2002). Expert systems in finance—A cross-section of the field. *Expert Systems With Applications, 23*(1), 49–66.
81. Matsasinis, F. N. (2002). CCAS: An intelligent decision support system for credit card assessment. *Journal Of Multi-Criteria Decision Analysis, 11*, 213–235.
82. Zhang, M., Gu, Y., and Zhu, J. (2009). Analysis of the framework for financial decision support system. *Proceedings of 2009 International Conference on Wireless Networks and Information Systems, Shanghai*, pp. 241–244.
83. Mahmoud M., Algadi N., and Ali A. (2008) Expert System for Banking Credit Decision. *2008 International Conference on Computer Science and Information Technology*, 813–819.
84. Tsaih, R., Liu, Y.-J., Liu, W., & Lien, Y.-L. (2004). Credit scoring system for small business loans. *Decision Support Systems, 38*(1), 91–99.
85. Kotsiantis, S. B., Kanellopoulos, D., Karioti, V., and Tampakas, V. (2009). An ontology-based portal for credit risk analysis. *2nd IEEE International Conference on Computer Science and Information Technology* (pp. 165–169).
86. Open Knowledge. (2015). *Open knowledge: What is open*? https://okfn.org/opendata/, date accessed 2015.08.20.

87. Bank for International Settlements (BIS) (2015). *International regulatory framework for banks* (Basel III), http://www.bis.org/bcbs/basel3.htm, date accessed 2015.08.20.
88. Balthazar, L. (2006). *From Basel 1 to Basel 3: The integration of state-of-the-art risk modeling in banking regulation*. UK: Palgrave Macmillan.
89. Patni A. (2015) Basel III is here—What are the implications for your business? *Infosys whitepaper*, http://www.infosys.com/industries/financial-services/white-papers/Documents/basel-here.pdf, date accessed 2015.08.20
90. Garsva, G., & Danenas, P. (2011). XBRL integration into intelligent system for credit risk evaluation. *Transformations in Business & Economics, 10, No. 2*(23), 88–103.
91. XBRL Inc. (2010). XBRL International Inc. Extensible Business Reporting Language (XBRL) 2.1. http://xbrl.org/Specification/XBRL-RECOMMENDATION-2003-12-31+Corrected-Errata-2008-07-02.htm, date accessed 2015.07.04
92. XBRL Inc. (2006). XBRL International Inc. XBRL Dimensions 1.0. http://www.xbrl.org/Specification/XDTREC-2006-09-18.htm, date accessed 2015.07.04.
93. XBRL Inc. (2009). XBRL International Inc. Formula Specifications. http://www.xbrl.org/Specification/formula/REC-2009-06-22/index.htm, date accessed 2015.07.04
94. XBRL Inc. (2007) XBRL International Inc. XBRL Rendering Requirements, Public Working Draft. https://www.xbrl.org/technical/requirements/REN-REQ-PWD-2007-07-24.htm, date accessed 2015.07.04.
95. XBRL Inc. (2007). XBRL International Inc. XBRL Versioning Specification 1.0, Public Working Draft. http://www.xbrl.org/Specification/Versioning/XVS-PWD-2007-11-28.htm, referred on 24/11/2010, date accessed 2015.07.04
96. Danenas P. and Garsva G. (2012) Domain driven development and feature driven development for development of decision support systems. Proceedings of Information and Software Technologies—18th International Conference ICIST 2012. *Communications in Computer and Information Science* (Vol. 319, pp. 187–198).
97. Danenas P. and Garsva G. (2011) SVM and XBRL based decision support system for credit risk evaluation. *Proceedings of the 17th International Conference on Information and Software Technologies* (*IT 2011*), Kaunas, Lithuania, 190–198.
98. Cao, J., Lu, H., Wang, W., & Wang, J. (2012). A novel five-category loan-risk evaluation model using multiclass LS-SVM By PSO. *International Journal of Information Technology & Decision Making, 11*(4), 857–874.

8

Artificial Intelligence for Islamic Sukuk Rating Predictions

Tika Arundina, Mira Kartiwi, and Mohd. Azmi Omar

1 Introduction

Credit rating has become an important metric in the modern financial services industry. It assesses the credit worthiness of a security and its issuer, most often based on the history of borrowing and repayment for the issuer, its underlying assets, its outstanding liabilities and its overall business performance. Credit rating fulfils a key function of information transmission in the capital market, especially to improve risk management and market liquidity. With the presence of credit rating, the issue of asymmetric information among market

T. Arundina (✉)
Faculty of Economics and Business, University of Indonesia,
Depok 16424, Indonesia
e-mail: tikarundina@gmail.com

M. Kartiwi
Kulliyyah of Information, Communication and Technology, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Selangor, Malaysia
e-mail: mira@iium.edu.my

M.A. Omar
Islamic Research and Training Institute, Islamic Development Bank Group,
Jeddah, Saudi Arabia
e-mail: azmiomar@isdb.org

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets*,
New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0_8

players can be reduced. For companies, having a good rating is beneficial as it may improve market trust in their business. Furthermore, many investors and banks rely on the ratings to make investment and financing decisions.

All of the benefits mentioned above promote development of capital markets, including the Islamic capital market. Islamic finance and capital market is one of the fastest growing segments of international financial markets. Recent innovations in Islamic finance and capital markets have changed the landscape of the financial industry. One of them is Islamic securities which are known as Sukuk. The use of Sukuk as the alternative to the existing conventional bond has become increasingly popular in the past few years. Based on the International Islamic Financial Market (2013), the total global Sukuk attained a very respectable issuance totaling USD 472.68 billion from 2001 to January 2013. This remarkable growth is shown from the significant increase in terms of Sukuk issuance by more than 100 times since the year 2001, from USD 1172 million to USD 137 billion in the year 2012. These Sukuk are used as a means of raising government finance through sovereign Sukuk issues, and a way through which companies raise funds by issuing corporate Sukuk.

The development of the Sukuk market has raised the issue of Sukuk ratings. The rating not only reflects risk and expected performance but is also beneficial and assists the investor, specifically banks that invest in the particular security when measuring capital charge for the investment. Similar to conventional banks, there is capital adequacy ratio requirement that has to be fulfilled by Islamic banks. The ratio is measured by dividing the Islamic bank's eligible capital with its risk weighted asset. The Risk Weighted Asset (RWA) is calculated as follows:

$$RWA = \sum_{i=1}^{n} w_{\mathrm{i}} a_{\mathrm{i}} \tag{8.1}$$

In this case w_{i} is the risk weight of asset category i, a is the quantity of asset i and n is the number of asset categories. For Sukuk, the risk weighted asset is measured by calculating Islamic bank exposure on the Sukuk with the risk weight of the Sukuk itself and the risk weight of the Sukuk varying based on Sukuk rating. Nonetheless, not all of the Sukuk have been rated. This situation leads the bank has to assign 100 % risk weight to the non-rated Sukuk. However, this practice is not reflecting the true risk attached to respective Sukuk.

Basel Committee for Banking Supervision has now instituted capital charges for credit risk based on credit ratings. Basel III framework allows the bank to establish capital adequacy requirements based on ratings provided by external credit rating agencies or to determine rates of its investment internally. With regards to this requirement, the model is expected to lead Islamic banks to develop an internal rating based approach.

The need to develop a model to predict Sukuk rating is becoming more important with regards to high cost and the fact that not all securities are being rated by agencies. Huang et al. [1] argue that credit ratings are very costly to obtain due to the large expense associated with performing credit assessments. Rating agencies require a large amount of time and human resources to perform an in-depth analysis of the company's risk status based on various aspects. As a result, some companies cannot afford to update their credit ratings regularly through rating agencies, which makes credit rating prediction particularly valuable to the investment market.

There are several studies on rating prediction such as, Hajek and Olej [2], Mizen and Tsoukas [3], Novotna [4], Dainelli et al. [5], Hajek and Michalak [6], Muscettola and Naccarato [7], and Doumpos et al. [8]. However, Sukuk is a very recent introduction and new to the financial world. The studies and analysis conducted on Sukuk are very limited, especially on Sukuk ratings. Arundina and Mohd Azmi [9] is one of study that examined the determinant of corporate Sukuk rating in Malaysia.

This study will extend the previous research on Sukuk rating using several theoretical variables adapted from previous studies on bond rating prediction with the inclusion of some additional variables such as Sukuk structure, guarantee status, the industrial sector and macroeconomic variables. This study attempts to observe whether bond rating determinant variables can predict Sukuk ratings, incorporating Sukuk structure variable as one of Sukuk specific variables.

2 Literature Review

What Is Sukuk

Sukuk and conventional bond securities have some similarities such as fixed-term maturity, coupons and they are both traded on the secondary market. Tariq [10] mentioned that Sukuk has a similar function to bonds, which is to enable companies to raise capital according to Islamic principle and at the same time expand the investor base and offer investment opportunities for new groups. Theoretically, to some extent, there should be some differences in rating methodologies for bonds and Sukuk because these two instruments are different in nature. Bonds are contractual debt obligations whereby the issuer is contractually obliged to pay bondholders, on certain specified dates, interest and principal. On the other hand, according to Accounting and Auditing Organization for Islamic Financial Institutions (AAOIFI) Standard no.17, Sukuk are certificates of equal value that represent an undivided interest in the ownership of an underlying asset, usufruct and services or assets of particular projects or special investment activity. The Sukuk holder, also has a claim to the underlying assets as is also the case with normal conventional bonds should the issuer default on payments. Consequently, Sukuk holders are entitled to share in the revenues generated by the Sukuk assets, as well as to share in the proceeds of the realization of the assets. Sukuk certificates are unique in the way that the investor becomes an asset holder, hence should bear the risk of its underlying assets. Sukuk certificate holders carry the burden of these unique risks. Table 8.1 provides a simple definition and comparison of conventional bonds and Sukuk bonds.

Another major difference between Sukuk and conventional bond is in terms of the investment risks. Sukuk is associated with a risk termed a Shariah compliance risk, which is essential during the structuring stage based on the available Islamic finance contracts. Nonetheless, Sukuk has some similarities to conventional bonds because they are structured with physical assets that generate revenue. The underlying revenue from these assets represents the source of income for payment of profits on the Sukuk. According to AAOIFI [11], Sukuk are issued on various transaction contracts. These Sukuk are Ijara, Murabaha, Salam, Istisna, Mudaraba and Musharaka, Muzara'a (share-cropping), Muqasa (irrigation) and Mugharasa (agricultural partnership).

Table 8.1 Conventional bonds versus Sukuk

	Bonds	Sukuk
Definition	Debt security in which borrowed money is repaid in an agreed amount at a specified time.	A fixed income certificate that represents an ownership of underlying assets, services or investments.
Expected cash flows	Interest payments	Profit sharing
Concept	□ Proof of debt □ Pre-determined periodic repayment of cash flows with pre-set amounts	□ Proof of ownership □ Income derived from asset/service/investments based on original contractual commitment
Structure and principles	□ Zero coupon bond □ Coupon bearing bonds □ Convertible bonds, etc.	□ Profit and loss sharing □ Leasing

Source: www.bpam.com.my [60]

However, the last three types are rarely used in the market. Appendix 1 summarizes the description of each structure, the relation with the Sukuk holders and the expected return. Those structures will affect the coupon payment method as well as the risk characteristics. The different nature of bonds and Sukuk in terms of their respective credit risk exposure causes the need for different ratings assessment.

In addition, Sukuk share similar risks with that of conventional bonds, namely credit risk (default risk), liquidity risk, business risk, market risk, interest rate risk (rate of return risk), foreign exchange risk, equity price risk and commodity risk [10]. As all these risks are commonly associated with corporate Sukuk, these risks are reflected in the specific ratings provided by the rating agencies. Similar to corporate bond ratings, the higher the rating the better quality the corporate Sukuk will be, hence the lower the risk of holding it [12].

Sukuk Rating Methodology Based on Recourse of the Underlying Asset

Related to Sukuk ratings, rating agencies, however, only give an opinion on the credit aspect linked to the instruments. Since the rating agencies argued that Sharia-compliant nature of Sukuk is neutral from a credit perspective, the rating assigned to Sukuk does not imply any confirmation on Sharia compliance [13, 14, 15, 16] and the Malaysian Agency Rating Corporation (MARC) [17]. Most of the Sukuk rated by rating agencies are structured with the approval of the Sharia board. The board evaluates the structure of the transaction and determines its compliance with Shariah prior to the launch of the Sukuk. Consistent with its position on addressing only the credit aspects of the transaction, rating agencies neither review the role or the composition of the Sharia board, nor comment on the validity of the board's recommendations and decisions.

Besides underlying contracts, Sukuk can also be differentiated into two types of mode due to recourse of the underlying asset: the asset-backed Sukuk and the asset-based Sukuk. These two types of Sukuk semantically have similar descriptions but mask significant differences in terms of credit risk.

Asset-backed Sukuk represent a true sale of assets because the underlying asset has been legitimately transferred to the Special Purpose Vehicle (SPV). In the event of default, therefore, the underlying assets will remain completely separate from the originator. The Sukuk holders or the investors have recourse to the underlying asset of the Sukuk, that is, the investors have the full claim over the underlying asset, without any risk of the sale subsequently being inverted by the local or Sharia courts. Fitch [13], S&P [14], Moody's [15], RAM [16] and MARC [17] argue that the rating assessment of asset-backed Sukuk is dependent on the performance of the underlying asset in how it generates cash to meet its contracted obligations.

Asset-based Sukuk, however, are structured such that investors only have a beneficial ownership in the underlying asset instead of legal ownership of the underlying asset. In this structure, assets are generally sold by the originator to an SPV in the form of a trust. The trustee issues certificates showing the investor's ownership interest, while the proceeds are used to purchase the assets. The investor receives a distribution income representing a share of the return generated by the underlying assets or from any source from the originator. In an asset-based Sukuk, it is clearly distinguished that the credit risk of Sukuk reflects the credit risk of the originator rather than the underlying assets. This is because the investors do not have any recourse to the underlying assets in the event of default since the investor is not the legal owner of the underlying asset. In the event of bankruptcy, in asset-based Sukuk, the investor(s) have an unsecured claim, because Sukuk claim will be ranked at the same level as other unsecured creditors. Sukuk investors, in the case of senior unsecured obligations, have no priority over the Sukuk assets when compared to other creditors and will be rated equivalent to them. Otherwise, in the case of subordinated Sukuk, ratings are notched down. In this case, rating agencies apply fundamental rating methodologies or corporate credit rating (CCR) to rate the Sukuk [13, 14, 15, 16, 17].

Since most Sukuk issuance is based on a corporate asset-based mode, this implies that the rating agencies adopt the common conventional ratings method. Due to this circumstance, this study uses bond rating prediction studies to determine the predictors of asset-based Sukuk rating.

Previous Studies on Rating Prediction

Numerous research studies have been conducted to study the prediction of ratings. Some of the earliest studies on the prediction of bond ratings were carried out in 1966 and 1970 by Horrigan and West respectively. Horrigan [18] and West [19] employed ordinary least squares (OLS) regression to predict Moody's and S&P's bond ratings. Pinches and Mingo [20] started to apply Multiple Discriminant Analysis (MDA) to improve the statistical fit in order to develop an accurate model. Following this research, Belkaoui [21] applied the same method using information from financial statements and macroeconomic variables to predict bond rating. The model correctly predicted 62.5 % of the ratings in the experimental sample and 65 % of bond ratings in a validation sample using MDA.

Since the late 1990s, researchers employed artificial intelligence (AI) models for predicting bond ratings. Artificial Intelligence Neural Network (AINN) was introduced the first time by Dutta and Shekhar [22]. They used non-statistical methods in the form of neural networks (NN) to predict ratings of corporate bonds. Their work was inspired by a lack of theories and methodologies for rating bonds. NN do not require a prior specification of a functional domain model. They attempt to learn underlying domain models from the training input-output examples [22, 23]. Dutta and Shekhar [22] compared NN with regression-based techniques using six financial variables based on the results produced by Horrigan [18] and Pinches and Mingo [20]. This study finds that NN outperform regression in bond rating prediction.

Singleton and Surkan [24] applied Back Propagation Neural Network (BPNN) to classify bonds of the 18 Bell Telephone companies divested by American Telephone and Telegraph Company (AT&T) in 1982 and compared an NN model with an MDA model. This study demonstrated that NN achieved better performance in predicting the direction of a bond rating than discriminant analysis.

Kwon et al. [25] compared a conventional NN method to a new NN training approach: Ordinal Pairwise Partitioning (OPP). During the same period, Chaveesuk et al. [23] also compared the prediction rate and accuracy of three AINN methods: Radial Basis Function (RBF), Learning Vector Quantization (LVQ) and Back Propagation (BP) with three different logistic regression models (first order logistic, second order logistic and stepwise logistic regression methods). The experimental result showed that the predictive performance of the proposed OPP approach exceeded conventional NN as well as multivariate discriminant analysis.

Subsequently, Huang et al. [1] introduced a new machine learning technique called Support Vector Machine (SVM). This chapter compares a BPNN as a benchmark with the newer evolutionary technique. Both methods obtained around 80 % for bond rating data sets from the United States and Taiwan markets from 1991–2000 and 1998–2002, respectively.

Lee [26] also applied SVM to predict bond rating of 3017 Korean companies that have been rated from the year 1997 to 2002. Lee et al.'s paper compared SVM to BPNN, MDA and Case Base Reasoning (CBR). SVM showed its superiority for bond rating prediction accuracy in both the training data and validation data. The second best model to predict bond ratings was the BPNN followed by the MDA model.

More recently, Hajek and Olej [2] utilized kernel-based approaches as well as the design of SVM to predict corporate and municipal credit rating classification. This result found that the model can predict 87 % and 91 % for two and seven class CCR. Kors et al. [27] attempted to predict the bond rating of 500 S&P listed firms for the years of 2008, 2009 and 2010. In order to estimate the best model, the study applied MDA, ordered logit and ordered probit. Ordered logit outperformed the other two models with 63.9 % accuracy. This result is slightly higher than the ordered probit model that classified 61.7 % of cases correctly. MDA produced the lowest predictive abilities with 57.9 % classification accuracy.

Dianelli, Giunta and Cipollini [5] used the Basel III approach to determine small- and medium-sized enterprises' (SME) credit worthiness. They studied the SME's credit history as well as the financial ratios to develop a failure prediction logit model of 187 Italian SMEs. Muscettola and Naccarato [7] also tries to predict SME credit risk using logistic regression for 5000 north Italian SMEs in the period 2007–2010. Bauer and Agarwal [28] compared z-scores, contingent claims-based models and hazard models in the UK to predict bankruptcy and found that z-scores provided better results than the other methods. One of the most recent researches was carried out by Galil and Sher [29]. This study explored the different characteristics of financial distress associated with companies using CDS spreads and financial ratios. In particular, this research used logistic regression to explain default model probabilities from 5542 de-listings, 1495 bankruptcies, 1503 drawdowns and 1098 cases when companies experienced defaults. However, this study found there is no accuracy improvement by switching from a logit model to hazard model as suggested by previous research.

The only study that predicts Sukuk rating is Arundina and Omar [9]. This study used Ordered Logistic Regression and Multinomial Logistic Regression to create a model of Sukuk ratings from several theoretical variables adapted from Touray [30]. The results show 70 % accuracy when using the Ordered Logistic model and 78.3 % prediction accuracy when using the Multinomial Logistic model. Due to a limited number of samples, this study could not validate the model using the validation sample. Another limitation of this study is that it did not take into account the different Sukuk structures that theoretically have different credit risk. Hence, the current study tries to extend this previous study on Sukuk rating prediction incorporating various Sukuk structures as well as industrial categories.

Variable Selection

This study follows Altman's [31] approach to select the most common variables used by previous studies that are relevant for application in Sukuk rating prediction. Based on the summary of variables used (see Appendix 1) it is found that size variable is the most frequent variable used in previous research. Additionally, Table A1 shows that liquidity, profitability and leverage ratios are also constantly used and considered to be important indicators in the prediction of bond rating. Some studies consider qualitative variables as additional information for the company, such as, subordination, guarantee status or tax burden.

Market variables such as credit spread, stock price volatility or GDP are rarely used in previous bond rating studies. However, Du [32], Hajek and Michalak [6], Doumpos et al. [8] believe that a market variable is an important indicator that captures the situation of a company or particular security. Thus, this current study also employs a market variable in the prediction model.

3 Data and Research Method

Data and Sample Selection

In this section, we will describe the sample selection process. For the statistical method all data is utilized to build the model. However, for the NN method, the data is divided into two sets: the training data, which is used to build the model, and the validation or holdout data, which is utilized to validate the model. The validation method is important to avoid the 'overfitting' problem. The financial variables were collected quarterly. This study uses Huang et al.'s [1] and Cao et al.'s [33] approach in using financial variables two quarters before the rating release date as the basis for rating prediction.

Dependent and Independent Variables

Sukuk rating is used as a dependent variable. Since the dependent variable has a polychotomous nature the study divided the dependent variable into four group categories. Table 8.2 presents all four categories that consist of rating AAA, AA, A and BBB.

Meanwhile, the list of the 24 independent variables is presented in Table A.3. All financial ratios represent size, leverage, liquidity, coverage, profitability and market variables. The data is obtained from Bloomberg and the Bond Pricing Agency Malaysia database. Credit enhancement and Sukuk structure variables are taken from Rating Agency Malaysia and MARC website.

Table 8.2 Number of all rating classes

Classes	Number
AAA	35
AA	120
A	150
BBB	9
Total	314

Research Method

This study employs Multinomial Logit Regression, Decision Tree (DT) and AINN to create a model of rating probability from several theoretical variables obtained from company financial ratios and/or financial characteristics. The choice of these methodologies was inspired by prior studies conducted on ratings prediction models.

Multinomial Logit Regression

If the categorical dependent variable is ordered or unordered in nature and if the problem involves more than two categories, an extended version of the binary logit model (known as a polychotomous or multinomial regression model) can still be applied to the problem [30]. Multinomial logit also provides stepwise models to find the best-fit model according to its significances. The multinomial logit strategy usually involves allowing one category to assume a specific value, for instance Y = h0, where h0 = 0. This category is then used as the reference category for the rest of the other categories. The method is also known as the base-line category type, which is explained by Menard [34], Agresti [35], Hosmer and Lemeshow [36] and Touray [30]. In this case, the baseline category logistic model as opposed to the other rating classes is presented in the equation 8.2:

$$\log \frac{p(group\ j)}{p(group\ J)} = \alpha_{i0} + \beta_{i1}X_1 + \beta_{i2}X_2 + \beta_{i3}X_3 + \beta_{i4}X_4 + \beta_{i5}X_5 + \cdots\cdots + \beta_{in}X_n * (Dp) \tag{8.2}$$

Where:

J refers to the baseline category,
j refers to any other category,
α_{i0} is a constant term,

β is a respective coefficient on predictor X,
$X_1 - X_n$ are predictor (independent) variables,
Dp dummy/binary variables

Decision Tree

Decision tree algorithms are relatively stable and accurate [37, 38]. DTs also do not require any assumptions about frequency distributions of target variable data and are relatively insensitive to outlying values [37]. A DT is a decision support system that utilizes a tree-like-graph decision and their possible classification, including a chance event problem or a classification problem. A decision tree, or a classification tree, is used to derive a classification function that determines the value of a dependent attribute (variable) given the values of independent attributes (input variables). This overcomes a problem known as supervised classification because the information on dependent variables and the targeted classes are given [39]. Rokach and Maimon [40] and Novotna [4] stated that a DT is a predictive model that can be used for both decision and classification problems. Classification trees can be used to classify an object, such as companies, into a pre-defined set of classes or rating groups. The companies are firstly classified according to most relevant variable(s), then into sub-groups according to other variable(s) until it is classified into several classes or groups [4, 41]. In order to construct classification trees this requires several phases/steps. First is the selection of attributes to place at the root node. This root node has external edges and no incoming edges. Branch nodes have one incoming edge and the leaves or decision nodes are the nodes without external nodes. Figure 8.1 below is a simple illustration of DT chart.

In the decision tree, each internal node splits up the input set into several sub-sets, one for every value of the attribute [42]. The process is then repeated recursively for each branch. At the end, the last nodes are assigned to one class representing the most appropriate target value (a leaf). The samples (instances) are classified by navigating them from the root n to a leaf, according to the outcome of the tests along the path.

Artificial Intelligence Neural Network

Artificial neural systems emulate the human learning process. The interconnected system of nodes/neurons learns relationships between inputs and outputs by repeatedly sampling input/output information sets. NNs have a

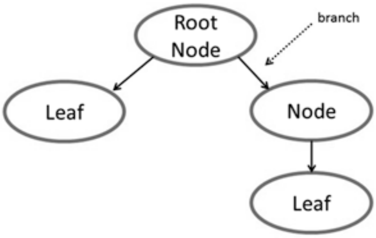


Fig. 8.1 Simple decision tree

particular advantage over expert systems when data are noisy or incomplete. NNs are characterized by three architectural features: inputs, weights and hidden units.

The NN method is trained to learn to classify patterns [43, 44]. Multilayer Perceptron (MLP) is a feed forward NN model that maps sets of input data into a set of appropriate output. The model adjusts weights for connections between layers until the forecast error is minimized and a steady state is reached [43, 44, 45]. In order to build the model, neural networks form three nodes (namely input nodes, hidden nodes and output nodes). The input nodes have linear activation functions and no threshold. Each hidden unit node and each output node have thresholds associated with them in addition to weights [46].

As a general overview Fig. 8.2 displays how the NN training process works. The inputs represent the data received by the system (for example, company financial ratios). Each piece of information is assigned a weight (w1,1; w2,1; ...wn1) that designates its relative importance to each hidden unit. These weights are 'learned' by the network over the course of the 'training' process using one a range of training algorithms. However, the backpropagation algorithm is the most widely applied. Each hidden unit computes the weighted sum of all inputs and transmits the result to other hidden units. In parallel, the other hidden units are weighting their inputs so as to transmit their signal to all other connected hidden units. Receipt of the signal from other hidden units further transforms the output from each node, the system continues to iterate until all information is incorporated.

This model incorporates complex correlation among the hidden units to improve the model fit and reduce type 1 and type 2 errors. Hair et al. [47] stated that a Type I error is the probability of incorrectly rejecting the null hypothesis. In most cases it means that a difference or correlation exists when it actually does not, it is also known as alpha. In this case, a Type I error occurs when the predictor has a higher rating than the actual rating. Type II errors mean the probability of incorrectly failing to reject the null hypothesis,

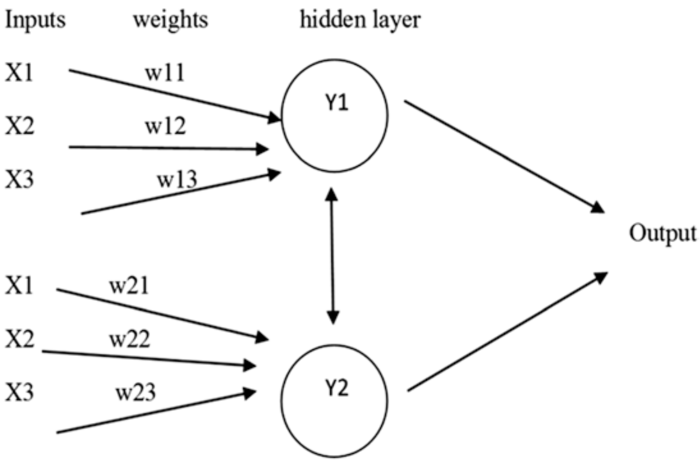


Fig. 8.2 An artificial neural network

known as beta (1 – alpha). The Type II error is inversely related to Type I error. Related to rating, Type II errors occur when the prediction is lower than the actual rating.

4 Result and Analysis

Data Screening

According to Hair [47] and Hosmer and Lemeshow [36], logistic regression is popular in part because it enables the researcher to overcome many of the restrictive assumptions of OLS regression such as linearity, normality and heteroscedasticity. In general, there are three initial steps for screening data prior to conducting a multivariate analysis. The first step deals with missing data. The second is assessing the effect of extreme values (i.e. outliers) for the analysis and finally the evaluation of the adequacy of fit between the data and the assumptions (i.e. statistical testing such as MSE, MAE, MAPE).

Due to incomplete data and influential outliers, the sample is reduced from 340 initial data to 314 samples. Logistic regression methods cannot handle multi-collinearity problems. A high level of this problem will create difficulties in the detection of significant relationships between variables. Hence, Menard [34] suggests performing a collinearity diagnostic test that is used to assess the inter-relationships between indicators. According to this test, it was found that the collinearity problem among size variables is high. Therefore, we included one variable that represented size and excluded other variables.

In this case, only total asset is utilized in the model. Total asset is selected from other size variables since this variable is mostly used to represent the size of the company 4, 6, 8.

Multinomial Logistic Result

The test of the overall relationship between the independent variables and groups defined by the dependent variable was conducted and the result is presented in Table 8.3. The likelihood ratio of the model is used to the test whether all predictors' regression coefficients in the model are simultaneously zero. We can reject the null hypothesis, which leads us to conclude that at least one of the regression coefficients in the model is not equal to zero and state that our final model is significantly better than the intercept model. Subsequently, Table A2 in Appendix 2 shows three pseudo R-square values that are presenting the fit index of the model's adequacy. The result shows the efficacy of the model to capture the data. Norusis [48] and Touray [30] acknowledge that this method attempts to mimic the OLS R^2 measure.

Before proceeding to the logistic test, this study performs stepwise method forward entry selection. This stepwise method allows the system to enter significant variables only into the model according to a chi-square value based on the likelihood ratio test. This step is used to objectively determine the significant variables in order to predict Sukuk rating.

Multinomial logistic regression compares multiple groups through a combination of binary logistic regression. The group comparisons compare the dependent variables (A, AA and AAA) with one reference group as a baseline comparison. The multinomial logit strategy usually involves allowing one category to assume a specific value, say Y = h0, where h0 = 0. This category is then used as the reference category for the rest of the other categories. Here, we use BBB as the reference category. This translates into the equations 8.3–8.5, in this case.

$$\begin{aligned}&\text{Predicted Logit}\left(\frac{\text{AAA}}{\text{BBB}}\right),\text{rating AAA contrast to BBB}\\&=1278.89-6.78\ \text{log share price}+2.18\ \text{ROA}-231.35\ \text{log GDP}\\&\quad+1.81\ \text{lt debt to total asset}-1.70\ \text{tot debt to tot asset}+15.05\ \text{cash ratio}\end{aligned} \tag{8.3}$$

$$\begin{aligned}&\text{Predicted Logit}\left(\frac{\text{AA}}{\text{BBB}}\right),\text{rating A contrast to BB}\\&=1261.44+2.32\ \text{ROA}-227.88\ \text{log GDP}+1.75\ \text{lt debt to total asset}\\&\quad-1.65\ \text{tot debt to tot asset}+15.51\ \text{cash ratio}\end{aligned} \tag{8.4}$$

Table 8.3 Multinomial logistic parameter estimates

Model	Action	Effect	Likelihood ratio tests	Wald test		
			Sig.	$\frac{p(A)}{p(BBB)}$	$\frac{p(AA)}{p(BBB)}$	$\frac{p(AAA)}{p(BBB)}$
0	Entered	Intercept	.	1278.891	1261.436	1270.37
1	Entered	LPX_LAST	0.000***	(−6.777)*	−0.73	(10.849)**
2	Entered	structure	0.000***			
		[structure=BBA]		−16.247	−18.09	−26.92
		[structure=Ijarah]		14.106	44.655	47.617
		[structure=Mudharaba]		8.442	11.011	48.398
		[structure=Murabahah]		17.713	19.256	−7.3
		[structure=Musharakah]		0^c	0^c	0^c
3	Entered	Industrial Sector	0.000***			
		[construction and engineering]		−5.365	−3.161	−18.45
		[consumer product]		0.062	−14.295	10.494
		[diversified holdings]		−7.122	14.489	−9.668
		[industrial product]		−3.661	−3.547	3.953
		[infrastucture and utilities]		−24.139	−17.742	−7.122
		[mining and petroleum]		−52.349	−35.139	−84.9
		[plantation and agriculture]		7.8	−3.403	−9.819
		[property]		−16.753	−19.227	4.3
		[trading/service transportation]		0^c	0^c	0^c
4	Entered	guar_status	0.000***			
		[guar_status=.00]		−8.314	−15.06	−35.221
		[guar_status=1.00]		0^c	0^c	0^c

(*continued*)

Table 8.3 (continued)

Model	Action	Effect	Likelihood ratio tests	Wald test		
			Sig.	$\frac{p(A)}{p(BBB)}$	$\frac{p(AA)}{p(BBB)}$	$\frac{p(AAA)}{p(BBB)}$
5	Entered	RETURN_ON_ASSET	0.001***	(2.182)*	(2.32)*	1.685
6	Entered	LLGDP	0.003***	(−231.359)*	(−227.882)*	(−227.237)*
7	Entered	LT_DEBT_TO_TOT_ASSET	0.000***	(1.813)*	(1.747)*	0.946
8	Entered	subordinate	0.000***			
		[subordinate=.00]		12.586	−7.896	−53.269
		[subordinate=1.00]		0[c]	0[c]	0[c]
9	Entered	TOT_DEBT_TO_TOT_ASSET	0.000***	(−1.701)**	(−1.654)*	−1.182
10	Entered	CASH_RATIO	0.007***	(15.049)*	(15.51)**	(17.284)**

Note: *,**,*** implies 10 %, 5 %, 1 % significant level

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

[a]The reference category is: BBB.

[b]Floating point overflow occurred while computing this statistic. Its value is therefore set to system missing.

[c]This parameter is set to zero because it is redundant.

$$\text{Predicted Logit}\left(\frac{\text{A}}{\text{BBB}}\right), \text{rating A contrast to BBB}$$
$$= 1270.37 + 10.85 \log \text{share price} - 227.24 \log \text{GDP} + 17.28 \text{ cash ratio} \quad (8.5)$$

The models are estimated for Sukuk rating A relative to BBB, rating AA to BBB and rating AAA relative to BBB. The Wald test in Table 8.3 presents results that show return on equity, current ratio, total debt to total equity, long-term debt to total asset, cash ratio and GDP is found statistically significant as it affected the classification probability between A, AA and AAA to BBB ratings. The coefficient shows the relationship's probability between one ratings class relative to the baseline category. In this case, it is shown that one unit increase in return on equity will increase the company's log-odds of getting an A rating as opposed to a BBB by 0.524 times, keeping all other variables constant. The same interpretation applies with other significant variables. The negative value indicates the negative correlation between independent and dependent variables.

The likelihood ratio test evaluates the overall relationship between an independent variable and the dependent variable. As shown in Table 8.3, the likelihood ratio test result shows that out of 24 variables: share price, structure, industrial sector, existence of guarantor, return on asset, GDP, long-term debt to total asset, subordinate, total debt to total asset and cash ratio are variables that significantly determine the Sukuk rating model. The null hypothesis that the coefficients associated with those variables were equal to zero is rejected. Hence, there is an existence of a relationship between the rating and the aforementioned variables.

However, if an independent variable has an overall relationship to the dependent variable, it might or might not be statistically significant in differentiating between pairs of groups defined by the dependent variable.

Table 8.4 presents the prediction accuracy test of the estimated multinomial logistic model, the final test of this method. The result shows that an overall 91.72 percent (268/314) of all valid cases are correctly classified into their original ratings classes.

The result indicates that the A rating category has the highest correct classification rate with 93.33 % (140 cases out of 150 cases) of individual cases correctly classified. The AA rating category can correctly predicted 110 cases out of 150 cases (93.3 %) and rating AAA category attains a 88.57 % accuracy rate. Meanwhile, the lowest accuracy rate is obtained by BBB rating class with 77.78 % accuracy rate.

Table 8.4 Multinomial logistic classification accuracy

Multinomial Logistic						
	Predicted					
Observed	AAA	AA	A	BBB	Number Correct	% Correct
AAA	31	4	0	0	31/35	88.57
AA	9	110	1	0	110/120	91.67
A	1	8	140	1	140/150	93.33
BBB	0	1	1	7	7/9	77.78
Number of prediction/ Total sample	13.06 %	39.17 %	45.22 %	2.55 %	288/314	91.72

Decision Tree and Artificial Intelligence Neural Network Result

Kim et al. [49] state that ANNs and DTs were the most common methods used in AI and data mining applications. DTs have an advantage compared to NNs because DTs can tolerate a small size of target variable data for the learning phase (training data set) to gain an induced rule among multivariate data [50]

Decision trees are a machine learning tool for building a tree structure from a training data set. The model produced by the DT is visualized in Fig. 8.3. There are 65 leaves that classify 314 samples into four classes of rating. The tree shows that out of 26 input variables, only 15 variables are used to build the tree. The variables are: guarantee status, subordinate, total asset, long-term debt to total capital, total debt to total asset, cash ratio, quick ratio, ROA to ROE, total market value, share price, operating margin, GDP, interest rates, Sukuk structure and the industrial sector.

It should be noted that NNs suffer a few disadvantages including computational intensity and an inability to explain conclusions and how they are derived [44] hence being classified as a 'black box' methodology. Therefore, the usefulness for social research is limited because the NN presents only prediction results and does not present features of the underlying process relating to the inputs and output [51]. Pre-processing and choice of the input variables will be decisive in how the NN model performs and in showing how robust its results are. Hence, this study utilizes selected variables from the DT to build an NN model and predict Sukuk ratings.

To predict the performance of both classifier models, this study utilizes two sets of data: the general training data and the validation data. The training data is used to build the model that creates a set of classifiers. The validation

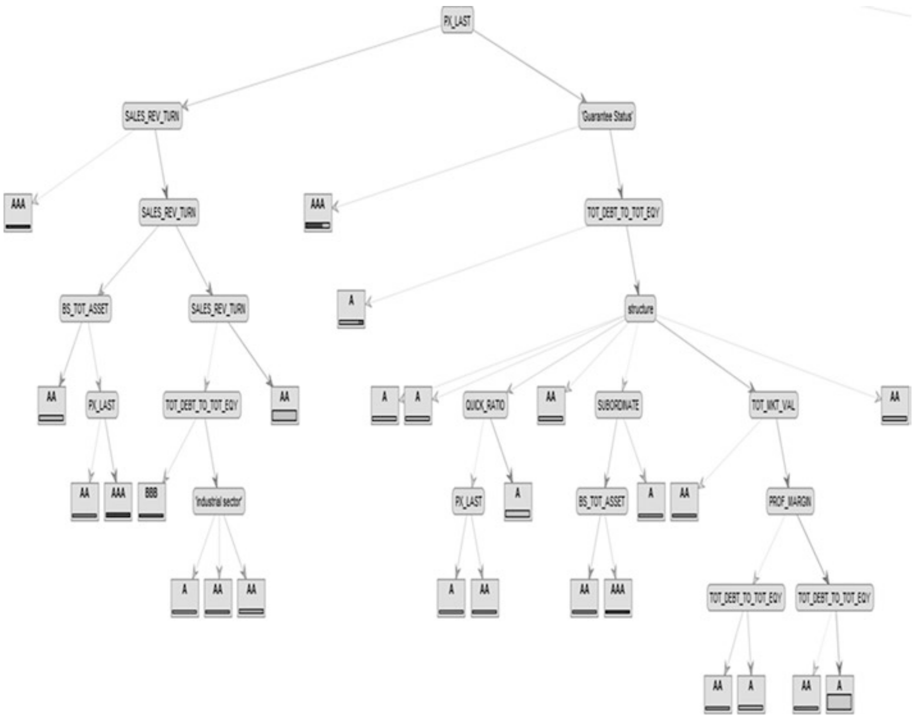


Fig. 8.3 Decision tree visualization

data is used to optimize parameters of those classifiers. The tests are divided into several phases discussed below

Phase one: General training

The average hit accuracy of training samples of both the DT and the multi-layer perceptron NN model are shown in Table 8.5. The columns show the predicted values and the rows present the actual or the observed values.

All 314 Sukuk samples are being used to build the model. The classification model is produced from the training samples. Those samples are identified as a training group in order for the DT and NN machine to learn the data and classify the patterns of the data. Following this, the output is classified into four classes; AAA, AA, A, and BBB. The classifier predicts the class of each instance and generates a classification accuracy which shows the performance of the model.

Table 8.5 Decision tree versus neural network classification accuracy

Decision tree training set						Neural network-decision tree variables					
	Predicted						Predicted				
Observed	AAA	AA	A	BBB	Percent correct	Observed	AAA	AA	A	BBB	Percent correct
AAA	33	1	1	0	94.29	AAA	34	1	0	0	97.14
AA	2	114	2	2	95.00	AA	1	119	0	0	99.17
A	2	0	148	0	98.67	A	0	1	149	0	99.33
BBB	0	0	1	8	88.89	BBB	0	1	2	6	66.67
No Predicted/ Tot Sample	11.78 %	36.62 %	48.41 %	3.18 %	96.50	No Predicted/ Tot Sample	11.15 %	38.85 %	48.09 %	1.91 %	98.09

The result the table shows that 96.5 % of the total cases (303 cases out of a total of 314 cases) are correctly classified by the DT model, whereas, the multilayer perceptron NN model correctly classified 98.09 % (308 cases out of total 314 cases). The highest prediction rate is once again obtained by the A rating category in the NN result with 99.33 % classification accuracy rate of individual rating (149 cases out of total 150 A rating cases) and the lowest rate is also attained by NN BBB rating category that only able to correctly classify 6 cases out of 9 individual rating cases in this group.

Phase Two: Validation test

The validation test evaluates the classifier on how well it predicts a certain percentage of the data, which is held out for testing. The overall sample of Sukuk is randomly split into two groups consisting of training and validation Sukuk. Of the Sukuk cases 80 % (251 instances) are taken as training data in order to build the model while the remaining 63 instances are used for validating the model.

Table 8.6 presents the hit ratio of the methods. The DT model was able to predict 77.78 % of the validation test for both validation cases (49 instances out of total 63 validation instances). The result shows that the highest accuracy rate is obtained by the A rating class with 82.76 %, followed by AA rating class with an 80.77 % accuracy rate. However, the random sample validation test does not show good results for the BBB rating class. The validation did not predict two instances in that class.

The validation test result for the NN shows better results. The overall accuracy reaches 79.37 %, 50 correctly predicted cases out of 63 validation data. Furthermore, the model misclassified 20.63 % (13 cases out of 63 validation data) Sukuk in the validation test. Rating BBB gets a 100 % accuracy rate, creating a huge gap with the DT result. The accuracy of rating AAA of NN model supersedes the DT model for that category. Whereas, the prediction accuracy of the DT method is better than the NN method for AA and A rating classes.

Table 8.6 Decision tree and neural network validation test

Decision tree-Validation Test						Neural network-Validation Test					
	Predicted						Predicted				
Observed	AAA	AA	A	BBB	Percent Correct	Observed	AAA	AA	A	BBB	Percent Correct
AAA	4	1	1	0	66.67	AAA	5	1	0	0	83.33
AA	2	21	2	1	80.77	AA	0	20	4	2	76.92
A	0	4	24	1	82.76	A	0	5	23	1	79.31
BBB	0	1	1	0	0.00	BBB	0	0	0	2	100.00
No Predicted/Tot Sample	9.52 %	42.86 %	44.44 %	3.17 %	77.78	No Predicted/Tot Sample	7.94 %	41.27 %	42.86 %	7.94 %	79.37

Table 8.7 Comparison of significant variables

Multinomial Logistic	Decision Tree/Neural Network
Share Price	Share Price
Sukuk Structure	Guarantee Status
Industrial Sector	Industrial Sector
Guarantee Status	Sukuk Structure
Return On Asset	Long Term Debt to Total Capital
GDP	Total Asset
Long Term Debt toTotal Asset	Interest Rate (KLIBOR)
Subordinate	Total Market Value
Total Debt to Total Asset	GDP
Cash Ratio	Operating Margin
	Quick Ratio
	Book Value Per Share
	Sales
	Subordinate
	Long Term Debt toTotal Asset
	Cash Ratio

Result comparison

This section discusses the comparison of variable selection and prediction accuracy using multinomial logistic, DT and NN methods. Using the approach of numerous previous studies related to bond prediction and financial distress, 24 independent variables were selected to develop a model in predicting Sukuk rating in the Malaysian market. Among all the variables, it is empirically found that some variables can significantly determine the Sukuk rating based on of multinomial logistic and DT approach. Table 8.7 presents a comparison of both lists of significant variables.

As shown in the Table 8.7, the significant variables derived from both methods are largely similar. However, the list variables from data mining are more extensive than logistic variables. It needs to be highlighted that both results show that share price, Sukuk structure, industrial sector and guarantee status are the key factors that determine Sukuk rating.

As for share price, Standard & Poor's [14] argue that share price gives information beyond the fundamental condition of companies. It also reflects the current condition of the company. Therefore, the movement in share price

Table 8.8 Comparison classification accuracy

General Training				Validation Test		
Observed	M-Logit	Decision Tree	Neural Network	Observed	Decision Tree	Neural Network
AAA	88.57 %	94.29 %	97.14 %	AAA	66.67 %	83.33 %
AA	91.67 %	95.00 %	99.17 %	AA	80.77 %	76.92 %
A	93.33 %	98.67 %	99.33 %	A	82.76 %	79.31 %
BBB	77.78 %	88.89 %	66.67 %	BBB	0.00 %	100.00 %
Percent Correct	91.72	96.50	98.09	Percent Correct	77.78	79.37
Type I Error	20	5	5	Type I Error	8	5
Type II Error	6	6	1	Type II Error	6	8

has an impact ono its credit rating. Unlike financial ratios, Sukuk structure and industrial sector have a different impact on ratings. Both have an indirect impact on credit risk, yet they are very important in the ratings assessment. Sukuk structure affects Sukuk rating due to the different nature of credit risk embedded in each of the specific structures. Industrial sector shows the business risk of the company. It indicates the level of competition and market concentration of the industry. However, guarantee status also contributes significant aspects in credit rating as it provides strong protection to the company in the event of the inability to make payments. Apart from this, profitability ratios, leverage ratios, liquidity ratios, macroeconomic variables and credit enhancement are also found to be significantly related to Sukuk rating according to both methods.

Table 8.8 shows that the NN model can classify the general training data with the highest accuracy, 98.09 %. The decision tree, on the other hand, correctly classified 96.5 % of all valid cases into their original rating classes. These two methods are followed by Multinomial Logistic, which obtains the lowest prediction rate among these three methods with an 84.5 % accuracy rate.

With regards to the misclassification cost of Type I and Type II errors, in our case, the lowest Type I error for the baseline training model is obtained by both NN and DT methods with five misclassification case. The multinomial logistic model generates the highest Type I error with 20 incorrectly predicted cases. On the other hand, for Type II errors, NNs have only one Type II error, while the DT and multinomial logistic model have six Type II errors.

Koh [52] also adds that the misclassification cost for Type I error is higher than Type II error. From these two methods, it is shown that the multinomial logistic and the DT result have more Type I and Type II errors than the NN. This leads us to a conclusion that the NN method will give more robust results in terms of accuracy.

Many previous studies of bond rating prediction show that AI and data mining methods can perform better than the traditional statistical methods as these models require some assumptions that may not be appropriate for bond rating [1]. For instance, Dutta and Shekar [22], Ederington et al. [53], Kim et al. [49], Kwon et al. [25], Maher and Sen [54], Chavesuuk et al. [23] and Hu and Ansell [55] found that NNs have a much better predictive ability than linear regression, MDA and/or logistic regression. This conclusion is consistent with the empirical results of this study, which finds that data mining and AI methods are superior to multivariate analysis methods when applied to the task of predicting ratings in the case of Malaysian Sukuk.

5 Conclusion

This study compares a multinomial logistic, a DT and a NN model to predict corporate Sukuk ratings. From the classification accuracy, we can conclude that non-statistical methods, in this case a NN and DT are more powerful than the multinomial logistic statistical method when predicting Sukuk ratings using the samples from the Malaysian Sukuk market. Furthermore, these findings are expected to enrich the literature and have practical implications. This model is expected to be useful for rating agencies to perform an initial rating shadow rating and for issuing companies and fund managers to conduct their own credit analysis for risk management and trading purposes. In addition, these models can be of use to banks that rely on rating systems in order to improve risk-assessment techniques, pricing strategies and provisioning levels as required in Basel III

Empirical results are also expected to contribute a wealth of knowledge to the development of Islamic finance while encouraging analysts and academic researchers to develop other potential research related to this topic. It is expected that further research efforts in this area could benefit from this study. On the issue of data collection, we would suggest more observations are collected. The analysis will be more comprehensive if the study incorporates both types of Sukuk securities to compare the asset-backed Sukuk and the asset-based Sukuk (as discussed here) in its discussion. For dependent variables, researchers can try to use more specific levels of rating categories rather than general levels of rating categories as we have used in this study. One interesting issue would be the guarantee status variable. There are various types of guarantee status or binding agreements in accordance with the structure of Sukuk. The study that considers the various types of this guarantee status in the model will give a better picture with regard to Sukuk credit risk profiles.

6 Appendices

Appendix 1

Table A1 Summary of variables, profiles and techniques used in bond and corporate ratings previous studies

No.	Author, Date	Size Variable	Liquidity Variable	Profitability Variable	Coverage Variable	Leverage Variable	Market Variable	Qualitative Variable	Method
1	Belkaoui [21]	√	√		√	√		√	MDA
2	Ederington et al. [54]	√		√	√	√		√	MDA
3	Dutta and Shekar [22]		√	√		√	√	√	AI ANN
4	Singleton and Surkan [24]	√	√	√	√	√			AI ANN MDA
5	Kwon et al. [25]	√	√	√	√	√	√	√	OPP
6	Chaveesuk et al. [26]	√	√	√	√	√			RBF, LVQ,BP,Logit
7	Kamstra et al. [56]	√		√	√	√		√	COMBINED LOGIT
8	Huang et al. [1]	√	√	√	√	√			SVM, BP
9	Touray [30]	√	√		√	√		√	MDA-MLOGIT
10	Kim [57]	√	√	√		√	√	√	AI ANN
11	Cao et al. [33]	√	√	√	√	√	√		Logit, OPR, SVM
12	Brabazon and Neill [58]	√	√	√	√	√	√		Grammatical Evolution
13	Lee [26]	√	√	√	√	√	√		SVM, BP, MDA
14	Hwang et al. [59]	√	√	√	√	√		√	OLPM
15	Hájek and Olej [2]	√	√	√		√			SVM, ANN
16	Körs et al. [27]	√	√	√	√	√	√		MDA, Logit, Probit
17	Mizen and Tsoukas [3]	√	√	√	√	√		√	OPM
18	Novotna [4]	√	√	√	√	√			MDA, M-logit, Decision Tree
19	Hajek and Michalak [6]	√	√	√		√	√	√	MLP, RBF, SVM, NB, RF, LDC, NMC
20	Doumpos et al. [8]	√	√	√	√	√		√	MCDM

Appendix 2

Table A2 Model fitting information and pseudo R-square

	Model fitting criteria	Likelihood ratio tests		
Model	-2 Log Likelihood	Chi-Square	Df	Sig.
Intercept Only	667.233			
Final	131.891	535.342	60	0.000***
Pseudo R-Square		Cox and Snell		0.818
		Nagelkerke		0.928
		McFadden		0.799

Note: *** implies 1% significance level

Appendix 3

Table A3 List of independent variables

Variable	Abbreviations	Formula
Size		
X1 BS_TOT_ASSET	Total asset	Total current assets plus total long-term assets
X2 SALES_REV_TURN	Sales/revenue/turn over	Total revenue
X3 TOTAL_EQUITY	Total equity	Total shareholder equity
X4 TOT_MKT_VAL	Total market value	Total market capitalization
Leverage		
X5 LT_DEBT_TO_TOT_ASSET	Long-term debt to total asset	Long-term debt/total asset
X6 LT_DEBT_TO_TOT_CAP	Long-term debt to total capital	Long-term debt/total capital
X7 TOT_DEBT_TO_TOT_ASSET	Total debt to total asset	Total debt/total asset
X8 TOT_DEBT_TO_TOT_EQY	Total debt to total equity	Total debt/total equity
Liquidity		
X9 CUR_RATIO	Current ratio	Current asset/current liabilities
X10 QUICK_RATIO	Quick ratio	Current asset-inventory/ current liabilities
X11 CASH_RATIO	Cash ratio	Cash + marketable securities/current liabilities
Coverage		
X12 INTEREST_COVERAGE_ RATIO	Interest coverage	EBITDA/Interest expense
Profitability		
X13 PROF_MARGIN	Net profit margin	Net income/sales
X14 RETURN_ON_ASSET	Return on asset	Net income/total asset
X15 RETURN_COM_EQY	Return on equity	Net income/equity
X16 OPER_MARGIN	Operating margin	Operating profit/sales

(*continued*)

Table A3 (continued)

Variable	Abbreviations	Formula
Market		
X17 BOOK_VAL_PER_SH	Book value per share	Book value/last price
X18 PX_LAST	Share price	Share price
X19 GDP	Gross domestic product	Gross domestic product
X20 INTEREST RATES	Klibor 3 month	Klibor 3 month
X21 Industrial Sector	Industrial sector	Industrial product, property, plantation and agriculture, construction and engineering, trading/ service transportation, infrastructure and utilities, consumer product, diversified holdings, mining and petroleum
Credit Enhancement		
X22 Guarantee Status	Guarantee status	With/without guarantor
X23 Subordinate	Subordination	Senior/junior Status
Specific Variables		
X24 Structure	Sukuk structure	Musharakah, Ijarah, Mudharaba, Murabahah, BBA

References

1. Huang, Z., Chen, H., Hsu, C.-J., Chen, W.-H., & Wu, S. (2004). Credit rating analysis with support vector machine and neural networks: A market comparative study. *Decision Support System, 37*, 543–558.
2. Hájek, P., & Olej, V. (2011). Credit rating modelling by kernel-based approaches with supervised and semi-supervised learning. *Neural Computing and Applications, 20*(6), 761–773. doi:10.1007/s00521-010-0495-0.
3. Mizen, P., & Tsoukas, S. (2012). Forecasting US bond default ratings allowing for previous and initial state dependence in an ordered probit model. *International Journal of Forecasting, 28*(1), 273–287.
4. Novotna, M. (2012, Sept). *The use of different approaches for credit rating prediction and their comparison*. In Proceedings 6th International Scientific Conference Managing and Modelling of Financial Risks (pp. 448–457). Ostrava.
5. Dainelli, F., Giunta, F., & Cipollini, F. (2013). Determinants of SME credit worthiness under Basel rules: the value of credit history information. *PSL Quarterly Review, 66*, 21–47.
6. Hajek, P., & Michalak, K. (2013). Feature selection in corporate credit rating prediction. *Knowledge-Based Systems, 51*, 72–84. doi:10.1016/j.knosys.2013.07.008.

7. Muscettola, M., & Naccarato, F. (2013). Probability of default and probability of excellence, an Inverse Model of Rating. One more tool to overcome the crisis: An empirical analysis. *Business System Review, 2*(2).
8. Doumpos, M., Niklis, D., Zopounidis, C., & Andriosopoulos, K. (2014). Combining accounting data and a structural model for predicting credit ratings: Empirical evidence from European listed firms. *Journal of Banking & Finance, 50*, 1–9.
9. Arundina, T. and Mohd Azmi Omar. (2010). *Determinant of Sukuk Rating*; *The Case of Malaysia*. Thesis Kulliyah of Economics and Management Science, International Islamic University Malaysia.
10. Tariq, A.A. (2004). *Managing financial risks of Sukuk structures*. Unpublished Dissertation, Loughborough University, Leicestershire, United Kingdom.
11. AAOIFI. (2005). *Accounting and Auditing Organization of Islamic Financial Institutions*, *Shari'ah Standards*.
12. Maya Puspa Rahman and Mohd Azmi Omar. (2012, July). *Factors influencing the excess returns of Sukuk*: *An empirical analysis on USD denominated Sukuk Issued by Corporations in the Gulf Cooperation Council* (*GCC*). Paper presented in the 3rd Gulf Research Meeting, University of Cambridge, Cambridge, United Kingdom.
13. Fitch, R. (2011). *Rating Sukuk Cross-Sector Criteria*. London: Author.
14. Standard and Poor's. (2007). *Standard and poor's approach to rating Sukuk*. Standard and Poor's Ratings Direct. Paris: Author.
15. Moody's Investor Service. (2006) *Shari'ah and Sukuk*: *A Moody's Primer*. Moody's Investors Service. International Structured Finance. London: Author
16. RAM Rating Service Berhad. (2011). *General approach to Sukuk Rating. Criteria and methodology*. Kuala Lumpur: Author.
17. Malaysian Rating Corporation Berhad (MARC). (2012). *Islamic Financial Institutions*. *MARC Rating Methodology*. Kuala Lumpur: Author.
18. Horrigan, J. (1966). The determination of long term credit sharing with financial ratios. *Journal of Accounting Research, 4*, 44–62.
19. West, R. (1970). An alternative approach to predicting corporate bond rating. *Journal of Accounting Research, 7*, 118–127.
20. Pinches, G., & Mingo, K. (1973). A multivariate analysis of industrial bond ratings. *Journal of Finance, 28*, 1–18.
21. Belkaoui, A. (1980). Industrial bond ratings: A new look. *Financial Management. Autumn, 99*, 44–51.
22. Dutta, S., & Shekar, S. (1988). Bond rating: A non-conservative application of neural networks. *IEEE International Conference on Neural Networks, 2*, 443–450.
23. Chaveesuk, R., Srivaree-ratana, C., & Smith, E. (1997). Alternative neural network approaches to corporate bond rating. *Journal of Engineering Valuation and Cost Analysis, 2*(2), 117–131.
24. Singleton, J. C., & Surkan, A. J. (1995). Bond rating with neural networks. In A. Refens (Ed.), *Neural networks in the capital markets*. New York: John Wiley.

25. Kwon, Y. S., Han, I., & Lee, K. C. (1997). Ordinal pairwise partitioning (OPP) approach to neural networks training in bond rating. *Intelligent Systems in Accounting, Finance and Management, 6*, 23–40.
26. Lee, Y.-C. (2007). Application of support vector machines to corporate credit rating prediction. *Expert Systems with Applications, 33*(1), 67–74. doi:10.1016/j.eswa.2006.04.018.
27. Körs, M., Akta, R., & Do, M. M. (2012). Predicting the bond ratings of S & P 500 firms. *IUP Journal of Applied Finance, 18*, 83–96.
28. Bauer, J. and Agarwal, V. (2014). Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking & Finance, 40,* 432–442.
29. Galil, K., & Sher, N. (2015). *Predicting default more accurately: To proxy or not to proxy for default?* Beersheba, Israel: Ben-Gurion University of the Negev.
30. Touray, A. K. (2004). *Predicting a bond rating*: *Multivariate analysis of corporate bonds*, *a new look at Malaysian corporate bonds*. Unpublished master thesis, International Islamic University of Malaysia, Kuala Lumpur.
31. Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*(4), 589–609.
32. Du, Y. (2003). *Predicting credit rating and credit rating changes* : *A new approach*. Unpublished master thesis, Queen's School of Business, Kingston, Ontario.
33. Cao, L., Kian, L., & Jingqing, Z. (2006). Bond rating using support vector machine. *Intelligent Data Analysis, 10*(70501008), 285–296.
34. Menard, S. (1995). *Applied logistic regression analysis. Sage University Paper series on Quatitative Applications in the Social Sciences* (pp. 07–106). Thousand Oaks, CA: Sage.
35. Agresti, A. (1996). *An introduction to categorical data analysis*. New York: John Wiley & Sons, Inc..
36. Hosmer, D. L., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York: John Willy & Sons Inc..
37. Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth.
38. Scheffer, T. (2001). Finding association rules that trade support optimally against confidence. In L. de Raedt & A. Siebes (Eds.), *Proceedings of the Fifth European Conference on Principles of Data Mining and Knowledge Discovery* (pp. 424–435). Freiburg, Germany, Berlin: Springer.
39. Bhargava, N., Bhargave, R., & Mathuria, M. (2013). Decision tree analysis on J48 Algorithm for Data Mining. *International Journal of Advanced Research in Computer Science and Software Engineering, 3*(6), 1114–1119.
40. Rokach, L., & Maimon, O. (2008). *Data mining with decision trees. Theory and applications*. Singapore: World Scientific Publishing.
41. Witzany, J. (2010). *Credit risk management and modelling*. Praha: Oeconomica.
42. Witten, I. H., Frank, E., & Mark, A. H. (2011). *Data mining practical machine learning tools and techniques* (3 ed.). Burlington, USA: Elsevier.

43. Coats, P. K., & Fant, L. F. (1993). Recognizing financial distress patterns using a neural network tool. *Financial Management (Autumn), 22*, 142–155.
44. Gambrel, P. A. (2004). *The impact of cash flow on business failure*. FL: Nova Southeastern University.
45. Altman, G. M., & Varetto, F. (1994). Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking and Finance, 18*, 505–529.
46. Nazzal, J. M., El-emary, I. M., Najim, S. A., & Ahliyya, A. (2008). Multilayer Perceptron Neural Network (MLPs) For Analyzing the Propoerties of Jordan Oil Shale. *World Applied Sciences Journal, 5*(5), 546–552.
47. Hair Jr., J. F., et al. (2006). *Mutivariate Data Analysis* (6 ed.). New Jersey: Prentice-Hall, Inc.
48. Norusis, M. J. (2009). *SPSS Regression Models 17.0*. SPSS Inc.
49. Kim, J. W., Weistroffer, H. R., & Redmond, R. T. (1993). Expert systems for bond rating: a comparative analysis of statistical, rule-based and neural network systems. *Expert System, 10*(3), 167–172.
50. Kim, K and Han, I. (2001). *The cluster-indexing method for case-based reasoning using self-organizing maps and learning vector quantization for bond rating cases*. Elsevier Science-Expert System with Applications. pp. 147–156.
51. Intrator, O., & Intrator, N. (2001). *Interpreting neural network results*: *A simulation study*. *Computational Statistics & Data Analysis, 37*, 373–393.
52. Koh, H.C. (1992). The sensitivity of optimal cutoff points to misclassification costs of type I and II errors in the going-concern prediction context. *Journal of Business Finance & Accounting, 19*(2), 187–197.
53. Ederington, L. H. Yawitz, J. B. and Roberts, B. E. (1984). *The informational content of bond ratings*. Working Paper Series National Bureau of Economic Research. Cambridge.
54. Maher, J. J., & Sen, T. K. (1997). Predicting bond ratings using neural networks: A comparison with logistic regression. *Intelligent Systems in Accounting, Finance and Management, 6*, 59–72.
55. Hu, Y. and Ansell, J. (2005). *Developing Financial Distress Prediction Models*: *A study of US*, *Europe and Japan Retail Performance*. Working paper series, University of Edinburgh. http://papers.ssrn.com (2009).
56. Kamstra, M., Kennedy, P., & Suan, T. (2001). *Combining bond rating forecasts using Logit*. *The Financial Review, 37*, 75–96.
57. Kim, K. S. (2005). Predicting bond ratings using publicly available information. *Expert System with Applications, 29*, 75–81.
58. Brabazon, A. and Neill, M. O. (2006). *Bond rating using grammatical evolution*. Springer Link. Volume 3005
59. Hwang, R., Cheng, K. F., & Lee, C. (2009). On multiple-class prediction of issuer credit ratings. *Applied Stcochastic Models in Business and Industry, 25*, 535–550. doi:10.1002/asmb.
60. Bond Rating Agency Malaysia. http://www.bpam.com.my.

Part V

Portfolio Management, Analysis and Optimisation

9

Portfolio Selection as a Multi-period Choice Problem Under Uncertainty: An Interaction-Based Approach

Matjaz Steinbacher

1 Introduction

The chapter is a simulation-based survey that applies an interaction-based approach to examine portfolio selection as a multi-period choice problem under uncertainty. Methodologically, it builds on individual agents and their subjective character. Our fundamental premise is that local interaction and information sharing are essential in agents' portfolio selections, while agents' selections are generated from simple behavioural rules on an individual level and local interaction over the social network.

Financial markets are inherently occupied with issues that involve time and uncertainty. The existence of uncertainty is essential to portfolio selection and the main reason for making a portfolio. In a world of certainty investors, we call them agents, would (or should) simply take a single asset with the highest return and solve their objective in the best possible way. Following the pioneering work of Markowitz [1] and the subsequent multi-period extension of Merton [2, 3], the portfolio selection research has developed in different directions. One strand of research has focused on solving an explicit solution to the dynamic choice problem [4–11] (see Campbell and Viceira [12] for an

M. Steinbacher (✉)
Kiel Institute for the World Economy, Kiellinie 66, 24105 Kiel, Germany
e-mail: matjaz.steinbacher@gmail.com

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets*,
New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0_9

overview). Others have examined the effects of various parameters to agents' optimization problems. Garlappi et al. [13] extend the mean-variance model to explicitly account for uncertainty about the estimated expected returns. Liu [14] includes multiple risky assets and predictable returns into the Merton model. Ang and Bekaert [15] examine portfolio choice with regime shifts. Benzoni et al. [16] study portfolio selection when labour income and dividends are cointegrated. Benartzi and Thaler [17] incorporate a behavioural aspect into the static portfolio problem. Berkelaar et al. [18] derive closed-form solutions for optimal portfolio choice under loss aversion. Approximate solutions of the portfolio selection problem based on perturbation methods have been proposed by Campbell and Viceira [19, 20]. Shefrin and Statman [21] develop a positive behavioural portfolio theory. These equilibrium models have given us many helpful insights and reduced the sensitivity of models to the parameter estimates, but have left many empirical facts unanswered, for which they have been subject to severe critique.

We bring the micro-structure into the portfolio selection process, which was missing in the previous models. Our primary aim is to integrate portfolio selection with the agent-based approach and to include a behavioural perspective into agents' choice functions. Knowledge is not fragmented but dispersed between agents and is evolving over time (see Hayek [22] for a discussion). We assume that each agent is only able to follow the return of a portfolio he has but is ignorant about the dynamics that drive the returns. Then, these bounded rational agents are constantly engaged in local interactions with adjacent counterparts, affecting the decisions of others and being affected by the decisions of others. The selection process is a perpetual activity over a four-stage process that starts with the observation of past returns of one's portfolio, continues through an agent selection and a comparison of the two portfolios, and ends in a decision.[1] The selection procedure reflects a sort of the reinforcement learning in which agents tend to adopt portfolios that have yielded high returns in the past [31]. We extend this concept by introducing agents' imperfect choices and denote it the level of suspiciousness. An agent is said to be suspicious if there is a strictly positive probability that he would not adopt a portfolio with a higher return. The description implies that suspicious agents are prone to making poor decisions, be they intentional or accidental. Such a behaviour could be explained by several agent-specific features, representing the behavioural aspect of the portfolio choice problem (see Hirshleifer [32] for a discussion). Finally, a small proportion of agents never change their

[1] Information diffuses over the network by the word-of-mouth. Communications-based models in economics often employ the word-of-mouth communication as a source for idea diffusion [23–30].

initial portfolio and they are referred to as liquidity agents. Therefore, the problem of a portfolio selection is not just technical but complex.

The model is simulated on real data. Dataset covers daily returns for stocks of five companies listed on the Dow Jones Industrial Index, from which agents can make portfolios. We consider three different time periods, the baseline with no specific trend, the bear and the bull market. The latter two are motivated by the work of Kahneman and Tversky [33, 34], which argues that agents behave differently in the domain of losses than in the domain of gains. In addition, Barberis et al. [35] argued that agents are less prone to taking risk in a bear market, as they first start to recognize and then also evaluate the risk.

The selected portfolios are analyzed in the context of the efficient frontier theory. We do not intend to challenge the theory as such. Instead, we are interested in whether the interacting agents can attain mean-variance portfolios in an uncertain and dynamic environment. Following the series of simulation-based experiments, we show that they can. Even though agents constantly make decisions upon the realized returns, we find the risk to be a decisive factor in portfolio selection within both cohorts with the lowest-risk mean-variance portfolios being highly preferred. Further, agents tend to synchronize their selections and distinguish the winners from the losers. Selections of suspicious agents are slightly more dispersed than that of unsuspicious, while the two cohorts mostly identify the same winners. This conclusion is supported in both bull and bear markets. However, we find that agents are much more susceptible to risk in a bear trend and extremely synchronize their choices across the least risky portfolios. A slight deviation from the efficient frontier portfolios can be perceived in a bull market, while the level of synchronization is weaker than in the bear market. Further, we use the Monte Carlo (MC) method and the coefficient of variation (CV) to examine the consistency in agents' selections and find that agents behave the most consistently on the most desired or the least desired choices. Unsuspicious agents are more consistent in their selections than the suspicious. The comparison of selections of the two cohorts depicts the sensitivity of a portfolio selection to the selection pattern, which may be substantial.

The chapter contributes to the growing field of agent-based models in economics and finance (see Tesfatsion and Judd [36] and Steinbacher et al. [37] for an overview of agent-based models in economics and finance). The very brief list includes: Zeeman [38], Kim and Markowitz [39], Brock and Hommes [40, 41], Lux [42], Lux and Marchesi [43], Boswijk et al. [44], Levy et al. [45], Cont and Bouchaud [46], Iori [47], Palmer et al. [48], Arthur [49] and LeBaron et al. [50]. Early attempts to use networks in economics were due to Myerson [51] and Kirman [52, 53].

The remainder of the present chapter is organized as follows. In Sect. 2, we construct the model. Results are described in Sect. 3, while consistency in portfolio selection is examined in Sect. 4. The chapter concludes with a discussion and concluding comments.

2 The Model

Agents

The model is a discrete-time and discrete-state model. It consists of a set of agents, who are represented by the nodes and connected to one another with undirected links into a network similar to that proposed by Watts and Strogatz [54].[2] Such a network is referred to as a small-world network in which each agent is connected only to a small fraction of counterparts; mostly to the closest neighbours and also to some distant ones. Hence, agents have much stronger local contacts than global. A schematic representation of a small-world network with 30 nodes is illustrated in Fig. 9.1.

Agents use the network as an infrastructure for communication with adjacent counterparts with whom they exchange portfolio-related information. Although only connected agents can communicate with one another, the chains of links allow the agents to receive indirect information from those with whom they are not directly linked. For instance, if we look at the figure and assume that agent 1 communicates with agent 29 and adopts his portfolio and then agent 2 communicates with agent 1 and adopts his portfolio, agent 2 has actually adopted portfolio of agent 29, even though they are not directly connected. The actions of each agent thus influence the others, whether directly or indirectly.

The present model consists of $n = 5000$ agents. Each agent is initially linked with the 10 closest neighbours ($k = 10$), five on each side of a ring lattice. Each link is then rewired to the arbitrarily chosen node with probability $p = 0.1$, which may give the agents some long-range connections. The resulting network remains locally connected and has a small diameter.[3] By assumption, the network is static. Therefore, agents are not allowed to form new connections over time nor severe the existing connections. Each game starts with a unique network.

[2] By definition, a mutual consent is required to establish an undirected link. Therefore, if agent i is connected with agent j, then agent j is also connected with agent i. This is not the case in directed networks in which one of the two agents is not aware of the link. An extensive review of social networks and network models is given in Wasserman and Faust [55], Boccaletti et al. [56], Goyal [57] and Jackson [58].

[3] The diameter of a network is the largest distance between any two nodes on the network.

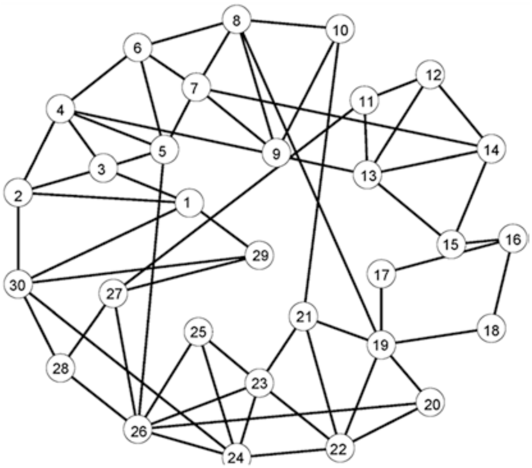


Fig. 9.1 Schematic representation of a network

At the start of each game, each agent is assigned an initial portfolio at random to the amount of $W_0 = 1$.[4] The principal objective of each agent then is to accumulate wealth over time.

$$
\begin{aligned}
&W_{t+1}(A_i) = W_t(A_i)\left[1 + R_t(\bullet)\right] \\
&W_0\ (A_i)\ = 1;\ q_t^i \geq 0
\end{aligned}
\tag{9.1}
$$

W_t and W_{t+1} represent the wealth of an agent i in time intervals t and $t+1$, and $R_t(\bullet)$ denotes the returns of the alternative $(\bullet)$ used by an agent in t. We assume that the entire return on portfolio is translated into agents' wealth and reinvested. Short sales or borrowings are not allowed for which $q_t^i \geq 0$. Agents are also not allowed to purchase costly information signals. The simplicity of the objective function illustrates the apparent simplicity of the problem that agents face, that is, in every time period to select a portfolio according to the following four stage procedure (Fig. 9.2).

Agents meet in a sequential fashion. Each period t starts with agent 1, who observes $t-1$ value of his portfolio. Each agent only has information about a portfolio he holds. Stage 1 is then followed by stage 2 in which an agent picks one of his adjacent agents at random. By random selection we assume that none is capable of beating the market systematically thus acknowledging that agents have no forecasting rules. After a counterpart is selected, the two compare

[4] By referring to games, we do not have the usual game theoretical framework in mind but computer-based experiments. Alternatively, the games on networks could also be referred to as the activities on networks.

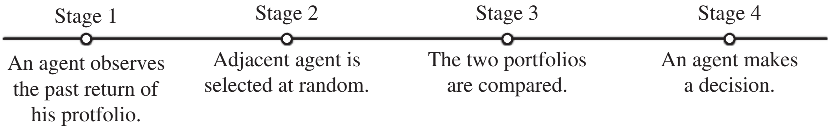


Fig. 9.2 Decision-making process

their portfolios. Specifically, the two compare their $t-1$ portfolio structures and the corresponding values. In the last stage, each of the two agents decides whether to continue with his current portfolio or to switch to the counterpart's. Contrary to the complex decision rules that can be found in many other agent-based models, we assume that each of the two agents, i and j, compares the two payoffs and makes a decision according to the following logistic function:

$$\Phi = \frac{1}{1+\exp\left[\frac{W_{t-1}\left(A_i\right)-W_{t-1}\left(A_j\right)}{\kappa}\right]} \tag{9.2}$$

The selection rule incorporates the idea that agents tend to adopt portfolios that have performed well in the past [31, 59, 60].[5] As Eq. (9.2) implies, agents' selections are bounded by the suspiciousness parameter $\kappa \in (0, 1)$.[6] The parameter is an exogenous factor that affects the probability that an agent does not choose the better of the two portfolios he compares. To execute the selection, we use a random number generator that is implemented as follows: if $ran > \Phi$, an agent keeps his portfolio, otherwise an agent adopts the portfolio of adjacent agent. Parameter $ran \sim U(0, 1)$ is a uniformly distributed IID (independent and identically distributed) random number [64, 65]. The probability that a portfolio with a smaller return is selected depends negatively upon the difference in the compared payoffs $W_{t-1}(A_i) - W_{t-1}(A_j)$ and positively upon the parameter κ. The lower the κ the higher the probability that an agent

[5] Strictly speaking, reinforcement learning is not learning as we know, but imitation. It has been extensively used in the 'cheap-talk' communication games of social learning to describe the result of some types of 'emulation' by economic agents.

[6] The rationale for including the suspiciousness parameter is manifold. Selten [61] argues that errors may arise between the decision to take a certain action and the action itself. The level of suspiciousness may also relate to a (dis)trust, suspiciousness of the data, temporary mood, or some other agent-specific factors, even luck, etc. In addition, the switch to the alternative with small potential benefits may not be preferred. The agents' unwillingness to do a switch is even more likely if transaction costs are present. Agents may also fail to have persistent preferences and rather use heuristics before acting, which may again lead to 'errors' [62]. Thaler et al. [63] argue that short evaluation periods force investors to make poorer decisions. Contributions of these factors are very indeterminate, but cannot be neglected. The variable thus includes all the factors for which an investor might behave differently than expected. From a technical perspective, it may also be considered a noise.

adopts the portfolio with a higher return, and vice versa. By definition, the fully unsuspicious agent has $\kappa = 0$ and fully suspicious has $\kappa = 1$.[7]

After each of the two agents updates his portfolio, the process repeats for agent 2, and so on until the last agent. Following the decision of the last agent in a given period, stock returns for the given period are reported and the system proceeds to the next period and repeats until $t = T$. It must be emphasized that although agents interact with one another, they make decisions simultaneously and autonomously without knowing what others have selected. Hence, an agent cannot be forced to adopt a certain decision. Agents also do not play against each other.

Proposition 1: *An unsuspicious agent shows absolute preference for portfolios with a higher expected return.*

Proof: Let $\Phi : R \to R$ s.t. $\Phi(A) = [1 + \exp[(A_i - A_j)\kappa^{-1}]]^{-1}$; $A_i, A_j \in R$ and $\kappa \in (0, 1)$. If we denote $ran \sim U(0, 1)$ simply as U, we can provide f as:

$$f = \begin{cases} A_i & U > \Phi \\ A_j & U < \Phi \\ 0 & U = \Phi \end{cases}$$

As $\kappa \downarrow 0$ and $A_i > A_j$ there exists $\lim_{\kappa \to 0}(\Phi) = 0 = \inf(\Phi)$. Say we pick $\varepsilon \in U$ s.t. $\varepsilon \le \Phi$. It follows that $\Pr(\varepsilon \le \Phi) = \int_0^{\varepsilon \le 0} du = 0 \to \Pr(\varepsilon > \Phi) = 1$ and $f(\varepsilon > \Phi, \Phi) = A_i$ for all $A_i > A_j$. Here, an unsuspicious agent will always choose A_i.

As $\kappa \downarrow 0$ and $A_i < A_j$ there exists $\lim_{\kappa \to 0}(\Phi) = 1 = \sup(\Phi)$. However, we can always pick $\delta \in U$ s.t. $\delta \ge \Phi$. It follows that $\Pr(\delta \ge \Phi) = \int_0^{\delta \ge \Phi} du = 0 \to \Pr(\delta < \Phi) = 1$ and $f(\delta < \Phi, \Phi) = A_j$ for all $A_i < A_j$ and unsuspicious agent will again choose a better of the two alternatives, which is A_j.

The level of suspiciousness κ does not affect the selection of two equally profitable portfolios as the choice between the two is made at random: $\Phi = 0.5$ for each $\kappa \in [0, 1]$ and $A_i = A_j$. If we pick $\beta \in U$, s.t.

[7] In order to make the proofs more tractable, we use A_i and A_j to designate agents' wealth; $W_t(A_i)$ and $W_t(A_j)$, respectively.

$\beta =\ \Phi$, it follows that $\Pr(\beta = \Phi = 0.5) = \int_0^{0.5} du = 0.5$. The result shows up also in a close proximity of the two portfolios as $\lim_{A_j \to A_i} (\Phi) = 0.5$. □

Proposition 2: *Unsuspicious and suspicious agents randomly choose among portfolios with equal expected returns.*

Proof: See the last part of the prior proof. □

Proposition 3: *A strictly suspicious agent shows a relative preference for portfolios with a higher expected return.*

Proof: We keep the setting from the proof of Proposition 1. By definition, a fully suspicious agent has $\kappa = 1$.

As per the mapping *f* from the previous proof, a suspicious agent holding A_i will start showing a relative preference to changing to A_j once $\Phi\ > 0.5$, which is in turn true for all $A_i < A_j$. Say we pick $\varepsilon \in U$ s.t. $\varepsilon <\ \Phi$. Probability that a suspicious agent switches to a better of the two alternatives is then $\Pr(\varepsilon > 0.5) = \int_0^{\varepsilon > 0.5} du = \varepsilon > 0.5 \to \varepsilon \in (0.5,1)$, true for all $A_i < A_j$. The probability that an agent remains with A_i is $q = 1 - \Pr \to q < \Pr$, showing a relative preference to a better of the two alternatives.

For $A_i > A_j$ an agent shows a relative preference to keeping a better of the two alternatives once $\Phi\ < 0.5$, which is in turn true for all $A_i > A_j$. Say we pick $\delta \in U$ s.t. $\delta >\ \Phi$. Probability that a suspicious agent remains with a better of the two alternatives is then $q(\delta > 0.5) = \int_0^{\delta > 0.5} du = \delta > 0.5 \to \delta \in (0.5,1)$, true for all $A_i > A_j$. However, the probability that an agent adopts A_j is $\Pr = 1 - q \to \Pr < q$, showing again a relative preference to a better of the two alternatives. In both cases, an agent's relative preference to a better of the two alternatives increases exponentially in return differential of the two, never falls below 0.5 and never reaches unity. □

Figure 9.3 plots the probability that a better of the two portfolios is selected as a function of the difference in the two alternatives (spread) and the suspiciousness level. Spread is provided in basis points.[8]

[8] 100 basis points = 1 percentage point.

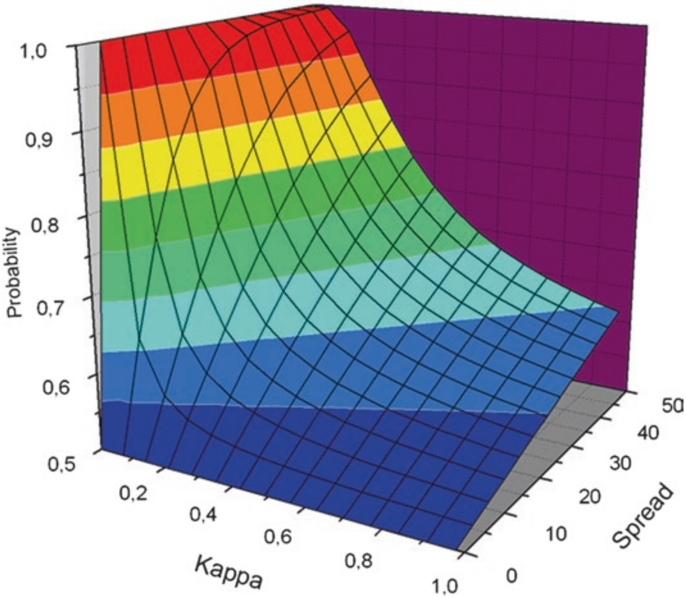


Fig. 9.3 Probability that the better of the two portfolios is selected as a function of the spread and the suspiciousness parameter

Interaction-based games in which agents share information with one another induce herding. Herding is one of the unavoidable consequences of imitation and among the most generally recognized observations in financial markets. It promotes synchronization. The side-effect of synchronization may be that portfolios with only occasionally unfavourable returns are eliminated from the sample set. Recall that only portfolios that are possessed by one or more agents are feasible. Bala and Goyal [23] show that when agents prefer higher payoffs and learn from their neighbours, they synchronize their selections with probability one, regardless of their initial states. Unsuspicious agents who perfectly rebalance their portfolios should be particularly prone to synchronization. To avoid such an outcome, we include liquidity agents. A liquidity agent refers to an agent who keeps the initial portfolio irrespective of the returns.[9] Liquidity agents may illustrate passive investors who participate in the market but are unwilling to change their portfolios, or even extremely loyal investors [66]. We call them liquidity agents because they keep the liquidity of each portfolio. Equipped with this feature, they assure that the selection process works smoothly. Liquidity agents occupy a small fraction of agents and are placed into several homogeneous groups,

[9] The initial portfolio is randomly assigned to each liquidity agent in the same way as it is to all the rest.

numbered from 700 to 719, from 1000 to 1019, from 1200 to 1219, from 1500 to 1519, from 2500 to 2519, from 3500 to 3519, and from 4800 to 4819. As the games repeat, liquidity agents might possess different portfolios and be adjacent to different agents.

Securities and Portfolios

A portfolio is a set of investments. Units of assets might include savings accounts, equities, bonds and other securities, debt and loans, options and derivatives, ETFs, currencies, real estate, commodities.

If $M = (1, 2, \ldots, n)$ represents a finite set of units of assets, then a portfolio is composed of the holdings of these n units of assets. If the number of non-zero holdings is larger than 1, the portfolio is called a diversified portfolio. We exclude short sales, which means that agents can only possess a non-negative quantity of different assets, hence $q_t^j \geq 0$. Diversified portfolios can only include an equal proportion of i assets from the total number of n available assets, which makes a maximum number of $K(n) = \sum_{i=1}^{n} \binom{n}{r} = 2^n - 1$ different portfolios.[10]

We use stocks of five companies from different sectors listed on the Dow Jones Industrial Index: C, KFT, MSFT, AA and XOM.[11] Hence, $n = 5$, which makes a sample set of 31 portfolios (Table 9.1).

Agents are price takers. All stocks are infinitesimally divisible and liquid from which agents can buy or sell any quantity of stocks quickly with no price impacts.[12] Without this assumption, agents with the worst-performing portfolio after the first iteration would need to bring new money into the game if they wanted to buy any other portfolio. Because this is not possible, we would thus confine these agents' selection activities. When an agent switches to a portfolio of an adjacent agent, an implicit assumption is made that he sells all stocks of the current portfolio and buys a mix of stocks in the portfolio of an adjacent agent. For simplicity, we further assume that portfolio transformation induces no transaction costs or any other trade-related costs, including taxes.[13]

[10] This is a simplification. In reality, agents can put a different proportion of assets into their portfolios, thus the number of different portfolios approaches infinity.

[11] Effective as of June 8, 2009, Citigroup (NYSE: C) was delisted from the index and was replaced by its sister insurance company Travelers (NYSE: TRV).

[12] Illiquidity has been found to have substantial effects on stock markets [67–69].

[13] This assumption is limited by the effects such costs might have on the trading policy. Constantinides [70] and Lo et al. [71] argue that agents accommodate large transaction costs by reducing the frequency

Table 9.1 Description of portfolios

S1	AA	S12	MSFT-KFT	S23	MSFT-XOM-KFT
S2	MSFT	S13	XOM-C	S24	MSFT-C-KFT
S3	XOM	S14	XOM-KFT	S25	XOM-C-KFT
S4	C	S15	C-KFT	S26	AA-MSFT-XOM-C
S5	KFT	S16	AA-MSFT-XOM	S27	AA-MSFT-XOM-KFT
S6	AA-MSFT	S17	AA-MSFT-C	S28	AA-MSFT-C-KFT
S7	AA-XOM	S18	AA-MSFT-KFT	S29	AA-XOM-C-KFT
S8	AA-C	S19	AA-XOM-C	S30	MSFT-XOM-C-KFT
S9	AA-KFT	S20	AA-XOM-MSFT	S31	AA-MSFT-XOM-C-KFT
S10	MSFT-XOM	S21	AA-XOM-KFT		
S11	MSFT-C	S22	MSFT-XOM-C		

Data

The real data for stock returns is used. Daily returns are calculated as the relative difference between the closing prices of two consecutive days. The return of a portfolio in time t is given as the weighted sum of the stock returns of the portfolio, hence $R_t^S = \sum_{j=1}^{n} q_t^j R_t^j$, with $\sum_{j=1}^{n} q_t^j = 1$.

The model is run in three different time spans. The baseline framework covers the period of 2 January 2009 to 21 January 2010. The bear market started on 22 September 2008 and ended on 13 March 2009. The bull market succeeded the bear, starting on 16 March 2009 and ending on 11 January 2010.There were two non-trading days: 14 and 15 March 2009. The baseline framework consists of 264 intervals, the bear market includes 120 intervals and the bull market 209.

Table 9.2 reports cumulative returns for each stock within both sub-periods. As the table shows, the worst investment would have been a dollar invested in C at the beginning of the bear market as it ended in only 8.90 cents. In the bull market, a dollar invested in AA produced a yield of a dollar and 85.13 cents. XOM was just slightly shaped by the bear market but did not exhibit any large positive move in the bull market, either. It would lead to only a 14.81 % loss in the bear market and would produce less than a 5 % yield in the bull market.

Next, we shall characterize the risk of a portfolio. Usually, the portfolio beta is a measure of portfolio risk. For each portfolio i, its beta is estimated from the market model:

$$\begin{gathered} R_{i,t} = \alpha_i + \beta_i R_{M,t} + \varepsilon_{i,t}, \text{ with} \\ E\left(\varepsilon_{i,t}\right) = 0;\ Var\left(\varepsilon_{i,t}\right) = \sigma_i^2 \end{gathered} \tag{9.3}$$

and volume of trade. Moreover, transaction costs also direct the choice towards riskless securities and determine the number of securities in a portfolio.

Table 9.2 Cumulative returns in a bear and a bull market (in %)

	AA	C	KFT	MSFT	XOM
Bear	−78.62	−91.10	−31.91	−34.45	−14.81
Bull	185.13	55.79	25.93	86.28	4.79

A β coefficient reflects portfolio risk against market risk, while *R*-squared measures the degree of the portfolio diversification in relation to the market portfolio. The difference up to unity is the portfolio idiosyncratic risk that could have been diversified. In our case, portfolio S31 represents a market portfolio because it consists of all available stocks and represents the highest possible degree of diversification.[14] Hence, variance of S31 represents the market risk σ_M^2 . Table 9.3 reports betas for the portfolios and the corresponding *R*-squared values.

3 Simulation Results

We simulate the games under two different assumptions, first with unsuspicious agents ($\kappa = 0.01$) and then with suspicious agents ($\kappa = 0.1$). Although we refer to suspicious agents, they do not make blind guesses but retain some capability to select a better of the two alternatives which they compare, particularly if they are sufficiently apart. For instance, with $\kappa = 0.1$ and a 10-basis-points spread in return between two portfolios, the probability that the one with the lower return is selected is 26.9 %; with the spread of 25-basis points, the corresponding probability is 7.6 %.

Each simulation run is repeated for 30 times and the average results are reported.[15] Two types of results are reported. The average endgame results (hereafter: endgame) present the average proportion of agents per portfolio of 30 independent repetitions in the last time period. The average average-game results (hereafter: average-game) present the average proportion of agents per portfolio over all time periods and over all 30 repetitions. The average-game results provide information on the desirability of individual portfolios throughout game developments, while endgame decisions indicate the desirability of single portfolios in the end and hence present the convergence patterns. Let us stress that we do not presume equilibrium, but do not exclude it, either.

[14] The choice for a market portfolio would designate a naïve $1/n$ allocation [72].

[15] Fragmented pseudo-code is provided in the Appendix.

Table 9.3 Beta coefficients for the portfolios and the corresponding *R*-squared values in the three sub-periods

	Baseline		Bear		Bull	
	Beta	*R*-squared	Beta	*R*-squared	Beta	*R*-squared
S1	1.246	0.574	1.205	0.680	1.461	0.540
S2	0.551	0.431	0.711	0.650	0.584	0.352
S3	0.413	0.490	0.724	0.615	0.439	0.385
S4	2.475	0.753	1.907	0.651	2.167	0.641
S5	0.314	0.264	0.453	0.552	0.530	0.241
S6	0.899	0.681	0.958	0.794	1.022	0.655
S7	0.830	0.670	0.965	0.764	0.950	0.647
S8	1.861	0.936	1.556	0.893	1.814	0.902
S9	0.780	0.642	0.829	0.769	0.905	0.624
S10	0.482	0.587	0.718	0.706	0.511	0.497
S11	1.513	0.848	1.309	0.833	1.378	0.775
S12	0.433	0.551	0.582	0.733	0.467	0.484
S13	1.444	0.835	1.316	0.858	1.303	0.749
S14	0.364	0.499	0.588	0.706	0.394	0.448
S15	1.395	0.813	1.180	0.769	1.258	0.719
S16	0.737	0.732	0.880	0.807	0.828	0.713
S17	1.424	0.972	1.274	0.962	1.404	0.959
S18	0.704	0.729	0.790	0.830	0.798	0.715
S19	1.378	0.965	1.279	0.968	1.355	0.947
S20	0.658	0.703	0.794	0.805	0.750	0.700
S21	1.345	0.961	1.188	0.937	1.326	0.937
S22	1.147	0.897	1.114	0.931	1.063	0.837
S23	0.426	0.633	0.629	0.754	0.458	0.568
S24	1.113	0.891	1.024	0.892	1.034	0.830
S25	1.068	0.871	1.028	0.909	0.985	0.799
S26	1.171	0.988	1.137	0.992	1.163	0.982
S27	0.631	0.760	0.773	0.831	0.708	0.753
S28	1.147	0.992	1.069	0.982	1.140	0.985
S29	1.112	0.980	1.072	0.985	1.104	0.969
S30	0.938	0.924	0.949	0.955	0.885	0.873
S31	1.000	1.000	1.000	1.000	1.000	1.000

Baseline Framework

As demonstrated in Fig. 9.4, the average-game decisions of unsuspicious agents are highly concentrated around the minimum variance point. Portfolios chosen by more than 5 % of agents are marked with grey squares. Black triangles in the figure designate portfolios chosen by less than 1.5 % of agents; black circles designate portfolios chosen by more than 1.5 % but less than 5 % of agents. The curve depicts the efficient frontier and is fictitious. The figure is a mean-standard deviation diagram.

The figure demonstrates that agents select the efficient low-risk portfolios that are clustered in the neighbourhood of the minimum-variance point. The

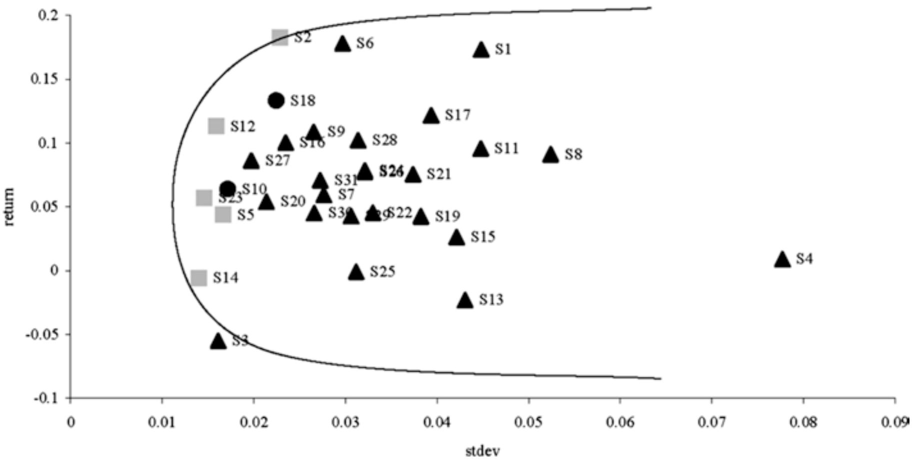


Fig. 9.4 Fraction of unsuspicious agents per portfolio in an average-game setting

most desired portfolio is a single-asset portfolio S5, chosen by 29.48 % of unsuspicious agents. It is followed by S12 (27.18 %), S23 (11.07 %), S2 (8.78 %) and S14 with an average share of 8.24 % of unsuspicious agents. The most desired portfolios are highly under-diversified with large idiosyncratic components. S12 is a two-asset portfolio made of the two most-desired single stocks S5 and S2. Interestingly, S23 is a three-asset portfolio of S2, S5 and S3; S3 is the lowest-return asset and among the least-desired alternatives. Some further results are given in Table 9.4.

S14 is the minimum variance portfolio, while S2 is the highest mean portfolio. S3 is the lowest mean portfolio and S4 is the riskiest portfolio, while both were strictly avoided in the average-game setting. The two most desired portfolios were chosen by 56.66 % of unsuspicious agents and the first five by 84.75 % of all unsuspicious agents. Only liquidity agents selected portfolio S4.

The average-game decisions of suspicious agents slightly differ from that of unsuspicious. If the transition from the most- to the least-desired portfolios was almost discrete within the unsuspicious agents, it goes through the group of relatively desired portfolios within the setting of suspicious agents, as indicated by the dashes shown in Fig. 9.5. Suspicious agents do not synchronize their selections as much as the unsuspicious, hence the number of relatively desired portfolios is accordingly higher in this setting and the distribution much more even than in the setting of unsuspicious agents. The most desired portfolios of suspicious agents are S2 (9.87 %), S12 (7.94 %), S23 (6.47 %), S5 (6.40 %) and S10 (5.73 %), all of which are efficient frontier portfolios.

Table 9.4 Fractions of unsuspicious (US) and suspicious (S) agents per portfolio in the baseline framework

	Average-game		Endgame	
	US	S	US	S
S1	0.18	1.39	0.28	2.28
S2	8.78	9.87	14.49	15.07
S3	1.07	3.58	0.32	0.92
S4	0.17	0.25	0.08	0.17
S5	29.48	6.40	22.02	3.19
S6	0.46	4.15	0.81	7.36
S7	0.19	2.09	0.18	2.07
S8	0.15	0.49	0.11	0.46
S9	0.53	3.78	0.59	4.91
S10	3.60	5.73	3.33	4.11
S11	0.23	1.37	0.11	1.51
S12	27.18	7.94	36.64	7.68
S13	0.18	0.82	0.10	0.59
S14	8.24	5.14	3.33	2.06
S15	0.33	0.97	0.10	0.68
S16	0.49	4.15	0.64	4.96
S17	0.20	1.14	0.21	1.68
S18	1.61	4.80	2.24	6.53
S19	0.15	1.38	0.11	1.48
S20	0.75	3.41	0.66	2.94
S21	0.19	1.51	0.13	1.68
S22	0.24	2.19	0.12	1.97
S23	11.07	6.47	9.91	4.64
S24	0.59	2.85	0.30	2.98
S25	0.39	1.73	0.13	1.09
S26	0.20	1.93	0.18	2.29
S27	1.42	4.73	1.58	4.60
S28	0.28	2.23	0.22	2.83
S29	0.24	1.66	0.14	1.38
S30	1.08	2.84	0.69	2.35
S31	0.33	3.00	0.23	3.53

Although suspicious agents are willing to bear greater risk than unsuspicious agents, they avoid the riskiest portfolios. Desirability of portfolios is decreasing with the level of risk, which is the same as in the case of unsuspicious agents.

The endgame results of unsuspicious agents exhibit synchronization patterns. Unsuspicious agents mostly ended with portfolios S12 (36.64 %) and S5 (22.02 %), which were followed by S2 (14.49 %) and S23 (9.91 %). S12 is a two-asset portfolio of S5 and S2, and S23 is a three-asset portfolio of S2, S5 and S3. These are low-risk portfolios from

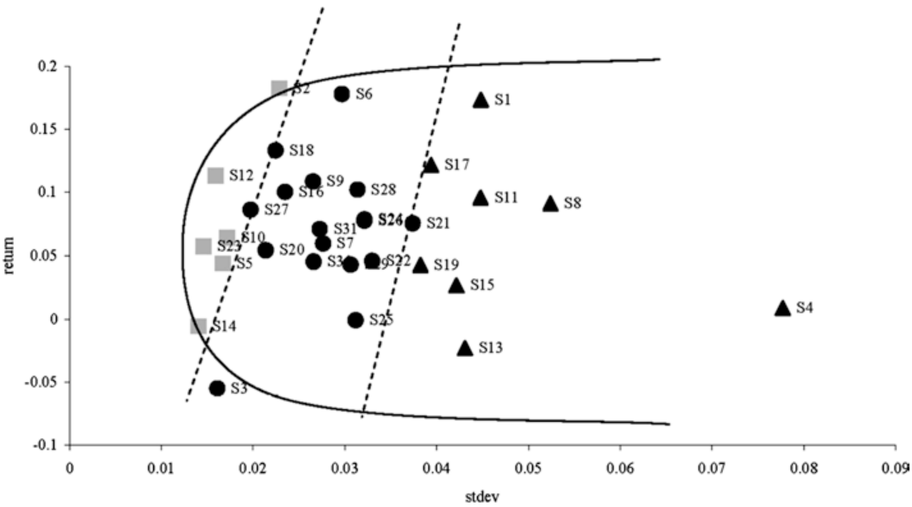


Fig. 9.5 Fraction of suspicious agents per portfolio in an average-game setting

the efficient frontier. There are plenty of high-risk and unprofitable portfolios that were strictly avoided: S4 (0.08 %), S13 and S15 (0.10 %), and S8, S11 and S19 (0.11 %). Except for S4 and S19 (a three-asset portfolio of S1, S3 and S4), the rest are two-asset portfolios that include the riskiest stock S4. Because of the riskiness of S4, the inclusion of an additional stock cannot reduce the risk sufficiently to make diversified portfolios more desirable. The two leading stocks are S5 and S2. The top five portfolios were chosen by 86.4 % of all unsuspicious agents. In the endgame setting of unsuspicious agent (see Fig. 9.6), the transition from the most- to the least-desired portfolios is very straight with only four portfolios in-between. These are portfolios S10 (3.33 %), S14 (3.33 %), S18 (2.24 %) and S27 (1.58 %).

The results of the runs with suspicious agents show that they also prefer less risky portfolios (Fig. 9.7). Again, the transition from the most- to the least-desired portfolios develops in stages as indicated by the dashes. The group of the most-desired portfolios consists of the high-return and low-risk portfolios: S2 (15.07 %), S12 (7.68 %), S6 (7.36 %) and S18 (6.53 %). The vast majority of the portfolios belong to the second group of moderate-to-high return and moderate-to-high-risk portfolios. The third group consists of the least-desired portfolios: S4 (0.17 %), S13 (0.59 %) and S15 (0.68 %), along with the portfolio S8 (0.46 %), which are low-return and moderate-to-high-risk portfolios. This group also includes S3, which is the lowest-return and (almost) the lowest-risk portfolio. The second most

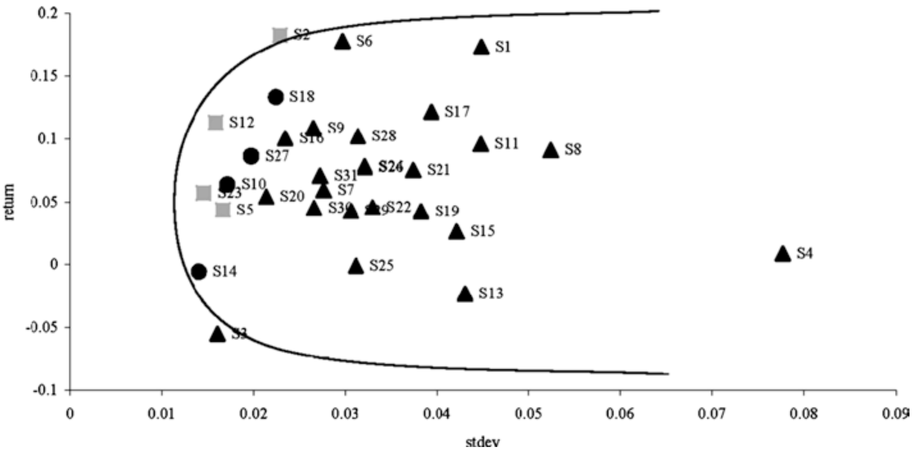


Fig. 9.6 Fraction of unsuspicious agents per portfolio in the endgame setting

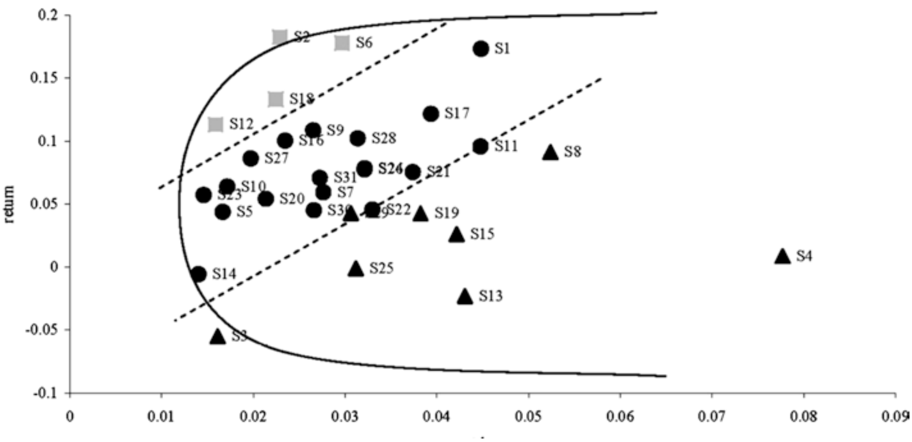


Fig. 9.7 Fraction of suspicious agents per portfolio in the endgame setting

desired portfolio of unsuspicious agents S5 was placed in the second group of moderate-to-high return and moderate-to-high-risk portfolios by suspicious agents.

Portfolio Selection in a Bear Market

The second time span relates to the bear trend. The bear market reflects a general fall in prices. In our case, the bear trend is characterized by large standard deviations and some highly negative mean returns.

Table 9.5 reports the fractions of unsuspicious and suspicious agents per portfolio in the average-game setting. The results show that both cohorts are

Table 9.5 Fractions of unsuspicious (US) and suspicious (S) agents per portfolio in the bear market

	Average-game		Endgame	
	US	S	US	S
S1	0.22	0.30	0.09	0.11
S2	2.99	4.76	2.47	4.16
S3	41.97	23.20	62.19	37.74
S4	2.11	1.32	0.08	0.09
S5	0.42	3.81	0.38	4.04
S6	0.23	0.58	0.10	0.16
S7	0.26	0.89	0.09	0.30
S8	0.31	0.47	0.09	0.09
S9	0.18	0.57	0.09	0.23
S10	14.22	11.50	17.90	15.74
S11	2.37	1.82	0.10	0.10
S12	1.21	4.55	1.11	4.83
S13	6.10	2.80	0.10	0.24
S14	5.72	10.28	7.14	14.75
S15	0.92	1.33	0.09	0.10
S16	0.31	1.72	0.14	1.12
S17	0.34	0.62	0.09	0.11
S18	0.21	0.89	0.10	0.48
S19	0.38	0.92	0.10	0.14
S20	0.26	1.51	0.14	0.89
S21	0.27	0.64	0.10	0.10
S22	5.22	3.22	0.21	0.59
S23	5.06	8.37	5.97	10.29
S24	1.23	1.82	0.09	0.19
S25	2.81	2.83	0.14	0.48
S26	0.42	1.16	0.09	0.23
S27	0.31	1.68	0.18	1.08
S28	0.28	0.83	0.07	0.13
S29	0.32	1.01	0.08	0.22
S30	2.97	3.26	0.37	0.94
S31	0.40	1.35	0.11	0.32

capable of selecting less risky portfolios and avoiding highly risky portfolios of S4 and S1. Almost 42 % of all unsuspicious agents possess the portfolio with the highest average return, S3, followed by S10 (14.22 %), S14 (5.72 %), and S23 (5.06 %) that also lie on the efficient frontier or in its closest neighbourhood, and S13 (6.10 %) and S22 (5.22 %) which do not. These later two could signify risk seeking in choices when prospects are negative when agents try to obtain profits on the variance [33, 73]. Risk seeking could also be noticed when comparing portfolio S13 to S28. Both portfolios have similar means, while S13 is riskier. S13 was chosen by 6.10 % of unsuspicious and 2.80 % of suspicious agents, while only liquidity agents held S28 in both cases. Interestingly, S4, a portfolio with the smallest mean and the highest variance, for which it could be considered the worst alternative, was chosen by 2.11 % of unsuspicious and 1.32 % of suspicious agents. The selected portfolios of suspicious agents were more evenly distributed than those of unsuspicious agents, with S3 (23.20 %), S10 (11.50 %), S14 (10.28 %) and S23 (8.37 %) from the efficient frontier still being the most desired.

The endgame results show that synchronization is highly present in unsuspicious agents' selections as only four of 31 alternatives are chosen by more than 5 % of agents. These are S3 (62.19 %), S10 (17.90 %), S14 (7.14 %) and S23 (5.97 %), and together account for 93.20 % of all selections by unsuspicious agents. They all are portfolios from the efficient frontier and have S3 as the leading stock. Of the remaining portfolios, only S2 (2.47 %) and S12 (1.11 %) do not end with liquidity agents. The picture is not much different for the suspicious agents, where the synchronization pattern is also present. The most desired portfolios are similar to that of the unsuspicious agents; S3 (37.74 %), S10 (15.74 %), S14 (14.75 %) and S23 (10.29 %).

In the bear market, agents concentrate their endgame decisions on the low-risk and high-return portfolios.

Portfolio Selection in a Bull Market

In contrast, a bull trend is characterized by positive shifts in returns. S1 and S2 exhibit the highest mean return of 0.50 % and 0.30 %, respectively, S3 the smallest of 0.023 %. S4 is the riskiest with a standard deviation of 8.65 %, while S3 and S5 have the smallest standard deviations of 7.05 %.

As reported in Table 9.6, the most-desired portfolios of unsuspicious agents are S8 (41.70 %), S17 (13.70 %), S4 (11.05 %), S21 (7.37 %), S19 (6.27 %) and S1 (3.36 %). Of these, only S1 is from the efficient frontier, while S17 is very close to it. S8 clearly lies outside the efficient frontier; it is riskier than S6. In addition, S8 is a multiple-stock portfolio. Generally, agents select riskier portfolios in the bull market. The most-desired endgame decisions of the unsuspicious agents are S8 (50.11 %), S17 (17.37 %), S21 (6.37 %), S1 (5.66 %) and S6 (2.25 %) together accounting for 84.83 % of all their selections. From these only S1 and S6 lie on the efficient frontier, with S17 being very close to it.

Suspicious agents are also capable of synchronization, although to a lesser extent. However, if the selection of five winning portfolios is nearly the same as within the unsuspicious agents, there are differences in the degree of desirability of each. The five most desired endgame portfolios of suspicious agents are S1 (24.59 %), S8 (16.77 %), S6 (9.52 %), S17 (9.16 %) and S21 (4.36 %), which account for 64.38 % of all selections. This is far less than within the unsuspicious agents. For example, the winning portfolio of suspicious agents, S1, was selected by less than 6 percent of unsuspicious agents in the endgame setting.

Table 9.6 Fractions of unsuspicious (US) and suspicious (S) agents per portfolio in the bull market

	Average-game		Endgame	
	US	S	US	S
S1	3.36	14.91	5.66	24.59
S2	0.13	1.59	0.13	1.79
S3	0.19	0.38	0.11	0.10
S4	11.05	6.11	0.10	3.35
S5	0.17	0.39	0.25	0.15
S6	1.58	6.30	2.25	9.52
S7	0.54	2.20	0.29	1.54
S8	41.70	14.75	50.11	16.77
S9	0.43	3.29	0.43	3.49
S10	0.13	0.75	0.09	0.34
S11	2.83	3.60	2.17	2.77
S12	0.14	0.70	0.09	0.32
S13	0.89	1.51	0.11	0.57
S14	0.17	0.38	0.09	0.12
S15	0.76	1.93	0.17	0.86
S16	0.37	2.18	0.28	1.78
S17	13.70	7.93	17.37	9.16
S18	0.31	2.67	0.26	2.91
S19	6.27	4.15	4.00	2.75
S20	0.22	1.30	0.12	0.69
S21	7.37	4.89	6.37	4.36
S22	0.46	1.36	0.16	0.61
S23	0.15	0.61	0.09	0.26
S24	0.46	1.80	0.20	1.24
S25	0.30	1.05	0.08	0.35
S26	2.09	3.19	1.29	2.43
S27	0.21	1.11	0.13	0.68
S28	2.02	3.61	1.72	3.14
S29	1.13	2.19	0.44	1.40
S30	0.24	0.99	0.09	0.40
S31	0.65	2.18	0.30	1.57

4 Consistency in Selection

It remains to examine how consistent agents are in their selections. Stochastic interaction-based games with bounded rational agents induce path-dependent outcomes. As a consequence, game repetitions do not necessarily replicate the results despite unchanged external conditions. A consistently chosen portfolio should exhibit little variability in its holdings in each time period over independent repetitions. We should acknowledge that we use proportions of agents per portfolio as a measure since we do not collect data on particular portfolios that were possessed by individual agents throughout the games.[16] Consistency in selection is tested with two different measures: CV and MC methods.

Coefficient of Variation

The CV is defined as the ratio between the standard deviation and the mean. We use it to measure the variability of portfolio holdings within single time periods over game repetitions. To estimate the coefficient, we apply the following procedure. First, we display proportions of agents per each portfolio in a given period over repetitions in the matrix form of 30 columns, representing repetitions, and T rows, representing periods. Then, we estimate the average proportion of agents per portfolio in every period across repetitions and the corresponding standard deviation, from which the row CV for a single period is obtained. The reported CV value for a single portfolio is then averaged over T periods.[17] If a portfolio is chosen consistently over repetitions, then the standard deviation is small, or close to zero, and the corresponding CV is also small, or close to zero. On the other hand, an inconsistently chosen portfolio over repetitions has a large standard deviation and, hence, large CV value.

To allow for the proportion of liquidity agents per portfolio and to de-link it from the portfolio desirability, we truncate the bottom line of portfolio holdings to 0.5 %. Therefore, if less than 0.5 % of agents possess a given portfolio in a given time unit, then the value is set to 0.5 %, by which we avoid the potentially high variability in the proportion of liquidity agents, which might not have been meaningful. For example, if a portfolio is possessed only by liquidity agents, whose proportion is set to 0.1 % in the first realization and 0.2 % in the second, this would signify 100 % variability in the holdings of a portfolio, although it would not relate to the portfolio desirability but to the proportion of liquidity agents. In both cases, a portfolio, which is possessed only by liquidity agents, is consistently avoided for which the true variability should be zero. Results for unsuspicious (US) and suspicious agents (S) for all three settings are reported in Table 9.7.

The least desired portfolios of unsuspicious agents within the baseline setting (S4, S13, S8, S11, S15 and S19) exhibit the smallest average row variability in holdings. This means that the average per-period holdings of these

---

[16] It might be the case that even though the series repeated itself, different agents might hold the same portfolios.

[17] For this reason, we also refer to row variability in the sequel to be explicit that we measure the consistency within different time units, while not also between time units. This is measured with MC.

Table 9.7 CVs of unsuspicious and suspicious agents

	Baseline		Bear		Bull	
	US	S	US	S	US	S
S1	17.47	80.02	0.51	1.68	31.13	16.89
S2	35.11	34.92	44.51	28.60	0.24	58.40
S3	52.94	49.12	14.03	16.52	0.46	11.16
S4	7.06	37.13	8.27	9.48	20.78	29.79
S5	20.71	34.35	25.37	39.15	0.43	28.80
S6	37.44	61.10	0.99	22.86	40.88	27.20
S7	23.16	56.01	0.90	40.90	20.42	44.99
S8	9.11	59.13	1.28	5.47	6.61	19.53
S9	53.13	58.68	0.67	29.52	16.54	39.36
S10	38.82	42.94	24.56	19.29	0.53	44.87
S11	9.32	75.83	9.43	25.41	38.35	34.16
S12	24.51	31.05	57.89	40.67	0.51	40.95
S13	7.77	67.57	21.25	36.61	14.36	52.08
S14	30.46	39.03	31.61	22.37	0.45	16.41
S15	10.38	64.04	9.83	20.95	23.86	48.82
S16	66.47	45.48	3.04	40.57	12.89	49.96
S17	22.53	80.11	1.76	11.55	16.07	29.40
S18	57.17	46.22	1.23	45.87	14.20	45.36
S19	10.86	74.43	1.63	30.55	28.07	38.15
S20	58.92	50.82	1.96	43.76	2.13	56.16
S21	13.91	75.67	1.55	17.18	25.84	31.46
S22	15.88	58.93	29.11	35.97	12.82	47.88
S23	25.09	38.00	33.74	29.38	0.47	38.00
S24	42.10	46.93	22.28	35.24	19.68	55.33
S25	27.52	62.16	27.58	40.86	1.91	54.43
S26	18.33	57.69	3.52	32.23	38.56	39.42
S27	52.94	44.01	13.36	41.95	1.94	50.20
S28	26.59	56.87	2.15	30.41	46.97	39.52
S29	18.35	57.64	2.46	29.92	39.70	41.23
S30	58.02	58.44	42.84	34.59	3.70	42.50
S31	39.93	47.93	4.86	35.52	27.34	45.73

portfolios were the most stable within the game repetitions. These portfolios lay the furthest from the portfolios of the efficient frontier and were the least desired. This indicates that agents might be capable of, first, identifying the least-preferred portfolios and then persistently avoiding them. In fact, agents are slightly more capable of avoiding the least-desired portfolios than of being consistent on the most desired. Namely, the variability of most-desired portfolios, S5, S12, S23, S2 and S14, was very similar for both groups of agents, ranging from 20.71 % (S5) to 35.11 % (S2) for unsuspicious agents and from 31.05 % (S12) to 39.03 % (S14) for suspicious agents. From the most-desired portfolios of unsuspicious agents, the smallest row variability was exhibited by the most desired portfolio S5, followed by the second most S12 (24.51 % variability), and then S23 (25.09 %), S14 (30.46 %) and S2 (35.11 %). However, portfolios S14 (8.24 %) and S2 (8.78 %) were similarly desired.

Because the suspicious agents were not as capable as the unsuspicious in selecting winning and avoiding losing portfolios, their selections exhibit larger row variability than that of unsuspicious agents. By suspicious agent, the smallest row variability is exhibited in the most desired portfolios from the efficient frontier, S12, S5, S2, S23 and S14, as well as the least-desired portfolio S4. All of these portfolios are followed by the neighbouring portfolios S10, S27, S16 and S18. Suspicious agents are thus far more capable of being persistent on the most-desired portfolios and little less for the least-desired, although they too are able to persistently avoid the latter. However, they are not so consistent regarding portfolios that lie in-between. In contrast to unsuspicious agents, this result indicates that suspicious agents fail to identify properly the least-desired portfolios, and for this reason they may either hold them too long or trade them too much. The implication is similar to that identified in Odean [74, 75]. Altogether, unsuspicious agents are more persistent in their selections than suspicious agents, which is not surprising.

In the bear market, decisions of unsuspicious agents exhibit a small degree of variability. This is especially true for the least-desired portfolios. Note that portfolios possessed by liquidity agents are truncated. The most-desired portfolios (S3, S10, S14 and S23) exhibit quite a small variability that does not exceeding 34 %. The least consistent portfolios are those that either do not end with liquidity agents or are not the most desired, such as S2 and S12. In a bull market, the unsuspicious agents were also the most consistent for the least risky portfolios and the least desired portfolios, which ended up with the liquidity agents. As before, suspicious agents' CV values exceed that of the unsuspicious, indicating that, in this respect, they behave less consistently than the unsuspicious.

Monte Carlo

Although the analysis of CVs is compelling, it is by no means complete. CV measures the row dispersion of portfolio holdings in each time period over repetitions, which might fail to acknowledge the potential linear dependence of game repetitions.

The idea underlying MC is that if a portfolio is chosen consistently over all repetitions of the games, then for its transition from one period t to the successive period $t+1$ state it should be irrelevant from which of the

30 repetitions of the game observations were taken. In perfectly consistent decision making, an agent would always opt for the same choice regardless of the transition from one time period to the next being made within the same repetition of the game or any other. If this is not the case, then this would mean that repetitions matter, that is, an indication of agents' inconsistent selections.

To compute the MC method, we employ the following six-step procedure. First, we transform the data because initial holdings of portfolios over repetitions may differ (initial distribution of portfolios was done randomly with a variable 'seed'), which would prevent proper inter-period comparability. The last time period is used as the base so as to minimize the influence of the initial set-up, while the data for each portfolio in every repetition of the game is expressed as a ratio to the value of portfolio holdings in the last period of that repetition as $\tilde{X}_t = \frac{X_t}{X_T}$.[18] Then, one from the 30 repetitions per selected portfolio is randomly chosen. Starting with the initial time period t, the computer algorithm is set to choose randomly one repetition from all 30 repetitions per selected portfolio (A), and then, irrespective of the first selection, one repetition from the subsequent $t+1$ time period (B). Subsequently, the value of the portfolio holding under (B) is compared to the value of (A) in time period $t+1$ and the difference is squared. If these portfolio holdings are chosen consistently, there should then be no difference between both compared values irrespective of the repetition number from which (B) was taken. Then, the squared differences are summed over all 10,000 runs for every time period and, finally, the median of all time periods is reported as the persistence factor of the portfolio.[19] By the rule, the lower the value, the larger is the persistency level, and vice versa. Table 9.8 reports results of the MC for all three settings.

In the baseline setting, unsuspicious agents most consistently choose S4 with a median of 0, which indicates a perfect fit. Not only do agents consistently avoid S4, but they do it in the very early stages of the games. Portfolio S4 is followed by portfolios S12 and S2, which still exhibit very large levels of persistence, and then by portfolios S23 and S5 with a bit lower (still very high) persistence levels. The grouping is very similar to that of the CV analysis above with one exception. Namely, in the CV analysis portfolio S4 follows the group of the efficient frontier portfolios but does not lead the group as in the present case.

[18] In a baseline case $T=264$, in a bull market $T=209$ and $T=120$ in a bear market.

[19] Median is used instead of the mean to reduce the influence of extreme values.

Table 9.8 Medians of sum of squares of the difference

	Baseline		Bear		Bull	
	US	S	US	S	US	S
S1	2608	3735	951	7667	309	265
S2	87	334	3575	1333	5518	1531
S3	804,228	489,909	307	340	7431	48,318
S4	0	816	87,687	3282	9291	9250
S5	312	142,580	58,912	1463	4504	103,892
S6	1645	490	6316	9595	1004	676
S7	4269	10,603	5133	1,150,530	12,060	37,204
S8	560	11,231	917	79,428	59	812
S9	3498	1987	1457	621,128	48,888	5412
S10	1499	10,142	846	809	2424	1,059,915
S11	1310	896,429	117,730	177,042	12,463	2998
S12	39	1468	75,968	48,811	6615	73,898
S13	325	44,102	5,585,360	54,851	78,240	100,291
S14	4829	79,743	1408	648	7061	261,609
S15	784	179,501	25,602	9103	81,889	959,069
S16	5066	3532	28,150	7575	450,027	8536
S17	2611	5934	5898	39,218	313	2668
S18	385	796	10,889	227,040	13,822	1315
S19	623	21,631	12,430	56,259	7387	8894
S20	2199	86,009	933	12,366	35,987	25,621
S21	2281	22,005	2487	510,585	4927	3654
S22	10,378	76,915	2,466,410	19,956	134,300	2,182,880
S23	213	5324	3168	1196	1328	1703
S24	7712	51,176	478,540	485,670	152,316	4994
S25	12,509	36,583	181,188	770,585	2034	17,258
S26	4872	5621	33,154	25,409	36,027	6402
S27	2709	398,750	69,253	1996	33,213	22,804
S28	5638	2618	146,572	4,633,220	4606	9460
S29	3735	56,333	9334	6240	17,734	3068
S30	5657	38,936	3,897,070	3884	54,837	11,137
S31	5327	31,181	53,111	730,177	3169	6355

The persistency level of other portfolios decreases with their distance from the most consistently chosen portfolios. The least consistently chosen portfolio is S3, which lies on the inefficient frontier. Suspicious agents exhibit a large degree of consistency for high-return and high-risk portfolios S6, S18 and S9. However, the behaviour of suspicious agents is much less consistent than that of unsuspicious agents, implying that suspicious agents in general exhibit a much lower preference over portfolio choice.

In the bear market, the most consistently chosen portfolios of unsuspicious agents were S3, S10, S8, S20, S1, S14 and S9, while those of suspicious agents were S3, S14, S10, S23, S5 and S27. Of these, S3, S10 and S14 were the most-desired portfolios, with S23 under the suspicious agents also being so.

The least consistent portfolios were those whose proportion was not stable over time (S13, S11, S22, S24, S28, S31, S7 for suspicious agents, and S30 for unsuspicious agents), for which they exhibited huge differences between the average-game and the endgame settings. These portfolios were the furthest apart from the most consistently chosen portfolios and were found to be in the centre. Of the two cohorts, suspicious agents were less consistent in their selections, which resulted from their inability to select the winners to the same extent as unsuspicious agents.

In the bull market, unsuspicious and suspicious agents most consistently chose portfolio S8, followed by portfolios S1, S6 and S17, which still exhibited very high levels of persistence. These three were the most-desired portfolios in both the average and the endgame settings. Portfolios S22, S15, S10, S13, S14, S16 and S5 exhibit the lowest levels of persistence. These are portfolios with the lowest mean returns and the lowest variance, and were among the least desired. Again we observe that the persistency level of other portfolios decreases with their distance from the most consistent portfolios. As before, the behaviour of the suspicious agents was much less consistent than that of the unsuspicious agents also in the bull market.

5 Discussion

Table 9.9 summarizes the results from all settings. It reports the average-game (AVG) and the endgame (END) results for all three time spans for both cohorts. For each setting, we first report the percentage of agents with five of the most desired portfolios, then the number of portfolios (out of 31) that are possessed by the last decile of agents, then weighted betas of portfolios held, then weighted betas for five the most desired and ten the least desired portfolios and finally lambda values from the power-law distribution.

The average and the weighted beta are calculated as $\overline{\beta} = \frac{1}{n}\sum_{i=1}^{n}\beta_i$ and $\beta_W = \frac{1}{n}\sum_{i=1}^{n}\beta_i h_i$, respectively, with h_i representing a fraction of agents with portfolio i and $\sum_{i=1}^{n} h_i = 1$. For sub-samples, in which $\sum_{i=1}^{n} h_i \neq \sum_{i=1}^{n} q_i$, a fraction of selected portfolios is re-calculated so as to reach the required condition $\sum_{i=1}^{n} h_i = 1$. As for the lambda values, we presume the power law distribution of the type $y = Ax^{-\lambda}$. Parameter λ refers to the densification parameter. The distribution of a random variable follows a power law if the frequency of

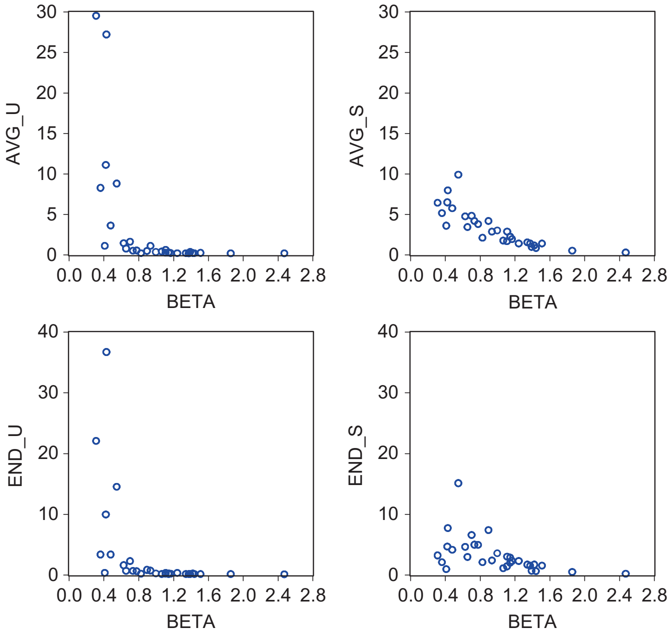
Table 9.9 Overview of results

	Unsuspicious		Suspicious	
	AVG	END	AVG	END
Baseline				
Top 5 (%)	84.75	86.40	36.41	41.60
Least 10 % (No./31)	24/31	25/31	10/31	11/31
Weighted beta	0.459	0.467	0.731	0.778
Top 5	0.396	0.419	0.451	0.637
Least 10	1.450	1.424	1.391	1.280
Lambda	1.826	2.030	0.838	0.945
Bear				
Top 5 (%)	73.23	95.67	57.90	83.35
Least 10 % (No./31)	21/31	28/31	13/31	25/31
Weighted beta	0.842	0.714	0.808	0.696
Top 5	0.789	0.706	0.684	0.679
Least 10	1.487	1.160	1.088	1.257
Lambda	1.672	2.017	1.885	2.037
Bull				
Top 5 (%)	80.00	84.83	50.01	64.38
Least 10 % (No./31)	23/31	25/31	12/31	17/31
Weighted beta	1.600	1.610	1.317	1.346
Top 5	1.712	1.692	1.587	1.471
Least 10	0.600	0.674	0.677	0.841
Lambda	1.820	2.088	1.077	1.883

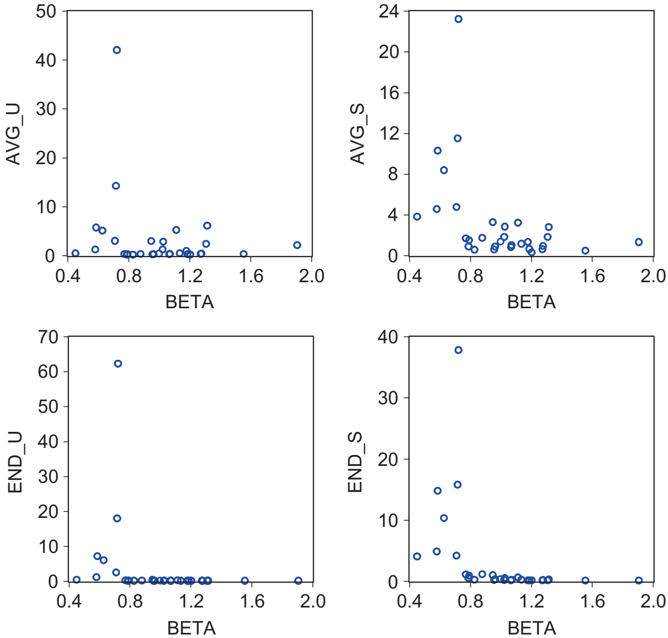
an event decreases to the inverse of some exponential degree as the size of the event increases. In our case, the parameter measures the level of synchronization in agents' selections. High lambda values indicate a striking synchronization pattern, which means that a small number of portfolios is possessed by a large number of agents, while the vast majority of remaining portfolios are possessed by only a few agents. Synchronization occurs for $\lambda > 1$. The corresponding power-law exponents were estimated with the OLS regression.[20]

Figure 9.8 plots the unsuspicious and suspicious agents' average-game and endgame selections against the beta coefficients of portfolios for all three time spans. Dots in the plots represent different portfolios according to their betas (*X*-axes) and the fraction of agents having selected them (*Y*-axes). 'S' and 'U' in the plots designate suspicious and unsuspicious agents, respectively; and 'AVG' and 'END' designate average-game and the endgame selection, respectively.

[20] The correlation coefficient has typically been used as an informal measure of the goodness of fit of a distribution to a power law [76]. For mathematical derivations of a scale-free distribution see Mitzenmacher [77] and Newman [78].

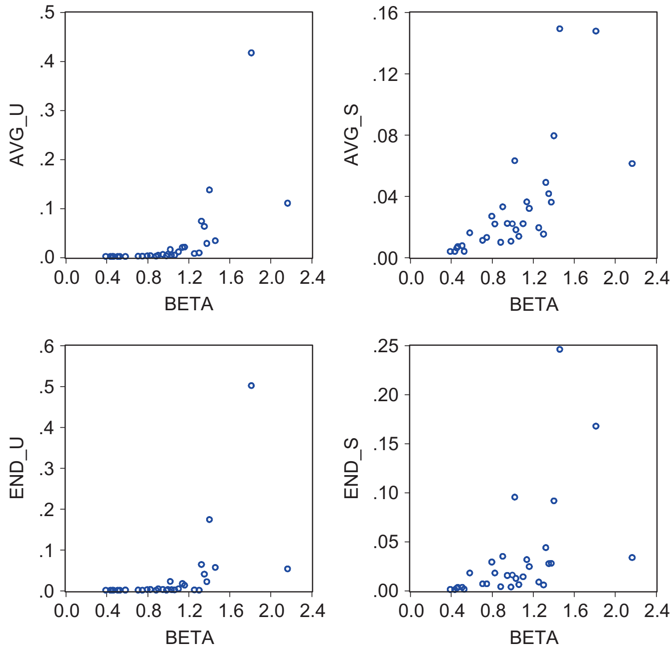


a) baseline



b) bear market

Fig. 9.8 Scatter graphs of unsuspicious and suspicious agents' average-game and endgame selections against the beta coefficients of portfolios



c) bull market

Fig. 9.8 (continued)

Because portfolios are selected upon their returns, we would expect that the portfolio desirability is increasing in return. But this is not the case. The results suggest that the riskier the portfolio, the more likely it is that it will be avoided. Agents could thus be regarded as risk-averse. This is true for both cohorts. However, the main difference in the selections of these cohorts is that the transition from most desired to least desired is very discrete in the case of unsuspicious agents, while the transition goes through the group of moderately desired portfolios within suspicious agents. In both cases, portfolios S2 and S3 demonstrate that agents are capable of selecting high return portfolios and avoiding low-risk low-mean portfolio. Diagonal dashes (Fig. 9.8) in the case of suspicious agents demonstrate that they are less risk averse, although they require higher returns for bearing additional risk. However, the agents did not select high-mean high-risk portfolios, such as S1 within the baseline framework. The steeper dashes signify that average-game decisions are more motivated by variance and less by returns. A general observation would be that unsuspicious agents are more focused on the risk than the returns and try to minimize it. The implication is instructive, yet encouraging, because portfolios are selected upon returns and not the risk, while the latter is crucial for the portfolio selection. On the other hand, although the suspicious agents prefer more risk, they still weigh returns against risk. In addition, agents hold very under-diversified portfolios, which is consistent with some empirical facts [79, 80]. Moreover, diversified portfolios are among the least consistently chosen portfolios.

Selected portfolios of unsuspicious agents of the baseline framework can be grouped in two clusters. The first one consists of five portfolios from the efficient frontier, which were chosen by more than 5 % of the agents each (S10 can also be added to this group, as it lies on the efficient frontier and was chosen by 3.6 %). The second one includes all the remaining portfolios. S3 is among the least risky portfolios but has the lowest return, while S18 and S27 lie very close to the efficient frontier but are slightly riskier. S1 and S6 are both high-return and high-risk portfolios and were avoided by unsuspicious agents. Hence, portfolios were either possessed or avoided. On the other hand, selected portfolios of suspicious agents can be grouped into three clusters, as represented by the dashes (Fig. 9.8): the six portfolios from the efficient frontier, being chosen by more than 5 % of the agents; the four least-desired portfolios as chosen by less than one percent (S4, S8, S13 and S15); and the portfolios that lie between the two groups. The least-desired portfolios from the second group are the riskiest portfolios and also exhibit the lowest returns. Most portfolios of the third group are riskier than those of the first group and less risky from those of the second. There are three outliers to this apparent linearity-in-risk rule: S1, S11 and S3. S1 and S11 are as risky as those in the second group, but exhibit higher returns. S3 is among the safest but exhibits the lowest return. All three belong to the third group of portfolios.

Suspicious agents are much more inclined towards riskier portfolios than the unsuspicious, which is reflected in their weighted beta (0.731), which is much higher than that of unsuspicious agents (0.459). Still, both values are much below market risk; an indication that both cohorts behave risk averse. The most-desired portfolio of unsuspicious agents (S5) has the smallest beta of only 0.314, while the second most desired portfolio (S12) also has very small beta (0.433). The five most-desired portfolios of unsuspicious agents have an average beta of 0.418 and weighted beta of 0.396. The five most-desired portfolios of suspicious agents also have small betas on average (0.441), with a weighted beta of 0.451. This is a logical consequence of the fact that the unsuspicious agents were highly capable of selecting winners, which are the lowest-beta portfolios. As a consequence, the least desired portfolios have larger betas. More precisely, the ten least-desired portfolios of unsuspicious agents have an average beta of 1.469 and weighted beta of 1.450, while those of suspicious agents have an average beta of 1.519 and weighted beta of 1.391. Under both settings, only liquidity agents selected portfolio S4, which is the portfolio with the largest beta of 2.475. Given the results, one could argue that the losses from high-beta portfolios are too large for agents to hold.

To see how agents weigh between the risk and returns, three cases seem of particular interest because the alternatives exhibit a lot of similarities but are differently possessed. One is S2 in relation to S6 and S18. S2 is the portfolio with the highest mean and a moderate variance. S2 has similar risk as S18 but a much higher mean return, while it had a mean return very similar to S6 but a substantially smaller risk. S2 was chosen by 8.78 % of unsuspicious agents and 9.87 % of suspicious agents, which made it the fourth most-desired portfolio. On the other hand, S18 was chosen by only 1.61 % of unsuspicious agents and 4.80 % of suspicious agents on average, while S6 was chosen by just 0.46 % of unsuspicious and as much as 4.15 % of suspicious agents. Obviously, unsuspicious agents were not willing to give up the additional 'riskless' return of S2, while the suspicious agents were also inclined towards riskier and less profitable alternatives. The second case is S3 in relation to S14. S3 was slightly more volatile than S14 but exhibited a significantly lower mean return; the lowest among all portfolios. However, S3 was on the average chosen by 1.07 % of unsuspicious agents and 3.58 % of suspicious agents. On the other hand, S14 was chosen by 8.24 % and 5.14 % of unsuspicious and suspicious agents, respectively. The last one is S20 to S23. Both portfolios have similar mean returns and significantly different variance. Unsuspicious agents largely chose S23 (11.07 %) and left the riskier S20 mostly to liquidity agents. On the contrary, 3.41 % of suspicious agents took S20 and 6.47 % took S23, which implies that suspicious agents prefer higher risk. Following these three cases, we could say that when agents select portfolios, they first make a group of satisfying portfolios in relation to their returns and risk. The width of this area depends on the level of agents' suspiciousness. In contrast to unsuspicious agents, who very accurately select portfolios with higher returns, suspicious agents' decisions are distributed among many portfolios that are close together. The reason for this is that when two suspicious agents compare their outcomes, the probability that they would take a less lucrative portfolio is different to zero. This is more likely to occur when the two returns are close together (see Proposition 3). Desirability of portfolio S3 indicates that agents may not want to opt for the least risky portfolios that also fail to yield a satisfactory return.

Agents' endgame decisions share many similarities with that of the average-game in the main conclusion that the desirability of a portfolio decreases with its risk. In both cases, the transition from the most to the least desired portfolios is gradual. However, less sloping dashes signify that endgame decisions are more affected by returns and not so much by risk. Clearly, the transition from the most-desired portfolios to the least-desired is diagonal from the high-mean and low-risk portfolios towards the low-mean and the high-risk portfolios.

The corresponding endgame weighted beta value of 0.467 for the unsuspicious agents is much lower than that of suspicious agents (0.778). The five most-desired endgame portfolios of unsuspicious agents have a weighted beta of 0.419, which is below the corresponding beta of the five most-desired portfolios of suspicious agents (0.637). Weighted beta of the least desired portfolios was 1.424 for unsuspicious agents and 1.280 for suspicious. We have seen before that unsuspicious agents highly prefer less risky portfolios, much more than suspicious agents whose selections are more dispersed, and the betas demonstrate this. Moreover, not only are unsuspicious agents more risk averse as the games proceed, they are also risk averse in their final decisions. The greatest part of this difference in the behaviour of the two groups does not lie in the selection of winners, but rather in the proportion of these winners and also in that of the losers. In addition, they are capable of selecting winners, while the suspicious agents are not. The corresponding lambda values of the average-game decisions equal $\lambda = -1.826$ ($R^2 = 0.97$) for unsuspicious agents and $\lambda = -0.838$ ($R^2 = 0.75$) for suspicious agents. In the endgame setting, the corresponding lambdas are $\lambda = -2.030$ ($R^2 = 0.98$) for unsuspicious agents and $\lambda = -0.945$ ($R^2 = 0.77$) for suspicious agents. Five of the most desired portfolios were chosen by 84.75 % of unsuspicious agents and 24 of all 31 possible portfolios were chosen by the last decile of unsuspicious agents on average. This is an implication of herding, which was highly pronounced in the setting of unsuspicious agents. The absolute herding was prevented by liquidity agents. There are no densification patterns in the selection by suspicious agents because their selections are more evenly distributed across the set of available portfolios. Five of the most desired portfolios account for 36.41 % of all suspicious agents with the last decile of suspicious agents having 10 out of the 31 possible portfolios on average.

It can be noted from the MC values that suspicious agents are much less consistent in their inter-period selections than the unsuspicious. The unsuspicious agents most consistently choose portfolios from the efficient frontier and those that are the farthest apart from the first group. The first could be considered consistently desired portfolios, while the second consistently avoided. An exception is S14, the inconsistently held portfolio from the efficient frontier. There is no appealing explanation for this. Portfolios S10 and S27 can be found alongside the efficient frontier portfolios but they exhibit significantly lower levels of persistence than the efficient frontier portfolios. An intuitive explanation for this may be that when agents decide which portfolio to acquire they opt for the first best alternative, which in our case lies on the efficient frontier, rather than their closest neighbouring portfolios. The latter are portfolios that an agent would not be eager to change once owning one, but also one that other agents would not be eager to obtain. Such is the case of S10, which is held by 3.6 % of agents. Portfolios from the efficient frontier are among the most consistently chosen. S14 is highly desirable with an average holding rate above 8 % and with a very small CV value. However, its MC result implies that agents exerted a highly inconsistent trade policy towards it. Even though the portfolio as such yields highly stable returns and has the lowest volatility, the returns are very low, fluctuating around zero. Unsuspicious agents, who have been found to weigh returns against volatility, valued the portfolio's low volatility, but the portfolio's poor return was a likely reason that it was not persistently desired.

Another interesting case (portfolio S4) comes from the position of suspicious agents. The possession of the portfolio was fairly stable within particular repetitions but highly volatile between repetitions. Higher row volatility on the one hand and lower column volatility on the other might indicate that the holding of such a portfolio would be stabilized at different levels over time in individual repetitions. Stabilization at different levels preserves the row volatility between repetitions but at the same time eliminates the column variability. Liquidity agents, who persist on their initial portfolios and do not trade them regardless of the returns, would very likely hold such portfolios. Further, possession of portfolio S5 by suspicious agents and, to a lesser extent, portfolio S3 of unsuspicious agents exhibited the second lowest level of row variability but the largest between-period variability. The MC test indicated that agents did not have any consistent preference for the portfolio, while the small row variability suggests that the holdings of the portfolio had not stabilized. S1 was also among the most consistently chosen portfolios by suspicious agents, but on the other hand, even though exerting the third largest average return, the portfolio was one of the least-desired by suspicious agents. The most desired portfolio of suspicious agents (S2) was also the most consistently chosen portfolio by suspicious agents. Portfolio S2 is profitable and risky.

The results do not differ greatly, in the bear and the bull market. A huge synchronization has been perceived in the bear trend where 62.19 % of all ended with the winning portfolio S3, 80 % with the first two portfolios, and 95.7 % with five the most desired portfolios. The winning portfolio is a one-asset portfolio S3, and the next two, two-asset portfolios that included S3 and either S2 and S5, which are the next two least-losing stocks. The weighted beta of endgame settings is much higher during a bull trend (1.61/1.35 for unsuspicious/suspicious agents) than during a bear (0.714/0.696). The same is true for the average-game settings (0.84/0.81 in the bear trend and 1.60/1.32 in the bull trend for unsuspicious/suspicious agents). The results demonstrate that agents respond differently to risk in different time trends. More specifically, agents take more risk in the bull market than in the bear, which is consistent with observations of Barberis et al. [35].

Agents' decisions are also much more synchronized in the bear trend than in the bull trend, which is an implication of the previous result. The winning portfolio has been selected by more than 62 % of unsuspicious agents in the bear trend comparing to 50.11 % in the bull trend. In the bear market, the power law parameter of the average-game setting is $\lambda = -1.672$ ($R^2 = 0.95$) for unsuspicious agents and $\lambda = -1.885$ ($R^2 = 0.83$) for suspicious agents. The corresponding parameters of the endgame setting are $\lambda = -2.017$ ($R^2 = 0.91$) for unsuspicious agents and $\lambda = -2.037$ ($R^2 = 0.97$) for suspicious agents. Higher values in the endgame setting than in the average-game indicate that over the course of time agents approach their desired portfolios. In the endgame settings, both cohorts display a striking synchronization pattern. In the bull market, the power law parameter of the average-game setting is $\lambda = -1.820$ ($R^2 = 0.98$) for unsuspicious agents and $\lambda = -1.077$ ($R^2 = 0.89$) for suspicious agents. The corresponding values of the endgame setting are $\lambda = -2.088$ ($R^2 = 0.97$) for unsuspicious and $\lambda = -1.883$ ($R^2 = 0.81$) for suspicious agents. Slightly smaller lambda values of the bull trend compared to the bear trend indicate the agents' weaker ability to select the winning portfolios in the bull trend compared to that in the bear trend. This is especially true for the suspicious agents' average-game decisions in the bull trend, where the synchronization pattern is barely met. However, the endgame decisions of both cohorts exhibit a striking synchronization pattern. Although the behaviour of agents is more risk aversive in a bear market than in a bull, the selection of portfolios S13 and S22 could signify the presence of a reflection effect, that is, risk seeking in the domain of losses or when agents are trapped inside the risk, the phenomena that is then followed by loss recognition and the retraction from such alternatives. These two risky portfolios were relatively largely possessed in an average-game setting of the bear market, but ended on liquidity agents in the endgame. Both effects are present only with the unsuspicious agents. A reflection effect is an important outcome of these simulation games, because it has been singled out as an anomaly that has also been documented as such.

Differences in the choices of the two cohorts indicate that agents' portfolio selections are in some instances highly sensitive to the choice parameters. Not only do the suspicious agents select different portfolios than the unsuspicious, they are also less consistent in their selections. Recall that, by definition, as the agents' suspiciousness increases, the probability that the less profitable of the two portfolios is adopted increases. One noteworthy implication of the increasing suspiciousness is that fewer liquidity agents are required in order to prevent synchronization in selections.

6 Conclusion

In the present chapter we examined portfolio selection under uncertainty and related it to the simulation-based games on social networks. Central to the model are bounded rational agents with pairwise relations. Portfolio selection is thus considered a synthesis between human cognition and social networks in a dynamic and emergent environment. We should add that a positive decision analysis is applied, addressing the question of which portfolios are selected and not which should be.

Although portfolios are selected upon the stock returns, it has been demonstrated that the interacting agents are capable of selecting the mean-variance portfolios and that they behave in a risk-averse way. In addition, the riskier the portfolio, the more likely it is that agents will avoid it. Further, selections of unsuspicious agents exhibit heavy synchronization patterns, both over the course of time and in the final selections. Higher lambda values in the endgame setting than in the average-game setting signify the synchronization process over the course of the game. Agents' suspiciousness can have a substantial effect on portfolio selection. A bigger dispersion in the selection of suspicious agents is a consequence of their slight failure to synchronize the selections, even though they also identify the same winners as do unsuspicious agents. This conclusion is supported in both bull and bear markets, with the exception that agents choose riskier portfolios in the bull market. The assumption that agents are less willing to take risk in a domain of losses (bear market) has been often confirmed by the behavioural studies. Highly preferred portfolios are two-asset portfolios with an additional stock added to the most desired one. Stocks that are added either sufficiently reduce the risk of a dominant stock or improve its profitability or both, thereby making it more desirable. In a bear market, synchronization is extreme. Although agents are more risk averse in a bear market, the model is able to single out the reflection effect as an anomaly, which has been documented as such. Consistency tests have demonstrated that unsuspicious agents are much more consistent in their selections than suspicious agents. In addition, the most consistently chosen portfolios are the most desired and the least desired portfolios. The latter could be considered as consistently avoided. Again, this conclusion is also supported in bull and bear markets. Highly preferred are the under-diversified portfolios, while the market portfolio is often left to liquidity agents. In addition, diversified portfolios are among the least consistently chosen choices. All this shows that portfolio selection may be, in some aspects, highly sensitive to the change in selection parameters.

Appendix: Fragmented pseudo-code

```
// import the data and create the network

// distribute initial strategies and set the payoffs
for (x=0;x<L;x++) {
      pay[x]=1.0;
      strat[x]=randl(31);
      stratnew[x]=strat[x];
}

// define the strategies / portfolios / and their dynamics
for (x=0;x<L;x++) {
      if (strat[x]==0) {
            pay[x]+=pay[x]*rAA;
      } if (strat[x]==1) {
            pay[x]+=pay[x]*rMSFT;
      } ... {
      } if (strat[x]==30) {
            pay[x]+=pay[x]*(0.2*rAA+0.2*rMSFT+0.2*rXOM+0.2*rC+0.2*rKFT);
      }

// selection procedure
for (x=0;x<L;x++) {

      if (x==lA) {} // do nothing if an agent is a liquidity agent

      else {
      center=x;
      k=randl(con[x]); // an adjacent agent is chosen at random
      neigh=vertex[center][k];
      prob=1.0/(1.0+exp(-(pay[neigh]-pay[center])/kapa));
            if (randd()<prob) {
                  adopt=strat[neigh];
            } else {}
      }
}

// calculate fractions of agents per portfolio

// export results
```

Acknowledgements The present chapter is a part of my doctoral dissertation *Simulating Portfolios by using Models of Social Networks*. I would like to thank to my supervisor Harry M. Markowitz and my brothers Matej and Mitja for their valuable comments. I thank seminar participants at ETH Zürich, the Kiel University, the 18th International Conference on Computing in Economics and Finance and the Conference on Behavioral Aspects in Macroeconomics and Finance, and an anonymous referee for helpful comments and suggestions. All errors remain my sole responsibility.

References

1. Markowitz, H. (1952a). Portfolio selection. *Journal of Finance, 7*, 77–91.
2. Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics, 51*, 247–257.
3. Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory, 3*, 373–413.
4. Barberis, N. (2000). Investing for the long run when returns are predictable. *Journal of Finance, 55*, 225–264.
5. Brennan, M. J., Schwartz, E. S., & Lagnado, R. (1997). Strategic asset allocation. *Journal of Economic Dynamics and Control, 21*, 1377–1403.
6. Buraschi, A., Porchia, P., & Trojani, F. (2010). Correlation risk and optimal portfolio choice. *Journal of Finance, 65*, 393–420.
7. Cocco, J. F. (2005). Portfolio choice in the presence of housing. *Review of Financial studies, 18*, 535–567.
8. Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance, 58*, 1651–1684.
9. Kan, R., & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis, 42*, 621–656.
10. Liu, J., Longstaff, F. A., & Pan, J. (2003). Dynamic asset allocation with event risk. *Journal of Finance, 58*, 231–259.
11. Longstaff, F. A. (2001). Optimal portfolio choice and the valuation of illiquid securities. *Review of Financial Studies, 14*, 407–431.
12. Campbell, J. Y., & Viceira, L. M. (2002). *Strategic asset allocation: Portfolio choice for long-term investors*. New York: Oxford University Press.
13. Garlappi, L., Uppal, R., & Wang, T. (2007). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies, 20*, 41–81.
14. Liu, J. (2007). Portfolio selection in stochastic environments. *Review of Financial Studies, 20*, 1–39.
15. Ang, A., & Bekaert, G. (2002). International asset allocation with regime shifts. *Review of Financial Studies, 15*, 1137–1187.
16. Benzoni, L., Collin-Dufresne, P., & Goldstein, R. S. (2007). Portfolio choice over the life-cycle when the stock and labor markets are cointegrated. *Journal of Finance, 62*, 2123–2167.
17. Benartzi, S., & Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics, 110*, 73–92.
18. Berkelaar, A. B., Kouwenberg, R., & Post, T. (2004). Optimal portfolio choice under loss aversion. *Review of Economics and Statistics, 86*, 973–987.
19. Campbell, J. Y., & Viceira, L. M. (1999). Consumption and portfolio decisions when expected returns are time varying. *Quarterly Journal of Economics, 114*, 433–495.
20. Campbell, J. Y., & Viceira, L. M. (2001). Who should buy long-term bonds? *American Economic Review, 91*, 99–127.

21. Shefrin, H., & Statman, M. (2000). Behavioral portfolio theory. *Journal of Financial and Quantitative Analysis, 35*, 127–151.
22. Hayek, F. (1945). The use of knowledge in society. *American Economic Review, 35*, 519–530.
23. Bala, V., & Goyal, S. (1998). Learning from neighbours. *Review of Economic Studies, 65*, 595–621.
24. Banerjee, A. V. (1992). A simple model of herd behavior. *Quarterly Journal of Economics, 107*, 797–817.
25. Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy, 100*, 992–1026.
26. Blume, L. E. (1993). The statistical mechanics of strategic interaction. *Games and Economic Behavior, 5*, 387–424.
27. Castellano, C., Fortunato, S., & Loreto, V. (2009). Statistical physics of social dynamics. *Reviews of Modern Physics, 81*, 591–646.
28. Ellison, G., & Fudenberg, D. (1995). Word-of-mouth communication and social learning. *Quarterly Journal of Economics, 110*, 93–125.
29. Hong, H., Kubik, J. D., & Stein, J. C. (2005). Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. *Journal of Finance, 60*, 2801–2824.
30. Shiller, R. J. (2005). *Irrational exuberance*. Princeton: Princeton University Press.
31. Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge: MIT Press.
32. Hirshleifer, D. (2001). Investor psychology and asset pricing. *Journal of Finance, 56*, 1533–1597.
33. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.
34. Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics, 106*, 1039–1061.
35. Barberis, N., Huang, M., & Santos, T. (2001). Prospect theory and asset prices. *Quarterly Journal of Economics, 116*, 1–53.
36. Tesfatsion, L., & Judd, K. L. (2006). *Handbook of computational economics* (vol. 2). North Holland: Elsevier.
37. Steinbacher, M., Steinbacher, M., & Steinbacher, M. (2014). Interaction-based approach to economics and finance. In M. Faggini & A. Parziale (Eds.), *New economic windows: Complexity in economics: Cutting edge research* (pp. 161–203). Switzerland: Springer.
38. Zeeman, E. C. (1974). On the unstable behaviour of stock exchanges. *Journal of Mathematical Economics, 1*, 39–49.
39. Kim, G.r., & Markowitz, H. M. (1989). Investment rules, margin, and market volatility. *Journal of Portfolio Management, 16*, 45–52.
40. Brock, W. A., & Hommes, C. H. (1997). A rational route to randomness. *Econometrica, 65*, 1059–1095.

41. Brock, W. A., & Hommes, C. H. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control, 22*, 1235–1274.
42. Lux, T. (1998). The socio-economic dynamics of speculative markets: Interacting agents, chaos, and the fat tails of return distributions. *Journal of Economic Behavior & Organization, 33*, 143–165.
43. Lux, T., & Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature, 397*, 498–500.
44. Boswijk, H. P., Hommes, C. H., & Manzan, S. (2007). Behavioral heterogeneity in stock prices. *Journal of Economic Dynamics and Control, 31*, 1938–1970.
45. Levy, M., Levy, H., & Solomon, S. (1994). A microscopic model of the stock market: Cycles, booms, and crashes. *Economics Letters, 45*, 103–111.
46. Cont, R., & Bouchaud, J. P. (2000). Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics, 4*, 170–196.
47. Iori, G. (2002). A microsimulation of traders activity in the stock market: The role of heterogeneity, agents interactions and trade frictions. *Journal of Economic Behavior & Organization, 49*, 269–285.
48. Palmer, R., Brian Arthur, W., Holland, J. H., LeBaron, B., & Tayler, P. (1994). Artificial economic life: A simple model of a stockmarket. *Physica D: Nonlinear Phenomena, 75*, 264–274.
49. Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *American Economic Review, 84*, 406–411.
50. LeBaron, B., Arthur, W. B., & Palmer, R. (1999). Time series properties of an artificial stock market. *Journal of Economic Dynamics and Control, 23*, 1487–1516.
51. Myerson, R. B. (1977). Graphs and cooperation in games. *Mathematics of Operations Research, 2*, 225–229.
52. Kirman, A. (1983). Communication in markets: A suggested approach. *Economics Letters, 12*, 101–108.
53. Kirman, A. (1997). The economy as an evolving network. *Journal of Evolutionary Economics, 7*, 339–353.
54. Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*, 440–442.
55. Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.
56. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports, 424*, 175–308.
57. Goyal, S. (2012). *Connections: An introduction to the economics of networks*. Princeton: Princeton University Press.
58. Jackson, M. O. (2010). *Social and economic networks*. Princeton: Princeton University Press.
59. Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review, 88*, 848–881.

60. Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior, 8*, 164–212.
61. Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory, 4*, 25–55.
62. Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.
63. Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The effect of myopia and loss aversion on risk taking: An experimental test. *Quarterly Journal of Economics, 112*, 647–661.
64. Knuth, D. E. (2006). *Art of computer programming*. Reading, MA: Addison-Wesley.
65. Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). Numerical recipes. 3rd edition: The art of scientific computing. Cambridge: Cambridge University Press.
66. Cohen, L. (2009). Loyalty-based portfolio choice. *Review of Financial Studies, 22*, 1213–1245.
67. Acharya, V., & Pedersen, L. H. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics, 77*, 375–410.
68. Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets, 5*, 31–56.
69. Pastor, L., & Stambaugh, R. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy, 111*, 642–685.
70. Constantinides, G. M. (1986). Capital market equilibrium with transaction costs. *Journal of Political Economy, 94*, 842–862.
71. Lo, A. W., Mamaysky, H., & Wang, J. (2004). Asset prices and trading volume under fixed transactions costs. *Journal of Political Economy, 112*, 1054–1090.
72. Benartzi, S., & Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review, 91*, 79–98.
73. Markowitz, H. (1952b). The utility of wealth. *Journal of Political Economy, 60*, 151–158.
74. Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance, 53*, 1775–1798.
75. Odean, T. (1999). Do investors trade too much? *American Economic Review, 89*, 1279–1298.
76. Chakrabarti, D., & Faloutsos, C. (2006). Graph mining: Laws, generators, and algorithms. *ACM Computing Surveys, 38*, 1–69.
77. Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet Mathematics, 1*, 226–251.
78. Newman, M. E. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics, 46*, 323–351.
79. Goetzmann, W. N., & Kumar, A. (2008). Equity portfolio diversification. *Review of Finance, 12*, 433–463.
80. Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance, 42*, 483–510.

10

Handling Model Risk in Portfolio Selection Using Multi-Objective Genetic Algorithm

Prisadarng Skolpadungket, Keshav Dahal, and Napat Harnpornchai

1 Introduction

The Markowitz model is also called the mean-variance analysis model. It has named in this way because it takes the mean (average) returns of assets in question and the variances of these returns as model inputs. As a model attempts to explain human behaviour, it assumes, for the sake of simplicity, homogenous characteristics of an investor, as a risk averse and avarice (prefer more money to less) person. Such a person tries to find a combination of assets (called a portfolio of assets) that will yield maximum return and minimum risk (i.e. minimum variance of the portfolio's returns). Furthermore, the model also assumes that returns of all investable assets in question are normally distributed. However, if this assumption is not true, especially in the short term, the

P. Skolpadungket (✉)
Thanachart Bank PCL, Bangkok, Thailand
e-mail: prisadarng@yahoo.com

K. Dahal
Chair in Intelligent Systems, School of Computing, University of the West of Scotland, High Street, Paisley, PA1 2BE Glasgow, Scotland, UK
e-mail: keshav.dahal@uws.ac.uk

N. Harnpornchai
Chiang Mai University, Chiang Mai, Thailand
e-mail: tomnapat@camt.info

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets*,
New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0_10

means of asset returns may vary from time to time. Therefore, those expected returns and volatilities need to be forecasted based on some mathematical or computational models using their previous values and some previous economic variables to find a short-term optimal portfolio. Forecasting models are far from perfect and it is their inaccuracy that is a source for model risk. In this chapter, we choose a genetic algorithm that is an approximate algorithm to solve the problem of portfolio optimization with some realistic constraints. Novel algorithms are proposed to take this model risk into consideration.

This chapter is organized as follows. Section 2 discusses the limitations associated with Markowitz's [1] portfolio optimization model. Section 3 briefly presents the concepts of model risk. The Multi-Objective Genetic Algorithms selected for this research are described in Sect. 4. Models for stock return and volatility forecasting are described in Section 5. Section 6 describes the parameters and methodology of the experimentation. The empirical results from the experiment and their analyses are discussed in Section 7. Finally, Section 8 summarizes the findings and highlights future research.

2 Portfolio Optimization and Modern Portfolio Theory

Markowitz's seminal paper [1] founded modern portfolio theory. The theory is based on an economic assumption that economic agents are rational beings. When they make their investment decision they maximize their expected utility under budget constraints. The Markowitz Mean-Variance Model assumes that investors make their decisions of portfolio construction by choosing assets that are expected to maximize their portfolio returns at the end of an investment period. By assuming that all investors are risk averse, the simplest model with a number of unrealistic constraints namely, perfect market without taxes, no transaction costs and assets are infinitely divisible, the Markowitz portfolio optimization can be stated mathematically as follows:

$$\mathrm{Min}_{x_i}\, \sigma_p^2 \tag{10.1}$$

Subject to

$$\sigma_p^2 = \sum_i \sum_j x_i x_j \sigma_{ij}$$

$$r_p = r^*$$

$$r_p = \sum_i x_i r_i$$

$$\sum_i x_i = 1$$

$$x_i \geq 0$$

Where,

σ_{ij} is covariance between asset i and j, if $i = j$, and variance of asset i, if $i = j$.
σ_p^2 is variance of the portfolio of assets.
r_i is expected return of asset i.
r_p is a portfolio's expected return.

The Markowitz model is simplified in such a way that it can be solved for closed-form solutions.[1] Its assumptions do not represent the realistic constraints that are found in capital markets. Beside imposing some unrealistic constraints and ignoring some realistic constraints, the objective function is also considered unrealistic. Portfolio optimization's objective is to find a combination of assets that yield minimum risk at a given level of portfolio return. The risk of the portfolio given an expected return was minimized to optimize a portfolio of assets. There are two problems of definition here. First, how 'risk' is measured since the original meaning of risk is subjective and depends on individual's risk appetite. Second, how expected return be estimated since all of returns of assets are to be realized in the future. This will ultimately lead us to an important assumption of the model of whether the return in the future is deterministic or stochastic. We take the estimated returns of assets for granted and optimize the model based on them hoping that the estimations will be realized in the future. Alternatively, we assume that the returns of assets in the future are not certain but follow some stochastic rule. Another aspect of modelling is whether the investors care only for the next period or also for the series of outcomes in the future. To accommodate the investors, the objective functions need to be adjusted for single-period optimization or for multi-period optimization accordingly.

There are also some motivations to alter the objective function of the portfolio optimization problem for other purposes besides the aforementioned 'theoretical' issues. The objective function may be modified for the convenience of

[1] Closed-form solutions are solutions that can be expressed in mathematical formulas.

solving the problem, for instance, in order to reduce the complexity of computation or to make it compatible with some known solving methods or algorithms.

A crucial issue for portfolio optimization is how to select the model's inputs. There are two kinds of inputs, namely representations of risk and estimations of asset returns. These inputs must represent the 'not-yet-known-for-sure' values and are expected to be realized in the future. Optimizing the portfolio selection model with erroneous estimations of the inputs will produce incorrect combinations of assets and ultimately result in an inefficient portfolio. This is model risk that we attempt to handle.

3 The Concepts of Model Risk

A review of the existing literature shows that there are many definitions of model risk. As far as we are concerned, especially for the purpose of this research, we define model risk in a broad sense. Model risk is a risk inherent in the use of inaccurate models to make decisions in order to assist in a process of decision making and by making such decisions it leads to financial loss or to experience unexpected risk—it is the subjectivity involved in the modelling process. For other definitions, Kato and Yoshiba [2] define model risk separately in the area of pricing models and of risk measurement models. In pricing models, model risk is defined as 'the risk arising from the use of a model which cannot accurately evaluate market prices, or which is not a mainstream in the market'. In risk measurement models however, model risk is defined as 'the risk of not accurately estimating the probability of future losses' [2]. In the same paper, the sources of model risk in pricing model are described as the utilization of wrong assumptions, errors in estimations of parameters, errors resulting from discretization and error in market data. The sources of model risk in risk measurement models are the differences between assumed and actual distribution and errors in the logical framework of the model.

In another perspective, Derman [3] classifies financial models into three categories: fundamental models, phenomenological models and statistical models. Different categories of models are prone to different sources of model risk. For the fundamental models, which are mathematical models based on a set of postulates, the sources of a models risk are incorrect assumptions and wrongly identified inputs. For phenomenological models, which are based on observations of the underlying behaviours, the main source of risk is attempting to apply them beyond their validity ranges and situations. Unlike both of the aforementioned categories, which embody some elements of causality, the statistical model relies on correlation rather than causation. They are built based on variable correlation that is assumed to be stable overtime. The main source of model risk for these kinds of models is mis-specification, that is, constructing a model with incorrect variables or erroneous relationship functions and unstable correlations. The latter may look correctly specified in the sample however this may not be the replicated during the out-of-sample period

In portfolio optimization based on equation (10.1) it is crucial to forecast portfolio return and standard deviation accurately. Forecasted portfolio return is calculated from the weighted average of forecasted return of all stocks included and portfolio standard deviation is calculated from standard deviation of the stock's past returns and their correlation as stated in equation (10.1). Note that the original Markowitz's model uses means and standard deviations of asset returns as the best prediction of the future returns and volatilities of assets under the assumption that all asset returns follow a certain statistical distribution (i.e. normal distribution). Most models use forecasted stock returns and standard deviations however, their predictability may suffer from the instability of correlations overtime. This kind of model risk in turn causes the outcomes of portfolio optimization to be suboptimal. As a result, an investor will experience short falls in expected portfolio returns and/or larger than expected volatility leading to a reduced risk/return trade off.

There are a number of measures to aid in the management or mitigation of model risks. These measures are complimentary rather than substitutable. Firstly, any models used in decision making need to be reviewed by independent model controllers and reported to the management or an investment committee. Secondly, the person(s) who utilizes the models needs to be aware of the limitations of the models. Thirdly, models need to be thoroughly examined for their validity and limitations before being put to use. Fourthly, models in use should be subjected to regular reviews [2].

We propose a measure to handle model risk by embedding stock selections based on the predicted accuracy or validity of their forecasting models in the portfolio selection process. Forecasting models are mostly statistical, which often suffer from instability. If we combine models and continuously evaluate their accuracy in order to make a bias selection based on their accuracy we would be able to mitigate model risk. In our case, we have a number of substitutable assets (stocks) that have different expected returns, expected volatilities and expected forecasting capabilities as a measure of model risk. As we incorporate the minimization of model risk into the portfolio selection process, we are potentially managing many of the aforementioned model risks.

4 Multi-Objective Genetic Algorithms for Portfolio Optimization

The Multi-Objective Genetic Algorithm (MOGA) used for portfolio optimization here is based on the algorithm initially proposed by Fonseca and Fleming in 1993 [4]. MOGA relies on Pareto rankings to assign the smallest ranking value to all non-dominated individuals. On the other hand, for those of dominated individuals, they are ranked by how many individuals in the population actually dominate them. Thus, the raw fitness of an individual is an inverse function of its Pareto rank. MOGA for two objectives portfolio optimization is tasked to rank individuals in the population by both portfolio return (to maximize) and portfolio standard deviation (to minimize). Mathematically, the two objectives can be stated in equation 10.2.

$$\text{Max}: \; r_p = \sum_i^N x_i r_i \tag{10.2}$$

$$\text{Min}: \qquad \sigma_p = \sum_i^N \sum_j^N x_i x_j \rho_{ij} \sigma_i \sigma_j \tag{10.3}$$

where,

x_i is a proportion of the asset i in the portfolio of assets.
x_j is a proportion of the asset j in the portfolio of assets.
ρ_{ij} is the correlation coefficient between asset i and j.
σ_p is expected standard deviation of the portfolio of assets.
σ_i is expected or forecasted standard deviation of asset i.
σ_j is expected or forecasted standard deviation of asset j.
r_p is the expected return of the portfolio of assets.
r_i is the expected or forecasted return of asset i.

For the purpose of handling model risk from forecasting asset returns and standard deviations, we add a third objective to accompany the original two-objectives MOGA to rebrand the model as the MOGA3O model. The third objective is based on equations 10.2 and 10.3 and can be stated as follows:

$$\text{Min}: \qquad SRE_p = \sum_i^N x_i \left(\frac{\partial ESR_p}{\partial x_i} \right)^2 \tag{10.4}$$

where,

SRE_p is the approximated Sharpe ratio error of the portfolio of assets resulting from inclusion of all of the assets.

$\dfrac{\partial ESR_p}{\partial x_i}$ is the approximated impact to portfolio Sharpe ratio error of inclusion of asset i into the portfolio. The term is squared to eliminate the sign (see Sect. 5).

In order to distribute the individuals in the population evenly along the Pareto front, the overall fitness function is then adjusted by the sum of sharing distance. The sharing distance between individuals i and j is given by:

$$SF_{ij} = 1 - \frac{d\left(x_i, x_j\right)}{\sigma_{\text{share}}}, \quad \text{if} \quad d\left(x_i, x_j\right) < \sigma_{share}$$
$$SF_{ij} = 0, \qquad \text{if} \quad d\left(x_i, x_j\right) \geq \sigma_{share} \tag{10.5}$$

Where,

$d(x_i, x_j)$ is a metric distance between two individuals in objective domain,

σ_{share} is a predefined sharing distance.

And, the overall fitness is defined by

$$F_i = \frac{Fit(i)}{\sum_j SF_{ij}} \tag{10.6}$$

Where, Fit(i) is the inverse of Pareto rank (i) (1/rank (i) in this setting).

The overall fitness values of individuals are to be used in the probabilistic selection process by comparing each individual's overall fitness to that of the individual that has maximum overall fitness. Each of individual comparative fitness values then compare with a random number. If they exceed the random number, the individual will be selected (roulette selection method) [4]. MOGA usually has O (n^2) for a single round, because it needs to compute Pareto ranks and also to calculate the sharing distance for all individuals (based on equation 10.5). The sharing distance is a measure of degree of difference of an individual to other individuals by Euclidian distance. There may be individuals that have the same or similar profiles (in this case, return and volatility). The algorithm is prone to select individuals that have distinct profiles that have the same or a similar fitness to make the boundary, or selected set, of individuals with evenly distributed solutions.

The problem is represented by hybrid encoding [5, 6]. A pair of genetic strings stands for a particular portfolio (an individual of population). The binary value string represents which stocks (or assets) are included in portfolio (0 stands for not included and 1 stands for included.) The real value string represents the weights of particular stocks in the portfolio. The lengths of both strings are equal to the number of stocks in the selection set (stock of interest.) The strings are generated with their real elements normalized in such a way that the summation of all elements of each combined string is always one. Before a round of repair process begins both strings are combined by a scalar product of the binary string and the real value string. Once this is complete then the repair process ends, the combined string is then normalized to ensure that the summation of all elements is one. Finally, the combined string separates into the new, normalized binary string and the real value string prior to crossing over and mutation operations (see Fig. 10.1).

Crossover and mutation operations are performed independently for both strings. However, before evaluation, both strings need to be combined so that the objective values can be calculated. The crossover operation for all Generic Algorithms (Gas) in this setting is a three-point crossover by randomly selecting three points for each string independently. The mutation operation for all algorithms in this study is one-point mutation by randomly selecting the mutation point. For the binary strings, the mutation is a flip-flop mutation by changing from 1 to 0 and 0 to 1 respectively. For those of real value strings, mutation points are added by random numbers (between 0 and 1) multiplying each by 0.1.

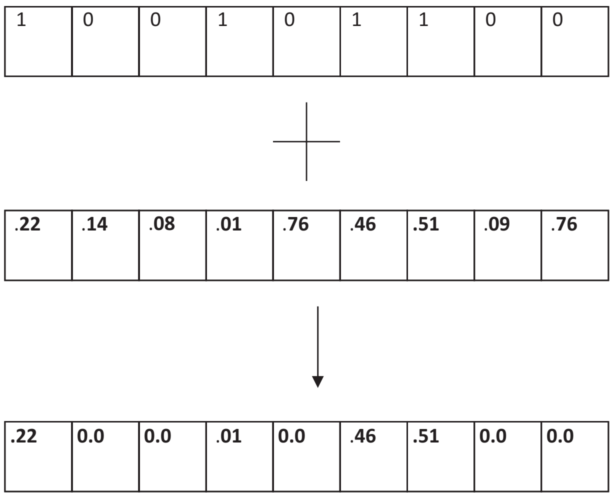


Fig. 10.1 Problem representation: binary string, real value string and combined string

All constraints are handled through a repair algorithm. The algorithm is proposed and used by Fonesca and Flemming and Streichert et al. [5–7]. The constraints in this setting are: unity constraint (the sum of weights must be equal to 1), cardinality constraint, floor (buy-in) constraints and round-lot constraint.

The Markowitz model is a simplified model to focus only on a theoretical point of view. In the real world of investment management, portfolio managers face a number of realistic constraints such as those which arise from normal market practices, practical matters and industry regulation. The realistic constraints, which are of practical importance, include integer constraints, cardinality constraints, floor and ceiling constraints, turnover constraints, trading constraints, buy-in threshold and the inclusion of transaction costs. Integer constraints or round-lot constraints impose the number of units of an asset included in the portfolio to be at least a certain round number imposed by stock exchanges (i.e. there cannot be an odd number of normal trading lots). This may not experience GA (genetic algorithm) optimizations because they are combinatorial but experience traditional optimization methods, which require continuity of input variables. The integer constraint (or round-lot constraint) can be expressed as

$$x_i = \frac{n_i}{\sum_{i=1}^{N} n_i} \tag{10.7}$$

$$n_i \bmod l_i = 0 \tag{10.8}$$

where, n_i is the number of unit of asset (share) and l_i is trading lot of the asset i.

Cardinality constraints are the maximum and the minimum number of assets that a portfolio manager wishes to include in the portfolio due to monitoring, diversification or transaction cost control reasons [8]. The constraints can be expressed as follows:

$$C_l \le \sum_{i=1}^{n} b_i \le C_u, \tag{10.9}$$

where, $b_i = 1$ if $x_i > 0$, else $b_i = 0$. C_l and C_u are the lowest number of assets and the highest number of assets that may be included in a portfolio respectively.

Floor and ceiling constraints define lower and upper limits on proportions of assets, which can be held in a portfolio. These constraints may result from institutional policies in order to diversify portfolio or to rule out negligible holding of assets for the ease of control [9]. The constraints can be expressed mathematically as follows:

$$f_i \leq x_i \leq c_i \qquad \forall i \tag{10.10}$$

where, f_i and c_i are the lowest proportion and the highest proportion that asset i can be held in the portfolio respectively.

The repair algorithm is used to ensure that all individual representations of portfolios comply with the aforementioned constraints. It firstly handles the cardinality constraints by setting smaller (S K) values (from S values) of the combined string to zero, where S is the number of selectable stocks (equal to the length of the strings) and K is the maximum number of assets (stocks in this case) permitted in a portfolio (cardinality constraint.) Then, it manages the floor constraint (buy-in threshold) by setting stocks whose weights are below the buy-in threshold to zero. Next, it normalizes those remaining non-zero weights to make all weights sum up to 1 by setting $w_i' = l_i + (w_i - l_i)/\Sigma(w_i - l_i)$, where w_i is a non-zero weight of stock i and l_i is the buy-in threshold (the minimum weight amount that can be purchased) for stock i. Following this, the round-lot constraints are handled by rounding all non-zero weights to the next round-lot level such that $w_i'' = w_i' - (w_i' \bmod c_i)$, where, c_i is the smallest volume that can be normally purchased from the stock market for stock i. Finally, the remainder from the rounding process $(\Sigma w_i' \bmod c_i)$ is allocated in quantity of c_i to w_i'', which has the biggest value of $(w_i' \bmod c_i)$ until all of the remainder is depleted.

All pairs of strings are firstly filled with a random number, so, they need to be repaired by the repair algorithm. Since crossover and mutation operations alter sequences of the strings and may make them not comply with the imposing constraints, the repair algorithm needs to be applied again to preserve the conformation of portfolios (represented by GA strings) to the constraints before evaluation and selection stages.

There are also a number of portfolio constraints not addressed here, for example, ethical constraints or blacklist constraints, which can be handled by setting weights of the appropriate stock to zero.

5 A Portfolio's Sharpe Ratio Error

The Sharpe ratio is a metric that is used to measure an asset's or portfolio's returns adjusted by risk. Risk is measured by standard deviation of assets' or portfolio's return [10]. The Sharpe ratio is defined as:

$$SR_p = \frac{R_p - R_f}{\sigma_p} \tag{10.11}$$

Where,

SR_p is Sharpe ratio of a portfolio,
R_p is return of the portfolio,
σ_p is standard deviation of the portfolio.

The Sharpe ratio is essentially a standardized excess return above risk free rate per unit of absolute risk, as measured by standard deviation. In a sense, we cannot compare any returns of assets or of portfolio directly because they may have different levels of inherent risk. However, the Sharpe ratio adjusts the level of risk so that it represents a standardized return that enables different assets and portfolios to be compared. If we need to summarily compare two assets or two portfolios the Sharpe ratio is one of the better metrics to achieve this. When comparing two portfolios, the portfolio with the larger Sharpe ratio is the portfolio with the best risk return trade off.

Our objective is to find a portfolio that can handle model risk effectively. The proposed algorithm requires a proof that it renders better portfolios, that is, a portfolio with a larger Sharpe ratio for more of the time. Portfolio optimization is deployed to construct a portfolio that satisfies two objectives simultaneously, for instance, minimization of risk (as measured by the portfolio's standard deviation) and maximization of the portfolio's return. This means maximizing the Sharpe ratio at a given level of returns or at a given level of risk. Consider two slightly different portfolios that hold marginally different lots of the same stock i: the difference in the Sharpe ratio is given by:

$$ESR_p = \left(\frac{R_p' - R_f}{\sigma_p'}\right) - \left(\frac{R_p - R_f}{\sigma_p}\right) \tag{10.12}$$

$$\begin{aligned}\frac{\partial ESR_p}{\partial x_i} = &\left(\frac{1}{\sigma_p'}\frac{\partial R_p'}{\partial x_i} - \frac{R_p'}{\sigma_p'^2}\frac{\partial \sigma_p'}{\partial x_i} + \frac{R_f}{\sigma_p'^2}\frac{\partial \sigma_p'}{\partial x_i}\right) \\ &-\left(\frac{1}{\sigma_p}\frac{\partial R_p}{\partial x_i} - \frac{R_p}{\sigma_p^2}\frac{\partial \sigma_p}{\partial x_i} + \frac{R_f}{\sigma_p'^2}\frac{\partial \sigma_p}{\partial x_i}\right)\end{aligned} \tag{10.13}$$

For simplicity, let assume that $R_f = 0$.

$$\frac{\partial ESR_p}{\partial x_i} = \left(\frac{1}{\sigma_p'}\frac{\partial R_p'}{\partial x_i} - \frac{R_p'}{\sigma_p'^2}\frac{\partial \sigma_p'}{\partial x_i}\right) - \left(\frac{1}{\sigma_p}\frac{\partial R_p}{\partial x_i} - \frac{R_p}{\sigma_p^2}\frac{\partial \sigma_p}{\partial x_i}\right) \tag{10.14}$$

Let us consider,

$$\frac{\partial R_p'}{\partial x_i} = \frac{\partial \sum x_i R_i'}{\partial x_i} = R_i' \tag{10.15}$$

$$\frac{\partial R_p}{\partial x_i} = \frac{\partial \sum x_i R_i}{\partial x_i} = R_i \tag{10.16}$$

$$\frac{\partial \sigma_p'}{\partial x_i} = \frac{\partial \sum_i \sum_j x_i x_j \rho_{ij} \sigma_i' \sigma_j'}{\partial x_j} = \frac{1}{2\sqrt{\sigma_p'}} \left(\sum_{j \neq i} x_j \rho_{ij} \sigma_i' \sigma_j' + 2 x_i \sigma_i'^2 \right) \tag{10.17}$$

$$\frac{\partial \sigma_p}{\partial x_i} = \frac{\partial \sum_i \sum_j x_i x_j \rho_{ij} \sigma_i \sigma_j}{\partial x_j} = \frac{1}{2\sqrt{\sigma_p}} \left(\sum_{j \neq i} x_j \rho_{ij} \sigma_i \sigma_j + 2 x_i \sigma_i^2 \right) \tag{10.18}$$

Substitute (10.15), (10.16), (10.17) and (10.18) in (10.14),

$$\begin{aligned} \frac{\partial ESR_p}{\partial x_i} = & \left(\frac{R_i'}{\sigma_p'} - \frac{R_p'}{2\sigma_p'^{5/2}} \left(\sum_{j \neq \mathrm{i}} x_j \rho_{ij} \sigma_i' \sigma_j' + 2 x_i \sigma_i'^2 \right) \right) \\ & - \left(\frac{R_i}{\sigma_p} - \frac{R_p}{2\sigma_p^{5/2}} \left(\sum_{j \neq i} x_j \rho_{ij} \sigma_i \sigma_j + 2 x_i \sigma_i^2 \right) \right) \end{aligned} \tag{10.19}$$

By rearranging terms, we have got

$$\begin{aligned} \frac{\partial ESR_p}{\partial x_i} = & \left(\frac{R_i'}{\sigma_p'} - \frac{R_i}{\sigma_p} \right) + \left(\frac{R_p}{\sigma_p^{5/2}} x_i \sigma_i^2 - \frac{R_p'}{\sigma_p'^{5/2}} x_i \sigma_i'^2 \right) \\ & + \left\{ \frac{R_p}{2\sigma_p^{5/2}} \left(\sum_{j \neq i} x_j \rho_{ij} \sigma_i \sigma_j \right) - \frac{R_\mathrm{p}'}{2\sigma_p'^{5/2}} \left(\sum_{j \neq i} x_j \rho_{ij} \sigma_i' \sigma_j' \right) \right\} \end{aligned} \tag{10.20}$$

Let us consider equation (10.10), according to the assumptions, the original portfolio is optimal and the second portfolio is only infinitesimal changed from the original, thus we can approximate that

$$R_p^{'} \approx R_p \tag{10.21}$$

And,

$$\sigma_p^{'} \approx \sigma_p \tag{10.22}$$

The third term (inside the {} brackets) represents correlations of stocks' returns. Since we assume that correlations between pair of assets are the same from both cases, thus we can assume that:

$$\sum_{j \neq i} x_j \rho_{ij} \sigma_i \sigma_j \approx \sum_{j \neq i} x_j \rho_{ij} \sigma_i' \sigma_j' \tag{10.23}$$

From the equations (10.21) and (10.22), the third term is eliminated and substitute both equations into (10.20), we have:

$$\frac{\partial ESR_p}{\partial x_i} = \frac{R_p}{\sigma_p^{5/2}} x_i \sigma_i^2 - \frac{R_p^{'}}{\sigma_p^{'5/2}} x_i \sigma_i^{'2} \tag{10.24}$$

Equation (10.24) is used for the third objective for MOGA by substituting in equation (10.4).

6 Stock Forecasting Models

We used the Box-Jenkins Methodology to select models for forecasting Stock returns and standard deviations [11]. The methodology uses the series auto-correlation function (ACF) and the partial correlation function (PCF) as guidelines to decide how many lags of both autocorrelation (AR) and moving average (MA) should be included into the forecasting equations. In response to this the models are built, estimated and statistically tested. We looked at the significant tests of the AR and MA coefficients to finally determine the lag structure of the equations. If we fail to find an ARMA equation either there are no ACF and PCF indications or there is no statistical significance of any lag coefficients, we then try a GRACH structure [12]. If all else fails, we will resort to the use of its mean as the estimated value.

Analysis is performed both graphically (scrutinize ACFs and PCFs) and statistically (run regressions and examine the significant tests) from the monthly time-series of stock returns and volatilities from 1980 to 1999 (240 months.) The stock series are originally closing prices adjusted for corporate actions (such as dividends and stock splits). Monthly returns can be calculated from the closing price on the last day of the month however monthly volatilities as represented by standard deviations of the period that are calculated from daily closing prices. The stock series consists of 17 stock series, namely, Alcoa Inc. (AA), Boeing Company (BA), Caterpillar Inc. (CAT), E.I. du Pont de Nemours (DD), Walt Disney Company (DIS), General Electric Company (GE), General Motors Corporation (GM), Honeywell International Inc. (HON), Hewlett-Packard Company (HPQ), International Business Machines (IBM), Johnson & Johnson (JNJ), Coca-Cola Company (KO), 3M Company (MMM), Merck & Company Inc. (MRK), Procter & Gamble Company (PG), United Technologies Corporation (UTX) and Exxon Mobil Corporation (XOM.) The suffix 'RTN' represents the stock return series and the suffix 'S' represents stock return volatilities. We chose these stocks because they are all long-standing US companies and have a long-standing record on the NYSE. Our experiment requires an adequate amount of historical data in order to estimate forecasting models. This data is used to train the model selection algorithm and also to evaluate portfolio optimization models. Therefore our model only considers stocks with 26 years of historical data and for this experiment this is our only criteria for stock selection. However, the experiment can be expanded to any set of stocks, an entire market or just blue-chip stocks. Hence, the algorithm can be adapted to apply to any number of stocks and universes given that the investors and the computer can handle such complexity.

7 The Experiment

We conducted an experiment to find out whether the modification of MOGA described in Sect. 4, to include the portfolio Sharpe ratio error optimization as a third objective improves the outcome of actual portfolios in terms of the portfolios Sharpe ratio. We compared the actual portfolio in the following three cases, MOGA with two objectives (forecasted portfolio return and forecasted portfolio standard deviation) using stock forecasting models according to Table 10.1, MOGA with two objectives using stock mean returns and standard deviations as stock forecasting values and the finally our proposed MOGAO3 model with three objectives (forecasted portfolio return, forecasted portfolio standard deviation and estimated portfolio Sharpe ratio error as

Table 10.1 Forecasting models for stock returns and volatilities

Stock returns series		Stock volatility series	
Series name	Selected model	Series name	Selected model
AA_RTN	GARCH-M(1,1)	AA_S	ARMA(3,3)
BA_RTN	Mean	BA_S	ARMA(3,3)
CAT_RTN	EGARCH-M(1,1)	CAT_S	ARMA(3,3)
DD_RTN	GARCH-M(1,1)	DD_S	ARMA(3,3)
DIS_RTN	ARMA(1,1)	DIS_S	ARMA(3,3)
GE_RTN	GARCH-M(1,1)	GE_S	ARMA(3,3)
GM_RTN	Mean	GM_S	ARMA(2,2)
HON_RTN	Mean	HON_S	ARMA(1,1)
HPQ_RTN	Mean	HPQ_S	ARMA(3,3)
IBM_RTN	Mean	IBM_S	ARMA(3,3)
JNJ_RTN	GARCH-M(1,1)	JNJ_S	ARMA(1,1)
KO_RTN	Mean	KO_S	ARMA(1,1)
MMM_RTN	Mean	MMM_S	ARMA(1,1)
MRK_RTN	Mean	MRK_S	ARMA(3,3)
PG_RTN	ARMA(1,1)	PG_S	ARMA(1,1)
UTX_RTN	ARMA(1,1)	UTX_S	AR(1)
XOM_RTN	ARMA(1,1)	XOM_S	ARMA(1,1)

(Eq. 10.24)). The three objective Pareto ranking in Java code is shown Fig. 10.3. We ran and evaluated the outcomes of 72 periods from 2001 (January) to 2006 (December), ten times for each period. All MOGAs had 400 as a population size and with 1000 generations. We also used 20 years of historical monthly data (240 observations) to estimate the stock forecasting models and calculate the sample stock returns and their standard deviations. We calculated the third objective by using equation (10.24). It is an average of the differences of 12 out-of-sample forecasted portfolio Sharpe ratios and the actual portfolio Sharpe ratios. For example, in the first year of 2001, we used 240 observations from 1980 to 1999 to estimate the forecasting models (or train the models). We then used the estimated models to forecast stock returns and deviations for the following 12 months of 2000 and then used the forecasted values with the actual values of the same periods to calculate the estimated stock portfolio Sharpe ratio errors (squares of the numerical results of equation (10.14)). The aforementioned estimated models and estimated portfolio Sharpe ratio errors are used for portfolio selections for the next 12 months of 2001 The forecasting models used actual month by month observations while the estimated stock portfolio Sharpe ratios error remained the same for all 12-month periods. For the next year, the window of observations will shift for one year to 1981–2000 so that the estimation of the forecasting models moves to 2001 in order to estimate Sharpe ratio errors for the portfolio, and then to 2002 for portfolio selections and evaluations. The windows of observations rolled on every year until 2006 the final year of our sample period.

Note that for simplicity we set the portfolio return to the average return of S&P 500 from 1980 to 1999 (as represent the market return) and set the portfolio standard deviation to the standard deviation of S&P 500 returns from 1980 to 1999. The values are 0.011874 and 0.043027 accordingly and will remain the same for all later periods. Again, for simplicity and estimation purposes, since the proportion of each stock in portfolio varies, we assume that the optimal portfolio consists of equal proportion of 17 stocks, thus for each x_i, x_i is 1/17 or 0.0588.

8 Empirical Results and Analyses

Table 10.2 shows the average of portfolio Sharpe ratios from selected portfolios developed using the three different algorithms. Namely, MOGA with 2 objectives (portfolio returns and portfolio standard deviation) using the stock forecasting models according to Table 10.1 (MOGA), MOGA with 2 objectives using means and standard deviations of stocks as forecasting values (MOGA_MEAN), and MOGA with 3 objectives (MOGA3O) in which the third objective, the estimated portfolio Sharpe ratio error, is incorporated

Figure 10.2 shows graphically the average portfolio Sharpe ratios from Table 10.1. The horizontal axis is the number of periods (1–72) and the vertical axis is the value of the portfolio Sharpe ratio. We can see that MOGA_MEAN's are mostly the lowest all of the three algorithms. While MOGA3O's and MOGA's are similar to each other however MOGA3O's are slightly higher. When we compare only those of MOGA3O and MOGA using area graphs as shown in Fig. 10.5, it is apparent that MOGA3O's are higher than MOGA's over the 72 periods hence the lighter shaded area mostly occludes the darker shade. We found that MOGAs are likely to outperform MOGA3Os when the outcomes of portfolio Shape ratio are high, that is, in bull markets. Since MOGAs stick to a single forecasting model for each stock return and each volatility that performs best for a particular stock, they are not adapted to economic and market environmental changes, therefore, only 175 out of 720 samples that MOGAs' realized Sharpe ratio are more than those of MOGA3Os (Figs. 10.3, 10.4, 10.5).

We want to determine from the results whether handling model risk by incorporating the third objective, the approximation of portfolio Sharpe ratio error, does indeed improve the actual outcomes of portfolios at the end of the investment periods. We have the results from 72 periods, ten samples for each period, thus 720 data points in total.

Table 10.2 Average portfolio Sharpe ratios from the outcomes of the different algorithms

P#	MOGA	MOGA_MEAN	MOGA3O	P#	MOGA	MOGA_MEAN	MOGA3O
1	2.651694	3.002966058	3.80421191	37	16.50464	−0.977052306	15.9592374
2	1.599297	−15.30412667	2.90921088	38	4.920148	−5.622387378	10.9416912
3	11.72089	9.737898982	12.5398981	39	16.26268	−0.485199681	17.8465377
4	8.077454	1.468998179	8.44886132	40	16.91984	0.435208791	17.1918426
5	9.055536	−10.04106948	10.6681083	41	26.90916	13.97791692	29.545814
6	13.41024	5.942963992	14.7826241	42	4.727717	−20.80481464	9.97438897
7	3.777822	−5.490921177	6.11300114	43	21.62514	12.03720491	23.4815468
8	−1.22276	−6.710955757	2.51333237	44	11.23973	−2.456988529	8.31890222
9	3.22673	1.651728232	2.07981293	45	30.59138	−5.363379776	34.9955662
10	16.52173	8.236407078	16.3208842	46	27.47407	21.82218814	29.2644133
11	13.82264	7.165711399	13.062293	47	28.28008	11.00086483	50.2561757
12	5.227387	−6.327691162	12.6721137	48	5.730111	−1.082302826	9.98593583
13	16.58808	15.55835688	16.9655906	49	47.78343	39.85016557	64.52086
14	22.05061	14.72102276	19.8126013	50	15.25964	−2.902600778	17.706636
15	11.00846	2.118459466	12.7090738	51	−15.1335	−12.07221998	21.0945988
16	8.917686	−4.688920043	11.0800128	52	16.40335	2.179405446	17.1205299
17	−0.77115	−17.03096526	4.38403375	53	−0.49858	−7.944061709	7.47265331
18	0.796028	−4.793950273	5.28216702	54	38.84528	1.98534631	45.9040463
19	1.613768	−0.981061023	1.7787698	55	17.82275	2.336407419	19.4503365
20	−6.37475	−8.354406996	−3.47696226	56	25.64921	9.600937271	29.7207576
21	17.91655	6.707948579	18.009378	57	18.78099	−21.76925019	23.510262
22	7.397799	0.136132173	7.62045694	58	20.05064	7.094870556	20.6078077
23	−3.21423	−5.503765875	−3.01929848	59	13.10278	−9.604570324	15.9180695
24	3.47741	−9.245737519	5.14168637	60	33.59123	11.38431056	34.5891038
25	2.741908	−3.230165085	3.58539562	61	17.31738	10.60068244	18.2345987

(*continued*)

Table 10.2 (continued)

P#	MOGA	MOGA_MEAN	MOGA3O	P#	MOGA	MOGA_MEAN	MOGA3O
26	19.85944	16.82725464	19.463751	62	23.10565	0.669988332	23.5894884
27	10.33042	3.412619621	11.1333695	63	45.21119	16.21852386	50.1093051
28	20.00226	3.274277358	21.0129769	64	14.26807	−4.275074836	9.71341146
29	9.889057	1.706191899	11.7125052	65	6.768207	−1.563089795	12.6601002
30	24.94557	7.709092685	27.5060266	66	26.03123	0.722956158	16.5820101
31	29.44655	1.051722096	28.5665364	67	32.00274	7.238998374	26.0553043
32	18.99602	−3.638976901	17.6767992	68	44.13047	5.240865724	45.5185342
33	36.82491	15.79729071	41.6978745	69	57.69881	32.55427308	50.8725289
34	6.468593	1.736355272	6.8218802	70	23.89759	2.48221826	24.1934988
35	97.60954	44.55992593	74.98473	71	54.7616	10.10853516	45.4241925
36	36.95656	−12.633463	27.4243178	72	30.75154	7.994603544	32.3795528

```
Initialize generation counter: n = 0
Create a population, Pop.
Repeat while stopping criteria is not met (n=N).
Normalise Pop.
    Evaluate Pop for all objective function values for all F(i).
    Evaluate Pareto Rank for Pop.
    Evaluate Sharing Distance for Pop.
    Evaluate Sharing Fit for Pop.
    Probabilistically Select Pop to half as ¼ Pop(1) and ¼  Pop(2).
    Crossover Pop(1) and Pop(2) to Create CPop
    Pop = Pop(1) + Pop(2) + CPop.
Probabilistically Mutate Pop.
    Shuffle the individual sequence in Pop.
n = n + 1.
End Repeat
Evaluate Pop for all objective function values for all F(i).
Return (Pop, all F(i),…)
```

Fig. 10.2 Code of Multi-Objective Genetic Algorithm (MOGA)

```
// compute Pareto Rank for each port
for(Portfolio ith: pop) {     // for ith
int q = 1;
for(Portfolio jth: pop) {     // for jth
          if( (ith.getPortYield() <jth.getPortYield()) && (ith.getPortStd() >jth.getPortStd()
          &&(ith.getPortPSRError() >jth.getPortPSRError()))||
          (ith.getPortYield() <jth.getPortYield()) && (ith.getPortStd() >jth.getPortStd()  &&
     (ith.getPortPSRError() == jth.getPortPSRError()))||
          (ith.getPortYield() <jth.getPortYield()) && (ith.getPortStd() == jth.getPortStd()
          && (ith.getPortPSRError() >jth.getPortPSRError()))||
          (ith.getPortYield() == jth.getPortYield()) && (ith.getPortStd() >jth.getPortStd()
          && (ith.getPortPSRError() >jth.getPortPSRError()))||
          (ith.getPortYield() == jth.getPortYield()) && (ith.getPortStd() == jth.getPortStd()
          && (ith.getPortPSRError() >jth.getPortPSRError()))||
          (ith.getPortYield() == jth.getPortYield()) && (ith.getPortStd() >jth.getPortStd()
          && (ith.getPortPSRError() == jth.getPortPSRError()))||
          (ith.getPortYield() <jth.getPortYield()) && (ith.getPortStd() == jth.getPortStd() &&
          (ith.getPortPSRError() == jth.getPortPSRError())))
   q++;
   }    // for jth
ith.setParetoRank(q);
ith.setFitness((1/(double)q));
```

Fig. 10.3 Pareto ranking in Java code (3 objective MOGA)

In order to determine this we conducted two statistical tests, namely testing the difference between two means and testing hypotheses about a proportion [13]. Testing of the difference between two means is to identify whether mean values from the two sample groups, which are represented by two different populations, are equals or statistically significantly different within a given confidence level. We needed to test whether mean of MOGA3O's Sharpe ratios is statistically more than those of MOGA and MOGA_MEAN. The appropriate significant test is:

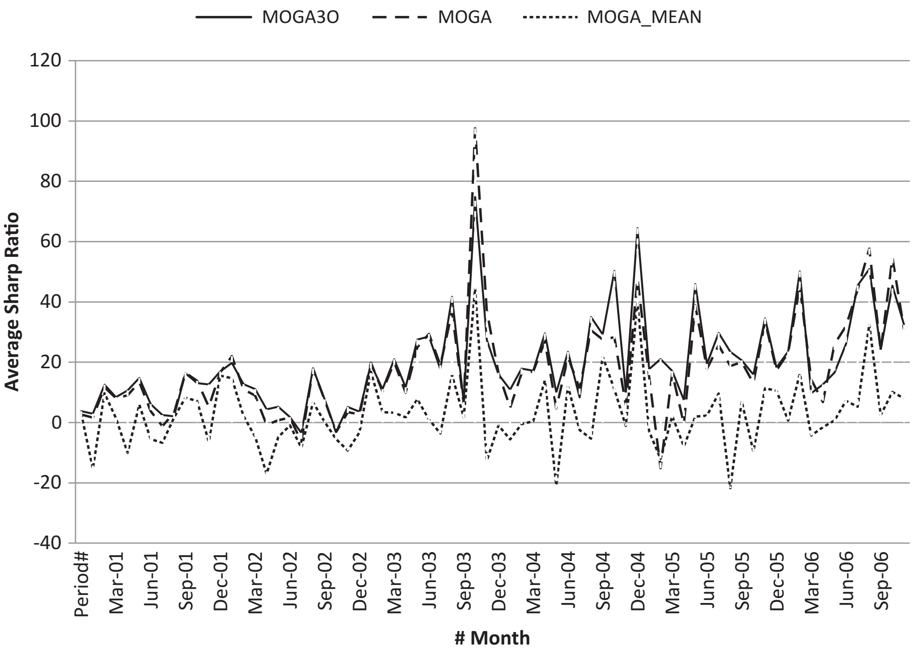


Fig. 10.4 Average portfolio Sharpe ratios of the outcomes of different algorithms (monthly periods from 2001 to 2006)

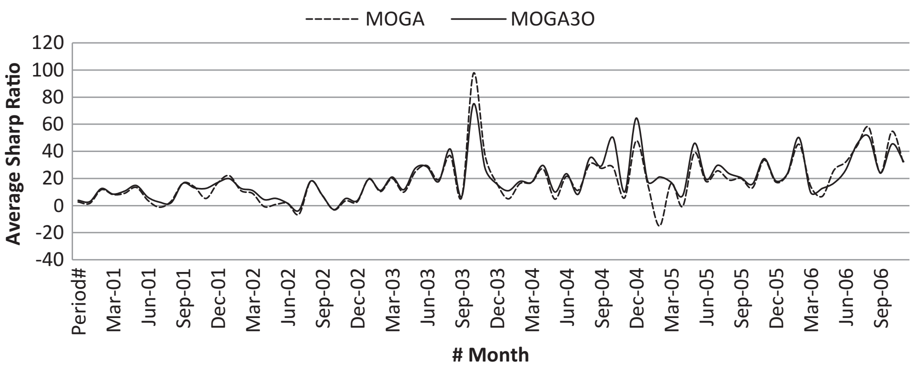


Fig. 10.5 Average portfolio Sharpe ratios of the outcomes of MOGA3O and MOGA (*Y* axis shows Sharpe ratio values and *X* axis shows monthly period from 2001 to 2006)

$$H_0 : \mu_1 \leq \mu_2$$
$$H_1 : \mu_1 > \mu_2$$
$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{10.25}$$

Where,

t	is student's t score of the statistical test.
$\overline{x}_1$	is the mean of the first population or sample (in our case are those of MOGA3O).
$\overline{x}_2$	is the mean of the second population or sample (MOGA or MOGA_MEAN respectively).
s_1^2	is the standard deviation of the population or sample of the first population.
s_2^2	is the standard deviation of the population or sample of the second population.
n_1	is the number of the first population.
n_2	is the number of the second population.

For the testing of the hypotheses about a proportion, we needed to know that, at a given level of confidence, whether the outcomes for the MOGA3O algorithm produced better results more frequently than MOGA and MOGA_MEAN. If the outcomes of MOGA3O are not better for most of the time the proportion of samples in which MOGA3O beats MOGA or MOGA_MEAN would be ranged from less than, to around 50:50 or 0.5, otherwise the proportion would be significantly greater than 0.5. These tests do not concern the means of portfolio Sharpe ratio values but instead the number of periods where MOGA3O's outcomes are better than those of MOGA or MOGA_MEAN. The appropriate significant test is as follows:

$$H_0 : p '' \ p_0$$
$$H_1 : p > p_0$$

Table 10.3 Summary of important statistical values for the testing of different between two means and the testing hypothesis about a proportion

Algorithms	n	$\bar{x}$	s^2	> MOGA3O	> MOGA	> MOGA_MEAN
MOGA3O	720	19.7846981	258.202959	0 (0.0 %)	545 (75.69 %)	717 (99.58 %)
MOGA	720	18.0573914	291.832071	175 (24.3 %)	0 (0.0 %)	702 (97.50 %)
MOGA_MEAN	720	2.62723092	137.342731	3 (0.42 %)	18 (2.50 %)	0 (0.0 %)

Table 10.4 Results of two tests whether MOGA3O's outcomes are better than those of MOGA

MOGA3O > MOGA	t/z	Degree of freedom	One-tail prob.	Reject Ho (at % level)
$H_0:\ p \leq p_0$ $H_1:\ p > p_0$	z = 13.7867	N/A	1.53229E-43	Yes (0.1 % level)
$H_0:\ \mu_1 \leq\ \mu_2$ $H_1:\ \mu_1 > \mu_2$	t = 1.97624	719	0.02425	Yes (5 % level)

Table 10.5 Results of two tests whether MOGA3O's outcomes are better than those of MOGA_MEAN

MOGA3O > MOGA_MEAN	t/z	Degree of freedom	One-tail prob.	Reject Ho (at % level)
$H_0:\ p \leq p_0$ $H_1:\ p > p_0$	z = 26.60903	N/A	2.6683E-156	Yes (0.1 % level)
$H_0:\ \mu_1 \leq\ \mu_2$ $H_1:\ \mu_1 > \mu_2$	t = 23.14841	719	2.55154E-89	Yes (0.1 % level)

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0\left(1 - p_0\right)}{n}}} \tag{10.26}$$

Where,

z is z score of the statistical test.

$\bar{p}$ is the sample estimate of population proportion (in our case, is that of MOGA3O beating).

p_0 is the hypothetical value to test against (in our case, 0.5).

n is the number of the second population.

Table 10.3 summarizes important statistics for conducting tests according to equations (10.15) and (10.16). We can see that the outcomes of MOGA3O are

Table 10.6 Results of two tests whether MOGA's outcomes are better than those of MOGA_MEAN

MOGA > MOGA_MEAN	t/z	Degree of freedom	One-tail prob.	Reject Ho (at % level)
$H_0: p \le p_0$ $H_1: p > p_0$	$z = 25.4912$	N/A	2.6683E-156	Yes (0.1 % level)
$H_0: \mu_1 \le \mu_2$ $H_1: \mu_1 > \mu_2$	$t = 20.1982$	719	1.67465E-72	Yes (0.1 % level)

comparatively better than those of MOGA and MOGA_MEAN. MOGA3O's Sharpe ratios have an average of 19.7846981 while both of MOGA and MOGA_MEAN are 18.0573914 and 2.62723092 respectively.

Table 10.4 concludes the results of the two tests in equations (10.15) and (10.16) that compare the outcomes of MOGA3O and MOGA. For the tests of hypotheses about a proportion described in the first row, the calculated z-score is 13.7867. With reference to the normal distribution table, the null hypothesis of the proportion of the outcomes of MOGA3O are worse or equal to those of MOGA is rejected at the 0.1 % level of confidence (with the probability that we wrongly rejected the hypothesis less than 1 out of 1000 times). Thus, we can safely conclude that the outcomes of MOGA3O are indeed better than those of MOGA. For the test of the difference between two means, the t-score is 1.97624. Referencing to the t-score probability table with 719 degree of freedom (N–1), the null hypothesis of the means of Sharpe ratio of the outcomes of MOGA3O is less than or equal to those of MOGA is rejected at 5 % level of confidence. For instance, we can wrongly reject the null hypothesis 5 times out of 100. Therefore, we are somewhat confident to say that on average the Sharpe ratio of the outcomes of MOGA3O is better than those of MOGA.

Table 10.5 shows the results of the two tests of equation (10.15) and (10.16) comparing the outcomes of MOGA3O and of MOGA_MEAN. For the test hypotheses about a proportion described in the first row, the calculated z-score is 26.60903. Referencing to the normal distribution table, the null hypothesis of the proportion of the outcomes of MOGA3O are worse or equal to those of MOGA_MEAN is rejected at the 0.1 % level of confidence. Therefore, we safely conclude that the outcomes of MOGA3O, are indeed better than those of MOGA_MEAN for most of the time both in bull and bear market situations. Only 3 out of 720 samples that MOGA_MEAN's Sharpe ratios are better than those of MOGA3O's for the test of the difference between two means, the t-score is 23.14841. Referencing to the t-score probability table with 719 degree of freedom (N–1), the null hypothesis of the means of Sharpe ratio of the outcomes of MOGA3O is less than or equal to those of MOGA is rejected at 0.1 % level of confidence, that is, we can wrongly rejected the null hypothesis once out of 1000. Thus, we can statistically say that in 999 cases out of 1000 the Sharpe ratios of the outcomes of MOGA3O are better than those of MOGA_MEAN.

Table 10.6 concludes the results of the two tests comparing the outcomes of MOGA and of MOGA_MEAN. For the test hypotheses about a proportion described in the first row, the calculated *z*-score is 25.4912. With reference to the normal distribution table, the null hypothesis of the proportion of the outcomes of MOGA are worse or equal to those of MOGA_MEAN is rejected at the 0.1 % level of confidence. Thus, we can safely conclude that the outcomes of MOGA are indeed better than those of MOGA_MEAN. For the test of the difference between two means, the *t*-score is 20.1982. Referring to the *t*-score probability table with 719 degree of freedom (N–1), the null hypothesis of the means of Sharpe ratio of the outcomes of MOGA is less than or equal to those of MOGA_MEAN is rejected at 0.1 % level of confidence. Hence, we are safe to say that, on average, the Sharpe ratios of the outcomes of MOGA are better than those of MOGA_MEAN.

9 Conclusions

In this chapter, a modified version of MOGA to handle model risk in the portfolio optimization problem with realistic constraints is proposed. Experiments and all relevant statistical tests have been conducted for the proposed algorithm with real data from the US stock market over 72 short-term investment periods. The results of the proposed algorithm (MOGA3O) are compared with those of original two-objective optimization (MOGA) and also with MOGA using stock mean returns and their standard deviations (MOGA_MEAN).

Experimental results indicated that the outcomes of MOGA3O are comparatively better for all three algorithms and statistical tests confirmed this conclusion. Incorporating a third objective, the approximation of portfolio Sharpe ratio error, helps to reduce the inherent model risk in the forecasting process. MOGA3O optimizes portfolio selections by inclining to choose stocks that are more accurately forecast given the same level of return and volatility. The third objective is added to capture the uncertain nature of forecasting models. It also has dynamic and adaptive elements when new information arrives. As a result, the model risk of forecasting models is mitigated. By altering the weights of pooled forecasting models according to prevailing economic, financial and market environment, we are able to forecast the future better. However, forecasting the future is extremely difficult as it is largely unforeseeable. There are many factors and variables involved in rendering a future event. Forecasting the future is disappointing at best. Also, our method for predicting the future is based on the forecasting of errors of particular forecasting models. It also has the same limitations as the forecasting models, that is, it needs to forecast future values based on past values, which are based on correlations not on causation. The forecasts are almost inaccurate. Outputs of any models that have inaccurate inputs are inaccurate. The third objective is only a measurement of errors of forecasting models. The errors could be measureable if and only if the error exhibits patterns that can be learned and thus predicted based on these patterns. If the errors exhibit no comprehensible patterns, they cannot be predicted and the third objective would not help to improve the portfolio selection.

Although, we believe that the proposed algorithm can be applied to any set of selectable stocks, our set was quite small (18 stocks) due to our experiment being limited to stocks with adequate historical data and the increased computational complexity resulting from the inclusion of more data. In future work, it may also be beneficial to use the MOGA platform to explore other optimization objectives such as the third moment of distribution (skewness), the fourth moments (kurtosis) or even to include other risk measures such as VaR (Value at Risk), CVaR (Conditional Value at Risk) and expected shortfall. An empirical comparison of the results here with results from an algorithm using these additional objectives would be interesting and would enable us to see whether the ex-post or out-of-sample performances could be improved. In order to see whether the algorithm also works for a larger number of selectable of assets, the number of stocks should be expanded in future research. The immediate objectives of future research are to include the third (skewness) and the fourth moments (kurtosis) as they may help to predict the future realization of stock returns. The multi-objective optimization platform can include, in theory, any number of objectives if they help to solve the Pareto optimization problems and here we tried to balance all objectives in non-discriminatory ways.

References

1. Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*(1), 77–91.
2. Kato, T., & Yoshiba, T. (2000). Model risk and its control. *Monetary and Economic Studies*, 18(2).
3. Derman, E. (1996). Model risk. *Goldman Sachs Quantitative Strategies Research Notes.*
4. Mitchell, M. (1998). *An introduction to genetic algorithms*, Paperback edition. Cambridge, MA: MIT Press, pp. 166–167.
5. Streichert, F., Ulmer, H., & Zell, A. (2005). Hybrid representation for compositional optimization and parallelizing MOEAs. Accessed July 12, 2008, from http://drops.dagstuhl.de/opus/volltexte/2005/251/pdf/04461
6. Streichert, F., Ulmer, H., & Zell, A. (2004). Evaluating a hybrid encoding and three crossover operators on the constrained portfolio selection problem. In *Proceedings of the 2004 congress on evolutionary computation,* IEEE Press (pp. 932–939).
7. Fonseca, C. M., & Flemming, P. J. (2001). Genetic algorithm for multiobjective optimization: Formulation, discussion and generalization. In S. Forrest (Ed.), *Proceeding of the fifth international conference on genetic algorithms* (pp. 416–423). San Mateo, CA: Morgan Kaufmann.
8. Stein, M., Branke, J., & Schmeck, H. (2004). Portfolio selection: How to integrate complex constraints. Accessed September 9, 2008, from http://symposium.fbv.unikarlsruhe.de/10th/papers/Stein_Branke_Schmeck%20%20Portfolio%20Selection%20How%20to%20Integrate%20Complex%20Constraints.pdf
9. Crama, Y., & Schyns, M. (2003). Simulated annealing for complex portfolio selection problems. *European Journal of Operational Research, 150*, 546–571.
10. Shape, W. E. (1964). Capital assets prices: A theory of market equilibrium and conditions of risk. *Journal of Finance, 19*(3), 425–442.
11. Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis*: *Forecasting and control*, revised ed. San Francisco: Holden-Day.
12. Asteriou, D., & Hall, S. G. (2007). *Applied econometrics*: *A modern approach*, revised ed. New York: Palgrave Macmillan.
13. Utts, J. M., & Heckard, R. F. (2006). *Statistical ideas and methods*. Belmont, CA: Thomson Higher Education.

11

Linear Regression Versus Fuzzy Linear Regression: Does it Make a Difference in the Evaluation of the Performance of Mutual Fund Managers?

Konstantina N. Pendaraki and Konstantinos P. Tsagarakis

1 Introduction

Policy makers and researchers are interested in whether superior returns can be generated by active managers who have the ability to collect and interpret information that helps forecast the returns on securities [1]. However, the increasing transaction growth generated worldwide by institutional portfolio managers highlights how important it is to know whether professional managers as a group add value to the portfolios they manage, or whether they create excessive transaction costs through their active management.

Researchers can use information on significant evidence of superior forecasting skills, if found, in order to earn superior returns that would violate the efficient market hypothesis. In the case where such violations are found, they would have far-reaching implications for the theory of finance with respect

K.N. Pendaraki (✉)
Department of Business Administration of Food and Agricultural Enterprises, University of Patras, Agrinio 30100, Greece
e-mail: dpendara@upatras.gr

K.P. Tsagarakis
Department of Environmental Engineering, Democritus University of Thrace, Xanthi 67100, Greece
e-mail: kandila@her.forthnet.gr

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets*,
New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0_11

to the optimal portfolio holdings of investors, the equilibrium valuation of securities and other decisions taken in corporate finance [2].

Therefore, the evaluation of the performance of investment managers is an area of considerable interest to both investors and researchers. According to Fama [3], the forecasting skills of a fund manager are distinguished through their micro-forecasting and macro-forecasting abilities. The key forecasting skills relate to the analysis of price movements of individual stocks relative to stocks generally, called security analysis (ability to select the right security in order to obtain high portfolio returns), and to price movements of the general stock market relative to fixed income securities, called market timing (ability to time the market). The first component refers to prediction of the non-systematic part of stock returns, while the second component refers to the systematic part versus the performance of the risk-free asset.

In the empirical literature on the evaluation of the performance of mutual fund managers, several models have been developed in order to investigate whether the performance of mutual funds is due to occasional events or due to superior investment management. Such models are the Capital Asset Pricing Model (CAPM; [4, 5]), the Treynor-Mazuy model [6], the Jensen model [7], the Henriksson-Merton model [2], the Grinblatt-Titman models [8, 9], the Fama-French model [10], the Carhart model [11] as well as several optimization models.

The literature on market-timing and the stock selection abilities of mutual fund managers is extensive and highly controversial, while the resulting evidence is mixed. Most empirical studies seem to suggest that significant timing ability is rare, and there is more evidence to suggest negative rather than positive market timing.

More precisely, Chen [12] and Coggin et al. [13] find that some fund managers demonstrate positive selection skills and negative timing abilities. Sehgal and Jhanwar [14] conclude that fund managers do not seem to possess significant market timing ability, while they exhibit significantly positive selectivity coefficient. Similarly, Oueslati et al. [15] find that the strong performance of fund managers comes from their selectivity abilities and not from their ability to demonstrate good market timing. Henriksson [16], Chang and Lewellen [17], Sinclair [18], Connor and Korajczyk [19], Lhabitant [20], Cesari and Panetta [21], Jiang [22] and Romacho and Cortez [23] conclude that there is no evidence of timing or selectivity in mutual funds returns. On the other hand, Ferson and Schadt [24] find evidence of the role played by the ability of fund managers. Lehmann and Modest [25], Cumby and Glen [26] and Lee and Rahman [27] find that there is some evidence of mutual fund manager's ability. It is clear that the results depend heavily on the methodology chosen, the different parameters used and the observation period under consideration.

Many studies on Greek mutual funds' performance evaluation based on traditional fund performance measures have been applied. See for example Philippas [28], Sorros [29] and Artikis [30]. Moreover, Pendaraki et al. [31, 32] and Babalos et al. [33, 34] evaluate Greek mutual funds' performance through multi-criteria analysis, Pendaraki and Spanoudakis [35] through argumentation-based decision-making theory, Alexakis and Tsolas [36], Babalos et al. [33, 34] and Pendaraki [37, 38] through data envelopment analysis. Literature on market timing ability of Greek fund managers is limited [39–44] and refers largely to the classical models. An overview of the literature in the area of mutual fund performance evaluation is presented in the early work of Pendaraki and Zopounidis [45].

The present study presents two different methodological approaches—ordinary least squares (OLS) and fuzzy linear regression (FLR)—that provide insights into the performance of mutual fund managers. We demonstrate how the performance results related to both selectivity and market-timing skills are modified when we use FLR instead of OLS. These two approaches are applied to both the Treynor-Mazuy and Henriksson-Merton models.

The application of FLR to the evaluation of the proposed mutual fund evaluation measures, along with the comparison of the obtained results using OLS is a research area, which to the best of our knowledge is missing from the literature, and it has never been undertaken in the mutual fund industry. Fuzzy logic architecture has been applied to other areas of mutual fund evaluation: mutual fund portfolio selection [46], the evaluation of mutual fund strategies [47] and the evaluation of mutual fund management companies' core competence [48].

The rest of the chapter is organized as follows. Section 2 outlines the main features of the proposed methodological framework used in this study. Section 3 gives a brief description of the data set used, while Sect. 4 describes and discusses the empirical results. Finally, Sect. 5 concludes by summarizing the main findings of the study and suggests avenues for future research.

2 Methodology

This study investigates the performance of mutual fund managers using the two most commonly used methods, the Treynor-Mazuy and the Henriksson-Merton models through OLS and FLR using two fitting measures; the root mean square error (RMSE) and the mean absolute error (MAE).

Treynor-Mazuy model

The majority of studies on the performance of mutual funds have employed the return-generating model developed by Jensen [7, 49]. Jensen [7] used the *alpha* coefficient in order to examine the ability of fund managers to achieve higher than the expected returns through the successful prediction of security prices, given the level of riskiness of the portfolio managed.

The first to examine the market timing and selectivity performance of mutual fund managers were Treynor and Mazuy in 1966. In order to test whether or not a mutual fund manager had actually outguessed the market, Treynor and Mazuy [6] in effect asked: 'Is there evidence that the volatility of the fund was higher in years when the market did well than in years when the market did badly?'

Treynor and Mazuy added a quadratic term to the Jensen equation to test for market-timing skills and suggested that portfolio return is a non-linear function of the market return:

$$\text{Treynor-Mazuy model:} \\ \left(R_{\mathrm{pt}} - R_{\mathrm{ft}}\right) = \alpha_{\mathrm{p}} + \beta_{\mathrm{p}}\left(R_{\mathrm{Mt}} - R_{\mathrm{ft}}\right) + \gamma_{\mathrm{p}}\left(R_{\mathrm{Mt}} - R_{\mathrm{ft}}\right)^2 + \varepsilon_{\mathrm{p}} \tag{11.1}$$

where, R_{pt} is the daily return on the fund's portfolio at time t, R_{ft} is the daily return of treasury bill (daily risk-free interest rate), R_{Mt} is the daily return of the portfolio market, α_{p} is the intercept term (estimated selectivity performance parameter), β_{p} is the estimated risk parameter, γ_{p} is the second slope coefficient (estimated market-timing performance parameter) and ε_{p} is the error term (independent normally distributed random variable with $E(\varepsilon_{\mathrm{p}}) = 0$).

The evaluation of the performance of a portfolio manager is shown using the two estimated parameters, α_{p} and γ_{p}. The parameter α_{p} shows the stock-selection ability of the portfolio manager, the parameter β_{p} shows the fund's systematic risk while the parameter γ_{p} shows the market-timing ability of the portfolio manager. Positive values of these parameters show the forecasting ability of the portfolio manager, while negative values show the forecasting inability of the portfolio manager. When the values of these parameters are near to zero or zero, this means that the portfolio manager has no forecasting ability at all.

Treynor and Mazuy [6] examined empirically the market-timing ability of 57 open-ended mutual funds for the period 1953–1962. The results obtained showed no statistical evidence that the managers of the examined funds had successfully outguessed the market.

Henriksson-Merton model

Henriksson [16] examined the forecasting ability of mutual fund managers using both the parametric and non-parametric techniques presented by Henriksson and Merton [2] and Merton [50]. The parametric tests require the assumption of either the capital asset pricing model or a multi-factor return structure. In this case and based strictly on observable returns, the tests permit the identification of the separate contributions of market-timing ability and micro-forecasting. On the other hand, the non-parametric tests require knowledge of the actual forecasts or a good proxy of them.

Under the conditions of the non-parametric approach, the independence is tested between the market timer's forecast and whether or not the return on the market portfolio is greater than the return of riskless securities. The non-parametric test performed on this null hypothesis takes advantage of the situation where those conditional probabilities of a correct forecast are sufficient to measure the forecasting ability. If the manager's forecasts are observable, then the non-parametric test can be used without further assumptions about the distribution of security returns on the market or on any particular model for security price valuation [2].

Under certain conditions, it is possible to figure out from the portfolio return series alone, what the manager's forecasts are; such conclusions generally give noisy estimates of the forecasts. In case the manager's portfolio positions are influenced by his/her forecasts for individual securities, the estimates will be extremely noisy [2].

The evaluation of the selectivity and timing abilities of portfolio managers using the Henriksson-Merton model is given by the following equation:

$$\left(R_{\mathrm{pt}} - R_{\mathrm{ft}}\right) = \alpha_{\mathrm{p}} + \beta_{\mathrm{p}}\left(R_{\mathrm{Mt}} - R_{\mathrm{ft}}\right) + \gamma_{\mathrm{p}} Z_{\mathrm{Mt}} + \varepsilon_{\mathrm{p}} \tag{11.2}$$

where, $Z_{\mathrm{Mt}} = \max\left[0, (R_{\mathrm{Mt}} - R_{\mathrm{ft}})\right]$.

Similar to the previous model, the parameter α_p shows the stock-selection ability while the parameter γ_p shows the market-timing ability of the portfolio manager. In the case where the parameter γ_p takes values greater than or equal to zero, this means that the fund manager changes the structure of his/her portfolio, resulting in a riskier portfolio in up markets and a less risky portfolio in a down market. On the other hand, when the parameter γ_p takes values of less than or equal to zero, this means that the fund manager does not take into account market movements and the structure of his/her portfolio does not change, or it is in the opposite direction of that of the markets.

According to this model an investment manager attempts to forecast when the market portfolio return exceeds the risk-free rate. In the case where the forecast is for an up market, the manager adjusts the portfolio to a higher target beta while in the case where the market forecast is pessimistic, a lower target beta is used.

Henriksson [16] examined the market-timing ability of 116 open-end mutual funds for the period 1968–1980. The results obtained did not support the hypothesis that mutual fund managers are able to follow an investment strategy that successfully times the return on the market portfolio. In addition, he found that only three funds had significantly positive estimates of market-timing ability. Henriksson found evidence of dynamic heteroscedasticity, but correcting it, did not alter his conclusion.

Fuzzy Linear Regression

In contrast to the OLS regression, a fuzzy regression has no disturbance term and thus the differences between the observed and the estimated value are reflected in the parameter fuzziness. The larger the width of a parameter, the less we know about the contribution of the variable to the model, but still we allow this incomplete knowledge to be included in the model. For this reason, a fuzzy regression provides a tool for analyzing large-scale and complex systems, such as economic and financial systems. Elements of the FLR theory are provided but for more information on the analysis below as well as calculations see Papadopoulos and Sirpi [51, 52] and Papadopoulos et al. [53]. An FLR model has the following form:

$$Y = A_0 + A_1 X_1 + A_2 X_2 + \cdots + A_n X_n \tag{11.3}$$

where A_i, $i = 0,1,\ldots,n$ are symmetrical triangular fuzzy numbers.

According to Tanaka et al. [54] and Tanaka [55], we assume that input data is a vector of non-fuzzy numbers and output data is a fuzzy number, and the deviations are caused by the 'indefiniteness' or ambiguity of the system structure. Then we have the following system:

$$Y = X\,A \tag{11.4}$$

where, A and Y are vectors of fuzzy numbers and X is the matrix of independent variables.

In particular, for the parameters' vector A, we assume that it consists of symmetrical fuzzy triangular numbers with membership functions:

$$\mu_{A_i}\left(a_i\right) = L_1\left(\frac{a_i - r_i}{c_i}\right), \quad c_i > 0 \tag{11.5}$$

where, r_i is the centre of the triangular number, and c_i is the width of the triangular number as shown in Fig. 11.1.

The reference function $L_1(x) = \max(0, \ 1 - |x|)$, satisfies the following conditions [54]:

$L_1(x) = L_1(-x)$, $L_1(0) = 1$ and $L_1(x)$ is strictly decreasing in $[0, \ 1)$.

When the parameters of the fuzzy linear regression model are symmetrical fuzzy triangular numbers, then by using the extension principle, it is proven that the outputs Yj of the system (11.4) are also symmetrical fuzzy triangular numbers [54] with membership functions:

$$\mu_{Y_j}\left(y_j\right) = L_1\left(\frac{y_j - \left(r_0 + \sum_{i=1}^{n} r_i x_{ij}\right)}{c_0 + \sum_{i=1}^{n} c_i \left|x_{ij}\right|}\right) \tag{11.6}$$

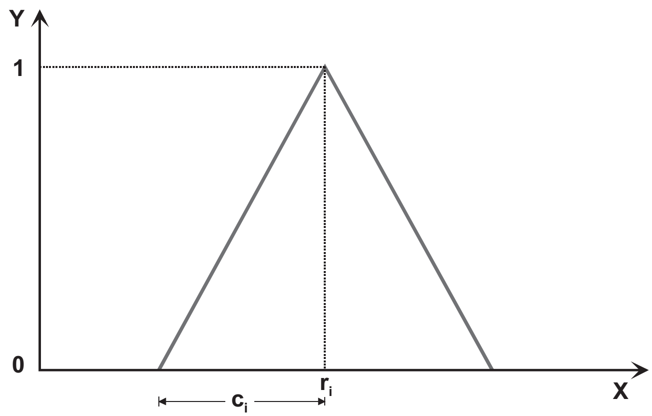


Fig. 11.1 Symmetrical triangular fuzzy number

The problem of finding the parameters of the linear possibility system is converted into a linear programming problem as described by Papadopoulos and Sirpi [51, 52].

Firstly, the above model is considered:

$$Y_j = A_0 + A_1 x_{1j} + A_2 x_{2j} + \cdots + A_n x_{nj}, \quad j = 1,2,\ldots,N$$

where $A_i = (r_i, c_i)_{L_1}$ are symmetrical triangular fuzzy numbers.

Then, we determine the degree h to which we wish the given data $((x_{1j}, x_{2j}, \ldots, x_{nj}), y_j)$ to be included in the inferred number Y_j, that is $\mu_{Y_j}(y_j) \geq h$ for $j = 1,2,\ldots,N$, which because of equation (11.6) takes the following form:

$$\mu_{Y_j}(y_j) = L_1\left(\frac{y_j - \left(r_0 \sum_{i=1}^{n} r_i x_{ij}\right)}{c_0 + \sum_{i=1}^{n} c_i \left|x_{ij}\right|}\right) \geq h \tag{11.7}$$

The fuzzy coefficients are found so that the total spread of the fuzzy output of all the data sets to be minimal, through the following objective function:

$$J = \min\left\{ N c_0 + \sum_{j=1}^{N}\sum_{i=1}^{n} c_i \left|x_{ij}\right| \right\} \tag{11.8}$$

Finally, by combining equations (11.7) and (11.8) we have the following linear programming problem:

$$J = \min\left\{ N c_0 + \sum_{j=1}^{N}\sum_{i=1}^{n} c_i \left|x_{ij}\right| \right\} \tag{11.9}$$

$$y_j \geq r_0 + \sum_{i=1}^{n} r_i x_{ij} - (1-h)\left(c_0 + \sum_{i=1}^{n} c_i \left|x_{ij}\right|\right) \tag{11.10}$$

$$y_j \leq r_0 + \sum_{i=1}^{n} r_i x_{ij} - (1-h)\left(c_0 + \sum_{i=1}^{n} c_i \left|x_{ij}\right|\right) \tag{11.11}$$

$$c_i \geq 0, \quad i = 1, 2, \ldots, n \tag{11.12}$$

If we have a fuzzy liner regression system $Y = A_0 + A_1X_1 + A_2X_2 + \cdots + A_nX_n$ where $A_i = (r_i, c_i)$ the following positive number is defined as the measure of fuzziness of Y:

$$m(Y) = \sqrt{\sum_{i=0}^{n} (c_i)^2} \tag{11.13}$$

From two given sets of data d and d', we extract the FLR models that have the same variables. The similarity ratio of these two models is defined as follows [51]:

$$\lambda(d, d') = \min\left\{ \left(\frac{\sum_{i=0}^{n} (c_i)_d^2}{\sum_{i=0}^{n} (c_i)_{d'}^2} \right)^{\frac{1}{2}}, \left(\frac{\sum_{i=0}^{n} (c_i)_{d'}^2}{\sum_{i=0}^{n} (c_i)_d^2} \right)^{\frac{1}{2}} \right\} \tag{11.14}$$

It can be seen that λ is related to the measure of fuzziness and obviously, $0 < \lambda \leq 1$.

For the calculations of the fuzzy numbers, fuzzy regression analysis software was used [56], while OLS regressions were estimated with EViews 7.

3 Data Set Description

The aim of this study is to analyze the estimates obtained from two different approaches: OLS and FLR. Discriminating among various risk levels that financial markets could be characterized as and capturing the uncertain variability of funds returns from the total number of the domestic equity mutual funds we select a small number of funds of different risk levels according to their systematic risk. We selected six funds from the high- and low-beta group classification. Specific thresholds that determine the MFs grouping are developed for each one of the examined sub-periods and beta coefficient. The purpose of this selection was to establish the differences in the MFs' behaviour (shifts in betas), and to investigate how these differences influence the results of our models in both approaches.

The empirical investigation of the domestic equity MFs was examined in three 4-month periods. Regarding the first sub-period (t1), 79 observations were included, while 85 observations were included in the second (t2) as well as in the third sub-period.

Daily records of domestic equity mutual funds were provided by the Association of Greek Institutional Investors. Further information was derived from the Athens Stock Exchange and the Bank of Greece, regarding the returns of the market portfolio (Athens Stock Exchange-General Index: ASE-GI) and the returns of the three-month Treasury bill (risk-free interest rate) respectively. The variations of the returns in the ASE-GI, were expressed through the variation of the prices of the stocks while also taking into account the fluctuations and the risk of the financial environment.

4 Empirical Application

Thirty-six equations were estimated both with OLS and FLR (a total of 72 equations): 18 equations using the Treynor-Mazuy model and 18 equations using the Henriksson-Merton model.

Results and Discussion

The Treynor-Mazuy and the Henriksson-Merton models were estimated by regressions presented in Tables 11.1 and 11.2 respectively. The estimates are given using the Newey-West method for heteroscedasticity correction [57], and the regression results appear not to have such a deficiency. In cases where necessary, we proceeded with the correction of autocorrelation. We apply Durbin-Watson test [58] statistic to test that the residuals from our ordinary least-squares regression are not autocorrelated against the alternative of positive first-order autocorrelation, since positive autocorrelation is seen much more frequently in practice than negative autocorrelation. The hypotheses we considered are

$$H_o : \rho = 0$$

$$H_1 : \rho > 0$$

The test statistic is

$$DW = \frac{\sum_{i=2}^{n}\left(e_i - e_{i-1}\right)^2}{\sum_{i=1}^{n} e_i^{\,2}}$$

where $e_i = y_i - \hat{y}_i$ and y_i and $\hat{y}_i$ are, respectively, the observed and predicted values of the response variable for individual *i*. The Durbin-Watson statistic ranges in value from 0 to 4. A value near 2 indicates non-autocorrelation; a value toward 0 indicates positive autocorrelation; a value toward 4 indicates negative autocorrelation. Because of the dependence of any computed Durbin-Watson value on the associated data matrix, exact critical values of the Durbin-Watson statistic are not tabulated for all possible cases. Instead, Durbin and Watson established upper and lower bounds for the critical values. Formally, decision rules are as follows:

If DW < dL Reject Ho
If DW > dU Fail to Reject Ho
If dL < DW < dU Test is inconclusive

The Durbin-Watson test is conditioned by the following assumptions: explanatory variables are non-stochastic, the error terms are assumed to be normally distributed and the regression models do not include the lagged values of the regression. The test measures first order serial correlation only.

Autocorrelation was found for both models in the third sub-period for MF2, in both the second and the third sub-periods for MF5, and in the first sub-period for MF6, for 5 % significance points of dL and dU.

The coefficient estimates are followed by *p*-value and R^2 for OLS, and *r* (fuzzy numbers-parameter estimates), *c* (width of the parameter estimates) and *m*(*Y*) (measure of fuzziness) for FLR. In the Treynor-Mazuy and Henriksson-Merton models the constant term is α (estimated selectivity performance parameter), the first slope is β (estimated risk parameter) and γ is the second slope coefficient (estimated market-timing performance parameter) in OLS estimates, while a_i is the fuzzy triangular number with α = (r_{ao}, c_{ao}), β = (r_{a1}, c_{a1}) and γ = (r_{a2}, c_{a2}), in FLR estimates.

To make the comparison of OLS and FLR, two fitting measures were used. These were RMSE and the MAE as presented in Tables 11.1 and 11.2, for both OLS and FLR. Although these fitting measures were applied to the

Table 11.1 Results of the Treynor-Mazuy model

		Ordinary least squares					Fuzzy linear regression				
Variable		Coef.	*p*	R^2	RMSE	MAE	*r*	*c*	*m*(*Y*)	RMSE	MAE
MF1t1	α	−0.0160	0.6000	0.9601	0.2118	0.1675	0.0408	0.3876	0.4170	0.2845	0.2389
	β	1.0559	0.0000				0.8669	0.1539			
	γ	0.0206	0.2583				0.0066	0.0000			
MF1t2	α	0.0001	0.9957	0.9640	0.1567	0.1191	−0.0140	0.3591	0.3664	0.1923	0.1518
	β	1.0279	0.0000				1.1412	0.0726			
	γ	0.0175	0.3757				0.0648	0.0000			
MF1t3	α	−0.0092	0.6126	0.9616	0.1345	0.1092	−0.0056	0.2889	0.2895	0.1506	0.1237
	β	0.9935	0.0000				1.0899	0.0195			
	γ	−0.0180	0.3900				−0.0463	0.0000			
MF2t1	α	0.0132	0.2895	0.9918	0.0867	0.0639	0.0759	0.1816	0.1909	0.1117	0.0906
	β	0.9552	0.0000				0.9240	0.0589			
	γ	−0.0080	0.2791				−0.0081	0.0000			
MF2t2	α	0.0187	0.0897	0.9890	0.0810	0.0483	0.1981	0.2561	0.2561	0.1700	0.1587
	β	0.9809	0.0000				0.9814	0.0000			
	γ	0.0027	0.7856				−0.0676	0.0000			
MF2t3	α	−0.0024	0.5359	0.9975	0.0340	0.0205	0.0381	0.0627	0.0967	0.0709	0.0604
	β	0.9938	0.0000				1.0611	0.0736			
	γ	0.0006	0.9127				−0.0184	0.0009			
MF3t1	α	0.0613	0.0788	0.8597	0.2403	0.1937	0.1809	0.5941	0.5941	0.2866	0.2275
	β	0.5607	0.0000				0.6676	0.0000			
	γ	−0.0693	0.0011				−0.1044	0.0000			
MF3t2	α	0.0396	0.2041	0.8262	0.2309	0.1824	0.0861	0.6187	0.6187	0.3093	0.2470
	β	0.6506	0.0000				0.8949	0.0000			
	γ	−0.0236	0.4076				−0.0263	0.0000			
MF3t3	α	0.0954	0.0011	0.8256	0.2101	0.1663	0.1114	0.5969	0.5969	0.2705	0.2171
	β	0.6772	0.0000				0.4295	0.0000			
	γ	−0.0434	0.1864				−0.0067	0.0000			
MF4t1	α	0.0438	0.3041	0.7664	0.2959	0.2405	0.0604	0.6372	0.6372	0.3078	0.2495

	β	0.5162	0.0000				0.5571	0.0000			
	γ	−0.0473	0.0645				−0.0876	0.0000			
MF4t2	α	0.0076	0.8657	0.6773	0.3328	0.2672	−0.0229	0.7596	0.7596	0.3427	0.2723
	β	0.6065	0.0000				0.5968	0.0000			
	γ	0.0359	0.3823				−0.0156	0.0000			
MF4t3	α	0.0139	0.6835	0.7497	0.2544	0.2052	0.0204	0.5107	0.5155	0.2575	0.2073
	β	0.6527	0.0000				0.6719	0.0698			
	γ	−0.0179	0.6514				−0.0676	0.0000			
MF5t1	α	0.1266	0.0372	0.7405	0.4169	0.3332	0.1785	0.8674	0.8683	0.4268	0.3440
	β	0.6310	0.0000				0.5425	0.0404			
	γ	−0.1292	0.0005				−0.1747	0.0000			
MF5t2	α	0.0607	0.3161	0.6036	0.3889	0.3150	0.0471	0.7790	0.8030	0.3961	0.3187
	β	0.6280	0.0000				0.5420	0.1950			
	γ	−0.0477	0.3370				−0.0610	0.0000			
MF5t3	α	0.0319	0.6098	0.5117	0.3992	0.2900	−0.2950	1.3607	1.3607	0.6069	0.5050
	β	0.6084	0.0000				0.1155	0.0000			
	γ	−0.0083	0.8922				0.1799	0.0000			
MF6t1	α	0.1065	0.0198	0.7118	0.3123	0.2626	0.0630	0.6127	0.6127	0.3342	0.2877
	β	0.4370	0.0000				0.5447	0.0000			
	γ	−0.0939	0.0007				−0.0961	0.0000			
MF6t2	α	0.0563	0.2860	0.5781	0.3912	0.3095	0.0095	0.8164	0.8167	0.4103	0.3236
	β	0.5830	0.0000				0.7114	0.0215			
	γ	0.0197	0.6828				−0.0381	0.0000			
MF6t3	α	0.0545	0.1513	0.7463	0.2807	0.2176	0.2055	0.6507	0.6507	0.3348	0.2733
	β	0.7117	0.0000				0.8259	0.0000			
	γ	0.0095	0.8274				0.0114	0.0000			

Table 11.2 Results of the Henriksson-Merton model

Variable		Ordinary least squares					Fuzzy linear regression				
		Coef.	p	R^2	RMSE	MAE	r	c	$m(Y)$	RMSE	MAE
MF1t1	α	−0.0421	0.3121	0.9604	0.2109	0.1663	−0.0207	0.4487	0.4654	0.2555	0.2120
	β	1.1122	0.0000				1.0291	0.0000			
	γ	0.1160	0.1672				0.1983	0.1240			
MF1t2	α	−0.0013	0.9673	0.9637	0.1571	0.1203	−0.0156	0.3562	0.3666	0.1619	0.1247
	β	1.0502	0.0000				1.0424	0.0000			
	γ	0.0371	0.6308				−0.0375	0.0866			
MF1t3	α	−0.0027	0.9109	0.9615	0.1345	0.1087	−0.0106	0.1985	0.2597	0.1466	0.1160
	β	0.9664	0.0000				1.0518	0.1099			
	γ	−0.0560	0.4227				−0.0551	0.1262			
MF2t1	α	0.0274	0.1076	0.9919	0.0858	0.0636	0.0793	0.1825	0.1920	0.1130	0.0904
	β	0.9276	0.0000				0.8457	0.0000			
	γ	−0.0552	0.1071				−0.0910	0.0598			
MF2t2	α	0.0227	0.1466	0.9890	0.0810	0.0487	0.2073	0.2584	0.2584	0.1756	0.1665
	β	0.9784	0.0000				0.9219	0.0000			
	γ	−0.0074	0.8503				−0.0979	0.0000			
MF2t3	α	−0.0024	0.6638	0.9975	0.0340	0.0205	0.0409	0.0620	0.1014	0.0730	0.0618
	β	0.9943	0.0000				1.0398	0.0803			
	γ	0.0010	0.9552				−0.0634	0.0000			
MF3t1	α	0.1121	0.0212	0.8568	0.2427	0.1978	0.2314	0.6031	0.6031	0.2916	0.2332
	β	0.4235	0.0000				0.4745	0.0000			
	γ	−0.2985	0.0026				−0.2003	0.0000			
MF3t2	α	0.0560	0.2077	0.8263	0.2308	0.1812	0.0884	0.6186	0.6186	0.2786	0.2157
	β	0.6002	0.0000				0.3817	0.0000			
	γ	−0.0956	0.3939				−0.5522	0.0000			
MF3t3	α	0.1138	0.0032	0.8256	0.2101	0.1656	0.1125	0.5971	0.5971	0.2691	0.2159
	β	0.6072	0.0000				0.4232	0.0000			
	γ	−0.1453	0.1842				−0.0198	0.0000			

MF4t1	α	0.0866	0.1396	0.7668	0.2956	0.2404	0.1171	0.6377	0.6377	0.3112	0.2512
	β	0.4111	0.0000				0.4142	0.0000			
	γ	−0.2238	0.0587				−0.1198	0.0000			
MF4t2	α	0.0108	0.8658	0.6748	0.3341	0.2668	−0.0117	0.7502	0.7502	0.3431	0.2727
	β	0.6435	0.0000				0.5556	0.0015			
	γ	0.0577	0.7214				−0.0757	0.0000			
MF4t3	α	0.0052	0.9091	0.7491	0.2547	0.2047	−0.0452	0.4235	0.4664	0.2610	0.2094
	β	0.6518	0.0000				0.6300	0.1403			
	γ	0.0009	0.9945				0.0853	0.1360			
MF5t1	α	0.1895	0.0276	0.7242	0.4298	0.3417	0.2721	0.8831	0.8844	0.4495	0.3648
	β	0.4200	0.0002				0.2347	0.0000			
	γ	−0.4775	0.0062				−0.5912	0.0469			
MF5t2	α	0.0778	0.3410	0.6016	0.3894	0.3138	0.0671	0.7620	0.7958	0.3970	0.3235
	β	0.5488	0.0000				0.6859	0.0000			
	γ	−0.1426	0.4679				0.0548	0.2293			
MF5t3	α	0.0299	0.6989	0.5116	0.3990	0.2903	−0.5986	1.1256	1.2247	0.6824	0.5603
	β	0.6053	0.0000				0.7628	0.0000			
	γ	−0.0066	0.9741				1.5172	0.4826			
MF6t1	α	0.1498	0.0264	0.7046	0.3212	0.2709	0.0935	0.6323	0.6323	0.3309	0.2781
	β	0.2930	0.0003				0.3791	0.0000			
	γ	−0.3358	0.0070				−0.3070	0.0000			
MF6t2	α	0.0711	0.3446	0.5773	0.3916	0.3111	0.0223	0.8111	0.8116	0.4127	0.3246
	β	0.5850	0.0000				0.6666	0.0295			
	γ	−0.0088	0.9631				−0.1234	0.0000			
MF6t3	α	0.0246	0.6227	0.7486	0.2794	0.2144	0.1945	0.6487	0.6488	0.3334	0.2721
	β	0.7713	0.0000				0.8527	0.0000			
	γ	0.1282	0.3770				0.0560	0.0083			

Table 11.3 Summary results of the Treynor-Mazuy model using OLS

Parameter	Positive	Negative	Statistically Significant*		Statistically insignificant	
			Positive	Negative	Positive	Negative
Selectivity (α)	15	3	3	0	12	3
Market timing (γ)	7	11	0	3	7	8

**at 5% level*

Table 11.4 Summary results of the Henriksson-Merton model using OLS

Parameter	Positive	Negative	Statistically Significant*		Statistically insignificant	
			Positive	Negative	Positive	Negative
Selectivity (α)	14	4	4	0	10	4
Market timing (γ)	6	12	0	3	6	9

**at 5% level*

overall model, explicit interpretations can be made for the estimated coefficients as compounds of this model. Furthermore, Tables 11.3 and 11.4 present the significance summary results of Treynor-Mazuy and Henriksson-Merton models for the OLS estimates respectively.

The Performance of Mutual Funds Managers

According to models in Tables 11.1 and 11.2, the OLS estimates of the beta coefficient were statistically significant in all cases, namely the *p*-value was almost zero. This fact is in accordance with the FLR estimates, which resulted in estimates with no or very low width for the estimated fuzzy numbers. However, this does not mean that mutual funds managers reveal any significant forecasting ability.

The empirical results of OLS of Treynor-Mazuy (TM) model presented in Tables 11.1 and 11.3 and the Henriksson-Merton (HM) model presented in Tables 11.2 and 11.4 showed that 15 TM and 14 HM fund managers had positive selectivity coefficients (α), whereas in three TM and four HM cases the coefficients were negative. Out of the 15 TM and 14 HM positive coefficients 12 TM and four HM were not significant at the 5 % level, whereas none of the negative coefficients were significant. As far as the market timing coefficient (γ) of the TM model and the HM model is concerned, it has a positive sign in seven TM and six HM cases of which, none were statistically significant whereas it was negative for the rest of which three were statistically significant. Thus, for both models, in only a few cases did fund managers show the ability to select the right securities in order to obtain high portfolio returns. This means that the examined fund managers were on average better stock pickers than market timers. More precisely, the average selectivity measure was positive and the average timing measure was negative, and only in a few cases statistically significant. This means that the market timing and selectivity performance for the examined mutual funds were in general poor.

Regarding the Henriksson-Merton model, this fact is in accordance with the FLR estimates (Table 11.2) for the selectivity timing coefficients, where c took values greater than zero, showing that there was uncertainty in our estimates, maybe due to the market conditions that are characterized by major fluctuations and considerable instability in all stock prices.

Figures 11.2 and 11.3 present the selectivity and market timing coefficients respectively for all models calculated by OLS and FLR. Coefficients calculated with FLR are followed by their width, as expressed by the error bars. In Figs. 11.2 and 11.3 fuzzy calculated coefficients were confirmatory over the OLS calculated coefficients as in all cases the fuzzy width included the OLS coefficient.

Regarding the Treynor-Mazuy model for the FLR estimates (Table 11.1) in almost all timing coefficients (γ) there was no spread in our estimates ($c = 0$). However, regarding the Henriksson-Merton model, for the market timing coefficient in half of the cases, c equalled zero, which leads to different conclusions (Fig. 11.3). According to FLR estimates there was no uncertainty in these cases for the market timing ability of specific fund managers.

Additional information can be drawn when the fuzzy number has a non-zero width, in the case where the OLS estimate is not statistically significant.

Furthermore, based on the results of the OLS estimates, in three mutual funds (MF1, MF2 and MF4), the selectivity and timing coefficients changed signs between the examined sub-periods and were not statistically significant in all the three sub-periods. The selectivity and timing coefficients for MF3 and MF5 did not change signs between the examined sub-periods however, they did change their significance. Finally, the selectivity and timing coefficients for MF6 changed both signs and significance between the examined sub-periods. The results given by FRL estimates showed that the selectivity and timing coefficients changed signs between the examined sub-periods for MF1, MF4, MF5 and MF6, while only for MF2 and MF3 the signs did not change. These changes in both OLS and FRL estimates are due to the fact that during the examined sub-periods, mutual funds characteristics on performance and risk were also changed. These results are the same for both Treynor-Mazuy and Henriksson-Merton models.

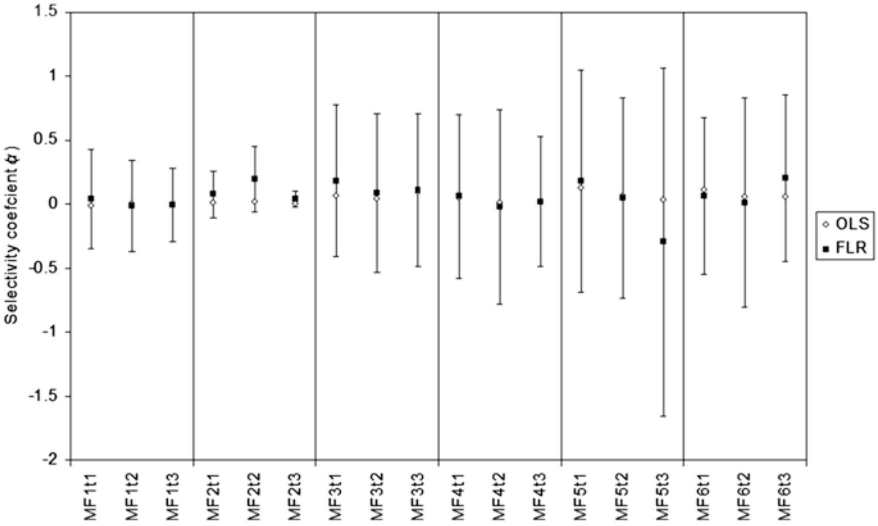


Fig. 11.2 Selectivity coefficient according to the Treynor-Mazuy model

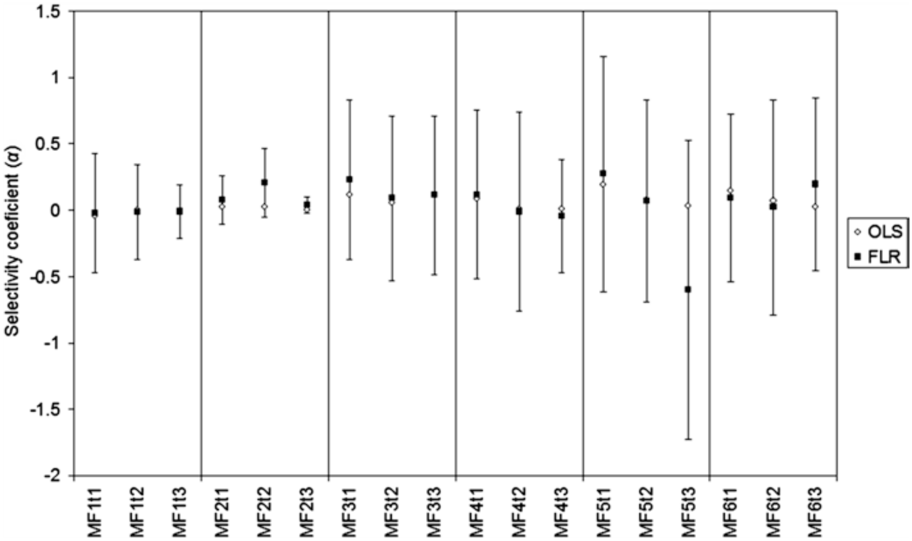


Fig. 11.3 Selectivity coefficient according to the Henriksson-Merton model

For both the Treynor-Mazuy and Henriksson-Merton models, the selectivity coefficient (α) did not change signs between the OLS and the FLR estimates in almost all of the cases. For the Treynor-Mazuy model, in five cases, a changes sign and are the following: MF1 in the first and second sub-period, MF2 and MF5 in the third sub-period, and MF4 in the second sub-period. For the Henriksson-Merton model, in four cases, a changes sign: MF2 in the first sub-period, MF4 in the second and third sub-period, and MF5 in the second sub-period.

For the Treynor-Mazuy model, the five cases where the market coefficient (γ) changed sign are as follows: MF2 in the second and third sub-period, MF4 and MF5 in the third sub-period, and MF6 in the second sub-period. For the Henriksson-Merton model, the five cases where the market coefficient (γ) changed sign are as follows: MF2 in the third sub-period, MF4 in the second sub-period, and MF5 in the second and third sub-period. For both models, in the cases where the selectivity and the market coefficients did not change signs between the OLS and the FLR estimates, their size was almost the same.

For both Treynor-Mazuy and Henriksson-Merton models, as regards OLS estimates, the R^2 has high values in all cases (good fit of our model) except for three (MF5 in the second and third sub-period and MF6 in the second sub-period). These three exceptions are in accordance with the FLR estimates when the values for the measure of fuzziness, $m(Y)$, were the highest ones.

Furthermore, for both Treynor-Mazuy and Henriksson-Merton models the highest values of R^2 in OLS estimates kept up with the lowest values of the measure of fuzziness in the case of MF2 in all the three sub-periods ($m(Y)$ had the lowest values). A Pearson correlation coefficient for R^2 and $m(Y)$ has been proven to be statistically significant with $r = 0.911$ ($p < 0.001$).

The similarity of RMSE and MAE measures for all cases is obvious. It can be observed that in all cases, the OLS had lower RMSE and MAE than FLR. Thus, the prediction value of our models is higher in OLS estimates than in FLR estimates.

Fuzzy Similarity Ratios

The similarity ratio (SR) of the FLR models for the Treynor-Mazuy and Henriksson-Merton models is presented in Tables 11.5 and 11.6 respectively. Here the similarity ratios show how similar the estimations were in terms of their fuzziness. Thus both models with high similarity ratios will both have either high or low fuzziness, that is, high or low uncertainty expressed with the width of the fuzzy numbers. On the other hand models with low similarity ratios will have one high and one low fuzziness.

In the Treynor-Mazuy model, MF5 and MF2 in the third sub-period were the least similar with SR = 0.07, followed by MF6 and MF2 with SR = 0.15, and MF3 and MF2 with SR = 0.16, in the same sub-period. These funds belonged to different risk groups (high and low systematic risk). The most similar funds, as it was expected, were MF6 and MF5 in the second sub-period, with SR = 0.98, followed by MF6 and MF4 with SR = 0.96, and MF6

»	MF1t1	MF1t2	MF1t3	MF2t1	MF2t2	MF2t3	MF3t1	MF3t2	MF3t3	MF4t1	MF4t2	MF4t3	MF5t1	MF5t2	MF5t3	MF6t1	MF6t2	MF6t3
MF1t1	1																	
MF1t2	0.88	1																
MF1t3	0.69	0.79	1															
MF2t1	0.46			1														
MF2t2		0.70		0.75	1													
MF2t3			0.33	0.51	0.38	1												
MF3t1	0.70			0.32			1											
MF3t2		0.59			0.41		0.96	1										
MF3t3			0.49			0.16	1.00	0.96	1									
MF4t1	0.65			0.30			0.93			1								
MF4t2		0.48			0.34			0.81		0.84	1							
MF4t3			0.56			0.19			0.86	0.81	0.68	1						
MF5t1	0.48			0.22			0.68			0.73			1					
MF5t2		0.46			0.32			0.77			0.95		0.92	1				
MF5t3			0.21			0.07			0.44			0.38	0.64	0.59	1			
MF6t1	0.68			0.31			0.97			0.96			0.71			1		
MF6t2		0.45			0.31			0.76			0.93			0.98		0.75	1	
MF6t3			0.44			0.15			0.92			0.79			0.48	0.94	0.80	1

Table 11.5 Similarity ratio for the Treynor-Mazuy model

»	MF1t1	MF1t2	MF1t3	MF2t1	MF2t2	MF2t3	MF3t1	MF3t2	MF3t3	MF4t1	MF4t2	MF4t3	MF5t1	MF5t2	MF5t3	MF6t1	MF6t2	MF6t3
MF1t1	1																	
MF1t2	0.79	1																
MF1t3	0.56	0.71	1															
MF2t1	0.41			1														
MF2t2		0.70		0.74	1													
MF2t3			0.39	0.53	0.39	1												
MF3t1	0.77			0.32			1											
MF3t2		0.59			0.42		0.97	1										
MF3t3			0.43			0.17	0.99	0.97	1									
MF4t1	0.73			0.30			0.95			1								
MF4t2		0.49			0.34			0.82		0.85	1							
MF4t3			0.56			0.22			0.78	0.73	0.62	1						
MF5t1	0.53			0.22			0.68			0.72			1					
MF5t2		0.46			0.32			0.78			0.94		0.90	1				
MF5t3			0.21			0.08			0.49			0.38	0.72	0.65	1			
MF6t1	0.74			0.30			0.95			0.99			0.71			1		
MF6t2		0.45			0.32			0.76			0.92			0.98		0.78	1	
MF6t3			0.40			0.16			0.92			0.72			0.53	0.97	0.80	1

Table 11.6 Similarity ratio for the Henriksson-Merton model

and MF3 with SR = 0.95, in the first sub-period. These funds belonged to the same risk groups (high systematic risk).

Finally, in Henriksson-Merton model, MF5 and MF2 in the third sub-period were the least similar with SR=0.08, followed by MF6 and MF2 with SR = 0.16, and MF3 and MF2 with SR = 0.17, in the same sub-period. The most similar funds in Henriksson-Merton models were MF6 and MF4 in the first sub-period with SR= 0.99, followed by MF6 and MF5 in the second sub-period with SR = 0.98, and MF4 and MF3 in the first sub-period, with SR = 0.95.

5 Conclusions and Future Perspectives

Using equity mutual fund data, we have used the conventional analysis from the OLS approach and have introduced the FLR approach in order to see similarities and novel approaches of mutual fund performance measures.

Furthermore, discriminating between various risk levels by which financial markets could be characterized, the combination of the two approaches will give a sharper picture of the actual behaviour of fund managers. In this contribution we view our results as a step in this direction. The main conclusions of our study are summarized as follows:

1. OLS and FLR models show great similarities, as observed by the estimated models of the data analyzed. Proof of this is provided by the two fitting measures used: the RMSE and the MAE. These measures had very similar values in most cases in the examined models. However, better (lower) values were observed for all OLS models.
2. Our overall result indicates that the Greek fund managers are characterized by poor market timing and selectivity ability in the examined period.
3. For FLR estimates there was no width in market timing ability of the examined fund managers. This result suggests that this approach has an advantage over OLS, where non-statistically significant coefficients are calculated.
4. It was noticed that there are cases where FLR estimates can be used: (a) confirmatory, when the fuzzy numbers range include OLS statistically calculated coefficients, (b) alternatively, when there are no OLS statistically calculated coefficients and fuzzy width is clearly placed, and (c) additional investigation needed when FLR and OLS numbers lead to different conclusions. However, further empirical analysis should be done to support the validity of this kind of information.
5. We have introduced $m(Y)$ as a fuzzy related fit towards R^2. It was found that R^2 and the measure of fuzziness behave in the same way. When R^2 is high in OLS models, the fuzziness is low and vice versa.
6. The similarity ratio of fuzziness provides a similarity measure for uncertainty based on the predicted values. This could become a new measure used in financial analysis of mutual funds for evaluating one mutual fund against another and one OF against itself in another period.

Further research should be undertaken on this fuzzy application to do with the optimization of this approach and the new measures in order to provide alternative evaluation scenarios. It would be of interest to investigate the performance of mutual fund managers and the comparison of OLS and FLR upon different investment choices, such as stocks, bonds and derivatives, or by incorporating further information such as fund age, fund size.

References

1. Grinblatt, M., & Titman, S. (1995). Performance evaluation. In R. Jarrow et al. (Eds.), *Handbooks in Operation Research and Management* (vol. 9, pp. 581–609). Elsevier Science.
2. Henriksson, R., & Merton, R. (1981). On market timing and investment performance. *Journal of Business, 54*, 513–534.
3. Fama, E. (1972). Components of investment performance. *Journal of Finance, 27*, 551–567.
4. Lintner, J. (1965). The valuation of risk assets and selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*, 13–37.
5. Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.
6. Treynor, J. L., & Mazuy, K. K. (1966). Can mutual funds outguess the market. *Harvard Business Review* (July/August), 131–136.
7. Jensen, C. M. (1968). The performance of mutual funds in the period 1945–1964. *Journal of Finance, 23*, 389–416.
8. Grinblatt, M., & Titman, S. (1989). Portfolio performance evaluation: Old issues and new insights. *The Review of Financial Studies, 2*, 393–416.
9. Grinblatt, M., & Titman, S. (1994). A study of monthly mutual fund returns and performance evaluation techniques. *Journal of Financial and Quantitative Analysis, 29*, 419–444.
10. Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3–56.
11. Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*(1), 57–82.
12. Chen, C. (1992). A cross-sectional analysis of mutual funds' market timing and mutual fund investment performance. *Journal of Business, 57*, 57–72.
13. Coggin, T. D., Fabozzi, F. J., & Rahman, S. (1993). The investment performance of US equity pension fund managers: An empirical investigation. *Journal of Finance, 48*, 1039–1055.
14. Sehgal, S., & Jhanwar, M. (2008). On stock selection skills and market timing abilities mutual fund managers in India. *International Research Journal of Finance and Economics, 15*, 307–317.
15. Oueslati, A., Hammami, Y., & Jilani, F. (2014). The timing ability and global performance of Tunisian mutual fund managers: A multivariate GARCH approach. *Research in International Business and Finance, 31*, 57–73.
16. Henriksson, R. (1984). Market timing and mutual fund performance: An empirical investigation. *Journal of Business, 57*, 73–96.
17. Chang, E. C., & Lewellen, W. G. (1984). Market timing and mutual fund investment performance. *Journal of Business, 59*, 585–598.

18. Sinclair, N. A. (1990). Market timing ability of pooled super-annuation funds January 1981 to December 1987. *Accounting and Finance, 30*, 51–65.
19. Connor, G., & Korajzcyk, R. (1991). The attribute, behaviour and performance of US equity pension fund managers: An empirical investigation. *Journal of Finance, 48*, 1039–1055.
20. Lhabitant, F. (2001). On Swiss timing and selectivity: In the quest of alpha. *Financial Markets and Portfolio Management, 15*, 154–172.
21. Cesari, R., & Panetta, F. (2002). The performance of Italian equity funds. *Journal of Banking and Finance, 26*, 99–126.
22. Jiang, W. (2003). A nonparametric test of market timing. *Journal of Empirical Finance, 10*, 399–425.
23. Romacho, J. C., & Cortez, M. C. (2006). Timing and selectivity in Portuguese mutual fund performance. *Research in International Business and Finance, 2*, 348–368.
24. Ferson, W. E., & Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. *Journal of Finance, 51*, 425–461.
25. Lehmann, B. N., & Modest, D. M. (1987). Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. *The Journal of Finance, XLII*(2), 233–265.
26. Cumby, R. E., & Glen, J. D. (1990, June). Evaluating the performance of international mutual funds. *Journal of Finance, XLV*(2), 497–521.
27. Lee, Ch. F., & Rahman, S. (1999). New evidence on timing and security skill of mutual fund managers. *The Journal of Portfolio Management*, (Winter), 80–83.
28. Philippas, N. D. (1999). *Mutual funds and stock exchange environment*. Athens: Globus Invest (In Greek).
29. Sorros, J. N. (2003). Return and risk analysis: A case study in equity mutual funds operating in the Greek financial market. *Managerial Finance, 29*(9), 21–28.
30. Artikis, G. P. (2004). Performance evaluation of the bond mutual funds operating in Greece. *Managerial Finance, 30*(10), 1–13.
31. Pendaraki, K., Doumpos, M., & Zopounidis, C. (2003). Assessing equity funds' performance using a multicriteria methodology: A comparative analysis. *South Eastern Europe Journal of Economics, 1*, 85–104.
32. Pendaraki, K., Zopounidis, C., & Doumpos, M. (2005). A multicriteria methodology and an application to the Greek market of equity mutual funds. *European Journal of Operational Research, 163*, 462–481.
33. Babalos, V., Caporale, G. M., & Philippas, N. (2012b). Efficiency evaluation of Greek equity funds. *Research in International Business and Finance, 26*, 317–333.
34. Babalos, V., Philippas, N., Doumpos, M., & Zopounidis, C. (2012a). Mutual fund performance appraisal using stochastic multicriteria acceptability analysis. *Applied Mathematics and Computation, 218*, 5693–5703.

35. Pendaraki, K., & Spanoudakis, N. (2012). An interactive tool for mutual funds portfolio composition using argumentation. *Journal of Business, Economics and Finance, 1*(3), 33–51.
36. Alexakis, P., & Tsolas, I. (2011). Appraisal of mutual equity fund performance using data envelopment analysis. *Multinational Finance Journal, 15*(3/4), 273–296.
37. Pendaraki, K. (2012). Mutual fund performance evaluation using data envelopment analysis with higher moments. *Journal of Applied Finance and Banking, 2*, 97–112.
38. Pendaraki, K. (2015). Mutual fund performance benchmarking using a quadratic directional distance function approach. *International Journal of Financial Engineering and Risk Management, 2*(1), 30–47.
39. Handjinicolaou, G. (1980). The performance of Greek mutual funds in the period 1973–76: A case of internationally diversified portfolios. *Spoudai, 11*(3/4), 381–391.
40. Milonas, N. (1999). Greek mutual funds-Theory and practice. *Greek Bank Association*, *Sakkoulas Publishing*, Athens.
41. Koulis, A., Beneki, C., Adam, M., & Botsaris, C. (2011). An assessment of the performance of Greek mutual equity funds selectivity and market timing. *Applied Mathematical Sciences, 5*(4), 159–171.
42. Philippas, N. (2001). An empirical evaluation of the performance of the Greek fund managers, *Spoudai*. *University of Piraeous, 51*(1/2), 142–161.
43. Sorros, J. (2001). Equity mutual fund managers performance in Greece. *Managerial Finance, 27*(6), 68–74.
44. Thanou, E. (2008). Mutual fund evaluation during up and down market conditions: The case of Greek equity mutual funds. *International Research Journal of Finance and Economics, 13*, 84–93.
45. Pendaraki, K., & Zopounidis, C. (2001). Evaluating mutual fund performance: An overview. In C. Zopounidis (Ed.), *New Trends in Banking Management* (pp. 251–275). Berlin: Physica Verlag.
46. Gladish, B. P., Jonesb, D. F., Tamizb, M., & Bilbao Terol, A. (2007). An interactive three-stage model for mutual funds portfolio selection. *Omega, 35*, 75–88.
47. Wang, S. Y., & Lee, C.-F. (2010). Fuzzy multi-criteria decision-making for evaluating mutual fund strategies. *Applied Economics, 1-10*.
48. Zhao, X., & Yue, W. (2012). A multi-subsystem fuzzy DEA model with its application in mutual funds management companies' competence evaluation. *Procedia Computer Science, 1*, 2469–2478.
49. Jensen, C. M. (1969). Risk, the pricing of capital assets, and evaluations of investment portfolios. *Journal of Business, 42*, 167–247.
50. Merton, R. C. (1981). On market timing and investment performance: An equilibrium theory of value for market forecasts. *Journal of Business, 54*, 363–406.
51. Papadopoulos, B. K., & Sirpi, M. (1999). Similarities in fuzzy regression models. *Journal of Optimization Theory and Applications, 102*, 373–383.

52. Papadopoulos, B. K., & Sirpi, M. (2004). Similarities and distances in fuzzy regression models. *Soft Computing, 8*, 556–561.
53. Papadopoulos, B., Tsagarakis, K. P., & Yannopoulos, A. (2007). Cost and land functions for wastewater treatment projects: Typical simple linear regression versus fuzzy linear regression. *Journal of Environmental Engineering, 133*, 581–586.
54. Tanaka, H., Uejima, S., & Asai, K. (1982). Linear fuzzy regression analysis with fuzzy models. *IEEE Transactions on Systems Man, and Cybernetics: SMC, 12*, 903–907.
55. Tanaka, H. (1987). Fuzzy data analysis by possibilistic linear models. *Fuzzy Sets and Systems, 24*, 363–375.
56. Kechagias, P. S., & Papadopoulos, B. (2006). Fuzzy regression analysis software. Modelling of economic & management systems and engineering systems using fuzzy linear regression. General Secretariat of Research and Technology.
57. Newey, W., & West, K. (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*, 703–708.
58. Durbin, J., & Watson, G. (1951). Testing for serial correlation in least squares regression. *Biometrika, 38*, 159–178.

Index

C.L. Dunis et al. (eds.), *Artificial Intelligence in Financial Markets*, New Developments in Quantitative Trading and Investment,
DOI 10.1057/978-1-137-48880-0