Improving risk-adjusted performance in high-frequency trading: The role of fuzzy logic systems

Vincent Vella

A thesis submitted for the degree of

Doctor of Philosophy

Centre for Computational Finance and Economic Agents (CCFEA)

University of Essex

CCFEA January 2017

Abstract

In recent years, algorithmic and high-frequency trading have been the subject of increasing risk concerns. A general theme that we adopt in this thesis is that trading practitioners are predominantly interested in risk-adjusted performance. Likewise, regulators are demanding stricter risk controls.

First, we scrutinise conventional AI model design approaches with the aim to increase the risk-adjusted trading performance of the proposed fuzzy logic models. We show that applying risk-return objective functions and accounting for transaction costs improve out-of-sample results. Our experiments identify that neuro-fuzzy models exhibit superior performance stability across multiple risk regimes when compared to popular neural network models identified in AI literature. Moreover, we propose an innovative ensemble model approach which combines multiple risk-adjusted objective functions and dynamically adapts risk-tolerance according to time-varying risk.

Next, we extend our findings to the money management aspects of trading algorithms. We introduce an effective fuzzy logic approach which dynamically discriminates across different regions in the trend and volatility space. The model prioritises higher performing regions at an intraday level and adapts capital allocation policies with the objective to maximise global risk-adjusted performance.

Finally, we explore trading improvements that can be attained by advancing our type-1 fuzzy logic ideas to higher order fuzzy systems in view of the increased noise (uncertainty) that is inherent in high-frequency data. We propose an innovative approach to design type-2

models with minimal increase in design and computational complexity. As a further step, we identify a relationship between the increased trading performance benefits of the proposed type-2 model and higher levels of trading frequencies.

In conclusion, this thesis sets a framework for practitioners, researchers and regulators in the design of fuzzy logic systems for better management of risk in the field of algorithmic and high-frequency trading.

I would like to dedicate this thesis to my loving parents …

Acknowledgements

My most sincere thanks must go to my supervisor, Dr. Wing Lon Ng. His deep insight, admirable dedication, disciplined and structured approach, and keen eye for detail, were not just key attributes that assisted me in the completion of this thesis, but have also inspired me with invaluable life lessons. For this I am forever in his debt.

I would also like to thank Prof. Edward Tsang for his assistance and direction, especially towards the end of my journey. I will remain grateful for his valuable feedback.

My appreciation also goes towards Prof. Hani Hagras for his helpful pointers and discussions.

Last but not least, this thesis would not have been possible without the person who was always behind me - my wife Mary Ann. Her unconditional support was always a source of encouragement that kept me going during the research journey.

Related Publications

Vella, V. and Ng, W. L. (2014a). Enhancing risk-adjusted performance of stock market intraday trading with Neuro-Fuzzy systems. *Neurocomputing*, 141:170 – 187.

Vella, V. and Ng, W. L. (2014b). Enhancing intraday trading performance of neural network using dynamic volatility clustering fuzzy filter. In *Proceedings of the 2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, London, pages 465–472. IEEE.

Vella, V. and Ng, W. L. (2015). A Dynamic Fuzzy Money Management Approach for Controlling the Intraday Risk-adjusted Performance of AI Trading Algorithms. *Intelligent Systems in Accounting, Finance and Management*, 22(2):153-178.

Vella, V. and Ng, W. L. (2016). Improving Risk-adjusted Performance in High Frequency Trading Using Interval Type-2 Fuzzy Logic. *Expert Systems with Applications*, 55:70-86.

Table of Contents

List of Figures xv

List of Tables xvii

1 Introduction 1

1.1 Background and motivation . . . 1

1.1.1 Trading in financial markets . . . 1

1.1.2 The architecture of a trading algorithm . . . 6

1.1.3 Aligning algorithmic and investors' risk preferences . . . 13

1.2 Research objectives and thesis structure . . . 16

1.3 Contributions . . . 19

2 Fuzzy logic background 25

2.1 Introduction . . . 25

2.2 Type-1 fuzzy sets and fuzzy membership functions . . . 27

2.3 Type-1 TSK fuzzy logic systems . . . 29

2.4 Designing fuzzy logic systems . . . 31

2.4.1 Structure identification . . . 31

2.4.2 Parameter tuning using least squares and back-propagation . . . 33

2.5 Type-2 fuzzy sets and fuzzy membership functions . . . 38

2.6 Extending type-1 TSK fuzzy logic systems to type-2 . . . 40

2.7 Conclusion 43

3 Improving the risk-adjusted performance of trading algorithms 45

3.1 Introduction 46

3.2 Method 51

3.2.1 Technical trading and moving averages 52

3.2.2 Neural network model 56

3.2.3 ANFIS model 57

3.2.4 DENFIS model 60

3.2.5 ANFIS ensemble 63

3.2.6 Trading algorithm 65

3.2.7 Model training and evaluation 67

3.2.8 Experiment data 74

3.3 Experiment results 77

3.3.1 Results for experiment 1 77

3.3.2 Results for experiment 2 86

3.4 Conclusion 93

4 Time-varying intraday volatility and risk-adjusted performance 97

4.1 Introduction 98

4.2 Part 1 - Enhancing the intraday trading performance of neural networks using a dynamic volatility clustering fuzzy filter 101

4.2.1 Method 102

4.2.2 Technical indicators and standard neural network model 104

4.2.3 Standard trading algorithm 109

4.2.4 Intraday volatility and its challenges 111

4.2.5 Enhancing trading model with a dynamic volatility filter 115

4.2.6 Model selection and experiment approach 121
4.2.7 Analysis and results 122
4.2.8 Conclusion 125
4.3 Part 2 – A dynamic fuzzy money management approach for controlling the intraday risk-adjusted performance of AI-trading algorithms 127
4.3.1 Method 130
4.3.2 Enhancing money management decisions by considering time-varying intraday risk 133
4.3.3 Experiment approach and considerations 137
4.3.4 Results and evaluation 140
4.3.5 Conclusion 149

5 Technical rules and uncertainty at higher trading frequencies 153

5.1 Introduction 154
5.2 Method 160
5.2.1 High-frequency data and technical indicators 161
5.2.2 Designing and tuning a TSK type-1 fuzzy model 163
5.2.3 Generalizing ANFIS model to T2 FLS 169
5.2.4 Trading algorithm 177
5.3 Experiment approach 178
5.3.1 Data 179
5.3.2 Performance measures 181
5.3.3 IT2 design considerations 182
5.3.4 Model training and testing 182
5.3.5 Controlling different levels of uncertainty 184
5.4 Results and analysis 185
5.4.1 Experiment 1: Comparison against standard ANFIS model 185

5.4.2 Experiment 2: Comparison of T1 and T2 models under different trading frequencies 191
5.5 Conclusion 195

6 Conclusion 199
6.1 Summary of presented work 200
6.2 Contributions 202
6.3 Limitations and future work 208

References 213

Appendix A Further results from research chapters 223

List of Figures

1.1 General components of a trading algorithm 7

2.1 Components of a fuzzy rule-based system 26

2.2 T1 Gaussian MF 29

2.3 T1 TSK fuzzy model example 30

2.4 Projection of identified clusters on the input space for MF identification . . 33

2.5 T2 primary MF 38

2.6 General T2 secondary MF 39

2.7 Interval T2 secondary MF 40

2.8 T2 fuzzy rule-based system 41

3.1 Experiment setup 51

3.2 Our proposed eANFIS architecture 64

3.3 Moving window approach adopted for model training and testing 69

3.4 Heat map identifying the sensitivity of the ANFIS model 78

3.5 Intraday trading pattern of our trading algorithm using ANFIS and Sharpe ratio optimisation 91

4.1 Experiment setup showing our proposed locally enhanced NN model 102

4.2 Neural network setup 108

4.3 Effect of applying different return thresholds on the number of trades 110

4.4 Realised volatility signature plots 114
4.5 Decision surface example as identified by our volatility clustering fuzzy filter 119
4.6 Effect of applying different levels of return thresholds on the number of trades 131
4.7 Example of clusters found in Hammerson (HMSO) data 134
4.8 Decision surface example for Hammerson (HMSO) as identified by our fuzzy controller showing different degrees of Sharpe ratio regions. 135
4.9 Intraday trading pattern of our trading algorithm using the standard NN model 146
4.10 Average reduction of trades placed by the NN-FMM model 148

5.1 Experiment setup showing our proposed T2 extension 161
5.2 Conversion of T1 fuzzy set to IT2 with fixed mean and uncertain standard deviation 174
5.3 An example showing the conversion of T1 MFs for one input variable to IT2 175

List of Tables

3.1 Parameters tested for FFN 57

3.2 Parameters tested for ANFIS 59

3.3 Parameters tested for DENFIS 63

3.4 Models applied in the experiments 68

3.5 Descriptive statistics of 5-minute returns for stocks utilised in the experiments. 75

3.6 Average 5-minute statistics for selected stocks. 76

3.7 Model performance using Sharpe ratio optimisation over the 100-day out-of-sample period 80

3.8 Trading performance profile by optimising the Sharpe ratio 81

3.9 Model performance using Sortino ratio optimisation over the 100-day out-of-sample period 83

3.10 eANFIS performance using intraday dynamic switching over the 100-day out-of-sample period 85

3.11 Model performance using Sharpe ratio optimisation with no transaction costs over the 100-day out-of-sample period 87

3.12 Model performance using RMSE optimisation over the 100-day out-of-sample period 88

3.13 Model performance using period return optimisation over the 100-day out-of-sample period 89

3.14 Model performance using moving average (MA) rules over the 100-day out-of-sample period 90
3.15 Results obtained from the random model following 100-day out-of-sample period. 92

4.1 Parameters tested for NN model 108
4.2 Additional parameters tested for the NN and the fuzzy volatility filter 121
4.3 Trading performance of the Buy & Hold model following the 100-day out-of-sample period. 122
4.4 Trading performance following the 100-day out-of-sample period. 124
4.5 Combined parameters tested for the NN model and trading algorithm 132
4.6 Additional parameters tested for the NN and fuzzy volatility filter 137
4.7 Trading performance of tested models following the 100-day out-of-sample period 142
4.8 Trading performance of neural network and NN-FMM models following 50 repeated trials over the 100-day out-of-sample period 144
4.9 Trading results following 50 independent trials of the random model 149

5.1 Parameters tested for ANFIS 169
5.2 Models applied in the experiments 180
5.3 Standard ANFIS performance after the 100-day out-of-sample period. . . . 186
5.4 Technical indicators performance after the 100-day out-of-sample period. . 188
5.5 ANFIS/T2a and ANFIS/T2b performance improvements against standard ANFIS after 100-day out-of-sample period 189
5.6 ANFIS Sharpe ratio performance after 100-day out-of-sample period across different levels of return threshold (uncertainty). 192
5.7 ANFIS excess return performance after 100-day out-of-sample period across different levels of return threshold (uncertainty). 193

5.8 Summary statistics for the standard ANFIS and the corresponding variations in the ANFIS/T2 models over a 100-day out-of-sample period across different levels of return threshold (uncertainty) 195

A.1 Additional performance measures for neural network and ANFIS using Sharpe ratio optimisation 224
A.2 Additional performance measures for neural network and ANFIS using Sortino ratio optimisation 225
A.3 Sortino ratio and number of trades for neural network, FCM1 and FCM2 following out-of-sample period 226
A.4 Win ratio and profit ratio for neural network, FCM1 and FCM2 following out-of-sample period 227
A.5 Sortino ratio and number of trades for neural network and NN-FMM following out-of-sample period 228
A.6 Additional risk-adjusted performance measures for neural network and NN-FMM following out-of-sample period 229
A.7 Sortino ratio, number of trades and profit per trade for standard ANFIS following out-of-sample period 230
A.8 Sortino ratio, number of trades and profit per trade for ANFIS/T2a following out-of-sample period 231
A.9 Sortino ratio, number of trades and profit per trade for ANFIS/T2b following out-of-sample period 232
A.10 ANFIS number of trades after 100-day out-of-sample period across different levels of return threshold (uncertainty). 233
A.11 ANFIS profit per trade after 100-day out-of-sample period across different levels of return threshold (uncertainty). 234

Chapter 1

Introduction

In this introductory chapter, we describe the adopted rationale which was used to conduct this research. First, we present reflections from three knowledge domains, namely high-frequency trading, financial risk and artificial intelligence, with a special focus on fuzzy logic techniques. The intention is to highlight the key concepts from each domain as a reflection of the current state of literature. In our review, we take a critical approach by identifying a number of open questions and literature gaps that motivate our research. Secondly, we present the main research objectives and how these are addressed in the subsequent chapters.

1.1 Background and motivation

1.1.1 Trading in financial markets

This section sets the scene of our research problem domain by highlighting the mechanisms of the trading environment as well as the decisions and actions with which traders are faced on a daily basis. The understanding of these elements, together with the interdependence with other key variables, such as price, volume, liquidity and volatility, are key in order to identify opportunities where fuzzy logic can be of beneficial assistance to traders, including the possibility of (intelligently) automating whole processes.

Johnson (2010) divides the trading process into three stages: (i) price formation, (ii) price discovery or trade execution, and (iii) reporting, clearing and settlement. In this thesis, our research overlaps with the first two stages.

Price formation is the process through which the price of an asset is determined. The price can be perceived as the consolidated investors' view of the future value of an asset. Different investors tend to have different views in terms of asset valuation, and hence, price formation is the result of supply and demand conditions. An understanding of market structure elements is crucial because they determine what people can know, and hence, what drives the actions and relations among the different investors, ultimately resulting in determining who will trade profitably.

The type of mechanism adopted in a specific market is typically either quote-driven or order-driven. In this thesis, our research is conducted using trade data from the London Stock Exchange's electronic order book – the Stock Exchange electronic Trading Service (SETS). SETS is one of the most liquid electronic order books in Europe, hence benefiting from lower transaction costs. Algorithmic, especially higher frequency, trading is typically characterised by a large number of orders with small order quantities, short holding periods and no overnight positions (Aldridge, 2013; Brogaard et al., 2014). This justifies our selection of SETS (offering high liquidity and volumes) for our experiments. Moreover, as shown by Aitken et al. (2015), HFT studies using SETS data are scarce.

Similar to most exchanges, LSE employs a continuous auction mechanism with orders entered between 8:00 and 16:30. Price discovery happens when supply and demand requirements intersect, hence determining the price of an asset. The constant tension between buyers and sellers dictates the dynamic (non-stationary) movement of market prices. Traders' actions are driven by the information available and the deductions they extract from it. To formulate these deductions, two approaches are commonly adopted, namely fundamental and technical analyses. Fundamental analysis focuses on company fundamentals information, such as financial forecasts and positioning with respect to competitive products. On the other hand, technical analysis is a methodology which primarily considers price and volume movements. The main aim of both approaches is to assist traders with the analysis of the information they receive and thus be in a position to formulate an opinion on the possible shifts in supply and demand, hence determining the direction in which prices are likely to move.

In this thesis, our interest is in short-term intraday trading, hence it is more speculative in nature. This makes technical analysis a more adequate tool and rests on the underlying assumption that prices move in trends (hence asserting price momentum) (Murphy, 1987). Harris (2002) defines speculators as traders who trade to profit from information and predictions about future prices. This also depends on the accuracy (and interpretation) of the signals generated from the information received. Typically, technical traders apply a set of technical indicators, which normally consist in an arsenal of statistical measures and charts, to assist with the analysis of the noisy price time series and with the prediction of short-term market moves. From the concluded analysis, speculators choose between buy or sell actions based upon their projected price movements. However, Bao and Yang (2008) show that capturing a true price turning point can be a difficult challenge since most of the technical indicators are price followers and, hence, it is difficult to infer whether a trend will continue or break. The difficulty is intensified due to the prevalence of higher market volatility. This presents the key challenge for trend-following trading strategies, which we investigate in this thesis.

The debatable predictability of financial markets is rooted in the Efficient Market Hypothesis (EMH) (Fama, 1965, 1970). According to EMH, markets are efficient and all attempts at predicting market prices are futile as the prices already incorporate all the information that could affect them. The theory interprets price movements as a random walk by hypothesising that since only new information moves stock prices significantly, and since new information arrives at random, hence future movements in stock prices follow a random path. EMH (in its three forms) makes two important assumptions, that all public information is immediately available to the market, and that many traders are perfectly rational (given the same information, traders reach the same optimal asset valuation). However, in the literature we find a number of theoretical contributions which attribute departures from EMH. Simon (1972) points out that decision makers act within bounded rationality, hence their rational choice is within the space limited by their knowledge and processing capacity. Tsang (2009) argues that even in the case of two perfect decision makers, the amount of time taken to arrive at the optimal decisions might be different. For example, the time taken for a first agent to pick up an investment opportunity might be quicker than others. Moreover, a number of authors associate decision making with human emotions and psychological factors, such as overconfidence, herding, fear and regret – all sources of irrationality that lead to market inefficiencies (see Lo et al., 2005; Shiller, 1999).

These theoretical conflicts on market efficiency have fueled incessant empirical research in the search for approaches that can glean trading profitability. Gençay (1996) and LeBaron (1999) demonstrated that simple technical rules can predict market returns. An interesting study by Schulmeister (2009) found that beyond the 1990s, using daily data, technical trading rules became unprofitable. The same study, however, showed that by shifting to intraday 30-minute data, technical rules registered profitability again until the year 2000. Beyond the year 2000, another decline in technical rules' profitability followed till 2006. Schulmeister suggested that beyond the year 2006, either markets became more efficient or stock price trends moved to higher frequency prices. However, Kearns et al. (2010) claimed that, after taking account of transaction costs, aggressive high-frequency trading leads to relatively low profitability, far below the expected high excessive returns. Later, Holmberg et al. (2013) found empirical evidence of intraday trending in stock prices, linking the profitability of technical rules with days which fall during periods exhibiting higher levels of volatility. In another study, Rechenthin and Street (2013) conducted empirical tests that indicated that, in general, a timespan of 30 minutes is required for stock prices to become efficient; however, no indication of possible profitability was investigated. In this thesis, we aim to contribute new findings to this ongoing debate.

There are also conflicting opinions amongst researchers in terms of the positive and negative market effects of algorithmic and high-frequency trading. Brogaard et al. (2014) recognise the wider adoption of algorithmic and high-frequency trading as a natural evolutionary path in market development, driven by advances in technology, and attribute this advancement with improvements in price discovery, increased liquidity, and with no negative effects on volatility. Zhang (2010) argues that high-frequency trading amplifies price reactions and volatility. Following incidents like the 2008 financial crisis and the flash crash of 2010, algorithmic trading and high-frequency trading began to attract more attention from regulatory bodies. Regulators' concern was heightened due to the development of special order types like hidden orders, icebergs, and the use of dark pools. In 2014, the European Commission introduced MiFID II (MiFID, 2014), a new regulatory regime for firms which engage in algorithmic or high-frequency trading. MiFID II and delegated acts under MiFID II will apply from January 2018 (this follows a recent one-year extension proposed by the Commission). Any firm carrying out algorithmic trading (including high-frequency trading) is subject to numerous organisational requirements. In particular, under Article 17 (1)(2), investment firms practising algorithmic trading are subjected to new rules which ensure resiliency, sufficient capacity, adequate risk controls, effective backtesting, and business continuity arrangements. Moreover, these firms are required to be more transparent regarding the type of algorithmic trading strategies adopted, the range of trading parameters or limits to which their systems are subjected, the risk controls applied and details of backtesting results. These new regulations underscore the importance of improving risk controls and the risk-adjusted performance of trading algorithms. In this thesis, we aim to present innovative techniques, based on fuzzy logic, that can assist in addressing these objectives.

In the next section, we analyse the main building blocks of trading algorithms and how AI can contribute to improving them.

1.1.2 The architecture of a trading algorithm

The advancement in computing power and general network (internet) bandwidth, coupled with increased direct access trading, have boosted the opportunity for the adoption of sophisticated algorithms to search for quick intraday profit opportunities. In this new reality, Tsang (2009) states that traders' rationality is measured by the solution (algorithm) optimality, hence from a computational point of view, traders' rationality is reflected by their computational intelligence, the better algorithms translating into increased trading opportunities. This puts into perspective the importance of our area of research.

In the literature, one finds different definitions of algorithmic and high-frequency trading. We adopt the definitions as stated in MiFID II. Article 4(1)(39) of MiFID II defines algorithmic trading as:

> "[... *trading in financial instruments where a computer algorithm automatically determines individual parameters of orders such as whether to initiate the order, the timing, price or quantity of the order or how to manage the order after its submission, with limited or no human intervention, and does not include any system that is only used for the purpose of routing orders to one or more trading venues or for the processing of orders involving no determination of any trading parameters or for the confirmation of orders or the post-trade processing of executed transactions.*"

Similarly, Article 4(1)(40) of MiFID II defines high-frequency algorithmic trading as:

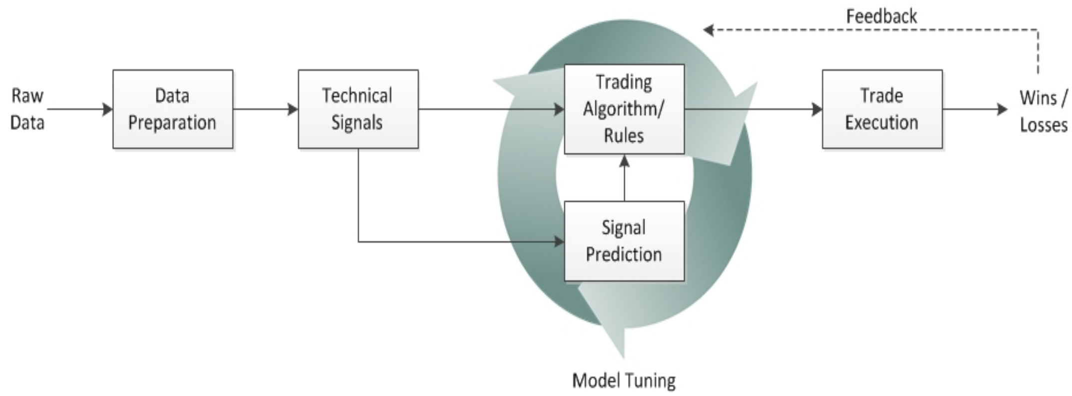


Fig. 1.1: General components of a trading algorithm

> “[*... an algorithmic trading technique characterised by: (a) infrastructure intended to minimise network and other types of latencies, including at least one of the following facilities for algorithmic order entry: co-location, proximity hosting or high-speed direct electronic access; (b) system-determination of order initiation, generation, routing or execution without human intervention for individual trades or orders; and (c) high message intraday rates which constitute orders, quotes or cancellations.*”

These definitions place high-frequency trading as essentially a subset of algorithmic trading but where the focus is on higher trading speeds. In Figure 1.1 we present a high level architecture of the trading solutions that we adopt in this thesis. From a computational perspective, an algorithmic trading solution can be viewed as a series of interconnected blocks of functions, each representing a loosely decoupled software module, with defined inputs and outputs, that permit independent development, updates and testing. Apart from being good software practice, loosely decoupling the modules is ideal in view of MiFID II obligations, which state that any substantial update of an algorithmic trading system requires rigorous testing and approvals.

Below, we provide a description of each module and the corresponding gaps that we identified from the literature:

Data preparation module

This module receives raw price data, typically in the form of ticks (stock, quantity, price) from data providers, and transforms the data into the format which will later be required by the model. Although it depends on the type of model adopted, a typical transformation converts irregular time interval points into regular ones (daily, hourly, 30-minute, 10-second, etc.) by taking the last price point before the end of each time bar. In spite of the persisting contention about the feasibility of trading in the high-frequency space, surveys (Cavalcante et al., 2016; Krollner et al., 2010; Tsai and Wang, 2009) show that many many former studies still utilise daily data. In this thesis our focus is to identify innovative techniques of how technical trading can be combined with fuzzy logic techniques for intraday trading.

Signal generation module

This module takes as an input the time series from the data preparation module and generates another time series of either transformed or smoothed variables. Kaastra and Boyd (1996) suggest that most common data transformations include first differencing and taking natural logs. The same authors advise to use smoothing techniques for both input and output data (e.g. using moving or exponential moving average), since trend prediction using only price changes or no noise filtering can be a difficult challenge for the underlying model. Tsai and Wang (2009), Krollner et al. (2010) and Cavalcante et al. (2016) indicate that the many former studies make use of lagged index data, including the use of technical indicator signals like moving averages, RSI, Bollinger bands, etc. Typically, the data output from this module is also converted into a format as required by the prediction model. For example, if the prediction model takes, as its input, three technical signals to predict the next return, the data is organised as a matrix consisting of four columns, where three columns represent the inputs (lagged signals) and one column represents the output. The inputs might also include feedback from the model output - this typically helps the prediction model to adapt to changing market conditions. A similar combination of signals, in conjunction with fuzzy logic, is used in this thesis.

From the AI literature, a common approach that is adopted when training a machine learning model (with the objective of identifying a model with good generalisation capabilities) involves splitting the data into three distinct data sets that are used for training, testing and validation (out-of-sample). Bailey et al. (2014) warn that many computational finance studies only use in-sample data or else short out-of-sample periods, hence increasing the possibility of obtaining spurious results. In this thesis, we note these concerns and follow the suggestions from Kaastra and Boyd (1996) and Pardo (2011), who suggest the adoption of a moving window (walk forward) approach. This approach is typically used for testing trading systems and constitutes a more rigorous and realistic methodology. This involves splitting the data into overlapping training-test-validation sets, and on each cycle moving each set forward through the time series. This approach tends to result in more robust models due to more frequent retraining and large out-of-sample data set (increasing training processing requirements but also resulting in models which adapt more quickly to changing market conditions).

Prediction module

Foucault and Roşu (2016) state that a stylised fact of high-frequency traders is that their aggressive (marketable) orders anticipate very short-run price changes. Brogaard et al. (2014) claim that high-frequency traders predict short-term price changes; by using marketable orders, trade in the direction of true price changes and filter out transitory pricing errors, sufficient informational advantage is generated in order to break the underlying costs threshold. Gençay et al. (2001) and Johnson (2010) state that the ability to predict trends, for example using data mining or artificial intelligence techniques, can offer an edge to traders. In view of the stochastic nature of market prices, increased volatility and unexpected market shifts (hence increased uncertainty), this poses a difficult challenge.

By increasing the granularity at which a time series is analysed, the more trends can be identified, hence increasing trading opportunities; however, being excessively sensitive to short-term fluctuations (hence resulting in position times which are too short) can result in losses, especially due to transaction costs (Kearns et al., 2010). The objective for an optimal algorithm is to identify trends as early as possible until the model signals that the trend has reversed. In this thesis we aim to contribute to existing literature by proposing new robust prediction models combining fuzzy logic and technical rules.

In the literature, we find that popular time series models include regression methods and the ARIMA models (Box et al., 2015), but these models exhibit limitations particularly because of their linear composition (see Lin et al., 2002). Surveys by Tsai and Wang (2009) and Krollner et al. (2010) indicate the popularity to use established Artificial Neural Networks (ANNs) in stock price forecasting and enhance them with new training algorithms or combine ANNs with evolutionary and optimisation techniques into hybrid systems. Following the emergence of fuzzy logic (Zadeh, 1973), neural networks and fuzzy inference systems were brought together as general structures for approximating non-linear functions and dynamic processes. A popular model in this category is the Adaptive Neuro-Fuzzy Inference Fuzzy System (ANFIS) Jang (1993). Fuzzy logic models were designed to better manage the prevalent uncertainties of the underlying complex systems, which is the case of financial markets. The learning capabilities of neural networks and the increased ability of fuzzy logic models to handle uncertainties are the key features which drove our interest to identify innovative hybrid models that can improve the performance of trading algorithms.

Following the surveys conducted by Tsai and Wang (2009), Krollner et al. (2010) and Cavalcante et al. (2016), we argue that the criteria that is used by many AI studies to evaluate and compare prediction models do not necessarily translate into profitable trading solutions. For example, Brabazon and O'Neill (2006) state that achieving the best root mean squared error does not indicate profitability for trading purposes, since smaller errors, but wrong side predictions, will still result in losses. Similarly, a high directional accuracy model might still suffer from fewer, but larger, losses. We address this gap by proposing better fuzzy logic model design methods for algorithmic trading purposes, especially in higher frequency intraday trading scenarios.

Trading algorithm module

The trading algorithm can be considered as a rule engine which, based on the technical signals and predictive signals as inputs, automates trading decisions. Covel (2009) identifies a number of decisions that are typically managed by a trading algorithm: (i) what to buy or sell, (ii) position sizing, (iii) market entries, (iv) trading stops, and (v) trading exits. Each condition is typically conditional on specific thresholds which need to be tuned as part of the model calibration process. For example, Vanstone and Finnie (2009, 2010) suggest a fixed return threshold filter to avoid small unprofitable movements; however, Holmberg et al. (2013) warn that although the use of higher fixed thresholds increase prediction accuracies, this also results in reduced trading opportunities. Similarly, Brabazon and O'Neill (2006) suggest a volatility filter which limits trading during high volatility (risk) periods, hence avoiding possibly strong adverse market movements. On the other hand, a number of authors (Han et al., 2013; Holmberg et al., 2013; LeBaron, 1999; Schulmeister, 2006) indicate that during high volatility periods, markets tend to become less efficient, hence the better chances for technical rules' profitability. We conclude that more research is required that investigates the automation of adaptive thresholds along intraday time-varying risk. We address this gap in our thesis using innovative fuzzy logic techniques.

Moreover, we propose a methodology which jointly calibrates the trading algorithm parameters with the prediction model parameters, and hence the adopted performance measure considers the joint performance of both the prediction model and the trading algorithm. Although Vanstone and Finnie (2009) and Brabazon and O'Neill (2006) indicate approaches which combine the tuning of the prediction model and the trading algorithm parameters jointly, rather than separately, we note that many published studies stop short at the prediction model (see survey by Cavalcante et al., 2016).

Pardo (2011) states that many AI studies focus solely on trend prediction, but decisions like order position sizing are often not appreciated and poorly understood in trading strategy design. This possibly leads to sub-optimal solutions. Additionally, firms engaging in algorithmic trading need to incorporate rules in their system that ensure that they abide by MiFID II regulations – for example, based on the capital base, there are limitations on trading style, maximum order value, maximum order volume, and maximum long-short positions. Vanstone and Finnie (2010) argue that many published studies do not consider real-world intraday market characteristics like trading costs, realistic trading hours and no overnight positions, possibly leading to biased results (Brabazon and O'Neill, 2006; Pardo, 2011). In this thesis we propose innovative techniques to address several of these conditions using fuzzy logic.

Trade execution module

This module handles the decisions passed on by the trading algorithm. It typically consists in a set of instructions to execute orders in a specific way. Johnson (2010) states that hundreds of trading algorithms exist in this category (for example, time-weighted average price, percentage of volume and implementation shortfall). Although AI algorithms can be used to improve these execution algorithms (for example, see Kablan and Ng, 2010), it was our conscious decision to apply a simple execution mechanism and focus on improving aspects related to the prediction and algorithmic trading modules only (specifically, trend-following strategies). This does not exclude the need for possible future research in other areas.

1.1.3 Aligning algorithmic and investors' risk preferences

The new regulations, like the upcoming MiFID II, emphasise the importance that trading systems are adequately designed and tested to mitigate the risks to which they are exposed. This has to be demonstrated by the (resilient) performance obtained through simulations and backtesting of algorithms. MiFID II lists a number of risks of algorithmic trading, like the overloading of systems at trading venues, the generation of duplicative or erroneous orders, and overreactions to market events, hence boosting volatility. We take the perspective of an investor and revert to a fundamental concept in standard economics literature which states that investors are concerned with both risk and return (Markowitz, 1952). Markowitz argues that if an investor knew the future returns with certainty, then choosing an investment option would merely boil down to picking the security which offers the highest return. However, in the presence of uncertainty, which is inherent in financial markets, the Markowitz mean-variance model allows an investor to seek the highest return at an acceptable level of risk – risk being measured by variance. Motivated by the mean-variance model, Sharpe (1966) proposed the Sharpe ratio as a risk-adjusted measure to compare investment performance. To date, it remains one of the most popular risk-adjusted performance measures due to its practical use, applying standard deviation as the measure of risk. Eling (2008), Prokop (2012) and Auer and Schuhmacher (2013) show the proliferation of risk-adjusted performance measures and their tests demonstrate that despite it's shortcomings, the Sharpe ratio indicates similar performance rankings of the more sophisticated performance ratios. In view of these findings, in this thesis the Sharpe ratio is adopted as one of the key risk-adjusted performance measures. This is also complemented by other risk-adjusted measures, mainly the Sortino ratio and the Calmar ratio, which we utilise in various experiments.

In finance, volatility is the statistical measure of stock price dispersion, hence the movement of a stock price without regard to direction. Large (small) average daily stock price movements mean high (low) volatility. From one perspective, volatility is linked with unstable market scenarios with the possibility of abrupt larger losses. However, from a different perspective, higher volatility can be interpreted as an opportunity for higher profits. This dilemma is what links volatility with market uncertainty (risk) and more challenging forecasting. Andersen et al. (2009) state that measuring and forecasting volatility is a core topic in finance literature. The ARCH model, introduced by Engle (1982), followed by the GARCH model, generalised by Bollerslev (1986), promoted time-varying volatility to a very active area of research. Over the years, numerous improvements were proposed, such as the ability to incorporate the nonlinearity, asymmetry and long memory properties in the volatility process. Popular extensions include the EGARCH model (Nelson, 1991) and the GJR-GARCH model (Glosten et al., 1993) (for a review on ARCH models, see Hansen and Lunde, 2005; Poon and Granger, 2003).

Merton (1980) noted that the volatility of a Brownian motion can be approximated to an arbitrary precision using the sum of intraday squared returns, a measure which was coined as realised variance. Later, Andersen et al. (2000c) showed that ex-post daily foreign exchange volatility is best measured by aggregating intraday squared five-minute returns. Consequently, volatility literature has steadily progressed towards the use of higher frequency data. This was facilitated by the fact that high-frequency data became much more accessible. Liu et al. (2015) identify around 400 combinations of realised variance estimators from the literature, and also refer to various parameters that practitioners must fine tune including the sampling size, sampling scheme, whether to use transaction prices or mid-quotes, and other measure specific parameters such as the kernel length.

Despite these advancements, we still identify many AI studies with application to finance (Brabazon and O'Neill, 2006; Cavalcante et al., 2016; Krollner et al., 2010; Pardo, 2011; Tsai and Wang, 2009) that clearly indicate that risk is given much less attention, especially in intraday trading scenarios. Many former studies tend to focus on the application of models solely to predict market movements (a more recent example can be found in Son et al.,

2012) and, for model evaluation, it is common practice that error measures, win ratios or profitability are applied, but with little evaluation of risk-adjusted performance. Moreover, although new measures of variation make it possible to measure and predict volatility with a good degree of accuracy up to an intraday level, the use of this information for intraday trading purposes is rarely considered. We identify that this is even more scarce in the fuzzy logic domain with application to finance.

We extend the thoughts of Lo et al. (2005) and Shiller (1999) with regard to the effect of human emotions in trading to a computational perspective. We argue that algorithms are programmed encapsulations of the thought processes of the traders who design them, hence the same human elements can be incorporated in the algorithms (e.g. applying different objective functions and risk considerations). Consequently, in line with this thesis, researching the application and performance of risk-adjusted objective functions exhibits higher resonance with investors' risk sensitive preferences.

Gradojevic and Gençay (2013) indicate the benefits of combining type-1 fuzzy logic with standard technical indicators in view of increased trading uncertainties. Other recent applications of fuzzy logic models in finance have been demonstrated by a number of authors (Boyacioglu and Avci, 2010; Chang et al., 2011; Chen, 2013; Kablan and Ng, 2011; Tan et al., 2011; Wei et al., 2014). We argue that although one of the key claimed advantages of fuzzy logic, when compared to other AI techniques, is the better management of uncertainty, the link (and feasibility) between fuzzy logic and risk-adjusted trading performance gains, to our knowledge, was never explored in previous studies. Mendel et al. (2006) and Wagner and Hagras (2010) demonstrate the increased capabilities in managing uncertainty that is possible by using type-2 fuzzy logic when compared to the earlier type-1. However, this comes at the cost of additional complexity, which is typically the main reason why many practitioners still favour the use of type-1 (Wu and Mendel, 2014). Aladi et al. (2014) suggest that it is not clear what level of uncertainty warrants the (feasible) use of type-2 fuzzy logic (in some cases, no benefits were achieved). Moreover, we have not identified any previous studies which explore the relationship between the move from type-1 to type-2 fuzzy logic and any potential additional gains in profitability and risk-adjusted performance in an intraday trading scenario. In this thesis we investigate the benefits of both type-1 and type-2 fuzzy models.

In the spirit of the Sharpe ratio, we contend that investors' risk is related to the performance stability of the underlying trading models across continuously changing market (risk) conditions. From a computational perspective, Marsland (2009) suggests that systems have to adapt, and hence evolve, to address recurring and changing patterns in the intraday environment which are driven by the actions of informed and uninformed traders. Implementing evolution requires an ability to balance learning about new information while still respecting past accumulated knowledge. We also note issues with regard to approaches that only compare model performance at a single point in time following a defined out-of-sample period (Kaastra and Boyd, 1996; Vanstone and Finnie, 2009, 2010). We contend that this approach does not necessarily reflect the performance stability of the underlying models across the different market conditions during the out-of-sample period. Pardo (2011) emphasises the importance of backtesting algorithms using a set of stocks with different trends – this avoids any bias towards particular trends and hence examines the stability of the tested models across different market scenarios. From the reviewed studies and surveys, we identify that these conditions are rarely considered in the domain of fuzzy logic in the control of trading algorithms, leaving a research gap for further investigation on the profitability and risk-adjusted performance of fuzzy logic controlled trading algorithms under stricter (and time-varying) conditions.

1.2 Research objectives and thesis structure

The approach developed in this thesis is rooted in finding better ways of how to manage the intimately intertwined concepts of uncertainty and risk as an integral part in the design of algorithmic and high-frequency trading solutions. The adopted framework cuts across three knowledge domains, namely high-frequency trading, financial risk and artificial intelligence, with special focus on fuzzy logic techniques.

Uncertainty (risk) is a core consideration in the design of trading systems. The source of this uncertainty originates from the underlying dynamic and complex market processes. This makes the problem of predicting financial markets, and achieving consistent profitable results, a very difficult one and the subject of ongoing research.

The management of uncertainty is a central theme within the broad framework of fuzzy set theory. Zadeh (1997, p. 123) states that “the guiding principle of soft computing is to exploit the tolerance of imprecision, uncertainty, and partial truth to achieve tractability, robustness, low solution cost, and better rapport with reality”. This precludes using conventional approaches that require a detailed description of the problem being solved. In this thesis, we investigate the use of soft computing as a decision-making support for short-term investment strategies.

We summarise our primary research objectives as follows:

1. Identify innovative fuzzy logic based techniques, when applied in order to control trading algorithms, that result in a positive impact on risk-adjusted trading performance.

2. Investigate aspects in the design and architecture of fuzzy logic models that contribute to risk-adjusted trading performance.

From these primary objectives, we identify a number of secondary objectives which need to be addressed in order to achieve the primary objectives:

1. Investigating the risk-adjusted performance of money-losing algorithms is pretty much pointless. Hence, we aim to explore the debated profitability of technical trading rules, in conjunction with neuro-fuzzy models, focusing particularly on high-frequency data in an intraday trading scenario.

2. Profitability and risk can be considered as two sides of the same coin. In our research, we intend to identify techniques of how fuzzy logic can assist in improving the risk-adjusted performance of intraday trading models and the explore corresponding effect on profitability.

3. Identify the effect of including (or omitting) real-world intraday trading constraints from fuzzy logic controlled trading algorithms like trading costs, realistic trading hours and no overnight positions, which are typically ignored in existing studies, possibly leading to biased results.

4. Identify a measure of time-varying intraday risk and how neuro-fuzzy models can benefit from this information to dynamically adapt the risk profile and improve the risk-adjusted performance of trading algorithms.

5. Identify whether (and when) advancing from type-1 (T1) to type-2 (T2) fuzzy logic systems provides viable benefits for trading purposes.

6. Increasing model complexity can limit the widespread adoption. In our approach, we consciously aim to identify stepwise improvements on existing models, using fuzzy logic, with minimal increase in model complexity. This also helps in clearly identifying the advantages gained from the proposed approaches (rather than obtaining performance improvements through increased model complexity).

   To address these objectives, we take the following approach:

   In Chapter 2 we cover the core mathematical and theoretical aspects of fuzzy logic as a preliminary background to the more advanced fuzzy models and optimisation techniques that we will make use of in subsequent chapters.

   The objective of Chapter 3 is twofold. Firstly, we analyse the out-of-sample performance of fuzzy logic controlled trading models based on our exploration of different objective functions, including risk-adjusted performance measures based on the Sharpe ratio and the Sortino ratio. We address issues resulting from conventional AI methodologies and present an approach which optimises risk-adjusted performance whilst considering realistic costs and constraints. Secondly, we explore the positive performance and stability contributions that can be attained by adopting fuzzy logic models in comparison to the more popular standard neural networks.

Next, in Chapter 4 we extend our research by addressing the problem of time-varying intraday volatility. We explore risk-return optimisations at a more granular intraday level and investigate the overall performance gains that can be attained by applying fuzzy logic controllers. Our proposed models automatically identify intraday risk scenarios and adapt money management decisions to varying degrees of risk.

As a further step, in Chapter 5 we expand our research by addressing modelling problems resulting from the technical rules uncertainty which is more pronounced in high-frequency trading. In our investigation we analyse the advantages that can be gained by advancing from T1 to T2 fuzzy logic by exploring the effects on risk and return measures at various levels of trading frequencies.

Finally, in Chapter 6 we consolidate our findings and bring together the final reflections emerging from this thesis. We conclude by providing suggestions for further research.

1.3 Contributions

We will now briefly highlight the contributions of the three research chapters in this thesis.

- Chapter 3

  Our first contribution is the introduction of neuro-fuzzy techniques to improve both the profitability and, in particular, the risk-adjusted performance, of standard neural network driven trading algorithms in an intraday trading scenario. As far as we know, this is the first time that fuzzy logic and risk-adjusted performance has been used in uncertainty management. Our results demonstrate that the proposed dynamic moving average approach in combination with ANFIS outperforms the risk-adjusted performance obtained from standard moving average technical rules, popular neural networks models and DENFIS. Our proposed model is further validated using a random model.

As our second contribution, we investigate the adoption of risk-adjusted performance for model selection rather than the error-based objective functions that are more common in the AI field (Brabazon and O'Neill, 2006). In addition, we also analyse the effect of omitting transaction costs in the training process (see surveys in Krollner et al., 2010; Tsai and Wang, 2009). To our best knowledge, a clear assessment of these model design aspects and constraints have never been investigated from the domain of fuzzy logic in an intraday high-frequency setting (as indicated by Bahrammirzaee, 2010; Cavalcante et al., 2016; Krollner et al., 2010, and references therein). We demonstrate that neuro-fuzzy models optimised using risk-adjusted performance functions, in our case utilising the Sharpe ratio and the Sortino ratio, lead to better out-of-sample trading performance. Additionally, we identify further improvements from the introduction of an ANFIS ensemble which combines multiple risk-return objective functions.

Thirdly, we investigate the trading stability of neuro-fuzzy models using the cumulative Sharpe ratio. Although it is very common, for AI studies with application to finance, to take the approach of comparing model performance at a the end of the out-of-sample period, in this case our interest is to evaluate the cumulative performance on a day-to-day basis during out-of-sample testing. We identify that the ANFIS models offer superior performance stability when compared to neural networks and DENFIS.

- Chapter 4

  In this chapter, we address the challenges posed by intraday volatility. As our first contribution, we extend our model improvements, as identified in Chapter 3, and introduce a novel model extension based on clustering techniques and fuzzy logic. This produces a dynamic decision surface that prioritises regions in the trend-volatility space according to trading performance. Our approach goes contrary to studies suggesting the use of fixed return and volatility thresholds (Holmberg et al., 2013; Kaastra and Boyd, 1996; Vanstone and Finnie, 2009, 2010). From our results we show the increase in trading performance that can be obtained by our proposed fuzzy logic enhanced neural network when compared to the standard neural network model.

  As our second contribution we extend our model improvements to the money management decisions of trading algorithms. In line with Pardo (2011), we argue that although effective money management can leave a huge impact on trading performance, however, a number of existing studies do not consider AI-driven money management decisions at an intraday level, particularly in view of risk-adjusted performance improvements (as reflected from surveys in Cavalcante et al., 2016; Krollner et al., 2010; Tsai and Wang, 2009). We present a novel fuzzy logic approach to dynamically adjust trading frequency and capital allocation depending on the varying degrees of risk at an intraday level. Our proposed optimisations show significant improvements in risk-adjusted performance when compared to standard neural network and buy-and-hold methods. We further validate the effectiveness of our model against a random model.

  As our third contribution, we extend the research from other studies that claim a relationship between the profitability of technical trading rules and volatility (Gradojevic and Gençay, 2013; Han et al., 2013; Holmberg et al., 2013; LeBaron, 1999; Schulmeister, 2006). Our contribution stems from the fact of exploring this link from a fuzzy logic perspective at a more granular intraday level and at higher trading frequencies with the objective to improve risk-adjusted performance. To our best knowledge this was not previously addressed in the literature.

- Chapter 5

  In our last research chapter, we explore further improvements by advancing our T1 fuzzy logic models adopted in previous chapters to T2. As our first contribution we propose two innovative methods of how the popular ANFIS model can be improved for trading purposes by introducing Interval T2 (IT2) fuzzy sets with a minimal increase in complexity. Our approach addresses the increased complexity that is typically attributed to the design of T2 fuzzy models, hence increasing the possibility of wider adoption (see Wu and Mendel, 2014).

  As our second contribution, we investigate the performance of T2 fuzzy models in conjunction with technical rules for intraday trading purposes with the objective to improve risk-adjusted performance. To our best knowledge, this is the first time that T2 fuzzy logic is applied in this context. We identify that T2 methods show a significant increase in risk-adjusted trading performance when compared to standard ANFIS and a technical trading approach. Additionally, our results extend the findings of a number of authors (Holmberg et al., 2013; Rechenthin and Street, 2013; Schulmeister, 2009) who claim possible breaks in market efficiency at short time frames. Our trading window of interest lies between 2 to 10 minutes. From our literature review, we have not identified any previous research which investigates risk-adjusted performance at this time interval. As a result of this, we manage to identify a positive link between higher order fuzzy systems and risk-adjusted trading performance.

  Finally, as our third contribution, we provide deeper insight on the benefits of adopting T2 models in conjunction with technical signals from the perspective of different levels of trading risk (uncertainty) and trading frequency. Although a number of authors (e.g. Aladi et al., 2014; Sepulveda et al., 2006) demonstrate the increased capability of IT2 models to handle increased uncertainty when compared to T1, to our best knowledge, we have not identified any previous studies that investigate at what level of uncertainty can T2 translate into tangible trading benefits. From our results, we identify that T2 models provide a significant increase in performance in high-frequency trading scenarios. Less positive impact is observed during lower-frequency trading.

Chapter 2

Fuzzy logic background

This chapter provides a preliminary mathematical background on fuzzy logic with a specific focus on Type-1 Takagi-Sugeno-Kang (T1 TSK) FLSs (Sugeno and Kang, 1988) and Interval Type-2 (IT2) FLSs (Mendel et al., 2006; Zadeh, 1975). The selected review is presented in such a way so as to set out the necessary building blocks for the more sophisticated fuzzy logic techniques which are extensively used in the subsequent chapters.

2.1 Introduction

Fuzzy logic saw its inception in the proposal of Fuzzy Set Theory by Zadeh (1965). The main motivation behind fuzzy logic can be attributed to the *Principle of Incompatibility* (Zadeh, 1973) which, in the face of the increasing complexity of underlying systems, supports a departure from the traditional quantitative techniques in favour to models which can model and minimise the effect of uncertainty. In essence, fuzzy logic and fuzzy set theory offer an alternative approach to conventional system theory, which relies on crisp mathematical models of systems, such as algebraic or difference equations.

Fuzzy Rule-Based Systems (FRBS) are a popular application of fuzzy logic and fuzzy set theory and present a competitive alternative to other classic models and algorithms in

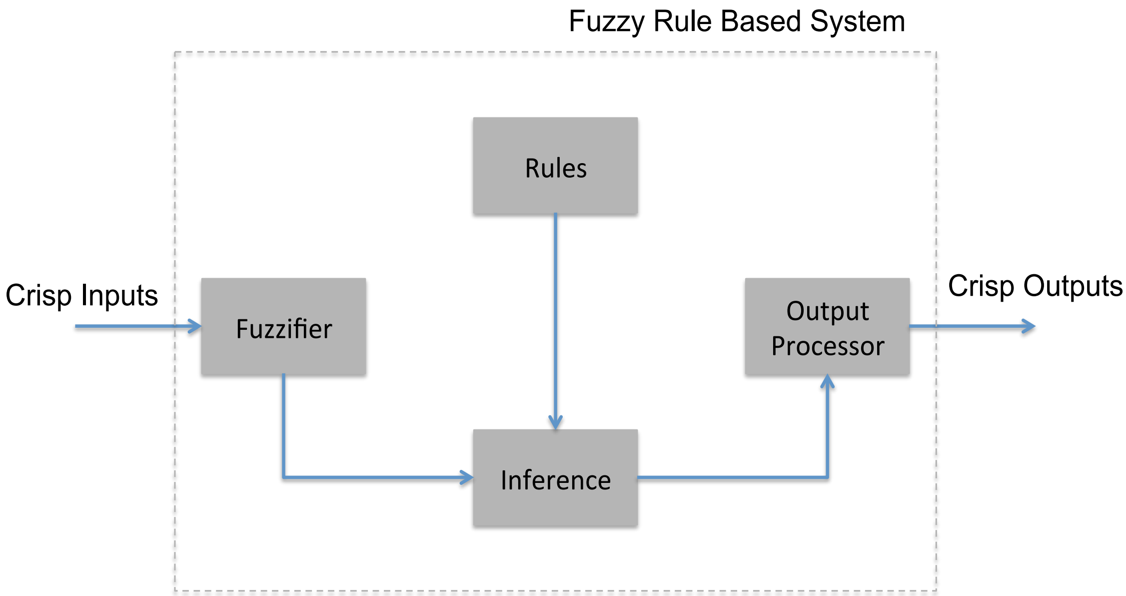


Fig. 2.1: Components of a fuzzy rule-based system

classification and regression problems. An FRBS (Figure 2.1) typically consists of four components (Mendel et al., 2014):

- a *fuzzifier* which takes crisp values as inputs and converts them into degrees of membership of the fuzzy terms of each variable.
- a *knowledge base* which consists of a database holding fuzzy set definitions, and a rule base storing a list of fuzzy IF-THEN rules.
- an *inference engine* responsible for inference operations by applying the fuzzy IF-THEN rules.
- a *defuzzification component* to convert fuzzy values into crisp outputs.

The first fuzzy models were focused on linguistic fuzzy modelling and their semantic interpretability capabilities. One of the most popular of these models is the Mamdani Model (Mamdani, 1977). In these systems, words or sentences are used in place of numbers to describe relationships which are too complex to be described in quantitative terms. Another popular type of fuzzy logic models follow the TSK structure (Sugeno and Kang, 1988) which replaces the fuzzy sets as defined in the consequent of a Mamdani rule with a function, typically linear, of the input variables. TSK rules are much more popular in practice due to their simplicity and flexibility (Wu and Mendel, 2014). Secondly, by avoiding the defuzzification process of the Mamdani approach makes the TSK model more computationally efficient. The TSK model was successfully applied in numerous financial applications and still the subject of ongoing research (Boyacioglu and Avci, 2010; Chang et al., 2011; Chen, 2013; Kablan and Ng, 2011; Tan et al., 2011; Wei et al., 2014). It is for this reason that in this thesis we adopt the TSK approach to enhance contemporary literature by identifying innovative solutions to improve risk-adjusted trading performance.

2.2 Type-1 fuzzy sets and fuzzy membership functions

Fuzzy set theory lies at the heart of FRBSs. In classical set theory, the membership of elements in a set is a bivalent condition defined by a characteristic function which can take crisp values of 0 or 1. Boolean operators *AND*, *OR*, and *NOT* are used to perform the intersect, union and complement operations. A fuzzy set extends this idea to multi-valued logic and expresses the degree to which an element belongs to a set, hence allowing the representation of vagueness and uncertainty. If X is a collection of objects denoted generically by x, then a fuzzy set A in X is defined as a set of ordered pairs (Zadeh, 1965, 1973):

$$A = \{(x, \mu_A(x)), x \in X\}, \tag{2.1}$$

where μ_A is called a *membership function* (MF) denoting the *degree of membership* of an element in a given fuzzy set A. Hence, rather than a crisp 0 or 1, the function is allowed to return values between 0 and 1. When X is continuous, A is commonly written as

$$A = \int_X \mu_A(x)/x, \tag{2.2}$$

where the integral operator does not denote integration, but it denotes the collection of all points, $x \in X$, and the slash $(/)$ operator associates the elements in X with their membership grades using membership function $\mu_A(x)$ (Zadeh, 1965, 1973). If X is discrete, A is commonly written as

$$A = \sum_X \mu_A(x)/x, \tag{2.3}$$

where the summation sign does not denote arithmetic addition, but the collection of all points $x \in X$ with associated MF $\mu_A(x)$ (Zadeh, 1965, 1973). Corresponding to the ordinary set operations of union, intersection, and complement, fuzzy sets have similar operations. Mathematical definitions of the generalised *AND* and *OR* operators are called t-norm ($\star$) and t-conorm ($\oplus$) respectively (Zadeh, 1965, 1973).

In the following chapters, we apply the algebraic product t-norm for fuzzy intersection and the maximum t-conorm for fuzzy union. As suggested by Mendel et al. (2014), these operators are the ones which are most commonly applied in engineering applications of fuzzy logic.

There are several classes of parameterised functions commonly used to define MFs. These parameterised MFs play an important role in adaptive fuzzy inference systems. In our models proposed in the next chapters, we adopt Gaussian MFs, where each fuzzy set is represented by

$$\mu_A(x; \bar{x}, \sigma) = e^{-(x-m)^2 / 2\sigma^2}. \tag{2.4}$$

Unlike other MFs, this shape has only two parameters (the mean m and variance σ) and it always spreads out over the entire input domain (Wu and Mendel, 2014).

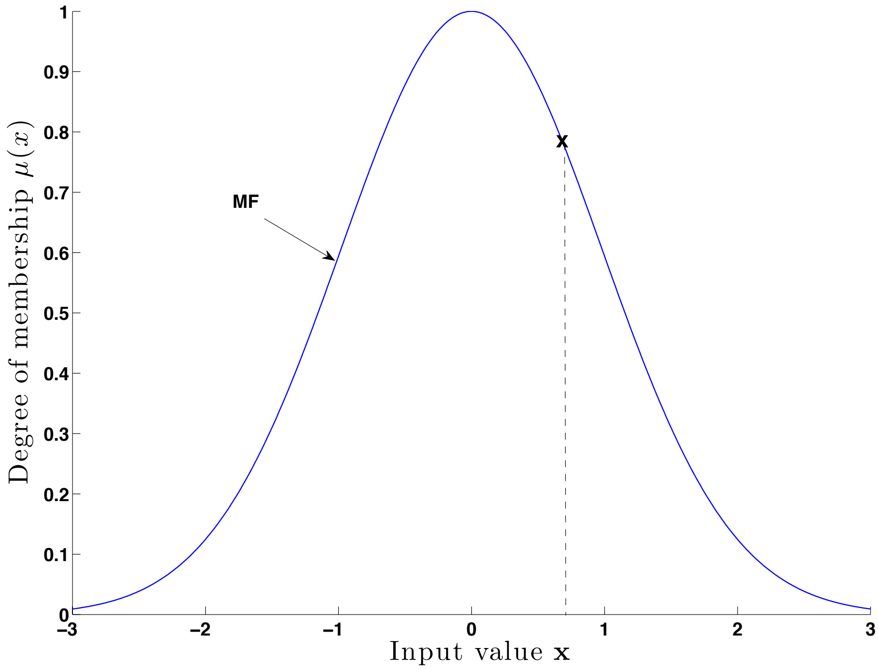


Fig. 2.2: T1 Gaussian MF

2.3 Type-1 TSK fuzzy logic systems

TSK fuzzy models are characterised by a rule base consisting of rules with fuzzy sets in the antecedents and functions, typically linear, in the consequents. A typical T1 TSK model, with k inputs and M rules, has rules in the following form (Takagi and Sugeno, 1985):

$$\begin{aligned} IF \quad & (x_1 \; is \; A_{i,1}) \; AND \; (x_2 \; is \; A_{i,2}) \; AND \; ... \; AND \; (x_k \; is \; A_{i,k}) \\ THEN \quad & y_i(\mathbf{x}) \; = \; w_{i,0} + \sum_{j=1}^{k} w_{i,j} x_j \end{aligned} \tag{2.5}$$

where $i = 1, ..., M$ is the rule number, $w_{i,v}(v = 0, 1, ..., k)$ are the consequent parameters, $y_i(\mathbf{x})$ is the output of the ith rule, and $A_{i,l}(l = 1, 2, ..., k)$ are T1 antecedent fuzzy sets. The output is computed as (Takagi and Sugeno, 1985):

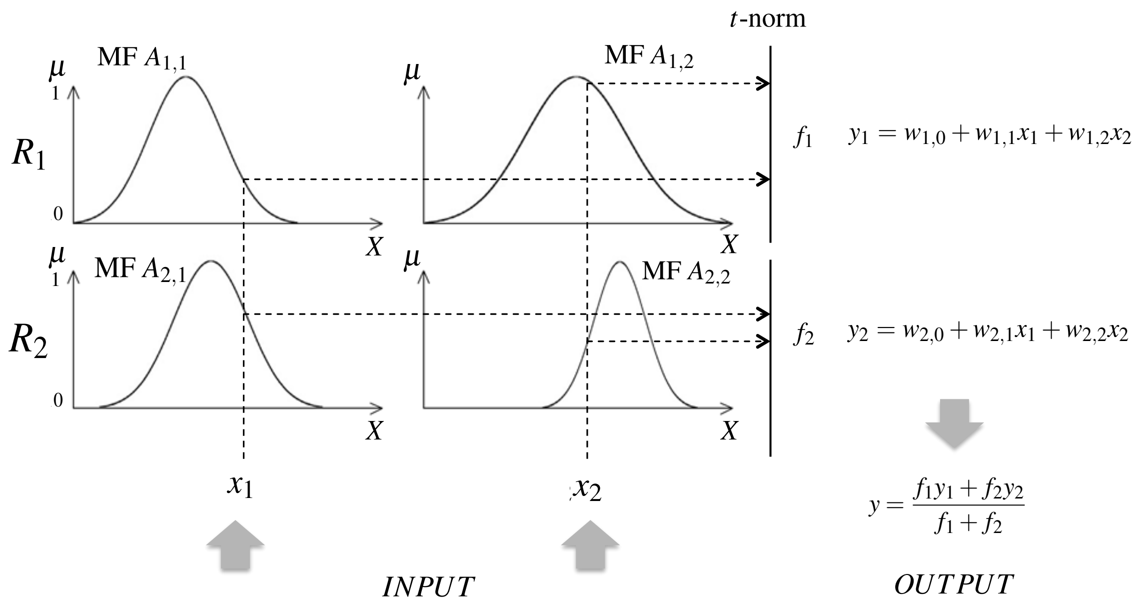


Fig. 2.3: T1 TSK fuzzy model example

$$y_{TSK,1}(\mathbf{x}) = \frac{\sum_{i=1}^{M} f_i(\mathbf{x}) y_i(\mathbf{x})}{\sum_{i=1}^{M} f_i(\mathbf{x})}, \tag{2.6}$$

which can be expanded to

$$y_{TSK,1}(\mathbf{x}) = \frac{\sum_{i=1}^{M} f_i(\mathbf{x})(w_{i,0} + w_{i,1}x_1 + w_{i,2}x_2 + ... + w_{i,k}x_k)}{\sum_{i=1}^{M} f_i(\mathbf{x})}, \tag{2.7}$$

where

$$f_i(\mathbf{x}) = \mu_{A_{i,1}} \star \mu_{A_{i,2}} \star ... \star \mu_{A_{i,k}}. \tag{2.8}$$

Figure 2.3 represents an example of a T1 TSK model consisting of 2 inputs $\{x_1, x_2\}$, 2 rules $\{R_1, R_2\}$ and 2 MFs $\{A_{i,1}, A_{i,2}\}$ describing each input space, where $i = \{1, 2\}$.

2.4 Designing fuzzy logic systems

A popular fuzzy modelling approach, originally proposed by Sugeno and Kang (1988), involves mapping input-output data. The procedure mainly consist of two parts: structure identification and parameter identification. The first step for a model designer is to define the structure of the fuzzy logic model. This consists in defining the number and position of antecedent membership functions and corresponding consequents. Ruano et al. (2002) recognized the similarity between neural networks and neuro-fuzzy systems, as they share a common structure: they can be envisaged as a two-stage model, the first performing a nonlinear mapping from an input space to an intermediate space, usually of greater dimensionality and using basis functions, and a latter stage, consisting of a linear mapping between the intermediate space and the output space.

2.4.1 Structure identification

From the literature we identify two common sources of information for building fuzzy models, prior expert knowledge and directly from data. Traditionally, fuzzy rules would have been elicited via discussions with domain experts. The disadvantage of this approach is that typically the elicitation process requires long discussions with the domain experts. Moreover, subjective or conflicting opinions from domain experts can lead to imprecise information (Buchanan and Wilkins, 1993). Gradojevic and Gençay (2013) and Cheung and Kaymak (2007) demonstrated examples of such an approach for the construction of trading algorithms. In their approach, fuzzy logic and technical analysis were combined by incorporating fuzzy rules based on expert knowledge.

For a data-driven approach, core to the rule induction process is data partitioning, each rule being a local model representation of the identified subspace. Published studies suggest numerous approaches for learning fuzzy rules from data, such as, directly from data (Wang and Mendel, 1992), decision trees (Chen et al., 2001), neural networks (Nayak, 2009) and support vector machines (Chiang and Hao, 2004). The focus of earlier data-driven methods was on achieving rule extraction automation using simple methods. However, the partitioning algorithms applied resulted in a large number of membership fuzzy sets, with the number of rules generated in many cases being proportional to the cartesian product of each membership fuzzy set. As noted by Mohammadian (1995), as the number of parameters of a system increase, the number of fuzzy rules of the system grows exponentially. Apart from the computational penalty, this also reduces the interpretability of the model.

A survey by Dutta and Angelov (2010) indicates that data clustering is one of the most popular approaches that is used to generate fuzzy rules automatically from input-output data. Advancement in clustering techniques lead to various improvements. These can be split into three categories, offline, online and evolving clustering techniques. The distinguishing factors between these categories are mainly the learning approach and the ability of the clusters to adapt (evolve) as opposed to using a fixed number of clusters. In this thesis we make extensive use of clustering methods to generate fuzzy rules. In Chapter 3 we use an evolving clustering method to generate fuzzy rules from data. In Chapter 4 we use a fuzzy clustering method to identify trading rules based on risk-return regions. This is again used in Chapter 5 to generate an initial T1 rule base which is then extended to IT2.

In order to generate fuzzy rules, fuzzy clustering can be applied in the input data space only, the output data space only or jointly together (Dutta and Angelov, 2010). Here we present an example to describe the concept of generating fuzzy rules using clustering. We consider a scenario where we have two inputs $\{x_1, x_2\}$. As indicated in Figure 2.4, Gaussian MFs can be extracted by projecting clusters identified in the input space onto the respective axes. In the above example two data clusters are identified resulting in two MFs $\{C1, C2\}$ on each input space. In the TSK model, each obtained cluster is represented by one rule. In this

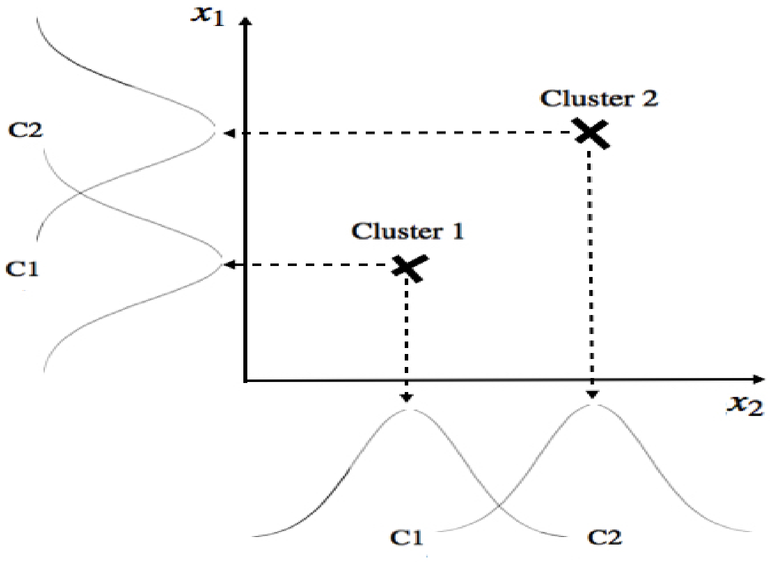


Fig. 2.4: Projection of identified clusters on the input space for MF identification

example, the generated rule base, consisting of two rules, will be defined as follows:

$$\begin{aligned} \textbf{R1: } IF\ (x_1\ is\ C1)\ AND\ (x_2\ is\ C1)\ THEN\ y_1(\mathbf{x}) &= w_{1,0} + w_{1,1}x_1 + w_{1,2}x_2 \\ \textbf{R2: } IF\ (x_1\ is\ C2)\ AND\ (x_2\ is\ C2)\ THEN\ y_1(\mathbf{x}) &= w_{2,0} + w_{2,1}x_1 + w_{2,2}x_2 \end{aligned} \tag{2.9}$$

It is evident that each rule generates a sub-model for each approximate position in the relevant input space. When a new input is presented, the rule (model) closest to the input position will be given more weight. Further parameter identification and rule-base tuning follows. This is described in the next section.

2.4.2 Parameter tuning using least squares and back-propagation

Fuzzy logic literature promotes both gradient-based and evolutionary algorithms (Wu and Mendel, 2014) for parameter optimisation. In this thesis we follow the former approach. Our decision is based on strong published evidence about the successful implementation of gradient-based optimisation in finance (Boyacioglu and Avci, 2010; Chang et al., 2011; Chen, 2013; Kablan and Ng, 2011; Tan et al., 2011; Wei et al., 2014). This also provided us with a good benchmark approach to explore possible risk-adjusted performance improvements. Two popular gradient-based optimisation approaches in machine learning literature are the method of Least Squares (LS) and the Back-Propagation (steepest descent) (BP) algorithms. Both methods can be used separately, however they are not mutually exclusive. In Chapters 3 and 4 we make extensive use of the Adaptive Neuro-Fuzzy Inference System (ANFIS) (Jang, 1993) which tunes model parameters by iteratively switching between the two algorithms. This results in a more efficient training algorithm.

Least squares method

In the LS approach, the MF shapes and parameters of the rules antecedent are fixed, however the consequent parameters are optimised using training data (Jang, 1993; Mendel et al., 2014). The approach is model-free, with the objective to completely specify the FLS using training data. The process starts from a given collection of N input-output data training pairs, $\left(\mathbf{x}^{(1)} : y^{(1)}\right), \left(\mathbf{x}^{(2)} : y^{(2)}\right), \ldots, \left(\mathbf{x}^{(N)} : y^{(N)}\right)$. Because the FLS is linear in the consequent parameters, this leads to a LS optimisation problem. In this case, the problem consists in optimising $w_{i,0}, w_{i,1}, w_{i,2}, \ldots, w_{i,k} : k+1$ consequent parameters (see Equation (2.5)) for M rules, for a total of $(k+1)M$ parameters.

By following Mendel et al. (2014), the algorithm in the LS optimisation approach starts by re-expressing Equation (2.7) as

$$\begin{aligned} y_{TSK,1}(\mathbf{x}) = g_{1,0}(\mathbf{x})w_{1,0} + g_{1,1}(\mathbf{x})w_{1,1} + g_{1,2}(\mathbf{x})w_{1,2} + \ldots + g_{1,k}(\mathbf{x})w_{1,k} + \ldots \\ + g_{M,0}(\mathbf{x})w_{M,0} + g_{M,1}(\mathbf{x})w_{M,1} + g_{M,2}(\mathbf{x})w_{M,2} + \ldots + g_{M,k}(\mathbf{x})w_{M,k}, \end{aligned} \quad (2.10)$$

where

$$g_{i,j}(\mathbf{x}) = \frac{f_i(\mathbf{x})x_j}{\sum_{i=1}^{M} f_i(\mathbf{x})}, \tag{2.11}$$

where $i = 1, ..., M, j = 0, 1, ..., k$, and $x_0 = 1$. Equation (2.10) can be expressed in vector format as

$$y_{TSK,1}(\mathbf{x}) = \mathbf{g}^T(\mathbf{x})\mathbf{w} \tag{2.12}$$

where $\mathbf{g}(\mathbf{x})$ and $\mathbf{w}$ are $(k+1)M \times 1$ vectors. If Equation (2.12) is used to represent all the N elements in the training set, we get

$$y_{TSK,1}\left(\mathbf{x}^{(t)}\right) = \mathbf{g}^T\left(\mathbf{x}^{(t)}\right)\mathbf{w} \qquad t = 1, ..., N \tag{2.13}$$

The N equations in the form of Equation (2.13) can be grouped together in a vector-matrix format as

$$\mathbf{y}_{TSK,1} = \mathbf{G}\mathbf{w} \tag{2.14}$$

where $\mathbf{G}$ is an $N \times (k+1)M$ matrix. The parameter vector $\mathbf{w}$ of the linear system of equations can be solved using the standard LS approach by minimising the cost function

$$\mathbf{J}(\mathbf{w}) = \frac{1}{2}[\mathbf{y} - \mathbf{G}\mathbf{w}]^T[\mathbf{y} - \mathbf{G}\mathbf{w}] \tag{2.15}$$

where the solution can be expressed as

$$\left[\mathbf{G}^T\mathbf{G}\right]\mathbf{w} = \mathbf{G}^T\mathbf{y}. \tag{2.16}$$

Back-propagation method

In the second approach, both the antecedent and consequent parameters are optimised (Jang, 1993; Mendel et al., 2014). Since the FLS antecedent parameters are not linear, in this approach optimisation can be solved using back-propagation. In addition to the consequent parameters, in this case the problem consists in optimising 2 parameters per antecedent Gaussian MF ($m_{A_{i,j}}$ and $\sigma_{A_{i,j}}$), k antecedents, and M rules, for a total of $2kM$ antecedent parameters. When combining both antecedent and consequent parameters, this results in a total of $3kM + M$ parameters.

As demonstrated by Mendel et al. (2014), to optimise both antecedent and consequent parameters, we use Gaussian MFs (defined in Equation (2.4)) and start from Equation (2.10) and Equation (2.11) but expand $g_{i,j}$ in more detail as

$$g_{i,j}\left(\mathbf{x}^{(t)}\right) = \frac{\exp\left(-\frac{1}{2}\sum_{c=1}^{k}\left(\frac{x_c^{(t)} - m_{A_{i,c}}}{\sigma_{A_{i,c}}}\right)^2\right) x_j^{(t)}}{\sum_{i=1}^{M}\exp\left(-\frac{1}{2}\sum_{c=1}^{k}\left(\frac{c_j^{(t)} - m_{A_{i,c}}}{\sigma_{A_{i,c}}}\right)^2\right)} \tag{2.17}$$

where $j = 0, 1, ..., k, i = 1, ..., M, t = 1, ..., N$ and $x_0 = 1$. Using our training set of N input-output data pairs, $\left(\mathbf{x}^{(1)} : y^{(1)}\right), \left(\mathbf{x}^{(2)} : y^{(2)}\right), ..., \left(\mathbf{x}^{(N)} : y^{(N)}\right)$, the objective is to minimise the sum of squares function

$$J^{(t)} = \frac{1}{2}\left(y_{TSK,1}\left(\mathbf{x}^{(t)}\right) - y^{(t)}\right)^2 = \frac{1}{2}e^2 \tag{2.18}$$

The gradient descent algorithm involves a number of iterations. Using Equation (2.10) and Equation (2.17), on each epoch the objective function J is minimised in steps that are defined by the learning rate α and the gradient of J with respect to each model parameter as follows (Mendel et al., 2014):

- Coefficients of the consequent function update:

$$w_{i,j}(t+1) = w_{i,j}(t) - \alpha \frac{\partial J^{(t)}}{\partial w_{i,j}} \tag{2.19}$$

which yields

$$w_{i,j}(t+1) = w_{i,j}(t) - \alpha e g_{i,j}\left(\mathbf{x}^{(t)}\right) \tag{2.20}$$

- Mean of the Gaussian antecedent MFs update:

$$m_{i,j}(t+1) = m_{i,j}(t) - \alpha \frac{\partial J^{(t)}}{\partial m_{A_{i,j}}} \tag{2.21}$$

which yields

$$m_{A_{i,j}}(t+1) = m_{A_{i,j}}(t) - \alpha e \left(w_{i,0} + w_{i,1}x_1 + w_{i,2}x_2 + \ldots + w_{i,k}x_k - y_{TSK,1}\left(\mathbf{x}^{(t)}\right)\right) \times \frac{\left(x_j^{(t)} - m_{A_{i,j}}(t)\right)}{\sigma^2_{A_{i,j}}(t)} g_{i,j}\left(\mathbf{x}^{(t)}\right) \tag{2.22}$$

- Variance of the Gaussian antecedent MFs update:

$$\sigma_{A_{i,j}}(t+1) = \sigma_{A_{i,j}}(t) - \alpha \frac{\partial J^{(t)}}{\partial \sigma_{A_{i,j}}} \tag{2.23}$$

which yields

$$\sigma_{A_{i,j}}(t+1) = \sigma_{A_{i,j}}(t) - \alpha e \left(w_{i,0} + w_{i,1}x_1 + w_{i,2}x_2 + \ldots + w_{i,k}x_k - y_{TSK,1}\left(\mathbf{x}^{(t)}\right)\right) \times \frac{\left(x_j^{(t)} - m_{A_{i,j}}(t)\right)^2}{\sigma^3_{A_{i,j}}(t)} g_{i,j}\left(\mathbf{x}^{(t)}\right) \tag{2.24}$$

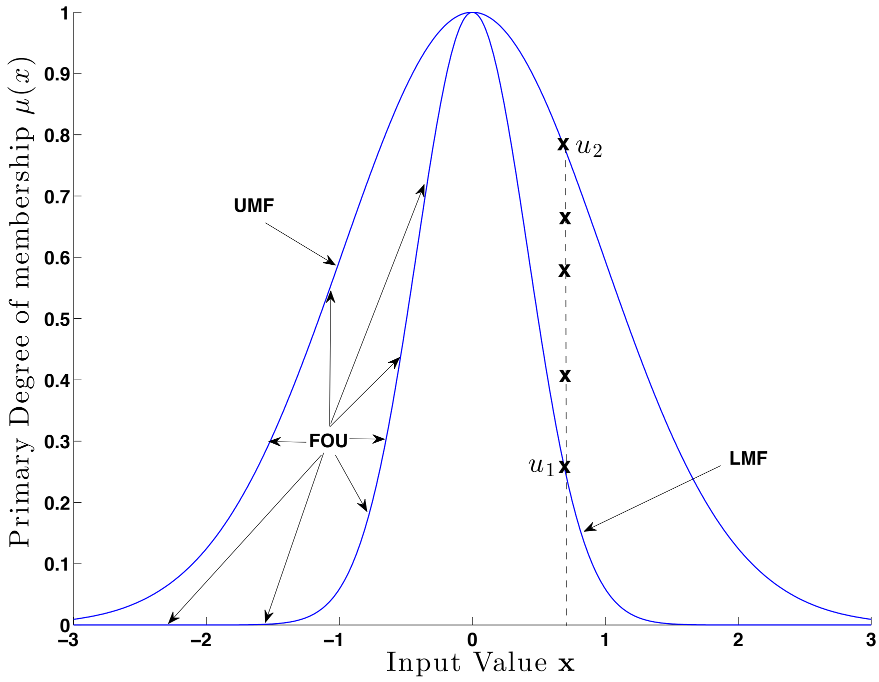


Fig. 2.5: T2 primary MF

2.5 Type-2 fuzzy sets and fuzzy membership functions

Type-2 (T2) fuzzy sets were introduced by Zadeh (1975) and are a generalisation of T1 fuzzy sets. This gives the ability to T2 fuzzy sets to handle higher levels of uncertainty. T2 fuzzy sets are characterised by their 3D structure, where the third dimension represents the membership degree of its 2D space which is defined by the footprint of uncertainty (FOU). As shown in Figure 2.5 the FOU is defined by an Upper Membership Function (UMF) and a Lower Membership Function (LMF). The FOU represents a blurring effect on top of a T1 MF, with the size of the FOU representing an additional dimension to represent the degree of uncertainty. The FOU is a key parameter in T2 FLSs, and in a later chapter we specifically focus on tuning this parameter.

When comparing the T1 MF (Figure 2.2) and the T2 MF (Figure 2.5), it can be identified that in the case of T2 MF, an input acts like a vertical slice between the UMF and LMF.

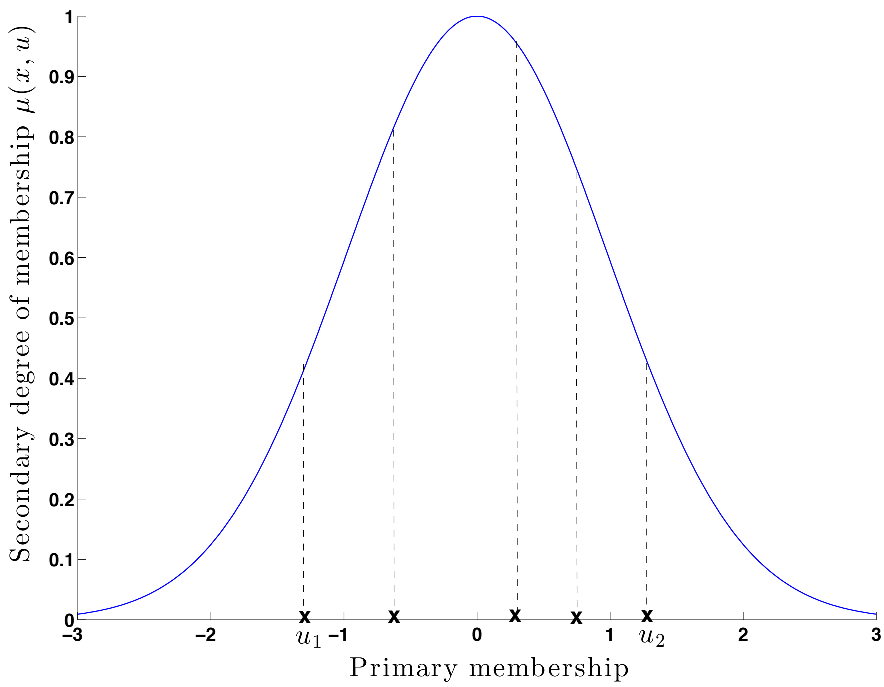


Fig. 2.6: General T2 secondary MF

Hence unlike the T1 case where the input maps to a single MF value, in the case of T2 it translates into a serious of points which we call the primary membership. For general T2 MFs, each point on the primary MF maps to a secondary membership degree which is represented by another MF (Figure 2.6).

More formally, a T2 fuzzy set (Zadeh, 1975), denoted as $\widetilde{A}$, is defined as

$$\widetilde{A} = \{((x,u), \mu_{\widetilde{A}}(x,u)) | \forall x \in X, \forall u \in J_x \subseteq [0,1]\}, \tag{2.25}$$

which shows that T2 sets depend on pairs of the two variables x and u. A T2 MF is typically represented as

$$\widetilde{A} = \int_{x \in X} \int_{u \in J_x} \mu_{\widetilde{A}}(x,u)/(x,u) \qquad J_x \subseteq [0.1] \tag{2.26}$$

where $\int\int$ denotes union over all x and u (Zadeh, 1975).

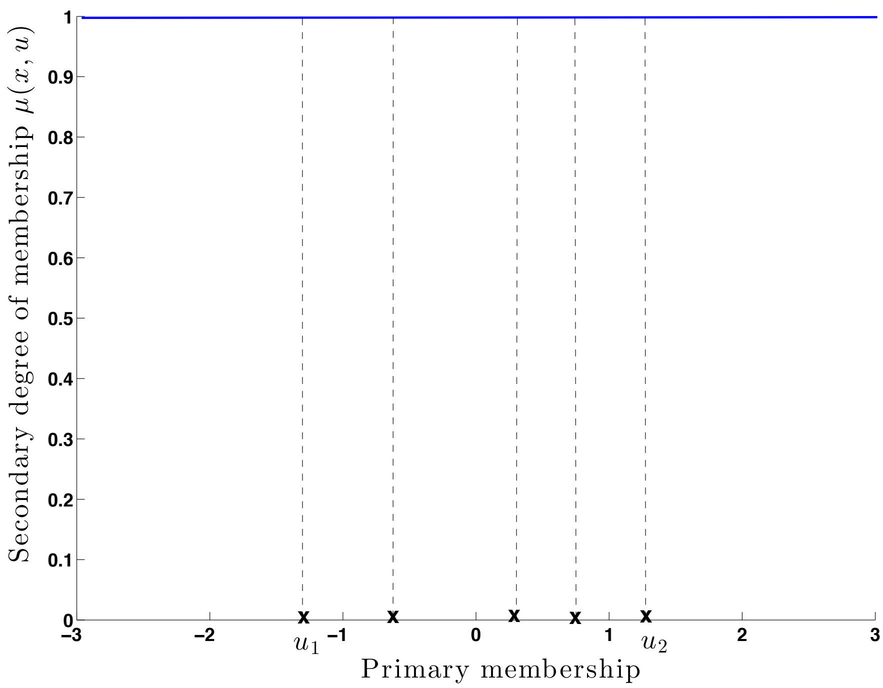


Fig. 2.7: Interval T2 secondary MF

In the special case when all $\mu_{\widetilde{A}}(x,u) = 1$ then $\widetilde{A}$ is called an interval T2 set (IT2) (see Figure 2.7). In this thesis we consider IT2 fuzzy sets since they are much less computationally intensive and more popular than the generalised T2 (Mendel et al., 2014; Wu and Mendel, 2014).

2.6 Extending type-1 TSK fuzzy logic systems to type-2

In the case of a T2 FLS (Figure 2.8), the high-level structure is very similar to a T1 FLS, except in the output processing, which consists of two components, namely the type-reducer and the defuzzifier. Although T1 rules can be extended to T2 in various forms, in this thesis we make use of rules in which the consequent consists of crisp numbers (see Mendel et al.,

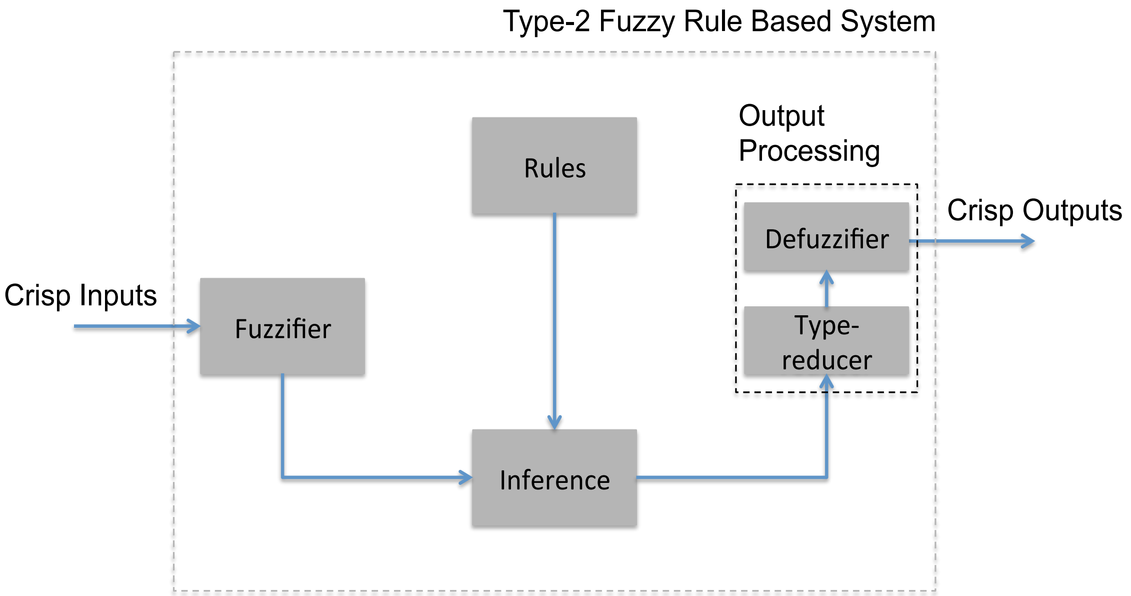


Fig. 2.8: T2 fuzzy rule-based system

2014). In this form, the rules in a T2 TSK model take the following form

$$\begin{aligned} IF \quad & \left(x_1 \; is \; \widetilde{A}_{i,1}\right) \; AND \; \left(x_2 \; is \; \widetilde{A}_{i,2}\right) \; AND \; ... \; AND \; \left(x_k \; is \; \widetilde{A}_{i,k}\right) \\ THEN \quad & y_i(\mathbf{x}) \;=\; w_{i,0} + \sum_{j=1}^{k} w_{i,j} x_j \end{aligned} \tag{2.27}$$

where $i = 1, ..., M$ is the rule number, $w_{i,v}(v = 0, 1, ..., k)$ are the consequent parameters, $y_i(\mathbf{x})$ is the output of the ith rule, and $\widetilde{A}_{i,l}(l = 1, 2, ..., k)$ are T2 antecedent fuzzy sets.

The final output of the model is as follows (Mendel et al., 2014):

$$\mathbf{Y}_{TSK,2}(\mathbf{x}) = \int_{f_1} ... \int_{f_M} \tau_{i=1}^{M} \mu_{F_i}(f_i) / \frac{\sum_{i=1}^{M} f_i y_i}{\sum_{i=1}^{M} f_i} \tag{2.28}$$

where M represents the number of rules, $f_i \in F_i$ where F_i represents the firing strength, and τ represents the t-norm. The integral sign represents the fuzzy union operation and the slash operator $(/)$ associates the elements of the rules' output and firing strength with their secondary membership grade.

The computational complexity derives from the enormous number of embedded sets that have to be individually processed in order to effect type-reduction and defuzzification. In this thesis we make use of IT2 MFs, resulting in A2-C0 models (Mendel et al., 2014). In this case, the combined firing interval of the rules and the corresponding rule consequents (Equation (2.28)) are simplified as

$$\mathbf{Y}_{A2-C0}(\mathbf{x}) = [y_l, y_r] = \int_{f_1 \in [\underline{f}_1, \overline{f}_1]} \cdots \int_{f_\alpha \in [\underline{f}_\alpha, \overline{f}_\alpha]} 1 / \frac{\sum_i f_i y_i}{\sum_i f_i} \tag{2.29}$$

where in this case the association between the elements of the rules' output and firing strength with their secondary membership grade is simplified to 1. The firing strength for each rule i, where $i = 1, 2, ..., M$, is calculated as

$$\underline{f}_i(\mathbf{x}) = \underline{\mu}_{A_{i,1}}(x_1) \star \underline{\mu}_{A_{i,2}}(x_2) \star \ldots \star \underline{\mu}_{A_{i,k}}(x_k) \tag{2.30}$$

$$\overline{f}_i(\mathbf{x}) = \overline{\mu}_{A_{i,1}}(x_1) \star \overline{\mu}_{A_{i,2}}(x_2) \star \ldots \star \overline{\mu}_{A_{i,k}}(x_k) \tag{2.31}$$

where, like in the case of T1, $\star$ represents the product t-norm. In an IT2 FLS, $\mathbf{Y}_{A2-C0}$ is an interval T1 set defined by y_l and y_r where

$$y_l = \min_{L \in [1, M-1]} \frac{\sum_{i=1}^{L} \underline{y}^i \overline{f}^i + \sum_{i=L+1}^{M} \underline{y}^i \underline{f}^i}{\sum_{i=1}^{L} \overline{f}^i + \sum_{i=L+1}^{M} \underline{f}^i} \tag{2.32}$$

$$y_r = \min_{R \in [1, M-1]} \frac{\sum_{i=1}^{R} \overline{y}^i \underline{f}^i + \sum_{i=R+1}^{M} \overline{y}^i \overline{f}^i}{\sum_{i=1}^{R} \underline{f}^i + \sum_{i=R+1}^{M} \overline{f}^i} \tag{2.33}$$

It is evident from Equation (2.32) and Equation (2.33) that computing y_l and y_r using this method requires an iterative approach. An evaluation of the most common algorithms that follow this approach can be found in Greenfield and Chiclana (2013) and Wu and Mendel (2014). Once y_l and y_r are computed, the final defuzzified output can be calculated as follows (Mendel et al., 2014):

$$y_{TSK,2}(\mathbf{x}) = \frac{y_l + y_r}{2}. \tag{2.34}$$

More recently, it was shown that closed-form approaches for calculating the output offer a good compromise between speed and complexity (Greenfield and Chiclana, 2013; Wu and Mendel, 2014). Closed-form approaches bypass the type-reduction step and the defuzzified output is calculated directly. In this thesis we make us of the Nie-Tan method (Nie and Tan, 2008). This method presents an efficient type-reduction method for interval type-2 fuzzy sets, which involves taking the mean of the lower and upper MFs of the interval set, thus creating a type-1 fuzzy set. In this case the output computation is simplified as

$$y_{TSK,2}(\mathbf{x}) = \frac{\sum_{i=1}^{M} y_i \left(\underline{f}_i + \overline{f}_i \right)}{\sum_{i=1}^{M} \left(\underline{f}_i + \overline{f}_i \right)}. \tag{2.35}$$

2.7 Conclusion

In this chapter we presented the mathematical foundations for the more sophisticated fuzzy logic techniques that we will be presenting in the following chapters. The next chapters are motivated by the essential idea behind fuzzy logic which supports a departure from the traditional quantitative techniques in favour of models which can model and minimise the effects of uncertainty. In our research, we use uncertainty as a proxy for risk and hence present innovative techniques for how to improve the risk-adjusted performance of trading algorithms.

Chapter 3

Improving the risk-adjusted performance of trading algorithms

The research problem that we investigate in this chapter addresses the challenges involved in fuzzy logic model design and selection processes when applied for trading purposes. Applying a wrong approach poses a high possibility of poor out-of-sample trading performance or, possibly worse, the risk of falling far short of the extraordinary profit expectations when applied in real trading scenarios.

This chapter presents a coherent framework for the design and tuning of neuro-fuzzy controlled trading algorithms with the aim of improving the level and stability of risk-adjusted performance. We limit the investigation of our approach to three milestone models from neuro-fuzzy systems literature, namely Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and Dynamic Evolving Neuro-Fuzzy Inference Systems (DENFIS).

As our first contribution, we propose an approach to address a typical problem which is faced when using popular moving average signals which tend to be over/under reactive to price movements. Contrary to the more common approaches in literature focusing on daily predictions, we seek model improvements in an intraday stock trading scenario using high-frequency data. This is further enhanced with a model validation methodology using heat maps to analyse favourable profitability in specific holding time and signal regions. Secondly, we investigate the effect of realistic constraints (risks) such as transaction costs and intraday trading hours in an intraday trading scenario with specific focus on neuro-fuzzy models. We demonstrate the impact of ignoring these trivial constraints and the risk of overestimated profitability. Finally, we investigate improvements resulting from combining neuro-fuzzy models with financially more relevant risk-adjusted objective functions and extend this to a deeper analysis on model stability and time-varying performance profiles. We also introduce an innovative approach which combines a set of risk-return objective functions using an ANFIS ensemble dynamic selection method, and show how it can improve the intraday trading performance of the popular standard ANFIS models.

3.1 Introduction

The profitability of technical trading rules is an incessant debate. By examining 30-minute prices, Schulmeister (2009) claimed that beyond the 1990s the profitability of technical trading rules has possibly moved to higher frequency prices as a result of faster algorithmic trading and more efficient markets. Kearns et al. (2010) presented opposing views, claiming that aggressive high-frequency trading (HFT) does not lead to the expected high excessive returns. In spite of this persisting contention in the high-frequency space, surveys (Krollner et al., 2010; Tsai and Wang, 2009) showed that the majority of former studies still focus on daily price forecasts. In a recent study, Cavalcante et al. (2016) reviewed a number of computational intelligence studies published from 2009 to 2015, which cover techniques for preprocessing and clustering of financial data, for forecasting future market movements, and for mining financial text information, among others. The study indicates that many former studies seem to disregard the fact that investors are more interested in risk-adjusted performance rather than just price predictions themselves (Choey and Weigend, 1997; Xufre Casqueiro and Rodrigues, 2006). From our research we identified that the study of fuzzy logic techniques in conjunction with risk-adjusted trading performance is scarce.

Tsai and Wang (2009), Krollner et al. (2010) and Cavalcante et al. (2016) identified that Artificial Neural Networks (ANNs) are the dominant machine learning technique in AI-based financial applications. On the downside, ANNs are regarded as black boxes that cannot describe the cause and effect. Moreover, hybrid models were again found to provide better forecasts compared to ANNs used alone or traditional time series models. Following the emergence of fuzzy logic (Zadeh, 1975), neural networks and fuzzy inference systems were brought together as general structures for approximating non-linear functions and dynamic processes. A popular cited technique in non-stationary and chaotic time series prediction is the Adaptive Neuro-Fuzzy Inference System (ANFIS) by Jang (1993). The successful application of ANFIS in trading applications by predicting stock price was demonstrated in Gradojevic (2007) and Kablan and Ng (2011) and many others. With a focus on the dynamic learning of rules from data, Kasabov and Song (2002) introduced pioneering work on evolving neuro-fuzzy systems with the introduction of the Dynamic Evolving Neuro-Fuzzy Inference System (DENFIS) and its application for time-series prediction. The proposition of evolving models is to keep systems continuously adapting, and hence evolving, and address recurring and changing patterns in the underlying environments. To our best knowledge DENFIS was not previously applied in a high-frequency setting.

Recently, Melin et al. (2012) and Lei and Wan (2012) identified that ANFIS ensembles (which we denote as eANFIS in this chapter) provided better generalisation and a reduction in the mean squared error when compared to conventional ANFIS and other chaotic time series models. Their ensemble integration approach was based on applying the average or weighted average methods across the regression predictions of all components of the ensemble (see also Soto et al. (2013) who extended this by applying a fuzzy integrator approach). Despite these claims, Faulina et al. (2012) showed that ensemble models involving more complex

ANFIS combinations do not necessarily lead to improved performance. In this chapter, we explore potential trading performance improvements in a high-frequency trading setting with the application of ANFIS ensembles and the identification of integration methods driven by risk-return objectives.

This chapter investigates a number of approaches to improve the design and tuning of neuro-fuzzy models and their performance stability for trading purposes. In our experiments we focus on a number of objectives:

1. We explore the debated profitability and an augmentation of moving average rules, particularly focusing on high-frequency data in an intraday trading scenario rather the more common day ahead predictions.

2. In contrast to common trading system designs that focus on fixed target returns, we investigate return bands in the region between 0.1% and 0.5% which act as a threshold for unprofitable small trades.

3. We evaluate the profitability of less aggressive HFT strategies, with a holding time of trading positions (PT) in the region between 10 minutes to 1 hour, in view of stated claims of unattainable high excessive returns from more aggressive HFT strategies.

4. We consider real-world intraday trading constraints like trading costs, realistic trading hours and no overnight positions, which are often ignored in existing studies dealing only with daily trading frequencies.

5. We investigate the risk-adjusted performance attained from three representative milestone models in neurocomputing, namely ANNs, ANFIS and DENFIS models, and also explore the effectiveness of the more sophisticated eANFIS architecture.

6. We analyse model stability by comparing the time series of risk-adjusted performance measures obtained using different model optimisation functions such as single risk-return functions, an innovative combination of different risk-return functions via an

ensemble, Root Mean Squared Error (RMSE), period return and models optimised without considering transaction costs.

In line with Tsang (2009), our models try to answer questions of the following form: "*Will the price go up (or down) by r% in the next t minutes?*" An important challenge in this study is the choice of moving average window length. For example, if the price over an interval is, in general, trending up, there are also several short-term downtrends in the price data. Some of them are real trend reversal points and others are just noise. The trend identifying mechanism should not be overly sensitive to short-term fluctuations, hence applying a too short moving average would result in falsely reporting a break in trend. On the other hand, choosing a too long moving average will result in late reaction to price movement.

Our first contribution is the simple, yet effective, extension of common technical trading strategies by considering a dynamic "portfolio" of moving average prediction models controlled by neuro-fuzzy systems. To our best knowledge, the profitability and risk-adjusted trading performance of technical rules and neuro-fuzzy systems in an intraday trading scenario was never explored. This is further enhanced by applying dynamic rules for return bands and trade position times. We suggest a combination of multiple moving average rules as input to the prediction models. Further investigation is carried out by applying a model validation methodology using heat maps to analyse favourable risk-return regions that identify profitability in specific holding time and signal regions. Our results demonstrate that the combination of fuzzy logic models with technical indicators outperform standard technical indicators and popular models like NNs.

Krollner et al. (2010) found that although more than 80% of the published studies in their survey stated that their proposed framework surpassed the benchmark model, most of them ignored trivial constraints that are typically incurred in the real world, possibly introducing bias in the results (see also Álvarez Díaz, 2010). This introduces a risk of overestimated profitability when realistic constraints are applied. Additionally, when training and evaluating a trading system, one frequently finds studies that have a very limited view of what constitutes successful investment decisions, defining performance on the grounds of forecast accuracy and win ratios, and often choose to minimise the forecast error of the price prediction, setting this as the objective function (Alves Portela Santos et al., 2007; de Faria et al., 2009; Enke and Thawornwong, 2005; Medeiros et al., 2006). This is also demonstrated by Cavalcante et al. (2016) who review a number of AI studies with application to finance that were published during the period 2009 and 2015. However, small forecast errors do not guarantee trading profits (Brabazon and O'Neill, 2006).

Driven by these important trading algorithm design aspects, as our second major contribution, we explore the effect of including more realistic trading constraints and how neuro-fuzzy models can be better trained to improve risk-adjusted performance. To our best knowledge this problem has not been studied in an intraday high-frequency setting before (see also comprehensive surveys by Bahrammirzaee (2010) and Cavalcante et al. (2016)). From our analysis, we demonstrate that neural network and neuro-fuzzy models selected on the basis of risk-adjusted functions, specifically the Sharpe ratio and the Sortino ratio, outperform other optimisation functions such as profitability, Root Mean Squared Error (RMSE) and models not considering transaction costs. In our experiments we identify the superior performance of ANFIS against neural network, DENFIS and a random model. We also introduce an innovative method which combines multiple risk-return functions using an ANFIS ensemble dynamic selection method, and show how it can further improve the intraday trading performance of the standard ANFIS model.

As our third contribution, we analyse the time series of risk-adjusted performance measures obtained by our neuro-fuzzy models. In contrast to other approaches in the literature which evaluate models using performance measures at an arbitrary single point in time (e.g. only at the end of the sample period), our goal is to provide a deeper understanding of the time-varying performance profile of the investigated neuro-fuzzy models. Our results

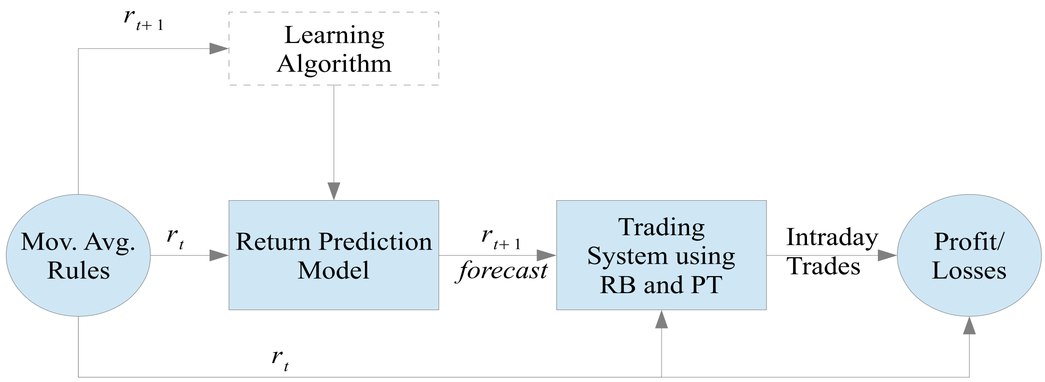


Fig. 3.1: Experiment setup

show that ANFIS coveys more stable performance when compared to the neural network model and DENFIS.

The remainder of the chapter is structured as follows. In Section 3.2 we first introduce the moving average signals and explain how these can be combined to model stock returns. We then discuss our experiment and describe model components and underlying prediction and trading algorithms. Section 3.3 presents the data, our findings and a discussion in the light of existing literature. Section 3.4 concludes.

3.2 Method

A central theme in the technical trading approach is the ability to recognise patterns in market prices that supposedly repeat themselves and hence can be used for predictive purposes. A number of authors showed the predictive capabilities of simple trading rules in conjunction with the application of ANNs. For a survey, e.g. see Vanstone and Finnie (2009) and Vanstone and Finnie (2010), and the references therein. This body of research showed the predictive ability of simple trading rules on daily returns with the application of ANNs and contrasted the weaknesses with traditional econometric models which fail to give satisfactory forecasts for some series because of their linear structure and some other inherent limitations such as the underlying distribution assumptions.

Our experiment setup consists of two core modules (see Figure 3.1). Sections 3.2.1 to 3.2.5 describe our return prediction models. Section 3.2.6 explains our trading algorithm. In Section 3.2.7 we explain how we measure and evaluate model performance.

3.2.1 Technical trading and moving averages

Traders typically employ two classes of tools to decide what stocks to buy and sell: fundamental and technical analysis, both of which aim at analysing and predicting shifts in supply and demand and hence determining the direction that prices are likely to move. While fundamental analysis involves the study of company fundamentals, such as revenues and expenses, market position, annual growth rates, and so on, technical analysis is solely concerned with price and volume data, particularly price patterns and volume spikes. Consider a set of n historical prices $\{p_t, p_{t-1}, ..., p_{t-n+1}\} \in \mathbb{R}^n_+$. Similarly to technical analysis, we aim to find a function $d : I_t \rightarrow \Omega$ that maps the information set I_t at time t to a set of trading decisions $\Omega = \{short, 0, long\}$, indicating short, neutral or long positions, respectively.

It is known from finance and economics literature that intraday prices can be very volatile over the course of the trading day (McAleer and Medeiros, 2008) due to market microstructure effects and trading behaviour (such as a change of trading sessions, lunch breaks, time zone effects, etc.). Another often reported source of (perceived) volatility is the bid-ask bounce (e.g. Roll, 1984), where the price appears to heavily fluctuate due to the random arrival of buyers and sellers. To minimise the biasing effect of this "technically" introduced noise and its propagation in our modelling approach, we apply a moving average filter to obtain a smoothed signal. This is a popular approach in technical analysis as it represents a simple yet effective means to account for the stochastic nature of the trading process. For a survey on the application of moving averages in trading, see Krollner et al. (2010) and Ahmed et al. (2010). In contrast to the common framework, however, we will augment the moving average model by jointly considering different lag structures to account for market microstructure effects arising from different intraday trading horizons.

Essentially, a moving average represents a low pass filter which removes higher frequency "noise", allowing the investor to more clearly identify the lower frequency trend. A typical moving average, MA, is calculated as:

$$MA_t^n = \frac{1}{n}\sum_{i=0}^{n-1} p_{t-i}, \tag{3.1}$$

where $i = 0, 1, 2, ..., n-1$ represents the "memory span" of the rule. One popular application of this rule is to trigger a buy signal if the price goes beyond its moving average, and a sell signal if it falls below it. Consider the signal at time t defined as

$$s_t = \begin{cases} long & \text{if} \quad p_t \geq (1+\phi)MA_t^n \\ 0 & \text{if} \quad (1-\phi)MA_t^n \leq p_t < (1+\phi)MA_t^n \\ short & \text{if} \quad p_t < (1-\phi)MA_t^n \end{cases} \tag{3.2}$$

where ϕ is the bandwidth of the rule for whiplash reduction. Another popular variation of the rule is to consider moving averages of different lengths instead, i.e. buy if the short moving average is above the long moving average, and sell if otherwise:

$$s_t^{n_1,n_2} = \begin{cases} (MA_t^{n_1} - MA_t^{n_2}) & \text{if} \quad |MA_t^{n_1} - MA_t^{n_2}| > \phi \\ 0 & \text{else} \end{cases} \tag{3.3}$$

where MA^{n_1} and MA^{n_2} are the short and long moving averages, respectively.

We investigate whether intraday high-frequency returns can be predicted by making use of buy and sell signals as inputs and use this information to build a profitable trading algorithm. For our time series we use 5-minute continuously compounded returns as they have much better statistical properties than price levels. These intraday returns are defined as:

$$y_t = \log(p_t) - \log(p_{t-1}), \tag{3.4}$$

where $\log(.)$ denotes the natural logarithm.

Recently, Gradojevic and Gençay (2013) have shown the effectiveness of combining fuzzy logic with moving average signals over conventional moving average filters due to their non-zero phase shift nature. Gencay et al. (2002) suggested that since a short moving average has a smaller phase shift than a long moving average, it would also indicate a turning point earlier than a long moving average. Although the authors argue that both filters would still indicate a turning point after the event due to their inherent nonzero phase shift nature, their combination is convenient since the trend signals from longer moving averages further confirm signals indicated by the shorter moving averages. In our experiments we propose a model that predicts the next 5-minute stock return by taking a combination of moving average rules as input variables. In our preliminary experimentation we test various moving average signals. Whilst on one hand we try to identify the correct moving average feature that helps us to reduce bias, however we limit our inputs to three variables in order to reduce model complexity and hence the possibility of overfitting (higher variance). The three moving average signals utilised are $(s^{n_1,n_2}) = \{s^{1,5}, s^{5,10}, s^{10,15}\}$, where n_1 and n_2 are in 5-minute time bars.

By combining these k input signals, the linear specification of the return y_t prediction model is defined as:

$$y_t = \theta_0 + \sum_{k=1}^{3} \theta_k s_{k,t-1} + \varepsilon_t \tag{3.5}$$

with the error term $\varepsilon_t \sim N(0,\rho)$ and

$$s_{k,t} = \begin{cases} MA_t^{1,5} & \text{for} \quad k=1 \\ MA_t^{5,10} & \text{for} \quad k=2 \\ MA_t^{10,15} & \text{for} \quad k=3 \end{cases} . \tag{3.6}$$

Kearns et al. (2010) noted that when taking transaction costs into account, aggressive HFT strategies considering holding periods between 10 milliseconds and 10 seconds can have surprisingly modest profitability. For this reason, we investigate the effect of (a) longer holding periods ranging from 10 minutes to 1 hour, and (b) the application of a return band ranging from 0.1% to 0.5%. This approach is more versatile when compared to common approaches in the literature that calibrate their trading systems based only on an arbitrary target return. Brabazon and O'Neill (2006) indicated that similar use of extended close in intraday trading scenarios can perform better than standard stop-loss, take-profit and buy-and-hold strategies.

In the following sections, we describe how the feed-forward network (FFN), ANFIS, DENFIS and eANFIS models are adapted for our experiments. The process starts from a given collection of N input-output data training pairs, $\left(\mathbf{x}^{(1)} : y^{(1)}\right), \left(\mathbf{x}^{(2)} : y^{(2)}\right), \ldots, \left(\mathbf{x}^{(N)} : y^{(N)}\right)$ where

$$\begin{aligned} &\mathbf{x}^{(1)} = [s_{1,t-N}, s_{2,t-N}, s_{3,t-N}], && y_{t-N+1}^{(1)} \\ &\mathbf{x}^{(2)} = [s_{1,t-N+1}, s_{2,t-N+1}, s_{3,t-N+1}], && y_{t-N+2}^{(2)} \\ &\ldots \\ &\mathbf{x}^{(N)} = [s_{1,t-1}, s_{2,t-1}, s_{3,t-1}], && y_t^{(N)}. \end{aligned} \tag{3.7}$$

In Equation (3.7), for each data instance at a specific time t, $\mathbf{x}$ is a vector consisting of $\{x_1, x_2, x_3\}$ input elements which represent the $\{s_{1,t-1}, s_{2,t-1}, s_{3,t-1}\}$ technical indicator signals (Equation (3.6)), and y represents the mean return over the next 5 minutes (Equation (3.4)).

3.2.2 Neural network model

Hudson et al. (1996), Gençay (1996) and Fernandez-Rodrıguez et al. (2000) showed that, under general regularity conditions, a sufficiently complex single hidden-layer feed-forward network can approximate any member of a class of functions. For a more recent survey on the applications of ANNs to model moving averages, see Vanstone and Finnie (2009) and Ahmed et al. (2010), and the references therein. Following Equation (3.5), we design a single-layer feed-forward network (FFN) regression model with 3 lagged buy and sell signals and with d hidden units. We test incremental levels of hidden units in order to identify the ideal model complexity that offers the best generalisation. Each hidden neuron, z_j, can be represented mathematically as

$$z_j = f\left(w_{1,0} + \sum_{k=1}^{3} w_{1,k} x_k\right), \tag{3.8}$$

where $w_{1,0}$ represents the bias parameter of a hidden neuron and $w_{1,k}$ represents the weight parameter between the k-th input and hidden neuron z_j. The output neuron can then be represented as

$$y_t = f\left(w_{2,0} + \sum_{j=1}^{d} w_{2,j} z_j\right), \tag{3.9}$$

where $w_{2,0}$ represents the bias parameter of the output neuron, $w_{2,j}$ represents the weight parameter between the j-th hidden neuron and output neuron, and

$$f(u) = \frac{1}{1+exp(-u)}, \tag{3.10}$$

Table 3.1: Parameters tested for FFN

Parameter	Parameter Value Set
Training Data Size (data points)	{510, 1020, 2040}
Number of Hidden Units	{5, 10, 20}
Max Training Epochs	{1,...,1000}

where $f(u)$ is the activation function in our application (see also Gençay, 1996). For our model identification, a number of model configuration parameters are considered during the in-sample training (see Table 3.1). We test and compare all $3 \times 3 \times 1 = 9$ permutations of the parameter combinations in our sensitivity analysis of FFN models.

3.2.3 ANFIS model

Neuro-fuzzy techniques synergise ANNs with fuzzy logic techniques by combining the human-like reasoning style of fuzzy systems with the learning and connectionist structure of neural networks. The model was specifically selected because it represents a neural network architecture but it also introduces type-1 fuzzy sets. This helps us to attribute more clearly any stepwise improvements originating from the better uncertainty management that is introduced using by type-1 fuzzy sets. In Chapter 5 this is extended further to type-2 fuzzy logic.

Algorithms for the acquisition or tuning of fuzzy models from data typically focus on one or all the following aspects: (i) rule consequent parameter optimisation, (ii) membership function parameter optimisation and (iii) rule induction. The tested systems will be taking input from a number of moving average rules, $\mathbf{x}$, and predict the stock return, y, as defined in Equation (3.7).

The Adaptive Neuro-Fuzzy Inference System (ANFIS) is a popular technique suggested by Jang (1993). ANFIS is a T1 TSK model, and although the rules and mathematical underpinnings follow the traditional T1 TSK models (explained in Section 2.3), the model structure is formulated in a connected layered approach to permit ANN learning techniques. Using the standard ANFIS model and Equation (3.5), we apply the ANFIS architecture layers, denoting the output of the i-th node in layer l as $O_{l,i}$, as follows:

Layer 1 Since we have 3 inputs, this layer contains $3 \times \alpha$ adaptive nodes, where α represents the number of MFs used to defined each input space. The output of each node $O_{1,i}$ is the membership grade for the input moving average signals $\{x_1, x_2, x_3\}$. For instance, the nodes with connections from the first input x_1 are in the form

$$O_{1,i} = \mu_{A_{1,k}}(x_1) \text{ for } k = 1, 2, ..., \alpha, \tag{3.11}$$

where $A_{1,k}$ represents the degree of membership of x_1 to the k-th MF. Different shapes and numbers of input MFs are tested in our model calibration process (see Table 3.2).

Layer 2 Every node in this layer is fixed and represents the firing strength of each rule. Using grid-based partitioning, this layer enumerates all possible combinations of membership functions of all inputs, in the case of our experiment resulting in α^3 nodes. Each node $O_{2,i}$ in this layer calculates the product t-norm to "AND" the membership grades,

$$O_{2,i} = f_i = \mu_{A_{i,1}}(x_1) \star \mu_{A_{i,2}}(x_2) \star \mu_{A_{i,3}}(x_3), \tag{3.12}$$

where $\mu_{A_{i,k}}$ represents the degree of membership calculation for the k-th input coming from the previous layer.

Layer 3 This layer contains fixed nodes which calculate the normalised firing strengths of the rules:

$$O_{3,i} = \hat{f}_i = \frac{f_i}{\sum_i f_i}. \tag{3.13}$$

Layer 4 The nodes in this layer are adaptive and perform the consequent of the rules:

$$O_{4,i} = \hat{f}_i y_i = \hat{f}_i \left(w_{i,0} + \sum_{k=1}^{3} w_{i,k} x_k \right) , \tag{3.14}$$

where $\hat{f}_i$ is the normalised firing strength from the previous layer and y_i is the rule consequent linear function for the i-th rule, $i = 1,...,\alpha^3$, with parameters $w_{i,0}$ and $w_{i,k}$.

Layer 5 This layer consists of a single node that computes the overall output. The nodes in this layer are adaptive and perform the consequent of the rules:

$$O_{5,i} = r_t = \sum_i \hat{f}_i y_i = \frac{\sum_i f_i y_i}{\sum_i f_i}. \tag{3.15}$$

Jang (1993) proposed premise and consequent parameters learning using a combination of gradient descent (GD) and least squares estimation (LSE) (see Section 2.4.2 for a mathematical explanation). The total parameter set is split into two sets, a Set_1 of premise (nonlinear) parameters (in Layer 1) and a Set_2 of consequent (linear) parameters (in Layer 4). ANFIS learning uses a two-pass algorithm. In a forward pass Set_1 is unmodified and Set_2 is computed using the LSE algorithm. This is followed by a backward pass where Set_2 is unmodified and Set_1 is computed with a GD algorithm (e.g. back-propagation).

Although the application of ANFIS in finance has been widely studied (Boyacioglu and Avci, 2010; Chang et al., 2011; Chen, 2013; Kablan and Ng, 2011; Tan et al., 2011;

Table 3.2: Parameters tested for ANFIS

Parameter	Parameter Value Set
Training Data Size (data points)	{510, 1020, 2040}
Input Membership Functions Shape	{Gaussian, Generalised Bell}
Number of Input Membership Functions	{2, 3}
Training Epochs	{10, 20, 40, 80}

Wei et al., 2014), most studies only employ daily data, whereas applications to intraday HFT are still scarce. Kablan and Ng (2011) successfully applied ANFIS to predict price movements from intraday tick data sampled at high frequency. Due to the intraday seasonality of volatility, they applied a volatility filter based on a directional changes threshold to filter out training data from the specific time-bins that do not exceed the specific activity threshold. Iteratively choosing the right number of epochs was also identified as an important step to avoid over-fitting. In their experiment, Kablan and Ng (2011) had the actual membership functions pre-defined and, consequently, the number of rules were fixed, hence limiting model adaptation to membership function and consequent parameter tuning. We argue that fuzzy logic models could be further improved for trading purposes by automatic updates in terms of rule base, membership function parameters and consequent parameters in view of new data. Hence, in our case we explore a number of model calibration parameters for the in-sample training (Table 3.2). We test and compare all $3 \times 2 \times 2 \times 4 = 48$ permutations of the parameter combinations in our sensitivity analysis for ANFIS models. In our in-sample training and testing we utilise the same training sample size options as those used for standard NN for consistency and better comparison. Whilst ensuring that the options cover the sample size requirements of both models, based on their specific number of model parameters, this allows the model selection process to identify the best configuration for the underlying models.

3.2.4 DENFIS model

Kasabov and Song (2002) introduced a new type of TS fuzzy inference systems, denoted as DENFIS, for adaptive on-line and off-line dynamic time series prediction. We specifically select DENFIS as our third model because it represents another milestone in the neuro-fuzzy literature representing fuzzy logic models which are more dynamic in nature due to the application of clustering algorithms to dynamically update the fuzzy rule-base (whilst in the case of ANFIS the number of fuzzy rules remains fixed). This helps us to identify whether dynamic fuzzy rule set architectures can actually lead to improved trading performance.

Research on evolving intelligent systems (refer to Kasabov and Filev, 2006) stems from the fact that data streams are often non-stationary or influenced by various conditions and modes of the corresponding environment, hence motivating models that are structurally dynamic, that is, they evolve. In relation to the models described earlier in this chapter, evolving systems can also represent an augmentation of the on-line identification of fuzzy rule-based models and on-line learning neural networks with flexible structures.

Core to DENFIS is a so-called evolving, on-line, maximum distance-based clustering method, called the Evolving Clustering Method (ECM) (for a detailed algorithm description refer to Kasabov and Song, 2002). The rule base is continuously updated and new fuzzy rules are created with new cluster identification (see Section 2.4.1 for a description of how clusters are projected on the input space to generate rules). In our study, the distance between the input examples is calculated by using the normalised Euclidean distance of a new sample to the cluster centres

$$dist_i = ||\mathbf{x}_i - \mathbf{C}_i||, \tag{3.16}$$

where $\mathbf{x}_i$ is the current example consisting of a moving average vector $\{x_1, x_2, x_3\}$ as defined in Equation (3.7) and $\mathbf{C}_i$ is the centre of the i-th cluster. A threshold value $Dthr$ that limits the cluster size is identified during our in-sample training.

The inference in DENFIS considers M fuzzy rules of the form:

$$\begin{aligned} IF \quad & (x_1 \ is \ A_{i,1}) \ AND \ (x_2 \ is \ A_{i,2}) \ AND \ (x_3 \ is \ A_{i,3}) \\ THEN \quad & y_t = f_i(\mathbf{x}), \end{aligned}$$

where $i = 1, 2, ..., M$ and $j = 1, 2, 3$; $(x_j \ is \ A_{i,j})$ are $M \times 3$ fuzzy propositions as M antecedents for M fuzzy rules respectively; and $A_{i,j}$ are fuzzy sets defined by their fuzzy membership functions $\mu_{A_{i,j}} : x_j \rightarrow [0,1]$. In the consequent part, linear functions f_i with $i = 1,2...,M$ are employed.

Using a similar principle to that employed by the clustering rule projection example in Section 2.4.1, here all fuzzy membership functions are triangular type functions with three parameters $\{a,b,c\}$ which represent the left, top, right triangle edges and the membership degree is calculated as

$$\mu(x,a,b,c) = max\left(min\left(\frac{x-a}{b-a},\frac{c-x}{c-b}\right),0\right), \tag{3.17}$$

where b is the value of the cluster centre on the x dimension, $a = b - d \times Dthr$, and $c = b + d \times Dthr$, $d \in [1.2;2]$. For a given input vector $\mathbf{x}$ the result of the inference, which is the predicted return y_t, is calculated as the weighted average of each rule's output using

$$y_t = \frac{\sum_{i=1}^{M} w_i f_i(\mathbf{x})}{\sum_{i=1}^{M} w_i}, \tag{3.18}$$

where the rule firing strength $w_i = \mu_{A_{i,1}}(x_1) \star \mu_{A_{i,2}}(x_2) \star \mu_{A_{i,3}}(x_3)$. Following a similar online learning approach to that presented in Takagi and Sugeno (1985) and Jang (1993), the linear functions in the consequent parts of the rules are updated using a recursively weighted LSE, applying also a forgetting factor. For each example, the weights are defined as the distance between the example and the corresponding cluster centre. Kasabov and Song (2002) demonstrated that DENFIS can effectively capture complex time-varying sequences in an adaptive manner and outperform some well-known existing neuro-fuzzy models (including ANFIS). To our best knowledge, DENFIS has not been applied to high-frequency trading yet. In our in-sample training we again consider a number of different model calibration parameters (see Table 3.3). For consistency and better comparison, we apply the same training sample length options in line with previous models whilst ensuring that the DENFIS

Table 3.3: Parameters tested for DENFIS

Parameter	Parameter Value Set
Training Data Size (data points)	{510, 1020, 2040}
ECM Clustering Threshold	{0.04, 0.06, 0.08, 0.1, 0.12}
Number of Rules in Dynamic FIS	{3, 5 ,6}

optimisation requirements are met with optimal generalisation. We test and compare all $3 \times 5 \times 3 = 45$ permutations of the parameter combinations in our sensitivity analysis of DENFIS models.

3.2.5 ANFIS ensemble

Melin et al. (2012), Lei and Wan (2012) and Soto et al. (2013) have recently shown the superiority of the ANFIS ensemble over singular ANFIS for chaotic time series prediction. Since their research was focused on evaluating different integration methods of the underlying ensemble regressors and on reducing statistical error, this still left an open question on whether such an architecture can be used to improve risk-adjusted performance when applied to an intraday stock trading scenario. This lead to our hypothesis that rather than optimising a single risk-adjusted objective function, better overall performance can be attained by combining multiple objective functions in an ensemble architecture.

Following these recently claimed successful results, in our experiment we adapted this architecture by combining it with the dynamic selection approach originally studied by Puuronen et al. (1999) for classification applications and later adapted for regression as in Rooney et al. (2004). The latter applied the dynamic selection method using a localised selection based on the lowest cumulative error for the nearest neighbours to the test instance. For our ensemble (eANFIS) we decided to combine the 3 ANFIS models (see Figure 3.2) that were obtained for each stock with respect to the maximisation of the Sharpe ratio, Sortino ratio and period return (details about these measures are provided in the next section).

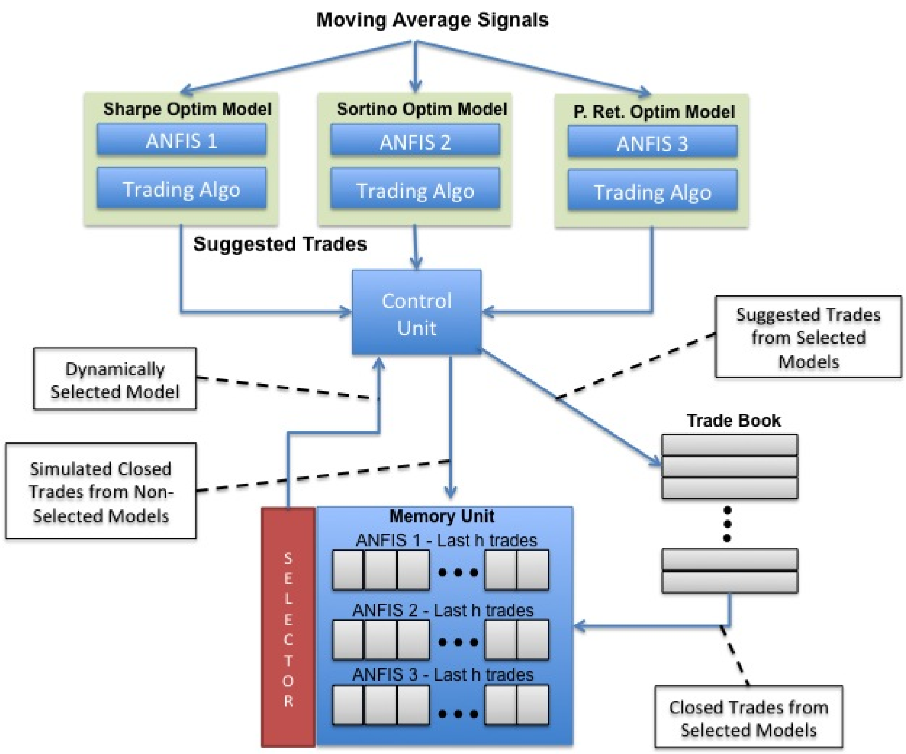


Fig. 3.2: Our proposed eANFIS architecture

Effectively, these measures represent different degrees of risk aversion where the Sharpe ratio penalises high variance on both wins and losses, the Sortino ratio penalises high variance on losses only, whilst period return does not take into consideration any risk. Hence creating an ensemble which dynamically selects between these base models permits the final ensemble model to switch to different degrees of risk on an intraday level based on the results attained from the most recent simulated trades. The Sharpe ratio is used to evaluate the underlying ensemble models. Hence if, for example, during a specific time slot a Sortino ratio-optimised ANFIS model generated higher returns for the same level of returns variance than the one optimising Sharpe ratio, it would be the selected model. At the same time the application of Sharpe ratio as the deciding measure used to shift between base models with various degrees of risk ensures that this is done in a "safe" manner by switching to riskier models when a specific time slot is identified, resulting in less abrupt movements or larger losses.

A memory component is introduced in our proposed ensemble architecture to keep track of the most recent trades for each base model. For our memory capacity, we define a new parameter, h, which is a measure defining the history length (in tick time) of closed trades that the model should store. Only newly suggested trades from the currently selected base model move to the actual trade book. However, newly suggested trades from non-selected models are still simulated based on the new arriving prices and are then passed to the memory unit. For each stock, a memory capacity value between 2 and 12 past trades was tested in the 100-day in-sample period. The value resulting in the best Sharpe Ratio performance was then applied in our out-of-sample period. Although Tsymbal et al. (2008) warn about the possible issues of using a fixed window approach, in our case this was addressed by the fact that the underlying base ANFIS models with different parameters, including the return band and position holding time parameters, suggest trades with a different rate intensity and hence time slots. Thus although the memory capacity parameter h is fixed, since this is reflected in tick time (keeping the last h trades) rather than calendar time (keeping the trades closed in a defined number of minutes), the actual time window covered by h trades used in our dynamic selection method represents different overlapping time periods.

3.2.6 Trading algorithm

The second module of our trading system (see Figure 3.1) consists of a trading and money management algorithm that takes the return prediction for the next 5 minutes from the first module and executes trades based on specific rules (see also Tan et al., 2011; Vanstone and Finnie, 2009, 2010). In Chapter 1 we explained how an HFT algorithm has to automate a number of decisions: (i) what to buy or sell (markets), (ii) how much to buy or sell (position sizing), (iii) when to buy or sell (entries), (iv) when to go out of a losing position (stops), (v) when to go out of a winning position (exits), and (vi) how to buy or sell (tactics). Our focus in this chapter is particularly on decisions (iii) to (v).

The objective of our trading algorithm is to generate *buy*, *sell* or *do-nothing* signals. For *buy* or *sell* signals the output signal has to be greater (smaller) than the upper (lower) limits of a specific return threshold (RT), otherwise the trade signal is set to *do-nothing*. This was introduced in order to filter whiplash effects when the short and long moving averages are close and also limit the number of small trades which even if profitable would result in a loss due to transaction costs. In contrast to the common approach in the literature focusing on a single target return, we consider different RTs between 0.1% and 0.5% for each stock to search for the optimal return during the in-sample period of 100 trading days. Based on the selected band size, the position taken at time t is:

$$position_t = \begin{cases} long & \text{if} \quad y_t > return\ band \\ short & \text{if} \quad y_t < return\ band \\ 0 & \text{if} \quad otherwise. \end{cases} \tag{3.19}$$

In our setup, we apply a constant transaction cost of 10 GBP per trade, per direction, and assume that a trader is willing to invest a fixed $50,000$ GBP per position. Every five minutes the trading algorithm takes a decision based on the predicted trading direction, the selected return band and the position holding time. If the signal is to go long (short), the system will buy (sell) $50,000$ GBP worth of stock at the current market price. A total of five open positions are allowed at one point in time, limiting the total investment to $250,000$ GBP. For this experiment, only positions in the same direction are allowed at the same time. This was done to specifically eliminate the hedging effect of opposing positions which, as a result, can bias the performance measure of the algorithm.

For a trade to be profitable, we defined each position to be held long enough for favourable price movement sufficient to overcome the trading costs. Different trade durations (TD) between 10 minutes to 1 hour are considered in the model selection process for each stock during the in-sample period. If after, a specific TD, the signal is still in the same direction,

Algorithm 1 Pseudo code of the trading algorithm, where RT is the predicted return threshold and TD is the trade duration.

```
Φ ← 0
if signal > RT and signal > prevsignal and balance-tradesize > 0 then
    Φ ← long
end if
if signal < −1 ∗ RT and signal < prevsignal and balance-tradesize > 0 then
    Φ ← short
end if
OPENTRADE(Φ)
for each open trade t do
    tradeDuration(t) ← tradeDuration(t)+1
    if tradeDirection(t) == Φ then
        tradeDuration(t) ← 0
    end if
    if tradeDuration(t) >= TD then
        CLOSETRADE(t)
    end if
end for
```

then the position is kept for another period of the same length. If, on the contrary, the signal has changed, then the position is closed (see Algorithm 1).

It is known that prices tend to be most volatile at market openining, due to the opening auction and the queuing and backlog of pre-opening orders in the order book. To overcome this period of volatility, our trading algorithm starts trading at 8:30. Furthermore, our trading algorithm places the last trade at 16:00, with all open positions closed at end of day, resulting in the system not holding any positions overnight. Since we are interested in active intraday trading algorithms, models configurations that do not generate trades daily over the 150-day in-sample period are excluded from the experiment.

3.2.7 Model training and evaluation

As identified in Chapter 1, although many algorithms minimise errors such as the mean squared error (MSE) or the root mean squared error (RMSE) (Alves Portela Santos et al., 2007; de Faria et al., 2009), trading systems optimised with respect to these criteria are not

Table 3.4: Models applied in the experiments

Experiment	Algorithms Tested	MA Model	Optimisation Criteria
1a	ANFIS, DENFIS, FFN	Dynamic	Sharpe Ratio
1b	ANFIS, DENFIS, FFN	Dynamic	Sortino Ratio
1c	ANFIS Ensemble (eANFIS)	Dynamic	Sharpe Ratio, Sortino Ratio, Period Return
2a	ANFIS, DENFIS, FFN	Dynamic	Sharpe Ratio, No Cost
2b	ANFIS, DENFIS, FFN	Dynamic	RMSE
2c	ANFIS, DENFIS, FFN	Dynamic	Period Return
2d	Standard Technical Rule	MA(1,5)	-
2d	Standard Technical Rule	MA(5,10)	-
2d	Standard Technical Rule	MA(10,15)	-
2e	Random Model	-	-

guaranteed to excel as the costs of prediction errors are assumed to be symmetric. Moreover, existing approaches in the literature that do not account for realistic transaction costs and trading hours possibly lead to biased results. Based on these findings, we seek to identify the effect of these conditions on the proposed neuro-fuzzy models in an intraday trading scenario. We construct a number of trading models by applying different AI algorithms and optimisation functions (see Table 3.4).

In our experiments we train the models on a daily rolling window basis (see Figure 3.3). At day_d, where $(d = 1, 2, ..., 150)$, the model is trained on 5-minute data points (Equation (5.3)), from day_{d-r} to day_{d-1}, where r represents the training sample size in days. In our training we apply random initial weights which helps to avoid any local optima. We train each model using three levels of training sample size in order to identify the right sample size which is neither too small and hence avoid underfitting (which depends on the number of model parameters), but at the same time not overly large that would dampen any shifts in the underlying data distributions. A range of model parameter combinations are tested (see Tables 3.1 to 3.3) in order to select the right model architecture complexity that offers good generalisation capabilities. The trained model is then validated by predicting, every 5

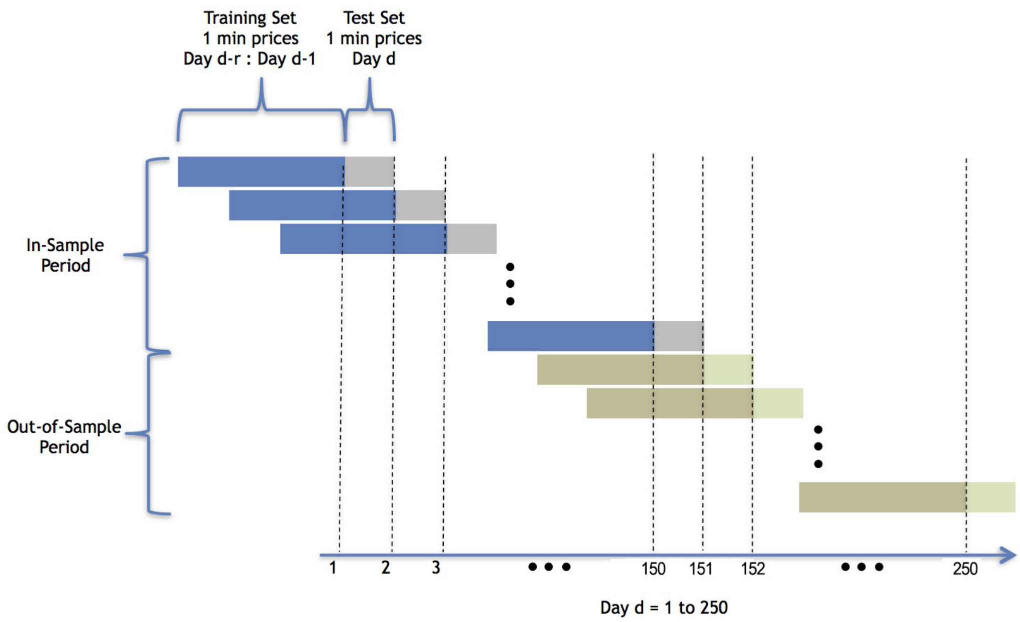


Fig. 3.3: Adopted moving window approach. The first 150 days, each day consisting of 5-minute price points, is reserved for the in-sample training and model selection process. The models are trained every day using 5-minute data points from the previous r days and then tested on the following day's prices. For out-of-sample testing, the same approach is applied and the selected model is moved forward, day-by-day, for the next 100 days.

minutes, the return over the next 5 minutes (Equation (3.4)) during day_d. Model results on validation data is stored and later used for model selection.

This process is repeated on a day-by-day basis for the whole 150-day in-sample period, for each parameter combination. The best performing model against each measure is finally tested on the following 100-day period out-of-sample on a moving window approach. In subsequent sections we describe the performance measures that we adopt in our experiments.

As part of the model calibration process and in order to ensure optimal results a number of further tests were done. One important aspect involved varying the number of epochs and step sizes in each run on the system. An epoch is defined as a single pass of the learning algorithm through the entire training dataset. On each epoch, new weights are calculated using the current learning rate. The higher the number of epochs, results in more evaluations, however, the more epochs we have, does not guarantee better results. In fact, beyond a certain number of epochs the performance will not improve. Additionally, a too large number of epochs would result in overtraining for the system, causing a decrease in performance due to overfitting. The right balance is identified by comparing the error variation in the results obtained from the training data set and the test data set and ensuring that the best calibration is obtained which minimises the error on the test data set (unseen data).

Another important consideration in the calibration process is the learning rate. If this is set too high, the algorithm can oscillate and become unstable. On the other hand, if set too small, the algorithm takes too long to converge. In all the models we configure the algorithm to allow the learning rate to change during the training process using an adaptive learning rate based on the complexity of the local error surface. This attempts to keep the learning step size as large as possible while keeping learning stable.

Furthermore, we also conduct a sensitivity analysis of the different models in order to investigate the uncertainty and stability in the model performance (see also Resta, 2009). The generation and visual inspection of heat maps was conducted in our model validation step. The objective of our sensitivity analysis was to perturb each model with small alterations in its parameters and check the possibility of any resulting large changes in the model's performance. Any such major changes could be identified as spurious results, possibly requiring changes in the model calibration, hence well-defined performance regions across our parameter space are desirable. A parameter sweep application was implemented in which the same code is run multiple times using unique sets of input parameter values (see Algorithm 2). This included varying one parameter at each step over the range of our parameter multi-dimensional space. Each individual run is executed independently of all other runs over the 150-day in-sample period applying a day-by-day moving window approach. By inspecting our 150 day-by-day trading results and analysing these across

Algorithm 2 Pseudo code for heat map data generation

```
P_1 ← set of vectors storing unique model parameter value combinations
P_2 ← set of vectors storing unique RB and PT parameter value combinations
for each parameter combination vector p_1, p_1 ∈ P_1 do
    for each parameter combination vector p_2, p_2 ∈ P_2 do
        trades ← prediction/trading algorithm (p_1,p_2)
        resultsGrid_{RB×PT} ← sharpe ratio (trades)
    end for
end for
```

the regions in the space of input factors, we can utilise a heat map approach to identify areas which maximised the Sharpe ratio criterion (for illustration, see Figure 3.4 in the next section). In particular, we are interested to see how the models behave across different levels of position time and return band parameters.

Many researchers claim positive results for their algorithmic trading models by analysing a set of performance measures at a the end of the out-of sample period. However, our interest is to validate our neuro-fuzzy models by looking at cumulated risk-return measures on a day-by-day basis. This method provides a clearer analysis of the models' performance stability over time. For this reason we take performance measurements for each model and stock at the end of each trading day.

In our first two experiments (1a and 1b) we combine the dynamic moving average model with the different AI methods and choose either the Sharpe ratio or the Sortino ratio (both defined below) as optimisation criteria. We then extend our risk-adjusted performance improvement ideas (experiment 1c) by investigating the effect of combining multiple rather than single objective function optimisation using an ensemble architecture approach. In our second part of the experiments, we apply the same models, but either (2a) do not account for transaction costs in the training period or (2b) optimise the system entirely on forecast accuracy or (2c) optimise the system to maximise return without accounting for risk, in order to see whether and how the ignorance of these constraints would have an impact on the trading performance. Next (2d), we assess the effectiveness of the dynamic moving average model, we also compare our models against the trading performance of standard fixed moving average models. This technical approach was adopted to evaluate the effectiveness of combining the same technical rules with our proposed models. Finally (2e), we benchmark our AI models against a random model in which we apply random signals however using similar trading frequency, intraday trade pattern and trade duration as the corresponding optimal Sharpe model.

To evaluate the trading system and compare the performance across different models, we apply two primary measures: Sharpe ratio and cumulative excess return. We further analyse our best models using four secondary measures: Sortino ratio, period return, profit ratio and win ratio.

The Sharpe ratio (Sharpe, 1966, 1994) is defined as

$$Sharpe\ Ratio = \frac{R - r_f}{\sigma}, \tag{3.20}$$

where R denotes the expected return, r_f the risk-free interest rate and σ the volatility of the return. Since our experiments are conducted on data from a period between year 2007 and 2008, we apply an annual risk-free rate of 5%. This falls in line with the average annual rate for UK Treasury 3-month bills for the specified period. By applying a risk-free rate we also benchmark our models against a risk-free investment.

The Sharpe ratio measures the risk premium per unit of risk in an investment. Investments with higher Sharpe ratios are often preferred due to their higher risk-adjusted performance. We note that the Sharpe ratio does not separate between variability in gains and losses, hence it attributes penalisation to both upside and downside variability. This might not represent the interest of investors who would rather welcome positive variability in gains. However, from our literature review we have identified that in previous studies (Auer and Schuhmacher, 2013; Eling, 2008; Prokop, 2012) the Sharpe ratio measures tend to indicate similar performance rankings of the more sophisticated performance ratios. Although we acknowledge that Sharpe ratio has limitations (refer to Chapter 6, Section 6.3 for more detail on these limitations), in balance, we decide in favour of the Sharpe ratio as our main risk-adjusted performance measure due to its simplicity and thus easy application, it's general tendency to indicate similar performance rankings to the more sophisticated measures, and also due to its widespread acceptance both in literature and in practice. To complement the Sharpe ratio, we also include the Sortino ratio. This is described below.

The cumulative excess return indicates the overall profitability of the strategy since the first trade (in excess of the risk-free rate) for the specified period. In all our experiments we test our strategies for a 100-day out-of-sample period.

As an additional risk-adjusted measure, we utilise the Sortino ratio (Sortino and Price, 1994; Sortino and Van Der Meer, 1991) which measures the risk premium per unit of downside risk in an investment and is defined as

$$Sortino\ Ratio = \frac{R - T}{TDD}, \tag{3.21}$$

where R denotes the expected return, T refers to the target or required rate of return for the investment strategy under considerations and TDD represents the target downside deviation which is defined as

$$TDD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Min(0, x_i - T))^2}, \tag{3.22}$$

where x_i represents the ith return, N is the total number of returns and T is the target return. In our case, we apply 5% as our target return which represents our benchmark for the minimum acceptable return. The main difference between Sharpe and Sortino ratios is that in the former standard deviation is a measure of dispersion of data around its mean, both above and below, whilst in the latter the target downside deviation is a measure of dispersion of data below some specified threshold with all above threshold returns treated as underperformance of zero. Albeit Sharpe ratio is one of the most popular risk measures used in practice, our inclusion of Sortino ratio addresses some of the weaknesses posed by Sharpe ratio especially in view of skewed return distributions. We believe that by presenting these two ratios, we provide a better ranking in terms of performance across the tested models.

The period return indicates the overall profitability of the strategy for the 100-day out-of-sample period.

The win ratio is the ratio between the number of winning trades and losing trades and is defined as

$$Win\ Ratio = \frac{Total\ Number\ of\ winning\ trades}{Total\ Number\ of\ losing\ trades}. \tag{3.23}$$

The profit ratio indicates a system's ability to generate profits over losses and is defined as

$$Profit\ Ratio = \frac{Total\ Gain/Number\ of\ winning\ trades}{Total\ Loss/Number\ of\ losing\ trades}. \tag{3.24}$$

It must be noted that although the profit and win ratios give an indication of the system's performance, it does not, however, take into consideration the underlying risk (a single loss of \$100 cannot compensate 99 winning trades of \$1). These ratios are however included in our evaluation to assess whether the models showing higher risk reflect higher returns or a higher number of wins.

3.2.8 Experiment data

The trading systems in this thesis are tested using high-frequency trade data for a set of 27 stocks listed on the London Stock Exchange (see Table 3.5) during a 250-day period between 28/06/2007 and 25/06/2008 (excluding weekends, holidays, half days and after-hours trading). The data set is composed of stocks from different industries. This helps to avoid using only stocks that follow a similar industry trend, possibly resulting in biased results.

For the experiments in this chapter, data is sampled at 5-minute intervals using the last trade price every 5-minute period. Since the London Stock Exchange operates between 8:00

Table 3.5: Descriptive statistics of 5-minute returns for stocks utilised in the experiments.

Stock	Ticker	Mean$\times 10^{-5}$	Std. Dev.	Skewness	Kurtosis
Alliance & Leicester	AL	-5.0353	0.0053	-1.4584	18.2761
Antofagasta	ANTO	0.3516	0.0042	1.2882	70.9506
Bhp Billiton	BLT	1.1087	0.0035	0.5683	23.4856
British Airways	BA	-2.7144	0.0035	-0.1592	44.5319
British Land	BLND	-2.5814	0.0031	0.2645	16.7137
Sky	SKY	-1.1155	0.0021	-0.3387	53.9515
Cable and Wireless	CWC	-1.0765	0.0024	-1.4433	146.9950
Aviva	AV	-1.4608	0.0027	0.2069	23.2902
Diageo	DGE	-5.0897	0.0022	-0.3658	47.0115
Schroders	SDR	-1.1002	0.0035	-1.3288	91.5887
Hammerson	HMSO	-2.0485	0.0033	0.4537	28.3607
Rexam	REX	-0.8891	0.0025	-1.2860	18.2506
Johnson Matthey	JMAT	0.2817	0.0024	0.0811	26.8757
HSBC Holdings	HSBA	-0.5993	0.0021	-0.0739	125.5446
Associated British Foods	ABF	-0.7105	0.0021	0.1958	9.5569
Intu Properties	INTU	-1.3083	0.0027	0.3749	43.5349
Rio Tinto	RIO	1.6696	0.0036	0.9732	25.4814
BP	BP	-0.1189	0.0021	-0.2667	45.9709
Sage Group	SGE	-0.4512	0.0024	-0.4696	17.2910
Lloyds Banking Group	LLOY	-0.2028	0.0029	0.0425	41.1450
Tesco	TSCO	-0.5911	0.0021	-1.8542	38.4018
GlaxoSmithkline	GSK	-0.5785	0.0021	-1.1249	10.9432
AstraZeneca	AZN	-0.8044	0.0022	0.8782	33.2434
HBOS	HBOS	-4.9525	0.0038	0.7131	25.2836
Xstrata	XTA	1.0409	0.0038	-1.8870	59.3569
ICAP	IAP	0.2627	0.0032	0.0814	22.0065
ITV	ITV	-3.2724	0.0028	0.2196	8.0839

and 16:30 GMT, this produced 102 price data points per day for each stock. In subsequent chapters we investigate trading algorithms at higher trading frequencies, hence we make use of 1-min trade prices. This results in 510 price data points per day for each stock.

Since in this thesis our algorithms start trading at 8:30 and avoid overnight positions, the data statistics do not consider returns that fall outside this trading range. From Table 3.5 one can identify that sample skewness and kurtosis indicate that the return distributions are far from being normal. The sample statistics also indicate that our set contains a mix of overall negative and positive trends over the selected 2007-2008 period.

Table 3.6: Average 5-minute statistics for selected stocks.

Stock	Trades	Trading Volume	Turnover (GBP)
AL	32.95	58738.67	4,012,198.45
ANTO	34.81	61809.52	4,392,559.97
BLT	109.42	210601.10	33,249,731.84
BA	40.98	176181.50	5,767,457.40
BLND	39.17	58496.55	6,106,645.82
SKY	28.89	84571.21	5,175,635.98
CWC	29.94	237598.21	4,144,406.32
AV	43.23	121746.06	8,180,508.91
DGE	44.78	115565.07	12,037,581.04
SDR	19.17	14411.58	1,688,941.14
HMSO	23.25	29159.82	3,386,577.06
REX	21.07	44697.39	2,125,137.42
JMAT	25.35	16287.67	2,887,054.58
HSBA	82.88	495076.56	42,457,796.64
ABF	22.30	27606.39	2,369,414.61
INTU	24.36	26077.22	2,731,624.84
RIO	101.83	79890.10	36,432,585.03
BP	73.81	671913.54	38,845,985.01
SGE	20.02	88894.86	1,990,645.18
LLOY	74.86	387382.39	18,402,448.33
TSCO	61.54	343251.71	14,622,519.59
GSK	64.81	209294.85	25,637,671.79
AZN	61.09	77568.86	17,828,709.65
HBOS	83.67	338586.71	22,418,314.28
XTA	87.85	79537.17	27,127,282.86
IAP	25.59	46001.83	2,608,385.95
ITV	19.44	247806.15	2,170,189.58

In Table 3.6 we provide volume statistics about the selected stocks. We defined N_t as the number of transactions per time unit. Trading volume per time interval (Q_t) is calculated as:

$$Q_t = \sum_{i=1}^{N_t} q_i, \tag{3.25}$$

where N_t denotes the number of trades between $t-1$ and t, q_i is the number of shares of trade i.

Similarly, the turnover (V_t) is calculated for a specific time interval as:

$$V_t = \sum_{i=1}^{N_t} p_i . q_i, \tag{3.26}$$

where p_i denotes the price of trade i. In the context of momentum and value strategies, these volume statistics provide an indication of liquidity - if the measure is high, this is a sign of high liquidity (Von Wyss, 2004).

3.3 Experiment results

In the following sections, results for experiment 1 are discussed in Section 3.3.1, and for (control) experiment 2 in Section 3.3.2 (see also the overview in Table 3.4).

3.3.1 Results for experiment 1

We first perform a sensitivity analysis of our models to identify the robustness of our models and also to investigate the effect of position time and return band on our results (see also Resta, 2009). For illustration, consider the heat map for ANFIS (Figure 3.4), which indicates concentrated regions of a higher Sharpe ratio in areas of higher holding position times and return bands.

This demonstrates the effectiveness of applying these two filters in our trading models. The heat map also provides an indication that, although a number of studies are mostly based on daily data (Krollner et al., 2010), technical rules do manage to identify pockets of profitability in the higher frequency range (Schulmeister, 2009). In view of the stated difficulty with aggressive high-frequency trading with position holding periods of between 10 milliseconds and 10 seconds (see Kearns et al., 2010), taking a less aggressive holding period of between 10 minutes to 1 hour can show very positive results. Of particular interest is the

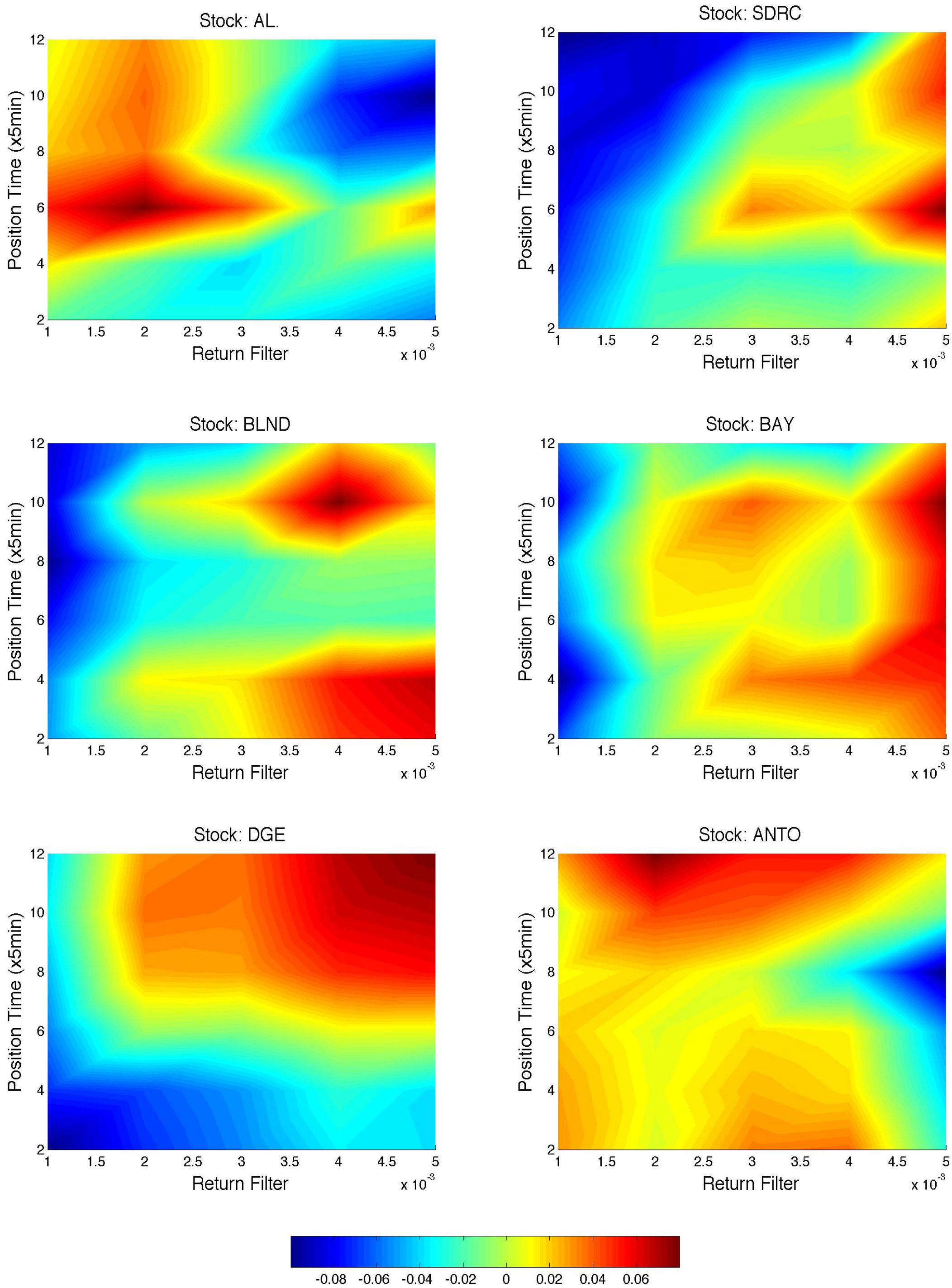


Fig. 3.4: Heat map for a set of sample stocks identifying the sensitivity of the ANFIS model and the highest Sharpe ratio for different position times and return band regions (in-sample).

fact that for specific stocks the heat maps identify more than one area of profitable regions, hence providing a clearer indication to traders on possible profitable trading strategies.

In experiment 1a, our model parameter identification was based on applying the Sharpe ratio as our objective function. From the out-of-sample results in Table 3.7, we see that LLOY was the only stock which has not generated a positive Sharpe ratio and excess return across all models. Positive results indicate that the respective models outperform the risk-free strategy that we defined as our benchmark for the period. ANFIS generated positive results in 20 out of 27 stocks (74%), followed by NN with 16 out of 27 stocks (59%), and DENFIS with 13 out of 27 stocks (48%). This demonstrates that the combination of moving average signals with neuro-fuzzy techniques can indeed be applied to generate profitable trading strategies in an intraday trading setting.

When comparing the results obtained by the different models, the results in Table 3.7 show that ANFIS generated the highest Sharpe ratio in 14 out of 27 stocks (52%), followed by NN with 7 out of 27 stocks (26%), and DENFIS with 6 out of 27 stocks (22%). In terms of profitability, ANFIS generated the highest excess return in 13 out of 27 stocks (48%), followed by DENFIS with 10 out of 27 stocks (37%), and NN with 4 out of 27 stocks (15%). These results provide a clear indication of the outperformance of ANFIS when compared to both NN and DENFIS. Another point to mention is that although DENFIS showed highest excess return in 10 stocks, however it showed losses in 14 stocks. In comparison, NN faired well and acted as a good benchmark against ANFIS.

Although this single point in time performance measurement is a very common approach adopted in literature, this comparison alone might not be sufficient to confirm the success of a model. Firstly, the Sharpe ratio for a certain sample on a certain day may be positive just by chance, in which case the reported result is not very conclusive. Secondly, even if the Sharpe ratio on a certain day or for a certain period is statistically significant and positive, different investors might not have the same holding period. The cumulative Sharpe ratio is therefore

Table 3.7: Model performance using Sharpe ratio optimisation over the 100-day out-of-sample period (bold font indicates the best result among the three AI methods for the specific stock).

	Sharpe Ratio			Excess Return		
Stock	**NN**	**ANFIS**	**DENFIS**	**NN**	**ANFIS**	**DENFIS**
AL	0.0702	**0.1259**	-0.0779	**0.0674**	0.0509	-0.2489
ANTO	0.0135	**0.0458**	-0.1845	0.0045	**0.0236**	-0.3141
BLT	**0.0855**	0.0492	0.0305	0.0310	0.0121	**0.0516**
BA	-0.0968	-0.0912	**0.0339**	-0.0777	-0.0193	**0.0207**
BLND	-0.0196	0.0313	**0.0874**	-0.0065	0.0108	**0.0222**
SKY	-0.1224	**0.1559**	-0.1636	-0.0147	**0.0252**	-0.0583
CWC	-0.0545	0.0191	**0.0235**	-0.0229	0.0107	**0.0258**
AV	0.0944	**0.1286**	-0.1324	0.0384	**0.0440**	-0.0623
DGE	**0.1770**	0.0674	0.1124	0.0460	0.0147	**0.0992**
SDR	-0.1939	**0.042**	-0.0902	-0.0519	**0.0107**	-0.0493
HMSO	**0.2571**	0.1291	0.2200	0.1116	0.0838	**0.1544**
REX	-0.0785	-0.0125	**0.0444**	-0.0179	-0.0118	**0.0519**
JMAT	-0.0618	**0.0115**	-0.2642	-0.0163	**0.0031**	-0.0953
HSBA	**0.0903**	-0.0658	-0.1840	**0.0108**	-0.0067	-0.1713
ABF	0.0001	**0.0627**	-0.0279	0.0002	**0.0271**	-0.0049
INTU	**0.1061**	-0.0988	0.0871	**0.0498**	-0.0353	0.0267
RIO	**0.0353**	-0.0888	-0.0649	**0.0084**	-0.0390	-0.0127
BP	0.0076	**0.0501**	0.0070	0.0120	**0.0188**	0.0027
SGE	-0.0539	**0.1188**	-0.1805	-0.0221	**0.0483**	-0.0174
LLOY	-0.0731	-0.0169	**-0.0088**	-0.0419	-0.0075	**-0.0043**
TSCO	0.0132	**0.0930**	-0.1854	0.0043	**0.0242**	-0.0263
GSK	0.0277	**0.0391**	-0.1994	0.0095	**0.0150**	-0.0634
AZN	-0.1738	**0.1295**	-0.1469	-0.0449	**0.0400**	-0.0131
HBOS	0.0572	**0.0725**	0.0332	0.0894	**0.1516**	0.0898
XTA	-0.1489	**0.2413**	0.0326	-0.0369	**0.0691**	0.0205
IAP	**0.1054**	-0.1529	0.0415	0.0201	-0.0582	**0.0336**
ITV	0.0510	0.1131	**0.1210**	0.0012	0.0332	**0.1040**

more informative as it indicates the overall performance of a certain strategy over a longer period. Let $measure_{t,t+n}$ represent the aggregated measure from day t to day $t+n$ of the out-of-sample period. Table 3.8 shows statistics on the distribution of $Sharpe\ Ratio_{1,20}$ up to $Sharpe\ Ratio_{1,100}$. The mean Sharpe ratio provides an indication of the average performance over the 100-day out-of-sample period, the higher the better. The standard deviation indicates

Table 3.8: Trading performance profile by optimising the Sharpe ratio. The statistics show the mean (Mean SR) and standard deviation (Std SR) of the cumulated Sharpe ratio for each stock over the 100-day out-of-sample period.

	NN		**ANFIS**		**DENFIS**	
Stock	**Mean SR**	**Std SR**	**Mean SR**	**Std SR**	**Mean SR**	**Std SR**
AL	0.0651	0.0210	**0.1084**	0.0320	-0.0427	**0.0187**
ANTO	0.0310	0.0230	**0.0553**	**0.0218**	-0.1717	0.0244
BLT	**0.0686**	0.0138	0.0248	**0.0084**	0.0479	0.0170
BA	-0.0716	0.0141	**-0.0441**	**0.0125**	-0.0461	0.0585
BLND	0.0025	0.0130	0.0640	0.0184	**0.0733**	**0.0083**
SKY	-0.1043	0.0174	**0.1137**	**0.0143**	-0.1965	0.0202
CWC	-0.0732	0.0151	**0.0002**	**0.0138**	-0.0232	0.0214
AV	0.0323	0.0171	**0.0966**	**0.0155**	-0.1284	0.0189
DGE	**0.1474**	0.0172	0.0309	0.0233	0.0940	**0.0093**
SDR	-0.1743	**0.0059**	**-0.0231**	0.0430	-0.0424	0.0234
HMSO	**0.2304**	0.0234	0.1539	0.0250	0.2021	**0.0102**
REX	0.0129	0.0610	**0.0653**	0.0413	0.0531	**0.0287**
JMAT	**-0.0522**	0.0256	-0.0623	0.0436	-0.2742	**0.0159**
HSBA	**0.0488**	**0.0103**	-0.0177	0.0302	-0.1549	0.0147
ABF	**-0.0317**	**0.0115**	-0.0986	0.0590	-0.0664	0.0228
INTU	**0.1497**	0.0191	-0.0661	0.0218	0.0214	**0.0073**
RIO	**0.0514**	0.0151	-0.0306	0.0241	-0.0740	**0.0121**
BP	-0.0288	0.0253	**0.0749**	**0.0213**	0.0211	0.0234
SGE	-0.1976	0.0200	**0.1053**	**0.0058**	-0.1996	0.0200
LLOY	-0.0560	0.0060	**-0.0191**	**0.0051**	-0.0316	0.0107
TSCO	-0.0016	0.0282	**0.0592**	**0.0117**	-0.1874	0.0124
GSK	0.0191	**0.0079**	**0.0451**	0.0111	-0.1735	0.0209
AZN	-0.1648	0.0158	**0.0936**	**0.0032**	-0.0960	0.0290
HBOS	0.0494	**0.0152**	**0.0965**	0.0206	0.0291	0.0364
XTA	-0.1527	**0.0051**	**0.2376**	0.0351	0.0159	0.0121
IAP	**0.0596**	0.0192	-0.1578	**0.0090**	0.0171	0.0160
ITV	-0.0495	0.0823	0.0709	**0.0271**	**0.1077**	0.0281

the stability of the Sharpe ratio over the same period, lower values indicating minimal fluctuations (reduced uncertainty).

From Table 3.8, we note that 20 out of 27 stocks (74%), on average, have a positive Sharpe ratio in at least one model over the 100-days out-of-sample period. This provides a stronger indication of consistent positive results across the whole out-of-sample period rather than just at the end of the period. ANFIS generates positive average Sharpe ratios in

18 stocks (67%) with highest average Sharpe ratios in 16 stocks (59%) when compared to the other models. This indicates that during the out-of-sample period, on average, ANFIS shows the highest Sharpe ratios in the majority of stocks. Moreover, ANFIS shows the highest occurrence of lower standard deviation. This indicates that in general ANFIS also shows more stable performance. NN exhibits positive average Sharpe ratio in 14 stocks (52%). This contrasts our earlier results where measurements are reported only at the end of the out-of-sample period and are positive in 16 stocks. NN scores, on average, the highest Sharpe ratios on 9 stocks (33%). However, the NN model suffers from higher standard deviations on the majority of stocks. This provides an indication that NNs suffer from high performance fluctuations. Less encouraging results were obtained in the case of DENFIS, with more than half of the stocks showing a negative average Sharpe ratio.

In experiment 1b, we base our model parameter identification process on the optimisation of the Sortino ratio. Results (Table 3.9) show that in many cases, setting the Sortino ratio as the objective function leads to similar performance rankings of the Sharpe ratio optimised models. In this case, 26 out of 27 stocks have at least one model which generates positive results over the full 100-day out-of-sample period, the only exception being RIO.

This again validates the possibility of achieving profitable trading strategies by applying a hybrid of moving average prediction models with neuro-fuzzy techniques on high-frequency data. ANFIS generates profitable trading results on 22 stocks (81%). In comparison to the other models, ANFIS obtains the highest Sortino ratio in 15 stocks (56%) and highest excess return in 15 stocks (56%). In the case of NN, performance is positive in 14 stocks (52%), with best Sortino ratio in 7 stocks (26%) and highest excess return in 8 stocks (30%). DENFIS shows positive results in only 13 stocks (48%), albeit it shows the highest Sortino ratio in 5 stocks (19%) and the highest excess return in 4 stocks (15%). From these results we conclude that in the case of the Sortino ratio optimisation, in line with the earlier results for the Sharpe ratio optimisation, ANFIS obtains the best results followed by NN.

Table 3.9: Model performance using Sortino ratio optimisation over the 100-day out-of-sample period (bold font indicates the best result among the three AI methods for the specific stock).

	Sortino Ratio			Excess Return		
Stock	**NN**	**ANFIS**	**DENFIS**	**NN**	**ANFIS**	**DENFIS**
AL	**0.1093**	0.0295	-0.0962	**0.0674**	0.0232	-0.2489
ANTO	-0.1670	**0.0655**	-0.1441	-0.0830	**0.0184**	-0.0792
BLT	0.0828	**0.0893**	0.0594	0.0627	**0.0731**	0.0993
BA	-0.2103	**0.3919**	-0.2797	-0.1608	**0.0195**	-0.2484
BLND	**0.3050**	0.2007	-0.1673	**0.0706**	0.0692	-0.0817
SKY	0.1886	**0.1922**	0.0844	**0.0180**	0.0138	0.0155
CWC	-0.3256	**0.1181**	0.0381	-0.0962	**0.0384**	0.0258
AV	**0.3197**	0.0003	-0.1650	**0.1118**	0.0001	-0.0623
DGE	-0.0952	0.2389	**0.2519**	-0.0134	**0.0391**	0.0345
SDR	-0.1694	**0.5914**	0.2458	-0.0562	**0.0604**	0.164
HMSO	**0.7296**	0.3640	0.2233	0.0761	0.0195	**0.1130**
REX	-0.094	-0.1695	**0.4979**	-0.0179	-0.0186	**0.0964**
JMAT	**0.1738**	0.0080	-0.2549	**0.0376**	0.0028	-0.0573
HSBA	0.0242	**0.1654**	-0.2234	0.0025	**0.0103**	-0.1713
ABF	0.0369	**0.3932**	0.0472	0.0126	**0.0176**	0.0049
INTU	0.1611	-0.1546	**0.2900**	**0.0659**	-0.0447	0.052
RIO	-0.2114	-0.0737	**-0.0622**	-0.0248	-0.0154	**-0.0094**
BP	0.0274	**0.2080**	0.0149	0.0050	**0.0134**	0.0030
SGE	-0.0336	**0.1519**	-0.1692	-0.0056	**0.0303**	-0.0431
LLOY	**0.2310**	-0.2224	-0.0995	**0.0652**	-0.0658	-0.027
TSCO	-0.0388	**0.1496**	-0.1156	-0.0101	**0.0242**	-0.0223
GSK	**0.4734**	0.0631	-0.2143	**0.0526**	0.0084	-0.0634
AZN	-0.1296	**0.5020**	-0.1865	-0.0105	**0.0667**	-0.0542
HBOS	-0.0600	**0.0862**	0.0527	-0.0796	**0.1644**	0.0898
XTA	-0.1518	**0.0839**	-0.0452	-0.035	**0.0129**	-0.0107
IAP	-0.0084	-0.1112	**0.0110**	-0.0024	-0.0305	**0.0070**
ITV	0.0288	**0.2202**	0.1539	0.0141	**0.0923**	0.0813

When comparing additional performance results for our best models, namely ANFIS and NN, (Tables A.1 and A.2), one can identify that in contrast to Kearns et al. (2010) our results show that the period return and win ratio for a number of stocks is considerably high. In line with our earlier results, ANFIS shows better performance than NN in terms of profitability and profit ratio. In the case of Sharpe optimisation, ANFIS shows higher period return than NN in 19 stocks (70%) and a higher profit ratio in 17 stocks (63%). The same trend is evident in the Sortino optimisation results where ANFIS achieves higher period return in 16 stocks (59%) and a higher profit ratio in 17 stocks (63%). Interestingly, NN shows more instances of higher win ratios in both Sharpe and Sortino optimisation. This demonstrates that although the win ratio is a common measure used in literature to measure performance, a higher win ratio does not necessarily result in a profitable model, hence albeit indicative, it cannot be used as a performance measure on its own (see also Brabazon and O'Neill, 2006). Our results indicate that since the profit ratio of ANFIS is in most cases higher than in NN, a higher win ratio in NN does not necessarily lead to higher profitability as a result of larger losses.

Following the results attained in the first set of experiments using the optimisation of a single risk-adjusted objective function, in experiment 1c we extend our ideas by investigating the possible risk-adjusted model performance improvement attained by combining multiple objective functions. Since from experiments 1a and 1b we identified that ANFIS performs well when compared to the other models both in terms of risk-return performance and stability, in our third experiment (1c) we decided to compare the conventional ANFIS models (1a and 1b) with our innovative ANFIS ensemble architecture, described earlier, which uses a novel dynamic selection method to integrate three risk-adjusted base ANFIS models optimised using Sharpe ratio, Sortino ratio and period return respectively. When compared against the singular ANFIS Sharpe optimised model, results (see Table 3.10) show a further improvement in Sharpe ratio in 17 stock instances (63%). A similar performance improvement is also obtained in terms of excess return.

Following the recent successful claims of reduced statistical error when applying ANFIS ensembles compared to singular ANFIS for time series prediction (Lei and Wan, 2012; Melin et al., 2012; Soto et al., 2013), our results extend these findings by showing that ANFIS ensembles composed of different risk-adjusted performance base models can outperform and show more robustness than the individual ANFIS base models that are optimised using a single risk-adjusted optimisation function. Our proposed ensemble integration method

Table 3.10: eANFIS performance using intraday dynamic switching over the 100-day out-of-sample period (bold font indicates improved performance of the combined objective function approach over the Sharpe optimised ANFIS).

	eANFIS	
Stock	**Sharpe Ratio**	**Excess Return**
AL	0.0509	0.0367
ANTO	**0.1339**	**0.1213**
BLT	**0.2282**	**0.2041**
BA	**0.0215**	**0.0145**
BLND	**0.1341**	**0.0716**
SKY	-0.0615	-0.0144
CWC	**0.2006**	**0.1117**
AV	0.0950	**0.0491**
DGE	**0.0945**	**0.0237**
SDR	**0.0902**	**0.0262**
HMSO	**0.2027**	**0.1251**
REX	**0.1053**	**0.0523**
JMAT	**0.1161**	**0.0654**
HSBA	-0.1561	-0.0159
ABF	0.0165	0.0020
INTU	**0.0250**	0.0241
RIO	-0.1553	-0.0531
BP	0.0096	0.0029
SGE	0.0023	0.0013
LLOY	**0.1749**	**0.1405**
TSCO	0.0032	0.0012
GSK	**0.0580**	**0.0184**
AZN	**0.1569**	**0.0444**
HBOS	**0.1750**	**0.3159**
XTA	0.1043	0.0353
IAP	**-0.0855**	**-0.0279**
ITV	**0.1451**	**0.1250**

introduces a dynamic short memory component in the ensemble architecture and allows for dynamic selection based on recent risk-adjusted performance at an intraday level. The results also show that following our dynamic selection approach (see also Puuronen et al., 1999) is an effective method that can be applied to further optimise ANFIS models in attaining improved profitability without incurring an additional risk penalty.

3.3.2 Results for experiment 2

We now present the results attained from benchmark models that are typically found in literature or used in practice (see overview in Table 3.4).

Surveys by Krollner et al. (2010) and Cavalcante et al. (2016) identified that many studies do not account for real-world constraints such as trading costs (see also Álvarez Díaz, 2010). In experiment 2a, we base our model selection criteria on the Sharpe ratio but exclude transaction costs in the training period. In our 100-day out-of-sample evaluation, we then apply transaction costs to the selected models as in our original Sharpe model in order to simulate realistic trading environments. As indicated in Table 3.11, all models show negative results in more than 50% of the stocks. The best model, ANFIS, shows positive results in 13 stocks (48%), whilst the worse results are shown in DENFIS with only 6 positive results (22%). The results provide a clear indication that not considering transaction costs during the training process has a negative effect on all three models with the possibility of lower than expected results when applied in real-world applications.

In experiment 2b, we apply the RMSE minimisation approach for our model selection process. This is a common optimisation approach that is applied in machine learning literature. However the results in Table 3.12 show that all three models show positive results in less than 50% of the stocks, with ANFIS and DENFIS showing positive results in only 7 and 4 stocks respectively. These results are in line with Brabazon and O'Neill (2006) and provide a clear indication that trading models based on risk-return selection criterion outperform those based on RMSE optimisation (as can be compared with both Sharpe and Sortino ratios optimisation in Tables 3.7 and 3.9).

In experiment 2c, we compare our risk-adjusted models with those optimised using period return, hence not considering risk in their objective function. When comparing the results in Table 3.13 (above) with those obtained for Sharpe ratio optimised models (Table 3.7), one can identify that in all three models the Sharpe ratio performance is less in more than 50%

Table 3.11: Model performance using the Sharpe ratio optimisation with no transaction costs consideration during training over the 100-day out-of-sample period.

	Sharpe Ratio			Excess Return		
Stock	**NN**	**ANFIS**	**DENFIS**	**NN**	**ANFIS**	**DENFIS**
AL	0.3443	-0.2159	-0.0779	0.0937	-0.2763	-0.2489
ANTO	-0.2024	-0.0888	-0.1283	-0.1603	-0.1115	-0.0792
BLT	-0.0913	0.0296	-0.2001	-0.0331	0.0281	-0.1966
BA	-0.1947	-0.0203	0.0049	-0.1608	-0.0029	0.0031
BLND	-0.0108	-0.0635	-0.1787	-0.0036	-0.0108	-0.0190
SKY	-0.0518	0.0023	-0.0647	-0.0146	0.0003	-0.0444
CWC	0.0171	-0.0679	-0.1704	0.0120	-0.0241	-0.2076
AV	0.0944	0.1260	-0.1324	0.0384	0.0354	-0.0623
DGE	0.1603	0.1399	0.0874	0.0211	0.0305	0.0187
SDR	-0.1407	0.0338	-0.1421	-0.0163	0.0083	-0.0642
HMSO	0.3203	-0.0293	0.1487	0.1124	-0.0024	0.0776
REX	-0.1077	0.0400	-0.0040	-0.0271	0.0264	-0.0058
JMAT	0.0695	-0.0021	-0.2448	0.0421	-0.0015	-0.1445
HSBA	0.0781	-0.2262	-0.2679	0.0225	-0.0288	-0.2686
ABF	0.1478	0.1089	0.2173	0.0322	0.0239	0.1194
INTU	0.1061	0.0096	0.0165	0.0498	0.0037	0.0054
RIO	0.1303	0.0959	-0.0583	0.0157	0.0234	-0.0193
BP	0.0201	-0.0208	-0.0557	0.0050	-0.0040	-0.0204
SGE	-0.0232	-0.0067	-0.1558	-0.0176	-0.0032	-0.1463
LLOY	-0.0718	-0.1545	-0.1817	-0.0220	-0.0553	-0.1075
TSCO	-0.0753	-0.1310	-0.2838	-0.0457	-0.0734	-0.0366
GSK	-0.1877	-0.1975	-0.3643	-0.0189	-0.0784	-0.1828
AZN	-0.1739	-0.1789	-0.2825	-0.0469	-0.0509	-0.0692
HBOS	-0.0692	0.0612	0.2101	-0.0211	0.1644	0.3445
XTA	-0.1493	0.0846	-0.0493	-0.0350	0.0573	-0.0447
IAP	-0.1021	0.0765	-0.1243	-0.0782	0.0403	-0.2011
ITV	-0.0546	0.1059	-0.1725	-0.0295	0.0614	-0.2555

of the models. In the case of NN, period return optimisation results in better Sharpe ratio in 12 stocks. Substantially worse performance is obtained in ANFIS and DENFIS where only 7 stocks result in better Sharpe ratio. Moreover, the excess return results show similar performance degradation when compared to Sharpe ratio optimisation. These results further reinforce our argument with respect to the effectiveness of using risk-adjusted optimisation functions as identified in experiment 2b.

Table 3.12: Model performance using RMSE optimisation over the 100-day out-of-sample period.

	Sharpe Ratio			**Excess Return**		
Stock	**NN**	**ANFIS**	**DENFIS**	**NN**	**ANFIS**	**DENFIS**
AL	0.0001	0.0304	-0.0779	0.0014	0.0561	-0.2489
ANTO	-0.1022	-0.1324	0.0086	-0.0069	-0.1617	0.0068
BLT	-0.1930	-0.1551	-0.1452	-0.0511	-0.1255	-0.0726
BA	-0.1063	-0.3169	-0.0413	-0.0866	-0.2880	-0.0310
BLND	-0.0879	-0.0154	-0.0929	-0.0139	-0.0068	-0.0084
SKY	-0.0518	-0.0778	-0.2129	-0.0149	-0.0353	-0.0316
CWC	-0.1246	-0.2187	-0.1062	-0.0228	-0.1584	-0.1450
AV	0.0215	-0.0371	-0.1954	0.0128	-0.0117	-0.1179
DGE	0.0657	0.0180	0.0090	0.0136	0.0048	0.0017
SDR	0.0763	0.0688	-0.1264	0.0127	0.0199	-0.1715
HMSO	0.1392	0.1119	-0.0031	0.0204	0.0932	-0.0035
REX	0.1125	-0.1003	-0.0870	0.0607	-0.0214	-0.0168
JMAT	-0.0917	-0.0021	-0.0943	-0.0512	-0.0015	-0.0080
HSBA	-0.3046	-0.3006	-0.1856	-0.1070	-0.0261	-0.0284
ABF	-0.5571	0.1059	-0.0001	-0.0149	0.0704	-0.0001
INTU	-0.0603	-0.1075	0.2033	-0.0091	-0.0134	0.0369
RIO	-0.2501	0.1333	0.0219	-0.0211	0.0467	0.0095
BP	0.0088	-0.0730	-0.0484	0.0024	-0.0132	-0.0078
SGE	-0.0041	-0.1323	-0.1805	-0.0022	-0.0435	-0.0174
LLOY	-0.2630	-0.0601	-0.1817	-0.1217	-0.0923	-0.1075
TSCO	-0.1814	-0.1257	-0.2838	-0.0768	-0.0539	-0.0366
GSK	0.0937	-0.1556	-0.1667	0.0415	-0.0525	-0.0419
AZN	0.0186	-0.1286	-0.2438	0.0036	-0.0501	-0.0995
HBOS	-0.0595	-0.1824	-0.0886	-0.0366	-0.0814	-0.1387
XTA	0.0848	0.1165	-0.0081	0.0271	0.0585	-0.0018
IAP	0.1350	-0.0122	-0.0933	0.0384	-0.0035	-0.0749
ITV	0.1538	-0.1035	-0.1612	0.1138	-0.1085	-0.1757

In our next benchmark comparison, experiment 2d, we investigate the application of standard moving average trading signals (e.g. Schulmeister, 2009) over the 100-day out-of-sample period. The results (see Table 3.14 above) show an evident poor performance in the majority of stocks for all moving average indicators. The applied moving average short and long lags represent those used in our dynamic moving average experiments, which however were optimised for best trade duration and return threshold during in-sample training.

Table 3.13: Model performance using period return optimisation over the 100-day out-of-sample period.

	Sharpe Ratio			Excess Return		
Stock	**NN**	**ANFIS**	**DENFIS**	**NN**	**ANFIS**	**DENFIS**
AL	0.2563	-0.1527	-0.0528	0.1014	-0.2114	-0.1673
ANTO	-0.1095	-0.0304	-0.1142	-0.1025	-0.0430	-0.1172
BLT	0.1814	0.0448	-0.1515	0.2103	0.0614	-0.1243
BA	-0.0259	0.2602	-0.0726	-0.0421	0.1586	-0.0593
BLND	0.1284	0.1870	-0.2762	0.0489	0.1261	-0.1206
SKY	0.0001	-0.0979	-0.0863	0.0001	-0.0317	-0.0357
CWC	-0.0978	-0.0019	-0.1062	-0.0289	-0.0098	-0.1450
AV	0.1732	0.0483	-0.2942	0.1118	0.0283	-0.3507
DGE	-0.0327	0.1745	0.0503	-0.0139	0.0518	0.0110
SDR	-0.1170	0.0443	-0.1129	-0.0577	0.0147	-0.1598
HMSO	0.2052	0.0252	-0.0071	0.1603	0.0262	-0.0093
REX	-0.0443	0.1005	0.0620	-0.0406	0.0648	0.0639
JMAT	0.0729	-0.0300	-0.2642	0.0507	-0.0220	-0.0953
HSBA	-0.1184	-0.2262	-0.1873	-0.0956	-0.0288	-0.1514
ABF	0.0001	-0.0113	0.1030	0.0001	-0.0030	0.0364
INTU	-0.0876	0.0096	-0.1076	-0.0411	0.0037	-0.0614
RIO	-0.0213	-0.0675	0.0350	-0.0132	-0.0151	0.0356
BP	0.0015	0.0040	-0.1018	0.0002	0.0012	-0.1014
SGE	-0.0733	-0.2313	-0.1805	-0.0835	-0.2290	-0.0174
LLOY	-0.0462	-0.1994	0.0001	-0.0008	-0.3843	0.0001
TSCO	-0.0145	-0.2191	-0.1854	-0.0052	-0.1320	-0.0263
GSK	-0.2035	-0.1560	-0.1994	-0.0408	-0.0268	-0.0634
AZN	-0.1739	0.0134	-0.1469	-0.0469	0.0044	-0.0131
HBOS	0.0382	-0.0542	-0.1644	0.0149	-0.1189	-0.3757
XTA	0.1297	0.1170	-0.1285	0.0989	0.0662	-0.0849
IAP	0.0457	-0.1529	-0.3119	0.0125	-0.0582	-0.5678
ITV	-0.1314	-0.1405	0.1026	-0.0381	-0.1941	0.0813

This brings to light the issues with moving average standard technical rules in trend identification due to their nonzero phase shift nature (as identified by Gencay et al., 2002) and provides evidence of the effectiveness of our combined dynamic moving average with neuro-fuzzy systems approach.

In our final benchmark, experiment 2e, we deploy a random model which generates random trades (buy or sell) however keeping a similar intraday trading pattern, trade frequency

Table 3.14: Model performance using moving average (MA) rules over the 100-day out-of-sample period.

	Sharpe Ratio			Excess Return		
Stock	**MA(1,5)**	**MA(5,10)**	**MA(10,20)**	**MA(1,5)**	**MA(5,10)**	**MA(10,20)**
AL	-0.0451	-0.0148	-0.0793	-0.1836	-0.0558	-0.3082
ANTO	-0.1141	-0.0911	-0.0909	-0.3022	-0.2384	-0.2565
BLT	-0.1390	-0.1521	-0.2110	-0.2832	-0.2842	-0.4478
BA	-0.0083	-0.0425	0.0481	-0.0237	-0.1198	0.1372
BLND	-0.0252	-0.0931	-0.1087	-0.0473	-0.1876	-0.2228
SKY	-0.1833	-0.0780	-0.1386	-0.2628	-0.0931	-0.2006
CWC	-0.1078	-0.1926	-0.0782	-0.1987	-0.3135	-0.1147
AV	-0.0525	-0.1729	-0.0359	-0.0972	-0.3171	-0.0693
DGE	-0.3461	-0.3091	-0.4111	-0.3948	-0.3135	-0.4441
SDR	0.0422	-0.0565	-0.0945	0.0783	-0.1281	-0.1895
HMSO	-0.1049	-0.2245	-0.5187	-0.1933	-0.3697	-0.9034
REX	-0.0802	-0.1175	-0.2504	-0.1373	-0.1646	-0.4465
JMAT	-0.0988	-0.0297	-0.1204	-0.1442	-0.0409	-0.1738
HSBA	-0.1884	-0.1126	-0.1332	-0.2725	-0.1751	-0.1902
ABF	-0.1742	-0.2944	-0.2386	-0.2125	-0.3375	-0.2489
INTU	-0.0092	-0.0227	-0.2534	-0.0156	-0.0379	-0.4274
RIO	-0.1875	-0.0314	-0.1774	-0.3642	-0.0578	-0.3249
BP	-0.3647	-0.2340	-0.2489	-0.4704	0.2762	-0.3283
SGE	-0.1101	-0.1448	-0.1995	-0.1781	-0.2626	-0.3545
LLOY	-0.0609	0.0637	-0.1486	-0.1396	0.1427	-0.3142
TSCO	-0.1174	-0.1486	-0.1625	-0.1685	-0.2216	-0.2269
GSK	-0.2717	-0.2690	-0.3068	-0.4654	-0.3780	-0.4630
AZN	-0.283	-0.3338	-0.0332	-0.4793	-0.4679	-0.0516
HBOS	0.1126	0.0162	-0.1474	0.4233	0.0588	-0.6037
XTA	-0.2722	-0.1513	-0.2948	-0.5397	-0.3058	-0.5713
IAP	0.0322	0.0327	0.0393	0.0793	0.0816	0.0936
ITV	0.1034	0.0947	-0.0540	0.2220	0.1980	-0.1133

and trade duration of the corresponding AI driven model for the respective stock. We decide to adopt the trading pattern of the ANFIS model since in earlier experiments it demonstrates the best performance when compared to the other models.

To identify the trading pattern we captured the average number of trades that our ANFIS model fired every hour (see Figure 3.5). It has to be noted that the first hour has the lowest number of trades, however since our algorithm starts to trade at 8:30am rather than 8:00am

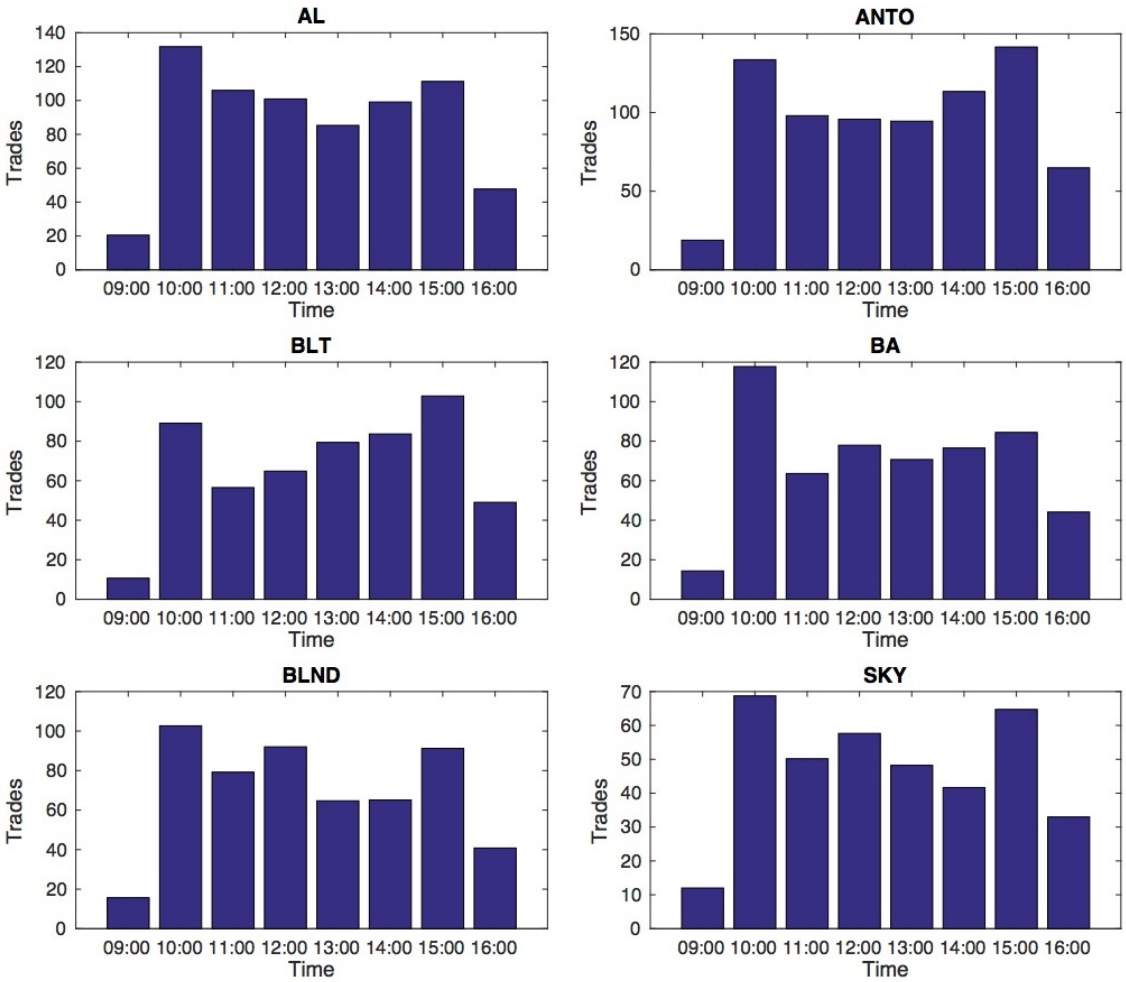


Fig. 3.5: Trading pattern of our trading algorithm on a set of stocks using ANFIS and Sharpe ratio optimisation. Each bar represents the average number of trades performed in the hour leading to the respective time slot (x-axis).

this represents only a 30 min duration. Moreover our algorithm generates new trades up to 16:00. We note that in general all stocks (as can be seen in the selected stocks in Figure 3.5) show very similar trade patterns, with the highest number of trades being placed in the time slots 9:00am to 10:00am and 14:00 to 15:00. In our random model we adopt the same hourly trade ratios for the specific stocks.

Our algorithm generates a series of uniformly distributed random numbers between -1 and 1. We apply an equal threshold on both the negative and positive side of the specified

Table 3.15: Results obtained from the random model following 100-day out-of-sample period.

	Random Model	
Stock	**Sharpe Ratio**	**Excess Return**
AL	-0.1810	-0.1220
ANTO	-0.0567	-0.0228
BLT	-0.1042	-0.0171
BA	-0.1604	-0.0670
BLND	-0.1022	-0.0226
SKY	-0.3726	-0.0389
CWC	-0.1767	-0.0353
AV	-0.1118	-0.0152
DGE	-0.2349	-0.0439
SDR	-0.3330	-0.0256
HMSO	-0.1857	-0.0387
REX	-0.2797	-0.0413
JMAT	-0.2252	-0.0445
HSBA	-0.3286	-0.0397
ABF	-0.4844	-0.0323
INTU	-0.0181	-0.0052
RIO	-0.4613	-0.0416
BP	-0.4277	-0.0538
SGE	-0.5140	-0.0574
LLOY	-0.2727	-0.0666
TSCO	-0.3719	-0.0704
GSK	-0.4996	-0.0399
AZN	-0.3608	-0.0476
HBOS	-0.1656	-0.0579
XTA	-0.3554	-0.0663
IAP	-0.0271	-0.0110
ITV	-0.3752	-0.0361

range. If the random number is smaller or greater than the threshold then a sell or buy signal is placed accordingly. By increasing (decreasing) the threshold we decrease (increase) the random trade instances in that specific time slot. From the results (Table 3.15) we identify that the random model does not produce good results on any of the stocks (considering that we are applying an annual risk-free rate of 5%). This indicates that our applied AI-models are indeed effective in generating valid trade signals.

3.4 Conclusion

In this chapter, our research problem stems from common model design and tuning recommendations in AI field but which do not fall in line with the risk concerns of investors when applied for trading purposes. We limit our investigation to three representative milestone models from neurocomputing literature, namely ANN, ANFIS and DENFIS. We corroborate our arguments by showing that the omission of risk-adjusted performance considerations leads to poor out-of-sample performance or, even worse, infers overestimated trading performance expectations.

In our experiments, we focus on a particular high-frequency trading (HFT) window that was chosen to gain more insight into the profitability of intraday trading with respect to the tension created between two literature findings: (i) the view that the profitability of trading rules has possibly moved to higher frequency prices (Schulmeister, 2009), and on the other hand, (ii) the view that aggressive HFT with position holding periods of between 10 milliseconds and 10 seconds does not reap the expected excess returns (Kearns et al., 2010).

Our first contribution is the simple yet effective extension of common technical moving average rules by considering a dynamic "portfolio" of moving average prediction models in conjunction with fuzzy logic models. To our knowledge, the investigation of fuzzy logic prediction models using moving average signals at short intraday time frames, in conjunction with risk-adjusted performance analysis, was never investigated before. A common challenge in selecting a trend identifying mechanism is that on the one hand, applying a too short moving average would result in falsely reporting a break in trend. On the other hand, choosing a too long moving average would result in a late reaction to price movement. In our approach, we use fuzzy logic techniques to combine different moving averages and dynamically tune the trend signals according to the changing speeds of the market. This is further enhanced by applying a model validation methodology using heat maps to analyse favourable risk-return regions that identify profitability in specific holding time and signal regions. Our results demonstrate that the proposed dynamic moving approach outperforms the risk-adjusted performance obtained from standard moving average technical rules and popular NN models.

Cavalcante et al. (2016) identify that a number of published studies focusing on the application of advanced AI techniques to control trading algorithms disregard the investor's risk-return trade-off. The same authors find that although the surveyed papers state that their proposed framework surpassed the benchmark model, most of them are actually ignoring trivial constraints in the real world, hence introducing a risk of overestimated profitability (see also Álvarez Díaz, 2010). From our research, we find that an investigation of these considerations, from the perspective of neuro-fuzzy models, are scarce.

As a second contribution, we present a thorough examination of the effect of realistic trading constraints and suggest better ways of how neuro-fuzzy models can be trained to improve risk-adjusted trading performance. When training and evaluating a trading system, many former studies have a very limited view of what constitutes successful investment decisions, defining success on the grounds of forecast accuracy, win ratios or minimum forecast error (Alves Portela Santos et al., 2007; de Faria et al., 2009; Enke and Thawornwong, 2005; Medeiros et al., 2006). While a number of studies investigate these aspects for other financial applications (e.g. Dempster and Leemans, 2006; Esfahanipour and Mousavi, 2011; Evans et al., 2013), the implementation of risk-adjusted performance control in the remit of neuro-fuzzy computing has to our best knowledge not been studied in an intraday high-frequency setting before. Our findings show that omitting transaction costs from the model selection process fall far short of the expected trading performance when tested out-of-sample with realistic constraints. From our experiments we conclude that neuro-fuzzy models using the Sharpe ratio and the Sortino ratio optimisation convey better trading performance than those using the more traditional objective functions in AI literature. Our results indicate the superior risk-adjusted performance of ANFIS when compared to standard neural network, DENFIS and standard moving average technical rules. We further demonstrate the effectiveness of the ANFIS model against a random model. Additionally, we propose a novel method of combining different risk-adjusted objective functions using an innovative ANFIS ensemble architecture (eANFIS). From the results we show that our proposed ensemble model outperforms conventional ANFIS using single risk-adjusted objective functions.

As our third contribution, we analyse in more detail the trading performance of the proposed neuro-fuzzy models by examining the time series of risk-adjusted performance measures. In contrast to a common approach recommended in the AI literature which evaluates models at the end of the out-of-sample sample period, we provide a deeper understanding of the time-varying performance profile of the applied models during the whole out-of-sample period. We address this by investigating the cumulative risk-adjusted performance on a daily basis. Our results show that the proposed ANFIS model exhibits the most stable performance when compared to standard NN and DENFIS models.

In the next chapter we extend our research problem to address the time-varying intraday volatility. Our ideas on the utilisation of heat maps are further extended to adapt models at more granular intraday time-windows.

Chapter 4

Time-varying intraday volatility and risk-adjusted performance

In Chapter 3, it was demonstrated that applying a global risk-adjusted objective function for trading algorithms controlled by neural networks and neuro-fuzzy is more appropriate than other more common objective functions found in conventional machine learning literature. By extending our ideas to a more granular intraday perspective, the research in this chapter emanates from the stylised fact that activity in markets is not constant throughout the trading day. Changes in the underlying conditions can introduce a substantial increase in error variance and large losses. The research problem addresses risk-adjusted performance maximisation of trading algorithms controlled by fuzzy logic systems in view of the challenge posed by continuously changing *intraday* market conditions as a reflection of time-varying volatility (risk).

A wide body of finance research claims that volatility exhibits strong intraday periodicity. Whilst sufficient market volatility is required to ensure price movement, on the other hand, higher volatility reflects higher uncertainty, more challenging forecasting, and the possibility of incurring larger losses. Although numerous models are presented in literature that are capable of measuring volatility at a good degree of accuracy up to an intraday level, the literature of how AI trading algorithms can be optimised to handle intraday time-varying volatility is scarce. In our proposed models, we utilise fuzzy logic methods to combine popular models like neural networks with realised volatility and, as a result, create hybrid models which can adapt to intraday time-varying volatility. The results indicate the effectiveness of our proposed models.

4.1 Introduction

Bearing in mind that essentially a model is a representation of a state, this chapter builds on the risk-adjusted model improvements presented in Chapter 3 by introducing a method to discriminate and adapt to the different market states (risk scenarios) that evolve during a typical trading day. Volatility is used as a statistical measure of the dispersion of returns to discriminate between the different market states.

Financial time series exhibit so-called stylised facts patterns (see Cont, 2001; Gencay et al., 2002). Volatility for the high-frequency trader poses a tricky scenario. Sufficient market volatility is required to ensure that changes in prices exceed transaction costs. On the other hand, higher volatility reflects higher uncertainty and poses the risk of adverse market movements resulting in losses. From a model design perspective, applying machine learning techniques to non-stationary time series to infer predictions becomes a more difficult challenge with the possibility of increased error variance.

This chapter is divided into two main parts:

In Part 1 (Section 4.2), we argue that although a wide body of research claims that volatility exhibits strong intraday periodicity, high persistence and can be predicted with a good degree of accuracy up to an intraday level (Andersen and Bollerslev, 1997; Andersen et al., 2000b), the literature of how AI algorithms can benefit from this information for trading purposes is lacking. Although a number of authors have suggested methodologies of how to design and tune AI models when applied for predicting financial time series, less consideration is given to the time-varying volatility (see Cavalcante et al., 2016; Kaastra and Boyd, 1996; Vanstone and Finnie, 2009, 2010), especially from the fuzzy logic domain. Other authors intentionally avoid higher periods of uncertainty by keeping out of the market especially during the initial and end intraday trading periods (e.g. Brabazon and O'Neill, 2006), which are well documented to exhibit the highest volatility (risk).

As our main contribution, we introduce a fuzzy logic based approach to better control trading algorithms in view of intraday realised volatility which we use as a proxy for uncertainty. Due to the popularity of Neural Networks (NNs) in non-linear times series applications (as identified in surveys by Bahrammirzaee, 2010; Cavalcante et al., 2016; Krollner et al., 2010; Tsai and Wang, 2009), as a motivating example we make use of a popular Neural Network (NN) model and use it as our benchmark to measure the effectiveness of our approach. By extending the concept of heat maps presented in Chapter 3 to identify preferable trade regions, we introduce a fuzzy logic hybrid model which consists of a standard NN enhanced with an innovative and efficient volatility filter based on fuzzy c-means (FCM) clustering (Bezdek, 1981). The proposed extension monitors the performance of the underlying NN model across various market volatility levels at an intraday level. Subsequently, it dynamically identifies unique intraday scenarios and creates a dynamic and visually apprehensible risk-return search space to control algorithmic trading decisions. This feature also mitigates the common black box criticisms that are typically attributed to standard NN techniques. The proposed technique does not limit the possible use of the proposed extension with other machine learning techniques, however in our case we focus on possible NN improvements. Our results show that this method can be successfully applied to support high-frequency intraday trading strategies, outperforming both standard NN and buy-and-hold models.

In Part 2 (Section 4.3), we extend the findings from Part 1 and as our first contribution we introduce a fuzzy logic approach to control risk-based money management decisions. Pardo

(2011) suggests that models that do not take into consideration effective money management decisions can lead to sub-optimal solutions. Our approach in this chapter also goes contrary to studies suggesting the use of *fixed* return and volatility thresholds (Holmberg et al., 2013; Kaastra and Boyd, 1996; Vanstone and Finnie, 2009, 2010). In line with Part 1, we propose an innovative fuzzy logic method which identifies the approximate regional performance across time-varying intraday risk-return states. Contrary to many studies that limit the application of AI to focus solely on market direction (as indicated by Cavalcante et al., 2016; Krollner et al., 2010; Tsai and Wang, 2009), we introduce an effective money management approach to dynamically adjust trading frequency and order position size by discriminating across different regions in the trend and volatility space with the objective to improve the overall risk-adjusted trading performance. This enhances the trading model presented in Chapter 3 which considers equally-sized trade positions and does not discriminate across different intraday volatility states. A key challenge is that our approach should not just act as a filter by just keeping the algorithm out of the market in adverse regions and hence underutilise available capital, but also improve model profitability by allocating more capital to preferable intraday trading scenarios. By adopting rigorous statistical tests, our results show significant performance improvements compared to both standard NN and buy-and-hold approaches.

From a different perspective, this argument contradicts the Efficient Market Hypothesis (EMH) (Fama, 1965, 1970) since, according to the theory, discriminating between different volatility scenarios should in the long term lead to no additional benefits. The EMH, which resulted in Eugene Fama being crowned with a Nobel Prize in Economics in 2013, remains without any doubt one of the most debated theories in economics. However, research claims from a number of authors (Gradojevic and Gençay, 2013; Han et al., 2013; Holmberg et al., 2013; LeBaron, 1999; Schulmeister, 2006) indicate a relationship between the possible periodic breakdown of market efficiency and volatility. To our knowledge, this has never been explored at intraday short term horizons.

In Part 2, our second contribution is that we provide tenable reason to infer a relationship between volatility levels and possible breaks in market efficiency during short intraday time horizons, hence we extend the literature by an investigation of this phenomena at more granular time frames. As a direct consequence of our first contribution, the resulting increase in trading performance using our proposed method suggests a possible martingale property breakdown during particular intraday trend-volatility states.

4.2 Part 1 - Enhancing the intraday trading performance of neural networks using a dynamic volatility clustering fuzzy filter

Andersen and Bollerslev (1997) and Andersen et al. (2000b) demonstrate a well-known stylised fact in finance literature that denotes the strong diurnal pattern exhibited by volatility. This is important from an algorithmic design perspective because it infers that risk is not constant during the course of a trading day. However in surveys like Krollner et al. (2010) and Cavalcante et al. (2016) it is evident that although numerous models are presented in finance literature that are capable of measuring volatility at a good degree of accuracy up to an intraday level, the literature of how AI-trading algorithms can be optimised to handle *intraday* time-varying volatility is scarce, if considered at all.

In Part 1 of this chapter, the objective is to investigate whether by discriminating trading performance across intraday trend and volatility states at short intraday time horizons, this can lead to an improvement in the risk-adjusted performance of the tested neuro-fuzzy systems. The aim is to augment trend prediction signals with information about preferable states and, as a result, enhance the control and performance of trading algorithms. To our best knowledge this has not been investigated in literature so far. This is enhanced further in Part 2 of this chapter by combining the proposed models with capital allocation decisions.

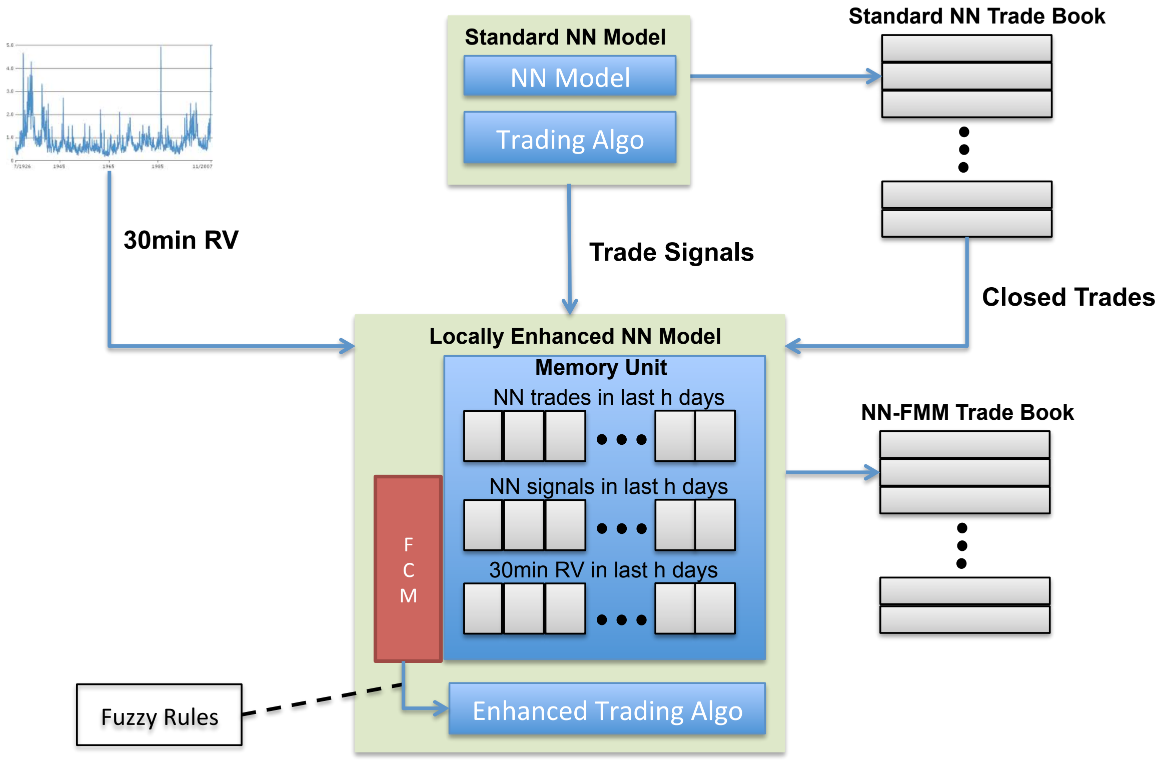


Fig. 4.1: Experiment setup showing our proposed locally enhanced NN model.

Part 1 is structured as follows. In Section 4.2.1 we introduce our experiment approach and describe the main model components. We then describe the base NN model and the data preparation process (Section 4.2.2), followed by a description of the applied trading algorithm (Section 4.2.3). We then discuss the challenges that arise due to intraday volatility (Section 4.2.4) and present our enhanced trading algorithm (Section 4.2.5). Section 4.2.6 presents our experiment approach, followed by our experiment results and discussions in the light of existing literature (Section 4.4). Section 4.2.8 concludes Part 1 of this chapter.

4.2.1 Method

Our experiment setup has two main components (see Figure 4.1). First, a standard NN, which is very popular in trading applications (Cavalcante et al., 2016; Choudhry et al., 2012; Krollner et al., 2010; Tsai and Wang, 2009), is employed to forecast the average return over the following 5 minutes, y_t, using standard trading rule signals. This model is similar to the NN model presented in Chapter 3, and on the same lines we also employ a global risk-return objective function for model calibration purposes. The standard NN model is also used as one of our benchmark models.

In this chapter we also introduce a second component, a locally enhanced NN model. The output from the standard NN model is passed as input to the locally enhanced NN model extension. We base our approach on fuzzy set theory (Zadeh, 1997). Since our problem deals with intraday uncertainty as a result of time-varying volatility, we present a hybrid model consisting of an additional dynamic fuzzy logic module. The module discriminates between different market states and enhances standard optimisation techniques that are typically applied to NN models by additionally taking into consideration different time-varying volatility states. We adopt a novel algorithm which takes past trade suggestions from standard NN as input, limited to the past h days, and infers the approximate regional (rather than global) risk-adjusted performance. Core to our controller is a fuzzy c-means clustering algorithm (Bezdek, 1981; Dutta and Angelov, 2010) that identifies unique trading performance regions across two dimensions: (a) the intraday realised volatility (Andersen and Bollerslev, 1997) which is used as a proxy for uncertainty, and (b) the predicted return size which indicates a trend direction. The identified fuzzy clusters allow the extraction of fuzzy rules and their combined result produces a decision surface across the trend and volatility space that is used to identify preferable trade regions.

We denote k as the number of clusters, $\mathbf{g}$ as a point in the trend-volatility space, C_j as the j-th induced fuzzy cluster in the trend-volatility space and $P_j(\mathbf{i})$ as a performance measure based on the profits $\mathbf{i}$ from trades which were executed *close* to that cluster region in the last h days. We define *close* as being a point whose degree of membership $\mu_{C_j}(.)$ to a specific cluster exceeds a certain threshold θ. For each cluster, a rule is extracted in the form

$$\begin{aligned} &IF \quad && \mathbf{g} \in C_j \\ &THEN \quad && i_{loc} = P_j(\mathbf{i}|\mu_{C_j}(\mathbf{g}) > \theta) \end{aligned} \tag{4.1}$$

for $j = 1, 2, ..., k$. The simple fuzzy rule (Equation 4.1) can be interpreted as *"if trend and volatility levels are close to region C_j, then the approximate regional risk-adjusted performance is i_{loc}"*. Detail on our fuzzy rule extraction and local performance calculation is presented in Section 4.2.5.

The full database of rules is updated on a daily basis and their combined output produces a decision surface that is used to approximate the trading algorithm risk-adjusted performance at a granular intraday level according to different trend and volatility states. The global trading performance is then optimised by dynamically accepting trade signals in regions that are identified to yield higher risk-return performance and reduces (or even stops) the trades in less favourable ones.

The following Sections 4.2.2 to 4.2.5 describe in more detail each model component and their calibration and selection.

4.2.2 Technical indicators and standard neural network model

The concept of a trend is absolutely essential to the technical approach (Murphy, 1987). In essence, one of the premises of technical trading is that prices move in trends. The objective is to identify a trend reversal at an early stage and ride on that trend until the underlying algorithm indicates that the trend has reversed. In our experiment, we intentionally make use of simple trading rules in order to investigate the simplest versions of trading rules used in common practice (Vanstone and Finnie, 2009, 2010). Our trading rules to derive a set of trading decisions $\Omega = \{long, short, 0\}$ are primarily based on n historical prices $\{p_t, p_{t-1}, ..., p_{t-n+1}\} \in \mathbb{R}^n_+$. Since the objective of this chapter is to be able to discriminate between different intraday market scenarios, we have to ensure that enough trade signals can be generated by our algorithm to be able to cover an adequate number of trade scenarios during the course of a trading day. For this reason, in this chapter we increase the price frequency adopted in Chapter 3 (5 minute) to 1-minute trade prices.

A moving average in essence represents a low pass filter which removes higher frequency noise, the combination of which provides a convenient signal (explained in Section 3.2.1). Short and long moving averages, MA_t^n, are calculated as defined by Equation (3.1) and the corresponding trade signal, $s_{t-1}^{n_1,n_2}$, defined by Equation (3.3).

In this experiment we include another popular trading signal which is defined using the Relative Strength Index (RSI) indicator

$$RSI_t = 100 - 100/(1 + RSF), \tag{4.2}$$

where the relative strength factor (RSF) is calculated by dividing the average of the gains by the average of the losses within a specified time period. A common RSI-based signal typically suggests that an asset is considered overbought if RSI_t reaches a value of 70, indicating a sell signal due to possible overvaluation. Similarly, if RSI_t drops to 30, the asset is likely to become oversold and undervalued, indicating a buy signal.

As indicated in Chapter 3, in our preliminary feature selection process we try to identify the correct moving average and RSI features that help us to reduce bias, however we limit our inputs to four variables in order to reduce model complexity and hence the possibility of overfitting (higher variance). We investigate whether the mean return for the next 5 minutes can be successfully predicted by making use of moving average and RSI signals and use this information to build a profitable trading algorithm. In our experiment, the mean return, y_t, is defined as

$$y_t = log(MA_{t+5}^5) - log(p_t). \tag{4.3}$$

The three moving average rules utilised have the lag structures $(n_1, n_2) \in [(1,5), (5,10), (10,20)]$ (see Equation (3.1)), where n_1 and n_2 are expressed in 1-minute time bars. For our RSI indicator we calculate this based on prices in the last 30-minute time slot. By combining these input signals, the linear specification of the return y_t prediction model is defined as:

$$y_t = \theta_0 + \sum_{k=1}^{4} \theta_k s_{k,t-1} + \varepsilon_t \tag{4.4}$$

with the error term $\varepsilon_t \sim N(0, \rho)$ and

$$s_{k,t} = \begin{cases} MA_t^{1,5} & \text{for} \quad k=1 \\ MA_t^{5,10} & \text{for} \quad k=2 \\ MA_t^{10,20} & \text{for} \quad k=3 \\ RSI_t & \text{for} \quad k=4 \end{cases} . \tag{4.5}$$

To generate the trend signals, we apply an NN model. We train our NN model with a set of moving averages and RSIs (relative strength indicators) in order to predict the next average 5-minute return every 1-minute time epoch. The process starts from a given collection of N input-output data training pairs, $\left(\mathbf{x}^{(1)} : y^{(1)}\right), \left(\mathbf{x}^{(2)} : y^{(2)}\right), ..., \left(\mathbf{x}^{(N)} : y^{(N)}\right)$ where

$$\begin{aligned} &\mathbf{x}^{(1)} = [s_{1,t-N}, s_{2,t-N}, s_{3,t-N}, s_{4,t-N}], && y_{t-N+1}^{(1)} \\ &\mathbf{x}^{(2)} = [s_{1,t-N+1}, s_{2,t-N+1}, s_{3,t-N+1}, s_{4,t-N+1}], && y_{t-N+2}^{(2)} \\ &\cdots \\ &\mathbf{x}^{(N)} = [s_{1,t-1}, s_{2,t-1}, s_{3,t-1}, s_{4,t-1}], && y_t^{(N)}. \end{aligned} \tag{4.6}$$

In Equation (4.6), for each data instance at a specific time t, $\mathbf{x}$ is a vector consisting of $\{x_1, x_2, x_3, x_4\}$ input elements which represent the $\{s_{1,t-1}, s_{2,t-1}, s_{3,t-1}, s_{4,t-1}\}$ technical indicator signals (Equation (4.3)), and y represents the mean return over the next 5 minutes (Equation (4.5)).

A number of authors indicate that claims of excessive returns from intraday trading is potentially based on spurious results (e.g. Kearns et al., 2010; Schulmeister, 2009). Our NN model consists of a single feed-forward network model (Figure 4.2). Although more signal combinations and complex NN models can be explored to gain higher accuracy, we decide to adopt a parsimonious approach to reduce the possibility of overfitting and spurious results (see Bailey et al., 2014).

Following Equation (4.4), we design a single-layer feed-forward network (FFN) regression model with 3 lagged buy and sell signals and 1 RSI signal as inputs with d hidden units. Each hidden neuron, z_j, can be represented mathematically as

$$z_j = f\left(w_{1,0} + \sum_{k=1}^{4} w_{1,k} x_k\right), \tag{4.7}$$

where $w_{1,0}$ represents the bias parameter of a hidden neuron and $w_{1,k}$ represents the weight parameter between the k-th input and the hidden neuron z_j. The output neuron can then be represented as

$$y_t = f\left(w_{2,0} + \sum_{j=1}^{d} w_{2,j} z_j\right), \tag{4.8}$$

where $w_{2,0}$ represents the bias parameter of the output neuron, and $w_{2,j}$ represents the weight parameter between the j-th hidden neuron and output neuron. Although in Chapter 3 we used the logistic function as our activation function, during preliminary testing we identified that in this configuration the hyperbolic tangent function presented better results. Hence, in this experiment the activation function $f(u)$ takes the form

$$f(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}}. \tag{4.9}$$

For our model identification and calibration, all $2 \times 3 = 6$ permutations of the parameter constellations combined with a range of up to 1000 training epochs are considered during the

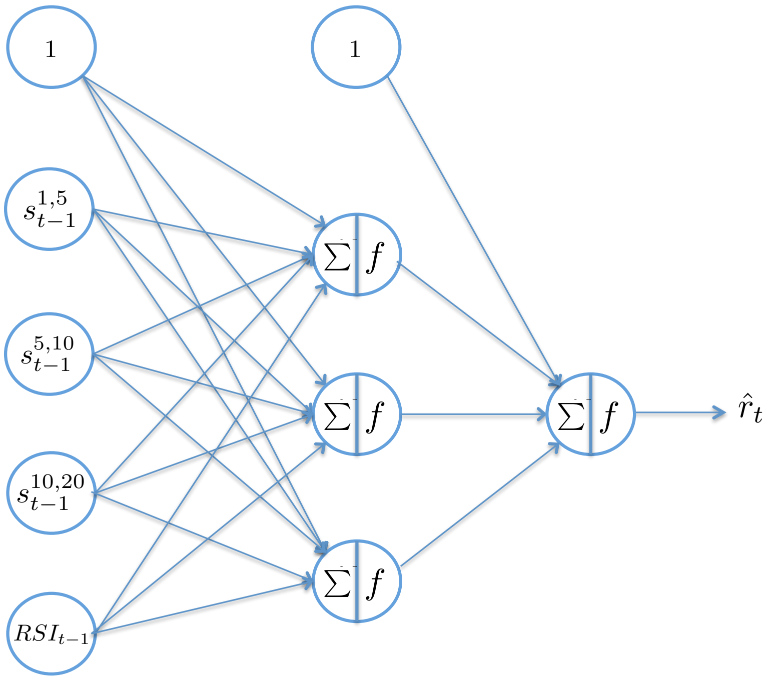


Fig. 4.2: Neural network setup, example showing 3 units in hidden layer.

Table 4.1: Parameters tested for NN model

Parameter	Parameter Value Set
Training Data Size (price points)	{510, 1020, 2040}
Number of Hidden Units, d	{5, 10, 20}
Training Epochs	{1,...,1000}

in-sample training (see Table 4.1). For efficiency purposes, we use the Levenberg–Marquardt algorithm for our backpropagation learning (Marquardt, 1963).

Model selection is performed by testing and comparing all settings with the different combinations of model parameters, identifying the model with the highest Sharpe ratio (see Equation (3.20)) during the in-sample period. Trading strategies with higher Sharpe ratios are always preferred as they translate into better risk-adjusted performance due to their higher risk premium for a given level of risk. This approach falls in line with our findings in Chapter 3 and is purposely selected to compare the improvements that can be attained by our new fuzzy logic extension which is discussed in Section 4.2.5.

4.2.3 Standard trading algorithm

The trading algorithm takes the next 5-minute average return prediction, every minute, from the NN model, and performs trades based on specific rules. We again use the stock data presented in Chapter 3, covering the period between 28/06/2007 and 25/06/2008. The general structure of the trading algorithm is kept similar to the approach presented in Chapter 3, Section 3.2.6, with trading activity starting at 8:30 and the latest possible trade generated at 16:00 with no overnight positions. In this chapter however, we perform a re-calibration due to a change in price frequency from 5 minutes to 1 minute. The similarity in approach was purposely kept in order to attribute any performance improvements to the proposed fuzzy logic extension.

The objective of our trading algorithm is to generate *buy* ($\Omega = long$), *sell* ($\Omega = short$) or *stand-by* ($\Omega = 0$) signals. For *buy* or *sell* signals the predicted return value has to be greater (smaller) than the upper (lower) return threshold RT, otherwise the trade signal is set to *stand-by*. We test different RTs between 0.05% and 0.1% for each stock. The lower threshold is selected based on the minimum return required to cover the fixed transaction costs defined in the algorithm. Similar to Chapter 3, we apply a constant transaction cost of 10 GBP per trade, per direction, and assume that a trader is willing to invest a fixed $50,000$ GBP per position. In our in-sample period we try to identify the least possible RT until the algorithm resulted in profitable results at the end of the in-sample period. An increased threshold results in a substantial reduction in intraday trades (Figure 4.3). Following our tests across our stocks, we set 0.07% as our fixed RT. We consciously set RT as fixed in Part 1 of this chapter in order to clearly attribute possible improvements to our proposed volatility fuzzy filter. A variable RT approach is then utilised in Part 2 of this chapter. In line with Vanstone and Finnie (2010), our trade decision takes account of the individual neural network threshold and also whether the signal is increasing in strength from its previous forecast. By

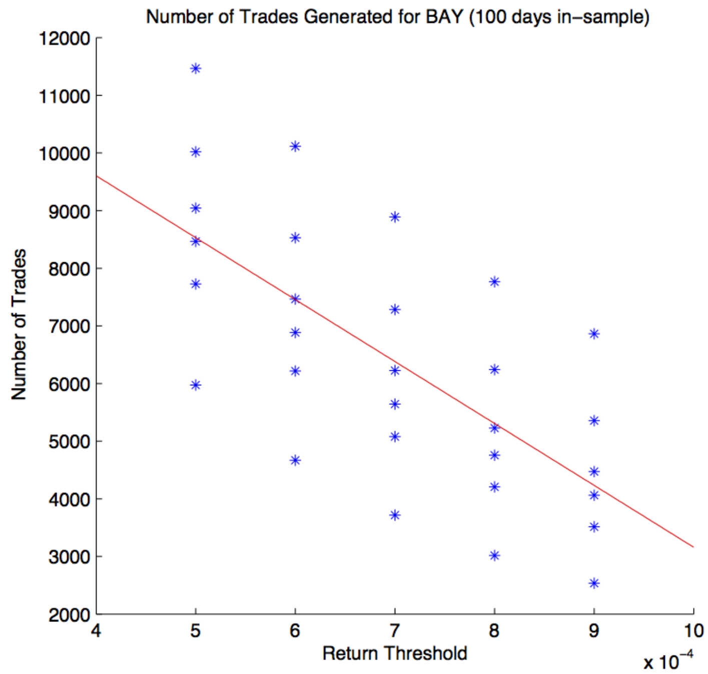


Fig. 4.3: Effect of applying different return thresholds on the number of trades

denoting the predicted mean return at time t as $\hat{y}_t$, the position Ω taken at time t is:

$$\Omega_t = \begin{cases} long: & \hat{y}_t > RT, \hat{y}_t > \hat{y}_{t-1} \\ short: & \hat{y}_t < -RT, \hat{y}_t < \hat{y}_{t-1} \\ 0: & otherwise. \end{cases} \tag{4.10}$$

Every 1 minute, starting from 8:30, the trading algorithm takes a decision based on the predicted trading direction, the selected return band and the elapsed position holding time. If the signal is to go long (short) the system will buy (sell) $50,000$ GBP worth of stock at the current market price. The total investment at each point in time is limited to $250,000$ GBP (which is also the amount of the initial capital allocated for trading on one specific asset). For this experiment only positions in the same direction are allowed at the same time. This is done to specifically eliminate the hedging effect of opposing positions which, as a result, can overestimate the performance of the algorithm.

When a trade is placed, if after a trade duration (TD) of 5 minutes the signal is still in the same direction, then the position is kept for another period of the same length. If, on the contrary, the signal has changed, then the position is closed. This short TD was selected to simulate real live intraday scenarios in which traders might wish to come out of the market quickly. Brabazon and O'Neill (2006) showed that the similar use of extended close in intraday trading scenarios can perform better than standard stop-loss, take-profit and buy-and-hold strategies. All open positions are closed at end of day, resulting in the system not holding any positions overnight.

4.2.4 Intraday volatility and its challenges

An important aspect which goes hand in hand with trend identification, and hence on the outlook of the potential return, is the risk involved, which is statistically represented by volatility.

Volatility brings about a dual contrasting agenda for traders. At face value, volatility brings negative risk associations indicating pull-back signs due to the possibility of cliff-like drops in prices in the underlying unstable scenarios. However, wide ranging moves can also be considered as profitable opportunities for traders to lock extraordinary gains. Adequate volatility is still required for favourable price movements sufficient to overcome the trading costs. As stated by Kearns et al. (2010), in the high-frequency trading domain, the typically shorter holding periods demand more extreme (and thus less frequent) relative price movements to break the costs barrier. Our objective is to identify a method can measure and subsequently categorise stock volatility across different levels during the course of a trading day. This will permit our trading algorithm to dynamically adapt to the different levels of intraday volatility.

The classical econometric tool to model the volatility of daily ("low-frequency") returns is the GARCH framework (Engle and Patton, 2001; Hansen and Lunde, 2005; Poon and

Granger, 2003, 2005). In a high-frequency setting, however, the literature suggests a different approach. The idea of using intraday daily data to measure volatility was first introduced by Merton (1980) who noted that under the diffusion assumption, volatility can be estimated to an arbitrary precision using the sum of intraday squared returns.

Andersen et al. (2001) introduced a natural estimator for the quadratic variation of the process which is commonly known as realised variance (RV). By considering a general jump-diffusion model for the log-price p of an asset (Liu et al., 2015)

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) + \kappa(t)dN(t) \tag{4.11}$$

where μ is the instantaneous drift, W is a standard Brownian motion, κ is the jump size, and N is a counting measure for the jumps. In the absence of jumps, the quadratic variation (QV) of the log-price process over period $t+1$ is defined as

$$QV_{t+1} = \underset{n \to \infty}{plim} \sum_{j=1}^{n} r^2_{t+j/n} = p_{t+j/n} - p_{t+(j-1)/n} \tag{4.12}$$

where

$$r_{t+j/n} = p_{t+j/n} - p_{t+(j-1)/n} \tag{4.13}$$

and the price series on period $t+1$ is assumed to be observed on a grid of n times $\{p_{t+1/n}, \ldots, p_{t+1-1/n}, p_{t+1}\}$. The standard realised variance estimator, which is the empirical analog of QV is calculated as

$$RV_{t+1} = \sum_{j=1}^{n} r^2_{t+j/n} \tag{4.14}$$

Contrary to classical models in mathematical finance where volatility is considered latent, the literature on realised volatility considers volatility as an observable variable and hence can be easily analysed and applied in our model (Andersen et al., 2001).

Although RV is approximately free of measurement error under general conditions (Andersen et al., 2001), it is susceptible to microstructure bias (Andersen et al., 2011; Hansen and Lunde, 2006) as well as abrupt jumps in financial markets (Barndorff-Nielsen and Shephard, 2004). A common problem with using RV is that if the interval is "too small", the RV estimator can suffer from market microstructure noise present in high-frequency financial data, which can make the estimator biased.

Andersen et al. (2000a) suggest what they call a "volatility signature plot" as a tool to indicate a suitable sampling frequency for calculating the realised volatility. A volatility signature plot provides an easy way to visually inspect the potential bias problems of RV-type estimators. We apply the same approach so that, firstly, identify whether microstructure bias manifests itself in our set of stocks. Secondly, the intention is to identify the sampling frequency which represents a reasonable tradeoff between minimising micro-structural bias and minimising sampling error. In Figure 4.4 we show the volatility plots for a sample of stocks from our data set. The stocks that we present exhibit different levels of liquidity (please refer to data statistics in Section 3.2.7) since the low liquidity stocks might exhibit a different volatility signature from stocks with high liquidity (see Andersen et al., 2000a). The signature plots that we present show that for all stocks the average realised volatility exhibits highest change in sampling frequencies less than 5 minutes, beyond which it tends to start stabilising. This falls in line with literature that usually suggests a sampling interval of 5 minutes (McAleer and Medeiros, 2008). Liu et al. (2015) also find that the standard 5-min RV provides an RV estimate which compares well with many other more sophisticated RV estimators.

Liu et al. (2015) investigate nearly 400 variations from 7 categories of realised variance estimators, and also refer to various parameters that practitioners must fine tune including the sampling size, sampling scheme, whether to use transaction prices or mid-quotes, and other measure specific parameters such as the kernel length. In the subsequent sub-section we

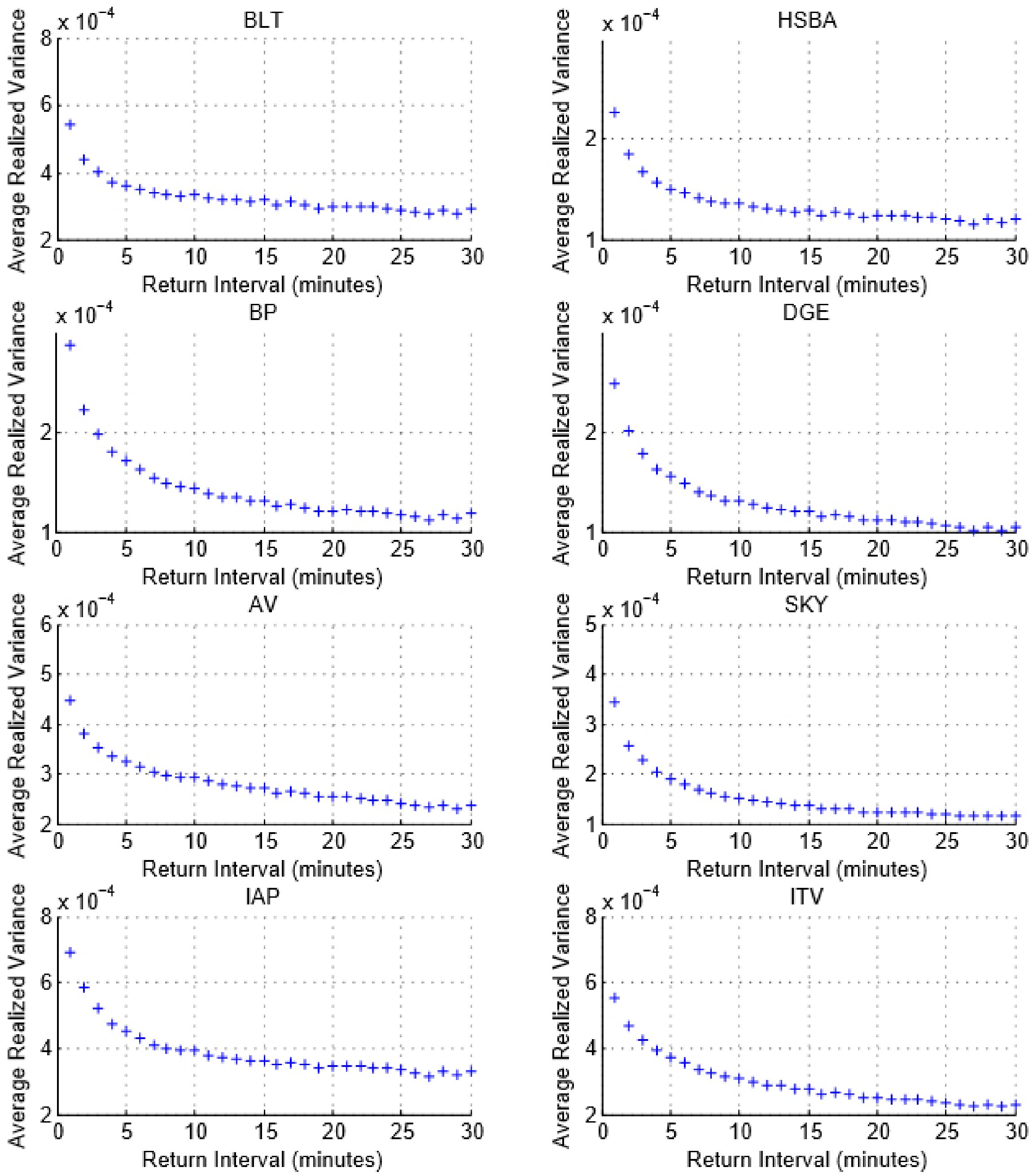


Fig. 4.4: Realised volatility signature plots for a selection of stocks from our dataset.

make use of the average RV estimator which was studied and identified as "second best" by Zhang et al. (2005). The authors show that subsampling and averaging result in a substantial decrease in the bias of the estimator. Subsampling involves using a variety of "grids" of prices sampled at a given frequency to obtain a collection of realised measures, which are then averaged to yield the "subsampled" version of the estimator. For example, 5-minute RV can be computed using prices sampled at 9:30, 9:35, etc. and can also be computed using prices sampled at 9:31, 9:36, etc. An average of the results is then calculated across the subgrids. Andersen et al. (2011) and Liu et al. (2015) have identified that the average RV estimator performs very well when compared to the more sophisticated estimators. In the next section we explain how we apply average RV to identify different volatility levels during the course of a trading day.

4.2.5 Enhancing trading model with a dynamic volatility filter

A number of studies do not consider volatility when designing NNs (see Kaastra and Boyd, 1996; Vanstone and Finnie, 2009, 2010). Others (e.g. Brabazon and O'Neill, 2006) simply suggest to avoid highly volatile periods. Although volatility is a proxy of uncertainty, merely avoiding volatility, however, can result in missed opportunities for large returns. The objective of our fuzzy volatility filter is to identify regions of predicted returns and intraday realised volatility where, based on a number of past trades, have achieved good risk-return results, and hence filter out those that result in losses. This is done by creating a simple learning mechanism to approximate the NN sub-regional trading performance which is subsequently used by our trading decision model to automatically control trade signals suggested by the NN and filter out those signals falling in regions with a lower likelihood of success.

We apply the average RV as our RV estimator to estimate the level of volatility in the last 30 minutes. Every 1-min we measure 5-min RV using 5 grids of 5-minute returns overlapping by 1 minute over the previous 30-minute slot. It has to be noted that this period covers the time slot from which our technical rules inputs are calculated. Since our trading algorithm starts trading at 8:30, our first RV calculation happens at 8:30. Our RV estimate from the first grid is computed using prices sampled at 8:00, 8:05, ...,8:25. The RV estimate from the second grid is computed using prices sampled at 8:01, 8:06, ..., 8:26. Once 5 grids are calculated our RV_{t+1} estimate is calculated as the average of the 5-min RV estimates across the grids. We take this measure on a minute-by-minute level during the course of a trading day. The advantage of using average RV is that it is a moving average itself, and hence consecutive values will be (by construction) highly autocorrelated, providing a good linear predictor in our NN model. After examining the high correlation between the RV measure at time t and the RV measure at time $t+5$ minutes (correlation coefficient of 0.85, significant at 95% confidence level), we decided to adopt a parsimonious approach and use the RV at time t as a proxy for RV in the next 5 minutes.

We apply the incompatibility principle of Zadeh (1973). In our context, reducing the complexity tends to also reduce the uncertainty, and hence trading risk, of the underlying NN model, which is something we aim to achieve. We apply clustering to detect the possible volatility groupings. Most common fuzzy systems define relationships between variables by means of fuzzy if-then rules, which can be viewed as a local description of the system under consideration. A survey by Dutta and Angelov (2010) indicates that data clustering is one of the approaches that have been applied extensively to automatically generate rules from data. One of the most popular methods for finding fuzzy partitions is FCM clustering (Bezdek, 1981). It has also been shown that FCM can be extended to support big data sets (see Havens et al., 2012), an important requirement when dealing with high-frequency trading data.

The input data is organised into data pairs from trade events which happened in the past h days. We denote the input to the fuzzy controller as a matrix $\psi = \{\mathbf{G}, \mathbf{i}\}$. We define $\mathbf{G}$ as a set of points $\mathbf{g}$ in the trend-volatility space, where as defined in Section 4.2.1 each point consists of our space coordinates $\{y_{t+1}, RV_{t+1}\}$. Each point $\mathbf{g}$ represents a past simulated

NN trade where y_{t+1} is the expected return signal over the next 5 minutes and RV_{t+1} is the realised volatility at the time the trade was placed. Vector $\mathbf{i}$ represents the corresponding profit resulting from each trade. Following our findings in Chapter 3, we define our trade profits *inclusive of transaction costs*. Since predicted return and volatility variables have different bases, these are standardised and rescaled to have a mean of zero and a standard deviation of one before being fed into the algorithm. This provides equal influence weight of each variable on the clustering algorithm, which is described next.

As a next step, we identify data clusters across the expected return and volatility dimensions. Depending on the memory size parameter defined in number of days, h, $\mathbf{G}$ will consist of l data points, each point represented by row vector $\mathbf{g}$, which is dependent on the number of executed trades covered by our memory span. Let z denote the fuzziness index. Furthermore, define α as the number of clusters, $d_{uv} = ||\mathbf{g}_u - \mathbf{c}_v||^2$ as the Euclidean distance between the u-th realisation and the current v-th cluster centre $\mathbf{c}_v$, and $d_{uo} = ||\mathbf{g}_u - \mathbf{c}_o||^2$ as the Euclidean distance from the u-th realisation and the other cluster centres $\mathbf{c_o}$. For each data point $\mathbf{g}_u$, $\forall u \in [1, l]$, and cluster $\mathbf{c}_v$, $\forall v \in [1, \alpha]$, the FCM algorithm iteratively updates the membership grade μ of the u-th data point to the v-th cluster

$$\mu_{uv} = \left(\sum_{j=1}^{\alpha} \left(\frac{d_{uv}}{d_{uo}} \right)^{\frac{2}{z-1}} \right)^{-1} , \tag{4.15}$$

and the centre of the v-th cluster

$$\mathbf{c}_v = \frac{\sum\limits_{u=1}^{l} \mu_{uv}^{z} \mathbf{g}_u}{\sum\limits_{u=1}^{l} \mu_{uv}^{z}} , \tag{4.16}$$

such that the objective function

$$J_z = \sum_{u=1}^{l} \sum_{v=1}^{\alpha} \mu_{uv}^{z} d_{uv} \tag{4.17}$$

is minimised.

The selection of the number of clusters, α, is a typical problem in cluster analysis. The literature proposes a number of approaches to construct α compact and distinct clusters that have small distances (variances) between points within the clusters and large distances between points belonging to different clusters Xu et al. (2005). Following our findings in Chapter 3, we take a different approach in selecting α by identifying the best number of clusters as part of our calibration process with the objective to maximise global risk-return ratios based on past trading performance.

A set of α rules are dynamically created using fuzzy sets derived directly from the components of the α centroids found by clustering on the input space and is updated on a daily basis using local fuzzy cluster risk-return trading information. These fuzzy sets are not projections of the clusters but fuzzy sets induced in the input space by the fuzzy clusters. The rules take the form of Equation (4.1). Since the summation of membership for each data point should be equal to 1 (Equation (4.15)), we decided that for our approximate calculation of the local regional performance function we apply a threshold of $\theta = 1/\alpha$. For our risk-return function $P(\mathbf{i})$ we applied the Sharpe ratio (Equation (3.20)). Hence, our rules take the form of

$$\begin{aligned} IF \quad & \mathbf{g} \in C_j \\ THEN \quad & i_{loc} = Sharpe(\mathbf{i} \mid \mu_{C_j}(\mathbf{g}) > 1/\alpha), \end{aligned} \tag{4.18}$$

where i_{loc} is the approximated local risk-adjusted performance.

It is well known that machine learning algorithms may suffer from spurious inferences and that precise relationships that might have been stable in the past may not hold in the longer term future, hence becoming unreliable (Aldridge, 2013). For this reason, when considering a new data point, rather than assigning the point to a specific cluster, we adopt a fuzzy approach by obtaining the degree of membership of the new point to the identified

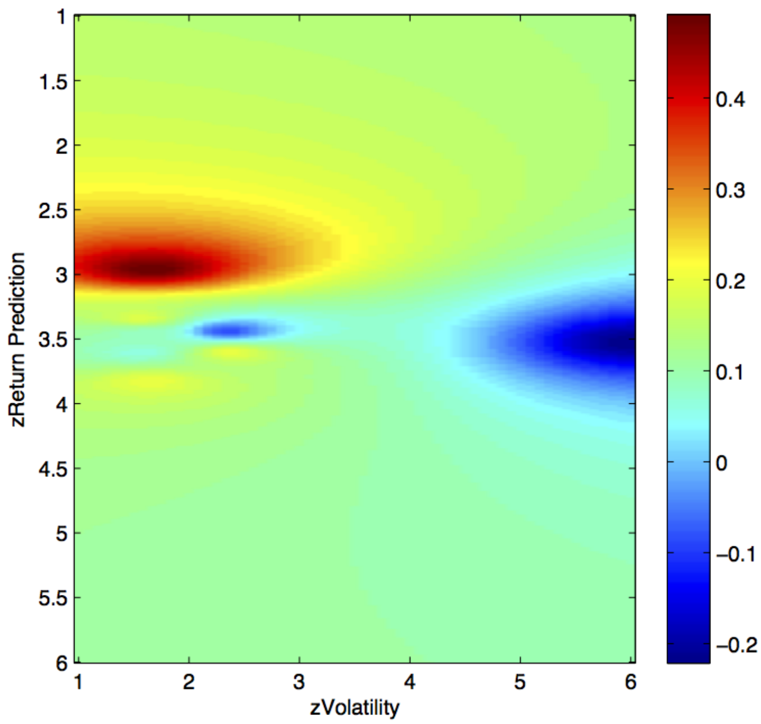


Fig. 4.5: Decision surface example as identified by our volatility clustering fuzzy filter showing different levels of Sharpe ratio regions (see colour bar).

clusters. By applying a threshold, rather than a crisp cluster boundary, the approximate local performance calculation in each cluster is influenced by the closest points to each cluster centre. This means that points close to cluster boundaries can be used in the approximate local performance calculation of multiple clusters. When projecting the cluster performance on a chart, this produces a smooth, and interpretable, decision surface showing the highest Sharpe ratio across the return prediction strength and realised volatility space (Figure 4.5).

For each new data point, the decision surface is used by our trading algorithm to take trade entry decisions. Every new, minute-by-minute, trend signal generated by the underlying NN module, together with the corresponding realised volatility measure, is passed as a new data point **g** to the fuzzy logic module to obtain the approximate regional performance as

$$S = \sum_{j=1}^{\alpha} \mu_{C_j}(\mathbf{g}) i_{loc,j}, \tag{4.19}$$

giving weight to the regional performance of the closest clusters.

Two sets of rules were tested for our trading algorithm. The first test (NN-FCM1) enhances the standard trading rules (Equation (4.10)) by allowing trades only if the regional performance, S, at point $\mathbf{g}$ is greater than the minimum positive risk-return performance attained amongst all clusters:

$$\Phi_t = \begin{cases} long: & \hat{y}_t > RT, \hat{y}_t > \hat{y}_{t-1}, S > min(i^{+}_{loc}) \\ short: & \hat{y}_t < -RT, \hat{y}_t < \hat{y}_{t-1}, S > min(i^{+}_{loc}) \\ 0: & otherwise. \end{cases} \tag{4.20}$$

In our second test (NN-FCM2) the algorithm was allowed to trade only if the regional performance, S, at point $\mathbf{g}$ is greater than the average risk-return performance which is attained across all clusters:

$$\Phi_t = \begin{cases} long: & \hat{y}_t > RT, \hat{y}_t > \hat{y}_{t-1}, S > max(\bar{i}_{loc}, 0) \\ short: & \hat{y}_t < -RT, \hat{y}_t < \hat{y}_{t-1}, S > max(\bar{i}_{loc}, 0) \\ 0: & otherwise. \end{cases} \tag{4.21}$$

Two additional parameters were included in these extended models. These are the number of clusters to be identified and the number of trades, based on the number of past trade days (see Table 4.6). Similar to the identification and calibration of our standard NN model, we again test and compare all possible combinations of model parameters for the in-sample training, but we now have to *additionally* consider the $4 \times 4 = 16$ permutations of the new parameters.

Table 4.2: Additional parameters tested for the NN and the fuzzy volatility filter

Parameter	Parameter Value Set
Trade History Data Size, h (days)	{5, 20, 40, 60}
Number of Clusters, α	{4, 6, 8, 10}

4.2.6 Model selection and experiment approach

The trading systems in this experiment are developed using the same data set for the 27 stocks described in Chapter 3, Section 3.2.8. In this chapter, we generate a time-series of 1-minute prices from which we produce a combination of moving average signals. This higher frequency time series was selected in order to increase the capability to generate more trade signals and hence ensure that we cover trading opportunities across all the trading day.

As in the previous chapter, model training and validation is performed by applying a day-by-day moving window approach for a period covering 150 in-sample days (approach and considerations described in more detail in Section 3.2.7). The model configuration with the highest Sharpe ratio is then evaluated on the following 100 days in the out-of-sample period. In line with Chapter 3, in our experiments we consider an annual risk-free rate of 5%. A number of statistics are collected to analyse model performance. Similarly, we adopt the Sharpe ratio and excess return as our primary performance measures. We complement these measures with the Sortino ratio, number of trades, win ratio and profit ratio (the definition of these ratios is presented in Section 3.2.7).

As a benchmark model, we implement a standard buy-and-hold strategy, buying at the opening price on a daily basis, holding it over the course of the trading day, and selling at the closing price. The inclusion of this zero-intelligence benchmark model is to assess the usefulness and potential outperformance of our AI-controlled algorithmic trading strategies in general. In Part 2, we also introduce a random model to further evaluate the effectiveness of our models.

Table 4.3: Trading performance of the Buy & Hold model following the 100-day out-of-sample period.

	Buy-and-hold Model	
Stock	Sharpe Ratio	Excess Return
AL	-0.1645	-0.1623
ANTO	-0.0513	-0.0326
BLT	0.0130	0.0072
BA	-0.1226	-0.0886
BLND	-0.2061	-0.0813
SKY	-0.1545	-0.0504
CWC	-0.0934	-0.0377
AV	-0.1252	-0.0562
DGE	-0.1561	-0.0400
SDR	-0.1083	-0.0554
HMSO	-0.1518	-0.0639
REX	-0.0974	-0.0378
JMAT	-0.1091	-0.0341
HSBA	-0.0598	-0.0176
ABF	-0.1843	-0.0486
INTU	-0.1998	-0.0691
RIO	-0.0206	-0.0110
BP	-0.0216	-0.0064
SGE	-0.0850	-0.0344
LLOY	-0.1222	-0.0746
TSCO	-0.1215	-0.0448
GSK	-0.0782	-0.0278
AZN	-0.0313	-0.0121
HBOS	-0.1836	-0.1872
XTA	-0.0393	-0.0204
IAP	-0.1188	-0.0706
ITV	-0.2347	-0.1147

4.2.7 Analysis and results

We first examine the results obtained from the buy-and-hold strategy. Table 4.3 shows that the only positive result is obtained in BLT and the remaining 26 stocks display negative results.

We then examine the Sharpe ratio and excess return obtained by our three intelligent algorithms. Table 4.4 shows that neural network has positive results in 22 out of 27 stocks

(81%). NN-FCM1 obtains a similar number of positive results. NN-FCM2 increases the positive results to 24 stocks (89%). These figures indicate that our three intelligent models are effective in providing better performance than B&H and the risk-free investment that we consider in our experiment.

When we compare the Sharpe ratio results, we identify that the neural network model achieved the highest score in 6 out of 27 stocks (22%). The rest of the highest scores are distributed across NN-FCM1 and NN-FCM2, with the highest number of occurrences residing under NN-FCM1. The average Sharpe ratio across all stocks, which we only use as an indicative statistic, also indicates that on average the highest Sharpe ratio is obtained by NN-FCM1, followed by NN-FCM2.

The excess return results show that in this case the highest scores are more distributed across the three models, mainly between NN and NN-FCM1. The average excess return indicates that on average the NN model obtained a higher score. In conjunction with the Sharpe ratio results, this suggests that in general NN was a more risky model but resulted in higher returns. We extend our investigation by examining additional performance measures.

From Table A.3 we see that the Sortino ratio results are consistent with the Sharpe ratio scores. This indicates that even in the case of Sortino ratio our proposed NN-FCM1 and NN-FCM2 models showed better risk-adjusted performance than the standard NN. In the same table we identify that the number of trades performed by the standard NN was substantially higher than in the case of our proposed models. This provided a first indication that in the case of the fuzzy models, the lower profitability is possibly the result of less trades, albeit at lower risk.

Additional statistics are displayed in Table A.4. Results show that in terms of win ratio, NN-FCM2 generates the highest score, followed by NN-FCM1. The table also shows that profit ratio is higher in our proposed models, the highest being NN-FCM2. In conjunction with previous results, results indicate that the proposed fuzzy models exhibit superior risk-

Table 4.4: Trading performance following the 100-day out-of-sample period. Bold figures represent the best performance for a given stock across all three models.

	Sharpe Ratio			Excess Return		
Stock	**NN**	**NN-FCM1**	**NN-FCM2**	**NN**	**NN-FCM1**	**NN-FCM2**
AL	0.2469	0.3408	**0.3588**	0.5410	**0.5854**	0.3940
ANTO	0.0091	0.0411	**0.1136**	0.0084	0.0275	**0.0685**
BLT	0.1244	**0.1490**	0.1300	0.0380	**0.0404**	0.0297
BA	0.1923	**0.2093**	0.1482	0.1207	**0.1235**	0.0497
BLND	0.0845	**0.1259**	0.0860	0.0495	**0.0610**	0.0280
SKY	0.2153	**0.2200**	0.1857	**0.0682**	0.0473	0.0271
CWC	0.0025	-0.0947	**0.0923**	0.0009	-0.0278	**0.0198**
AV	-0.1164	**-0.0564**	-0.1235	-0.0367	**-0.0118**	-0.0319
DGE	0.2506	**0.2923**	0.1733	0.0563	**0.0592**	0.0268
SDR	-0.0006	-0.1645	**0.0242**	-0.0044	**-0.0513**	0.0110
HMSO	**0.3124**	0.1968	0.1961	**0.1732**	0.1043	0.0813
REX	0.1436	**0.2424**	0.2153	**0.1420**	0.0706	0.0410
JMAT	**0.1597**	0.0175	0.0314	**0.2182**	0.0036	0.0064
HSBA	**0.2023**	0.1620	0.1222	**0.0346**	0.0308	0.0188
ABF	0.0759	**0.1138**	0.0877	0.0316	**0.0342**	0.0243
INTU	**0.4409**	0.4056	0.3193	**0.2599**	0.2252	0.1232
RIO	-0.1493	-0.1552	**-0.1072**	-0.0745	-0.0505	**-0.0317**
BP	0.0155	**0.3286**	0.0412	0.0072	**0.0918**	0.0051
SGE	**0.2552**	0.2462	0.2181	**0.0913**	0.0850	0.0479
LLOY	0.0433	**0.0722**	0.0437	0.0257	**0.0299**	0.0082
TSCO	0.1960	0.1720	**0.2519**	**0.0533**	0.0426	0.0473
GSK	**0.2819**	0.2496	0.2708	**0.1015**	0.0639	0.0529
AZN	0.1461	**0.1867**	0.1518	0.0423	**0.0431**	0.0285
HBOS	-0.0787	0.0109	**0.0410**	-0.1067	0.0112	**0.0420**
XTA	-0.2277	-0.2496	**-0.1540**	-0.1185	-0.0754	**-0.0433**
IAP	0.0085	**0.1315**	0.0667	0.0076	**0.0477**	0.0233
ITV	0.2561	**0.3751**	0.2222	**0.1822**	0.1735	0.0430
Average	0.1145	**0.1322**	0.1188	**0.0708**	0.0661	0.0423

adjusted performance and increased capability to generate more profits over losses when compared to standard NN. This indicates higher quality trades. However, due to the lower number of trades resulting from the increased signal filtering the overall profit of fuzzy models was inferior. This aspect is addressed in Part 2 of this chapter. Our results show that although volatility information in conjunction with NN models is scarcely researched in literature, combining a fuzzy volatility filter results in higher quality trades.

4.2.8 Conclusion

Following our findings in Chapter 3 with respect to model tuning and selection when applied to trading purposes, in this chapter we take a more granular perspective and seek to address the challenges presented by time-varying volatility (risk) during the course of a trading day.

In this first part of the chapter, we focus on a well-known stylised fact in finance literature which denotes the strong diurnal pattern exhibited by volatility (Andersen and Bollerslev, 1997; Andersen et al., 2000b). This infers that risk is not constant during the course of a trading day. Although numerous models are presented in literature that are capable of measuring volatility at a good degree of accuracy up to an intraday level, the literature of how AI-trading algorithms can be optimised to handle *intraday* time-varying volatility is scarce, if considered at all. From a different perspective, this argument contradicts the Efficient Market Hypothesis (Fama, 1965) since, according to the latter, discriminating between different volatility scenarios should in the long term lead to no additional benefits.

As our main contribution we introduce a novel approach based on clustering techniques and fuzzy logic which adapts the underlying neuro-fuzzy trading model to the time-varying intraday risk. To show the effectiveness of our approach, in our experiment we combine our extension with an NN model that time and again has proven its popularity in a number of surveys. Our approach extracts fuzzy rules dynamically using FCM clustering and makes use of local fuzzy cluster information which is subsequently used by our dynamic fuzzy model to approximate the NN trading performance across different volatility states. This produces a decision surface that is used by our trading algorithm to control intraday trade signals suggested by the NN. Results show that our model obtains better risk-return performance than standard NN and buy-and-hold methods, indicating that intraday time-varying volatility information can improve NN trading models.

Despite the long-standing Efficient Market Hypothesis, a number of authors claim a link between the profitability of technical trading rules and volatility (Gradojevic and Gençay, 2013; Han et al., 2013; Holmberg et al., 2013; LeBaron, 1999; Schulmeister, 2006). In Part 1, we have conducted the first steps to extend these claims to shorter term intraday horizons. To our best knowledge, this has never been investigated in earlier studies and, more specifically, how this can translate into improved risk-adjusted performance of intraday trading algorithms. Our positive results from Part 1 formulate the basis for more optimisation and rigorous testing of this finding. Consequently, in the second part of this chapter we seek further trading improvements by extending our proposed approach to automate money management decisions, supported by a number of statistical tests. The objective is to address the challenge of optimising capital allocation across different levels of intraday risk scenarios with the ultimate objective of increasing overall risk-adjusted performance.

4.3 Part 2 – A dynamic fuzzy money management approach for controlling the intraday risk-adjusted performance of AI-trading algorithms

Part 1 of this chapter provides initial evidence that by discriminating across intraday trend signals and volatility levels, and subsequently targeting trades in more profitable regions, helps to improve risk-adjusted performance. The corollary to this approach is to ensure that filtering trades from specific unprofitable trend-volatility regions does not reduce the overall intraday trading activity with the possibility of unused capital. As an additional optimisation step, in Part 2 we hypothesise that rather than just filtering trade signals by discriminating across different trend-volatility regions, improved overall trading performance can be obtained by allocating more (less) trading capital to regions which exhibit better (worse) results. Pardo (2011) warns that models that do not take into consideration trade position sizing possibly result in sub-optimal solutions. Hence, in this second part of this chapter we address the problem of optimising capital allocation for intraday trading using fuzzy logic money management decisions. This is further supported by a rigorous experiment approach and statistical tests which also corroborate our initial findings.

Our approach stems from two gaps in the computational finance literature. Firstly, although most AI literature focuses on identifying market direction, traders in financial markets, on a daily basis, are repetitively presented with a sequence of decisions, with market direction being only one piece of the puzzle. This presents a very limited view of the applicability of AI in trading scenarios. Moreover, many studies in the existing literature do not reflect the rigid constraints that typically govern the trading desk (see also Vanstone and Tan, 2003). In particular, Pardo (2011) states that position sizing is often not appreciated and poorly understood in trading strategy design.

Secondly, as was elaborated upon in Section 4.2, although the proliferation of high-frequency data led to new measures of variation (Andersen et al., 2001; Barndorff-Nielsen and Shephard, 2004, 2006), making it possible to be predicted with a good degree of accuracy up to an intraday level (Andersen et al., 2000b), the use of this information for intraday trading purposes is rarely considered. In our opinion, reverting to AI models solely for market movement predictions with little consideration to the time-varying market uncertainty (risk) (Son et al., 2012) portrays an incongruent view by financial market practitioners since investors are mostly interested in risk-adjusted performance.

Our goal in Part 2 of this chapter is to address these two, albeit interrelated, literature gaps with the ultimate objective to enhance the risk-adjusted performance of neuro-fuzzy controlled trading algorithms in an intraday stock trading scenario. In particular, we present an innovative method which identifies, in addition to trend direction signals, the optimum capital allocation across an intraday trading period by dynamically adapting the levels of trade frequency and position sizes (two common decisions taken by traders) based on different degrees of expected return and volatility (uncertainty) over short intraday horizons. This also sheds more light on the theoretical claims of Holmberg et al. (2013) that market inefficiency, and hence the profitability of technical rules, can be linked to different volatility periods.

Finding the optimum level of trade frequency and position size along continuously changing intraday market conditions can be a non-trivial task. Although volatility is typically linked to risk, sufficient market volatility is required to ensure that changes in prices exceed transaction costs. Several studies suggest a fixed return threshold filter to avoid small unprofitable movements (Kaastra and Boyd, 1996; Vanstone and Finnie, 2009, 2010). Holmberg et al. (2013) show that increasing the return filter size results in a better success rate and average return. However, the same authors show that this leads to a reduction in the number of trades, hence reducing the investors' potential profits. Other authors (e.g. Brabazon and O'Neill, 2006) prefer to avoid high volatility (uncertainty) completely by staying out of the market during periods when volatility goes beyond a specified fixed threshold, hence avoiding the risk of possibly strong adverse market movements. This, however, reduces the opportunity of possibly extraordinary gains. Also, Pardo (2011) claims that if a sound position sizing strategy or algorithm is not employed, the effective rate at which trading equity is compounded will remain sub-optimal.

These arguments cast the trade frequency and order position size decisions in the context of the better management of uncertainty. We make use of the model presented in Section 4.2 and extend the idea to identify fuzzy clusters which in turn allow for the extraction of fuzzy rules. Their combined result produces a decision surface across the trend and volatility space that is used to adapt trade frequency and position sizing levels based on local (rather than global) regional performance.

With respect to the AI literature, many authors (see Lawrence et al., 1997; Luengo et al., 2009; Prechelt, 1996) question the validity of the experimental and statistical frameworks adopted in many published academic studies. Recently, Bailey et al. (2014) stated that although there are many academic studies that claim to have identified profitable investment strategies, their reported results are almost always based on in-sample statistics. Overfitting a trading strategy on in-sample data can produce (seemingly) impressive results during simulation but often also a devastatingly poor performance during real-time trading (see Pardo, 2011). In this second part of the chapter, we address these criticisms and adopt a thorough experimental framework using a number of independent repeated trials conducted on a set of stocks listed on the London Stock Exchange. Our statistical tests on *out*-of-sample data reveal that the proposed fuzzy money management approach leads to significant improvements when compared to standard NN and buy & hold methods.

The structure of Part 2 of this chapter is organised as follows. In Section 4.3.1, we describe our experiment approach and explain the main model components and their underlying prediction and trading algorithms. In Section 4.3.2, we provide more detail on our proposed enhanced algorithm. This is followed by our experiment approach and considerations, discussed in Section 4.3.3. In Section 4.3.4, we present and discuss our results. Section 4.3.5 concludes.

4.3.1 Method

In this experiment, we re-use the same setup defined in Part 1 (see Figure 4.1). Again, the process starts from a given collection of N input-output data training pairs,

$$\left(\mathbf{x}^{(1)} : y^{(1)}\right), \left(\mathbf{x}^{(2)} : y^{(2)}\right), \dots, \left(\mathbf{x}^{(N)} : y^{(N)}\right)$$

where

$$\begin{aligned}
&\mathbf{x}^{(1)} = [s_{1,t-N}, s_{2,t-N}, s_{3,t-N}, s_{4,t-N}], && y^{(1)}_{t-N+1}\\
&\mathbf{x}^{(2)} = [s_{1,t-N+1}, s_{2,t-N+1}, s_{3,t-N+1}, s_{4,t-N+1}], && y^{(2)}_{t-N+2}\\
&\dots\\
&\mathbf{x}^{(N)} = [s_{1,t-1}, s_{2,t-1}, s_{3,t-1}, s_{4,t-1}], && y^{(N)}_{t},
\end{aligned} \tag{4.22}$$

where y_t is the average return over the next 5 minutes (defined in Equation (4.3)), and the input, $\mathbf{x}$, is defined as a set of moving averages and RSI trading rule signals $\{s_1, s_2, s_3, s_4\}$ (defined in Equation (4.5)).

A standard NN is again employed to forecast, every 1-minute time epoch, the average return over the next 5 minutes. The same calibration approach is used, using the Sharpe ratio as the global risk-adjusted objective function.

Every minute, the NN layer passes a signal to the trading algorithm (further detail in Section 4.2.3). The challenge is to identify a threshold that is high enough that can filter out small price movements, mostly as a result of microstructure effects (see McAleer and Medeiros, 2008), however not being so high that it eliminates trading opportunities. From further investigation on a set of stocks we analysed the drop in the number of trades as the threshold was increased (see Figure 4.6). We decided to improve on the standard trading algorithm described in Part 1 (Section 4.2.3) and introduce more flexibility by permitting 3

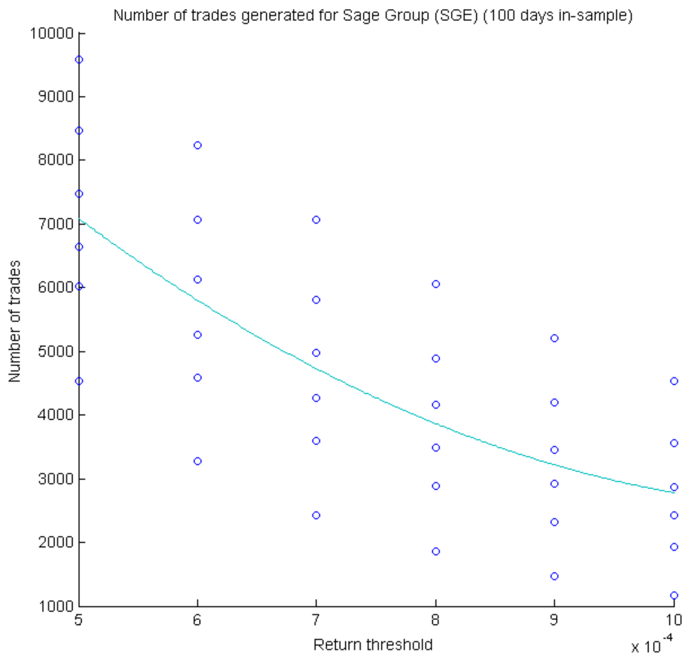


Fig. 4.6: Effect on the number of trades as the trading algorithm is tested with various levels of return threshold.

levels of RTs. The selection of RT is included as part of the parameter calibration process for each stock. Since models are trained daily on a moving window approach, having flexible return threshold also permits each model to adjust to the general volatility levels that are being experienced by the underlying stock across different time periods.

Since calibration is based on optimising the Sharpe ratio (using an annual risk-free rate of 5%), in our preliminary testing we identified that this can result in a substantial reduction in trades. Hence an additional constraint is introduced that sets a limit on the lowest number of trades for each stock during the calibration process. Based on the fact that in our experiment scenarios we start with an initial capital of 250,000 GBP and a standard position size of 50,000 GBP, we select a threshold that generates 20 daily trades (minimum) to ensure a wider intraday trading coverage. Hence, in conjunction with the NN parameters, we now consider $3 \times 3 \times 3 = 27$ possible combinations of the parameter settings for our benchmark NN model (Table 4.5).

Table 4.5: Combined parameters tested for the NN model and trading algorithm

Parameter	Parameter Value Set
Training Data Size (price points)	$\{510, 1020, 2040\}$
Number of Hidden Units	$\{5, 10, 20\}$
Return Threshold (RT)	$\{0.05\%, 0.08\%, 0.1\%\}$
Training Epochs	$\{1, ..., 1000\}$

Algorithm 3 Pseudo code for the extended close algorithm

```
if tradeDuration > 5 minutes then
    if {(signal == trade direction)
        or (price ≠ trade entry price)} then
        State ← keep open
        tradeOpenTime ← get current time
    else
        State ← close
    end if
else
    tradeDuration ← get current time - tradeOpenTime
end if
```

Another constraint that we added to the standard trading algorithm used in Part 1 (Section 4.2.3) is that to simulate real-world intraday trading environments, where durations can be very short, we consider a default holding period of five minutes for each trade (see Algorithm 3). Closing positions quickly in intraday trading does not necessarily represent stop-losses but can outperform common Buy-and-Hold strategies (Brabazon and O'Neill, 2006). However, we suppress any fast exits when the price does not exhibit any movement from the opening price in the 5-minute time window.

Apart from the dynamic capital allocation enhancements introduced by NN-FMM, the same trading conditions are applied for both NN and NN-FMM models.

4.3.2 Enhancing money management decisions by considering time-varying intraday risk

We enhance our approach presented in Part 1 and address time-varying intraday volatility by proposing an automated trading algorithm that identifies preferable pockets of intraday risk-adjusted profitability at different trend and volatility levels with the goal to increase *both* our position frequency and size in successful regions and reduce these in regions that are likely to result in losses.

By following the approach adopted in Part 1 (Section 4.2.5), we transform our set of signals (Equation (4.22)) into a second data set $\psi = \{\mathbf{G}, \mathbf{i}\}$. Matrix **G** consists of data points $\mathbf{g} = \{y_{t+1}, RV_{t+1}\}$ from simulated NN trades performed in the last h days, where y_{t+1} is the expected return over the next 5 minutes and RV_{t+1} is the realised volatility at the time the trade was placed. Vector **i** consists of the corresponding trade profits, inclusive of transaction costs. This data set is used by the money management algorithm that is described in the next section.

Again we employ FCM clustering on the data points in **G** (see Section 4.2.5) in order to identify approximate (rather than hard bound) trend and volatility regional spaces which more closely reflect the common terms used by traders in practice, such as “strong positive trend” or “low volatility”, which are not precise in nature. This is contrary to a single *fixed* threshold approach commonly adopted for return or volatility (Brabazon and O’Neill, 2006; Holmberg et al., 2013; Kaastra and Boyd, 1996; Vanstone and Finnie, 2009, 2010). For example, a return filter of 0.1% should not automatically disqualify a trade signal of 0.0999%. We argue that the underlying price process is too complex to model using hard bound filters and hence we believe that, in line with the incompatibility principle of Zadeh (1973), reducing the complexity by using a fuzzy approach also reduces uncertainty.

In addition to the approach presented in Part 1, as suggested by Pal and Bezdek (1995), further exploration is conducted with regard to the fuzziness index that is used by the FCM

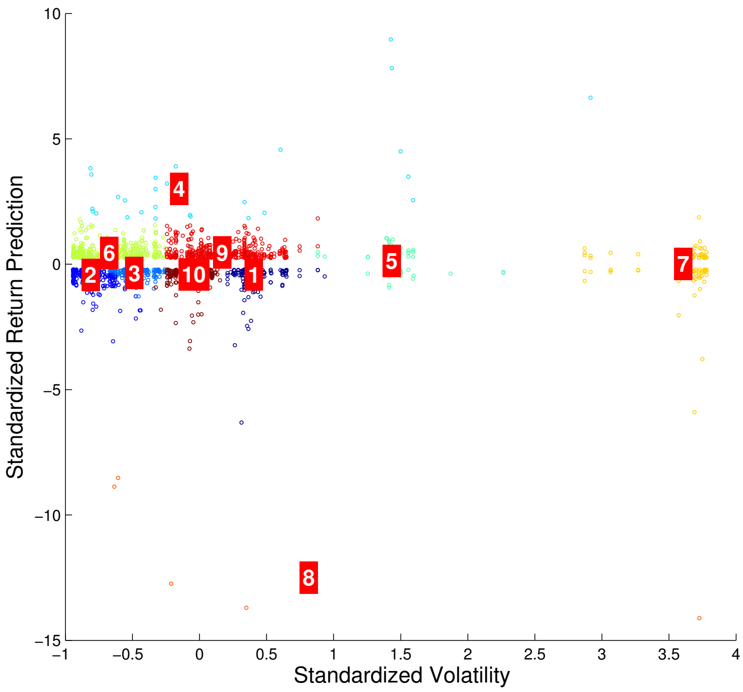


Fig. 4.7: Example of clusters found in Hammerson (HMSO) data as identified by FCM.

algorithm. Values between 1.5 and 2.5 are tested. After examining cluster plots (see for example Figure 4.7) on different stocks and time periods, a fuzzy index value of 1.7 is selected. The criteria are based on identifying a well-distributed set of clusters on the return and volatility space, whilst at the same time avoiding the heavy influence of possible outliers.

On a daily basis, a set of α rules are dynamically created using fuzzy sets derived directly from the components of the α centroids (see Figure 4.7) found by using clustering on the input space. The approximate local performance calculation in each cluster is influenced by the trade profits (from vector **i**) of the closest points (past trades defined by vector **g**) to each cluster centre rather than by a single crisp cluster boundary. This means that data points can be members of different clusters. The consolidated effect of the rules results in a decision surface which highlights regions of different Sharpe ratio performance across the trend-volatility space. The decision surface (e.g. Figure 4.8) shows smooth transitions between different performance regions and also the possibility of identifying multiple profitable

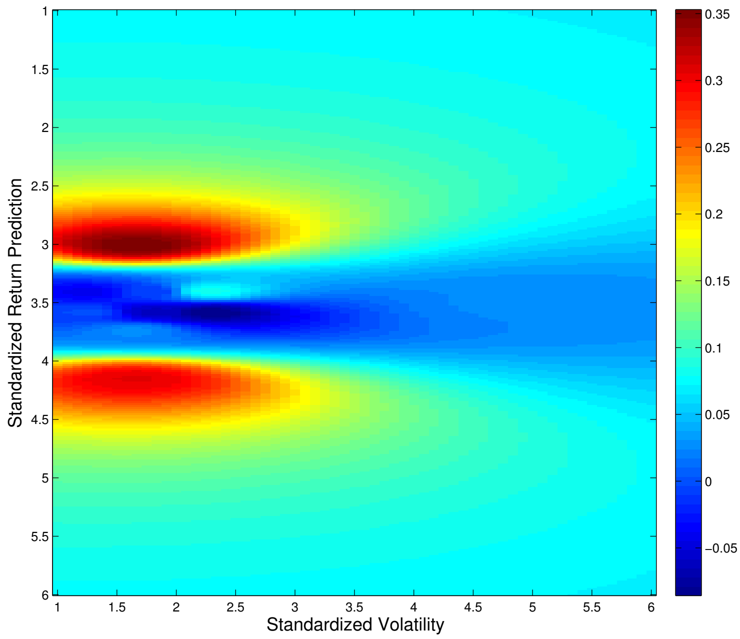


Fig. 4.8: Decision surface example for Hammerson (HMSO) as identified by our fuzzy controller showing different degrees of Sharpe ratio regions (see colour bar).

regions. We believe that this representation presents a more sensible and interpretable trading profile rather than by using a set of hard bound rules that would have resulted through the application of fixed return and volatility thresholds. This also alleviates common criticisms on the use of black box approaches for trading purposes.

Our proposed Neural Network – Fuzzy Money Management (NN-FMM) model enhances the standard trading rules by taking into consideration the approximate local regional performance on the space defined by our trend-volatility space. As noted in Part 1, Holmberg et al. (2013) show that increasing the return filter size results in a better success rate and average return, but possibly it also leads to a reduction in the number of trades, hence reducing the investors' potential profits. On these lines, our aim is to ensure that improving the risk-adjusted performance of the underlying trading model does not come at the cost of reduced profitability. As an improvement to our model to Part 1, this means that our money management approach should not simply act as a filter by only keeping the algorithm out of the market in adverse regions and hence possibly underutilise available capital, but also maximise model profitability by allocating more capital to preferable intraday trading scenarios (based on Sharpe ratio performance). Once we obtain the localised performance measure S, the position size ω_t is categorised as follows:

$$\omega_t = \begin{cases} large\ position: & S >= \lambda_1 \\ small\ position: & S >= \lambda_2, S < \lambda_1 \\ no\ position: & S < \lambda_2. \end{cases} \tag{4.23}$$

and

$$\begin{aligned} \lambda &= [0,1], \\ \lambda_1 &> \lambda_2 \ . \end{aligned} \tag{4.24}$$

To limit the number of possible combinations for λ values, in our experiment we opt for values of $\lambda_2 \in \{0, 0.1, 0.2, 0.3, 0.4\}$ and set $\lambda_1 = \lambda_2 + 0.3$. This effectively divides the stock risk-adjusted performance space, and our trade positions, into 3 categories. Contrary to our standard NN trading algorithm, which used equally sized positions of 50,000 GBP per trade, in our enhanced trading algorithm we open positions of 70,000 GBP for high-performance trend-volatility regions ($S > \lambda_1$), 30,000 GBP for medium ones ($S >= \lambda_2, S < \lambda_1$), and filter out trades in low (or negative) performing regions ($S < \lambda_2$). The remaining entry and exit conditions are applied as in the standard trading algorithm used in Part 1, including the same amount of initial capital of 250,000 GBP.

Three additional model parameters are included in this extended NN approach, which are α, the trade history data size, and λ_2 (see Table 4.6). Given this, *additional* $4 \times 4 \times 5 = 80$ permutations of the new parameters have to be compared in the model calibration. In the next section, we describe the experiment design process.

Table 4.6: Additional parameters tested for the NN and fuzzy volatility filter

Parameter	Parameter Value Set
Trade History Data Size (days)	$\{5, 20, 40, 60\}$
Number of Clusters (α)	$\{4, 6, 8, 10\}$
λ_2 threshold	$\{0, 0.1, 0.2, 0.3, 0.4\}$

4.3.3 Experiment approach and considerations

When designing our experiment approach, we paid particular attention to the harsh criticisms put forward by a number of authors (Bailey et al., 2014; Kearns et al., 2010; Lawrence et al., 1997; Pardo, 2011; Schulmeister, 2009; White, 2000) with respect to serious experimental flaws present in several published studies. In the context of our experiments, we have grouped and addressed these criticisms under five areas:

Experiment data

Research has shown that trend changes, which by nature occur swiftly, and large shifts in both volatility and liquidity, can have a large and often negative impact on trading performance. Of course, a good, robust model will be more capable of toughing out and trading profitably during such changes (Pardo, 2011). We again utilise our standard data set described in Chapter 3, Section 3.2.8. Similarly to Part 1, we utilise 1-minute trade records for the 27 stocks covering the same period between 28/06/2007 and 25/06/2008. The different trends in our data set ensure that we avoid possible bias in our experiment results which can occur by only picking stocks with a similar trend during the selected period.

Realistic constraints

Experiment findings in Chapter 3 show that overestimated profitability can be attained when non-realistic constraints are applied. In our experiments we account for realistic trading hours and no overnight positions. Since the London Stock Exchange operates between 8:00 and

16:30 GMT, this produced 510 1-minute prices per day. Similarly to previous experiments, our algorithm starts trading at 8:30am, with the latest possible trade placed at 16:00 and no overnight positions. In line with typical broker costs, in our experiments we apply a transaction cost of 10 GBP per trade per direction.

Model optimisation and performance measures

Many studies ignore the fact that the main interest of investors is risk-adjusted performance. Defining success solely on the grounds of forecast accuracy and win ratios has little practical value (Alves Portela Santos et al., 2007; de Faria et al., 2009; Enke and Thawornwong, 2005; Krollner et al., 2010; Medeiros et al., 2006). In our experiments, our primary interest is to optimise and then assess models according to their risk-adjusted performance. Following the Sharpe ratio improvements that we obtained using the NN fuzzy logic extension in Part 1 of this chapter, in Part 2 we also use this measure as our primary risk-adjusted performance measure and in addition we also include the Sortino ratio (described in Chapter 3, Section 3.2.7). In line with previous experiments, we consider an annual risk-free rate of 5%. We also introduce Calmar ratio as an additional risk-adjusted performance measure based on our interest to investigate whether discriminating between different trend-volatility states can also reduce drawdown risk. Unlike the Sortino ratio, which uses downside deviation as a proxy for risk (see Equation (3.22)), the Calmar ratio employs the maximum drawdown to penalise risk. This further strengthens our model ranking process in terms of risk-adjusted performance.

Model training, selection and validation process

Bailey et al. (2014) show that one can achieve almost any desired Sharpe ratio if one explores large parameter combinations or variations of a strategy, especially if backtesting is performed against an insufficiently large historical dataset. They further show that overfitted strategies are not only likely to disappoint, but, in the presence of memory (as real markets possess), they are actually prone to lose money. As an example, they show that if only five years of daily data are available, no more than forty-five independent model configurations should be tried or we are almost guaranteed to produce strategies with an annualised in-sample Sharpe ratio of 1 but an expected out-of-sample Sharpe ratio of zero.

To keep aligned with the approach adopted in previous experiments, we divide the 250-day data for each stock into 150 days in-sample and 100 days out-of-sample, consisting of 1-min price points. In the in-sample training and model selection process, for each parameter setting (see Tables 4.5 and 4.6), we roll forward the window for 150 days on a day-by-day basis, training the model on the previous days' data (defined as one of the parameters) and forecasting the current (unseen) day data (approach explained in Chapter 3, Section 3.2.7). Only positions in the same direction are allowed in order to avoid the spurious hedging effect of opposing positions that would overestimate the risk-adjusted performance of the algorithm. We select the specification with the highest Sharpe ratio obtained in the 150-day in-sample period. A positive result indicated that the model has surpassed the 5% annual risk-free rate, this acting as our lowest accepted benchmark. The selected model configuration was then tested out-of-sample for the next 100 days using the same rolling window approach.

Analysing model robustness and additional benchmark models

To avoid any positively biased conclusions on high-frequency trading returns based on possibly spurious results (as indicated by Schulmeister (2009) and Kearns et al. (2010)), we also run a second set of experiments to compare the risk-adjusted performance measures of 50 independent repeated trials for each stock for both NN and NN-FMM. In each trial we use a random initialisation for NN training in the case of the standard NN (see Lawrence et al., 1997; Luengo et al., 2009), and both neural network and fuzzy clustering in the case of NN-FMM. A total of 2700 models (50 models $\times$ 27 stocks $\times$ 2 AI models) are trained separately on the in-sample period, and then evaluated on the 100-day out-of-sample trading period, at the end of which we recorded the respective performance measures (obtaining $50 \times 27 \times 2$ sets of out-of-sample results). We also checked the distribution of the model results since Lawrence et al. (1997) stressed that one cannot assume that random NN initialisations always lead to a Gaussian distribution. This was done to validate our statistical tests as explained in Section 4.3.4 below.

Since in Part 1 we already tested the buy-and-hold model over the same period, in Part 2 we consider a random model as an additional benchmark model. In this case we implement 1350 models (50 random models $\times$ 27 stocks). The inclusion of this zero-intelligence benchmark model is to assess the effectiveness of our AI-controlled algorithmic trading strategies.

4.3.4 Results and evaluation

In this section, we first compare the results of NN and NN-FMM models following 100 out-of-sample trading days. Secondly we analyse the robustness of our models by investigating the results of 50 independent repeated trials for each stock covering the same out-of-sample period. This is supported by a number of statistical tests. Finally we benchmark the performance of our model against 50 repeated trials of the random model.

Risk-adjusted performance

In Table 4.7 we present the main results of our NN and NN-FMM models. In the case of NN, positive results are achieved in 23 out of 27 stocks (85%), the only negative results showing in AV, SDR, AZN and IAP. In general the model generates acceptable excess return results in the majority of stocks. Moreover, the mentioned 4 stocks which achieved negative results exhibit minor negative excess returns, the worse being SDR with -2.27% (considering that we are applying an annual risk-free rate of 5%). We note the improvement in the NN results when compared to the NN model that we present in Part 1 (Part 1 results available in Table 4.4). This is mainly attributed to the refinements that we include in the trading algorithm in Part 2, namely increased flexibility due to different return thresholds, a wider intraday trading coverage due to a minimum number of daily trades, and suppression of fast exits when price does not exhibit any price movement. Both NN and NN-FM utilise the same enhancements, hence any NN-FMM performance improvements are attributed to the fuzzy money management extension.

When we examine the results obtained by NN-FMM, we find that the model achieves positive results in 24 out of 27 stocks (89%), with the only negative results obtained in AV, HBOS and IAP. We also note that the instances where excess return results are negative, these are very minor with the only exception of IAP which shows a score of nearly 11%.

When comparing the results obtained by both models, we identify that the NN-FMM model exhibits a higher Sharpe ratio than NN in 21 out of 27 stocks (78%). In terms of excess return, NN-FMM shows a better score in 22 out of 27 stocks (81%). Although only indicative, we note that the average Sharpe ratio and excess return across the 27 stocks are higher for NN-FMM. This provides an initial indication of the trading optimisation gains offered by our risk-based money management fuzzy controller as an extension to NN design methods for trading purposes (as originally proposed by Kaastra and Boyd, 1996; Vanstone and Finnie, 2009, 2010).

The Sortino ratio results in Table A.6 show a similar outcome to Sharpe ratio. The NN-FMM model demonstrates a better score in 22 stocks, with the average across all stocks suggesting a clear improvement, on average, over the standard NN. From the same table we also see that in the majority of stocks, the NN-FMM model trades less. In conjunction with the above Sharpe ratio and excess return results this indicates that the NN-FMM model is indeed filtering out certain signal-volatility regions due to lower performance but allocating more capital in others.

Table 4.7: Trading performance of tested models following the 100-day out-of-sample period (bold font indicates the best result among the models for the specific stock).

	Sharpe Ratio		Excess Return	
Stock	**NN**	**NN-FMM**	**NN**	**NN-FMM**
AL	0.8727	**0.9126**	1.0119	**1.3135**
ANTO	0.2612	**0.3291**	0.1408	**0.2047**
BLT	0.2145	**0.2213**	**0.1018**	0.0760
BA	0.3240	**0.3686**	0.1581	**0.2079**
BLND	0.2292	**0.2714**	0.0660	**0.0826**
SKY	0.2717	**0.3371**	0.0765	**0.1032**
CWC	0.0026	**0.1128**	0.0012	**0.0429**
AV	-0.0718	**-0.0212**	-0.0170	**-0.0075**
DGE	**0.3564**	0.3041	0.0625	**0.0854**
SDR	-0.0312	**0.0560**	-0.0227	**0.0513**
HMSO	**0.4183**	0.3422	0.2339	**0.2562**
REX	**0.2352**	0.1680	**0.0836**	0.0606
JMAT	**0.0716**	0.0005	**0.0159**	0.0001
HSBA	0.1314	**0.1521**	0.0218	**0.0290**
ABF	0.1706	**0.2827**	0.0608	**0.0884**
INTU	0.6517	**0.6858**	0.2563	**0.3202**
RIO	0.0680	**0.0941**	0.0265	**0.0213**
BP	0.3586	**0.3679**	0.0930	**0.1363**
SGE	0.3235	**0.3494**	0.1057	**0.1346**
LLOY	0.1783	**0.2136**	0.0828	**0.1198**
TSCO	0.2585	**0.2680**	0.0893	**0.0986**
GSK	0.5202	**0.5727**	0.1395	**0.1807**
AZN	-0.0206	**0.0263**	-0.0063	**0.0092**
HBOS	**0.0332**	-0.0052	**0.0407**	-0.0063
XTA	0.0368	**0.0627**	0.0128	**0.0277**
IAP	**-0.0228**	-0.1682	**-0.0188**	-0.1084
ITV	0.4098	**0.6023**	0.2022	**0.4199**
Average	0.2315	**0.2558**	0.1118	**0.1462**

In the following section we further analyse and validate these initial results.

Model robustness

In this section we present a number of statistical tests to validate the hypothesis that the proposed NN-FMM optimisations significantly outperform the standard NN model. We do this by analysing the distribution of results following 50 independent trials of each model against each stock. Subsequently we compare the results against those obtained from 50 trials of a random model.

As a first step we analyse and compare the distribution of the obtained Sharpe ratio for NN and NN-FMM models. We run an Anderson-Darling (AD) test on the respective distributions for each stock and each model (see Table 4.8). As suggested by Lawrence et al. (1997), this is done to confirm a symmetric conversion to the mean, thus validating the approach of using a mean comparison. The AD test has rejected the null hypothesis that the Sharpe ratio distribution is from a normal distribution at a 5% significance level in only two instances —ANTO in the case of NN and LLOY in the case of NN-FMM. In the case of excess return, this is rejected in SDR and LLOY for NN, and only in GSK for NN-FMM.

For this reason we decide to apply both a parametric one-sided paired t-test and a non-parametric Wilcoxon rank-sum (RS) test to test the superiority of NN-FMM over the standard NN. When analysing the Sharpe ratio scores, we find that NN-FMM has a higher average score in 22 out of 27 stocks (81%). The tests indicate that the NN-FNN model is significantly superior in 18 out of 27 stocks (67%). In the case of RIO, the result was inconclusive since the tests indicated different outcomes. On the other hand, NN was significantly superior in only 4 stocks (HMSO, JMAT, HBOS, IAP). We consider the performance on 4 stocks (ANTO, DGE, SDR, BP) at par since the tests do not indicate any significant difference. From these results, we conclude that NN-FMM generally outperforms NN in terms of Sharpe ratio, and in few cases either performance is at par or else less than NN.

When we analyse excess return performance, we find that NN-FMM shows a higher average score in 24 stocks (89%). Tests indicate that NN-FMM is significantly superior in 17 stocks (63%). NN is significantly superior in only two stocks (HBOS and IAP). From these results we conclude that in general the additional optimisation resulting from our fuzzy money management module leads to improved excess return.

Table 4.8: Trading performance of neural network and NN-FMM models following 50 repeated trials over the 100-day out-of-sample period (bold font indicates the best result among the models for the specific stock). Each measure represents the average score from the 50 trials.

	Sharpe Ratio		Excess Return	
Stock	**NN**	**NN-FMM**	**NN**	**NN-FMM**
AL	0.8534	**0.9045**[bc]	0.8771	**0.9764**[bc]
ANTO	0.2783[a]	**0.2826**	0.1645	**0.1821**
BLT	0.1968	**0.2337**[bc]	0.0861	**0.1017**
BA	0.2283	**0.2839**[bc]	0.0952	**0.1415**[bc]
BLND	0.1986	**0.2209**[bc]	0.0654	**0.1031**[bc]
SKY	0.2109	**0.2875**[bc]	0.0693	**0.1248**[bc]
CWC	0.0107	**0.0840**[bc]	0.0008	**0.0521**[bc]
AV	-0.0247	**0.0148**[bc]	-0.0054	**0.0035**
DGE	0.2498	**0.2618**	0.0506	**0.0812**[bc]
SDR	-0.0209	**0.0009**	-0.0147[a]	**0.0412**[bc]
HMSO	**0.3851**	0.3550[bc]	0.2292	**0.2416**
REX	0.2562	**0.3980**[bc]	0.1076	**0.1775**[bc]
JMAT	**0.1020**	0.0672[bc]	**0.0325**	0.0201
HSBA	0.1143	**0.1675**[bc]	0.0302	**0.0410**
ABF	0.1967	**0.2712**[bc]	0.0618	**0.1256**[bc]
INTU	0.7326	**0.7783**[bc]	0.2854	**0.3445**[bc]
RIO	0.0227	**0.0407**[b]	0.0071	**0.0140**
BP	**0.4415**	0.4390	0.1408	**0.1627**[bc]
SGE	0.2401	**0.2747**[bc]	0.0739	**0.1004**[bc]
LLOY	0.1548	**0.1990**[abc]	0.0873[a]	**0.1212**[bc]
TSCO	0.2530	**0.3352**[bc]	0.0894	**0.1267**[bc]
GSK	0.4148	**0.4571**[bc]	0.1486	**0.1708**[abc]
AZN	-0.0310	**0.0319**[bc]	-0.0038	**0.0146**
HBOS	**0.0410**	-0.0076[bc]	**0.0587**	-0.0231[bc]
XTA	0.0310	**0.1005**[bc]	0.0198	**0.0554**[bc]
IAP	**-0.0378**	-0.1110 [bc]	**-0.0264**	-0.0724[bc]
ITV	0.3224	**0.6784**[bc]	0.2452	**0.3954**[bc]
Average	0.2156	**0.2611**	0.1099	**0.1412**

[a] Anderson-Darling test rejects the null hypothesis at the 5% sig. level that the distribution of specific measure is from a normal distribution.
[b] Two-sample t-test supports the alternative hypothesis at the 5% sig. level that the distribution of the specific measure for both models comes from populations with unequal means.
[c] Ranksum test supports the alternative hypothesis at the 5% sig. level that the distribution of the specific measure for both models comes from populations with unequal medians.

We also analyse addition risk-adjusted performance statistics displayed in Table A.6. In line with the Sharpe ratio results, the vast majority of the average scores for both the Sortino and Calmar ratio are in favour of NN-FMM. In the case of Sortino ratio, the NN model is significantly superior in only 2 stocks (HBOS and IAP), whilst in terms of Calmar ratio the NN model is significantly superior in only 3 stocks (HMSO, HBOS and IAP). When compared to previous results, one can notice that in the case of Sortino and Calmar ratio, the AD test indicated a higher number of occurrences of possible departure from normality in their distribution. However, we conclude that in the case of Sortino and Calmar ratio, similar to the Sharpe ratio results, the NN-FMM model, in general, shows a clear risk-adjusted performance superiority in the majority of stocks.

As our final benchmark, we deploy a random model which generates random trades (buy or sell) however keeping a similar intraday trading pattern, trade frequency and trade duration of the corresponding AI driven model for the respective stock. In line with the AI models in this chapter, we consider trading on a minute-by-minute basis. Additionally, similar to the approach adopted earlier for the AI models, the random model is executed independently 50 times for each stock for the same 100-day out-of-sample period.

First we analyse the pattern of the trades executed by our standard NN model. We note that in general all stocks (a subset is displayed in Figure 4.9) exhibit a similar pattern with the highest activity in general displayed during the first and last couple of hours and the lowest activity close to mid-day. Due to the increased number of trades and shorter trade durations when compared to our models in Chapter 3 (see Figure 3.5, the activity of our trading model more clearly depicts an activity pattern which follows the typical U-shaped day trading activity pattern established in literature (see Andersen and Bollerslev, 1997). This indicates that our model optimisation selects a signal threshold which automatically filters out periods of lower activity.

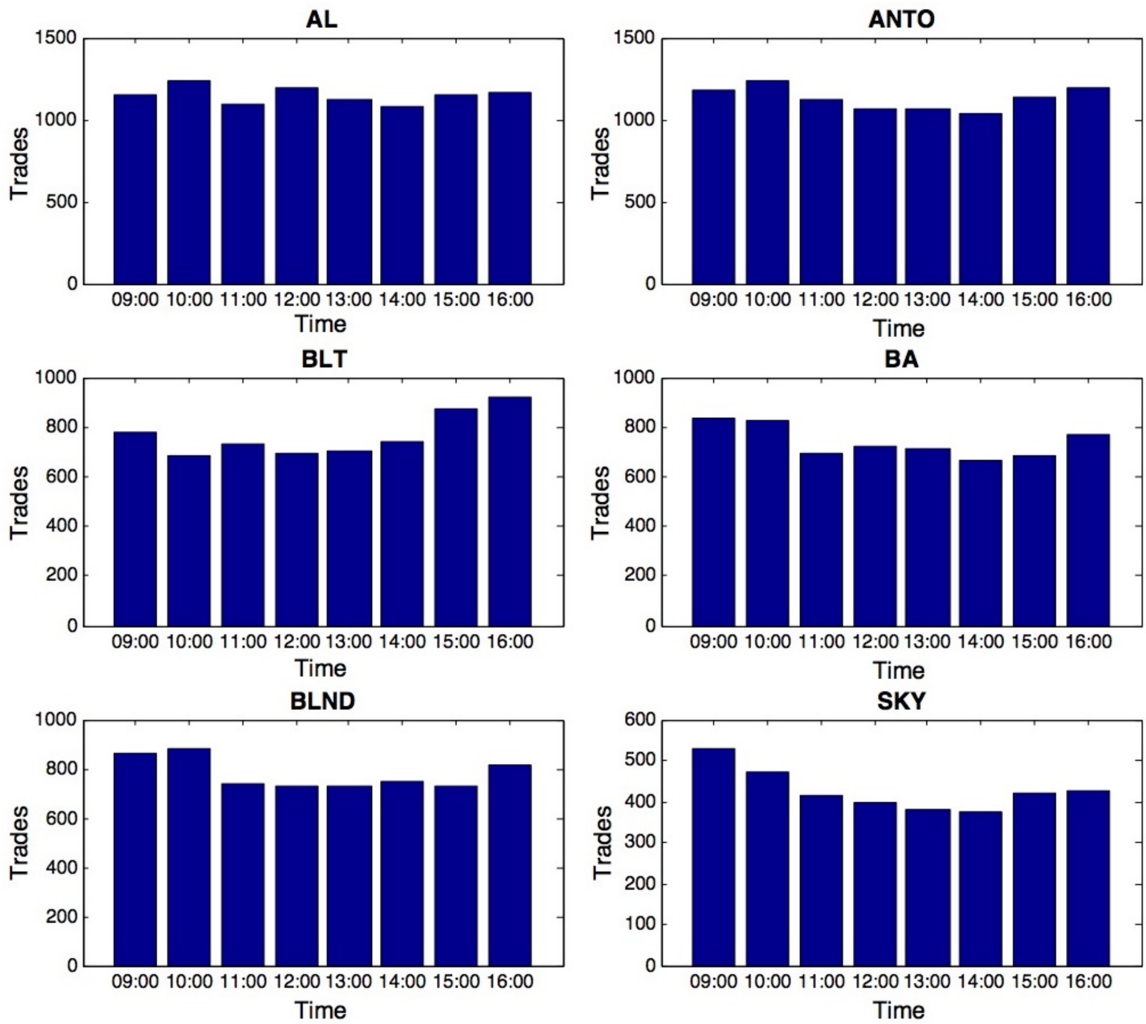


Fig. 4.9: Trading pattern of our trading algorithm on a set of stocks using the standard NN model. Each bar represents the average number of trades performed in the hour leading to the respective time slot (x-axis).

We then investigate the trading pattern of our NN-FMM model. As indicated in earlier results, our fuzzy filter results in lower number of trades, albeit more capital is allocated to preferred signal-volatility regions. We investigate the difference between the number of trades in our proposed NN-FNN model and the corresponding NN model for each stock. In Figure 4.10 we display the average difference in the number of trades between the two models across all stocks. The plot indicates that on average the highest trade reductions happen between 12:00 and 14:00, ranging approximately between 5% and 7.5%. This has the effect of further accentuating the U-shaped trading pattern of the standard NN model. We conclude that our fuzzy filter automatically gives preference to periods of higher activity due to the higher chances of generating enough price movement which can in turn generate trade profits.

Finally, based on this trading pattern information, we design a random model and adopt similar hourly trade ratios for the specific stocks. Using the same approach adopted in Chapter 3, we control this by generating uniformly distributed random numbers between -1 and 1. We apply an equal threshold on both the negative and positive side of the specified range. If the random number is smaller or greater than the threshold then a sell or buy signal is executed accordingly. Increasing (decreasing) the threshold tends to reduce (increase) the number of random trades in the specific time slot.

From the results, displayed in Table 4.9, we identify that the random model produces positive results in 7 stocks (taking into consideration that we are applying an annual risk-free rate of 5%). When we compare these results against those produced by the AI-models for the same number of trials, it is evident that our applied AI-models are indeed effective in generating valid trade signals.

The above results lead us to conclude in favour of the main objectives that we aim to validate in this chapter. Firstly, we find a significant link between intraday trading profitability and two interrelated dimensions, which are NN return prediction strength (profit) and realised volatility (risk), extending the findings in the literature (Gradojevic and Gençay, 2013; Han et al., 2013; Holmberg et al., 2013; LeBaron, 1999; Schulmeister, 2006) at a more granular intraday level. Our results show that our proposed fuzzy logic method to dynamically adjust trade position size to different degrees of risk is an effective money management approach which improves the risk-adjusted performance of intraday trading models. Secondly, our approach does not only act as a filter by merely keeping the algorithm out of the market in adverse regions, and hence under-utilising available capital, but also

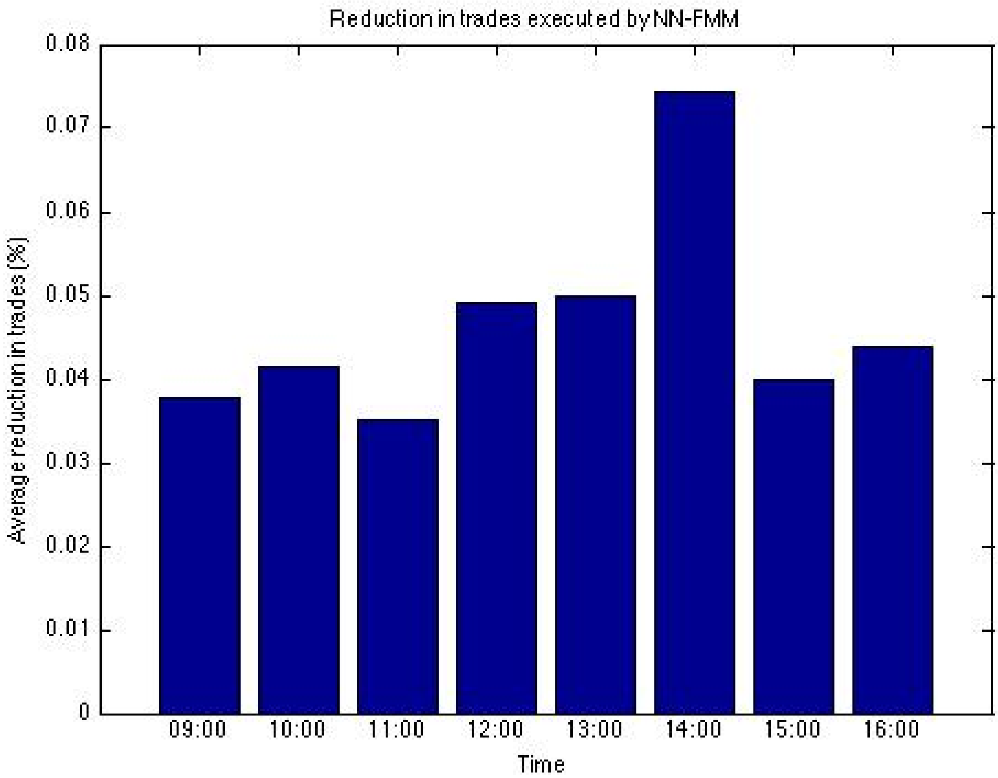


Fig. 4.10: The plot indicates the average reduction in trades placed by NN-FMM when compared to the standard NN model. Each bar represents the average reduction in number of trades performed in the hour leading to the respective time slot (x-axis).

improves model profitability by allocating more capital to preferable intraday states. This also validates the claims of Pardo (2011) in view of the disregarded effect of risk-based position sizing methods and should raise interest in the use of AI methods to address a wider set of trading decisions rather than limiting research primarily (if not solely) to market movement predictions (Cavalcante et al., 2016; Krollner et al., 2010; Son et al., 2012; Tsai and Wang, 2009; Vanstone and Tan, 2003).

Table 4.9: Trading results following 50 independent trials of the random model. Each measure represents the average over the 50 trials.

	Random Model	
Stock	**Sharpe Ratio**	**Excess Return**
AL	-0.0062	-0.0342
ANTO	-0.0415	-0.0484
BLT	-0.0581	-0.0571
BA	-0.0941	-0.1189
BLND	0.0052	0.0045
SKY	0.0389	0.0195
CWC	-0.0172	-0.0141
AV	-0.0396	-0.0376
DGE	0.0062	0.0007
SDR	-0.0738	0.0803
HMSO	0.0832	0.0626
REX	-0.0876	-0.0669
JMAT	-0.0489	-0.0300
HSBA	0.0055	0.0048
ABF	-0.0207	-0.0123
INTU	0.0729	0.0445
RIO	-0.0632	-0.0531
BP	0.0783	0.0426
SGE	-0.0431	-0.0386
LLOY	-0.0702	-0.0724
TSCO	-0.0981	-0.0646
GSK	-0.0307	-0.0226
AZN	-0.0926	-0.0692
HBOS	-0.0537	-0.1024
XTA	-0.0616	-0.0558
IAP	-0.0263	-0.0348
ITV	-0.0123	-0.0217
Average	-0.0278	-0.0257

4.3.5 Conclusion

In Part 2 of this chapter, we seek further model optimisations from the identified relationship between the intraday profitability of technical trading rules and volatility that was initially explored in Part 1. The research problem focuses on optimising capital allocation for intraday trading using fuzzy logic money management decisions. The problem is highlighted by

Pardo (2011) who claim that models that do not take into consideration effective money management possibly lead to sub-optimal solutions. In line with Zadeh's principle of compatibility, we resort to the fact that preciseness remains much less applicable in complex trading environments and hence we propose a novel fuzzy logic approach to control risk-based trading decisions.

As our first contribution from Part 2, we introduce an innovative fuzzy money management approach which dynamically adapts trading frequency and position sizing decisions across intraday trend (profit) and volatility (risk) states to improve the overall risk-adjusted performance. Contrary to many studies that suggest trading rules based on a fixed position sizing strategy (as reported by Pardo, 2011), fixed return thresholds (Kaastra and Boyd, 1996; Vanstone and Finnie, 2009, 2010) and fixed volatility thresholds (Brabazon and O'Neill, 2006; Holmberg et al., 2013), our approach dynamically evolves a continuous trading decision surface across the whole intraday trend-volatility space. We demonstrate the applicability of our fuzzy logic approach by presenting a hybrid fuzzy model as an extension to a popular neural network trading model (Choudhry et al., 2012; Krollner et al., 2010; Tsai and Wang, 2009). Our hybrid model calculates the approximate risk-adjusted performance across different trend-volatility states and automatically balances capital allocation according to preferable trend and volatility scenarios during the course of a trading day. Our results show evidence of superior risk-adjusted performance when compared to the popular NN model and also a random model. In comparison to our fixed position trading model in Part 1, we also demonstrate how our fuzzy logic money management extension leads to a more optimal solution, increasing the risk-adjusted performance without incurring a reduction in profitability. This is supported by our results which show significant improvements when compared to standard NN and random model methods. We further validate our results with a robust statistical analysis on a number of independent trials in order to address the numerous criticisms in the literature (see Bailey et al., 2014; Lawrence et al., 1997; Luengo et al., 2009;

Prechelt, 1996) that cast doubts on the experimental and statistical frameworks adopted by many published studies in computational finance literature.

As our second contribution, contrary to the EMH, our findings extend the support for studies that claim a relationship between the profitability of technical trading rules and volatility (Gradojevic and Gençay, 2013; Han et al., 2013; Holmberg et al., 2013; LeBaron, 1999; Schulmeister, 2006). The uniqueness of our work stems from the fact of exploring this link at a more granular intraday level and at higher trading frequencies with the objective to improve risk-adjusted performance. To our best knowledge this was not previously addressed in the literature. Our results indicate that by discriminating across different intraday trend-volatility states, trading algorithms can benefit from increased profitability and risk-adjusted performance. From a theoretical perspective this sheds more light on the possible breakdown of the martingale property of prices under certain market conditions.

Following the demonstrated advantages of applying global risk-adjusted objective functions at a daily level (Chapter 3), in this chapter we took a more granular perspective and presented innovative ways of how fuzzy logic can further improve trading algorithms by discriminating across time-varying intraday volatility. In the next chapter we endeavour to explore further improvements by investigating how type-2 fuzzy logic can improve technical signals on high frequency data.

Chapter 5

Technical rules and uncertainty at higher trading frequencies

In this chapter, our aim is to investigate the profitability and risk-adjusted performance of trading algorithms from a higher frequency perspective than that presented in Chapters 3 and 4. The challenge with technical indicator signals is to identify price trends, however at the same time avoid any ‘whiplash’ trades, which arise when the indicator generates too many trading signals in a short period of time. The occurrence of these problematic signals is more pronounced when a security price moves sideways instead of moving in clear trends, with the possibility of falsely indicating trend formations. The uncertainty is heightened in intraday trading since by nature the trading opportunities that one is after are typically of shorter duration and higher frequency with the possibility of being exposed to more noise in the underlying data. In this chapter, our interest is to introduce type-2 (T2) fuzzy logic in combination with technical rules and identify whether the move from type-1 (T1) to T2 can effectively improve the management of technical trading uncertainty. Our main aim is to identify whether this transposes into enhanced risk-adjusted performance.

This chapter extends our research findings that are presented in previous chapters from two aspects. Firstly, in Chapter 3 we show the effectiveness of ANFIS and also identify the increased stability of ANFIS in terms of risk-adjusted performance when compared to ANN alone. In this chapter we improve on these findings by investigating the possible refinements that can be obtained by generalising ANFIS to interval T2 (IT2) FLS. Secondly, in Chapter 4 we enhance model risk-adjusted performance by applying FCM clustering to discriminate between different intraday volatility scenarios. In this chapter we use a similar data partitioning approach but build on these findings by applying FCM clustering to identify the T1 fuzzy logic rules which are then extended to IT2. This not only results in a more compact and efficient model but also in increased risk-adjusted performance.

We propose an innovative approach to design an IT2 model which is based on a generalisation of the popular T1 ANFIS model. The significance of this work lies in the identification of risk-adjusted performance improvements that are obtained as a result of introducing T2 fuzzy sets in intelligent trading algorithms. This is achieved with a minimal increase in the design and computational complexity of the underlying T1 models. Overall, the proposed ANFIS/T2 model scores significant performance improvements when compared to standard ANFIS. As a further step, we identify a relationship between the increased trading performance benefits of the proposed T2 model and higher levels of microstructure noise. The results satisfy a desirable need for practitioners, researchers and regulators in the design of expert and intelligent systems for the better management of risk in the field of HFT.

5.1 Introduction

Most transactions in modern, highly computerised, financial markets are being largely controlled by specialised algorithms which incessantly sift through masses of data and take split-second trading decisions. According to a recent study by Brogaard et al. (2014), between 2008 and 2010 HFT algorithms accounted for 70% of dollar volume on the NASDAQ exchange. This tends to defy the long-standing Efficient Market Hypothesis (EMH) (Fama, 1965, 1970) that states that current prices incorporate all relevant information with no possibility of predictability or excess returns. A number of authors (e.g. Brogaard et al., 2014; Holmberg et al., 2013; Rechenthin and Street, 2013; Schulmeister, 2009; Zhang, 2010) insist that the presence of efficient pricing becomes more questionable when investigating short-lived (milliseconds to a few minutes) trades. However, Kearns et al. (2010) validate the EMH in their study and argue that generating profits from aggressive HFT is next to impossible. These debates keep this domain a very active area of research.

According to Johnson et al. (2013), this new machine-dominated reality highlights the need for new theories in support of sub-second financial phenomena during which the human traders lose the ability to react in real time. Due to the non-stationary characteristics of financial time series (see Fama, 1965), applying machine learning techniques to infer predictions is a challenging task and prone to increased error variance. Complexity is heightened given the level of noise in high-frequency stock price movements. Incidents like the "flash crash" of May 2010 stress the importance of risk management. As a result, in recent years HFT and algorithmic trading have been the subject of increasing global regulatory attention. The new regulations are intended to ensure that trading systems are adequately designed and tested to mitigate the risks to which they are exposed (see reference to MiFID II regulation in Chapter 1, Section 1.1.3, which will apply from January 2018).

In line with the main theme adopted in previous chapters, we stress again that this acts as a reminder for model or algorithm designers that both investors and regulators are more concerned with risk-adjusted performance rather than directional accuracies, error measures or solely profitability. Unfortunately, in previous chapters we highlight that the great majority of computational finance research disregards this fact (Bahrammirzaee, 2010; Cavalcante et al., 2016; Krollner et al., 2010; Tsai and Wang, 2009). Using a risk-adjusted performance measure is essential in order to compare relative trading performance. A higher return trading strategy does not necessarily outperform other strategies if the associated risk is also substantially higher. This is the reason why in this chapter we continue to seek further improvements in risk-adjusted performance.

Gradojevic and Gençay (2013) show that common technical indicators, such as moving averages, generate information that is possibly imprecise, incomplete or unreliable. This results in possibly indicating turning points which are out of synch with the underlying price movements, hence resulting in missed opportunities, or worse, substantial trading losses. The challenge accumulates with the sampling frequency and hence uncertainty management becomes of utmost importance. Gradojevic and Gençay (2013) demonstrate that type-1 (T1) fuzzy logic in conjunction with technical rules can offer better results than standard technical rules due to better management of uncertainty. In this chapter we try to extend this finding by investigating the additional improvements that can be gained using the more advanced type-2 (T2) fuzzy logic.

As we note in Chapter 3, one of the most popular combinations of ANN and T1 fuzzy logic is the Adaptive Neuro-Fuzzy Inference System (ANFIS) (Jang, 1993). The successful application and active continuous research dedicated to improving ANFIS-based techniques in trading applications is demonstrated by numerous publications (Boyacioglu and Avci, 2010; Chang et al., 2011; Chen, 2013; Gradojevic, 2007; Kablan and Ng, 2011; Tan et al., 2011; Wei et al., 2014). In Chapter 3 we extend this literature by presenting an ANFIS approach which can improve risk-adjusted performance and stability. However, T2 fuzzy logic has recently gained significant academic attention (see review in Melin and Castillo, 2014) and as of today it remains a primary area of research in the fuzzy logic domain (Mendel et al., 2014). To our best knowledge, the use of higher order fuzzy logic systems (FLSs) in a high-frequency trading environment has not been addressed in the literature before. This presents an opportunity, which we tackle in this chapter, to investigate further improvements for the standard ANFIS model by extending it to T2 fuzzy logic. However, in line with Wu and Mendel (2014), we argue that although T2 FLSs provide the researcher with extensive freedom in design options, the increased computational and design complexity can possibly hinder the wider application of such systems. This challenge is a key consideration that inspired our innovative and practical T2 approach that we present in this chapter.

The investigation of possible improvements using T2 in HFT is appealing in view of increased uncertainties which are inherent in high-frequency data. Although the concepts of risk and uncertainty have often been used interchangeably, economists have long distinguished between the two (e.g. Knight, 1921), also in recent literature (e.g. Heal and Millner, 2014; Nelson and Katzenstein, 2014). Our view is that overall risk can be divided into measurable risk (e.g. the flip of a fair coin), and uncertainty, which we categorise as the risk of events to which it is difficult to attach a probability distribution. Our aim is to further reduce the trading uncertainty by utilising T2 FLSs. We have not identified any existing literature that investigates the level of noise (uncertainty), indirectly reflecting the trading frequency, that would warrant the (feasible) use of T2 over T1 fuzzy logic methods for algorithmic trading purposes.

With respect to the above literature review and identified gaps, in this chapter the objectives can be summarised as follows:

1. To identify practical methods of how the popular ANFIS model, the stability strengths of which we show in Chapter 3, can be generalised to an interval T2 Takagi-Sugeno-Kang (IT2 TSK) fuzzy system. We aim to address this with a minimal increase in design and computational complexity.

2. To investigate the ability of higher order fuzzy systems to handle the increased uncertainties inherent in an HFT scenario. In this chapter our interest is to investigate higher trading frequencies than those we present in Chapters 3 and 4.

3. To identify *if* T2 FLSs provide a viable alternative for trading purposes in view of improving risk-adjusted performance.

4. To explore *when* (in what trading scenarios) T2 models can offer a more viable approach than T1 alternatives. Our interest is to analyse this from the perspective of different levels of trading frequencies (uncertainty).

A criticism that is sometimes attributed to IT2 models is that there are many, sometimes an overwhelming amount (Wu and Mendel, 2014), of design choices to be made. These include the shape of membership functions, the number of membership functions, the type of fuzzifier, the kind of rules, the type of *i*-norm, the method to compute the output, and the methods for tuning the parameters.

As a first contribution we propose an innovative, but at the same time a more accessible, way of how to design a T2 FLS from an optimised T1 neuro-fuzzy FLS (ANFIS/T2). We address this from a number of aspects. Firstly, we improve on the ideas in Chapter 4 by applying a fuzzy clustering algorithm to identify different trend and volatility data segments and use the clustering results to construct the fuzzy rule database. This approach reduces the number of rules and hence simplifies the final model. We apply simple rules where antecedents are T2 fuzzy sets and consequents are crisp numbers (A2-C0). Secondly, as our base structure for the T2 model we use the ANFIS model since, as we identified in Chapter 3, it acts as a solid benchmark model, is computationally fast and has also been successfully applied in high-frequency trading (see also Kablan and Ng, 2011). Thirdly, we reduce the training complexity by reducing the number of tuning parameters, limiting this to varying sizes of the Footprint of Uncertainty (FOU). Our parsimonious approach also reduces the possibility of overfitting and spurious results (see Bailey et al., 2014). Finally, we apply an efficient closed-form type reduction method.

Schulmeister (2009) points towards possible market inefficiencies and the profitability of technical trading rules at higher frequencies, this being driven by faster algorithmic trading. Recently, Rechenthin and Street (2013) claimed that when price shocks break the bid-ask spread, which was identified to happen anywhere between 5 to 10 seconds, price movements can be predicted for up to one minute. Beyond this point, prediction probabilities remained significant for about the next 5 minutes, dying out completely beyond 30 minutes.

As our second contribution, we shed more light on the theoretical market efficiency debate as regards HFT. More specifically, we are not aware of any previous studies which investigate the link between higher order fuzzy systems and risk-adjusted performance in trading windows of between 2 to 10 minutes. In our case, we make use of HFT trade data from a set of stocks listed on the London Stock Exchange and investigate a combination of technical rules on 2-minute returns with holding periods ranging between 2 to 10 minutes. This extends the market efficiency investigations that we present in Chapters 3 and 4 to a more microscopic perspective and hence we aim to seek possible pockets of profitability in shorter time windows. In line with previous chapters, we align ourselves with the priorities of investors and regulators and focus on comparing the proposed models using risk-adjusted performance (Choey and Weigend, 1997; Vanstone and Finnie, 2010; Xufre Casqueiro and Rodrigues, 2006). We conclude that pockets of predictability are possible in the investigated time horizon and that T2 fuzzy logic can contribute in reducing the negative effects emanating from microstructure noise.

Although previous literature has found that T2 models can perform better under increasing uncertainties (Aladi et al., 2014; Sepulveda et al., 2006), it is however not clear at which uncertainty level this would be reflected in a reasonable improvement in risk-adjusted trading performance. Birkin and Garibaldi (2009) have even shown that if the level of noise is too low, T2 models show no significant improvement on T1. A number of authors (Gençay, 1996; Holmberg et al., 2013; Vanstone and Finnie, 2009, 2010) suggest the use of a threshold on the predicted signals below which a trading action is not taken into consideration. This is done to reduce the effect of the underlying noise; however, this comes at the cost of reduced trades and hence possible return.

By extending the concept of the threshold technique, as our third contribution we aim to answer an important question which explores, from a trading performance perspective, *when* it is *viable* to apply T2 rather than T1 models. We propose an innovative experiment approach by extending the technique to analyse how T1 and T2 models cope with decreasing (increasing) levels of return thresholds, which are reflected in an increase (reduction) in uncertainty but also in increased (reduced) return potential. This ability to handle higher uncertainty is fundamental for HFT. Our evaluation of *out*-of-sample data demonstrates that the proposed ANFIS/T2 model outperforms the standard ANFIS and a standard technical approach using optimised technical indicators. Statistically significant improvements in trading performace are registered in higher trading frequency scenarios but disappear when trading activity is lowered.

The structure of this chapter is organised as follows. In Section 5.2, we introduce the main model components and design method. This is followed by a description of our experiment approach and model evaluation presented in Section 5.3. In Section 5.4, we present our results and analyse model performance. Finally, in Section 5.5 we draw our conclusions in view of the existing literature.

5.2 Method

Our experiment setup consists of five modules (see Figure 5.1). Section 5.2.1 explains our variable selection and data pre-processing. Sections 5.2.2 and 5.2.3 explain our Fuzzy Inference System (FIS) design approach using clustering which is later fed into ANFIS and ANFIS/T2 for tuning (this section presents our main model enhancements on previous chapters). In Section 5.2.4 we explain our trading algorithm.

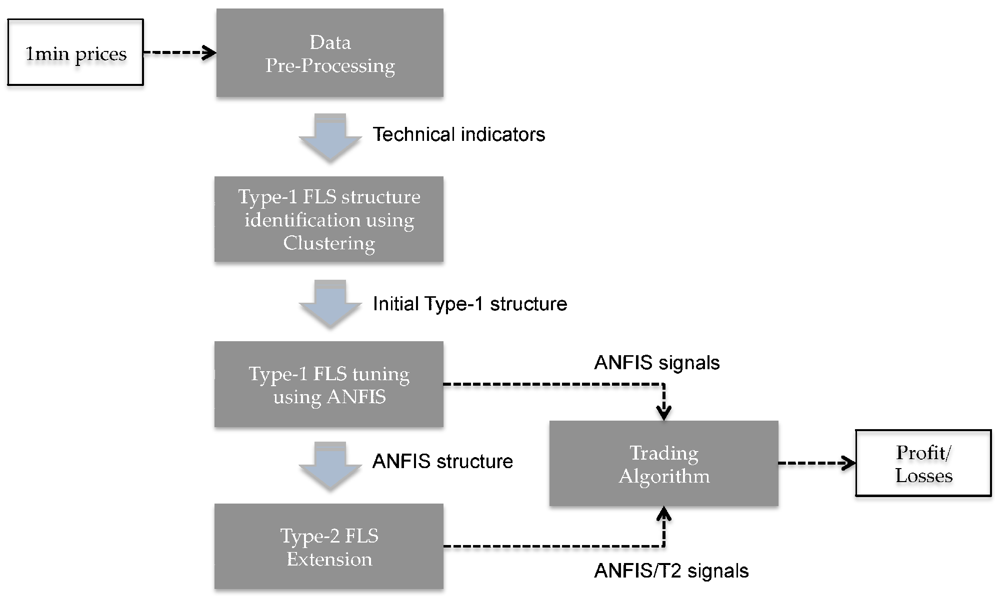


Fig. 5.1: Experiment setup showing our proposed T2 extension.

5.2.1 High-frequency data and technical indicators

In this section we explore evidence presented by a number of authors who claim that in HFT scenarios there exist short time windows where past prices can convey information which can be used for predictive purposes. Our interest is not to identify the determining factors of this claimed HFT phenomenon, but to identify candidate features that can be used by our trading models. We do however reduce our window of interest to a more granular (higher frequency) price structure than that we present in Chapters 3 and 4.

We investigate whether the average return for the next 2 minutes can be successfully estimated using 5 signals which can provide information on price trend, reversion and movement strength from a time window of previous prices ranging from 1 to 15 minutes. The set of signals that we select for this chapter is based on findings by Brogaard et al. (2014) and Zhang (2010), who identified that the main determinants of current HFT activity are past returns, liquidity, and HFT activity. The time-window selection is based on claims from Rechenthin and Street (2013), who stated that the stock price typically broke the price reversal pattern due to the bid-ask bounce after 5 to 10 seconds, and that traces of predictability existed for up to 30 minutes, beyond which markets became efficient.

Our findings in Chapters 3 and 4 (see also Gradojevic and Gençay, 2013; Naranjo et al., 2015) show the effectiveness of combining moving average signals with fuzzy logic to capture trend information. In this chapter we use 1-minute stock prices, p_t, and define the expected mean return, y, at time t, as

$$y_t = log(MA^2_{t+2}) - log(p_t), \tag{5.1}$$

where short and long moving averages, MA^n_t, are calculated as defined by Equation (3.1). We apply corresponding trade signals, MA^{n_1,n_2}_t, defined by Equation (3.3) with the lag structures $(n_1,n_2) \in [(1,2),(1,5),(1,10)]$ where n_1 and n_2 are expressed in 1-minute time bars. Compared to Chapters 3 and 4, here we apply shorter lag structures in order to generate a higher number of signals and hence increase our potential trade frequency.

In line with Chapter 4, for our mean-reversion indicator we use the popular Relative Strength Index (RSI) (Murphy, 1987). However, since in this chapter we wish to investigate higher frequency trading, to calculate RSI we consider 1-minute prices in the previous 15 minutes.

In Chapter 4 we identify model improvements by discriminating between different levels of intraday volatility. For this reason, here we utilise Realised Volatility (RV) as our fifth input variable. This presents a more efficient and compact approach to discriminate between intraday volatility levels by including it as part of the main prediction model (rather than at a subsequent second step). Similarly to Chapter 4 (see Section 4.2.4), we again utilise the average RV estimator. However, in this case we utilise the last 15 minutes average RV (rather than 30 minutes) at time t using 5-minute return intervals. This is done primarily due to our interest in higher frequency trading at shorter time intervals.

In summary, the identified k variables yield a linear regression model to describe the relationship with y_t as

$$y_t = \theta_0 + \sum_{k=1}^{5} \theta_k s_{k,t-1} + \varepsilon_t \tag{5.2}$$

with the error term $\varepsilon_t \sim N(0,\rho)$ and

$$s_{k,t} = \begin{cases} MA_t^{1,2} & \text{for} \quad k=1 \\ MA_t^{1,5} & \text{for} \quad k=2 \\ MA_t^{1,10} & \text{for} \quad k=3 \\ RSI_t & \text{for} \quad k=4 \\ RV_t^{avg} & \text{for} \quad k=5 \end{cases} . \tag{5.3}$$

5.2.2 Designing and tuning a TSK type-1 fuzzy model

A first consideration is to select the type of FLS to employ. The literature identifies two main types, namely Mamdani (Mamdani, 1974), where the rule consequents are fuzzy sets on the output space, and TSK (Sugeno and Kang, 1988), where the rule consequents are crisp functions of the inputs. We adopt the TSK approach due to their popularity in practice resulting from their simplicity and flexibility (Wu and Mendel, 2014). Moreover, for TSK rules, output calculation is less computationally intensive: the output is a weighted average of the crisp rule consequents, where the weights are the firing levels of the rules (the preliminary mathematical background is presented in Section 2.3).

In the following two sub-sections, we describe our approach for (i) the initial identification of a T1 TSK FIS model (Section 5.2.2) and (ii) model tuning (Section 5.2.2).

Initial FIS structure identification

We follow a model-free approach (Mendel et al., 2014) with the objective to completely specify the FLS using training data. The process starts from a given collection of q minute-by-minute input-output data training pairs, $(\mathbf{x}^{(1)} : y^{(1)}), (\mathbf{x}^{(2)} : y^{(2)}), ..., (\mathbf{x}^{(q)} : y^{(q)})$ where

$$\begin{aligned}
&\mathbf{x}^{(1)} = [s_{1,t-q}, s_{2,t-q}, \ldots, s_{5,t-q}], && y^{(1)}_{t-q+1} \\
&\mathbf{x}^{(2)} = [s_{1,t-q+1}, s_{2,t-q+1}, \ldots, s_{5,t-q+1}], && y^{(2)}_{t-q+2} \\
&\ldots \\
&\mathbf{x}^{(q)} = [s_{1,t-1}, s_{2,t-1}, \ldots, s_{5,t-1}], && y^{(q)}_{t}.
\end{aligned} \tag{5.4}$$

In Equation (5.4), for each data instance at a specific time t, $\mathbf{x}$ is a vector consisting of $\{x_1, x_2, ..., x_5\}$ input elements which represent the $\{s_{1,t-1}, s_{2,t-1}, ..., s_{5,t-1}\}$ technical indicator signals (Equation (5.3)), and y represents the mean return over the next 2 minutes (Equation (5.1)).

The idea of fuzzy inference systems can be broken down into a *divide-and-conquer* (Jang and Sun, 1995) approach. The first objective is to identify fuzzy regions that partition the input space using the antecedents of fuzzy rules. The second objective is to map a local behaviour within a given region using the rule consequents. The selection of the space partitioning scheme has two important effects on the resulting model. The first effect is that the more granular the space partitioning, the higher the number of rules, hence improving model accuracy. However, a resulting effect is the increased number of optimisation parameters and therefore computational complexity. Increased model complexity can also result in overfitting. The designer has to balance model accuracy and complexity depending on the structure of the underlying data and the specific context.

For this reason, we apply a clustering algorithm, namely fuzzy c-means (FCM) clustering, that we introduced earlier in Chapter 4. Similar to Chapter 4, we use FCM to identify fuzzy partitions in data (Bezdek, 1981; Dutta and Angelov, 2010); however, in this chapter we use it as a precursor (rather than a subsequent step) to our learning algorithm. By controlling the number of clusters, this gives us the opportunity to identify the best model structure which balances model accuracy and complexity (improving on both performance and complexity when compared to the rule database structure presented in Chapter 3). This is possible because each cluster centre essentially exemplifies a characteristic behaviour of the system in a specific region. Hence, each cluster centre can be used as the basis of a membership function for each input variable and these are combined in a rule that describes the local system behaviour.

Since our input variables have a different basis, these variables are standardised and rescaled to have a mean of zero and a standard deviation of one before being fed into the algorithm. This ensures equivalent influence weighting of each variable on the clustering algorithm. Let z denote the fuzziness index. Furthermore, define α as the number of clusters, $d_{uv} = \|\mathbf{x}_u - \mathbf{c}_v\|^2$ as the Euclidean distance between the u-th realisation and the current v-th cluster centre $\mathbf{c}_v$, and $d_{uo} = \|\mathbf{x}_u - \mathbf{c}_o\|^2$ as the Euclidean distance from the u-th realisation and the other cluster centres $\mathbf{c_o}$. For each data point $\mathbf{x}_u$, $\forall u \in [1,q]$, and cluster $\mathbf{c}_v$, $\forall v \in [1,\alpha]$, the FCM algorithm iteratively updates the membership grade μ of the u-th data point to the v-th cluster until the algorithm has converged (a detailed description of the FCM algorithm is presented in Section 4.2.5, page 117).

As suggested by Pal and Bezdek (1995), we test values between 1.5 and 2.5 for the fuzzy index z. For the number of clusters, α, we test values between 2 to 5 clusters. This range selection is based on identifying a well-distributed set of clusters representing the variable distribution, whilst at the same time avoiding the heavy influence of possible outliers. After examining cluster plots on different stocks and time periods, we select a fuzzy index value of 1.7. The number of clusters is used as a model parameter which is included as part of the model tuning process in combination with the ANFIS parameters (see next section).

In this chapter, we adopt Gaussian membership functions (MFs), where each fuzzy set is represented by

$$Gaussian(x;\bar{x},\sigma) = e^{-(x-\bar{x})^2 / 2\sigma^2}. \tag{5.5}$$

An MF returns the degree of membership, in the range [0,1], of a specific point in a particular variable region. We select this particular fuzzy set shape because, unlike other MFs, it has only two parameters (the mean $\bar{x}$ and the standard deviation σ) and it always spreads out over the entire input domain (Wu and Mendel, 2014).

Once the clustering process is complete, we follow two steps to create an initial T1 TSK fuzzy model (preliminary background on FIS structure identification using clustering is presented in Section 2.4.1). In the first step, the identified α clusters are projected on each input variable space. This results in α Gaussian MFs for each variable with the mean represented by the cluster centres **c**. The standard deviation σ is obtained by re-arranging Equation (5.5) and utilising cluster centres **c** and membership grades μ. As a second step, a set of α rules are created in the form

$$\begin{aligned} IF \quad & (x_1 \ is \ A_{i,1}) \ AND \ (x_2 \ is \ A_{i,2}) \ AND \ ... \ AND \ (x_5 \ is \ A_{i,5}) \\ THEN \quad & y_i \ = b_i + \sum_{k=1}^{5} w_{i,k} x_k \end{aligned} \tag{5.6}$$

where $A_{i,k}$ represents T1 Gaussian MFs, projected from the identified clusters, for the i-th rule and the k-th input $(i = 1,2,...,\alpha; k = 1,2,...,5)$. In the consequent, y_i is the rule output, defined as the mean return over the next 2 minutes (Equation 5.1), as a linear function of the input variables $\{x_1, x_2, ..., x_5\}$ with parameters b_i and $w_{i,k}$. Following the identification of the initial FIS using FCM clustering, we seek further model tuning which is described in the next section.

FIS tuning with ANFIS

From the literature we identify two major classes of optimisation algorithms for FLSs: gradient-based algorithms and heuristic algorithms, where in the latter case most studies focus on evolutionary computation (EC) algorithms (see the discussion in Wu and Mendel, 2014). ANFIS follows the former approach. It is due to the popular use of ANFIS in finance that we decide to use it as our optimisation technique and hence as a benchmark model to explore possible risk-adjusted performance improvements by extending the model to an IT2 TSK FLS. The performance stability of ANFIS is also demonstrated by our findings in Chapter 3. The optimisation algorithm is also similar to the approach that we adopted in previous chapters, hence we eliminate, as much as possible, any performance discrepancies resulting from alternative optimisation techniques and focus more on identifying performance gains as a result of model structure and identification improvements.

Based on our variable selection (Section 5.2.1), next we define the ANFIS configuration that we adopt in this chapter. The main differences when compared to the ANFIS configuration that we present in Chapter 3 (Section 3.2.3) are that here we have 5 input variables (rather than 3) and the number of membership functions and rules are based on the number of identified clusters α (rather than MF space partitioning). The ANFIS network is defined as follows:

Layer 1 Since we have 5 inputs, this layer contains $5 \times \alpha$ adaptive nodes, one node for every membership function associated with each input. For instance, the α nodes with connections from the first input x_1 are in the form

$$O_{1,i} = \mu_{A_{i,1}}(x_1) \text{ for } i = 1,2,...,\alpha, \tag{5.7}$$

where $\mu_{A_{i,k}}$ is the membership degree for the i-th T1 MF and the k-th input ($i = 1,2,...,\alpha; k = 1,2,...,5$). In our setup, α represents both the number of rules and also the number of MFs for each input variable. Different values for α are tested as part of our model calibration process (see Table 5.1).

Layer 2 This layer contains α fixed nodes. In each node, $O_{2,i}$, where $i = 1, ..., \alpha$, the product t-norm ($\star$) is used to "AND" the membership grades which are passed from the previous layer. The output is the firing strength, f_i, of each rule:

$$O_{2,i} = f_i = \mu_{A_{i,1}}(x_1) \star \mu_{A_{i,2}}(x_2) \star \ldots \star \mu_{A_{i,5}}(x_5). \tag{5.8}$$

Layer 3 In this layer, which consists of α fixed nodes, the normalised firing strengths, $\hat{f}_i$, are calculated using

$$O_{3,i} = \hat{f}_i = \frac{f_i}{\sum_i f_i}. \tag{5.9}$$

Layer 4 The nodes in this layer are adaptive and act as a function

$$O_{4,i} = \hat{f}_i y_i = \hat{f}_i (b_i + \sum_{k=1}^{5} w_{i,k} x_k), \tag{5.10}$$

where $\hat{f}_i$ is the normalised firing strength from the previous layer and y_i is the rule consequent linear function for the i-th rule, $i = 1, ..., \alpha$, with parameters b_i and $w_{i,k}$.

Layer 5 This layer consists of a single node and combines the output from all the nodes in the previous layer to calculate the overall output as

$$O_{5,i} = r_t = \sum_i \hat{f}_i y_i = \frac{\sum_i f_i y_i}{\sum_i f_i}. \tag{5.11}$$

Similar to the process described in Section 3.2.3, for ANFIS learning we follow an iterative two-pass algorithm. In a forward pass, the premise parameters (in Layer 1) defining the membership functions are unmodified and the consequent parameters (in Layer 4) are computed using the least squares algorithm. On completion of the forward pass, the

Table 5.1: Parameters tested for ANFIS.

Parameter	Parameter Value Set
Training Data Size (data points)	{510, 1020, 2040}
Number of Input Membership Functions and Rules (α)	{2, 3, 4, 5}
Training Epochs	{10, 20, 40, 60}

consequent parameters are unmodified and a backward pass feeds the errors back into the network using back-propagation to adjust the premise parameters (more detail is presented in Section 2.4.2). In our in-sample training and model selection process we test and compare all $3 \times 4 \times 4 = 48$ permutations of the parameter combinations (Table 5.1). The training sample size and training epochs options are kept in line with the ANFIS model that we utilise in Chapter 3. These parameters are tested in combination with an additional set of parameters that are defined for our trading algorithm (see Section 5.2.4). In the next section, we propose how the standard ANFIS model can be extended to a T2 TSK model.

5.2.3 Generalizing ANFIS model to T2 FLS

There is no mathematical proof that by changing a T1 FLS to T2 FLS, a T2 fuzzy logic controller (FLC) will automatically outperform a T1 FLC (Wu and Mendel, 2014). When considering T2, an initial step for an algorithm designer is to understand the underlying uncertainties and the sources thereof. In an HFT environment, one can identify a number of sources of uncertainty:

- Constantly changing market activity and volatility conditions.
- The use of non-precise terms: “rising steadily”, “high volatility”, “small loss”, “high activity”.
- Microstructure noise in the observations.

- Inconsistencies affecting trade execution (for example, execution time and transaction costs).

By identifying the levels and sources of uncertainty, the designer can formulate an initial opinion, firstly on the fit for T2 FLS, rather than using T1 FLS, and secondly on the numerous T2 design options to consider (see Wu and Mendel, 2014, for a summary of design options). In our scenario, all the identified uncertainties point towards T2 as being a good contender.

Our primary interest is to identify model improvements that can result from the better handling of the microstructure noise. This noise is attributed as one of the key modelling challenges and sources of uncertainty in HFT (Anderson, 2011; Anderson et al., 2013; Rechenthin and Street, 2013). T1 fuzzy sets cannot fully represent the uncertainty associated with the inputs since, as a contradiction, the membership function of a T1 fuzzy set has no uncertainty associated with it. To address this criticism, Zadeh (1975) introduced T2 fuzzy sets, which has been a very active area of research (see Wu and Mendel, 2014). We consider Interval T2 (IT2) fuzzy sets since they are much less computationally intensive and more popular than the generalised T2 (Mendel et al., 2014; Wu and Mendel, 2014). This projects our focus on investigating whether the introduction of T2 fuzzy sets in the antecedents of fuzzy rules can improve the risk-adjusted performance of trading algorithms, in which scenarios and to what extent. We consider a special IT2 TSK FLS when its antecedents are T2 fuzzy sets but its consequents are crisp numbers (referred to as A2-C0 by Mendel et al., 2014). Although other IT2 models exist (see Mendel et al., 2014), we select this model because it provides a good balance between the better management of uncertainty and increased model complexity. Preliminary mathematical background on IT2 models is provided in Sections 2.5 and 2.6.

IT2 FLS design approach

At this point, we decide on an important design consideration for our experiments. The decision is to select the method to adopt when it comes to tuning our IT2 FLS. In this section, we describe our rationale.

We consider two different approaches that are commonly adopted to design IT2 FLSs (Aladi et al., 2014; Mendel et al., 2014; Wu and Mendel, 2014): a partially dependent approach and a totally independent approach. In the former approach, the designer starts with an optimised T1 FLS, which is then used as a basis for the design of the IT2 FLSs. On the other hand, the totally independent approach is used to design IT2 FLSs from scratch, hence avoiding the use of an intermediate T1 FLS.

We adopt the partially dependent approach for a number of reasons. Firstly, although previous literature found that T2 models can perform better under increasing uncertainties (Aladi et al., 2014; Sepulveda et al., 2006), we do not seek to achieve an optimal performance in the error reduction; instead, the primary objective is to compare T1 FLSs and IT2 FLSs and to shed more light on the possible gains in risk-adjusted performance. Secondly, in line with Wu and Mendel (2014), the increased parameters and design options in IT2 FLS can be overwhelming and possibly limit more widespread use. With the adopted approach, our objective is to propose an incremental step from standard ANFIS, which as we highlighted earlier is a popular and already established technique in finance, and to contribute new improvements in this active area of research. Thirdly, our parsimonious approach also reduces the possibility of overfitting and spurious results (see Bailey et al., 2014). Whilst these reasons present our rationale for selecting a partially dependent approach, in Section 5.3.3 (IT2 design considerations) we highlight the strengths and weaknesses of this approach.

ANFIS/T2 models

In this section, we propose two methods, ANFIS/T2a and ANFIS/T2b, of how the T1 FIS structure resulting from the ANFIS training can be extended to an IT2 FLS (A2-C0) model. In line with the economic theories about the separation of overall risk between risk and uncertainty (Heal and Millner, 2014; Knight, 1921; Nelson and Katzenstein, 2014), our intention is to seek trading performance improvements that can result from the introduction of T2 fuzzy sets. This is obtained by minimising the uncertainty caused by microstructure noise present in high-frequency data, hence reducing the overall risk.

The A2-C0 rules are defined as follows:

$$\begin{aligned} IF \quad & (x_1 \; is \; \widetilde{A}_{i,1}) \; AND \; (x_2 \; is \; \widetilde{A}_{i,2}) \; AND \; ... \; AND \; (x_5 \; is \; \widetilde{A}_{i,5}) \\ THEN \quad & y_i \; = \; b_i + \sum_{k=1}^{5} w_{i,k} x_k \end{aligned} \tag{5.12}$$

where in the premise part of the rule, $\widetilde{A}_{i,k}$ are IT2 Gaussian MFs projected from the identified clusters for the i-th rule and the k-th input $(i = 1,2,...,\alpha; k = 1,2,...,5)$. In the rule consequent, y_i is a linear function of the input variables $\{x_1, x_2, ..., x_5\}$ with parameters b_i and $w_{i,k}$.

For our proposed ANFIS/T2a, the objective is to convert the T1 rules defined in Equation (5.6) to the A2-C0 rules defined in Equation (5.12). To do this, we start from the ANFIS-optimised T1 fuzzy sets. The next step is to introduce the footprint of uncertainty (FOU) for the MFs in the premise part of the rules whilst keeping the consequent part fixed. The FOU represents the blurring effect of a T1 membership function, $\mu_{A_{i,k}}$, and is completely described by two corresponding bounding functions, a lower membership function (LMF), $\underline{\mu}_{A_{i,k}}$, and an upper membership function (UMF), $\overline{\mu}_{A_{i,k}}$, both of which are T1 fuzzy sets. Unlike their T1 counterparts, whose membership values are precise numbers in the range [0, 1], the membership grades of a T2 fuzzy set are themselves T1 fuzzy sets. Therefore, T2 fuzzy sets offer the ability of modelling higher levels of uncertainty (John and Coupland, 2007; Mendel et al., 2006). Aladi et al. (2014) show how T2 fuzzy sets can handle increased noise and claim a direct relationship between FOU size and levels of noise. However, Benatar et al. (2012) warn that selecting too small an FOU will result in no improvements over the T1, whilst if too large, the T2 model will perform worse. In our case, complexity is compounded since noise levels are not fixed but time-varying due to time-varying market activity.

To define the size of the FOU, for ANFIS/T2a we adopt a parsimonious approach by introducing one additional parameter. This parameter, $\beta \in [0,1)$, determines the increase or decrease in the standard deviation, $\sigma_{i,k}$, of all the Gaussian T1 MFs across all input variables ($i = 1,2,...,\alpha; k = 1,2,...,5$), whilst keeping the mean, $\bar{x}_{i,k}$, fixed. Hence, for each T1 MF, the LMF and UMF are defined as follows:

$$\underline{\mu}_{A_{i,k}} = Gaussian(x_k; \bar{x}_{i,k}, (1-\beta)\sigma_{i,k}) \tag{5.13}$$

$$\overline{\mu}_{A_{i,k}} = Gaussian(x_k; \bar{x}_{i,k}, (1+\beta)\sigma_{i,k}) \tag{5.14}$$

where $\bar{x}_{i,k}$ and $\sigma_{i,k}$ are the parameters for the i-th T1 Gaussian MF and k-th input tuned by ANFIS. When applied, this results in a new set of IT2 MFs (Figure 5.2). We transform the complete T1 to IT2 rule base in this manner. The final output of the model is obtained as follows

$$\mathbf{Y}_{A2-C0} = [y_l, y_r] = \int_{f_1 \in [\underline{f}_1, \overline{f}_1]} \cdots \int_{f_\alpha \in [\underline{f}_\alpha, \overline{f}_\alpha]} 1 / \frac{\sum_i f_i y_i}{\sum_i f_i} \tag{5.15}$$

where the integral sign represents the fuzzy union operation and the slash operator $(/)$ associates the elements of rules output and firing strength with their secondary membership grade, which in the case of IT2 is simplified to 1. The firing strength for each rule i, where

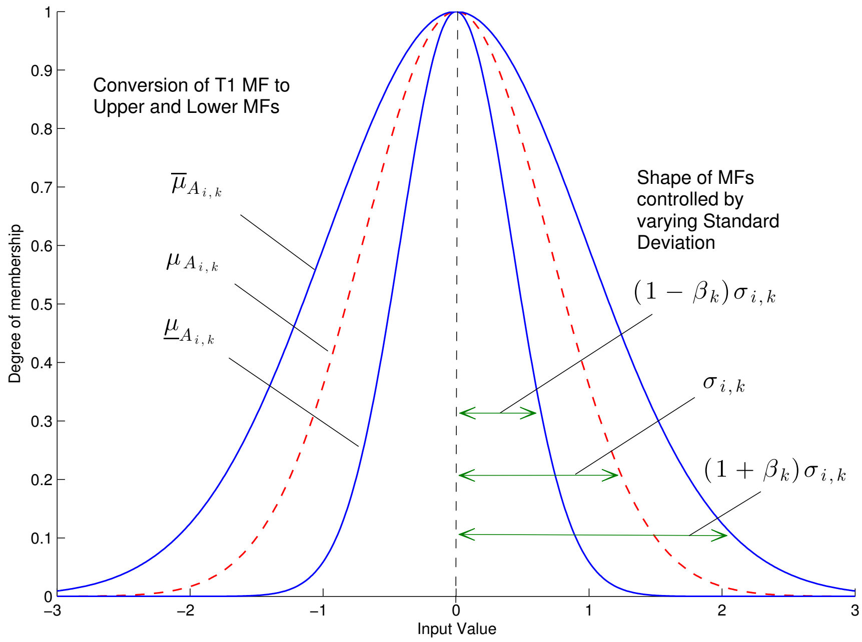


Fig. 5.2: Conversion of T1 fuzzy set to IT2 with fixed mean and uncertain standard deviation. Upper and lower MFs are defined using an additional parameter β_k, which represents the width of the IT2 MF as a percentage increase or decrease on the base T1 MF standard deviation respectively. In our first experiment, we train the model to identify and assign the same β value across all inputs $k = 1, 2, ..., 5$. In our second experiment, each β_k is tuned to different values.

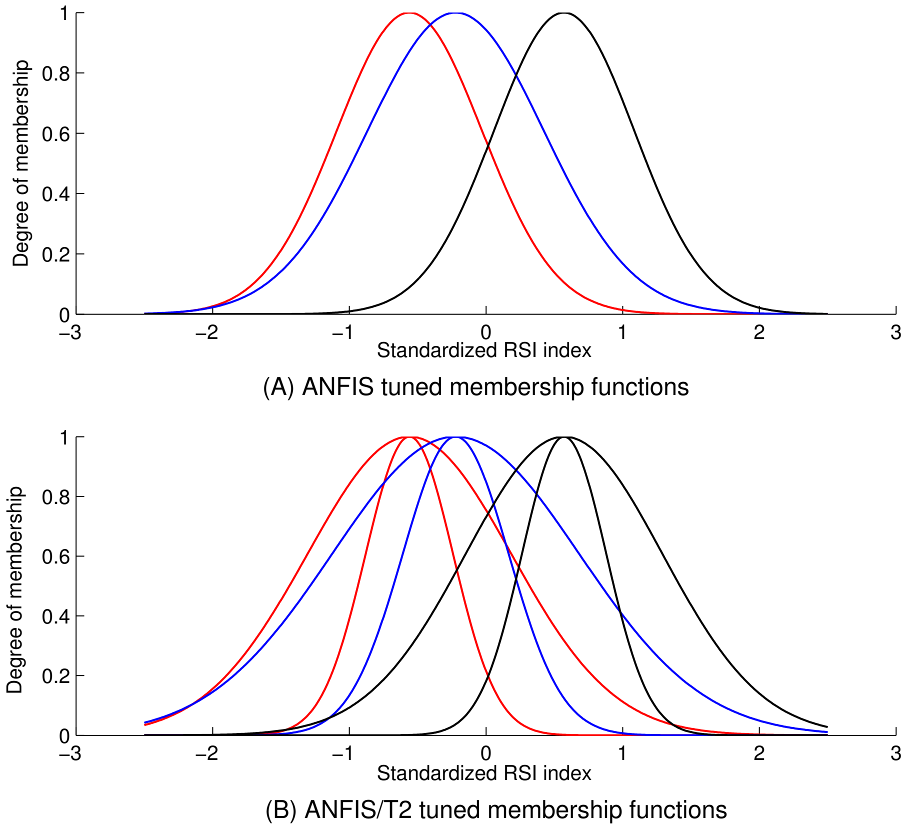
$i = 1, 2, ..., \alpha$, is calculated as

$$\underline{f}_i(\mathbf{x}) = \underline{\mu}_{A_{i,1}}(x_1) \star \underline{\mu}_{A_{i,2}}(x_2) \star \ldots \star \underline{\mu}_{A_{i,5}}(x_5) \tag{5.16}$$

$$\overline{f}_i(\mathbf{x}) = \overline{\mu}_{A_{i,1}}(x_1) \star \overline{\mu}_{A_{i,2}}(x_2) \star \ldots \star \overline{\mu}_{A_{i,5}}(x_5) \tag{5.17}$$

where, like in the case of ANFIS (Equation (5.8)), $\star$ represents the product t-norm. For further details, the interested reader is directed to Mendel et al. (2014).

Following the ANFIS optimisation, as a second training pass we train our model with values of β ranging from 0% to 40% in discrete steps of 5%. This range is selected from our testing on in-sample data. It is to be noted that we intentionally include 0% in our search space, which results in the reduction of the IT2 FLS back to the corresponding T1 FLS. This



(A) ANFIS tuned membership functions

(B) ANFIS/T2 tuned membership functions

Fig. 5.3: An example showing the MFs for the input variable RSI index following ANFIS training (top), and the corresponding ANFIS/T2 tuned MFs (bottom). Both ANFIS and ANFIS/T2 MFs are dynamically adapted on a daily basis, reflecting the price volatility in that period. During training, the latter IT2 MFs optimal thickness is identified by initially adopting the same Gaussian MFs parameters tuned by ANFIS and then increase the standard deviation by a factor ranging from 0% to 40%.

allows the training algorithm to select between T1 and IT2 FLS and to dynamically adapt the model FOU according to the level of market uncertainty during the specific training period (see the example in Figure 5.3). This also guarantees that during the training process, the model achieves at least the same level of performance as that of the T1 FLS.

For our second proposed method, ANFIS/T2b, we introduce more flexibility in the model by introducing 5 new parameters in the model, $\{\beta_1, \beta_2, ..., \beta_5\}$, where $\beta_k \in [0,1)$. These parameters represent the increase or decrease in the FOU for all T1 fuzzy sets defining the space of the individual 5 input variables. Hence, with this approach, the FOU can dynamically adapt to different levels of uncertainty across the different input variables. Therefore in the case of ANFIS/T2b, we convert every rule i, where $i = 1, 2, ..., \alpha$, by transforming the MFs of each input variable to IT2 MFs using β_k, where $k = 1, 2, ..., 5$. In this case, the lower and upper MFs are defined as:

$$\underline{\mu}_{A_{i,k}} = Gaussian(x_k; \bar{x}_{i,k}, (1 - \beta_k)\sigma_{i,k}) \tag{5.18}$$

$$\overline{\mu}_{A_{i,k}} = Gaussian(x_k; \bar{x}_{i,k}, (1 + \beta_k)\sigma_{i,k}). \tag{5.19}$$

In our training algorithm, we apply the same discrete range for possible β_k values. This approach results in a much larger search space of possible β_k combinations, hence rather than performing a parameter sweep, we speed up the training process by optimising the β_k values using a mixed integer genetic algorithm. This does not limit our approach to other possible optimisation methods. Following the rule base conversion from T1 to A2-CO, the final step is to decide on how to compute the output. This is described in the next section.

Computing the output

The T2 fuzzy logic community has proposed a number of methods for computing the output (the interested reader is directed to Mendel et al., 2014; Wu, 2013, for a review). At a high level, the methods can be divided into two groups. The first group requires an interim defuzzification process which reduces the output T2 fuzzy sets to T1 fuzzy sets. These are typically solved via iterative algorithms. The second group skips this step completely by calculating the output directly. Siding with less computational intensive methods, we chose the Nie-Tan (NT) method (Nie and Tan, 2008), which falls under the latter group. The NT method computes the output as follows:

$$y = \hat{r} = \frac{\sum_{i=1}^{\alpha} y_i(\underline{f}_i + \overline{f}_i)}{\sum_{i=1}^{\alpha}(\underline{f}_i + \overline{f}_i)}, \tag{5.20}$$

which effectively makes use of a vertical-slice representation of a T2 fuzzy set and involves taking the mean of the lower and upper membership functions, creating a type-1 fuzzy set. When compared to other defuzzification methods, it was demonstrated that the NT method provides a good balance between accuracy and complexity (Greenfield and Chiclana, 2013; Wu, 2013).

5.2.4 Trading algorithm

Our trading algorithm is kept similar to the approach presented in Chapters 3 and 4, with daily trading activity starting from 8:30, the latest possible trade is executed at 16:00, and no overnight positions. The similarity is purposely kept in order to attribute any performance improvements to the proposed T2 enhancements. Every minute, the prediction of the average return over the following two minutes is passed to the trading algorithm which in turn recommends a *buy* ($\Phi = long$), *sell* ($\Phi = short$) or *stand-by* ($\Phi = 0$) action. Similar to other chapters, we allocate a starting capital of 250,000 GBP for each stock, buy or sell positions of $50,000$ GBP, a transaction cost of 10 GBP per trade per direction, and a threshold of a 5% annual risk free rate. When a trade is closed, the net proceeds are added back to the capital balance, hence maximising the utilisation of the available capital.

However, in this chapter, the algorithm is calibrated to simulate an increased number of intraday trades. This is done primarily through the re-calibration of the return threshold (RT) and trade duration (TD) parameters. In our experiments, we explore four levels of the RT parameter for each stock, $\{0.08\%, 0.06\%, 0.04\%, 0.02\%\}$, to avoid small price movements mostly originating from microstructure effects. As suggested by Vanstone and Finnie (2009), in our algorithm we also take account of whether the signal is increasing in strength, or decreasing in strength from its previous forecast. Hence, before opening a position, the algorithm confirms the current signal by comparing this with the forecast signal generated in the previous 1-minute time bar.

For our second parameter, TD, which represents the duration of each trade in minutes, we consider values between 2 minutes to 10 minutes. Following evidence by Rechenthin and Street (2013), this range is selected to reduce the possible perceived movements resulting from possible noise. The time window is, however, short enough to capture any instances of market inefficiencies due to HFT. If a signal in the same direction is generated prior to closure, TD is reset back to zero. This form of extended close proved to be successful in previous studies (Brabazon and O'Neill, 2006).

The pseudo code presented in Algorithm 1 (presented in Chapter 3, page 67) describes the structure of the trading algorithm that we use in our experiments. Both RT and TD parameters are used in conjunction with the standard ANFIS parameters referred to earlier (see Table 5.1), and are applied in the model selection process to identify the base T1 model for each stock.

5.3 Experiment approach

We conduct two experiments (see Table 5.2). In the first experiment (1a-1d), each model is tuned at the optimum level of return threshold (and hence, trading frequency) that can result in the highest Sharpe ratio by the respective model. In the second experiment (2a-2c), the trading performance of our proposed T2 models is compared with standard ANFIS at different fixed levels of return thresholds with the objective to identify the best performing model at different degrees of uncertainty.

For reference and comparison, we also refer the reader to the Buy-and-Hold and random models that we present in Chapter 4. Based on the fact that we indicated the superiority of our proposed models against these benchmarks and also since these models use the same data set that is used in this chapter, these are not repeated here. However, we introduce a new benchmark model that is based on the use of optimised technical indicators (1b). In contrast to the technical models that we present in Chapter 3 in which we apply the same moving average signals that are applied in the AI models in order to compare the effectiveness with our fuzzy dynamic moving average approach, in this chapter we optimise the MA and RSI indicators on the same in-sample period and then tested out-of-sample. For our model tuning we applied a parameter sweep approach. In the case of MA, we identify the best lead and lag indicators in the range from 5 to 100 in steps of 5 (refer to Equation (3.3)). In the case of RSI we tested signals ranging from 25 to 35 in steps of 5 for the upper boundary and correspondingly from 65 to 75 for the lower boundary. The signals generated by the technical indicators are fed to the same trading algorithm that is used by our AI algorithms, hence adopting the same intraday trading conditions. The model with the highest positive Sharpe ratio is selected for out-of-sample testing. By including this benchmark we provide an approach to identify stepwise improvements that can be gained by moving from a standard technical rules approach to our proposed type-1 fuzzy logic models, followed by the move from type-1 to type-2 models.

In the following sub-sections, we describe the different aspects of the approach we adopt in our experiments and also aim to present a critical analysis in view of our decisions. Section 5.3.1 describes our data and underlying statistics. Section 5.3.2 explains the performance measures that we adopt to perform model selection and evaluation. A discussion about our IT2 model design considerations is presented in Section 5.3.3. This is followed by a description of our model training and testing process (Section 5.3.4) and our approach for controlling different levels of noise (Section 5.3.5).

5.3.1 Data

We again utilise the standard data set described in Chapter 3, Section 3.2.8. Similar to Chapter 4, we utilise 1-minute trade records for the 27 stocks covering the same period between 28/06/2007 and 25/06/2008. Data is sampled at 1-minute intervals using the last trade price every 1 minute.

Table 5.2: Models applied in the experiments.

Experiment	Algorithm	Adaptive Tuning	Return Threshold (Uncertainty)
1a	ANFIS	T1 MFs & Rules	Optimum
1b	Technical Rules	MA and RSI parameters	Optimum
1c	ANFIS/T2a	IT2 MFs (variable width) & Rules	Optimum
1d	ANFIS/T2b	IT2 MFs (variable width/MF) & Rules	Optimum
2a	ANFIS	T1 MFs & Rules	Fixed Levels
2b	ANFIS/T2a	IT2 MFs (variable width) & Rules	Fixed Levels
2c	ANFIS/T2b	IT2 MFs (variable width/MF) & Rules	Fixed Levels

In the selection of our data set size, we note the harsh criticism brought forward by Bailey et al. (2014) in view of the number of publications which base their study on a small backtest length given the number of model configurations tested. Bailey et al. (2014) prove that this easily gives rise to possible overfitting with the chance of spurious results (especially in in-sample tests). To mitigate this risk, in our case, each model is tested on a number of price points which would be equivalent to over 500 years of daily data per stock.

Another important consideration when selecting stocks for backtesting purposes is the importance of picking a mix of stocks which exhibit different trends. As it can be seen from the numerous machine learning and artificial intelligence studies surveyed in Krollner et al. (2010), Tsai and Wang (2009) and Bahrammirzaee (2010), this is rarely considered. Pardo (2011) warns that only including stocks that follow similar trends can lead to ungeneralised models which only work in specific scenarios only, hence introducing a bias in the experiment results. We note this risk when picking the stocks, and as we show in the data descriptive statistics presented in Chapter 3, Section 3.2.8, we include a mix of stocks with both positive and negative mean returns from diverse industries for the selected training and test period. Further testing on a wider stock selection, instruments and markets is left as future work.

5.3.2 Performance measures

Cavalcante et al. (2016) present a review of around 56 computational intelligence studies with application to finance published from 2009 to 2015. These studies cover techniques for preprocessing and clustering of financial data, for forecasting future market movements, for mining financial text information. The large majority of these studies focus on the minimisation of error functions, directional accuracy or else profitability (issues with these measures are discussed in Brabazon and O'Neill (2006) and Pardo (2011), and we explore them further in Chapter 3). The danger with error functions is that a small error does not necessarily translate into profitability since it does not reflect the direction. On the other hand, directional accuracy measures are not enough to ensure overall profitability since they do not incorporate the magnitude of the correct or incorrect predictions. Moreover, a high directional accuracy might be completely misleading since few large losses can still cancel out a higher number of wins, but which are smaller in size. Finally, focusing only on profitability does not incorporate the possible drawdowns that can be experienced during specific periods. This can be disastrous for an investor. This also reflects the rules that are being proposed by directives like MiFID II that are intended to ensure that trading algorithms show robustness with a lower risk of unexpected huge losses.

In this chapter we remain consistent with previous chapters and use the Sharpe ratio (Equation 3.20) as our key measure of risk premium per unit of risk in an investment, including the application of a 5% annual risk free rate. We also include the Sortino ratio as a risk-adjusted measure. In view of our interest to investigate model performance at higher trading frequencies, we present the profit per trade measure as an indication of quality of the trades placed by the respective models. This provides an indication of the existing spread between the average return per trade and the underling transaction costs. A higher spread would indicate the possibility of increasing the number of trades with the chance to increase overall profitability. We conclude that the combination of the above measures, in conjunction with overall profitability, provide a clear overall picture of model performance both in terms of profitability and risk.

5.3.3 IT2 design considerations

For our IT2 design approach, we consider two different options that are commonly adopted to design IT2 FLSs: a partially dependent approach and a totally independent approach. Although in Section 5.2.3 we present both options and the reasons why we opted for a partially dependent approach, it is important for model designers to understand the advantages and disadvantages of both options. In the partially dependent approach, the primary advantage is that it makes it easier to directly compare the T1 and IT2 FLSs. A second advantage is that the training of the IT2 FLS can be much faster since a number parameters would already have been optimised by the T1 model. On the other hand, the advantage of the totally independent approach is that it avoids the assumption that the optimised parameters of the T1 model; for example, the type and number of membership functions, are the best parameters to be inherited by the IT2 FLS, hence possibly leading to a sub-optimal IT2 FLS model. This is a conscious risk that we undertake in this study, the primary reason being that our main objective is the comparison of T1 and IT2 models.

5.3.4 Model training and testing

The AI model selection process is based on identifying the best model parameters (defined in Table 5.1 in conjunction with trading algorithm parameters RT and DT) that result in the highest performance during the 150-day in-sample period. Indirectly, this means that the optimum trading frequency is selected for each model that results in the highest Sharpe ratio, utilising as well our lowest benchmark of 5% annual risk free rate. In the case of ANFIS/T2 models, parameter selection is extended to the identification of β parameters (as described in Section 5.2.3) which define the optimum size of FOU.

Common practice in time series and machine learning literature is to divide the time-series into training, testing and validation sets. However, Kaastra and Boyd (1996) and Pardo (2011) argue that in the case of trading scenarios, a more rigorous approach is to adopt moving window (also known as walk-forward) testing which consists in a series of overlapping training-validation-test sets. Although the moving window approach requires more frequent model re-training, it tries to simulate real-life trading and also permits quicker model adaptation to changing market conditions. More detailed description of this approach and other training aspects are presented in Chapter 3, Section 3.2.7. In this case, at day_d, where $(d = 1, 2, ..., 150)$, the model is trained on 1-minute data points (Equation (5.3)) from day_{d-r} to day_{d-1}, and r represents the training data size in days. The trained model is then validated by predicting minute-by-minute mean returns (Equation (5.1)) during day_d. This is repeated for the whole 150-day in-sample period, for each parameter combination. Further detail and considerations on the training and calibration process are provided in Chapter 3, Section 3.2.7. A key concern is the fact that the increase in model complexity of our IT2 models might lead to possible overfitting (high variance). We address this concern by limiting the number of parameter optimisations required for both ANFIS/T2a and ANFIS/T2b (e.g. ANFIS/T2a FOU size is adjusted by introducing only 1 additional parameter with values of β ranging from 0% to 40% in discrete steps of 5%). The final selected model is then tested, using the same day-by-day moving window approach, over the next 100-day out-of-sample period. The size of the time series provides a sufficiently large historical dataset which reduces the possibility of over fitting or the production of spurious results during backtesting (Bailey et al., 2014).

5.3.5 Controlling different levels of uncertainty

An important decision that we need to take is to define the approach to use to simulate different levels of uncertainty. This will in turn enable us to compare T1 FLS with IT2 FLS under different degrees of uncertainty.

We propose an innovative approach which can allow us to adjust, in a controlled fashion, the level of uncertainty and hence be able to compare T1 and IT2 FLSs under different uncertainty conditions. A number of authors (e.g. Aladi et al., 2014; Sepulveda et al., 2006) take the approach of methodically generating and injecting synthetic noise in the data. We decide to take a different approach. The use of a return threshold is a common method of filtering model prediction signals (Gençay, 1996; Holmberg et al., 2013; Vanstone and Finnie, 2009, 2010). The filter helps to avoid unprofitable small price fluctuations and possible noise by discarding any signals that fall below a specified threshold. We follow this approach in both Chapter 3 and Chapter 4. Here, we extend the use of this method by hypothesising that this approach is effectively controlling for uncertainty (indirectly). An increased (reduced) signal strength effectively translates into reduced (increased) uncertainty that the predicted move is due to price fluctuations or noise, rather than true price trends, with the reduced (increased) risk to result in unprofitable trades. Hence, we test the models under different levels of uncertainty by adjusting different levels of the return threshold. We argue that this proposed approach is more practical and realistic for algorithmic trading scenarios rather than by injecting synthetic noise.

Based on this decision, in the second experiment (see 2a-2c in Table 5.2) the objective is to compare the trading performance of T1 FLS and IT2 FLS at different levels of uncertainty by controlling the return threshold, RT. For this reason, after testing the models during the in-sample period, using the same moving window approach that is applied in the first experiment, the model selection process is based on choosing the model which returned the highest Sharpe ratio at specific levels of RT. Subsequently, we conduct statistical tests on the 100-day out-of-sample results to determine whether there is enough evidence of IT2 superiority at different levels of RT. The results of these experiments are presented in Section 5.4.2.

5.4 Results and analysis

In Section 5.4.1, we conduct a performance comparison between the benchmark models, namely B&H methods and standard ANFIS, and the proposed IT2 FLS models during 100 out-of-sample trading days. In Section 5.4.2, we analyse the models' performance across increasing levels of uncertainty.

5.4.1 Experiment 1: Comparison against standard ANFIS model

In our first experiment we investigate whether the introduction of type-2 fuzzy logic leads to improvements on standard ANFIS. For each model and stock, we utilise the optimum signal threshold, hence trading frequency, that in the in-sample period results in the best Sharpe ratio for the specific stock and model. In our investigation we also compare the results with an optimised technical signals approach.

First we investigate the results obtained by standard ANFIS during the 100-day out-of-sample period. The results in Table 5.3 show that the algorithm scores positive results in 25 out of 27 stocks (93%), the only exceptions being SDR and IAP. We note that minor positive excess returns (<1%) were obtained in two stocks, namely JMAT and AZN. Considering that the results are based on a 100-day out-of-sample period, the model shows moderate excess returns (3% to 10%) as in the case of BLT, BLND, AV, DGE, HSBA, ABF, RIO, GSK and XTA, and excellent returns on all other stocks. These results validate the popularity of ANFIS in finance (Boyacioglu and Avci, 2010; Chang et al., 2011; Chen, 2013; Kablan and Ng, 2011; Tan et al., 2011; Wei et al., 2014) and the active research in improving the

Table 5.3: Standard ANFIS performance after the 100-day out-of-sample period.

	Standard ANFIS	
Stock	**Sharpe Ratio**	**Excess Return**
AL	0.8943	116.95%
ANTO	0.3084	13.29%
BLT	0.2733	7.63%
BA	1.5265	84.58%
BLND	0.4087	9.69%
SKY	0.4960	15.30%
CWC	1.2860	141.29%
AV	0.3444	6.81%
DGE	0.3133	6.36%
SDR	-0.1244	-2.63%
HMSO	0.4375	15.75%
REX	0.7213	35.09%
JMAT	0.0212	0.34%
HSBA	0.2366	5.93%
ABF	0.2272	4.90%
INTU	0.5033	14.91%
RIO	0.1477	3.59%
BP	0.9238	41.41%
SGE	1.4761	80.65%
LLOY	0.7065	65.82%
TSCO	1.5164	60.86%
GSK	0.1897	4.26%
AZN	0.0219	0.23%
HBOS	0.4554	24.93%
XTA	0.0940	4.57%
IAP	-0.0321	-2.03%
ITV	1.3787	63.06%
Average	0.5464	30.49%

model and application techniques. This also strengthens our findings in Chapter 3 and our proposal to use ANFIS as our main benchmark model and use it as a basis to seek further improvements by extending to T2.

From further investigation of the standard ANFIS results (Table A.7) we notice that at the individual stock level the Sortino ratio measures are in line with those obtained using the Sharpe ratio (correlation coefficient > 0.85). The number of trades performed over the

100-day out-of-sample period varies substantially across the different stocks, ranging from an average of 5 trades a day (BLT and AZN) up to 96 trades a day (LLOY). When we examine the number of trades in relation to returns, we can identify that in a number of stocks the algorithm obtained high excess returns in spite of lower profit per trade (e.g. ANTO, SKY, REX, BP and LLOY). This is a typical outcome of HFT, whereby higher overall profits are obtained from lower profits per trade but increased trading frequency. Contrary to Kearns et al. (2010) who claim the absence of profitability in HFT, our findings support the claims of Schulmeister (2009) who identified pockets of profitability in shorter time windows, in our case in the 2-minute to 10-minute range. Our results also validate the theoretical claims of Zhang (2010), Rechenthin and Street (2013) and Brogaard et al. (2014) regarding the possible market efficiency breakdowns in the high-frequency range.

Next we try to validate our ANFIS results against those obtained by the technical indicators alone without the use of an AI model. From the results (Table 5.4) we see that the technical indicators manage to return positive results in 9 out of 27 stocks (33%). For some stocks (e.g. BP, SGE and ITV) the technical indicators manage to obtain excellent results. However, the only stock which which exhibits better performance than the ANFIS model is IAP. These results further support the effectiveness of our proposed ANFIS model and reinforce the claims of Gradojevic and Gençay (2013) that the combination of technical rules with T1 fuzzy logic is superior to technical rules on their own.

Once we have established the superiority of T1 ANFIS on standard technical rules, next we explore any additional performance gains that can be obtained from our proposed IT2 TSK models, namely ANFIS/T2a and ANFIS/T2b. The two models are tested on the same 100-day out-of-sample period used in the benchmark models. From our performance comparison of the ANFIS/T2a model against standard ANFIS, the results in Table 5.5 show Sharpe ratio improvements in 19 out of 27 stocks (70%). For indicative purposes, we note that ANFIS/T2a shows an average percentage increase in Sharpe ratio of 2.91%. In terms of

Table 5.4: Technical indicators performance after the 100-day out-of-sample period.

	Technical Indicators	
Stock	**Sharpe Ratio**	**Excess Return**
AL	-0.0580	-0.60%
ANTO	-0.5073	-7.61%
BLT	0.1197	2.64%
BA	-0.0683	-0.52%
BLND	-0.5771	-8.27%
SKY	0.0845	1.09%
CWC	0.4425	18.91%
AV	-0.0100	-0.08%
DGE	-0.0545	-0.79%
SDR	-0.0375	-0.98%
HMSO	-0.1206	-2.91%
REX	-0.1565	-2.14%
JMAT	-0.0663	-1.30%
HSBA	-0.0003	-0.03%
ABF	-0.0151	-0.16%
INTU	-0.1345	-3.22%
RIO	-0.0402	-0.66%
BP	0.1690	2.04%
SGE	0.6337	21.28%
LLOY	0.1057	2.10%
TSCO	-0.2261	-2.90%
GSK	-0.1814	-3.61%
AZN	-0.9707	-36.21%
HBOS	0.0329	0.38%
XTA	-0.3866	-21.56%
IAP	0.0474	1.01%
ITV	0.7553	42.32%
Average	-0.0452	-0.07%

profitability, it can be noted that 17 out of 27 (63%) stocks show higher profitability. We note that when considering the portfolio of 27 stocks the average excess return per stock increases by 0.46 p.p. Considering that portfolios of financial institutions typically hold thousands of equities and millions in investments, minor increases in profitability can translate into significant monetary value. More importantly, this increase in profitability is achieved in

Table 5.5: ANFIS/T2a and ANFIS/T2b performance improvements against standard ANFIS after the 100-day out-of-sample period. The measures indicate variations from the results produced by ANFIS in Table 5.3. The Sharpe ratio is presented as a percentage difference. Excess return column shows the percentage point (p.p.) differences.

	Sharpe Ratio Improvements (%)		Excess Return Improvements (p.p.)	
Stock	**ANFIS/T2a**	**ANFIS/T2b**	**ANFIS/T2a**	**ANFIS/T2b**
AL	+0.12%	+0.88%	-1.14%	-0.47%
ANTO	+8.14%	-10.54%	+0.62%	-2.15%
BLT	+7.76%	+1.02%	+0.59%	-0.11%
BA	-0.05%	-0.15%	-3.02%	-2.34%
BLND	+3.57%	+10.15%	+0.49%	+0.69%
SKY	+3.51%	+1.45%	+0.46%	+0.12%
CWC	+0.73%	-2.15%	-1.22%	-0.53%
AV	+16.20%	+23.26%	+0.72%	+1.04%
DGE	+3.99%	+3.73%	+0.17%	+0.16%
SDR	-6.19%	+2.01%	-0.50%	+0.91%
HMSO	+12.71%	+15.18%	+1.49%	+1.94%
REX	+3.20%	+4.58%	-0.12%	+0.25%
JMAT	+21.23%	+47.64%	+0.07%	+0.14%
HSBA	-0.25%	-0.63%	-0.14%	-0.09%
ABF	-8.63%	-7.75%	-0.37%	-0.10%
INTU	+6.22%	+3.91%	+0.48%	+0.32%
RIO	+24.37%	+23.56%	+0.71%	+0.65%
BP	-2.88%	-1.18%	+2.93%	+1.95%
SGE	-1.77%	-1.45%	-5.37%	-5.36%
LLOY	+22.08%	-4.83%	+6.56%	+0.05%
TSCO	+0.49%	+0.40%	+0.32%	+0.31%
GSK	+1.27%	-1.53%	-0.08%	-0.07%
AZN	-9.13%	+6.85%	-0.02%	-0.19%
HBOS	+4.63%	+2.35%	+1.64%	+1.35%
XTA	+19.57%	-7.02%	+0.84%	-0.51%
IAP	-3.43%	+7.79%	+0.31%	+0.51%
ITV	+0.09%	+2.19%	+6.04%	+8.04%
Average	+2.91%	+1.43%	+0.46%	+0.24%

conjunction with higher a Sharpe ratio (lower risk). These results provide a clear indication of the improved model optimisations of our proposed ANFIS/T2a over standard ANFIS.

The performance of ANFIS/T2a is further assessed by investigating the summary statistics in Table A.8. Similar to Sharpe ratio, an increase in Sortino ratio is observed in the majority of stocks, indicating similar rankings in terms of performance across the 27 stocks. Another advantage of ANFIS/T2a is the fact that improved performance is obtained with a lower number of trades (on average, 3.12% less trades).

Holmberg et al. (2013) claim that increased risk-adjusted performance can result from increased signal filtering (reduced trades) but at the cost of reduced overall profitability. On the contrary, our results indicate that ANFIS/T2a shows lower overall trading activity but more efficient capital allocation by instigating higher quality trades. This is further confirmed by the improvement in profit per trade where ANFIST2/a establishes a better result in 20 stocks (74%).

As a next step, we compare the performance of ANFIS/T2b against standard ANFIS. The results are displayed in Table 5.5. The indications are that although, in general, ANFIS/T2b performed better than standard ANFIS, the increased overall performance was slightly less than that obtained by ANFIS/T2a. This is also demonstrated from the summary statistics in Table A.8 which show a lower average improvement in terms of the Sortino ratio; however, at a slightly improved average return per stock. As in the case of ANFIS/T2a, ANFIST2b achieves these improvements at a lower average number of trades, indicating that the model generates higher quality trades than the standard ANFIS. This also conveys an important message for model designers. Increased model complexity, as in the case of ANFIS/T2b when compared to ANFIS/T2a, does not guarantee a better risk-adjusted performance. Moreover, it provides an indication that our approach of adding incremental levels of model complexity results in identifying the best balance between complexity and risk-adjusted performance.

In the next section we try to analyse further the performance of our proposed models by examining their performance at different levels of trading frequency (uncertainty).

5.4.2 Experiment 2: Comparison of T1 and T2 models under different trading frequencies

In our second set of experiments, we investigate the performance of standard ANFIS and the proposed IT2 TSK models under increasing levels of uncertainty. As described in Section 5.2.3, we propose an innovative method to control the level of uncertainty using the signal threshold. We test fixed thresholds starting from 0.08% down to 0.02% in steps of 0.02%. For each threshold level, we apply the model configuration that performed best at the respective trading frequency for the respective model category and stock in the in-sample period. The models are then tested on the 100-day out-of-sample period.

In Table 5.6 we present the Sharpe ratio results for the Standard ANFIS model at decreasing levels of return threshold. When comparing the highest threshold column (0.08%) with the lowest one (0.02%) one can immediately notice that the majority of stocks experience a reduction in Sharpe ratio. This is also reflected by the average Sharpe ratio which indicates decreasing levels of Sharpe ratio. At the highest threshold (0.08%), standard ANFIS shows 2 negative results (SDR and IAP). This increases to 5 negative results at the 0.06% and 0.04% threshold level. A further increase to 9 negative results is shown at the 0.02% threshold level. This pattern reflects increasing levels of risk (uncertainty) in line with decreasing thresholds (higher frequency trading).

When we analyse the corresponding excess return results (Table 5.7), we identify that in general there is an increasing trend in excess return in line with decreasing return thresholds. Hence, in general, we see an average increase in excess return levels in line with decreasing average Sharpe ratios (higher risk). However, we note a steep drop in the average excess return at the 0.02% level. We attribute this drop to the fact that at the 0.04% level, the return threshold is equivalent to our transaction costs, hence making it much more difficult to obtain profitable trades beyond this level.

Table 5.6: ANFIS Sharpe ratio performance after 100-day out-of-sample period across different levels of return threshold (uncertainty).

	Standard ANFIS - Sharpe Ratio			
	Return Threshold			
Stock	**0.08%**	**0.06%**	**0.04%**	**0.02%**
AL	0.8943	0.9422	0.9085	0.8415
ANTO	0.3084	0.2418	0.3226	0.1276
BLT	0.2733	0.2276	0.1289	-0.1943
BA	1.5265	1.4034	1.3728	1.1116
BLND	0.4087	0.2144	0.2160	-0.1202
SKY	0.5581	0.4960	0.6487	0.3826
CWC	1.4663	1.2860	1.4276	1.3846
AV	0.3444	0.3982	0.1779	0.0229
DGE	0.2879	0.3133	0.2292	0.1634
SDR	-0.1244	-0.1266	-0.0386	-0.1943
HMSO	0.4375	0.4875	0.3173	0.1160
REX	0.4820	0.5606	0.7213	0.7067
JMAT	0.0212	-0.0452	-0.1119	-0.2420
HSBA	0.3579	0.2366	0.3109	0.0140
ABF	0.2272	0.2370	0.2418	0.5243
INTU	0.5033	0.5597	0.5552	0.1352
RIO	0.1477	0.0814	-0.1090	-0.3973
BP	0.7549	0.7708	0.9238	0.7528
SGE	1.4761	1.6455	1.3508	1.3498
LLOY	0.5383	0.6492	0.7065	0.6103
TSCO	0.9606	1.1907	1.5164	1.2506
GSK	0.1897	0.2111	0.1848	-0.0128
AZN	0.0219	-0.0535	-0.2337	-0.2201
HBOS	0.4554	0.4439	0.3611	0.2676
XTA	0.0941	-0.0463	-0.1996	-0.2536
IAP	-0.0338	-0.0321	0.0274	-0.0378
ITV	1.3587	1.3787	1.3290	1.2143
Average	0.5162	0.5064	0.4921	0.3446

From further analysis we identify the increase in the number of trades in line with decreasing return thresholds (see Table A.10). For example, in the case of ANTO, the number of trades increases from an average of 20 trades per day at a threshold of 0.08% up to 143 trades per day at a threshold of 0.02%. However from Table A.11 we note that the profit per trade reduces drastically in line with the reduction in return threshold. The pattern clearly

Table 5.7: ANFIS excess return performance after 100-day out-of-sample period across different levels of return threshold (uncertainty).

	Standard ANFIS - Excess Return			
	Return Threshold			
Stock	**0.08%**	**0.06%**	**0.04%**	**0.02%**
AL	1.1695	1.3803	1.7305	1.0871
ANTO	0.1391	0.1470	0.3393	0.1707
BLT	0.0763	0.0571	0.0404	-0.1361
BA	0.8458	1.0141	1.2517	1.3441
BLND	0.0969	0.0678	0.1189	-0.1031
SKY	0.1257	0.1530	0.2664	0.2651
CWC	1.2561	1.4129	1.5165	1.5790
AV	0.0681	0.1033	0.0681	0.0180
DGE	0.0543	0.0636	0.0669	0.0845
SDR	-0.0263	-0.0278	-0.0146	-0.1583
HMSO	0.1575	0.3096	0.2833	0.1696
REX	0.1103	0.2986	0.3509	0.5438
JMAT	0.0034	-0.0094	-0.0385	-0.1633
HSBA	0.0500	0.0593	0.0919	0.0066
ABF	0.0490	0.0761	0.1100	0.3294
INTU	0.1491	0.2482	0.3508	0.1659
RIO	0.0359	0.0254	-0.0439	-0.2720
BP	0.1993	0.2486	0.4141	0.4372
SGE	0.8065	0.9974	1.1909	1.0658
LLOY	0.2438	0.4175	0.6582	0.6959
TSCO	0.1825	0.4629	0.6086	0.6679
GSK	0.0426	0.0554	0.0704	-0.0085
AZN	0.0023	-0.0085	-0.0582	-0.1336
HBOS	0.2493	0.3828	0.4772	0.3769
XTA	0.0457	-0.0259	-0.1182	-0.2611
IAP	-0.0171	-0.0203	0.0179	-0.0712
ITV	0.7084	0.6306	0.9826	0.5783
Average	0.2527	0.3155	0.3975	0.3066

shows that lower return thresholds instigate higher number of trades but at lower profit per trade. Our results indicate that this can lead to an increase in overall profitability, however it is important to note that due to the lower profit per trade this brings a more difficult challenge to obtain positive results.

Next, we utilise these results to benchmark the standard ANFIS against the proposed T2 models. From Table 5.8 we note that in comparison to standard ANFIS, the proposed IT2 models show increasing improvements in the Sharpe ratio, the average excess return per stock and the average profit per trade. In the case of ANFIS/T2a, the Sharpe ratio, improvements range from an increase of 1.85% at the 0.08% threshold level, up to 10.48% at the 0.02% level. This also indicates that although, as expected, there is a decreasing trend in the Sharpe ratio in line with lower thresholds (higher frequency trading), however the decrease rate in the case of ANFIS/T2a is lower than that in the ANFIS model. Improvements in the average excess return per stock range from 0.2 p.p. at the 0.08% threshold level, up to 2.24 p.p. at the 0.02% level. A similar trend is obtained in the average profit per trade. In the case of ANFIS/T2b model, similar outperformance against ANFIS is observed, although to a lesser degree. The relative increase in all three measures, indicates the superior capabilities of the proposed ANFIS/T2 models when compared to standard ANFIS. The pattern in results indicates that this relative increase in performance gets more pronounced at lower thresholds which experience higher effects of uncertainty. This is due to the increasing exposure of the technical signals input features to smaller price fluctuations and possible noise.

As a final step, we validate our results using statistical tests. The tests are carried out on the average Sharpe ratio, the excess return per stock and the average profit per trade using a paired t-test on the results obtained from the 27 stocks. Table 5.8 shows that the difference in performance results at the lower return thresholds (0.04% and 0.02%) are all significant. The tests strengthen our earlier claims of the improved performance of our proposed ANFIS/T2 models against standard ANFIS models at levels of higher uncertainty. This makes our proposed models more suitable contenders for HFT environments. At the higher return thresholds (0.08% and 0.06%), both models experience some insignificant measures due to lower improvements against the standard ANFIS results. This indicates that at reduced uncertainty, the introduction of IT2 fuzzy sets has less effect on trading performance.

Table 5.8: Summary statistics for standard ANFIS and the corresponding variations in the ANFIS/T2 models. Results are obtained over a 100-day out-of-sample period across different levels of return threshold (uncertainty). Bold figures for performance measures average Sharpe ratio, average excess return/stock and average profit/trade indicate a rejected paired-sample t-test. The test applies the null hypothesis that the difference in results (ANFIS vs. ANFIS/T2) comes from a normal distribution with mean equal to zero and unknown variance at 5% sig. level.

	Return threshold level			
Measure	**0.08%**	**0.06%**	**0.04%**	**0.02%**
Standard ANFIS				
Average number of trades	1967	3130	5539	8742
Average Sharpe Ratio	0.5162	0.5064	0.4921	0.3446
Average Excess Return / Stock	25.27%	31.55%	39.75%	30.66%
Average Profit / Trade (£)	32.12	25.20	17.94	8.77
ANFIS/T2a Improvement				
Average number of trades	−2.85%	−2.73%	−2.48%	−1.15%
Average Sharpe Ratio	+1.85%	**+4.97 %**	**+5.70%**	**+10.48%**
Average Excess Return / Stock (p.p.)	+0.20	**+1.10**	**+1.03**	**+2.24**
Average Profit/ Trade (£)	**+3.75%**	**+6.39%**	**+5.03%**	**+8.55%**
ANFIS/T2b Improvement				
Average number of trades	−2.16%	−2.13%	−1.88%	−1.01%
Average Sharpe Ratio	+0.78%	**+3.77%**	**+2.44%**	**+7.13%**
Average Excess Return / Stock (p.p)	−0.15	+0.72	+0.43	**+1.88**
Average Profit / Trade (£)	**+1.62%**	**+4.51%**	**+3.01%**	**+7.21%**

5.5 Conclusion

In this chapter we have investigated the risk-adjusted performance of neuro-fuzzy models in conjunction with technical indicators from a higher frequency perspective. In particular we extend the research conducted by Gradojevic and Gençay (2013) who highlight the uncertainties resulting from technical trading rules. The challenge becomes more difficult at higher trading frequencies due to smaller, possibly technically induced, price movements.

We address this challenge by consolidating our research findings from previous chapters and extending this to new proposed models based on T2 fuzzy logic. As our first contribution, we present two innovative and practical methods of how the popular ANFIS model, a popular

AI technique in finance, can be improved by introducing IT2 fuzzy sets with a minimal increase in complexity. Our approach addresses the concerns of increased complexity that is typically attributed to the design of T2 fuzzy models, possibly limiting wider-spread adoption (Wu and Mendel, 2014). The proposed T2 models in this chapter also present a stepping stone from our ANFIS findings presented in Chapter 3. The main benefit is to minimise the effect of uncertainty caused by technical signals, hence reducing the overall risk. By extending the clustering data partitioning ideas that we use in Chapter 4, especially in view of different intraday volatility levels, here we apply a similar approach for model structure identification which is also used as a basis for our IT2 extensions. This approach results in more compact and efficient fuzzy models.

As our second contribution, we introduce T2 fuzzy logic models to high-frequency trading and investigate the effects of better management of technical signals on the profitability and risk-adjusted performance of trading algorithms. To our knowledge, this is the first time that T2 fuzzy logic has been used for this purpose. By utilising a combination of technical signals on 2-minute returns we particularly focus on holding periods ranging from 2 to 10 minutes, and extend our market efficiency investigations that we present in Chapters 3 and 4 to a more microscopic perspective. Our results show a significant increase risk-adjusted trading performance when compared to the popular T1 ANFIS. As a result of this, from our results we manage to identify a positive link between higher order fuzzy systems and risk-adjusted trading performance at short time horizons.

Finally, as our third contribution, although a number of authors (e.g. Aladi et al., 2014; Sepulveda et al., 2006) demonstrate the increased capability of IT2 models to handle increased uncertainty when compared to T1, we address the current literature gap of *when* can the introduction of T2 offer *feasible* improvements in profitability and risk-adjusted performance of T1 models. We provide deeper insight on the benefits of adopting IT2 models from the perspective of different levels of trading risk (uncertainty) and trading frequency. We conclude that the introduction of T2 fuzzy sets (in contrast to the models presented in Chapters 3 and 4) exhibit the highest tangible benefits in trading models seeking possible short intraday price trends, making our models more ideal (and feasible) for HFT environments. Less significant improvements are observed in lower-frequency trading scenarios.

Chapter 6

Conclusion

In this thesis we have proposed a number of fuzzy logic techniques applied to algorithmic and high-frequency trading. Our point of departure has originated from two stronghold perspectives in finance. Firstly, while the literature in finance and economics stands firm on the Efficient Market Hypothesis (EMH), many traders continue to make buy and sell decisions based on historical data. Whilst the debate continues, this thesis has stimulated discussion against a strict adherence to the EMH. Our interest has been in challenging the EMH in its weak form, and in particular our focus has concentrated on the application of technical trading rules (in conjunction with fuzzy logic techniques) at short intraday time horizons. The second perspective is that investors typically look beyond profits when evaluating trading strategies. The reason is that brilliant results could mask the fact that an algorithm might carry huge risks to achieve the claimed performance. From our reviewed literature we have identified that studies that investigate fuzzy logic models from the perspective of risk-adjusted performance in intraday trading scenarios are scarce.

These arguments have steered our research to identify innovative techniques directed towards intraday and neuro-fuzzy controlled trading algorithms with the objective to improve risk-adjusted performance. Our research falls in line with the upcoming regulatory regimes like MiFID II that increase the responsibility on algorithm designers to implement adequate risk controls. This chapter provides a summary of the presented work, lists the main contributions and discusses limitations and future research.

6.1 Summary of presented work

First, in Chapter 1, we have brought together the literature review from three knowledge domains, namely high-frequency trading, financial risk and artificial intelligence, with a special focus on fuzzy logic techniques. Our review has centred on highlighting key intersecting concepts from each domain, with the intention to present literature gaps and put forward arguments from which the research objectives of this thesis have been derived. Furthermore, as a precursor to the fuzzy logic techniques which we have used extensively in subsequent chapters, in Chapter 2 we have provided the necessary mathematical ground work on fuzzy logic with a specific focus on Type-1 Takagi-Sugeno-Kang (TSK) FLSs and Interval Type-2 (IT2) FLSs.

Following the introductory chapters, in Chapter 3 we have investigated a number of fuzzy logic modelling aspects that have an impact on risk-adjusted trading performance. In addition, we have explored the debated profitability of moving average rules using high-frequency data in an intraday trading scenario. In our case, we have investigated holding times of trading positions in the region between 10 minutes to one hour. We have tested a range of return bands in the region between 0.1% and 0.5%. These bands were applied to act as a threshold for unprofitable small trades. For our experiments, we have utilised a data set consisting of stock trades effected on the London Stock Exchange. We have investigated the performance attained from three representative milestone models in neurocomputing, namely Neural Network (NN), Adaptive Neuro-Fuzzy Inference System (ANFIS) and Dynamic Evolving Neuro-Fuzzy Inference System (DENFIS) models, and have also introduced the more sophisticated eANFIS architecture. In our experiments, we have tested the trading performance using different model optimisation functions, such as single risk-return functions, an innovative combination of different risk-return functions via an ensemble, Root Mean Squared Error (RMSE), period return and models optimised without considering transaction costs. The stability of our proposed models has been analysed in greater depth by comparing the time series (performance profile) of performance measures over the full out-of-sample period.

Following the risk-adjusted performance enhancements of the fuzzy logic models presented in the Chapter 3, in Chapter 4 we have explored further improvements by introducing a method on how to effectively adapt trading algorithms to the different market states (risk scenarios) that evolve during a typical trading day. This problem has been addressed in two parts. In Part 1 (Section 4.2) we have investigated whether fuzzy logic, in conjunction with intraday realised volatility as a proxy for uncertainty, can improve the risk-adjusted performance of standard NN trading algorithms. In Part 2 (Section 4.3), we have extended our research to explore whether the combination of fuzzy logic with intraday volatility information can improve risk-based money management decisions. The main challenge has been to identify a method which dynamically adjusts trading frequency depending on the varying degrees of risk at an intraday level. Another challenge that we have addressed in this chapter has been to optimise capital allocation decisions based on preferable intraday trading scenarios.

Finally, in Chapter 5, we have investigated the profitability and the risk-adjusted performance of trading algorithms at higher trading frequencies by addressing a key challenge presented by the increased uncertainty of technical indicators at very short time intervals. From the AI literature, T2 fuzzy logic has been identified to offer better management of uncertainty than T1. To this aim, we have extended the T1 techniques explored in Chapters 3 and 4 by seeking better performance from higher order fuzzy systems to handle the increased uncertainties inherent in an HFT scenario. In our approach, we have introduced practical methods of how the popular ANFIS T1 model can be generalised to an IT2 TSK fuzzy system. We have been conscious about the criticism of increased complexity that is normally attributed to T2 models. Consequently, we have addressed this problem by proposing methods which have a minimal impact on the design and computational complexity. In our experiments, we have investigated whether the proposed IT2 FLSs provide a viable alternative for trading purposes in view of improving risk-adjusted performance. As a further step, we have presented a deeper investigation to explore *when* it is more feasible to utilise IT2 FLSs in comparison to T1 models. We have addressed this question at different levels of intraday risk and trading frequencies.

6.2 Contributions

In this thesis, we have put forth a number of contributions. Within this section, we consolidate the main contributions under four different perspectives:

1. **The role of fuzzy logic in algorithmic trading**
   From this perspective, as our first contribution in Chapter 3 we have introduced fuzzy logic to uncertainty management in algorithmic and high-frequency trading. To our best knowledge, the relationship between fuzzy logic and risk-adjusted performance in intraday trading scenarios has never been investigated. We have examined the effectiveness of three milestone models from neuro-fuzzy literature, namely NN, ANFIS and DENFIS. We have identified that although NNs live up to their popularity in terms of performance, their combination with fuzzy logic using ANFIS provides a superior risk-adjusted trading performance. We have also demonstrated the outperformance of ANFIS when compared against DENFIS. In addition, we have validated the effectiveness of ANFIS against standard technical rules and a random model. We have also introduced an innovative ANFIS ensemble model (eANFIS) which combines and dynamically selects models which are optimised using different risk-return objective

functions. On the one hand, the model switches to riskier risk-adjusted objective functions during intraday periods of high performance. This avoids the possible adoption of an overly risk-averse model with the possibility of increased penalisation in profitability. On the other hand, the model dynamically reverts back to more risk-sensitive objective functions during unfavourable time windows, hence reducing the possibility of large losses. Our results show that the proposed eANFIS model outperforms the trading performance of standard neural network, ANFIS and DENFIS.

As our second contribution, in Chapter 4, we have extended our research by proposing fuzzy logic to improve money management decisions in intraday algorithmic trading. This goes contrary to many studies that suggest trading rules based on a fixed position sizing strategy, fixed return thresholds and fixed volatility thresholds. We have demonstrated the applicability of a fuzzy logic approach by presenting a hybrid fuzzy model as an extension to a popular neural network trading model. Core to our approach is a fuzzy clustering method that automatically identifies preferable pockets of intraday risk-adjusted profitability at different trend and volatility levels. Our innovative hybrid model dynamically adapts trading frequency and order position sizing decisions across intraday trend (profit) and volatility (risk) states to improve the overall risk-adjusted performance. We demonstrate that the presented fuzzy logic money management model applied to NN results in a more optimal solution than standard NN using fixed position sizing, hence increasing risk-adjusted performance without incurring a penalty in profitability. We have also validated the effectiveness of our hybrid model against a random model.

Finally, as our third contribution, in Chapter 5 we have extended our T1 fuzzy logic models by introducing two innovative and practical methods of how the popular ANFIS model in conjunction with technical signals can be improved by shifting from T1 to IT2 fuzzy sets. Although a number of existing studies claim the improved management of uncertainty of T2 models when compared to T1, to our knowledge we do not know about any studies that examine if and when can T2 models translate into enhanced risk-adjusted performance of intraday trading algorithms. Our main hypothesis is that minimising the effect of technical rules uncertainty reduces the overall risk. Since a number of authors argue that T2 models suffer from increased complexity, especially in the design process, we have proposed a simple, yet effective, approach of how the popular ANFIS model can be extended to IT2. The approach that we have suggested also promotes the design of more compact and efficient fuzzy models. We have demonstrated that the proposed T2 methods offer additional capabilities that result in a significant increase in trading performance during out-of-sample testing when compared to standard T1 ANFIS. Moreover, we have provided deeper insight into the benefits of adopting IT2 models from the perspective of different levels of trading risk (uncertainty) and trading frequency. From our results we have concluded that the introduction of IT2 fuzzy sets exhibit the highest tangible (investor) benefits at higher trading frequencies. Less significant benefits were observed at lower trading frequencies.

2. **Market efficiency during short intraday time windows**

   From this perspective, in Chapter 3 we have presented our first contribution in view of our investigation on the predictive performance of technical trading rules and subsequently the risk-adjusted performance of trading algorithms based on the generated signals. From our knowledge, the link between the profitability of intraday fuzzy logic controlled algorithms and the ways of improving risk-adjusted performance has never been investigated before. In our approach, we have applied fuzzy logic techniques to dynamically tune the trend signals based on the current intraday market speed. We have demonstrated a simple yet effective extension of common technical rules by considering a dynamic amalgamation of moving average signals. We have established that the proposed dynamic moving approach controlled using neuro-fuzzy models outperforms the risk-adjusted performance obtained from the equivalent standard technical rules. Our experiments indicate positive trading results on the majority of stocks, taking into consideration a risk-free rate of 5%. By applying heat maps, we have also identified concentrated regions of a higher Sharpe ratio in specific areas of holding position times and return bands. Our analysis has indicated that contrary to a number of claims in the literature, technical rules do manage to identify pockets of profitability in the high-frequency range, with holding periods of between 10 minutes to one hour. Of particular interest is the fact that for specific stocks, the heat maps have identified more than one region of profitability, hence the proposed method can provide a clearer indication to traders on the possible profitable trading strategies.

As our second contribution, in Chapter 4 we have examined the relationship between intraday volatility and the profitability and risk-adjusted performance of trading algorithms. Although a number of authors claim a relationship between the possible periodic breakdown of market efficiency and volatility, our originality resided in investigating these theoretical claims at a more granular intraday level and shorter intraday time horizons. In our models we have borrowed the concepts of realised volatility from economics literature and have combined this information with trading models using an innovative fuzzy logic controller approach. Although economics literature shows great advancements in volatility measurement up to an intraday level, to our best knowledge the application of this information to improve the trading performance of trading algorithms was never explored. In our experiments we have validated trading model performance optimisations by dynamically discriminating between different intraday trend and volatility states. From our statistical tests, we have identified a link between the profitability of technical rules with different levels of price volatility at short intraday horizons.

As our third contribution, in Chapter 5 we have introduced T2 fuzzy logic to high-frequency trading and, in particular, investigated the profitability of algorithmic trading at higher trading frequencies. In our experiments we have targeted a trading window of between 2 to 10 minutes. From our literature review, we have not identified any previous research which investigates risk-adjusted performance at this time interval. We have demonstrated that profitability in this range is possible and managed to identify a relationship between higher order fuzzy systems and risk-adjusted trading performance.

3. **Model design and evaluation for trading purposes**
From a model design perspective, as our first contribution we have presented a coherent approach for optimising neuro-fuzzy controlled trading algorithms in line with investors' risk concerns. To our best knowledge, an investigation on how to improve risk-adjusted performance in an intraday high-frequency setting using fuzzy logic has not been studied before. In Chapter 3, we have identified that neuro-fuzzy models offer superior out-of-sample trading performance when optimised using risk-adjusted performance functions. Although NN, ANFIS and DENFIS share commonalities in their model structures, the best performance was achieved by the ANFIS architecture. This was demonstrated both in case of the Sharpe ratio and the Sortino ratio model optimisations. From our results we concluded that when the models are optimised using non risk-adjusted measures like the win ratio and profitability perform poorly out-of-sample. Similarly, our results have provided clear indications that trading models based on the commonly used measures in the AI literature, such as RMSE, also lead to weak trading results. Our experiments have also identified that not considering transaction costs during model training can lead to biased results in real-world applications. All the above results offer a coherent framework for designers of fuzzy logic models that can improve risk-adjusted trading performance.

As our second contribution, in Chapter 3 we have introduced the application of the cumulative Sharpe ratio to evaluate the performance stability, and hence establish a performance profile, of the proposed fuzzy logic models across the full out-of-sample period (rather than just at the end of the out-of-sample period). We have demonstrated that this approach provides a deeper understanding of the time-varying performance profile of the applied neuro-fuzzy models. From our results we have indicated that ANFIS provides the most stable trading performance when compared to the NN model and DENFIS. Although in general the NN performed well, however we have identified more variability (less stability) in the results.

As our third contribution, we have investigated innovative fuzzy logic methods that permit trading models to dynamically adapt across various levels of market risk. To our best knowledge, the risk-adjusted performance benefits resulting from the dynamic adaptiveness of fuzzy logic controlled algorithms across intraday risk levels was never explored. In Chapter 3 and Chapter 4 we have introduced fuzzy logic methods that are capable of such adaptiveness. In both scenarios we have demonstrated improvements in both profitability and risk-adjusted performance when compared to the more rigid approaches that are common in the literature. Additionally, in Chapter 5 we have demonstrated that adopting models that can dynamically adapt to risk stemming from technical rules uncertainty, especially when trading at higher frequencies, results in additional trading improvements.

4. **Management insights**

   Overall, in our thesis we have also brought about contributions from a financial management perspective. As our first contribution, our results have indicated important outcomes that are of utmost interest for decision makers and also encourage further research and investment by firms employing algorithmic trading which will be impacted by new regulatory regimes, such as MiFID II. Albeit the new directives and regulations stress the demand that the employed trading systems must meet numerous new requirements, particularly when it comes to risk controls, the literature on *how* fuzzy logic can assist in this regard is still scarce. In our thesis, we have presented numerous (tangible) methods that demonstrate how fuzzy logic can be applied to improve the risk-adjusted performance and performance stability of trading algorithms and hence can assist in addressing these new requirements.

As our second contribution, in Chapters 3, 4 and 5 we have introduced a number of innovative fuzzy logic techniques of how existing algorithmic and high-frequency trading models can be improved by increasing risk-adjusted performance. Our proposed models demonstrate a stepwise approach by starting from the popular models used in finance literature, such as NNs and ANFIS, and show how by introducing incremental components on the base AI models (rather than a whole overhaul of existing investment), these can be enhanced to meet improved risk objectives.

6.3 Limitations and future work

In conclusion, this thesis opens up a number of avenues for further research. As a start, we propose future research paths that can be followed from a fuzzy logic perspective. In Chapters 3 and 4, we have identified increased model stability and risk-adjusted improvements by introducing T1 models. In Chapter 5, further improvements have been demonstrated in higher frequency bands with the introduction of IT2 fuzzy sets and adaptive FOU sizes. In our approach, we have adopted incremental enhancements on popular models in neuro-fuzzy literature with an application to finance. This choice of approach has been consciously followed in order for the proposed models to remain practical to use and easier to compare by keeping the additional parameters to a minimum. Our first research proposition is that we do not exclude the possibility of further improvements (in the management of uncertainty and hence trading performance) by exploring additional incremental steps in model architecture complexity, alternative rule extraction methods, different membership functions, and more complex T2 rules or defuzzification methods.

In this thesis, we have primarily applied Sugeno fuzzy models. In comparison to Mamdani fuzzy models, Sugeno models are known to provide higher accuracy but they do this at the cost of lower interpretability. In our case, the Sugeno approach has been selected since it acts as a more natural extension to NNs, the latter acting as a popular and strong benchmark in finance. However, we appreciate that in line with additional model complexity, this can mitigate wider adoption. An interesting research path can seek to mitigate this risk by investigating the performance of similar Mamdani models which tend to increase model interpretability. To meet this aim, the exploration might also be extended to the combination of fuzzy logic models with more interpretable machine learning techniques such as decision trees.

In terms of model optimisation, we have limited our approach to gradient based optimisation methods. Apart from being one of the classical optimisation methods in use in the neural network field, the approach is popularly extended to neuro-fuzzy models, thus enabling any model improvements to be attributed as much as possible to the fuzzy logic structures. However, we note that optimisation problems might have multiple solutions, some of which might be local optima. Repeated random trials, random initialisation and testing were applied in our experiments to avoid the possible problem of local optima. However, we direct future research to consider the use of advanced heuristic based optimisation methods that can offer global optimisation methods with more flexible application to complex optimisation functions. This can potentially identify new avenues for further improvement in trading performance. These algorithms typically fall under three categories, namely simulated annealing, traditional genetic algorithm and evolutionary algorithms. We refer the reader to Wu and Mendel (2014) and Mendel et al. (2014) for practical methods of applying these methods for fuzzy model parameter tuning.

We also identify a number of research avenues from a finance perspective. Firstly, in our experiments we have limited our data set to trade data from a list of 27 stocks listed on the London Stock Exchange. Albeit the findings from this thesis provide a clear direction in terms of fuzzy logic techniques that can lead to trading benefits, however, further research can be expanded to include additional portfolios and a variety of financial instruments and markets (including forex). This can help to validate further the contributions of fuzzy logic techniques in improving risk-adjusted trading performance and increase the viable applicability on a wider scale.

Thirdly, in this thesis we have adopted the Sharpe ratio as a key risk-adjusted performance measure. We have also compared our model performance on each stock against the risk-free rate. Two limitations of the Sharpe ratio are noted. Firstly, the Sharpe ratio does not separate between variability in gains and losses, hence it attributes penalisation to both upside and downside variability. This might not represent the interest of investors who would rather welcome positive variability in gains. In this thesis, however, we favour model stability and hence our interest lies more in identifying algorithmic trading models that can offer steady returns. Secondly, Lo (2002) warns that the Sharpe ratio highly depends on the distribution of the underlying returns. For example, in the case of positively skewed return distributions, this might result in risk overestimation. Conversely, in the case of negatively skewed return distributions, this can result in risk underestimation. Excess kurtosis can also be problematic for the ratio as standard deviation doesn't have the same effectiveness when this problem exists. However, a number of investigations in previous studies (Auer and Schuhmacher, 2013; Eling, 2008; Prokop, 2012) show that despite it's weaknesses with abnormality, Sharpe ratio measures tend to indicate performance rankings that are positively related to those obtained from more sophisticated performance ratios. In our case, this is also reflected in the similar performance rankings obtained using the Sharpe ratio with those obtained with other risk-adjusted measures, specifically the Sortino ratio and the Calmar ratio (this includes the models which are optimised using Sortino ratio in Chapter 3). Although the applied measures indicate clear rankings amongst the tested models, this however does not exclude further research to extend our presented framework to identify any additional improvements that can be gained by considering higher order moments using more sophisticated risk-adjusted measures or their combination (as indicated in the ensemble approach presented in Chapter 3).

Another limitation in this thesis is that we focus on optimising models and compare performances across different stocks independently. Although this investigation at individual stock level is a first step in identifying model components that improve risk-adjusted performance, further risk reduction can be explored by considering portfolio optimisation which does not necessarily reduce the expected return. This opens research opportunities to extend fuzzy logic techniques to portfolio optimisation in the high frequency domain (see Maringer, 2008, and references therein). In conjunction with our comments with regard to the Sharpe ratio mentioned earlier, this research can be extended to consider higher order moments in portfolio selection (e.g. Maringer and Parpas, 2009).

In this thesis we investigate the increased trading performance that can be obtained by combining fuzzy logic with intraday realised volatility information. The superior results obtained encourage more research in this field in view of intraday trading scenarios. In particular, in our study we limit our realised volatility estimator to the average RV. Further research is encouraged to establish whether fuzzy logic techniques in conjunction with other RV measures which are less susceptible to microstructure bias (Andersen et al., 2011; Hansen and Lunde, 2006) as well as abrupt jumps in financial markets (Barndorff-Nielsen and Shephard, 2004) can lead to further risk-adjusted trading improvements. Additionally, we encourage further research to investigate the effect of alternative FOU tuning frequencies, which in this thesis we limit to a daily basis. In particular, any possible beneficial relationship between the FOU tuning frequency and the diurnal patterns of market intraday activity, a stylised fact in finance, remains an open question.

Finally, we have focused on identifying the benefits of fuzzy logic techniques in trade order signalling and also position sizing in conjunction with intraday time-varying risk. Although these are core decisions that need to be tackled by a trader or an automated system, the process typically includes additional decision steps, such as the trade execution method, which are not covered within the scope of this thesis. These decisions can also contribute to the the improvement of risk-adjusted performance. In the case of trade execution, Johnson (2010) states that hundreds of trading algorithms exist in this category (for example, time-weighted average price, percentage of volume and implementation shortfall). However, research on how fuzzy logic can contribute in this area, especially in view of improving risk-adjusted performance and the recent MiFID II execution restrictions, is scarce.

References

Ahmed, N. K., Atiya, A. F., Gayar, N. E., and El-Shishiny, H. (2010). An Empirical Comparison of Machine Learning Models for Time Series Forecasting. *Econometric Reviews*, 29(5-6):594–621.

Aitken, M., Cumming, D., and Zhan, F. (2015). High frequency trading and end-of-day price dislocation. *Journal of Banking & Finance*, 59:330–349.

Aladi, J. H., Wagner, C., and Garibaldi, J. M. (2014). Type-1 or interval type-2 fuzzy logic systems — on the relationship of the amount of uncertainty and FOU size. In *Proceedings of FUZZ-IEEE 2014, Beijing, China*, pages 2360–2367.

Aldridge, I. (2013). *High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems*. John Wiley & Sons, Hoboken, NJ, USA.

Álvarez Díaz, M. (2010). Speculative strategies in the foreign exchange market based on genetic programming predictions. *Applied Financial Economics*, 20(6):465–476.

Alves Portela Santos, A., Carneiro Affonso da Costa, N., and dos Santos Coelho, L. (2007). Computational intelligence approaches and linear models in case studies of forecasting exchange rates. *Expert Systems with Applications*, 33(4):816–823.

Andersen, T., Bollerslev, T., and Diebold, F. (2000a). Great realizations. *Risk*, 13:105–108.

Andersen, T. G. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2):115–158.

Andersen, T. G., Bollerslev, T., and Cai, J. (2000b). Intraday and interday volatility in the japanese stock market. *Journal of International Financial Markets, Institutions and Money*, 10(2):107–130.

Andersen, T. G., Bollerslev, T., and Diebold, F. (2000c). Great realizations. *Risk*, 13:105–108.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96(453):42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2009). Parametric and nonparametric volatility measurement. *Handbook of Financial Econometrics*, 1:67–138.

Andersen, T. G., Bollerslev, T., and Meddahi, N. (2011). Realized volatility forecasting and market microstructure noise. *Journal of Econometrics*, 160(1):220–234.

Anderson, R. M. (2011). Time-varying risk premia. *Journal of Mathematical Economics*, 47(3):253–259.

Anderson, R. M., Eom, K. S., Hahn, S. B., and Park, J.-H. (2013). Autocorrelation and partial price adjustment. *Journal of Empirical Finance*, 24:78–93.

Auer, B. R. and Schuhmacher, F. (2013). Robust evidence on the similarity of sharpe ratio and drawdown-based hedge fund performance rankings. *Journal of international financial markets, institutions and money*, 24:153–165.

Bahrammirzaee, A. (2010). A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications*, 19(8):1165–1195.

Bailey, D. H., Borwein, J. M., de Prado, M. L., and Zhu, Q. J. (2014). Pseudomathematics and financial charlatanism: The effects of backtest over fitting on out-of-sample performance. *Notices of the AMS*, 61(5):458–471.

Bao, D. and Yang, Z. (2008). Intelligent stock trading system by turning point confirming and probabilistic reasoning. *Expert Systems with Applications*, 34(1):620–627.

Barndorff-Nielsen, O. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37.

Barndorff-Nielsen, O. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1):1–30.

Benatar, N., Aickelin, U., and Garibaldi, J. M. (2012). An investigation into the relationship between type-2 FOU size and environmental uncertainty in robotic control. In *Proceedings of FUZZ-IEEE 2012, Brisbane, Australia*, pages 1–8.

Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Norwell, MA, USA: Kluwer Academic Publishers.

Birkin, P. A. and Garibaldi, J. M. (2009). A comparison of type-1 and type-2 fuzzy controllers in a micro-robot context. In *Proceedings of FUZZ-IEEE 2009, Jeju Island, Korea*, pages 1857–1862.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, Hoboken, NJ, USA.

Boyacioglu, M. A. and Avci, D. (2010). An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: the case of the Istanbul stock exchange. *Expert Systems with Applications*, 37(12):7908–7912.

Brabazon, A. and O'Neill, M. (2006). *Biologically Inspired Algorithms for Financial Modelling*. Springer-Verlag, Heidelberg, Berlin, Germany.

Brogaard, J., Hendershott, T., and Riordan, R. (2014). High-frequency trading and price discovery. *Review of Financial Studies*, 27(8):2267–2306.

Buchanan, B. and Wilkins, D. (1993). *Readings in Knowledge Acquisition and Learning: Automating the Construction and Improvement of Expert Systems*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P., and Oliveira, A. L. (2016). Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, 55:194–211.

Chang, J.-R., Wei, L.-Y., and Cheng, C.-H. (2011). A hybrid ANFIS model based on AR and volatility for TAIEX forecasting. *Applied Soft Computing*, 11(1):1388–1395.

Chen, M.-Y. (2013). A hybrid anfis model for business failure prediction utilizing particle swarm optimization and subtractive clustering. *Information Sciences*, 220:180–195.

Chen, S.-M., Lee, S.-H., and Lee, C.-H. (2001). A new method for generating fuzzy rules from numerical data for handling classification problems. *Applied Artificial Intelligence*, 15(7):645–664.

Cheung, W. M. and Kaymak, U. (2007). A fuzzy logic based trading system. In *Proceedings of the Third European Symposium on Nature-inspired Smart Information Systems, St. Julians, Malta*, pages 141–148.

Chiang, J.-H. and Hao, P.-Y. (2004). Support vector learning mechanism for fuzzy rule-based modeling: a new approach. *IEEE Transactions on Fuzzy Systems*, 12(1):1–12.

Choey, M. and Weigend, A. S. (1997). Nonlinear trading models through Sharpe Ratio maximization. *International Journal of Neural Systems*, 8(04):417–431.

Choudhry, T., McGroarty, F., Peng, K., and Wang, S. (2012). High-frequency exchange-rate prediction with an artificial neural network. *Intelligent Systems in Accounting, Finance and Management*, 19(3):170–178.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236.

Covel, M. W. (2009). *The Complete TurtleTrader: How 23 Novice Investors Became Overnight Millionaires*. HarperCollins, NY, USA.

de Faria, E., Albuquerque, M. P., Gonzalez, J., Cavalcante, J., and Albuquerque, M. P. (2009). Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods. *Expert Systems with Applications*, 36(10):12506–12509.

Dempster, M. and Leemans, V. (2006). An automated FX trading system using adaptive reinforcement learning. *Expert Systems with Applications*, 30(3):543–552.

Dutta, B. and Angelov, P. (2010). Clustering as a tool for self-generation of intelligent systems : a survey. In *Proceedings of the International Symposium on Evolving Intelligent Systems, Leicester, UK*, pages 34–41.

Eling, M. (2008). Does the measure matter in the mutual fund industry? *Financial Analysts Journal*, 64(3):54–66.

Engle, R. and Patton, A. J. (2001). What good is a volatility model? *Quantitative Finance*, 1(2):237–245.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007.

Enke, D. and Thawornwong, S. (2005). The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with Applications*, 29(4):927–940.

Esfahanipour, A. and 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.

Evans, C., Pappas, K., and Xhafa, F. (2013). Utilizing artificial neural networks and genetic algorithms to build an algo-trading model for intra-day foreign exchange speculation. *Mathematical and Computer Modelling*, 58(5-6):1249–1266.

Fama, E. (1965). Random Walks in Stock Market Prices. *Financial Analysts Journal*, 21(5):55–59.

Fama, E. (1970). Efficient Capital Markets: A Review Of Theory And Empirical Work. *The Journal of Finance*, 25(2):383–417.

Faulina, R., Lusia, D. A., Otok, B. W., Kuswanto, H., et al. (2012). Ensemble method based on anfis-arima for rainfall prediction. In *Proceedings of the 2012 IEEE International Conference on Statistics in Science, Business, and Engineering (ICSSBE)*, pages 1–4.

Fernandez-Rodrıguez, F., Gonzalez-Martel, C., and Sosvilla-Rivero, S. (2000). On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market. *Economics Letters*, 69(1):89–94.

Foucault, Thierryand Hombert, J. and Roşu, I. (2016). News trading and speed. *The Journal of Finance*, 71(1):335–382.

Gençay, R., Dacorogna, M., Muller, U. A., Pictet, O., and Olsen, R. (2001). *An Introduction to High-frequency Finance*. Academic Press, San Diego, CA, USA.

Gencay, R., Selcuk, F., and Whitcher, B. (2002). *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Academic Press, San Diego, CA, USA.

Gençay, R. (1996). Non-linear prediction of security returns with moving average rules. *Journal of Forecasting*, 15(3):165–174.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 48(5):1779–1801.

Gradojevic, N. (2007). Non-linear, hybrid exchange rate modeling and trading profitability in the foreign exchange market. *Journal of Economic Dynamics and Control*, 31(2):557–574.

Gradojevic, N. and Gençay, R. (2013). Fuzzy logic, trading uncertainty and technical trading. *Journal of Banking & Finance*, 37(2):578–586.

Greenfield, S. and Chiclana, F. (2013). Accuracy and complexity evaluation of defuzzification strategies for the discretised interval type-2 fuzzy set. *International Journal of Approximate Reasoning*, 54(8):1013–1033.

Han, Y., Yang, K., and Zhou, G. (2013). A new anomaly: The cross-sectional profitability of technical analysis. *Journal of Financial and Quantitative Analysis*, 48(05):1433–1461.

Hansen, P. and Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH (1, 1)? *Journal of Applied Econometrics*, 20(7):873–889.

Hansen, P. R. and Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24(2):127–161.

Harris, L. (2002). *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford University Press, NY, USA.

Havens, T. C., Bezdek, J. C., Leckie, C., Hall, L. O., and Palaniswami, M. (2012). Fuzzy c-means for very large data. *IEEE Transactions on Fuzzy Systems*, pages 99:1–1.

Heal, G. and Millner, A. (2014). Reflections uncertainty and decision making in climate change economics. *Review of Environmental Economics and Policy*, 8(1):120–137.

Holmberg, U., Lönnbark, C., and Lundström, C. (2013). Assessing the profitability of intraday opening range breakout strategies. *Finance Research Letters*, 10(1):27–33.

Hudson, R., Dempsey, M., and Keasey, K. (1996). A note on the weak form efficiency of capital markets: The application of simple technical trading rules to UK stock prices-1935 to 1994. *Journal of Banking & Finance*, 20(6):1121–1132.

Jang, J. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics*, 23(3):665–685.

Jang, J.-S. and Sun, C.-T. (1995). Neuro-fuzzy modeling and control. In *Proceedings of the IEEE*, volume 83, pages 378–406.

John, R. and Coupland, S. (2007). Type-2 fuzzy logic: A historical view. *Computational Intelligence Magazine, IEEE*, 2(1):57–62.

Johnson, B. (2010). *Algorithmic Trading & DMA: An introduction to direct access trading strategies*. 4Myeloma Press, London, UK.

Johnson, N., Zhao, G., Hunsader, E., Qi, H., Johnson, N., Meng, J., and Tivnan, B. (2013). Abrupt rise of new machine ecology beyond human response time. *Nature Scientific Reports 3*, (Article Number: 2627).

Kaastra, I. and Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, 10(3):215–236.

Kablan, A. and Ng, W. (2011). Intraday high-frequency FX trading with adaptive neuro-fuzzy inference systems. *International Journal of Financial Markets and Derivatives*, 2(1):68–87.

Kablan, A. and Ng, W. L. (2010). Enhancing high-frequency order placement strategies with fuzzy logic and fuzzy inference. *IAENG International Journal of Computer Science*, 37(4).

Kasabov, N. and Filev, D. (2006). Evolving intelligent systems: methods, learning, & applications. In *Proceedings of the 2006 International Symposium on Evolving Fuzzy Systems, Ambleside, UK*, pages 8–18. IEEE.

Kasabov, N. and Song, Q. (2002). DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. *IEEE Transactions on Fuzzy Systems*, 10(2):144–154.

Kearns, M., Kulesza, A., and Nevmyvaka, Y. (2010). Empirical limitations on High-Frequency Trading Profitability. *Journal of Trading*, 5(4):50–62.

Knight, F. H. (1921). Risk, uncertainty and profit. *Boston: Houghton-Mifflin*.

Krollner, B., Vanstone, B., and Finnie, G. (2010). Financial time series forecasting with machine learning techniques: A survey. In *Proceedings of the European Symposium on Artificial Neural Networks: Computational and Machine Learning, Bruges, Belgium*, pages 28–30.

Lawrence, S., Back, A. D., Tsoi, A. C., and Giles, C. L. (1997). On the distribution of performance from multiple neural-network trials. *IEEE Transactions on Neural Networks*, 8(6):1507–1517.

LeBaron, B. (1999). Technical trading rule profitability and foreign exchange intervention. *Journal of International Economics*, 49(1):125–143.

Lei, K. S. and Wan, F. (2012). Applying ensemble learning techniques to anfis for air pollution index prediction in macau. In *Advances in Neural Networks–ISNN 2012*, pages 509–516. Springer.

Lin, C., Khan, H., and Huang, C. (2002). Can the neuro fuzzy model predict stock indexes better than its rivals? *CIRJE F-Series*.

Liu, L. Y., Patton, A. J., and Sheppard, K. (2015). Does anything beat 5-minute rv? a comparison of realized measures across multiple asset classes. *Journal of Econometrics*, 187(1):293–311.

Lo, A. W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, 58(4):36–52.

Lo, A. W., Repin, D. V., and Steenbarger, B. N. (2005). Fear and greed in financial markets: A clinical study of day-traders. *American Economic Review*, 95(3):352–359.

Luengo, J., García, S., and Herrera, F. (2009). A study on the use of statistical tests for experimentation with neural networks: Analysis of parametric test conditions and non-parametric tests. *Expert Systems with Applications*, 36(4):7798–7808.

Mamdani, E. (1977). Application of fuzzy logic to approximate reasoning using linguistic synthesis. *IEEE Transactions on Computers*, 100(12):1182–1191.

Mamdani, E. H. (1974). Application of fuzzy algorithms for control of simple dynamic plant. *Proceedings of the Institution of Electrical Engineers*, 121(12):1585–1588.

Maringer, D. (2008). Heuristic optimization for portfolio management [application notes]. *IEEE Computational Intelligence Magazine*, 3(4):31–34.

Maringer, D. and Parpas, P. (2009). Global optimization of higher order moments in portfolio selection. *Journal of Global Optimization*, 43(2-3):219–230.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1):77–91.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics*, 11(2):431–441.

Marsland, S. (2009). *Machine Learning: an Algorithmic Perspective*. Chapman & Hall/CRC, Boca Raton, FL, USA.

McAleer, M. and Medeiros, M. C. (2008). Realized volatility: A review. *Econometric Reviews*, 27(1):10–45.

Medeiros, M. C., Teräsvirta, T., and Rech, G. (2006). Building neural network models for time series: a statistical approach. *Journal of Forecasting*, 25(1):49–75.

Melin, P. and Castillo, O. (2014). A review on type-2 fuzzy logic applications in clustering, classification and pattern recognition. *Applied Soft Computing*, 21:568–577.

Melin, P., Soto, J., Castillo, O., and Soria, J. (2012). A new approach for time series prediction using ensembles of anfis models. *Expert Systems with Applications*, 39(3):3494–3506.

Mendel, J., Hagras, H., Tan, W.-W., Melek, W. W., and Ying, H. (2014). *Introduction to Type-2 Fuzzy Logic Control: Theory and Applications*. John Wiley & Sons, Hoboken, NJ, USA.

Mendel, J. M., John, R. I., and Liu, F. (2006). Interval type-2 fuzzy logic systems made simple. *IEEE Transactions on Fuzzy Systems*, 14(6):808–821.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4):323–361.

MiFID, II. (2014). Directive 2014/65/EC of the European Parliament and of the Council.

Mohammadian, M. (1995). Designing customized hierarchical fuzzy logic systems for modelling and prediction. *Computational Intelligence*, 1(2):5.

Murphy, J. (1987). *Study Guide for Technical Analysis of the Futures Markets: A Self-training Manual*. New York Institute of Finance.

Naranjo, R., Meco, A., Arroyo, J., and Santos, M. (2015). An intelligent trading system with fuzzy rules and fuzzy capital management. *International Journal of Intelligent Systems*, 30(8):963–983.

Nayak, R. (2009). Generating rules with predicates, terms and variables from the pruned neural networks. *Neural Networks*, 22(4):405–414.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2):347–370.

Nelson, S. C. and Katzenstein, P. J. (2014). Uncertainty, risk, and the financial crisis of 2008. *International Organization*, 68(02):361–392.

Nie, M. and Tan, W. W. (2008). Towards an efficient type-reduction method for interval type-2 fuzzy logic systems. In *Proceedings of FUZZ-IEEE 2008, Hong Kong*, pages 1425–1432.

Pal, N. R. and Bezdek, J. C. (1995). On cluster validity for the fuzzy c-means model. *IEEE Transactions on Fuzzy Systems*, 3(3):370–379.

Pardo, R. (2011). *The evaluation and optimization of trading strategies*, volume 314. John Wiley & Sons, New Jersey.

Poon, S. and Granger, C. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41(2):478–539.

Poon, S.-h. and Granger, C. (2005). Practical Issues in Forecasting Volatility. *Financial Analysts Journal*, 61(1):45–56.

Prechelt, L. (1996). A quantitative study of experimental evaluations of neural network learning algorithms: Current research practice. *Neural Networks*, 9(3):457–462.

Prokop, J. (2012). Further evidence on the role of ratio choice in hedge fund performance evaluation. *Journal of Finance and Investment Analysis*, 1(3):181–195.

Puuronen, S., Terziyan, V., and Tsymbal, A. (1999). A dynamic integration algorithm for an ensemble of classifiers. In *Z.W. Ras, A. Skowron (eds.), Foundations of Intelligent Systems, ISMIS 1999, Lecture Notes in AI*, volume 1609, pages 592–600. Springer-Verlag, Warsaw.

Rechenthin, M. and Street, W. N. (2013). Using conditional probability to identify trends in intra-day high-frequency equity pricing. *Physica A: Statistical Mechanics and its Applications*, 392(24):6169–6188.

Resta, M. (2009). Seize the (intra)day: Features selection and rules extraction for tradings on high-frequency data. *Neurocomputing*, 72(16-18):3413–3427.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance*, 39(4):1127–1139.

Rooney, N., Patterson, D., Anand, S., and Tsymbal, A. (2004). Dynamic integration of regression models. In *Lecture notes in computer science, MCS 2004: Multiple Classifier Systems, 5th International workshop, Cagliari, Italy*, volume 3077, pages 164–173. Springer, Berlin.

Ruano, A. E., Ferreira, P. M., Cabrita, C., and Matos, S. (2002). Training neural networks and neuro-fuzzy systems: A unified view. In *Proceedings of the 15th IFAC World Congress, Barcelona, Spain*, volume 15.

Schulmeister, S. (2006). The interaction between technical currency trading and exchange rate fluctuations. *Finance Research Letters*, 3(3):212–233.

Schulmeister, S. (2009). Profitability of technical stock trading: Has it moved from daily to intraday data? *Review of Financial Economics*, 18(4):190–201.

Sepulveda, R., Melin, P., Díaz, A. R., Mancilla, A., and Montiel, O. (2006). Analyzing the effects of the footprint of uncertainty in type-2 fuzzy logic controllers. *Engineering Letters*, 13(2):138–147.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, 39(1):119–138.

Sharpe, W. F. (1994). The sharpe ratio. *The journal of portfolio management*, 21(1):49–58.

Shiller, R. J. (1999). Human behavior and the efficiency of the financial system. *Handbook of Macroeconomics*, 1:1305–1340.

Simon, H. A. (1972). Theories of bounded rationality. *Decision and organization*, 1(1):161–176.

Son, Y., Noh, D., and Lee, J. (2012). Forecasting trends of high-frequency kospi200 index data using learning classifiers. *Expert Systems with Applications*, 39(14):11607–11615.

Sortino, F. A. and Price, L. N. (1994). Performance measurement in a downside risk framework. *the Journal of Investing*, 3(3):59–64.

Sortino, F. A. and Van Der Meer, R. (1991). Downside risk. *The Journal of Portfolio Management*, 17(4):27–31.

Soto, J., Melin, P., and Castillo, O. (2013). A new approach for time series prediction using ensembles of anfis models with interval type-2 and type-1 fuzzy integrators. In *Proceedings of the 2013 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, pages 68–73. IEEE.

Sugeno, M. and Kang, G. (1988). Structure identification of fuzzy model. *Fuzzy Sets and Systems*, 28(1):15–33.

Takagi, T. and Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics*, (1):116–132.

Tan, Z., Quek, C., and Cheng, P. Y. (2011). Stock trading with cycles: A financial application of ANFIS and reinforcement learning. *Expert Systems with Applications*, 38(5):4741–4755.

Tsai, C. and Wang, S. (2009). Stock price forecasting by hybrid machine learning techniques. In *Proceedings of the International MultiConference of Engineers and Computer Scientists, Hong Kong*, volume 1, pages 755–760.

Tsang, E. (2009). Forecasting - where computational intelligence meets the stock market. *Frontiers of Computer Science in China*, 1(3):53–63.

Tsymbal, A., Pechenizkiy, M., Cunningham, P., and Puuronen, S. (2008). Dynamic integration of classifiers for handling concept drift. *Information Fusion*, 9(1):56–68.

Vanstone, B. and Finnie, G. (2009). An empirical methodology for developing stockmarket trading systems using artificial neural networks. *Expert Systems with Applications*, 36(3):6668–6680.

Vanstone, B. and Finnie, G. (2010). Enhancing stockmarket trading performance with ANNs. *Expert Systems with Applications*, 37(9):6602–6610.

Vanstone, B. and Tan, C. (2003). A Survey of the Application of Soft Computing to Investment and Financial Trading. In *Proceedings of the Australian and New Zealand Intelligent Information Systems Conference*, volume 1(1), pages 211–216. Australian and New Zealand Intelligent Information Systems.

Von Wyss, R. (2004). *Measuring and predicting liquidity in the stock market*. Phd Dissertation, University of St. Gallen, Switzerland.

Wagner, C. and Hagras, H. (2010). Toward general type-2 fuzzy logic systems based on zslices. *IEEE Transactions on Fuzzy Systems*, 18(4):637–660.

Wang, L. and Mendel, J. (1992). Generating fuzzy rules by learning from examples. *IEEE Transactions on Systems, Man and Cybernetics*, 22(6):1414–1427.

Wei, L.-Y., Cheng, C.-H., and Wu, H.-H. (2014). A hybrid anfis based on n-period moving average model to forecast TAIEX stock. *Applied Soft Computing*, 19:86–92.

White, H. (2000). A reality check for data snooping. *Econometrica*, 68(5):1097–1126.

Wu, D. (2013). Approaches for reducing the computational cost of interval type-2 fuzzy logic systems: overview and comparisons. *IEEE Transactions on Fuzzy Systems*, 21(1):80–99.

Wu, D. and Mendel, J. M. (2014). Designing practical interval type-2 fuzzy logic systems made simple. In *Proceedings of FUZZ-IEEE 2014, Beijing*, pages 800–807.

Xu, R., Wunsch, D., et al. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3):645–678.

Xufre Casqueiro, P. and Rodrigues, A. J. (2006). Neuro-dynamic trading methods. *European Journal of Operational Research*, 175(3):1400–1412.

Zadeh, L. (1975). The concept of a linguistic variable and its application to approximate reasoning – Part I. *Information Sciences*, 8(3):199–249.

Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3):338–353.

Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(1):28–44.

Zadeh, L. A. (1997). Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets and Systems*, 90(2):111–127.

Zhang, F. (2010). The effect of high-frequency trading on stock volatility and price discovery. *Working paper, Yale University*.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411.

Appendix A

Further results from research chapters

Table A.1: Additional performance measures for neural network and ANFIS using Sharpe ratio optimisation following the 100-day out-sample period. Bold figures indicate highest score in the respective stock.

	Period Return		Profit Ratio		Win Ratio	
Stock	**NN**	**ANFIS**	**NN**	**ANFIS**	**NN**	**ANFIS**
AL	**0.0822**	0.0634	1.174	**1.3606**	0.5326	**0.5562**
ANTO	0.0185	**0.0379**	1.2398	**1.4041**	**0.4898**	0.4662
BLT	**0.0443**	0.0230	1.3539	**1.8878**	0.5238	**0.6444**
BA	-0.0627	**-0.0054**	0.7894	**0.9268**	0.4479	**0.4541**
BLND	0.0077	**0.0248**	1.0904	**1.1181**	0.4497	**0.5561**
SKY	-0.0009	**0.0383**	0.9857	**3.2727**	0.5385	**0.6081**
CWC	-0.0084	**0.0256**	0.9258	**1.1604**	**0.5562**	0.5410
AV	0.0516	**0.0566**	**1.5454**	1.3067	**0.5552**	0.5398
DGE	**0.0583**	0.0282	1.8357	**2.0224**	**0.6035**	0.5429
SDR	-0.0386	**0.0244**	**0.2840**	1.2916	0.3929	**0.5144**
HMSO	**0.1190**	0.0950	**2.3733**	1.4468	**0.6407**	0.5643
REX	-0.0040	**0.0063**	0.9516	**1.0139**	**0.5471**	0.5088
JMAT	-0.0023	**0.0170**	0.9777	**1.4961**	0.5166	**0.6028**
HSBA	**0.0243**	0.0070	**2.6442**	1.4418	**0.5949**	0.6000
ABF	0.0137	**0.0409**	**1.5551**	1.1828	**0.6273**	0.5406
INTU	**0.0626**	-0.0212	**2.0049**	0.8539	**0.5645**	0.5420
RIO	**0.0222**	-0.0247	**1.4063**	0.8409	**0.5027**	0.4850
BP	0.0171	**0.0327**	1.1500	**1.2268**	0.5102	**0.5421**
SGE	-0.0076	**0.0609**	0.9562	**2.5188**	**0.5311**	0.5304
LLOY	-0.0269	**0.0071**	0.8804	**1.0692**	0.4794	**0.5328**
TSCO	0.0184	**0.0375**	1.2571	**1.3918**	**0.5500**	0.5166
GSK	0.0235	**0.0284**	1.3122	**1.3704**	**0.5398**	0.4957
AZN	-0.0314	**0.0528**	0.2957	**1.7196**	0.4250	**0.5859**
HBOS	0.1093	**0.1717**	**1.2932**	1.1914	**0.5232**	0.4846
XTA	-0.0231	**0.0800**	0.5229	**2.4923**	0.4225	**0.5882**
IAP	**0.0335**	-0.0448	**2.0786**	0.7060	**0.6267**	0.4557
ITV	0.0181	**0.0467**	**1.7216**	1.4173	**0.5738**	0.5439

Table A.2: Additional performance measures for neural network and ANFIS using Sortino ratio optimisation following the 100-day out-sample period. Bold figures indicate highest score in the respective stock.

	Period Return		**Profit Ratio**		**Win Ratio**	
Stock	**NN**	**ANFIS**	**NN**	**ANFIS**	**NN**	**ANFIS**
AL	**0.0814**	0.0409	**1.4135**	1.059	**0.5321**	0.5241
ANTO	-0.0709	**0.0337**	0.4618	**1.2449**	**0.4651**	0.4602
BLT	0.0793	**0.0903**	**1.1605**	1.1178	0.5142	**0.5191**
BA	-0.1551	**0.0328**	0.6318	**3.2670**	0.4820	**0.5500**
BLND	**0.0818**	0.0810	**1.4889**	1.3368	0.5524	**0.5528**
SKY	**0.0314**	0.0273	1.9368	**2.3584**	**0.6040**	0.5676
CWC	-0.0856	**0.0518**	0.5563	**1.2887**	0.4658	**0.5504**
AV	**0.1201**	0.0154	**1.7144**	1.0542	**0.5345**	0.5297
DGE	0.0004	**0.0519**	1.0137	**1.3686**	**0.5798**	0.5635
SDR	-0.0428	**0.0721**	0.3986	**2.2870**	**0.5122**	0.4977
HMSO	**0.0864**	0.0327	2.4330	**2.7396**	**0.6154**	0.6000
REX	**-0.0040**	-0.0049	**0.9516**	0.8921	**0.5471**	0.5039
JMAT	**0.0514**	0.0179	**1.3439**	1.2534	**0.5688**	0.4545
HSBA	0.0163	**0.0239**	1.4835	**1.8178**	0.4844	**0.5057**
ABF	0.0272	**0.0308**	1.1068	**2.2605**	0.5326	**0.5965**
INTU	**0.0785**	-0.0309	**1.4052**	0.7730	**0.5744**	0.5311
RIO	-0.0111	**-0.0013**	0.4634	**0.9844**	0.4242	**0.5082**
BP	0.0188	**0.0268**	1.1944	**3.7832**	0.5724	**0.6333**
SGE	0.0083	**0.0443**	1.0841	**1.7342**	**0.5268**	0.4842
LLOY	**0.0791**	-0.0529	**1.5771**	0.6592	**0.5938**	0.4706
TSCO	0.0043	**0.0375**	1.0235	**1.3918**	**0.5249**	0.5166
GSK	**0.0657**	0.0223	**2.1519**	1.3037	**0.5291**	0.5083
AZN	0.0033	**0.0777**	1.2032	**2.0305**	0.5385	**0.6091**
HBOS	-0.0568	**0.1937**	0.8480	**1.1612**	0.4630	**0.4880**
XTA	-0.0213	**0.0266**	0.4659	**1.3122**	0.4675	**0.5031**
IAP	**0.0121**	-0.0165	**1.1139**	0.6767	**0.5506**	0.5152
ITV	0.0292	**0.1021**	1.0755	**1.3406**	0.5283	**0.5378**

Table A.3: Sortino ratio and number of trades for neural network, FCM1 and FCM2 following 100-day out-of-sample period. Bold figures indicate highest score in the respective stock.

	Sortino Ratio			No. of Trades		
Stock	**NN**	**FCM1**	**FCM2**	**NN**	**FCM1**	**FCM2**
AL	0.3288	0.5058	**0.7366**	**9120**	7019	3134
ANTO	0.0126	0.0623	**0.1791**	**8069**	3873	3637
BLT	0.2119	**0.2667**	0.2583	**1284**	819	602
BA	0.3319	**0.3847**	0.3327	**2889**	2446	632
BLND	0.1504	**0.2644**	0.1401	**2556**	1550	831
SKY	0.5120	0.4272	**0.9577**	**1151**	772	174
CWC	0.0028	-0.0995	**0.1654**	**1437**	566	581
AV	-0.1479	**-0.0762**	-0.1699	**1691**	544	537
DGE	0.5109	**0.6539**	0.3683	**1345**	1032	617
SDR	-0.0101	-0.1953	**0.0364**	**3008**	805	1372
HMSO	**0.7009**	0.4014	0.4533	**3574**	3036	1657
REX	0.2043	**0.6075**	0.5095	**3709**	1247	607
JMAT	**0.2570**	0.0283	0.0505	**2560**	654	390
HSBA	**0.4359**	0.3541	0.2375	492	**634**	272
ABF	0.1030	**0.1629**	0.1319	**2372**	1618	1236
INTU	**0.9743**	0.8970	0.8080	**3901**	3447	2132
RIO	-0.1807	-0.1864	**-0.1232**	**2114**	937	919
BP	0.0215	**0.9521**	0.0455	**1578**	859	383
SGE	**0.4482**	0.4292	0.3408	**2414**	2247	1369
LLOY	0.0634	**0.1151**	0.0772	**2770**	1856	351
TSCO	0.3189	0.2934	**0.4797**	**1275**	1073	588
GSK	**0.5892**	0.4630	0.5045	**1792**	1100	932
AZN	0.2509	**0.4314**	0.3474	**1399**	1026	573
HBOS	-0.1013	0.0175	**0.0667**	**7137**	2968	2600
XTA	-0.2691	-0.2645	**-0.1943**	**2616**	850	667
IAP	0.0137	**0.2075**	0.0981	**3961**	1775	2233
ITV	0.6155	**0.6780**	0.4227	**2745**	2406	837
Average	0.2351	**0.2882**	0.2689	**2924**	1747	1106

Table A.4: Win ratio and profit ratio for neural network, FCM1 and FCM2 following 100-day out-of-sample period. Bold figures indicate highest score in the respective stock.

	Win Ratio			Profit Ratio		
Stock	**NN**	**FCM1**	**FCM2**	**NN**	**FCM1**	**FCM2**
AL	0.5471	0.5521	**0.5625**	1.2109	1.3080	**1.4428**
ANTO	0.5038	0.5092	**0.5194**	1.0141	1.0473	**1.0958**
BLT	0.5210	0.5372	**0.5432**	1.1839	1.3121	**1.3401**
BA	0.4874	0.4947	**0.5047**	1.1970	1.2462	**1.4475**
BLND	0.5070	0.5155	**0.5172**	**1.2814**	1.2524	1.2526
SKY	0.4718	0.4741	**0.5230**	1.4911	1.5555	**3.1517**
CWC	0.5240	0.5159	**0.5318**	1.0556	0.8981	**1.3725**
AV	0.4849	**0.4926**	0.4655	0.9368	**1.0175**	0.8620
DGE	0.4981	0.5252	**0.5365**	1.3925	**1.5705**	1.4989
SDR	0.4508	**0.4892**	0.4875	0.8325	0.7795	**1.0882**
HMSO	**0.5129**	0.5011	0.5003	1.2742	1.1929	**1.3021**
REX	0.4999	0.5373	**0.5535**	1.2127	1.4532	**1.6234**
JMAT	**0.5219**	0.5061	0.5000	1.2926	1.1676	**1.3055**
HSBA	0.5305	0.5174	**0.5404**	1.6744	1.4611	**1.6883**
ABF	0.5223	0.5321	**0.5324**	1.1371	1.2222	**1.2244**
INTU	**0.5586**	0.5573	0.5549	**1.4672**	1.4501	1.4525
RIO	0.4782	0.4792	**0.4995**	0.8894	0.8520	**0.9223**
BP	0.5013	**0.5343**	0.5274	1.0123	**1.8266**	1.3056
SGE	0.5123	0.4798	**0.5268**	1.2981	1.3008	**1.3161**
LLOY	0.5229	0.5059	**0.5251**	1.0611	1.1076	**1.3121**
TSCO	0.5224	0.5200	**0.5527**	1.3437	1.3415	**1.7316**
GSK	0.5521	**0.5536**	0.5386	1.3858	1.5035	**1.5129**
AZN	0.5318	0.5429	**0.5620**	1.2697	1.3934	**1.5789**
HBOS	0.4768	0.4821	**0.5134**	0.9631	1.0302	**1.0680**
XTA	**0.4698**	0.4678	0.4587	**0.8336**	0.7206	0.8171
IAP	0.5128	0.5153	0.4863	1.0266	**1.1673**	1.0745
ITV	**0.5421**	0.4456	0.5312	1.4063	**1.4577**	1.4304
Average	0.5098	0.5105	**0.5220**	1.1905	1.2458	**1.3784**

Table A.5: Sortino ratio and number of trades for neural network and NN-FMM following 100-day out-of-sample period. Bold figures indicate highest score in the respective stock.

	Sortino Ratio		**No. of Trades**	
Stock	**NN**	**NN-FMM**	**NN**	**NN-FMM**
AL	3.2327	**3.9894**	**5761**	5703
ANTO	0.4189	**0.6057**	**2509**	2469
BLT	0.4226	**0.4780**	**975**	900
BA	0.7212	**0.8412**	**1313**	1295
BLND	0.4443	**0.5829**	**1178**	1148
SKY	0.5529	**0.8179**	**1583**	1549
CWC	0.0030	**0.1596**	**1990**	1345
AV	-0.0879	**-0.0277**	**476**	461
DGE	0.7664	**0.8230**	702	**948**
SDR	-0.0521	**0.1028**	**2005**	1901
HMSO	**1.0341**	0.8904	**2804**	1706
REX	**0.4456**	0.3374	**2125**	1375
JMAT	**0.1279**	0.0008	**612**	430
HSBA	0.2246	**0.3404**	**349**	294
ABF	0.2699	**0.5624**	**1273**	1232
INTU	2.2515	**2.4529**	**2041**	2014
RIO	0.1036	**0.1781**	**1141**	464
BP	1.0049	**1.3139**	711	**886**
SGE	0.8445	**1.0233**	**1237**	1194
LLOY	0.4140	**0.6520**	**1402**	1365
TSCO	0.5302	**0.5469**	**1935**	1905
GSK	1.6452	**1.9154**	**1810**	1740
AZN	-0.0242	**0.0333**	**1069**	1037
HBOS	**0.0432**	-0.0067	2949	**3297**
XTA	0.0477	**0.0820**	**912**	869
IAP	**-0.0318**	-0.1786	**3165**	1514
ITV	0.7594	**1.6373**	2091	**2804**
Average	0.5968	**0.7464**	**1708**	1550

Table A.6: Additional risk-adjusted performance measures for neural network and NN-FMM following 50 trials over the 100-day out-of-sample period. Bold figures indicate highest score in the respective stock. Each measure represents an average score from the 50 trials.

	Sortino Ratio		Calmar Ratio	
Stock	**NN**	**NN-FMM**	**NN**	**NN-FMM**
AL	2.5916[a]	**2.6553**[a]	16.3597[a]	**17.1582**[a]
ANTO	0.5213[a]	**0.5623**[a]	8.1682[a]	**9.1225**[a]
BLT	0.3821	**0.4453**[bc]	2.7642[a]	**3.3214**[abc]
BA	0.3955	**0.5368**[bc]	3.4181	**3.6273**[abc]
BLND	0.3929[a]	**0.4664**	4.3528[a]	**5.5768**[a]
SKY	0.3694	**0.6221**[bc]	4.1706	**7.2559**[abc]
CWC	0.0377	**0.1455**[bc]	0.1621[a]	**1.1701**[bc]
AV	-0.0268	**0.0315**[bc]	-0.0855[a]	**0.2850**[abc]
DGE	0.4892	**0.5847**[bc]	4.6872	**6.2567**[a]
SDR	-0.0199[a]	**0.0213**	0.0190[a]	**0.3882**[a]
HMSO	**1.1524**[a]	1.0113	**15.3612**	12.7256[bc]
REX	0.4910[a]	**0.7435**[bc]	3.8960[a]	**6.3349**[abc]
JMAT	**0.1999**	0.1476	**1.5696**	1.1716
HSBA	0.2626	**0.3285**[bc]	1.0436[a]	**1.3916**[abc]
ABF	0.3014	**0.5515**[bc]	3.1425	**3.7412**
INTU	1.9243[a]	**2.0837**[abc]	8.7378[a]	**10.9900**[bc]
RIO	0.0442[a]	**0.0788**[a]	0.6273[a]	**0.8058**[a]
BP	1.2721	**1.5777**	20.0934	**25.4096**
SGE	0.4687[a]	**0.5897**[abc]	5.0386[a]	**6.2753**[a]
LLOY	0.3060	**0.3948**[bc]	1.8534	**2.4356**[bc]
TSCO	0.4396	**0.6145**[bc]	4.5210	**6.5764**[bc]
GSK	0.7304[a]	**0.8469**[a]	6.3696[a]	**7.7643**[a]
AZN	-0.0278	**0.0517**[bc]	-0.0974	**0.0923**[bc]
HBOS	**0.0621**	-0.0025[bc]	**1.0243**[a]	0.0665[bc]
XTA	0.0534[a]	**0.1463**[abc]	0.9875[a]	**1.5764**[abc]
IAP	**-0.0321**	-0.1065[bc]	**-0.0852**[a]	-0.6639[abc]
ITV	0.5802	**1.3276**[bc]	8.5481	**12.5732**[abc]
Average	0.4949	**0.6095**	4.6907	**5.6825**

[a] Anderson-Darling test rejects the null hypothesis at the 5% sig. level that the distribution of specific measure is from a normal distribution.
[b] Two-sample t-test supports the alternative hypothesis at the 5% sig. level that the distribution of the specific measure for both models comes from populations with unequal means.
[c] Ranksum test supports the alternative hypothesis at the 5% sig. level that the distribution of the specific measure for both models comes from populations with unequal medians.

Table A.7: Sortino ratio, number of trades and profit per trade for standard ANFIS following 100-day out-of-sample period.

	Standard ANFIS		
Stock	**Sortino Ratio**	**No. of Trades**	**Profit / Trade**
AL	7.3629	6137	£47.64
ANTO	0.6430	2036	£16.32
BLT	0.4601	475	£40.16
BA	5.8219	6248	£33.84
BLND	0.8979	921	£26.30
SKY	1.3154	2606	£14.68
CWC	4.6573	4709	£75.01
AV	0.8282	1074	£15.85
DGE	0.7130	1203	£13.22
SDR	-0.1461	436	-£15.08
HMSO	0.9685	1103	£35.70
REX	2.5425	6381	£13.75
JMAT	0.0326	398	£2.14
HSBA	0.3087	1129	£13.13
ABF	0.3981	630	£19.44
INTU	1.7394	1463	£25.48
RIO	0.2469	535	£16.78
BP	4.1172	6978	£14.84
SGE	6.8817	4481	£45.00
LLOY	1.3071	9621	£17.10
TSCO	2.8888	7928	£19.19
GSK	0.2501	632	£16.85
AZN	0.0353	456	£1.26
HBOS	1.0497	4708	£13.24
XTA	0.1334	911	£12.54
IAP	-0.0361	2803	-£2.12
ITV	4.7751	3307	£47.67
Average	1.8590	2937	£25.95

Table A.8: Sortino ratio, number of trades and profit per trade for ANFIS/T2a following 100-day out-of-sample period. Percentage values show differences against the scores obtained by standard ANFIS.

	ANFIS/T2a performance improvements (%)		
Stock	**Sortino Ratio**	**No. of Trades**	**Profit / Trade**
AL	-20.72%	-1.84%	+0.88%
ANTO	+0.98%	-7.96%	+13.71%
BLT	+8.15%	-5.68%	+14.23%
BA	-46.57%	-1.52%	-2.08%
BLND	+4.91%	-2.39%	+7.63%
SKY	+8.27%	-1.27%	+4.33%
CWC	+2.63%	-0.38%	-0.48%
AV	+44.10%	-1.77%	+12.56%
DGE	+6.17%	-0.83%	+3.53%
SDR	141.00%	-16.97%	-43.34%
HMSO	+34.95%	-2.27%	+12.00%
REX	+5.38%	-3.62%	+3.40%
JMAT	+20.86%	-2.76%	+24.02%
HSBA	-0.32%	-1.59%	-0.78%
ABF	-9.44%	-1.90%	-5.76%
INTU	+26.88%	-6.02%	+9.83%
RIO	+37.14%	-3.36%	+23.95%
BP	+21.03%	-0.54%	+7.66%
SGE	-8.43%	-3.50%	-3.27%
LLOY	+48.66%	-1.14%	+11.24%
TSCO	+51.77%	-1.00%	+1.54%
GSK	-1.96%	-2.69%	+0.83%
AZN	-31.44%	-0.66%	-8.09%
HBOS	+2.84%	-6.46%	+13.94%
XTA	+19.42%	-2.63%	+21.58%
IAP	-18.56%	-4.32%	+9.10%
ITV	+52.45%	+0.88%	+8.63%
Average	+4.76%	-3.12%	+3.87%

Table A.9: Sortino ratio, number of trades and profit per trade for ANFIS/T2b following 100-day out-of-sample period. Percentage values show differences against the scores obtained by standard ANFIS.

	ANFIS/T2b performance improvements (%)		
Stock	**Sortino Ratio**	**No. of Trades**	**Profit / Trade**
AL	-8.29%	-1.42%	+1.03%
ANTO	-27.48%	-6.39%	-10.46%
BLT	-0.93%	-5.89%	+4.73%
BA	-24.69%	-0.59%	-2.19%
BLND	+9.17%	-1.30%	+8.53%
SKY	+1.70%	-0.77%	+1.56%
CWC	-0.75%	-0.83%	+0.46%
AV	+45.34%	-1.40%	+16.90%
DGE	+5.92%	-0.25%	+2.77%
SDR	+21.90%	-5.96%	+30.45%
HMSO	+41.12%	-1.72%	+14.29%
REX	+16.03%	-2.37%	+3.15%
JMAT	+42.94%	-2.51%	+44.82%
HSBA	-1.23%	-0.80%	-0.73%
ABF	-4.42%	-1.75%	-0.30%
INTU	+8.80%	-5.19%	+7.74%
RIO	+35.76%	-2.99%	+21.75%
BP	+16.99%	-0.40%	+5.13%
SGE	-28.04%	-3.46%	-3.30%
LLOY	-12.39%	-1.25%	+1.34%
TSCO	+49.96%	-0.77%	+1.29%
GSK	-1.64%	-1.90%	+0.26%
AZN	+1.70%	+0.44%	-82.68%
HBOS	+0.26%	-5.54%	+11.60%
XTA	-10.57%	-1.87%	-9.47%
IAP	+7.20%	-5.21%	+20.64%
ITV	+51.59%	+1.27%	+11.34%
Average	+3.52%	-2.25%	+2.66%

Table A.10: ANFIS number of trades after 100-day out-of-sample period across different levels of return threshold (uncertainty).

	Standard ANFIS - No. of Trades			
	Return Threshold			
Stock	**0.08%**	**0.06%**	**0.04%**	**0.02%**
AL	6137	8698	12208	16618
ANTO	2036	3687	6795	14270
BLT	475	1016	2652	2200
BA	6248	7830	9738	12520
BLND	921	1956	4462	10466
SKY	1197	2606	5481	9385
CWC	4308	4709	5260	6204
AV	1074	2429	6104	12930
DGE	480	1203	3348	10495
SDR	436	792	1804	2848
HMSO	1103	2048	4704	5464
REX	1368	2867	6381	11319
JMAT	398	870	2092	4910
HSBA	426	1129	3836	10153
ABF	630	1404	3789	9935
INTU	1463	2837	5930	4776
RIO	535	1071	2643	7534
BP	1577	3220	6978	11295
SGE	4481	5515	7273	8567
LLOY	3140	5849	9621	13693
TSCO	3211	5505	7928	10865
GSK	632	1265	3100	9361
AZN	456	867	1845	3154
HBOS	4708	7322	11467	16007
XTA	911	1715	3771	2992
IAP	1622	2803	6185	4123
ITV	3136	3307	4153	3960
Average	1967	3130	5539	8742

Table A.11: ANFIS profit per trade after 100-day out-of-sample period across different levels of return threshold (uncertainty).

	Standard ANFIS - Profit / Trade (£)			
	Return Threshold			
Stock	**0.08%**	**0.06%**	**0.04%**	**0.02%**
AL	47.64	39.67	35.44	16.35
ANTO	17.08	9.97	12.48	2.99
BLT	40.16	14.05	3.81	-15.47
BA	33.84	32.38	32.13	26.84
BLND	26.30	8.67	6.66	-2.46
SKY	26.25	14.68	12.15	7.06
CWC	72.89	75.01	72.08	63.63
AV	15.85	10.63	2.79	0.35
DGE	28.28	13.22	5.00	2.01
SDR	-15.08	-8.78	-2.02	-13.90
HMSO	35.70	37.79	15.06	7.76
REX	20.16	26.04	13.75	12.01
JMAT	2.14	-2.70	-4.60	-8.31
HSBA	29.34	13.13	5.99	0.16
ABF	19.44	13.55	7.26	8.29
INTU	25.48	21.87	14.79	8.68
RIO	16.78	5.93	-4.15	-9.03
BP	31.59	19.30	14.84	9.68
SGE	45.00	45.21	40.94	31.10
LLOY	19.41	17.84	17.10	12.71
TSCO	14.21	21.02	19.19	15.37
GSK	16.85	10.95	5.68	-0.23
AZN	1.26	-2.45	-7.89	-10.59
HBOS	13.24	13.07	10.40	5.89
XTA	12.54	-3.78	-7.84	-21.82
IAP	-2.64	-1.81	0.72	-4.32
ITV	56.47	47.67	59.15	36.51
Average	32.12	25.20	17.94	8.77