A NEW HYBRID FORECASTING MODEL USING WAVELET-PCA AND ARTIFICIAL NEURAL NETWORK FOR FUTURES MARKETS

MOHAMMADALI MEHRALIZADEH

FACULTY OF BUSINESS AND ACCOUNTANCY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018

A NEW HYBRID FORECASTING MODEL USING WAVELET-PCA AND ARTIFICIAL NEURAL NETWORK FOR FUTURES MARKETS

MOHAMMADALI MEHRALIZADEH

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE SCHOOL OF BUSINESS
FACULTY OF BUSINESS AND ACCOUNTANCY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Mohammadali Mehralizadeh

Matric No: CHA120016

Name of Degree: Doctor of Philosophy in Finance

Title of Thesis: A NEW HYBRID FORECASTING MODEL USING WAVELET-PCA AND ARTIFICIAL NEURAL NETWORK FOR FUTURES MARKETS

Field of Study: Finance, Banking

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 28/11/2018

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:
Designation:

ABSTRACT

The motivation behind the present study is to determine whether stock indices futures possess long-term memory and entropy, and whether they follow the adaptive market hypothesis (AMH). Another motivation is to specify whether innovatively combined new methods such as wavelets, principal component analysis (PCA), and artificial neural networks (ANNs) produce returns in excess of the threshold buy-and-hold strategy advocated by the random walk hypothesis (RWH). In addition, this study tries to determine if a novel hybrid forecasting model produces higher returns than traditional and intelligent technical trading strategies, such as the pure ANN forecasting method and combination of wavelets and ANN, in the current increasingly difficult and volatile futures markets. In accordance with the literature review and the limitations of ANNs in the context of noisy time series, this study proposes a hybrid ANN with wavelet transforms, which is a special tool for denoising signal, namely wavelet neural network (WNN). Moreover, in accordance with an enhancement on denoising process with wavelet transforms using principal component analysis, a novel hybrid intelligent system, namely wavelet principal component analysis neural network (WPCA-NN) model, is developed in order to denoise financial time series carefully and forecast them with greater accuracy.

The study shows evidence drawn from 2005 to 2014 of the predictive ability and profitability of the WPCA-NN model regarding the contracts of Hong Kong's Hang Seng futures, Japan's NIKKEI 225 futures, Malaysia's Kuala Lumpur Composite Index (KLCI) futures, Singapore's Morgan Stanley Capital International (MSCI) futures, the Korea Composite Stock Price Index 200 (KOSPI 200) futures, the Standard & Poor's 500 (S&P 500), and Taiwan's Stock Exchange TAIEX futures markets. Moreover, this study employs many technical analysis indicators, consisting

of the relative strength index (RSI), moving average convergence/divergence (MACD), MACD signal, stochastic fast %K, stochastic slow %K, stochastic %D. In addition, this study uses three years of historical data to train a network in order to forecast an incoming quarter, which we call the evaluation period. Because the investment trading community and fund managers evaluate their portfolio performances quarterly, and change their trading parameters for the next quarter, this study is conducted on a quarterly basis. There is also a quarterly evaluation for 10 years, from 2005 to 2014, in order to measure trading performance and trading profitability.

The advantage of the sustainable returns of the WPCA-NN model to threshold buy-and-hold strategy and other forecasting methods, would offer existence of the inefficiency in selected markets and will give suitable tools for weary traders to forecast financial markets. In addition, this superiority would support the adaptive market hypothesis (AMH). Moreover, the offered method of denoising can be considered an enhancement to the univariate wavelet denoising, not only in financial domain, but also in other fields of study.

ABSTRAK

Motivasi di sebalik kajian ini adalah untuk menentukan sama ada kontrak niaga hadapan indeks saham mempunyai memori jangka masa panjang dan entropi, dan sama ada ia mematuhi hipotesis pasaran yang adaptif (AMH). Motivasi selanjutnya adalah untuk menentukan sama ada kaedah perdagangan baru yang menggabungkan teknik-teknik seperti wavelet, analisis komponen utama (PCA) dan rangkaian saraf tiruan (ANN) dapat menghasilkan pulangan yang lebih tinggi daripada strategi membeli dan menyimpan, yang disokong oleh hipotesis perjalanan rawak (RWH). Disamping itu, kajian ini juga bertujuan untuk menentukan sama ada model ramalan gabungan yang baru ini dapat menjana pulangan yang lebih tinggi berbanding dengan strategi perdagangan teknikal tradisional dan bestari seperti kaedah peramalan ANN yang asal serta kombinasi di antara wavelet dan ANN, di dalam pasaran niaga hadapan yang kini semakin rumit dan tidak menentu. Selaras dengan hasil kajian sastera dan batasan ANN dalam konteks analisis siri masa yang hingar, kajian ini mencadangkan sistem gabungan ANN dan transformasi wavelet, yang merupakan suatu teknik khas untuk menyahbunyi isyarat dipanggil wavelet rangkaian saraf (WNN). Selanjutnya, kajian ini telah mengaplikasikan teknik transformasi wavelet yang bersendikan analisis komponen utama, dan menghasilkan sebuah sistem gabungan pintar baru yang dinamakan model rangkaian saraf - wavelet analisis komponen utama (WPCA-NN) untuk mengurangkan hingar siri masa dengan teliti dan menghasilkan ramalan yang berketepatan lebih tinggi.

Kajian ini telah menunjukkan keupayaan ramalan dan keuntungan model perdagangan WPCA-NN menggunakan beberapa kontrak niaga hadapan melibatkan indeks saham Hang Seng dari Hong Kong, NIKKEI 225 dari Jepun, indeks komposit

Kuala Lumpur (KLCI) dari Malaysia, indeks permodalan antarabangsa Morgan Stanley (MSCI) dari Singapura, indeks komposit harga saham Korea 200 (KOSPI 200) dari Korea Selatan, Standard & Poor's 500 (S&P 500) dan indeks pasaran saham Taiwan (TAIEX) dari Taiwan menggunakan data dari tahun 2000 hingga 2014. Kajian ini menggunapakai isyarat perdagangan dari beberapa petunjuk analisis teknikal yang antaranya terdiri daripada indeks bandingan kekuatan (RSI), penumpuan/ penyimpangan purata bergerak (MACD), isyarat MACD, stokastik pantas %K, stokastik perlahan %K, dan stokastik %D. Di samping itu, kajian ini juga menggunakan data bertempoh tiga tahun untuk melatih rangkaian saraf bagi tujuan meramal data pada suku tahun berikutnya yang dipanggil tempoh penilaian. Kajian ini telah menjalankan penilaian berdasarkan kekerapan masa suku tahunan, ekoran prosedur standard di kalangan komuniti pedagang dan pengurus dana yang menilai prestasi potfolio pada setiap suku tahun dan menukar tetapan parameter perdagangan pada suku berikutnya. Justeru, penilaian suku tahunan bagi tempoh 10 tahun bermula 2005 sehingga 2014 telah dilakukan bagi menilai prestasi dan keuntungan perdagangan.

Pulangan lebih mampan yang dihasilkan oleh model WPCA-NN berbanding strategi membeli-dan-menyimpan dan kaedah-kaedah ramalan lain mencerminkan wujudnya ketidakcekapan dalam beberapa pasaran kewangan, dan mampu menjadi alat yang sesuai di kalangan pedagang untuk meramal perubahan pasaran. Di samping itu, keunggulan yang ditonjolkan oleh model ini memberi sokongan kepada hipotesis pasaran yang adaptif (AMH). Tambahan pula, kaedah nyahbunyi multivariat yang ditawarkan boleh dianggap sebagai peningkatan kepada kaedah nyahbunyi wavelet univariat, bukan sahaja dalam ruang lingkup bidang kewangan tetapi juga dalam lapangan kajian yang lain.

ACKNOWLEDGEMENTS

I would like to thank my wife and parents for their emotional and spiritual support.

I would also like to express my sincere gratitude to my supervisor Dr. Jacinta for the continuous support of my Ph.D study.

I would like to thank Dr. Noor Fadiya for the helpful and kind comments about my work.

TABLE OF CONTENTS

Abstract ii

ABSTRAK iv

Acknowledgements vi

Table of Contents vii

List of Figures xii

List of Tables..xv

List of Symbols and Abbreviations..xviii

List of Appendices..xxii

CHAPTER 1: INTRODUCTION 1

1.1 Background of the Study 7

1.1.1 Random Walk Hypothesis.. 8

1.1.2 The Efficient Market Hypothesis.. 10

1.1.3 Adaptive Market Hypothesis 12

1.2 Problem Statement.. 14

1.3 Research Questions.. 15

1.4 Research Objectives.. 16

1.5 Research Hypotheses 17

1.6 Purpose of Study.. 18

1.7 Scope of the Study 19

1.8 Justification of the Study 23

1.9 Organization of the Study 25

CHAPTER 2: LITERATURE REVIEW 30

2.1 Introduction 30

2.2 Discussion of the Literature Review and Market Theories 32

2.3 The Adaptive Market Hypothesis 34

2.4 Long-Term Memory 40

2.4.1 The Hurst Exponent 44

2.4.2 The Method of Mandelbrot 46

2.5 Approximate Entropy 48

2.6 Technical Analysis Review 50

2.6.1 Technical Analysis in Contrast to Fundamental Analysis 51

2.6.2 Technical Analysis and Critiques 52

2.7 Technical Trading Rules 53

2.7.1 Buy and Hold Criterion 53

2.7.2 Mechanical Trading Rules 54

2.7.2.1 Moving Average Rules 54

2.7.2.2 Momentum and Oscillator Rules 58

2.7.2.3 Other Trading Rules 61

2.8 Artificial Neural Networks 63

2.8.1 Hybrid Intelligent Systems 67

2.8.2 ANNs in Financial Prediction and Planning 68

2.9 Wavelet Transforms and Artificial Neural Networks 72

2.10 Technical Variables in Artificial Neural Networks 76

2.11 The Financial Crisis 83

2.12 Development of Research Methodology 84

2.13 Literature Review Summary 87

CHAPTER 3: RESEARCH METHODOLOGY 108

3.1 Chapter Overview 108

3.2 Modeling the Financial Market 110

3.3 Sample Data 112

3.3.1 Data Analysis Techniques 113

3.3.1.1 Unit Root Test 113

3.3.1.2 Serial Correlation Test 116

3.3.1.3 Granger Causality Test 118

3.3.2 Data Analysis 120

3.4 Long-Term Memory 121

3.5 Approximate Entropy 124

3.6 Technical Trading Rules 125

3.6.1 The Buy-and-Hold Benchmark 126

3.6.2 The RSI Trading Rule 126

3.6.3 The MACD Trading Rule 126

3.6.4 The Stochastic Trading Rule 127

3.6.5 The OMA Trading Rule 127

3.6.6 The Return of Trading Rules 128

3.7 Developing a Modern Forecasting Technique 129

3.7.1 Wavelet Principal Component Analysis Denoising 130

3.7.1.1 Multiple Univariate Wavelet Denoising 131

3.7.1.2 Multivariate Denoising Using Wavelet and Principal Component Analysis (WPCA) 133

3.7.2 A Nonlinear Autoregressive Neural Network with Exogenous Inputs (NARX-NN) 134

3.7.3 Research Framework 136

3.7.3.1 Data Framework 137

3.7.3.2 Model Inputs 141

3.7.4 Architecture of the Models 143

3.7.5 Forecasting Performance 153

3.7.6 Profitability (Return) Performance 154

3.8 The Methodology of Wavelet-PCA Neural Network 155

3.8.1 Data Collection 155

3.8.2 Denoising OHLC Signal Using WPCA 157

3.8.3 Preparing Denoised Data to Feed to Neural Network 158

3.8.4 Training Network by Neural Network Toolbox 159

3.8.5 Selecting The Best-Performing Network in The In-Sample Data 160

3.8.6 Preparing The Out-sample Data for Evaluation 160

3.8.7 Application of The Trained Network For Out-Sample Data 161

3.8.8 Conclusion of the Methodology 161

3.9 Hypothesis Review 162

3.10 Summary 163

CHAPTER 4: RESULTS AND RESEARCH ANALYSES 166

4.1 Introduction 166

4.2 Hurst Exponent Index 166

4.3 Approximate Entropy 169

4.4 Results of Forecasting Performance: MAPE 172

4.5 Profitability Results 173

4.6 Analysis 175

4.7 Discussion of Results 181

4.7.1 Risks in Forecasting Results 181

4.7.2 T-Test and Comparison of Two Means 185

4.7.3 The Significant Alpha above Market Return 187

4.8 Research Results 194

4.8.1 Hang Seng Futures Market 194

4.8.2 KLCI Futures Market 199

4.8.3 KOSPI 200 Futures Market 203

4.8.4 NIKKEI 225 Futures Market 207

4.8.5 SiMSCI Futures Market 211

4.8.6 S&P 500 Futures Market 215

4.8.7 TAIEX Futures 219

4.9 Summary and Discussion of Results 223

CHAPTER 5: CONCLUSION 244

5.1 Summary 244

5.2 Conclusion 246

5.3 Significance of the Study 252

5.4 Suggestions for Future Studies 253

References 256

List of Publications 281

LIST OF FIGURES

Figure 1.1: Scope of the study 22

Figure 2.1: Structure of the literature review 31

Figure 3.1: Feedforward backpropagation neural network architecture 135

Figure 3.2. The arrangement of the continuous datasets for training and evaluation, 2002–2014 138

Figure 3.3: The continuous datasets of training and evaluation for the Hang Seng futures market, from 2002 to 2014 140

Figure 3.4: Feedforward backpropagation neural network architecture 145

Figure 3.5: Univariate (wavelet) and multivariate (wavelet-PCA) denoising of Hang Seng futures, 2002 148

Figure 3.6: Research framework 149

Figure 3.7: Time series response to training with the pure ANN model, the Hang Seng futures market, 2002–2004 151

Figure 3.8: Time series response to training with the WNN model, the Hang Seng futures market, 2002–2004 151

Figure 3.9: Time series response to training with the WPCA-NN model, the Hang Seng futures market, 2002–2004 152

Figure 3.10: A screenshot of Bloomberg terminal with RSI indicator 156

Figure 3.11: WPCA Denoising Code for Matlab 158

Figure 3.12: Using a saved neural network in Matlab 161

Figure 4.1: Hurst exponent for all futures markets from 2005 to 2014 167

Figure 4.2: Approximate entropy for all futures markets from 2005 to 2014 170

Figure 4.3: Histogram of return distribution, Hang Seng futures market 177

Figure 4.4: Histogram of return distribution, KLCI futures market 177

Figure 4.5: Histogram of return distribution, KOSPI 200 futures market 178

Figure 4.6: Histogram of return distribution, NIKKEI 225 futures market 178

Figure 4.7: Histogram of return distribution, SiMSCI futures market 179

Figure 4.8: Histogram of return distribution, S&P 500 futures market 179

Figure 4.9: Histogram of return distribution, TAIEX futures market 180

Figure 4.10: Returns of the models without transaction costs: the results for the Hang Seng futures market 196

Figure 4.11: Returns of the models and strategies with transaction costs: the results for the Hang Seng futures market 197

Figure 4.12: Forecasting results of NN, WNN and WPCA-NN models for the Hang Seng futures market in out-sample data 198

Figure 4.13: Returns of the models and strategies without transaction costs: the results for the KLCI futures market 200

Figure 4.14: Returns of the models and strategies with transaction costs: the results for the KLCI futures market 201

Figure 4.15: Forecasting results of NN, WNN and WPCA-NN models for the KLCI futures market in out-sample data 202

Figure 4.16: Returns of the models and strategies without transaction costs: the results for the KOSPI 200 futures market 204

Figure 4.17: Returns of the models and strategies with transaction costs: the results for the KOSPI 200 futures market 205

Figure 4.18: Forecasting results of NN, WNN and WPCA-NN models for the KOSPI 200 futures market in out-sample data 206

Figure 4.19: Returns of the models and strategies without transaction costs: the results for the NIKKEI 225 futures market 208

Figure 4.20: Returns of the models and strategies with transaction costs: the results for the NIKKEI 225 futures market 209

Figure 4.21: Forecasting results of NN, WNN and WPCA-NN models for the NIKKEI 225 futures market in out-sample data 210

Figure 4.22: Returns of the models and strategies without transaction costs: the results for the SiMSCI futures market 212

Figure 4.23: Returns of the models and strategies with transaction costs: the results for the SiMSCI futures market 213

Figure 4.24: Forecasting results of NN, WNN and WPCA-NN models for the SiMSCI futures market in out-sample data 214

Figure 4.25: Returns of the models and strategies without transaction costs: the results for the S&P 500 futures market 216

Figure 4.26: Returns of the models and strategies with transaction costs: the results for the S&P 500 futures market 217

Figure 4.27: Forecasting results of NN, WNN and WPCA-NN models for the S&P 500 futures market in out-sample data 218

Figure 4.28: Returns of the models and strategies without transaction costs: the results for the TAIEX futures market 220

Figure 4.29: Returns of the models and strategies with transaction costs: the results for the TAIEX futures market 221

Figure 4.30: Forecasting results of NN, WNN and WPCA-NN models for the TAIEX futures market in out-sample data 222

LIST OF TABLES

Table 1.1: Overview of the study 27

Table 2.1: Hybrid intelligent systems compared with conventional systems for financial planning and forecasting 70

Table 2.2: A comparison between a HIS and one or more single intelligent models in the context of financial planning and forecasting 71

Table 2.3: A comparison of hybrid models and single intelligent methods in the same studies and with the same data 72

Table 2.4: Summary of the literature review 89

Table 3.1: Research overview 109

Table 3.2: Results of the Unit Root (ADF) test 115

Table 3.3: Results of the Breusch–Godfrey serial correlation LM test 117

Table 3.4: Results of the Granger Causality Test for all markets and all variables . 120

Table 3.5. Wavelet families and subsets 146

Table 4.1: Hurst exponent index results for all markets from 2002 to 2014 167

Table 4.2: Approximate entropy for all futures markets from 2005 to 2014 170

Table 4.3: Results of the forecasting performance, using a MAPE ratio, from 2005 to 2014 172

Table 4.4: Results of average annual returns from 2005 to 2014 173

Table 4.5: Results of average annual returns with transaction costs from 2005 to 2014 174

Table 4.6: Settings of the best-performing networks 174

Table 4.7: Descriptive analysis of trading returns from 2005-2014 176

Table 4.8: Sharpe ratios with trading strategies and models, including the WPCA-NN model 183

Table 4.9: Satterthwaite-Welch t-test results for the differences in means 186

Table 4.10: Regression results for the WPCA-NN model and the buy-and-hold strategy 189

Table 4.11: Regression results for the WPCA-NN model and the MACD trading strategy 191

Table 4.12: Regression results for the WPCA-NN model and the RSI trading strategy 191

Table 4.13: Regression results for the WPCA-NN model and the stochastics trading strategy 191

Table 4.14: Regression results for the WPCA-NN model and the OMA trading strategy 192

Table 4.15: Regression results for the WPCA-NN model and the pure NN model. 193

Table 4.16: Regression results for the WPCA-NN model and the WNN model 193

Table 4.17: Performance of the models and strategies measured by the MAPE ratios of the evaluation results for the Hang Seng futures market 195

Table 4.18: Returns of the models and strategies without transaction costs (in percentages) for the Hang Seng futures market 195

Table 4.19: Returns of the models and strategies with transaction costs: the results for the Hang Seng futures market 197

Table 4.20: Performance of the models and strategies measured by the MAPE ratios of the evaluation results for the KLCI futures market 199

Table 4.21: Returns of the models and strategies without transaction costs: the results for the KLCI futures market 199

Table 4.22: Returns of the models and strategies with transaction costs: the results for the KLCI futures market 200

Table 4.23: Performance of the models and strategies measured by the MAPE ratios of evaluation results for the KOSPI 200 futures market 203

Table 4.24: Returns of the models and strategies without transaction costs: results for the KOSPI 200 futures market 203

Table 4.25: Returns of the models and strategies with transaction costs: the results for the KOSPI 200 futures market 205

Table 4.26: Performance of the models and strategies measured by the MAPE ratio of evaluation results for the NIKKEI 225 futures market 207

Table 4.27: Returns of the models and strategies without transaction costs: the results for the NIKKEI 225 futures market 207

Table 4.28: Returns of the models and strategies with transaction costs: the results for the NIKKEI 225 futures market 209

Table 4.29: Performance of the models and strategies measured by the MAPE ratio of evaluation results for the SiMSCI futures market 211

Table 4.30. Returns of the models and strategies without transaction costs: the results for the SiMSCI futures market 211

Table 4.31: Returns of the models and strategies with transaction costs: the results for the SiMSCI futures market 213

Table 4.32: Performance of the models and strategies measured by the MAPE ratio of evaluation results for the S&P 500 futures market 215

Table 4.33: Returns of the models and strategies without transaction costs: the results for the S&P 500 futures market 215

Table 4.34: Returns of the models and strategies with transaction costs: the results for the S&P 500 futures market 217

Table 4.35: Performance of the models and strategies measured by the MAPE ratio of evaluation results for the TAIEX futures market 219

Table 4.36: Returns of the models and strategies without transaction costs: the results for the TAIEX futures market 219

Table 4.37: Returns of the models and strategies with transaction costs: the results for the TAIEX futures market 221

Table 4.38: Summary of the results 243

Table 5.1: Summary of conclusion 245

LIST OF SYMBOLS AND ABBREVIATIONS

A/D	:	Accumulation/distribution
ADF	:	Augmented Dickey Fuller
AMH	:	Adaptive Market Hypothesis
ANN	:	Artificial Neural Network
ATR	:	average true range
BB	:	Bollinger bands
BPNN	:	Back Propagation Neural Network
CART	:	Classification And *Regression Tree*
CCI	:	commodity channel index
CHO	:	Chaikin Oscillator
CME	:	Chicago Mercantile Exchange
CMF	:	Chaikin money flow
Coif	:	Coiflets
D	:	Delays
db	:	Daubechies
DJIA	:	Dow Jones Industrial Average
DPO	:	Detrended price oscillator
EMA	:	Exponential Moving Average
EMH	:	Efficient Market Hypothesis
EMV	:	Exogenous Maturity Vintage
ES	:	Expert System
EUR	:	Euro Dollar
FAHP	:	Fuzzy Analytic Hierarchy Process

FBNN	:	Feedforward Backpropagation Neural Network
FDM	:	Fuzzy Delphi Method
FNN	:	Fuzzy Neural Network
FTSE	:	Financial Times Stock Exchange
GA	:	Genetic Algorithm
GARCH	:	Generalized Autoregressive Conditional Heteroscedasticity
GBP	:	British Pound
GP	:	Genetic Programming
H	:	Hidden neurons
HCGA	:	Hybrid Chaotic Genetic Algorithm
HiCEFS	:	Hierarchical Coevolutionary Fuzzy System
HIS	:	Hybrid Intelligent System
HMM	:	Hidden Markov Model
ISMF	:	International Statistics Model for Forecasting
JCT	:	Johansen Cointegration Test
JPY	:	Japanese Yen
KLCI	:	Kuala Lumpur Composite Index
KOSPI	:	Korean Composite Stock Price Futures Index
LR	:	Linear Regression
MA	:	Moving Average
MACD	:	Moving Average Convergence Divergence
MAPE	:	Mean Absolute Percentage Error
MDA	:	Modern Driven Architecture
MI	:	Mass Index
MLP	:	Multi Layer Perceptron

MSCI	:	Morgan Stanley Capital International
MT	:	Momentum
NARX	:	Nonlinear Autoregressive with Exogenous input
NARX-NN	:	Nonlinear Autoregressive with Exogenous input - Neural Network
NASDAQ	:	National Association of Securities Dealers Automated Quotations
NIKKEI	:	Tokyo Stock Exchange Futures Index
NN	:	Neural Network
NPM	:	net profit margin
NYSE	:	New York Stock Exchange
OBV	:	on-balance volume
OHLC	:	Open High Low Close
OMA	:	Optimized Moving Average
OT	:	Oscillator
P/E	:	price/earnings
PCA	:	Principal Component Analysis
PO	:	Price oscillator
PVT	:	price and volume trend
RBF	:	Radial Basis Function
RBF NN	:	Radial Basis Function Neural Network
RNN	:	Recurren Neural Network
ROC	:	rate of change
RSI	:	Relative Strength Index
RWH	:	Random Walk Hypothesis
S&P 500	:	The Standard & Poor's 500 futures
SEC	:	Security Exchange Commission

SiMSCI	:	Singapore's MSCI (Morgan Stanley Capital International) futures
SMA	:	Simple Moving Average
SO	:	Stochastic Oscillator
SOM	:	Self Organization Map
STAR	:	Smooth Transition Autoregressive
STOCH	:	Stochastic Oscillator
SVM	:	Support Vector Machine
SVR	:	Support Vector Regression
sym	:	Symlets
TAIEX	:	Taiwan Stock Exchange futures index
TAR-VEC	:	Threshold Autoregressive Vector Error Correction
TRIX	:	Triple EMA
UK	:	United Kingdom
US	:	United States
USD	:	US Dollar
VMA	:	Variable Moving Average
WMA	:	Weighted Moving Average
WNN	:	Wavelet Neural Network
WPCA	:	Wavelet Principal Component Analysis
WPCA-NN	:	Wavelet Principal Component Analysis Neural Network

LIST OF APPENDICES

Appendix A: Additional Time Series Tests 282

Appendix B: Methodology of WPCA-NN step by step in Matlab 287

Appendix C: Hurst Exponent Radar Charts 307

Appendix D: Structural Breaks Figures 311

Appendix E: Approximate Entropy Radar Charts 315

Appendix F: Trading Results Around 2008 Financial Crisis.................................. 319

CHAPTER 1: INTRODUCTION

Most of human activity is stimulated, influenced, and determined by forecasts. Consider this example from daily life. A student going to school on time every day goes through a subconscious prediction procedure in his brain. Therefore, the time required to reach the school is a forecasting variable. The precision of this forecast is dependent on the distance between home and school, traffic, weather, and speed of walking, among other factors and assumptions. Most of these factors change frequently, which can result in the student arriving late. However, some factors are stable through time. Not considering the factors relevant for a forecast is equivalent to no forecasting at all. Therefore, forecasting is highly interrelated to the concept of uncertainty.

However, the effects of incorrect predictions in financial decisions are substantially greater than those of being on time for school on one day. Moreover, predictions in the financial domain are affected by a multitude of factors compared to this simple example, since the financial domain is so complex. Investors endeavor to predict events that will affect a firm, and then decide on whether the share price will go up or down. Forecasts of expected returns or future cash flows might impact a business decision to borrow or lend money. Since future inflation or unemployment trends are important for monetary policy changes, their extrapolation is the focus of economists in central banks. Thus, the development of precise financial prediction models is extremely important, especially in conditions of market uncertainty and economic crisis. It is under these circumstances that financial markets are noisy, and structural breaks and nonlinearities influence the usual financial and macroeconomic variables. Because of this, forecasting becomes interesting and challenging for investors, academicians, traders, policy makers, and relevant market practitioners. Under these circumstances, these groups try to successfully forecast economic and financial movement, using intelligent computational techniques, where traditional statistical methods would disappoint.

Artificial intelligence is a scientific field that creates and models tools that could attain human intellectual abilities. These abilities can be categorized as reasoning, understanding, and learning. An artificial intelligence system has the ability to recognize patterns, adjust computations, and demonstrate error tolerance (Bezdek, 1994). At the same time, the speed with which an artificial intelligence system deals with unexpected changes and errors comes close to the performance of the human brain. Thus, computational models have been widely applied in forecasting developments. Neural Networks (NNs) is commonly used in the extensive research regarding financial predictions (see amongst others Chen et al., 2003; Chen et al., 2009; Conejo et al. 2005; Huang et al., 2009; Huang et al., 2007; Kim & Han,2000). In recent years, artificial intelligence has evolved exponentially in financial trading because of more rapid reactions to temporary mispricing, and easier price management using computational trading systems that can learn from thousands of information sources without the barrier of human emotions (Boboc & Dinică, 2013; Harris, 2013). Technical analysis, which is the technique and science of deciphering historical data to forecast price movements, has also grown to include artificial intelligence models such as NNs (Dempster & Jones, 2001).

The adaptive nature of intelligent techniques differs from statistical techniques. They can adopt several forms and have numerous potential variables. These techniques have the advantage where the exact nature of the time series being studied is unknown, since statistical techniques cannot measure nonlinearity and have reached their limitations (Atsalakis & Valavanis, 2009b; Bahrammirzaee, 2010). Critics argue that these approaches are useless in finance, because of the absence of a supporting formal statistical theoretical context. However, financial series are subjugated by factors (such as behavioral elements, politics, etc.) that statistical analysis is incapable of capturing in a single model. Therefore, a statistical model that could capture such a time series trend is impractical in the long term. There is an open debate regarding the ability of intelligent techniques to dominate complex computational problems, because of their fundamental engineering architecture and unique utilization of existing financial data. However, intelligent models exhibit astonishing results.

Over-fitting may be one of the problems that is occurring in statistical derivations employing adaptable intelligent models. Over-fitting occurs when a learning machine is trained to perform well in a training period, but performs poorly in the testing period. One of the solutions to this issue is to separate the dataset into two periods: in-sample and out-of-sample. Other famous solutions to over-fitting are cross validation (Sermpinis et al., 2012), early stopping procedures (Prechelt, 2012), and pruning parameter approaches (Lidan et al., 2010).

Because of the huge number of inputs to the model, intelligent techniques also suffer from dimensionality problems. This is closely connected to the process of optimum feature selection, where, to improve its parameters, the technique chooses only suitable data subsets from the entire field of data. This problem can be solved with other techniques, such as filtering methods (Mundra & Rajapakse, 2007), principal component techniques (Aminghafari et al., 2006; Jolliffe, 2002), and embedded techniques (Hsieh et al., 2011). Ultimately, the low degree of theoretical interpretation is a serious drawback of some artificial intelligence techniques. Due to their computational complexity, people often consider them "black boxes," that require expert knowledge. However, over-simplifying these models causes problems with performance. Feature selection is one of the solutions to this issue. Applying or integrating fuzzy logic into these approaches might be another effective solution (Hua et al., 2007).

NNs, which are capable prediction tools (Zadeh, 1994), outperform traditional statistical methods and various other intelligent models (Fernandez-Rodrıguez et al., 2000; Refenes et al., 1994; Yoon et al., 1993)). The encouraging empirical results from NNs make them the primary choice for the majority of financial research. Further, the inadequacies resulting from the previously mentioned computational problems suggest that these techniques perform well in a task-specific modeling situation. This creates limitations in term of generalizing their performance to a more worldwide modeling structure. Of course, there are similar deficiencies in most of the modern econometric and statistical approaches. A current solution for overcoming these deficiencies is to apply hybrid models that combine the features of each technique (Taha & Ghosh, 1997), diminish over-fitting effects, and optimally handle the problem of high dimensionality (see amongst others Aminghafari et al., 2006; Busse et al., 2012; Cheng et al., 2010; Fernandez-Rodrıguez et al., 2000; Lotric & Dobnikar, 2005).

Considering NNs more specifically, they are able to establish complex relationships among training variables and targets, thus, they improve the chances to more accurately predict highly complicated and volatile trends in the markets (Wang & Fu, 2006; Wang & Gupta, 2013; Zurada, 1992). Researchers use artificial neural networks (ANNs) because of their qualities, including efficiency, performance, reproducibility, consistency, completeness, breadth, consistency of decision-making, and, most importantly, timeliness (Bahrammirzaee, 2010). However, ANNs are limited because of the complexities of financial markets and the noise that such markets display (Chang et al., 2004). On the other hand, like other artificial intelligent models, ANNs can be used with other machine-learning models in parallel, transformational, or sequential methods so as to overcome their limitations and deficiencies (Chen & Leung, 2004; Garliauskas, 1999; Hassan et al., 2007; Hsieh et al., 2011, Kim & Shin, 2007, Zhang & Wu, 2009). Because of the vulnerability of NNs to futures' price noises, this study combines them with wavelet analysis. Wavelet transform is a mathematical tool for analyzing signals in the two main domains of frequency and time (Ramsey & Zhang, 1997). Wavelet transform has become a widespread technique for single-dimensional signal filtering and data mining in univariate denoising algorithms (Aminghafari et al., 2006; Donoho, 1995; Hsieh et al., 2011). Several studies have proved that the wavelet transform denoising process improves the performance of NNs time series forecasting (Hsieh et al., 2011; Jin & Kim, 2015; Lotrič, 2004; Lotric & Dobnikar, 2005; Moazzami et al., 2013). This results in one of the models being used in this study.

A study by Aminghafari et al. (2006) offers a multivariate denoising procedure using principal component analysis (PCA) and wavelet analysis. PCA is one of the best-known data analysis techniques, especially designed to simplify multi-scale signals by tracing new factors and obtaining the main features of data (Bakshi, 1999). Aminghafari et al. (2006) believe that denoising multivariate signals using Wavelet Principal-Component-Analysis, namely WPCA, outperforms univariate wavelet denoising by deriving the same noise at different frequencies from the elements of a multivariate signal. Another significance of this study and contribution to the existing literature is that it is the first attempt to use this multivariate denoising technique to denoise the open–high–low–close (OHLC) index as a multivariate signal to feed to a NN for forecasting future price movements. Through this solution, which is used to analyze the OHLC as a multivariate signal, there is an opportunity to extract the common noise components of these four signals more accurately. Thus, the proposed WPCA-NN model can overcome the limitations and consequently forecast more accurately compared to NN and WNN approaches. In this experiment, three different models, the NN, WNN, and WPCA-NN, are evaluated not only against each other, but also against the threshold buy-and-hold and best performing technical analysis strategies in predictive accuracy and trading performance results.

In accordance with the random walk hypothesis (RWH) and a weak form of the efficient market hypothesis (EMH), market indices or stock prices should exhibit no pattern that could enable future indices or prices to be predicted with any consistency. Thus, returns on assets are required to be normally distributed and consecutively independent. In other words, returns on assets should not exhibit any certainty or long-term memory of the series of prices that they follow. Therefore, the primary objective of this study is to assess the efficiency and predictability of certain markets. If the time series reveals either persistence or anti-persistence, then a trading strategy using that information can earn abnormal returns. In contrast, returns on assets that are neither persistent nor anti-persistent are volatile; hence, they cannot be used for profit, consistent with one of the assertions of the EMH. Although advocates of the RWH and EMH consider that financial markets are not predictable (Fama, 1965, 1991), many studies that are based on recent and historical data oppose this view and argue in favor of the predictability of financial time series (Taylor, 1986; Thaler, 1985). This is despite the fact that the evidence of the predictability of stock market returns persuades scholars to study the source of this predictability (Gencay, 1996).

The adaptive market hypothesis (AMH) is an alternative hypothesis to the RWH and EMH, and can coexist with them in a logically stable way (Lo, 2004). This theory applies the evolutionary approach of biology and a behavioral perspective to economic interactions and demonstrates that the behavior of market agents, and thus the market, is adaptive (Lo, 2004). Since markets are flexible and adaptable, and make progress over time, they switch between inefficiency and efficiency (Lo, 2004). In addition, the AMH suggests that the degree of a market's efficiency is related to the environmental determinants that represent the market ecology, including the extent of available profitable opportunities, the number of competitors, and the adaptability of the participants in the market (Lo, 2005). In the AMH framework, investors commit mistakes and learn to adapt their behaviors in accordance with the framework, which is not the case with investors in efficient markets. Overall, if the studied markets follow the AMH, they are inefficient and predictable. Given the foregoing perspective, this study also aims to investigate whether the AMH can provide a better explanation of futures indices.

In addition to traditional techniques for efficiency analysis, such as unit root and serial correlation tests, there are modern mathematical factors that are used in this study. Long-term memory and entropy are measurements for capital market efficiency and predictability (Kristoufek & Vosvrda, 2014). Long-term memory shows that market time series are portrayed by serial reliance as well as periodic long cycles (Cheung & Lai, 1995). When market trends are random walk, the variations are an entirely uncorrelated string of numbers (Zunino et al., 2010). In this regard, the string of data is totally disordered; thus, its entropy is maximized. Therefore, the value of entropy can represent the efficiency of the market. The value of long-term memory and entropy, and the changes in their values over time would provide a proof of market inefficiency for the particular markets in the particular period of time, according to the AMH. Consequently, the selected markets will be predictable in that particular time period. All the above are the significance of the study, contribute to the literature, and provide motivation for putting considerable effort in identifying the best forecasting techniques.

1.1 Background of the Study

In this section, a brief history of the financial theories and hypotheses regarding to the market behavior is introduced. There are different schools of thoughts in finance regarding to the forecasting domain. Followings are the financial theories, which support or are against prediction in the financial markets.

1.1.1 Random Walk Hypothesis

Early in the prior century, statisticians remarked that variations in stock prices appear to track a fair-game pattern. This finding led to the random walk hypothesis (RWH), first supported by the French mathematician Louis Bachelier in 1900 (Bahrammirzaee, 2010). The RWH indicates that stock prices are random, such as the steps taken by a drunken person, and are thus unforeseen.

Some further research in this area was undertaken in the 1930s; however, the RWH was not debated and considered intensively until the 1960s. The present consensus is that the RWH is justified by the weak form of the efficient markets hypothesis so called EMH (Malkiel & Fama, 1970).

For several years, economists, researchers, and financial analysts have been studying approaches to forecast the trend of individual stocks. Technical theorists and chartists consider that historical patterns can be applied to estimate future prices. However, the random walk theory states that such trends or movements cannot be precisely forecasted.

As part of the debate, this study compares the RWH to other common theories, such as the EMH, technical analysis, and fundamental analysis. Following this, the study presents the background of the RWH. Then, the study sums up with a concise debate of the implications of the theory. There are mainly two competing methods for forecasting the trends of financial markets: technical and fundamental analyses. Next, these competing approaches are discussed, followed by a consideration of how the RWH is associated with each.

Chartists that use technical methods consider that the historical data of a stock's price or market index can be employed to forecast the future price trend (Bahrammirzaee, 2010). Using such methods, the technical trader studies the structure of the upward and downward trends of a stock or market. The pattern of these trends helps the trader to determine what he or she imagines will be the future trend for the stock or market.

Fundamental analysts consider that a stock price is a function of its inherent value, which has a significant relationship to a company's likely future earnings (Atsalakis & Valavanis, 2009b). By considering fundamental parameters such as economic news, industry trends, and a company's earnings per share, a fundamental analyst can specify whether a stock's price is over or under its inherent value. Comparing a stock's price with its core value enables a fundamental analyst to forecast the probable future movement of the stock's price.

The next hypothesis to consider is the EMH. The EMH states that the price of a stock reflects all publicly known news and information about a firm (Malkiel & Fama, 1970). In fact, scholars subscribing to what is known as the "strong" EMH state that stock prices also reflect what insiders know. Since public and individual data about a company is immediately reflected in the market price of a stock, a trader or an investor finds it difficult to gain positive excess returns above the buy-and-hold benchmark. The principles of the RWH are generally the same as those of the EMH.

Ultimately, the RWH considers that prices of stocks cannot be forecasted. The stock price and the market are "informationally efficient," and the individuals buying and selling stocks include a huge number of rational traders and investors with access to this incoming data (Malkiel, 2007). Since long-term prices reflect the functioning of a company over time, short-term fluctuations in prices can best be explained as a random walk.

While the RWH can trace its basis to an early-twentieth century mathematician, today's hypothesis can be credited to Eugene Fama's 1965 doctoral thesis, "The Behavior of

Stock-Market Prices," along with the publication of his 1995 "Random Walks in Stock Market Prices." Since the RWH is rooted in an efficient market, prior patterns cannot be employed to forecast future trends in any type of meaningful way. The pattern of a stock is always unsystematic (Malkiel & Fama, 1970).

The RWH has been verified by substantial empirical research. Using actual data, researchers have concluded that there is no correlation between continuous price changes (Malkiel & Fama, 1970). In other words, the next direction of a stock is entirely independent of its prior trends. Indeed, the direction of individual stocks, as well as stock markets, is just as coincidental as the result of flipping a coin (Malkiel, 2007).

1.1.2 The Efficient Market Hypothesis

The EMH states that financial markets are efficient, that price movements reflect all previously released information about a stock or other security, and that prices quickly adjust to any fresh information (Malkiel & Fama, 1970). Information consists not only of what is recently known concerning a stock, but also any upcoming prospects, such as dividend yields or earnings. The EMH seeks to clarify the RWH by suggesting that only new information will affect stock prices significantly; since new data or information are currently unspecified and unpredictable, future trends in stock prices or indices are also unspecified and consequently move randomly (Malkiel, 2007). Thus, it is not possible to perform better than the market by selecting undervalued indices or stocks because the EMH states that there are no overvalued, or even undervalued, stocks.

The foundation of the EMH is that the market has many rational investors and traders who are always aware of the information and news and who react rapidly to any fresh and important news regarding a stock. There are also many funds whose managers are always aware of new reports and information. Moreover, with the help of high-grade computers, fund managers are continually searching through financial information for mispriced stocks. High-speed investors similarly use high-grade computers located close to exchanges to perform trades based on price differences between stocks on different exchanges or between related securities that have interconnected prices, such as stocks and options.

Thus, the EMH relies on the following.

1) News is broadly accessible to all investors.
2) Investors employ this data to study the economy, the markets, and specific stocks to make investment decisions.
3) Almost all events that have a major influence on securities' prices, such as workforce strikes, litigation, and coincidences, are random and usually unpredictable events; and when they do occur, they are rapidly transmitted to investors.
4) Traders react rapidly to any fresh data.

There are three levels or forms of the EMH that vary in accordance with the information that they reflect. In the weak form of the EMH, only historical market trading data, such as trading volumes, short interests, and stock price movements, are noted. Thus, even the weak form of the EMH indicates that technical analysis does not create profits because such analysis rests entirely on historical trading information to predict future price trends.

The semi-strong form of the EMH expands data to include public data in addition to market information, such as news, company management details, accounting reports, company products, patents, and analysts' recommendations (Malkiel, 2007).

The strong level of the EMH expands the data to comprise not only public data but also private data that are usually held by company insiders, such as a company's executives and officers (Malkiel, 2007). Clearly, insiders can make profits by trading their company's stock before a major change to the company is broadcast to the public. This is the reason insider trading is forbidden by the U.S. Securities and Exchange Commission (SEC). Company insiders can invest in their stock, but only if an investment is not made prior to an important change that only a few insiders know about, such as a new product line, a merger, or other significant activity. Since this study does not deal with information flow, it will only test whether the studied markets follow the weak form of EMH (also known as RWH).

1.1.3 Adaptive Market Hypothesis

Regardless of the significant amount of research on the EMH regarding developing and developed markets, agreement on whether analyses of markets are efficient remains elusive. Recently, notable evidence found that stock returns are not volatile, and retain certain elements of predictability; however, there is an absence of any robust alternative theoretical clarifications about the EMH (Atsalakis & Valavanis, 2009b; Bahrammirzaee, 2010). Nonetheless, by making use of an evolutionary approach to economic interaction, one study suggests that the adaptive market hypothesis (AMH) can coexist with the EMH in a logically stable way (Lo, 2004). The developing and emerging markets have greater propensities to refute the EMH because of numerous market frictions. The AMH accommodates such market frictions and suggests that markets advance or make progress over time. This perspective is not the case with the EMH, which only considers frictionless markets. Given the foregoing perspective, this current study aims to investigate whether the AMH can provide a better explanation of future stock indices. The AMH is an alternative market theory to the EMH from a behavioral perspective. The theory states that markets are flexible and adaptable and that they switch between inefficiency and efficiency at different times (Lo, 2004). The theory also applies the evolutionary approach of biology to economic interactions and demonstrates that the behavior of agents, and thus the market, is adaptive (Lo, 2004). Further, the theory suggests that the degree of a market's efficiency is associated with the environmental determinants that represent the market ecology, including the extent of available profitable opportunities, the number of competitors, and the adaptability of the participants in the market (Lo, 2005). However, the AMH highlights the law of the "survival of the richest," or natural selection, and regulates the evolution of institutions and markets in the real world with frictions (Charles et al., 2012). In the AMH framework, investors commit mistakes and learn to adapt their behaviors in accordance with the framework, which is not the case with investors in efficient markets (Lo, 2004). There are numerous practical repercussions under the AMH (Lo, 2004). First, the risk and reward relationship varies over time because of populations' preferences within a market. Second, the changes in past prices affect current preferences because of the forces of natural selection. Thus, the AMH compares to the weak type of efficiency, where there is no usage of historical prices. Third, time-to-time arbitrage opportunities exist in an adaptive market. From the evolutionary perspective, the opportunities to earn profits consistently appear and disappear. This situation requires investment strategies consistent with the prevailing market environment. Thus, the AMH entails "complex market dynamics," which makes active portfolio management necessary. Fourth, innovation is imperative for survival; consequently, the AMH recommends adapting to variable market situations to ensure consistency in expected returns. Finally, market efficiency is an attribute that constantly changes across markets and over time. Thus, financial markets may observe times of both inefficiency and efficiency.

According to the RWH, the future movements of stock prices are not possible to predict using historical data of stock prices or volume (Fama, 1965). Hence, technical analysis and any other trading strategies based on technical methods are unlikely to generate excess returns. Based on this hypothesis, the way to obtain the maximum return from a stock over a given period is to buy and hold it for that period (Fama, 1965). Thus, if any trading strategy based on technical analysis produces an excess return compared with a buy-and-hold strategy, it will challenge the support of the RWH and the weak form of the EMH in tested markets. Moreover, if any trading strategy significantly generates excess returns in different stocks or markets (generalizability) and over various periods (reliability), it will, with high probability, contradict the RWH and the weak form of the EMH. In such a context, this study offers the WPCA-NN model to address the contradiction.

1.2 Problem Statement

The ability to consistently and, as accurately as possible, forecast the future is essential to many decision procedures for the timing of financial transactions, planning, scheduling, organizing, policy making, strategy formulation, purchasing and supply chain management, and so on (Atsalakis & Valavanis, 2009b). Forecasting is a significant and active area of human activity and will continue to be in the future (Zhang et al., 2014). Recently, artificial intelligence systems are the most accepted methods implemented in financial markets (Bahrammirzaee, 2010). Indeed, one study shows that artificial intelligence models and machine-learning techniques, especially hybrid models, outperform conventional statistical models for analyzing financial issues in terms of times-series and nonlinear trends, although this superiority is uncertain (Bahrammirzaee, 2010). However, the RWH is at odds with market forecasting. The current study examines whether the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets follow the RWH or contradict it.

However, a problem arises because trading with technical rules has generated positive returns exceeding the threshold buy-and-hold for many markets over time (Brock et al. 1992; Irwin & Park, 2004; Lukac et al., 1988; Yao et al., 1999). The excess returns and presence of entropy and long-term memory represent the predictability and inefficiency of financial time series (Lo, 1989, 2004, 2005; Mandelbrot, 1971; Mandelbrot & Hudson, 2005; Mandelbrot & Wallis, 1969a; Sadique & Silvapulle, 2001; Shafie-Khah et al., 2011; Tolvi, 2003; Urquhart & Hudson, 2013; Rafal, 2002; Willinger et al., 1999). This situation contradicts the RWH and supports the AMH (Lo, 1989, 2004).

The current study seeks to resolve the controversy surrounding the problem of excess returns and noise in financial time series. It endeavors to do this with approximate entropy and long-term memory analyses and the support of the adaptive market hypothesis, in addition to the use of a hybrid model of wavelet and neural network, namely WNN, to generate significantly higher returns above the buy-and-hold threshold. Moreover, a new hybrid model of forecasting that employs wavelets, PCA, and an ANN (WPCA-NN) is offered in this study in order to enhance the denoising process and achieve higher forecasting performance and return than those by WNN. Because of the limitations of ANNs regarding futures' price noises, wavelet analysis is applied combined with an ANN in this study.

1.3 Research Questions

To address the formulated problem statement, this research seeks to find answers to the questions which can be constructed as follows.

1) Do futures markets, particularly stock index futures, have long-term memory and approximate entropy?
2) If so, can technical analysis indicators such as the newly created hybrid model, the WPCA-NN model, consistently generate higher abnormal returns than those advocated by the RWH, namely, the buy-and-hold strategy, for selected futures markets?

3) Does the WPCA-NN model consistently generate significantly higher returns for futures markets than the best performing technical analysis indicators, such as Moving Average Convergence/Divergence (MACD), Relative Strength Index (RSI), Stochastics, and Moving Average (MA)?
4) Does the WPCA-NN model consistently generate significantly higher returns and achieve higher forecasting performance for futures markets than the pure NN and WNN methods of forecasting?

This study tries to answer whether the present hybrid model achieves positive excess returns. Moreover, this study investigates whether the offered hybrid model, the WPCA-NN, is superior and more successful in denoising financial time series.

1.4 Research Objectives

Based on the formulated research questions, this study attempts to attain the following objectives.

1) To investigate whether the futures markets exhibit significant long-term memory.
2) To determine if the hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the RWH (a passive buy-and-hold strategy) for selected financial markets.
3) To investigate, test, and establish whether, the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other common technical analysis indicators (such as optimized moving average, MACD, RSI, and Stochastics).
4) To investigate, test, and establish whether:
    a. The WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other pure NN methods of forecasting; and

b. The WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than the WNN models.

1.5 Research Hypotheses

According to the research questions and objectives, followings are the hypotheses this research tests:

1) Hypothesis 1 or H1: Futures markets exhibit significant changes in predictability over time, taking into account long-term memory and approximate entropy.
2) Hypothesis 2 or H2: The hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the random walk hypothesis (a passive buy-and-hold strategy) for selected financial markets, as shown by significant mean differences and significant positive alphas in profit regressions.
3) Hypothesis 3 or H3: Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other best-performing technical trading rules (such as MACD, RSI, OMA and Stochastics).
4) Hypothesis 4 or H4: Significant mean differences and significant positive alphas in profit regressions show that:
    a. Hypothesis 4a or H4a: The WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other pure ANN forecasting methods.
    b. Hypothesis 4b or H4b: The WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than a WNN model.

1.6 Purpose of Study

The first purpose of this study is to investigate long-term memory and trends in futures time series as shown by the abnormal higher returns exceeding the passive buy-and-hold strategy advocated by the RWH. By investigating long-range dependence, or long-term memory, in the futures markets based on the Hurst exponent, and approximate entropy, this study's results will examine the predictability of future trends in accordance with historical data. Generating consistent returns exceeding the passive buy-and-hold strategy as advocated by the RWH indicates that market trends can be predicted through the analysis of historical data. To generate higher returns than the passive buy-and-hold and best-performing technical indicator strategies, this study offers a combination of wavelet denoising process and neural network forecasting technique. In addition to that, to enhance the denoising process and generate higher returns than WNN, this study offers a novel hybrid of multivariate wavelet denoising process with neural network forecasting technique, namely WPCA-NN. Moreover, the offered novel hybrid model is the first attempt to apply a combination of multivariate signal denoising with a feedforward backpropagation neural network (FBNN) model. It is believed that this denoising process, which is a known deficiency of artificial NNs (Chang et al., 2004), will add a great deal of accuracy to the prediction results compared with a simple FBNN model and a univariate-wavelet signal denoising model with an NN (WNN). This study posits that the offered novel hybrid model is superior to the FBNN and WNN models, the buy-and-hold strategy, and the best-known technical trading strategies presented in the literature (Chong et al., 2010; Chong & Ng, 2008; Fernández-Blanco et al., 2008; Rosillo et al. 2013).

Another purpose of this research is to offer a timely method for making better trading decisions to trading practitioners in financial markets where studies show that excess returns above the buy-and-hold strategy using the basic technical indicators of yesteryear are fast declining (Olson, 2004). In this context, empirical evidence suggests that financial markets are evolving and increasing their efficiency over time (Coronel-Brizio et al., 2007). Further, with increasing efficiency, abnormal profits are harder to come by (Pukthuanthong et al., 2006).

This research contributes to the literature of financial market forecasting and modern techniques. Thus, it will interest market traders, who, in the current increasingly difficult and volatile markets, find that establishing their trading decisions solely on traditional technical analysis signals is not as profitable as was previously. The empirical results of this research support prior academic literature (Bahrammirzaee, 2010; Barkoulas & Baum, 1996; Mandelbrot & Wallis, 1968; Behradmehr & Ahrari, 2015; Bildirici et al., 2010; Bildirici & Ersin, 2009; Chandwani & Saluja, 2014; Chang & Fan, 2008; Charbonneau & Kharma, 2009; Chavarnakul & Enke, 2008; Chen & Leung, 2004; Chen et al., 2009; Cheng et al., 2010; Cheung et al, 2004; Donoho, 1995; Garliauskas, 1999; Godfrey, 1978; Granger, 1969; Hassan et al., 2007; Huang et al., 2009; Kanas & Yannopoulos, 2001; Kim, 2003; Kristoufek, 2012; Kristoufek & Vosvrda, 2014; Lo, 2004; Luther, 1998; Ni & Yin, 2009; Shah & Murtaza, 2000; Taffese, 2007; Tsakonas et al, 2006; Wang, 2009). Such literature provides evidence of the successful forecasting of future price movements by using machine-learning methods such as ANNs, genetic programming, and wavelet analysis.

1.7 Scope of the Study

This study tests efficiency and predictability in the futures markets through the analysis of the unit root test, serial correlation, approximate entropy, and long-term memory. The approximate entropy is an estimation that enables different degrees of correlation to be distinguished, and Gaussian processes to be distinguished from non-Gaussian, which helps to analyze the efficiency of a financial time series (Zunino et al., 2010). The long-term memory analysis is conducted by the Hurst exponent (Mandelbrot & Hudson, 2005), which is used as a measure for long-term memory and predictability in financial time series. This study uses the denoising ability of wavelet transforms along with a nonlinear autoregressive neural network with exogenous inputs (NARX-NN) to establish a forecasting model, namely WNN. In addition to that, another forecasting model is used in this study that applies a denoising technique other than univariate wavelet denoising with ANN. In the latter denoising technique, namely wavelet principal component analysis, the open-high-low-close index is considered as a multivariate signal. The combination of wavelet transform and principal component analysis gives the ability to denoise a multivariate signal. This study is the first attempt to enable a hybrid of wavelet, PCA, and ANN.

The latter model, the WPCA-NN, is compared with WNN and pure NN models in seven futures markets (the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX) to compare forecasting performances and trading profitability. The comparison between the WPCA-NN, WNN, and pure NN models shows the performance of the noise reduction process that this study offers for futures indices. The following is a summary of what is undertaken in this study and Figure 1.1 comprehensively illustrates the scope of this study.

1) Testing predictability and efficiency through approximate entropy and long-term memory in all selected markets.
2) Examining the support of the RWH and AMH in selected futures markets.
3) Establishing technical trading strategies to obtain excess returns above a buy-and-hold strategy.
4) The pure NN model: using only prediction (NN) and then applying the trading strategy.

5) The WNN model: using a simple univariate denoising process with forecasting (NN) and then applying the trading strategy.
6) The WPCA-NN model: using a more complex denoising process with forecasting (NN) and then applying the trading strategy.
7) Using seven futures markets to compare and check the reliability and generalizability of the offered hybrid system.

Problem Statement (1.2)
To investigate the gap between market traders using technical rules and the researcher's knowledge of technical analysis.

Research Objectives (1.4)
1) To investigate whether the futures markets exhibit significant entropy and long-term memory.
2) To determine whether the hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the RW H (passive buy and hold) for selected financial markets.
3) To investigate and report if the hybrid WPCA-NN model consistently generates significantly higher returns for the given financial markets than the best performing technical strategies.
4) To investigate and report whether the WPCA-NN model consistently outperforms and generates significantly higher returns for the futures markets than the WNN and NN models.

Literature Review (Chapter 2)
A literature review is conducted to investigate the existing gaps in the topic and review existing theory to create an innovative methodology and form hypotheses.

Hypothesis 1 (3.7)
Hypothesis 1: All selected markets have entropy and long memory in accordance with their historical data.

Measurements (3.3, 4.1.4)
Calculating entropy by approximate entropy and long-term memory by the Hurst exponent index suggested by Mandelbrot and Hudson (2005).

Hypotheses 2, 3 and b (3.7)
Hypothesis 2: The WPCA-NN model produces higher returns than threshold passive buy-and-hold (required by RWH).
Hypothesis 3: The WPCA-NN model produces higher returns than other technical analysis indicators for futures markets.
Hypothesis 4a: The WPCA-NN model produces higher returns and performance than a pure NN model for futures markets.
Hypothesis 4b: The WPCA-NN model produces higher returns and performance than a WNN model for futures markets.

Data Tests (3.10, 4.1)
Data tests and analysis include a unit root test, serial correlation or autocorrelation test, and a Granger causality test.

Trading Systems (3.4, 3.8, 3.9)
Test trading profitability and forecasting performance by using the RWH, other technical strategies, NN, WNN, and WPCA-NN models.

Descriptive Statistics of Results (3.11, 4.2)
T-test (mean, skewness, and kurtosis), significant alpha and Sharpe ratio.

Results (4.3, 4.4)
According to the data tests, approximate entropy and long-term memory analysis; all selected markets are inefficient, predictable, and follow the adaptive market hypothesis. Forecasting by the WPCA-NN model outperforms the WNN, NN, buy-and-hold and best-performing technical trading strategies. Trading profitability with the WPCA-NN model is higher than returns from the WNN, NN, buy-and-hold, and other technical trading systems in all the tested markets. Moreover, the WPCA-NN model presents less sensitivity and better performance during a financial crisis. All the return results and comparisons between different strategies are significant in the descriptive statistics.

Conclusion (Chapter 5)
All selected futures markets possess long memory and not complete entropy so are inefficient and predictable, according to the study's results. The study's promising outcomes indicate the prediction power and high denoising performance of the WPCA-NN model. Ultimately, the results suggest that the adaptive market hypothesis gives better support for the selected futures markets.

Figure 1.1: Scope of the study

1.8 Justification of the Study

The main purpose of this study is to obtain positive excess returns and overcome the noises that exist due to the nature of financial time series.

Several traditional and statistical models have been used in order to model and forecast financial time series with the goal of obtaining positive excess returns (Bildirici et al., 2010; Bildirici & Ersin, 2009; Charbonneau & Kharma, 2009; Chavarnakul & Enke, 2008; Chen & Leung, 2004; Garliauskas, 1999; Hassan et al., 2007; Huang et al., 2009; Ni & Yin, 2009; Wang, 2009). Since traditional statistical methods have reached their limitations (Atsalakis & Valavanis, 2009b; Bahrammirzaee, 2010), machine-learning systems are currently used to forecast financial time series. Using modern intelligent technique is one of the main motivations behind this research. ANN, as a capable prediction tool (Zadeh, 1994), outperforms traditional statistical methods and various other intelligent models (Fernandez-Rodrıguez et al., 2000; Refenes et al., 1994; Yoon et al., 1993). ANNs are able to establish complex relationships between training variables and targets; thus, they improve the chances of predicting highly complicated and volatile trends in the markets (Wang & Fu, 2006; Wang & Gupta, 2013; Zurada, 1992). However, the capability of ANNs are limited because of the complexities of financial markets and the noise that they represent (Chang et al., 2004).

Hence, in order to use an ANN, its deficiency with regard to noise must be overcome by using a denoising technique (Wang & Gupta, 2013). Combining different machine-learning techniques into a new effective learning system that has the best-performing and strongest features of each approach, while eliminating the defects and weak points, can provide a better representation of machine-trading systems in terms of the ability to process and learn within non-symbolic and symbolic paradigms (Taha & Ghosh, 1997). Wavelet analysis is applied in combination with an ANN in this research because of the vulnerability of ANNs to futures price noises (Pindoriya et al, 2008; Shafie-Khah et al., 2011).

Several studies use wavelet transform and an ANN as a hybrid to achieve better forecasting performance (Jin & Kim, 2015; Lai et al., 2007; Lotrič, 2004; Ortega, 2012; Wang & Gupta, 2013); however, this current study adjusts a multivariate denoising technique (Aminghafari et al., 2006) with an FBNN model in order to establish a better denoising procedure and achieve higher forecasting performance and excess returns.

The previously mentioned futures markets were chosen because, first, they are typical markets that an Asian fund manager in the Asian time zone would trade in. Second, this study is testing a novel approach with a hybrid model; thus, it is necessary to test the model with a range of emerging markets and developed markets for generalizability and stability. Finally, these markets are those studied by other scholars (Chen et al., 2003; Dai et al., 2012). Additionally, Bursa Malaysia offers the highest growth of 10.5% in terms of the dividend yield of benchmark index performance, and possesses the largest amount of funds raised among countries of the Association of Southeast Asian Nations (Errunza, 1994). Because of the high returns in emerging markets, investors are attracted to them to enhance their performance and diversify their portfolios (Errunza, 1994). Malaysia has one of the largest market capitalizations among Asian countries, with more than 900 listed firms from 2009 to 2012. Bursa Malaysia is a typical Asian emerging market; thus, any study of this market adds to the literature on international investment.

Moreover, the WPCA-NN model is expected to outperform the WNN model in terms of forecasting performance and trading profitability. It is anticipated that the results of the offered model outperform the buy-and-hold benchmark proposed by the RWH (Fama, 1965), other commonly used technical trading indicators, and univariate wavelet denoising with ANN and FBNN models. All the above tied together in a broad framework, motivates this thesis.

To evaluate the performance of these models, this study tests these trading systems in seven markets: Hong Kong's Hang Seng futures, Malaysia's Kuala Lumpur Composite Index (KLCI) futures, Japan's NIKKEI 225 futures (NIKKEI), Singapore's Morgan Stanley Capital International (MSCI) futures (SiMSCI), the Korea Composite Stock Price Index (KOSPI) 200 futures of South Korea, the Standard & Poor's 500 (S&P 500), and Taiwan's Stock Exchange (TAIEX) futures markets. Apart from Japan's NIKKEI 225 and S&P 500 futures market, these are the futures markets of the Asian Tiger economies, which studies show are experiencing rapid growth, not only in monetary terms but also in importance to the current world economy in relation to the global trend toward diversification (Dunis & Shannon, 2005; Errunza, 1994; Li, et al., 2003). These markets are either developing or developed countries with different economic conditions, which consequently will result in more general and robust outcomes. Additionally, all of these markets have been affected by the 2008 financial crisis. The sample data of this study is collected from Bloomberg and consists of 10 years' worth of Open, High, Low, Close (OHLC) and popular technical indicators: Relative Strength Index (RSI), Moving Average Convergence/Divergence (MACD), MACD Signal, Stochastic fast %K, Stochastic slow %K, Stochastic %D, and Ultimate Oscillator.

1.9 Organization of the Study

In the second chapter, this study continues with a review of the literature that considers financial market forecasting in the context of different schools of thought. The review surveys the background of technical trading and describes the variables, tools, and methods that researchers use to forecast the market. Moreover, the chapter reviews studies that examine the RWH: studies that support the hypothesis and those that contradict the hypothesis. The chapter also reviews studies that investigate the AMH, which contradicts the RWH. The traditional and modern models that forecast financial time series and achieve excess returns are then presented. The hypotheses of this study are developed from the literature review and our objectives.

In the third chapter, the data sample, the selected futures markets, the methods of collecting data, the quality of the in-sample data and out-sample data, and the data analysis tests, such as the unit root test, serial correlation test, Johansen cointegration test, Granger causality test, and long-range dependence test (the Hurst exponent), are presented. Most importantly, this chapter shows this study's methods, such as wavelet transform, PCA, and ANN, and the quality involved in the creation of the novel hybrid model. The methods used to attain accuracy, performance measurement, and the trading profitability of the studied models are then presented.

Statistical tests of historical information, the results of the forecasting of the futures markets, and the significance of the results are described in the fourth chapter, together with a comprehensive explanation of the results. After this, the study's conclusion and the scope for future contributions are presented. Table 1.1 provides an overview of the study's organization.

Table 1.1: Overview of the study

Problem statement (Chapter 1)	This thesis seeks to resolve the controversy surrounding the problem of excess returns and noise in financial time series with the use of a new hybrid forecasting model using a wavelet, PCA, and an ANN (WPCA-NN) model to generate significantly higher returns above the buy-and-hold threshold. Because of the limitations of ANNs regarding futures price noises, wavelet analysis is applied combined with an ANN. The goals are: 1) To obtain higher returns than the RWH (positive excess returns); and 2) To overcome the noises that exist due to the nature of financial time series.
Objectives (Chapter 1)	1) To investigate whether the futures markets exhibit significant changes in predictability over time, considering long-term memory and approximate entropy. 2) To determine if the hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the RWH (a passive buy-and-hold strategy) for selected financial markets. 3) To investigate, test, and establish whether the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other technical analysis indicators (MACD, RSI, stochastics, and optimized moving average (OMA)). 4) To investigate, test, and establish whether: a) The WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other pure ANN methods of forecasting.

	b) The WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets compared with the WNN.
Hypotheses (Chapter 2)	1) Futures markets exhibit significant changes in predictability over time, considering long-term memory and approximate entropy. 2) The hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the RWH (a passive buy-and-hold strategy) for selected financial markets, as shown by significant mean differences and significant positive alphas in profit regressions. 3) Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other technical analysis indicators (3a. MACD, 3b. RSI, 3c. stochastics, and 3d. OMA). 4) Significant mean differences and significant positive alphas in profit regressions show the following: a) The WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other pure NN methods of forecasting. b) The WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets compared with the WNN model.
Tests (Chapter 3)	• Data analysis tests: 1. Augmented Dickey–Fuller (ADF) unit root test 2. Serial autocorrelation 3. Granger causality 4. Hurst exponent index test • Forecasting accuracy with mean absolute percentage error (MAPE) (for the WPCA-NN, WNN, and NN models)

	• Trading strategy return: buy and hold, common technical trading indicators (such as MACD and RSI), and the WPCA-NN, WNN, and NN models • Significance of results' tests: 1. Skewness and kurtosis 2. T-test: comparison of two means 3. Significant alpha above market return 4. Reliability of results and risk of trading (Sharpe ratio)
Empirical Results (Chapter 4)	Empirical results of long-term memory, forecasting performance, and trading profitability (of a buy-and-hold strategy, common technical trading indicators, and the WPCA-NN, WNN, and pure NN models) for seven futures markets. Data from 2005 to 2014.
Conclusions (Chapter 5)	Conclusions, implications, future contributions, and deficiencies of the study.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter discusses a review of the literature on market theories that contradict and support market predictability, and a review of the literature on prediction tools, such as ANNs, wavelet analysis, and other technical trading rules, which have been used by other academics. The purpose is to improve existing methods and to reduce any existing gap in the literature in order to develop the methodology of this current study. The chapter concludes with a summary of the literature review for the purpose of developing the research hypotheses. Figure 2.1 presents a diagram of the issues covered in this chapter and how these issues are organized.

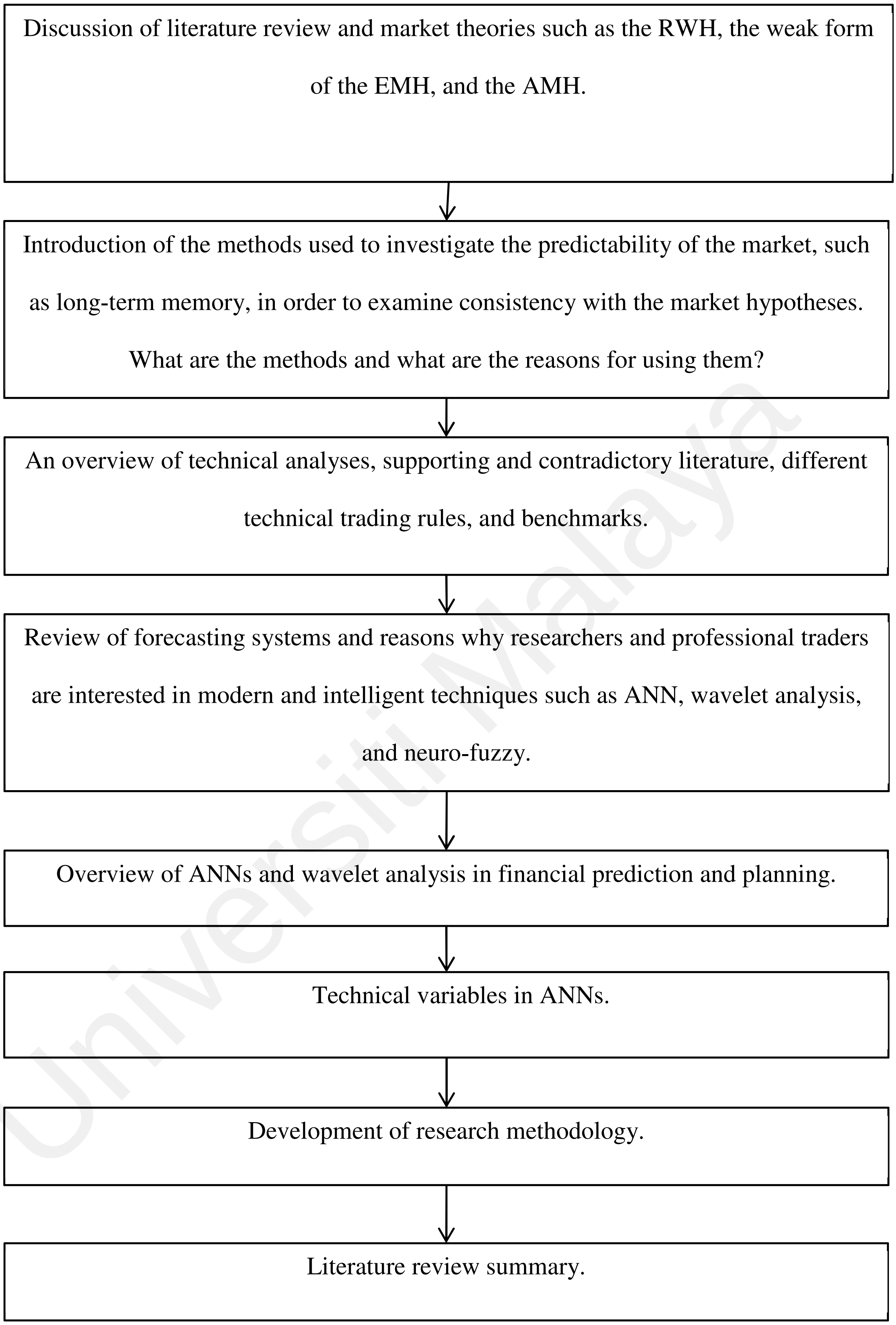


Figure 2.1: Structure of the literature review

2.2 Discussion of the Literature Review and Market Theories

The notion about stock returns being unpredictable is inherent and enduring within the academic study of finance. The idea that stock returns are volatile in nature was first anticipated by Regnault (1863), according to whom “l'écart des cours est en raison directe de la racine carrée des temps,” which means that price deviations are directly proportionate to the square root of the relevant time span. He also proposed a similar idea that price involves all significant information and further established that when “shares become publicly known in an open market, the value which they acquire may be regarded as the judgment of the best intelligence concerning them.” Malkiel and Fama (1970) combined these two concepts together and made the result the prime concept of the EMH.

For many years, the EMH was at the forefront of academic studies into financial markets. In this regard, Jensen (1978) said that he believed that there was no other hypothesis in the field of economics that had stronger, practical evidence to support it than the efficient market proposition. In the years since, various cracks started appearing in the façade of the EMH. Grossman and Stiglitz (1980) showed that informationally effective markets are unimaginable on the grounds that a certain level of inefficiency requirements exists with a specific end goal in order to influence arbitrageurs to be careful of costs and then reward them; thus, arbitrageurs should consider proficiency. Shiller, Fischer, and Friedman (1984) contended that conflating stock-price consistency with the parity of stock-price intrinsic value represents one of the most striking mistakes in the history of economic thought. Utilizing long-term chronicled information, Shiller (1980) calculated a progression of present values for US stocks that utilized discounted actual dividends from succeeding periods. He then contrasted these series and the market price series of each stock over a period. He affirmed that few changes occur to discounted (fundamental) value from year to year and that the market price displays unusual unpredictability when comparisons are undertaken. Poterba and Summers (1988) contended that financial markets are subject to fads, with prices deviating broadly from the intrinsic value of an asset because a specific asset may be more attractive compared with others for reasons that are separate from discounted future cash flow. Cheung et al. (2004) reviewed the conclusions of dealers in a foreign exchange market and discovered that most of the dealers trusted that crucial elements, which most macroeconomists worldwide accept are the principle drivers of price, have little significance for periods less than six months. Instead, dealers assume that technical trading factors, speculative trading, and bandwagon effects are the essential drivers of changes in price over a shorter time interval. Abreu and Brunnermeier (2003) said that a rational economic agent intentionally takes part in market rises because, with a burst bubble, basic value has little significance. Various studies (Lo & MacKinlay, 1988; Lo, 1989; Andersen et al., 2001) demonstrate that returns show noteworthy serial reliance and are thus predictable.

Elective hypotheses of market dynamics regarding the EMH have developed over the prior decade or so. These perspectives are efficient in the sense that the market tends toward them, as opposed to a state that consequently holds at all times. The foremost option for EMH is the heterogeneous market hypothesis of Dacorogna et al., 2001, which portrays how prices emerge from the connection of the participants in a market with various investment horizons. This model takes into account that mispricing emerges irregularly and that momentum in prices is clear because prices move toward what long-term investors see as the right level. Comparable properties are apparent in another EMH option, the AMH of Lo (2004), which consolidates behavioral finance ideas with elements of evolution and advancement. An essential conclusion of the approach is that, while it is not possible to have interminably profitable trading strategies, trading strategies are predominant and can be productive in certain market situations. However, this productivity only applies to whatever period during which the overarching market environment maintains stability.

In contrast to EMH, these two alternative speculations regarding market dynamics allow for the likelihood of serial reliance. According to the AMH paradigm, the opposing concepts of EMH and market inefficiency can simultaneously exist and be intellectually consistent (Ghazani & Araghi, 2014). There are many statistically based evaluations that can be performed to demonstrate this effect through the independence of price movements. These evaluations show that over a period of time, there are variations between the dependency and independence of returns, the reason for which could be that the AMH paradigm is actually true and that, clearly, the market is inefficient (Lo, 2004). Hence, the predictability of a market from historical data proves the adaptive nature of the market. In order to check the predictability of a market, various modern techniques and efficiency indices are used in the literature. Hence, this study proposes a novel hybrid model using artificial intelligent techniques to examine the predictability of markets. This study also investigates long-range dependence, so-called long-term memory, as a typical method (Cajueiro & Tabak, 2004a, 2004b, 2004c, 2005) to examine predictability and the adaptability of markets over time.

Thus, in the next sections, the literature regarding the AMH, long-term memory, technical trading, and modern techniques is reviewed.

2.3 The Adaptive Market Hypothesis

The adaptive market hypothesis (AMH) is an alternative hypothesis to the RWH and EMH, which applies the evolutionary approach of biology and a behavioral perspective to economic interactions and shows that the behavior of market agents, and thus the market, is adaptive (Lo, 2004). The AMH is still in its early stages and is commanding the attention of researchers. Ito and Sugiyama (2009) discovered time-changing market inefficiency in the US. Charles et al. 2012 maintained that the AMH is valid with regard to the foreign exchange rates of developing countries. Here, the authors discovered the consistency and predictability of returns in accordance with prevailing market conditions. Kim, Shamsuddin, and Lim (2011) examined whether the US stock market advanced or evolved over a specific period. They found that market situations are the driving elements of consistency and that the market was more productive after the 1980s than during prior periods. Investigating relative productivity, Noda (2012) suggested that the Tokyo Price Index (TOPIX) sustains the AMH while the Tokyo Stock Exchange (TSE) discards the AMH. Alvarez-Ramirez, Rodriguez, and Espinosa-Paredes (2012) provided evidence of the AMH and discovered that the US market was more efficient during 1973 to 2003. Urquhart and Hudson (2013) recorded compound results for the US, the UK, and Japan, and suggested that the AMH provides a persuasive portrayal of these markets.

Lim and Brooks (2011) stated as part of their research analysis that all the concepts that support the AMH are working toward their own interests and are individual in nature. The authors also stated that this adaptation, as well as innovation, is driven by competition. The market is shaped by the forces of natural selection, and the dynamics prevalent in the market are determined by evolution (Lo, 2005).

In other words, the AMH approach suggests that there is no single state of market efficiency; instead, market efficiency varies over time. The other conclusion that can be drawn regarding this approach is that market efficiency depends greatly on the context or the prevailing market conditions (Lo, 2004). Frequent attempts have been undertaken to confirm both of these statements in various stock markets (Kim et al., 2011; Lim et al. 2013).

Lim and Brooks (2011) have also stated that even though the concept of the AMH may appear to be highly abstract and qualitative, there are many practical scenarios where it can be observed. For example, the risk premium of equities varies over time depending upon the context of a particular stock market and the profile of the investors. Further, the existence of arbitrage opportunities in financial markets at different points of time is a sign of the existence of the AMH. In addition, Lim and Brooks (2011) argued that survival is the main objective for the evolution of financial instructions and markets.

In order to demonstrate the presence of the AMH concept, Lo (2005) calculated the rolling first order correlation for monthly returns of the S&P Index between January 1871 and April 2003. The author's analysis demonstrated that the efficiency of the markets has varied on a cyclic basis. For example, his findings showed that market efficiency was higher in the 1950s compared with the first half of the 1990s. In order to measure market efficiency, he used the first order autocorrelation coefficient. The author provided an explanation for the behavior of the markets through evolution (which causes market participants to change over time) and the dynamism in the behavior of participants. This explanation also shows that market behavior is not exactly supported by the EMH, in which the underlying assumption is that the markets are efficient and always in a state of equilibrium.

The AMH has garnered much recent attention from academic scholars. Indeed, a large number of studies have been conducted to identify the existence of this phenomenon in the stock markets. Lim and Brooks (2006) focused their attention on testing the AMH concept in emerging stock markets, and checked for any evolution in terms of overall efficiency. For this purpose, portmanteau bicorrelation statistics were employed, together with the rolling sampling method. This particular study also showed that the overall efficiency of the markets has cyclic variation.

Another study was conducted by Smith (2012) in which the focus was on observing the tenets of the AMH in the European stock markets. For this purpose, 15 stock markets were selected that were considered to be emerging in Europe. These were compared with the stock markets in the UK, Greece, and Portugal from February 2000 to December 2009.

In this study, the rolling window variance ration method was also employed. The study showed that the degree of efficiency in each market, when compared with another country, differs, some countries' markets being more efficient than others. However, even when considered individually, time-dependent variations of the returns do occur, a finding that clearly supports the existence of the AMH.

Todea et al. (2009) also studied the existence of the AMH, but in the context of Asian markets. In this study, the portmanteau test, together with the bicorrelation method subject to linear/non-linear correlations, were performed. The period over which the data were collected was from 1997 to 2008. The study showed that profitability strategies do not yield constant returns for a moving average value. Kim et al. (2011) found similar behavior in the UK stock markets. Their study used the automatic variance ratio (VR) test, automatic portmanteau test, and generalized spectral test on a century's worth of daily data from the Dow Jones Industrial Average (DJIA) Index. The results showed that the predictability of returns is directly associated with market conditions.

Lim et al. (2013) conducted a study to re-verify the foregoing phenomenon. For this purpose, the authors considered three major US-based stock exchange indices; namely, the New York Stock Exchange (NYSE) Composite Price Index, the S&P 500 Composite Price Index, and the DJIA Index. Samples were taken on a daily basis from 1969 to 2008. Even though all three indices belonged to the US markets, they all displayed a varying degree of predictability of returns, as indicated by the automatic portmanteau test and the automatic VR test. This particular study also highlighted that the predictability of returns decreased drastically over the last 20 years of the sample period. In other words, the efficiency of the markets in this period was found to have increased. Popovićet et al. (2013) conducted further research on the same topic to examine specifically how the observation point/period of time, the frequency at which the observations are made and analyzed, and the rolling time horizon impact overall market efficiency. The period of sampling contained market events such as euphoria and economic recession that could potentially impact the market returns. The authors used the Montenegro market for their analysis and employed the Monex20 Index. The study showed that according to the rolling window analysis on the first order serial autocorrelation coefficients and the p-value of the tests, all the three aforementioned factors are relevant for deciding market efficiency; moreover, market efficiency varies over time.

In another study, conducted by Urquhart and Hudson (2013), the degree of independence of the UK, US, and Japanese stock markets was investigated. Linear and nonlinear models are used in this study. The tests used included autocorrelation, runs, VR, McLeod Li, and Engle LM for the nonlinear model. These tests were performed on data collected over five decades, where the subsample interval was five years. The linear tests indicated that all three markets have phases of efficiency and inefficiency. The nonlinear tests showed that the three markets have inefficiency, which is time dependent. In a study on the Tehran Stock Exchange conducted by Ghazani and Araghi (2014), stock returns were shown to have a cyclical pattern. This study also used linear and nonlinear methods. Studies have been performed in the context of the Indian stock market as well. Hiremath and Kumari (2014) used linear as well as nonlinear models to examine the behavior of the Indian stock market between 1991 and 2013. The authors arrived at the same conclusion as the other studies mentioned above.

Manahov and Hudson (2014) attempted to approach the AMH problem from a different angle. In their study, the authors used a learning algorithm based on strongly typed genetic programming (STGP). They employed this algorithm to create artificial stock markets and traders, using the data from various actual indices such as the FTSE 100, S&P 500, and Russell 3000. The techniques differed from those described in the aforementioned studies. For example, econometric techniques such as the Kaplan test, Hurst Exponent, and the time-based dependence in series (BDS) Test were used. These metrics and techniques show concurrence with the AMH approach and confirm that the markets are indeed adaptive and that their efficiency is not a constant phenomenon. One of the factors on which market efficiency varies is the calendar anomaly. This anomaly comprises the day-of-the-week effect, the January effect, the turn-of-the-month effect, and the Halloween effect. Urquhart and McGroarty (2014) studied the effect of these anomalies on the DJIA Index from 1900 to 2013. The analysis undertaken through generalized autoregressive conditional heteroscedasticity GARCH(1,1) regression and the Kruskal–Wallis test showed that these calendar anomalies are variable and that their effect varies with time and the current market context.

Hull and McGroarty (2014) stated that it is only logical that the overall level of financial market efficiency is related to the economic development achieved by a country. The authors performed an extensive study in which 22 countries were examined and data were gathered over 16 years. Through the Hurst–Mandelbrot–Wallis rescaled range (used as a metric for efficiency), the returns generated, together with their fluctuations, were measured. While this particular study showed that the fluctuations have long memory persistence, it was unable to demonstrate conclusively that the rescaled ranges have a downward trend over time. In order to balance this, another metric for economic development was proposed; namely, the categorization of advanced and secondary development. The metric showed a higher efficiency level for markets that are advanced and emerging.

In order to check the adaptability and changing efficiency of a market over time, long-term memory analysis is applied in this current study. Long-term memory can be considered an efficiency index of a market (Kristoufek & Vosvrda, 2014). Hence, if a market's long-term memory index varies over time, which represents changing efficiency over time, this situation can be interpreted as consistent with the AMH (Hull & McGroarty, 2014). In the next section, a review of the literature of long-term memory in financial markets is presented.

2.4 Long-Term Memory

Long-term memory is an attribute of a financial time series, which reveals that the market does not immediately react to new data as proposed by the EMH; instead, the market reacts to such data slowly over a time frame (Mukherjee et al., 2011). Mandelbrot (1971) contended that serial reliance in a financial time series shows that recently arriving market data are not being completely arbitraged. Lo (1989) advised that numerous early thoughts of trade and business cycles are established around the definite assumptions that financial time series are long-term. As indicated by Cheung and Lai (1995), time series that show long-term memory are portrayed by serial reliance as well as aperiodic long cycles. Henry (2002) presumed that proof of persistence in equity returns implies that stock returns are measurable. Barkoulas et al. (2000) recommended that this situation could be exploited by speculators in order to gain profits.

Los (2006) demonstrated that standard and normally utilized (G)ARCH procedures are experimentally inadequate models because they do not show the right empirical long-term reliance. Barkoulas and Baum (1996) contended that long-term memory processes in financial time series hinder and conceivably refute conclusions, taking into account every customary procedure, including linear modeling, statistical testing methodologies in asset-pricing models, and forecasting, on the grounds that such results are based on the consideration of EMH-predictable Gaussian distributions.

Baillie (1996) analyzed financial time series as different because exchange rates, interest rates, and a wide assortment of products all display significant long-term memory behavior. Liu (2000) presumed that models that show long-term memory and overwhelming tails concentrate on the distribution of the returns of financial assets exceptionally well. Mukherjee et al. (2011) observed that Indian stock returns are autocorrelated. Lillo and Farmer (2004) showed that the indications of orders (buy/sell) on the London Stock Exchange comply with a long-term memory process, which they credited to the “herd mentality” of group-strengthening conduct amongst traders. Costa and Vasconcelos (2003) found that the Brazilian securities exchange shows a memory impact that lasts for up to six months. Barkoulas et al. (2000) and Panas (2001) discovered critical and vigorous confirmation of long-term memory processes in the Greek securities exchange. Tolvi (2003) discovered comparable confirmation for the Finnish securities exchange. Rege and Martín (2011) found that maintained long-term memory impacts the Portuguese securities exchange. Batten et al. (2005) found a relentless intraday structure for the levels of long-term memory in exchange rates, yet the overall direction of persistence stayed steady.

Henry (2002) proposed that long-term memory persistence in equity returns infers that stock returns are forecastable. Rasheed and Qian (2004) found that markets that have high measures of long-term memory perseverance create predominant forecasting measures. Barkoulas et al. (2000) noticed that experimental studies have paid almost no attention to the issue of serial reliance in developing markets. Various experimental papers on developing markets have appeared since the latter study. For example, Cajueiro and Tabak (2004b & 2004c) discovered confirmation of long-term memory perseverance in developing markets, which they noted as an inconsistency of EMH.

Although numerous writings have been published providing substantial proof of long-term memory processes in financial time series, numerous different papers challenge this phenomenon as spurious. Aydogan and Booth (1988) reconsidered the proof of earlier papers and presumed that recognition of the long-term memory pattern in American stock returns is deceptive. They said that the pattern emerged from the presence of "preasymptotic conduct" in statistical evaluations. Cheung and Lai (1995) discovered little proof for long-term memory processes in an assortment of worldwide stock returns. Chow et al. (1995) found no convincing confirmation about sustaining long-term memory in the equity returns they inspected. Barkoulas and Baum (1996) failed to locate any substantial support for long-term memory in the American securities exchange. Grau-Carles (2005) was unable to discover proof of long-term memory procedures from the log return series of either the S&P 500 or the DJIA, utilizing a wide assortment of statistical techniques and strategies. Zhuang et al. (2000) could reveal little confirmation of long-term memory processes in UK stock returns.

Sadique and Silvapulle (2001) showed that the level of improvement of a financial market seems to offer clarification for the clearly conflicting proof for and against long-term memory impacts, and that developed markets repudiate serial reliance. In comparison, smaller and less-developed markets generally discover proof for long-term memory. In their study, Sadique and Silvapulle (2001) discovered critical proof of the long-term memory process in economies that included Korea, Singapore, Malaysia, and New Zealand; however, the authors did not discover such proof for the larger propelled equity markets of the USA, the UK, and Japan. Cajueiro and Tabak (2004b & 2004c) detected a negative relationship between long-term memory and the capitalization of markets. They also found that proof of long-range reliance is more noteworthy in small rising economies than in expansive developed ones.

Hassler and Olivares (2008) presented an analysis of the German stock index, namely the DAX, in which more than 9000 daily absolute return figures were analyzed. The authors' results show that there is an increase in the estimated memory parameter from 0.33 to 0.45 and that the increase in persistence is significant.

Arouri et al. (2012) focused their research on identifying the long-term effects of variation in returns. Instead of stock markets, their research examined commodity markets and the prices of the four most commonly traded metal commodities on the COMEX market; namely, gold, silver, platinum, and palladium. In general, any random variable is assumed to demonstrate long-term memory behavior when its autocorrelation is unable to be integrated. The impact of any structural changes in this variable is generally manifested through sudden and strong changes in time series behavior. The result from the study shows that a strong long-term dependence exists in the daily returns of the metals under consideration. In addition, for each metal, this effect can be accurately seen with the help of an autoregressive fractionally integrated moving average-fractionally integrated generalized autoregressive conditional heteroscedasticity (ARFIMA-FIGARCH) model, which provides better out-of-sample forecasting accuracy than many other popular volatility models. The evidence in the study also indicates conclusively that conditional volatility of the traded metals is explained more successfully by long-term memory than by structural breaks.

Long-term memory analysis and structural breaks have also been studied extensively by other academics. For example, Kirkulak and Lkhamazhapov (2014) studied these properties with respect to the derivative markets in Turkey. The authors used the spot and future trades of gold from 2008 to 2013, when gold prices were at their peaks. The ARFIMA-FIGARCH model used in this study conclusively demonstrated that dual long-term memory exists in spot market transactions, while there is no long-term memory effect in futures' returns. This anti-persistence clearly indicates the overreaction of people to the all-time high prices of gold and is thus a clear indication of the market being inefficient or having weak efficiency. The authors also considered a structural break that is associated with the correction of gold prices in the global markets after the global economic crisis. This clearly demonstrated that long-term memory is a true phenomenon and not an after-effect of structural change.

In order to clarify further the distinction between true and spurious long-term memory effect, Ohanissian et al. (2008) proposed a statistical test. This test was based on invariance of the long-term memory factor for temporal aggregates of the process, under the null hypothesis of correct long-term memory. The estimate proposed by Geweke and Hudak (1983) for the long-term memory parameter that is received from different temporal aggregates of the time series on which it is based, is found to be asymptotically jointly normal. This finding implies that a test statistic should be constructed as the quadratic form of a demeaned vector of the estimates, which is quite simple to implement. Many simulations are performed on this statistic. These show that the statistic has good size and power properties, thus making it a useful alternative to true long-term memory. The asymptotic distribution of this analytical statistic is similarly valid for stochastic volatility with a Gaussian long-term memory model. In order to confirm its validity, this test is applied to foreign exchange rate data. Based on the information regarding all the different models that are presented in the aforementioned study, it can be stated that the long-term memory property in this case is actually created because of a true long-term memory process.

2.4.1 The Hurst Exponent

The Hurst exponent may be considered the best-known strategy that is used to quantify long-term memory in financial prices (Cajueiro & Tabak, 2004; Rasheed & Qian, 2004). This technique was initially created by Hurst (1951) to calculate the ideal size for a dam on the River Nile. The river's propensity to flood was not stable for some time. Hurst

(1951) discovered serial reliance in the time series that depicted vacillations in the River Nile's overflow. This technique was further refined by Mandelbrot and Wallis (1969b). Mandelbrot (1975) connected this measurement to security prices and observed that it is the better option for reliance. Mandelbrot (1975) called this new use of the measure the "rescaled range" or "R/S statistic." Lo (1989) highlighted several studies that showed the prevalence of rescaled range examination compared with more customary techniques for deciding serial reliance, including variance ratios, autocorrelation analysis, and spectral decompositions. Rescaled range investigation can distinguish serial reliance processes in a time series that is, to a great degree, non-ordinary. Further, the rescaled range method can incorporate time series with uniquely huge skewness and kurtosis coefficients, which are vital while considering financial time series. Elective techniques such as serial autocorrelation can dependably distinguish long-term memory in a time series in a manner that is close to Gaussian, which is unquestionably not the situation in the financial time series recognized in Lo's (1989) study. Mandelbrot (1971) additionally found that conventional rescaled range examination can distinguish nonperiodic cycles within a financial time series where the reliance proceeds for more than the specimen time frame being viewed. This finding is dissimilar to methodological approaches based on elective spectral analysis.

Lim (2007) proposed a nonlinear evaluation to rank market efficiency for financial markets. He examined how the ranking of stock markets according to the Hurst exponent progresses in time and concluded that market behavior can be rather disordered, which may mean changes in market efficiency over time. Kristoufek and Vosvrda (2014) measured capital market efficiency with three factors: long-term memory, fractal dimension, and approximate entropy. They utilized these factors as input variables for the efficiency index that they had previously presented (Kristoufek & Vosvrda, 2013). This approach made it possible to observe stock market efficiency after controlling for various types of inefficiency. By applying their model to 38 stock market indices around the world, the authors concluded that the most efficient markets are situated in European countries such as France, Germany, and the Netherlands, and that the least efficient markets are in Latin American countries such as Venezuela and Chile.

2.4.2 The Method of Mandelbrot

The Mandelbrot method is the way to connect the Hurst exponent measurement to financial time series and demonstrates that it is the better option for reliance (Mandelbrot, 1975). Mandelbrot (1975) called this method of the measure the “rescaled range” or “R/S statistic.” A significant criticism of the rescaled range method is that is vulnerable to confusing the effects of short-term memory and structural breaks with the effects of long-term memory. This issue was considered by Lo (1989), who proposed an alternative to the rescaled range, namely modified rescaled range statistics, in order to resolve the issue. Oh et al. (2006) introduced a similar accounting method for short-term memory known as detrended fluctuation analysis. Similarly, Cajueiro and Tabak (2004b) used an autoregressive (AR)-GARCH model in order to filter the effects of short-term memory. According to Cajueiro & Tabak (2004b), this filter severely enhanced the inconsistency in their time-varying rescaled ranges compared with the rescaled ranges that were calculated by making use of unfiltered data. The empirical evidence provided by Peters (1994) from financial markets mainly inclines toward the unique and original Mandelbrot (1971) model than toward the Lo (1989) model. Moreover, a number of researchers have used simulations in order to compare the models of Mandelbrot (1971) and Lo (1989). Willinger et al. (1999) and Kristoufek (2012) used the Monte Carlo simulation approach and demonstrated that the modified rescaled range, where long-term memory actually prevails, occasionally refutes long-term memory effects. Cont (2007) used a multi-agent simulation model and found that a similar false refutation issue prevails for long-term memory in terms of the unpredictability of prices. Mensi et al. (2014) studied the fluctuating and volatile levels of weak-form efficiency along with the presence of structural breaks for the crude oil benchmarks followed worldwide. The period was between January 1990 and September 2012. The authors utilized two different methods; namely, the Hurst exponent and the Shannon entropy methods. According to the findings, the Hurst exponent is a better indicator of performance compared with the Shannon entropy technique. The Hurst exponent was also found to be more effective than the Shannon entropy method in order to identify any financial or economic crises, or any strong structural breaks because of wars, attacks, etc.

In conclusion, these outcomes strongly recommend that the unfiltered or original Hurst–Mandelbrot rescaled range statistic is essentially the finest test for long-term memory. This is the reason behind choosing the statistic as the method for investigating the evidence of long-term memory in the data of markets. Hence, this current study uses the long-term memory test introduced by Rafal Weron (2002) as a known test of the true long-term memory process. This test is also coded in MATLAB programming language (Rafal Weron, 2011). Moreover, for additional consideration (Yalama & Celik, 2013), this current study investigates structural breaks in markets, analyzes long-term memory annually, and calculates annual Hurst exponents in different markets. Hence, this study measures a general Hurst exponent for each market and an annual Hurst exponent in order to compare structural breaks embedded in each year. In addition, this study uses Hurst exponents as an index for market predictability and compatibility with the AMH.

Other methods have been proposed to investigate predictability and rank stock markets in order to differentiate between emergent and developed markets (Coronel-Brizio et al., 2007; Giglio et al., 2008; Lan & Tan, 2007; Matia et al., 2004; Oh et al. 2007). This current study proposes to test the predictability and market efficiency of financial time series by using a recently introduced statistical instrument, approximate entropy. This estimation enables different degrees of correlation to be distinguished and Gaussian process to be differentiated from non-Gaussian (Zunino et al., 2010). Several studies (Kristoufek & Vosvrda, 2013, 2014; Zunino et al., 2010) have employed entropy to rank stock markets in order to illustrate that developing and emergent economies are actually less efficient than developed ones. This current study applies approximate entropy to investigate changes of market efficiency and predictability over time in order to check the concordance between futures markets and the AMH, and to act as a comparison with long-term memory results. The following section is a brief review of approximate entropy.

2.5 Approximate Entropy

Generally, entropy is disorder or uncertainty, and the definition of the entropy in financial domain refers to disorder or uncertainty in financial time series. The idea of entropy can significantly assist studies of stock market data because it obtains the disorder and uncertainty of a time series without inflicting any restrictions on theoretical probability distribution (Bentes et al., 2008; Darbellay & Wuertz, 2000). The variations are an entirely uncorrelated string of numbers when market trends are a random walk (Zunino et al., 2010). In this regard, Zunino et al. (2010) believe that such a string of data is totally disordered; thus, its entropy is maximized. Further, if the index variations are correlated, then the string of data is not completely disordered and the entropy does not reach its maximal amount. Consequently, negative entropy can be employed as a degree of predictability and thus of market inefficiency (Zhang, 1999). It should be stated here that Renyi, Shannon, and Tsallis entropy (Bentes et al., 2008; Dionisio, Menezes, & Mendes, 2006; Matesanz & Ortega, 2008), approximate entropy (Oh et al., 2007; S. Pincus & Kalman, 2004), transfer entropy (O. Kwon & Yang, 2008), and the Shannon entropy local approach (Risso, 2008) have been introduced to measure different features of financial time series. Approximate entropy is a measurement in financial time series, which demonstrates the predictability of the time series.

The complexity-entropy causality plane has been recently introduced as a powerful instrument for distinguishing the Gaussian from the non-Gaussian process and for distinguishing different degrees of correlation (Rosso et al., 2007). Zunino et al. (2010) proposed the use of the complexity-entropy causality plane to rank the stages of stock market development. Their empirical results showed that this statistical physics method is beneficial, achieving a more distinguished classification of stock market features. Indeed, the differences between developed and emergent stock markets can be clearly displayed with this statistical physics tool. The authors find that, for the studied stock markets, the temporal correlations are the principal elements of inefficiency. They point to the fact that the complexity-entropy causality plane is an approach-independent diagnostic tool, possessing more universal applicability than other generally applied alternatives such as the Hurst parameter. Thus, this current study concludes that the complexity-entropy causality plane can be considered a suitable physical model for predictability and market efficiency analysis.

Kristoufek and Vosvrda (2014) proposed three different efficiency measurements for capital markets, long-term memory with a Hurst exponent calculation, fractal dimension, and approximate entropy, as input variables for an efficiency index (Kristoufek & Vosvrda, 2013). Thus, stock market efficiency can be improved after controlling for different types of inefficiency. These measurements result in different efficiencies among capital markets: Some markets are more efficient, such as European markets, and some less efficient, such as Latin American markets. This result is consistent with the AMH. Comparing these three efficiency measurements, the authors found that the results of the approximate entropy ranking on 38 stock markets are the nearest to the proposed efficiency index.

In conclusion, this study uses approximate entropy as a successful statistical physics approach to investigate the predictability and market efficiency changes of futures markets over different periods. This approach helps to examine the adaptive nature of markets and can be used as a comparison or support tool for true long-term memory.

2.6 Technical Analysis Review

Being able to forecast or predict the behavior of the market has been the focal point of many studies conducted by academicians and industry experts. Even though a large amount of resources has been dedicated for this purpose, this activity still remains challenging and elusive. The main reason is the noisy and non-stationary nature of financial time series, particularly during any economic disturbances or distress. Thus, many different trading strategies are proposed that include the approach presented above and that help with forecasting. Some of these technical trading rules are discussed here.

Technical analysis is basically a financial market method that has the main objective of analyzing and then predicting actions taken by the markets (Bahrammirzaee, 2010). These actions are taken on the basis of price, volume, and futures' open interest. Such analysis is visual in nature because it relies on charts as one of the main tools. This analytical approach was first used by Charles Dow in the eighteenth century (Bahrammirzaee, 2010). Dow relied on price trends because he assumed that the market discounts all the available information. These trends in prices are further strengthened by trading volume.

2.6.1 Technical Analysis in Contrast to Fundamental Analysis

In order to understand the concept of technical analysis fully, it is necessary to distinguish it clearly from fundamental analysis. The main underlying assumption behind technical analysis is that the market discounts all the available information regarding a stock or commodity. Technical analysis also shows that prices have a tendency to follow trends and repeat themselves. However, fundamental analysis states that a market's action or a price change occurs purely because of the forces of demand and supply; thus, a greater focus is laid on what has caused the change in demand or supply. The focus of technical analysis is on the effect of such a change (Murphy, 1999).

Even though these two approaches may seem to conflict, there is no clear-cut separation that between the experts belonging to these fields. For example, those companies that require long-term assessments of their equity generally adopt a fundamental analysis approach. However, for short-term purposes, the technical analysis approach is usually found to be more effective. Technical analysis is more frequently utilized to determine the short-term share price of a stock; however, the long-term profitability of a company requires fundamental analysis. The biggest advantage of fundamental analysis is that it provides the ability to understand market dynamics; it does not just base the analysis on the panic that occurs when strong volatility exists in a market. Technical data, however, does not use any economic reports or financial data. It simply uses straightforward tools that are easy to understand in comparison with fundamental indicators. Moreover, these tools, because of their simplicity, can be adapted to any trading platform or medium, thus providing a trader with greater flexibility. To sum up, both techniques have pros and cons; moreover, in different situations, either of these techniques are more suitable (Atsalakis & Valavanis, 2009b).

2.6.2 Technical Analysis and Critiques

From the given description, it is clear that technical analysis contrasts with the notion of market efficiency. The prime reason for this conflict is that technical analysis conflicts with the recognized viewpoint of regarding profitability in an efficient capital market. The basis of technical analysis is the principal that investors can attain greater profits or returns than those achieved by holding a randomly chosen investment with comparable long-term risk. Hence, the market can be indeed beaten.

The claims of experts that the trading rules that are followed are correlated with profitability are quite justified. For instance, Brock et al. (1992) proved in their study that many of the trading methods that they used through the bootstrapping method are actually profitable. Their study was a pioneer in its field and was conducted with respect to the DJIA index. Bessembinder and Chan (1995) also found the same results in the context of the Asian stock markets. These studies, together with many others, for example Menkhoff and Schmidt (2005), Hsu and Kuan (2005), and Park and Irwin (2007), which were conducted in different contexts, showcased the effectiveness and utility of technical analysis. Indeed, based upon the literature, a distinction can be made between periods: the early period between the years 1960 and 1987 and a modern period from 1988 to 2004. The basis of this categorization is the analytical tools that were available for the period, the models, the tests, and some of the costs or disadvantages that the researchers had to bear. Some examples of these costs are the cost of transactions, risk factor analysis issues related to pattern recognition, optimization of parameters, out-of-sample verification processes, bootstrap and white reality checks, and NNs and Genetic Programming architectures. Park and Irwin (2007) also observed that most studies that were performed in the 1960s were published around three decades later, namely post 1990. There are two main justifications for this delay. First, adequate computational sources were available only during the latter period. Second, some of the main advantages of technical analysis came into public knowledge through many seminal papers, which until the 1990s were not in the public realm.

When all the above factors are taken into consideration, it becomes easy to understand the constant criticism that the technical analysis technique has received. In particular, academicians have strongly opposed this approach and stated that there are two reasons why academic researchers find faults with it. First, because the methods are totally false and second, the false methodology makes it easier for academicians to find faults with it (Malkiel, 2007). Some of the other main points of criticisms are as follows.

1) Technical analysis rejects the concept of the weak form of the EMH.
2) Many studies have proved that the technical analysis rules are not useful (Fama & Blume, 1966; Jensen & Benington, 1970).
3) Traders using this technique use the charts, which are very commonly known and employ the same signals. Thus, the actions define the trend of a market and thus the traders are working with a self-fulfilling prophecy (Malkiel, 2007).

2.7 Technical Trading Rules

There are diverse rules of technical trading implemented daily by market practitioners, technical analysts, and trading experts. This section is intended to provide an analysis of the "universe" of these rules and to categorize them into certain basic classifications.

2.7.1 Buy and Hold Criterion

The rule of "buy and hold" is a passive investment strategy and is considered to be the standard or benchmark of all the market-trading rules. Buy and hold is aligned with the weak form of the EMH. The prime principle of buy and hold is that investors purchase stocks and hold them for a longer term without becoming worried about short-term movements in prices, market volatility, and technical indicators. Although buy and hold is not sophisticated, the historical data reveals that it may be effective, mainly with long-term equities. Usually, buy-and-hold investors make use of the passive components, including index funds and dollar-cost averaging, and focus on establishing a portfolio rather than security research (Fernandez-Rodrıguez et al., 2000).

There are grounds for criticism, especially from technical analysts who, after the Great Recession of 1990, declared that the buy-and-hold rules were no longer valid. Corrado and Lee (1992), Jegadeesh and Titman (1993), Gencay (1998b), Levis and Liodakis (1999), Fernandez-Rodrıguez et al. (2000), O'Neill et al. (2001), Barber et al. (2006), and Szafarz (2012) have compared from a competitive perspective those trading strategies that are based on the buy-and-hold method as a benchmark. These comparisons have shown that other methods outperform the buy-and-hold strategy most of the time.

2.7.2 Mechanical Trading Rules

Charting is a process that depends upon how an expert interprets historically prevailing price patterns. This subjectivity allows experts' bias to affect the decisions that they take and the trading strategies that they use. Mechanical trading strategies attempt to limit the extent of these personal biases by introducing discipline in the process of decision-making, which is strictly based on trends. In other words, these rules attempt to bring greater objectivity to the decision-making process.

2.7.2.1 Moving Average Rules

Moving average (MA) rules (MAs) are common mechanical indicators that have been used for many years in decision-making systems. Essentially, an MA is the mean of a time series, which is recalculated on a daily basis. One of the most crucial parameters is the duration of the window or the number of days for which the average is calculated. These MAs can indicate trends over short or medium terms, depending upon a window's duration. For example, in the event of an MA window being short, ranging from one to five days, the trend indicated will be short term. However, with regard to predicting long-term trends, the window can be as long as 100 days. The main logic is that the buy signals (Ohanissian et al., 2008) are triggered only when the current closing price exceeds the calculated MA.

If it is assumed that the length of a window is *n* days, the current period is *t*, and the closing price is C_t, the MA can be further divided into three main categories as follows (Gencay, 1999).

Simple moving average (SMA):

$$SMA_{t+1} = \left(\frac{1}{n}\right)(C_t + C_{t-1} + \cdots C_{t-n+1}); \quad (1)$$

Weighted moving average (WMA):

$$WMA_{t+1} = [nC_t + (n-1)C_{t-1} \ldots + 2C_{t-n+2} + C_{t-n+1})]/[n(n+1)/2]; \quad (2)$$

Exponential moving average (EMA):

$$EMA_{t+1} = EMA_t + a(C_t - EMA_t). \quad (3)$$

The SMA is an average of values recalculated on a daily basis. The EMA adapts to the current market price with the help of a smoothing constant parameter. This parameter is an indicator of how responsive the EMA is to price changes. If this constant is low, the responsiveness is low and vice versa. The WMA assigns weightages to the price that are used as lags. The weightage for the current time period is higher so that the WMA reflects the current price conditions more accurately than the historical prices.

Normally, all these MAs use the closing price as the parameter for calculation; however, open prices, as well as high and lows, can also be used for the calculation. MAs are discussed at length in the literature. Brock et al. (1992) and Hudson, Dempsey, and Keasey (1996) examined the DJIA and the Financial Times Industrial Ordinary Index with MA analysis and concluded that if long series of data are taken into consideration, the MA technique has predictive ability. Further, the best approach is to consider a window of 50 days, through which it is possible to generate returns of 9.4%. These findings have been confirmed by artificial intelligence technologies such as NNs and fuzzy logic. For instance, Gencay (1999) studied the nonlinear predictability of foreign exchange and index returns through a combination of NNs and MA rules. The forecasts showed that the buy and sell indicators that are predicted by the MAs are able to forecast market timing. The indicators are also statistically significant in providing the predicted improvements for existing returns compared with the prior random walk model. LeBaron (1999) concurred with this finding and stated that a 150-day MA generates Sharpe ratios of 0.06–0.98 after considering the transaction costs in Deutsche Mark (DEM) and Japanese Yen (JPY) transactions from 1979 to 1992. Further, LeBaron (2000) evaluated profitability and found that in order to predict higher returns, more complex combinations of the MA should be considered. Gunasekarage and Power (2001) applied variable length MA together with fixed length MA and used them for forecasting the Asian stock markets. The first rule helps in identifying whether short-run MA is above (below) long-run MA, thereby suggesting that the general trend in prices is upward (downward). The second rule lays more emphasis on the crossover of long-run MA with short-run MA. These studies show that equity returns in these markets can be forecasted and that variable length MA is very successful.

In contrast, Fong and Yong (2005) attempted to confirm the prior findings by studying the fluctuations of Internet stocks with a recursive MA strategy applied to over 800 MAs.

However, the authors' findings did not provide any indication of significant trading profits. They stated that the reason for this result is the EMH. Chiarella et al. (2006) studied the impact of long-term MAs on market behavior. In the event of an unbalanced difference between fundamentalists and chartists, it can be inferred that the lag length or the window size of the MA rule has the potential to destabilize the market price. Zhu and Zhou (2009) studied the efficiency of MAs from the perspective of asset selection and found that the literature does not provide any guidance about optimal investments, even when MAs are forecasting clear trends. Consequently, academics often combine MAs with other rules that are fixed in nature so as to identify accurately when to time a market and when to move between assets and cash, especially when risky assets are considered. This method has been found to produce better results than the methods discussed so far. Milionis and Papanagiotou (2011) tested the significance of the predictive power of MAs on the NYSE, Athens Stock Exchange, and Vienna Stock Exchange. They stated that the performance of an MA relates to the length of the window considered. The authors also said that this technical strategy generally performs better than the buy-and-hold benchmark. This is particularly true when changes in an MA's performance occur around a mean level. Such a phenomenon is known as the rejection of the weak form of market efficiency. Lastly, Bajgrowicz and Scaillet (2012) studied the success that technical analysis has enjoyed by analyzing the DJIA Index from 1897 to 2011 and using the false discovery rate for data snooping. In their analysis, the authors found that the MAs are indeed profitable in this period; however, they also noticed that the economic value of the technical trading rules throughout this period are somewhat questionable.

Optimized moving average is the simple moving average with the parameters of successful trading results from the in-sample data. Consider the three years of in-sample data, and a SMA with various periods for calculating the average. Whatever is the number of days for moving average which generates the highest trading return in in-sample data, will be used in out–sample data for the next three months. This method of trading with SMA is so-called OMA in this thesis.

2.7.2.2 Momentum and Oscillator Rules

Another category of mechanical trading rules comprises the oscillator (OT) and momentum (MT) rules. OTs are methods that are not based on any analysis of trends. Instead, these methods try to identify when a trend is apparent for too long or is about to end (Gifford, 1995). As a result, they are called "non-trending market indicators." The main disadvantage of MTs is their inability to pick any rapid movements in the direction of prices, which may result in significant losses, because the MT technique presents the wrong signal (Edwards & Magee, 1997). This disadvantage is overcome by the OT indices. Their basic logic is that when prices are moving away from the average, a possible trend reversal is imminent (Edwards & Magee, 1997). In the case of a simple OT rule, the decision is based on the difference between two MAs (of varying lengths). A buy signal is generated (Ohanissian et al., 2008) when the prices are extremely high (or extremely low). Nevertheless, because simple OTs are in fact the differences between MAs, a change in the position indicator can also be generated when an index crosses the zero mark (Ohanissian et al., 2008). The boundaries between OTs and MTs can be fairly ambiguous at different times. The reason is that MTs can be applied to both MAs and OTs (Gifford, 1995). However, the major differentiating point here is that OTs are non-trend indicators, whereas MTs leverage a trend and its persistence in a market. An example of a simple MT rule is the difference between today's closing price and the closing price of *n* days ago. This persistence, which is also known as momentum, is the basis for generating a trading signal. In other words, the buy signal (Ohanissian et al., 2008) is provided when the current closing price is higher than the closing price *n* days ago. The value of *n* is generally decided based upon the initiation of the trade. Most commonly, it is set between five and 20 days.

Many varieties of OTs and MTs are currently being used. Some common examples are presented briefly below. They are also accompanied by some common research applications. The specifications and formulas for the following techniques can be seen in the works of Gifford (1995), Chang et al. (1996), and Edwards and Magee (1997).

1) Moving average convergence/divergence (MACD). MACD is calculated as the difference between short- and long-term EMAs. It is used to detect crossover points (indicating a change of trend) and whether or not a trend is diverging. This becomes the basis of generating buy and sell signals.
2) Accumulation/distribution (A/D). A/D is a commonly used momentum indicator that identifies whether investors are simply buying (accumulating) or selling (distributing) on the basis of the volume of trading, which in turn impacts price.
3) The Chaikin oscillator (CHO). The CHO is calculated as the MACD of A/D.
4) Relative strength index (RSI). The RSI is calculated in accordance with the average number of times a stock price moves up and down, and the magnitudes of such movements. The analysis is based on whether or not a particular stock is overbought or underbought. An overbought stock is identified when the value is greater than 70; in this case, a sell signal is generated. For an oversold stock, the value is less than 30.
5) Price oscillator (PO). PO identifies the momentum between two EMAs.
6) Detrended price oscillator (DPO). DPO removes any long-term trends in order to make it simpler to identify cycles. DPO also calculates the difference between a closing price and the SMA.

7) Bollinger bands (BB). BBs are based on the difference between closing prices and SMAs. They are also employed to identify whether stocks are overbought or oversold.
8) Stochastic oscillator (SO). The basic assumption for this technique is that as a stock price increases, the closing price also increases and reaches the high prices of prior periods.
9) Triple EMA (TRIX). TRIX is a momentum indicator. It uses three EMAs for the generation of trading signals.

These techniques are also commonly present in any literature containing technical analysis. The utility of these techniques has been known for some time. Brock et al. (1992) were pioneers in presenting the profitability of MACD, as well as for MAs and filter rules (FRs), as aforementioned. Kim and Han (2000) presented a hybrid genetic algorithm that consisted of an NN model based on OTs, for example PO, SO, A/D, and RSI, and simple momentum rules, all of which were used together to predict a stock market. Leung and Chong (2003) compared the profitability of MA envelopes and BBs. The results showed that BBs do not outperform the MA envelopes; however, the BBs are more efficient at depicting any sudden changes in price. Shen and Loh (2004) presented a trading system with rough sets that predicts trading signals for the S&P 500 index and that also outperforms the buy-and-hold rules. In order to create this hybrid trading system, the authors researched the rules that provide the most efficient results and that are also based on historical data; for example, MACD, RSI, and SO. Lento et al. (2007) provided evidence that BBs are unable to achieve high profit figures when compared with a buy-and-hold strategy. The authors' study was based on the S&P/TSX 300 composite Index and other popular indices such as the DJIA Index, the National Association of Securities Dealers Automated Quotations (NASDAQ) Composite Index, and the Canada/USD exchange rate. Chong and Ng (2008) studied the profitability of MACD and RSI by analyzing 60 years of data from the London Stock Exchange. The authors stated that the RSI and the MACD rules can generally generate returns in excess of the buy-and-hold benchmark.

Ye and Huang (2008) developed the research conducted by Frisch and Cassel (1933). Frisch and Cassel (1933) studied the damping of an OT with a non-classical OT. The non-classical OT caused quantum mechanics to be introduced in markets and is considered to be a method through which values can be accurately measured and prices can be predicted. Numerical simulations have shown that OT techniques are able to provide an explanation at a qualitative level of continuous changes in stock prices. Aggarwal and Krishna (2011) studied support vector machines (SVMs) and decision tree classifiers in an attempt to improve the direction and accuracy of prediction. In their study, the historical analysis of company stock takes into consideration daily OHLC prices and trading volumes. Usually, data for the last 5–10 years is considered for this model. Such a technique provided great forecasting results for the researchers and was even able to improve overall accuracy by more than 50%. Performance was also tested with several OTs and MTs (i.e. MACD, DPO, SO, A/D, and RSI). Dunis et al. (2011) and Sermpinis et al. (2013) presented a forecasting methodology with foreign exchange rates that applied adaptive NNs using radial-basis functions and particle swarm optimization. These were used with several NNs to predict exchange rates. In these applications, MACD was used but only as a benchmark or standard and not for generating actual trading or buy/sell signals.

2.7.2.3 Other Trading Rules

The factors and rules (FRs) discussed here are the most reliable indicators for technical analysis; however, their use is rather limited because of their restricted applicability to certain scenarios (Bahrammirzaee, 2010). FRs create a few signals for a short time when a market price increases or decreases; the market price is then multiplied by the percentage difference of the prior highest price (Chong & Ng, 2008). Moreover, the professionals who work as technical analysts try to make new rules so that more universal rules can be formed and the existing ones can be changed. The changes introduced so far are very common among the literature and are known by various names (Chong & Ng, 2008). These names are not related to the existing theories and mechanical rules that have been discussed in the earlier sections.

For instance, consider the approach of people who are against market trend, or, in simplified terms, the rules of such people; namely, contrarian rules. The logic of such people is very simple and can be specified easily. The rules state that for every rule of trading, there is a signal of sale, and for such a sale, there is a corresponding signal to buy. The analysts who use this approach usually believe that the market price changes only for a limited time and sooner or later it will return to its usual state (Gencay, 1999).

Trading range break rules are also tested in the literature of technical rules. These rules can be understood as MT indicators because they show that whenever a stock price increases or falls during several continuous days, it creates favorable or non-favorable momentum.

Another fascinating category of trading rules is known as the channel breakout and volatility breakout rules. These were introduced by a trading expert, Richard Donchian, who maintained a simple ideology to the effect that a market price channel is an integrated factor of a trading strategy (Kestner, 2003).

The final category is pattern rules. Head and shoulders, double tops and bottoms, triangles and rectangles are a few common examples of these rules. They help analysts to understand the patterns in pricing charts and can be regarded as extended classes of MAs or MTs (Dunis et al., 2011).

Many researchers and professional traders believe that traditional statistical methods have reached their limits (Atsalakis & Valavanis, 2009b; Bahrammirzaee, 2010) and that machine-learning systems are much better substitutes. These systems include ANNs, SVMs, and neuro-fuzzy systems. ANN, as a capable prediction tool (Zadeh, 1994), outperforms traditional statistical methods and various other intelligent models (Fernandez-Rodrıguez et al., 2000; Refenes et al., 1994; Yoon et al., 1993). In the next section, we present an introduction to ANNs and their application in financial time series.

2.8 Artificial Neural Networks

ANNs are devices of computational modeling that have achieved broad acceptance in many arrangements in order to model complicated and composite real-world issues (Basheer & Hajmeer, 2000). ANNs take their inspiration from human nervous systems and the structure of the brain, and have been the central sources of inspiration for many unique and novel techniques, including an infinite area of applications (Anderson et al., 1995). Generally, ANNs can be considered information processing systems that apply generalization and self-learning abilities, which are also extremely adaptive. As a result, ANNs produce significant solutions for subjective information processing (Touzet, 1997), predictions (Faller & Schreck, 1995), making decisions (Kwon et al., 1997), and associated issues. Thus, ANNs have become central to a growing variety of real-world and industrial applications.

ANNs provide sufficient learning capacity to make them more likely to capture the complex nonlinear associations that are embedded in financial markets. The advantages of ANNs are explained with clarity in the trading literature, and reviews are provided in Gooijer and Hyndman (2006). However, skeptics believe that NNs are inefficient in terms of the parameter-tuning process; as a result, they make performance unexceptional. Because of this reason, analysts who use NN algorithms have attempted to overcome such limitations (Ling et al., 2003). Moreover, see amongst others, Harrald and Kamstra (1997) and Teräsvirta et al. (2005).

Commonly used NN architecture is multilayer perceptron (MLP), which delivers suitable outcomes in time series financial forecasting (Makridakis et al., 1982). However, in certain instances, the evidence shows contradictory results. For example, Tsaih et al. (1998) tried to predict the S&P 500 futures market. In their analysis, reasoning neural networks (RNNs) outperform MLPs. Lam (2004) reviewed the systems that are used to forecast financial performance and found that the systems could not improve on the benchmarks. Ince and Trafalis (2006) used MLP with SVM to forecast multiple currencies. The result was not impressive because it showed that the MLP/SVM method is inaccurate. When outcomes are compared, it is clear that algorithms are more reliable than MPLs (Alfarot et al., 2008). This finding was tested and proven with forecasts of European companies' bankruptcies. Tenti (1996) and Dunis and Huang (2002) achieved some positive outcomes though RNNs to forecast exchange rates. However, the performance of pi-sigma networks (PSNs) is better. PSNs were formed by Shin and Ghosh (1991) to capture high-order correlations. Ghosh and Shin(1992) and Shin and Ghosh (1991) also showed that the PSN method is superior to others by demonstrating that function superiority is better than MLPs and higher-order neural networks (HONNs). Besides functional superiority, the PSN method was also tested for forecasting and trading shares. In this regard, it gave impressive results with accurate forecasts (Ghazali, 2007). The PSN method was also used by Hussain et al. (2006). The authors tested the method in many exchanges as inputs for their networks. However, in the context of simple trading, this tactic could not outperform others (Dunis et al., 2011).

This current study focuses on a specific type of RNN that has the benefits of MLPs and back-propagation neural networks (BPNNs) and is compatible with various training algorithms such as Bayesian and Levenberg–Marquardt algorithms, which are NARX-NNs. A NARX-NN is a type of recurrent dynamic NN with feedback links that connect some layers of a network (MathWorks, 2014b; Siegelmann et al. 1997). The NARX model is built on the linear autoregressive exogenous (ARX) method, which is generally applied in time series modeling. As per the findings of Siegelmann et al. (1997), it is certain that NARX networks are equally strong as fully recurrent networks when addressing a limited number of parameters. This means that NARX networks are as good as Turing machines. Although there is insufficient feedback on NARX models, the models can be used in preference to other conventional recurrent networks without any calculative loss. Busse et al. (2012) used a more variant forecasting approach, adopting NARX as an NN base, so as to calculate the price, one day in advance, for natural gas in the NetConnect market of Germany. The outcomes clearly indicate that the NARX method provides the most accurate forecast compared with others on the basis of five factors. Nemes & Butoi (2013) considered NNs that are specifically designed to forecast the stock exchange rates of Romanian stocks listed on the Bucharest Stock Exchange (BSE). In order to manage price flexibility in the short term, the authors used a multistep strategy in advance. The results of such a strategy are better when it is used with NARX compared with the MLP results. Abdulkadir et al. (2014) introduced a novel hybrid model, namely the unscented Kalman filter (UKF)-NARX, which was based on unscented Kalman networks including nonlinear and Bayesian regulation algorithms. When this model was compared with other models that are commonly used, the results were quite impressive because the proposed model outperforms the existing models. In addition, the forecasts are reliable in the long term (Abdulkadir & Yong, 2014) because the UKF-NARX model uses a parallel nonlinear with exogenous input model, commonly known as P-NARX, which is trained together with a Bayesian regulation algorithm. The final results showed that this network is better than the Levenberg–Marquardt network. Diaconescu (2008) analyzed another method with architecture for an RNN together with an embedded-memory NARX that shows results that promise much, including multiple qualities of dynamic system applications. The performance of the NARX has been under observation for many programs and has now been tested for multiple time series so as to be used as an NN. When analyzed, the training algorithms and the dimensions of the embedded memory had more influence compared with the number of neurons. Such research and analysis has given helped to implement traditional methods of statistics in order to obtain better methods for improving the efficiency of forecasts predicted with NARX. In addition, Shahbazi et al. (2016) developed a new strategy for forecasting movements and trends in a foreign exchange market. This strategy was based on the NARX-NN with changes in the bagging technique of time shifting and with different financial indicators; for instance, RSI and a stock indicator. The ability to learn through an NN is very noticeable; however, NNs often exhibit negative and unpredictable performance for noisy data. When results are compared with a static NN, the rate of error reduces and the overall quality of predictions improves. In a foreign exchange market, the method is implemented so as to determine the hourly rate. Moreover, experiments can be undertaken with predictions of two different rates of exchange.

In conclusion, this current study uses a NARX-NN in order to predict the movement of futures markets because NARX-NNs have promising outcomes in the reviewed literature and benefit from the features of RNNs, MLPs, and BPNNs and can be optimized with various training algorithms such as Bayesian and Levenberg–Marquard algorithms. In chapter 4, the architecture of a NARX-NN is introduced. Generally, ANNs can be applied in three forms: singular, expert system (ES), or hybrid intelligent. So far, this study has covered the singular form of an ANN. The following sections provide a brief review of ES applications and hybrid intelligent systems. The reasons for applying ANNs and any intelligent technique in a hybrid form, and the benefits, are also considered.

2.8.1 Hybrid Intelligent Systems

Although much still needs to be discovered regarding how the animal brain learns, trains, and self-forms itself in favor of mining and processing numerous and complex data, many modern advances in neurobiology allocate time to considering some significant mechanisms of this sophisticated system. The brain's "modular" construction and its hybridization abilities, such as combining different roles in order to execute a complex order, are certainly worth investigating. A hybrid intelligent system (HIS) is a productive and strong learning system that assembles corresponding features and overcomes the weak points of the representation and processing abilities of non-symbolic and symbolic learning paradigms (Taha, 1998). Further, a HIS is a technique that combines intelligent systems in order to solve problems (Lertpalangsunti, 1999). The technique also integrates intelligent systems with traditional computer systems and databases. Goonatilake and Khebbal (1995) and Lertpalangsunti (1999) introduce three general reasons for executing hybrid models as follows: improvements in technique, the diversity of application duties, and the recognition of multi-functionality. The interaction level between the two modules in a hybrid system can be classified at four levels: loosely coupled (separate models), transformational models, tightly or strongly coupled models, and fully coupled models. A loosely coupled hybrid intelligent model includes two detached components, such as an ANN and an ES, with no contact between them (Medsker, 2012). Another type of stand-alone or loosely coupled model is a transformational model. The model's characteristics are similar to those of a loosely coupled model although its internal information structure is totally different. Indeed, the operational nature of a transformational model is sequential. The process usually starts with one model and ends with another (Medsker, 2012). Handelman et al. (1990) merged knowledge-based techniques and ANNs based on transformational models for robotics. The two elements of a transformational model use only part of their internal information structure to interface rather than using external data, as in a tightly coupled model (Medsker, 2012). Fully coupled models have a hybrid structure of dual character in which the exclusive features of both models can still be seen (Metaxiotis & Psarras, 2003).

So far, different types of HIS have been described. Moreover, the superiority of a HIS over a single intelligent technique such as an ANN has been claimed. The next section reviews literature that compares different hybrid ANNs, the singular form of ANNs, and traditional statistics, in the context of predicting financial time series.

2.8.2 ANNs in Financial Prediction and Planning

In financial markets, Bahrammirzaee (2010) discusses that the three most significant artificial intelligence applicability fields are "credit evaluation" (including credit risk analysis, credit ranking and scoring, and bond rating), "portfolio management" (including optimal portfolio selection, asset portfolio selection, and equity selection), and "financial forecasting and planning" (including financial forecasting, bankruptcy prediction, and stock and foreign exchange rate forecasting). The focus in this study is financial forecasting and planning with the use of ANNs in a hybrid form. This section presents literature related to the reasons why a hybrid form of ANN is used in the current study, compared with a singular form of ANN or traditional methods.

Financial markets are compound nonlinear entities with complications and interrelations that are difficult for humans to understand. This is the general reason why ANNs have been widely employed in this field. ANNs have been designed to forecast foreign exchange markets, the liquidity of banks, inflation, and various other financial requirements. Researchers (see Table 2.1, Table 2.2 and Table 2.3) have undertaken many comparative studies in this field. The results show that ANNs and artificial intelligent models outperform traditional methods such as linear regression and the random walk model. For instance, Thawornwong and Enke (2004) compared conventional linear regression, the buy-and-hold strategy, and the random walk model with an ANN model that employed constant related variables in a financial and economic variables' selection. The results indicated that reconstructed NN techniques that employ recently related variables work better. Panda and Narasimhan (2007) applied an NN to forecast the weekly Indian Rupee/USD exchange rate. Gradojevic and Gençay (2013) believed that from the market microstructure perspective, technical analysis can be beneficial when uninformed traders have predictable effects on price or when informed traders make systematic mistakes. Further, chartists confront a substantial degree of trading uncertainty because technical indicators such as MAs are fundamentally deficient filters alongside a nonzero phase shift. Hence, conventional technical trading probably results in imprecise trading recommendations and considerable losses. In this regard, Gradojevic and Gençay (2013) offered the optimization of technical indicators through fuzzy logic in order to decrease the uncertainty level.

The most important point about successful financial prediction is achieving the best output with a minimum quantity of input data. In contrast with other areas, the application of hybrid artificial intelligent techniques in financial forecasting is more useful because hybrid techniques enable researchers to combine the abilities of different models. In this field, Kuo et al. (1996) created a hybrid model for stock market prediction that studies qualitative and quantitative factors at the same time. This model consists of a combination of an NN for the quantitative components and a fuzzy Delphi system for the qualitative components. The authors employed their hybrid model in the Taiwan stock market and achieved satisfactory results.

Table 2.1 displays the results of HISs compared with conventional systems for financial planning and forecasting in accordance with the following literature (Bildirici et al., 2010; Bildirici & Ersin, 2009; Charbonneau & Kharma, 2009; Chavarnakul & Enke, 2008; Chen & Leung, 2004; Chen et al., 2009; Garliauskas, 1999; Hassan et al., 2007; Huang et al., 2009; Luther, 1998; Ni & Yin, 2009; Shah & Murtaza, 2000; Taffese, 2007; Tsakonas et al., 2006; Wang, 2009).

Table 2.1: Hybrid intelligent systems compared with conventional systems for financial planning and forecasting

Author/s	Domain	Hybrid method/s used	Method/s compared with	Result
(Garliauskas, 1999)	Financial forecasting	ANN, kernel function approach and the recursive prediction error	Classical statistical methods	Hybrid model performs better
(Chen & Leung, 2004)	Foreign exchange forecasting and trading correction	Time series model and GRNN	Multivariate transfer function. GMM. and Bayesian vector autoregression without ANN error correction	Hybrid model performs better
(Hassan et al., 2007)	Stock market forecasting	HMM, ANN and GA	Conventional HMM. ARIMA	Hybrid model performs better
(Chavarnakul & Enke, 2008)	Stock trading	VAMA and EMV indicator with GRNN	VAMA and EMV indicator	Hybrid system performs better
(Ni & Yin, 2009)	Exchange rate prediction	Temporal SOM and SVR and GA	GARCH model	Hybrid models performs better
(Charbonneau & Kharma, 2009)	Stock price forecasting	Rule-based trading agents and GA	ARIMA and LR	Hybrid models performs better
(Bildirici & Ersin, 2009)	Stock exchange forecasting	GRACH model with ANN	GRACH model	Hybrid models performs better
(Huang et al., 2009)	Financial market trading system	HiCEFS: ISMF and HCGA	B&H strategy, trading system without and also without prediction and also with other predictive models (EFuNN. DENFIS and RSPOP)	Hybrid model performs better
(Wang, 2009)	Forecasting model for stock index option price	GJR-GARCH and ANN	GARCH and conventional GJR-GARCH	Hybrid model performs better
(Bildirici et al., 2010)	Exchange rates and stock returns	TAR-VEC-MLP. TAR-VEC-RBF and TAR-VEC-RHE models	TAR-VEC model	TAR-VEC-RBF performs the best
(Jin & Kim, 2015)	Forecasting natural gas prices	Wavelet and ANN	ARIMA and GARCH	Hybrid model performs better

As can be seen from the brief review in Table 2.1, all hybrid models outperform traditional models that have been compared in the same study and with the same input data. Table 2.2 illustrates the comparison between the results of a HIS and one or more single intelligent models in the context of financial planning and forecasting.

Table 2.2: A comparison between a HIS and one or more single intelligent models in the context of financial planning and forecasting

Author/s	Domain	Hybrid method/s used	Method/s compared with	Result
(Kim & Shin, 2007)	Temporal patterns in stock markets detection	1. ATNNs with GA 2. TNNs with GA	Conventional ATNN. TDNN and RNN	Hybrid model performs better
(Hua, Wang, Xu, Zhang, & Liang, 2007)	Corporate financial distress prediction	SVM and LR	Conventional SVM	Hybrid model performs better
(Li & Kuo, 2008)	Financial investment decision support	Integrating K-chart technical analysis, discrete wavelet transform and a novel two-level SOM network	Conventional SOM	Hybrid model performs better
(Zhang & Wu, 2009)	Stock market prediction	Improved bacterial chemo taxis optimization and BPNN	BPNN	Hybrid model performs better
(Hsu, Hsieh, Chih, & Hsu, 2009)	Stock price forecasting	SOM and then SVR	Conventional SVR model	Hybrid model performs better
(Tsai & Chiou, 2009)	Earnings management prediction	ANN and decision trees model	Conventional ANN	Hybrid model performs better
(Lee, 2009)	Stock trend prediction	SVM with F-score and F_SSFS	BPNN along with three commonly used feature selection methods	Hybrid model performs better
(Chen, Lin, Yu, & Chen, 2009)	Government bond yields forecasting	ANFIS	RPROP. RBF NN and BPNN	RBF NN performs better
(Chengzhao, Heiping, & Ke, 2015)	Forecasting three Asian stock indices	Empirical Mode Decomposition-NN	BPNN	Hybrid model performs better

As the results of the comparisons in Table 2.2 indicate, nearly all hybrid systems perform better than single intelligent methods. In this regard, Table 2.3 presents a comparison of hybrid models and single intelligent methods in the same studies and with the same data. The results of HIS analyses and the comparisons with conventional and single intelligent methods show that HISs outperform other methods because they employ the capabilities of various single intelligent models in the form of united systems, although there are a few exceptions.

Table 2.3: A comparison of hybrid models and single intelligent methods in the same studies and with the same data

Author/s	Domain	Hybrid method/s used	Method/s compared with	Result
(Kim, 2006)	Financial forecasting	ANN and GA with instance selection algorithm	ANN and GA without instance selection algorithm	Proposed hybrid model performs better
(Nikolaos, 2008)	Financial evaluation of corporation	Regular RBF with 3 layers, GA in all the layers	Neuro-genetic forms of RBF	Proposed hybrid system performs better
(Kamo & Dagli, 2009)	Financial forecasting	Candlestick method based on GRNN with rule-based fuzzy gating network	Candlestick method based on GRNN with simple gating network	Proposed hybrid model performs better
(Yu & Huarng, 2008)	Stock index forecasting	Bivariate NNs with fuzzy time series forecasting substitutes	Bivariate and univariate conventional RL, ANN and ANN-based fuzzy time series and univariate ANN with fuzzy time series forecasting substitutes	Proposed bivariate hybrid model performs better
(Fatima & Hussain, 2008)	Forecasting KSEI00 index	1. ANN- ARIMA 2. ANN-ARCH/GARCH	Conventional ARIMA, ARCH/ GARCH and ANN	Second hybrid model performs better
(Wen, Yang, & Song, 2009)	Stock price prediction	Hybrid BPNN, SVM and ANFIS	Separate BPNN, SVM and ANFIS	Hybrid model performs better
(Teoh, Chen, Cheng, & Chu, 2009)	Stock market prediction	Rough set theory and multiorder fuzzy time series	Weighted fuzzy time series models, high-order fuzzy time series, the partial autocorrelation function and autoregressive models	Hybrid model performs better

In conclusion, ANNs are developed in hybrid forms with other artificial intelligent techniques. In this regard, the ANNs have the best-performing and strongest features of each approach, while omitting the defects and weak points. Hybrid systems can provide a better representation of machine-trading systems and have the ability to process and learn within non-symbolic and symbolic paradigms (Taha & Ghosh, 1997). Hence, this study uses a novel hybrid system that includes ANNs in order to analyze futures markets' trends and obtain a better forecasting model for futures markets. The next section presents the factors that are combined with ANNs in order to forecast futures markets.

2.9 Wavelet Transforms and Artificial Neural Networks

According to the literature, researchers apply ANNs because of their qualities such as efficiency, performance, reproducibility, consistency, completeness, breadth, consistency of decision-making, and, most importantly, timeliness (Bahrammirzaee, 2010). Chang et al. (2004) believed that ANNs are limited because of the complexities of financial markets and the noise that they display. This current study adds a well-known denoising process, wavelet transform analysis, to an ANN in order to overcome this limitation regarding the financial markets' noise.

Because of the weakness of ANNs to noises that are latent in futures' prices, this current study applies wavelet analysis in combination with an ANN. Wavelet decomposition, so-called wavelet transform, is a technical tool for analyzing signals and is applied for its superiority in such analysis in two main domains: frequency and time (Ramsey & Zhang, 1997). The denoising and decomposing procedures of wavelet transform have become a widespread technique for single-dimensional signal filtering and data mining (Aminghafari et al., 2006; Donoho, 1995; Hsieh et al., 2011). Several studies have shown that wavelet transform denoising procedures increase the performance of time series predictions (Pindoriya et al., 2008; Shafie-Khah et al., 2011). Hsieh et al. (2011) proposed a hybrid technique whereby wavelet denoising and a recurrent NN are combined to forecast the DJIA, FTSE, NIKKEI, and TAIEX indices with astonishingly profitable trading results. Lotric and Dobnikar (2005) and Lotrič (2004) combined the wavelet denoising method with an ANN to improve denoising factors dynamically. The authors reported a performance increase in forecasting accuracy. Moazzami et al. (2013) applied a wavelet transform together with an ANN to predict the day-ahead peak load of Iran's national grid. The study presented promising outcomes. Jin and Kim (2015) suggested some hybrid models of wavelet approximation, ARIMA, GARCH, and ANN to forecast natural gas prices. The outcomes demonstrated not only that the performance of the wavelet hybrid is superior in all models, but also that the wavelet-ANN outperforms other hybrid systems. There has also been other research on wavelets and NNs where wavelets are applied in hidden layers and as neuron transfer functions, called "wavelet networks" (Fernandez et al., 2000; Wang & Fu, 2006). Instead of applying wavelet coefficients to training an NN, which is what "wavelet networks" do, this current study focuses on using wavelet transform for denoising signals and then feeds the signals into the NN.

Zhang and Dong (2001) combined the inherent abilities of wavelets and ANNs to capture non-stationary and nonlinear attributes of the electricity market. The authors proposed a forecasting model based on wavelet multi-scale decomposition and a multilayer perceptron (MLP) NN modeling of wavelet coefficients. In order to lessen the effect of noisy low-level coefficients, they applied the practical Bayesian model and automatic relevance determination methodology to choose the size of the MLPs. The authors reported successful results from the hybrid of a wavelet and NN. Conejo et al. (2005) predicted 24 market-closing prices of a day-ahead electric energy market using wavelet transform and an NN. The authors compared the results to ARIMA, transfer function, and dynamic regression, and claimed that the combination of an NN and wavelet outperforms other methods. Wang et al. (2011) investigated the forecasting accuracy of a novel WNN model for the prediction of stock indices. They offered a novel hybrid of wavelet denoising and ANN forecasting, and compared an ANN with wavelet denoising and the ANN. The authors stated that a combination of a wavelet and NN performs better than a simple ANN; thus, the proposed denoising procedure with wavelet decomposition achieved promising results. Pang et al. (2011) estimated the crude oil price through the use of a WNN and input variables such as the industrial petroleum inventory level. The authors used the WNN to forecast the crude oil price and claimed that their results are promising and that the WNN method is successful. Jammazi and Aloui (2012) proposed a wavelet decomposition and NN model to achieve prominent forecasting of the crude oil price because of the intrinsic complexity of the oil market's structure. The authors used a Haar and Trous wavelet for decomposition and denoising, then a BPNN for the prediction of the crude oil price. They reported promising results with their denoising and forecasting procedure. Behradmehr and Ahrari (2015) proposed a model to smooth and minimize the noise presented in crude oil prices. They examined the impact of wavelet smoothing on oil price prediction while using the group method of data handling (GMDH) NN as the prediction technique. The authors indicated a 40% increase in accuracy by using a wavelet as a denoising process. Babić (2015) examined and discussed the modeling and forecasting results of exchange rates with a wavelet and NN model. They used Daubiches and Haar for the wavelet and also coded a forecasting program from their study in MATLAB. The authors reported promising results with the denoising process of the wavelet compared with the singular NN.

The current study is the first attempt in the literature to develop a hybrid system of multivariate denoising using a wavelet and PCA (Aminghafari et al., 2006) in order to feed to a NARX-NN. Since wavelet denoising is a univariate algorithm, Aminghafari et al. (2006) suggested a multivariate denoising method, applying wavelets and PCA. PCA is a well-known data analysis technique and is especially considered to simplify multi-scale signals by tracing new elements and revealing the main features of data (Bakshi, 1999). WPCA analyzes multivariate signals with multiple univariate wavelets and then performs a PCA in order to choose a convenient number of useful principal components. Aminghafari et al. (2006) suggested that denoising multivariate signals by WPCA is a method that performs better than univariate wavelet denoising on each factor separately. Further, WPCA extracts the same noise at different frequencies from elements of a multivariate signal.

According to the literature, it is clear that wavelet decompositions (wavelet transforms) can control noise in a signal while an ANN can learn various movements in a nonlinear time series. Generally, hybrid models are established in three stages: preprocessing, modeling, and evaluation. This study proposes a novel forecasting approach that integrates WPCA and nonlinear autoregressive ANN with exogenous input (NARX-NN) in the preprocessing stage to develop an ensemble prediction model, a WPCA-NN, and a trading strategy.

2.10 Technical Variables in Artificial Neural Networks

Hitherto, this study has presented the literature regarding the hybrid system used in this study. Now, this section reviews the variables that are commonly used. Atsalakis and Valavanis (2009b) examined more than 100 related studies that considered neural and neuro-fuzzy methods that were derived and then applied to predict stock markets. They classified the studies in terms of input information, prediction methodology, performance assessment, and the measures employed. Consequently, they showed that soft computing algorithms are largely accepted in order to study and evaluate stock market activities.

Thawornwong and Enke (2004) examined whether the application of relevant financial and economic variables leads to improvements in stock return prediction. They applied NNs, including feedforward and probabilistic NNs, in order to predict the directions of stock returns. The results showed that redeveloped NN methods that apply recent applicable variables produce higher profits with lower risks than traditional linear regression, the buy-and-hold strategy, the random walk method, and the ANN methods that employ constant related variables.

Wu et al. (2001) proposed a hybridized neural network and fuzzy logic model called the feedforward neuro-fuzzy (FFNF) system, in order to undertake financial prediction. The authors used a consumer price index, total industrial production index, leading economic index, bank prime loan rate, federal funds rate, unemployment rate, S&P 500 lag(1), S&P 500 closing prices output, and prior S&P 500 data as input data.

Olson and Mossman (2003) compared logistic regression (logit) and ordinary least squares (OLS) techniques to forecast one-year-ahead Canadian stock returns, with the predictions achieved by applying an ANN. The input data or independent variables were 61 financial and accounting ratios for more than 2300 Canadian firms from 1976 to 1993. The study's outcomes indicated that BPNNs, taking into account nonlinear associations between input and target data, perform better than the best regression alternatives in terms of classifying firms that are expected to have either low or high returns and point estimation. The advantage of the ANN model is that it results in higher profitability by applying various trading rules. However, the results were not very promising because the method forecast stock prices to an accuracy of 46%–60%.

Kim (2003) believed that SVMs are promising models for forecasting financial time series trends because they employ a risk role, including empirical error and a regularized factor, which results from the principle of structural risk minimization. In this study, SVM is applied to predict stock price movements. Additionally, the author studied the practicability of using an SVM for financial prediction by comparing it with BPNNs and case-based reasoning. The experimental findings showed that SVM systems offer a promising approach to stock market forecasting. The input variables employed in the research were technical indicators and the trends in the KOSPI. Because the author tried to predict the direction of daily price movements, technical indicators were employed as input data. The author chose nine technical indicators to build the initial attributes: stochastics %k, %D, slow %D, Momentum, ROC, Williams' %R, A/D oscillator, CCI, and RSI.

Rodríguez et al. (2005) investigated whether it is promising to develop the nonlinear manners of daily Spanish Ibex-35 stock index returns in order to make predictions over short-term and long-term horizons. the authors analyzed the prediction performance of

ANNs in out-of-sample and smooth transition autoregressive (STAR) systems. The findings suggested an improved fit for ANN models, in terms of the Sharpe risk-adjusted ratio and average net profitability, by using one-step-ahead predictions. The results showed that there is a good probability of achieving a more precise fit and prediction of the daily stock index returns by applying nonlinear techniques and one-step-ahead predictors, but that these are intrinsically compound and provide an awkward economic explanation.

Pai and Lin (2005) applied SVMs and ARIMA to solve nonlinear regression estimation problems. Further, the authors proposed a hybrid methodology that develops the unique power of the SVM method and the ARIMA method in predicting stock price movements. The authors used real data collections of stock prices to analyze the prediction accuracy of the planned model and believed that the results of their computational system are promising.

Pan et al. (2005) presented a computational model for forecasting the Australian stock market index (AORD) by applying multilayer feedforward ANNs from financial time series data and different interrelated markets. This attempt aimed to discover an efficient NN model or a collection of adaptive NN models for the purpose of forecasting. The goal was to develop and model various dynamical swings and influences of co-moving markets derived from expert technical and quantitative analysis. Within a partial range determined by the empirical information, three features of efficiency on data selection were measured: an adequate collection of interrelated markets, an adequate collection of co-moving markets, and efficient data from the target market (AORD). Two conventional scopes of the ANN structure were also considered: the optimal quantity of hidden layers and the optimal quantity of hidden neurons per hidden layer. This approach achieved three significant results: a six-day sequence was revealed in the Australian share market during the research period; the time signature applied as further inputs gave practical data; and a simple ANN employing six daily returns of the Australian share market and one daily return of the S&P 500 Index, in addition to the day of the week as data, showed up to 80% directional forecast accuracy. The authors used technical variables, the OHLC index level and traded volume, as input data for the study.

Motiwalla and Wahab (2000) used a switching rule trained on out-of-sample forecasts of returns in order to establish an investment arrangement in either treasury bills or stocks. The authors evaluated the economic importance of any apparent trends of predictability by integrating transaction costs in the computer-generated trading rules and strategies. They discovered that ANN methods generate switching signals that may develop through shareholders in an out-of-sample environment in order to obtain greater growth and risk-adjusted incomes when contrasted with either a simple buy-and-hold strategy or conventional regression in a market's index. The strength of these results across a large number of stock market indices is promising. The authors used 20 inputs as technical data and as outputs of ANN from the monthly data of various US stock indices: the NYSE Composite Index (NYER), the NASDAQ Composite Index (HCCXR), the S&P 100 Index (OEXR), the NASDAQ Industrials Index (HCDXR), the S&P 500 Index (SPXR), the Russell 3000 Index (RUAZR), the NYSE Industrials Index (NYIR), the S&P Mid-Cap Index (SNPR), the Value Line Index (VALR), the Russell Large-Cap 1000 Index (RUIZR), and the Wilshire 5000 Index (WLSR).

It has been largely admitted by many studies that financial markets consist of nonlinearity and that ANNs can be effectively applied to uncover this association. Sadly, many of these studies have not considered alternative prediction techniques, the importance of input variables, or the performance of the methods while applying various trading approaches. Enke and Thawornwong (2005) offered an information-obtaining technique applied to a machine-learning system for data mining in order to assess the forecasting relationships of many economic and financial variables. NN methods for classification and level estimation were then analyzed for their capabilities to supply helpful predictions of future values and returns. A cross-validation algorithm was also used to develop the generalization capability of many methods. The results of the study showed that the trading approaches and strategies directed by the categorization models create higher risk-adjusted benefits than the buy-and-hold strategy and those directed by level-estimation regarding the predictions of the ANN and conventional regression methods.

Chandwani and Saluja (2014) tested three sets of indicators (fundamental variables, a technical variables model, and hybrid variables), applying the separate and combined machine-learning techniques of SVM, genetic algorithm (GA)-SVM, ANN, and GA-ANN models. The outcomes of all three sets of indicators were contrasted in the four models. The main objective of the study was to determine a set from the above assigned models that forecast the movement of Indian stock prices in the most effective manner. Chandwani and Saluja (2014) used these indicators and variables as inputs for their machine-learning system. The technical indicators ($n = 4$ in all the parameters) were: mass index (MI), average true range (ATR), momentum, Chaikin money flow (CMF), commodity channel index (CCI), EMA, RSI, stochastic oscillator, William's %R, rate of change (ROC), price and volume trend (PVT), on-balance volume (OBV), and MACD (6,13). The fundamental variables were: growth in net sales, growth in net profit, return on equity, price/earnings (P/E) ratio, net profit margin (NPM), and price/sales ratio.

Chavan and Patil (2013) believed that none of the prediction methods have been proved to create satisfactory outcomes. Machine-learning techniques seem to be better than other conventional models because of their capability of nonlinear mapping. The authors surveyed various input indicators that could be applied for stock market forecasting using an ANN. They tried to find the most effective input variables that have a significant effect on the accuracy of predictions. From the literature, they discovered that most machine-learning techniques apply technical indicators rather than fundamental variables for specific stock price forecasting, whereas financial variables are generally applied to forecast stock market indices. Additionally, combined indicators outperform the single type of input variables.

Adebiyi et al. (2012) believed that data mining methods could be applied widely in the financial markets to help investors make decisions. The authors mentioned the use of ANNs for forecasting stock markets and stated that the application of technical indicators for stock forecasting is popular. The authors offered a hybridized model that combines the use of fundamental and technical analysis indicators of stock markets for forecasting incoming stock prices in order to develop current models. The hybridized model was studied together with collected stock data. The results showed significant improvement over the application of a single type of variables. In addition, the forecasting from the hybridized model was found to be a sufficient guide for the decision-making of investors and traders.

Adebiyi et al. (2012) examined the effect of combined market variables/indicators for improved stock price forecasting. The combined market variables included expert opinion variables and technical and fundamental ratios as inputs for the ANN method. The empirical outcomes achieved with the stock price data of Nokia and Dell collected from the NYSE showed that the offered technique could be helpful for advancing the accuracy of stock price forecasting.

Jaruszewicz and Mańdziuk (2004) considered the issue of stock index forecasting. The data were collected at the NIKKEI stock market and the NASDAQ and DAX markets.

The input consisted of not only the unique statistical values of the markets but also the variables pre-refined in terms of technical analysis; for example, the oscillators were measured and the organization of an assessment chart was derived. Selected information was then put into an NN system that was operationally split into different modules. The forecasting purpose was next day opening rates of the NIKKEI Index with a match to the German and US stock market indices. The results showed that "the average forecasting error on the analysis collection equals 43 points and the average percentage forecasting error is equal to 0.27% whereas the average index volatility equals 0.96%" (Jaruszewicz & Mańdziuk, 2004).

A pre-processing step of data mining is feature selection, which aims to filter out deceiving indicators from a given data collection for effective forecasting. Because the application of various feature selection models leads to the selection of various features and thus causes an impact on forecasting performance, the aim of Tsai and Hsiao (2010) was to create a hybrid of multiple feature selection models in order to present more representative indicators for better forecasting. Specifically, three famous feature selection models, GA, PCA, and decision trees (CART), were applied in the authors' study. The combined models were used to filter out deceiving or unrepresentative indicators and were based on strategies of multi-intersection, intersection, and combination. For the prediction technique, the BPNN was used. The experimental outcomes showed that the intersection between GA and PCA, and the multi-intersection of GA, PCA, and CART, were most successful, with accuracy levels of 79% and 78.98% respectively. Additionally, these two hybrid feature selection models filtered out approximately 80% of the unreliable features from 85 novel indicators, resulting in 14 and 17 significant features respectively. These indicators are the key terms for stock forecasting and can be applied for future trading decisions. Tsai and Hsiao (2010) used three models to choose their indicators through several fundamental variables (accounting ratios) and macroeconomic variables. Then, they selected indicators with technical variables in an ANN in order to forecast prices.

In accordance with the literature (Boboc & Dinică, 2013; Marshall et al., 2008; Metghalchi et al., 2012; Preen, 2009; Varutbangkul, 2013), this current study uses some technical variables as inputs for the proposed hybrid model. The variables are defined and discussed in Chapter 3.

2.11 The Financial Crisis

Portfolio managers always seek a trading method to generate reliable and robust return, especially during structural breaks or a crisis. This study investigates the results of the proposed hybrid model, together with long-term memory and entropy, before, during, and after structural breaks and crises. The most notable crisis during the period of the study is the 2008 financial crisis. The financial crisis of 2007 to 2008, also known as the 2008 financial crisis and the global financial crisis, is considered by many economists and academics to have been the most significant financial crisis since the 1930s and the Great Depression (Eigner & Umlauft, 2015; Temin, 2010). However, our proposed model shows robust results even during financial crisis.

The 2008 financial crisis threatened to cause the collapse of large financial institutions and witnessed the bailout of banks by national governments (Williams, 2012). In addition, equity markets fell worldwide. In many regions, the property markets also suffered, causing foreclosures, evictions and prolonged unemployment. The crisis caused consumer wealth to decline by trillions of US dollars, led to the collapse of strategic businesses, and created a recession in economic activity that contributed to the European sovereign-debt crisis and resulted in the Great Recession of 2008–2012 (Williams, 2012).

The growth falloffs in emerging Asian economies during the 2008 financial crisis were not as drastic as those experienced by the five most affected Asian emerging economies and by the economies of the commonwealth of independent states (CIS) throughout Asia's own financial crisis of 1997–98. However, Korea, Hong Kong, Singapore, and Malaysia suffered significant falls in growth through the period of the 2008 crisis. Further, even India and China saw their economic growth rates fall to approximately half of their pre-crisis highs. At the peak of the crisis, emerging Asia economies resembled those of other emerging economies in terms of peak-to-through fluctuations in stock markets and in terms of exports and spikes in financial indices.

2.12 Development of Research Methodology

In this section the hypotheses of the study are build up based on the literature review. According to the literature review, forecasting financial time series is extremely interesting for researchers and professional market traders. A significant number of studies have been conducted in this field (Sections 2.2-2.9), and several hypotheses have been generated (Section 2.2). Different hypotheses have different perspectives, with various tests and analyses. Challenging market theories and generating high profits from the capital markets is certainly one motive behind such studies and hypotheses. The main motivation behind this current study is to investigate whether stock index futures possess long-term memory and to determine whether innovatively combined new techniques such as wavelets, PCA, and ANNs generate returns in excess of the buy-and-hold benchmark.

One of the objectives of this study is to investigate the problem of market movement and to establish if it is random. In accordance with the literature (Section 2.1) and the various tests and analyses that scholars have conducted (Section 2.2), this current study tests the random walk theory and the AMH (Section 2.3). Using unit root (stationarity), autocorrelation, and the Granger causality test, this study investigates long-term memory and entropy in futures markets (Sections 2.4 and 2.5). Further, in accordance with the review surrounding the issue of long-term memory, this study employs the Hurst exponent presented by Mandelbrot (1969) as a reliable index for true long-term memory through the use of financial time series (Sections 2.4.1 and 2.4.2). The Hurst exponent can capture the predictability of a market when there is long-term memory in the market; approximate entropy can capture the predictability and efficiency of a market when there is complexity and disorder. In addition, if the Hurst exponent and approximate entropy, as the efficiency indices of the market, vary over time, a market experiences changes in efficiency during different periods, a circumstance that is consistent with the AMH (Sections 2.3 and 2.4.2). Thus, this study posits hypothesis 1 as follows.

H1: Futures markets exhibit significant changes in predictability over time, taking into account long-term memory and approximate entropy.

Moreover, taking account of the surveyed empirical studies (Sections 2.8 and 2.9), this research aims to determine whether a novel hybrid forecasting system, the so-called WPCA-NN, generates better returns than existing modern technical trading rules, such as the singular ANN forecasting method and the combination of ANN with a wavelet, for trading in today's increasingly complicated and volatile futures markets. According to the literature review (Section 2.8), ANNs have deficiencies in terms of hidden noise in time series; consequently, this study offers a combination of an ANN with wavelet analysis, which is a special technique for denoising signals (Section 2.9). In addition, based on an enhancement for multivariate signals by wavelet transform, using PCA proposed by Aminghafari et al. (2006), this study develops a novel hybrid intelligent system, the so-called WPCA-NN. Using this hybrid technique, it is proposed here that not only can the model denoise financial time series delicately, it can also make forecasts with high performance, accurate results. Hence, another objective of this study is to investigate whether the proposed novel hybrid model, the WPCA-NN model, can produce consistently and significantly higher returns than the buy-and-hold benchmark (Section 2.7.1). Thus, we propose the following hypothesis.

H2: The hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the random walk hypothesis (a passive buy-and-hold strategy) for selected financial markets, as shown by significant mean differences and significant positive alphas in profit regressions.

In addition, this study compares the empirical results of the WPCA-NN model with the results of the best technical trading rules (Chong et al., 2010; Chong & Ng, 2008; Fernández-Blanco et al., 2008; Rosillo et al., 2013) according to the literature (Section 2.7.2). Thus, the study offers the following hypotheses.

H3: Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other best-performing technical trading rules (such as MACD, RSI, OMA and stochastics).

In order to examine the reliability of the proposed model, this study compares the empirical results of the WPCA-NN model's use with singular ANN and WNN models. Since the WPCA-NN model is constructed with a wavelet, PCA, and ANN, we examine the results with the results of its components, the pure ANN and WNN models. A singular ANN model is a pure forecasting technique without any denoising process (Section 2.8). Using univariate wavelet denoising and then an ANN produces a denoising-forecasting technique, the WNN (Section 2.9). This study's novel model applies a multivariate denoising process using a wavelet and PCA in order to provide inputs to the ANN forecasting machine. Thus, the components of the WPCA-NN model, which must be compared with this study's proposed model, are WNN and singular NN models.

H4: Significant mean differences and significant positive alphas in profit regressions show that:

a. H4a: The WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other pure ANN forecasting methods.

b. H4b: The WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than a WNN model.

According to the reviewed literature, no study has used wavelet and PCA as multivariate denoising for financial time series and forecast future values with an ANN. This study is the first attempt to apply WPCA and ANN models on futures markets' OHLC data as a multivariate signal with the purpose of capturing the denoised signal, which is the signal that fits best with the original signal. The model is a novel hybrid model that is applied to the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets in order to test its forecasting power in fluctuating financial markets, confirm its reliability, and assess its generalizability.

2.13 Literature Review Summary

In this section, we present a summary of the literature review as a comprehensive table. This table indicates the details of the studies; namely, their objectives, methodologies, sample data, statements of results, and influence on this current research. The table helps to identify the gap in the literature and contributes to this current study's development of an innovative and suitable research methodology. Table 2.4 presents a summary of the literature review.

Table 2.4: Summary of the literature review

No.	Authors	Year	Title	Journal	Ref.	Objectives	Techniques	Data Type	Results	Comments
1	Halbert White	1988	Economic prediction using NNs: The case of IBM daily stock returns	Neural Networks, 1988. IEEE International Conference	(White, 1988)	To search for and decode nonlinear regularities in asset price movements.	ANN	IBM daily stock returns	They report that an ANN outperforms statistical methods and achieves higher forecasting accuracy.	Very strong and comprehensive calculation and formulas in methodology part, which is good. They offer useful and general modifications to standard learning techniques. They explain how to make a study applicable and useful in other context.
2	Kuo, Lee, Lee	1996	Integration of ANNs and fuzzy Delphi for stock market forecasting	Systems, Man, and Cybernetics, IEEE International Conference	(Kuo et al., 1996)	To consider the quantitative factors as well as the qualitative factors in forecasting markets.	ANN and fuzzy	Taiwan stock market	This study showed that the proposed intelligent stock market prediction system is capable of handling quantitative and qualitative factors.	They use a hybrid model with an ANN for technical variables, and fuzzy Delphi for technical indicators and fundamental or qualitative variables. They show that the fundamental environment plays an important role in forecasting, as well as technical indicators. They use OHLC, MACD, stochastics, RSI, momentum, etc. as inputs for their method.

3	Hirotaka Mizuno, Michitaka Kosaka, Hiroshi Yajima	1998	Application of an NN to technical analysis of stock market prediction	Studies in Informatics and Control	(Mizuno , Kosaka, Yajima, & Komoda , 1998)	A forecasting model for technical analysis of a stock market, and the model's application to a buying and selling timing prediction model for stock prices.	ANN	Stock prices	The NN technique did not achieve the higher gains in buying positions than other cases; however, it made the highest gains in selling positions in terms of the meaning of a minimum loss.	The interesting point of this study is the use of 11 technical indicators of TOPIX, and the promising results for the application of an ANN. The study shows how to improve forecasting accuracy by controlling the number of samples in each category based on the importance of each category. The current study uses confirmed indicators with an ANN.
4	Luvai Motiwalla, Mahmoud Wahab	2000	Predictable variation and profitable trading of US equities: a trading simulation using NNs	Computers & Operations Research	(Motiwa lla & Wahab, 2000)	To set up investment positions in either stocks or treasury bills.	ANN, buy and hold	Stock market indices	They report that ANN models achieve superior cumulative and risk-adjusted returns when compared with either a simple buy-and-hold strategy or regression in the market indices. The robustness of their results across a large number of stock market indices is promising.	They use 20 inputs that they claim as technical data, and for the output of an ANN they use monthly data on a variety of US stock indices: the NASDAQ Composite Index, the S&P Mid-Cap Index, the NASDAQ Industrials Index, the NYSE Composite Index, the S&P 100 Index, the Russell 3000 Index, the NYSE Industrials Index, the Russell Large-Cap 1000 Index, the S&P 500 Index, the Value Line Index, and the Wilshire 5000 Index.

5	Angelos Kanas, Andreas Yannopoulos	2001	Comparing linear and nonlinear forecasts for stock returns	International Review of Economics & Finance	(Kanas & Yannopoulos, 2001)	To check the prediction performance of monthly return forecasts for two indices, the Dow Jones (Hussain et al.) and the Financial Times (Eigner & Umlauft) indices.	ANN	Stock indices	The results show that the inclusion of nonlinear terms in the relationship between stock returns and fundamental variables is important in out-of-sample forecasting.	This study represents linear and nonlinear forecasting. Trading volume and stationary transformations of dividends are considered as fundamental indicators in the linear model and the input variables in the ANN model. The forecasting performance comparison is done on the basis of forecast accuracy, employing the Diebold and Mariano test.
6	Xiaodan Wu, Ming Fung, Andrew Flitman	2001	Forecasting stock market performance using an HIS	International Conference on Computational Science	(Wu et al., 2001)	To check the prediction power of the offered technique.	Feedforward neuro-fuzzy (FFNF) model	S&P 500 Index	The hybrid intelligent technique they offer, an FFNF model, yields better performance in the sense of computational efficiency, forecasting accuracy, and generalization capability. The FFNF model also overcomes the black art method in conventional ANNs by applying "transparency."	Prior S&P 500 results and six economic indices are used as inputs: the consumer price index, total industrial production index, leading economic index, bank prime loan rate, federal funds rate, unemployment rate, S&P 500 lag(1), S&P 500 lag(1), S&P 500 lag(1), and S&P 500 closing prices output.
7	Bai-Ling Zhang, Zhao-Yang Dong	2001	An adaptive neural-wavelet model for short-term load forecasting	Electric Power Systems Research	(Zhang & Dong, 2001)	To combine the inherent ability of wavelets and ANNs to capture non-stationary and nonlinear attributes inherited in	Neural-wavelet	Electricity market	They report very promising results by applying a wavelet and neural network as an ensemble.	The prediction model is based on wavelet multi-scale decomposition and multilayer perceptron (MLP) neural network modeling of wavelet coefficients. To lower the influence of noisy low-level coefficients, they use the practical Bayesian model and automatic relevance determination

						financial time series.				technique to select the size of the MLPs.
8	Bai-Ling Zhang, Richard Coggins	2001	Multi-resolution forecasting for futures trading using wavelet decompositions	IEEE Transactions on Neural Networks	(Zhang et al., 2001)	To investigate the effectiveness of a financial time series prediction strategy that clears the multi-resolution attribute of the wavelet transform.	Wavelet and an NN	Exchange rate	They perform a trading strategy to check the results and measure forecasting accuracy with various scales. By applying an accurate trading method, their system displays promising profitability performance. Outcomes that compare the performance of the proposed model with MLP bond futures show a doubling in gain per trade and a Sharpe ratio increase, as well as a significant increase in the ratio of winning to losing positions; thus indicating significant potential profitability for live trading.	They use a wavelet and NN. They analyze the signal in high frequency and low frequency by using the Bayesian method of automatic relevance determination. They perform a trading strategy to check the results, and measure forecasting accuracy with various scales. There are encouraging results with the usage of the wavelet and NN.
9	Kim	2003	Financial time series forecasting using SVMs	Neurocomputi ng	(Kim, 2003)	To check the forecasting power of an SVM and compare it to an ANN.	SVM and ANN	Stock price index	The results show that an SVM is a suitable alternative for predicting financial time series.	In this study, the polynomial kernel and the Gaussian radial basis function are employed as the kernel function of an SVM. Since there are few general guidelines to specify the settings of an SVM, this study varies the parameters to choose the best values for the best forecasting performance.

10	Dennis Olsona, Charles Mossman	2003	NN forecasts of Canadian stock returns using accounting ratios	International Journal of Forecasting	(Olson & Mossman, 2003)	The forecasting power of a BPNN in Canadian companies.	BPNN, ordinary least squares, and logit	Stock prices of Canadian companies	The results show that a BPNN outperforms conventional methods such as ordinary least squares and logit.	They use 61 accounting ratios as input variables and a BPNN. The data are approximately 2300 Canadian companies. The results are not so promising. They can forecast 46–60 % correctly. This finding could be a good reason why researchers mostly use technical indicators and not accounting ratios.
11	Wei Huang, Yoshiteru Nakamoria, Shou-Yang Wang	2005	Forecasting stock market movement direction with an SVM	Computers & Operations Research	(Huang, Nakamori, & Wang, 2005)	To check the predictability of financial movement direction with AN SVM by predicting the weekly movement direction of the NIKKEI 225 Index.	SVM	Stock index	They propose a hybrid model of an ANN and SVM. This hybrid model outperforms all other models.	An SVM has been used for its capacity control of the decision function and the application of the kernel functions. They compare their results with linear discriminant analysis, quadratic discriminant analysis, and Elman BPNNs. They propose a combination of all the above methods..
12	An-sing Chen, Mark Leung	2004	Regression NN for error correction in foreign exchange forecasting and trading	Computers & Operations Research	(Chen & Leung, 2004)	To propose an adaptive forecasting approach that assembles the strengths of NNs and multivariate econometric methods in order to forecast the market better than the random walk model.	Regression NN	Exchange rate	They claim a very good result, with error correction through the use of an NN that outperforms random walk theory.	They use a hybrid model including Bayesian. The data relate to a foreign exchange.

13	Suraphan Thawornwong , David Enke	2004	The adaptive selection of financial and economic variables for use with ANNs	Neurocomputi ng	(Thawor nwong & Enke, 2004)	To examine whether using the current relevant variables guides achieves additional improvements in stock return prediction.	ANN	Stock return	They use an adaptive technique to select input variables and feed to an ANN. They report that their method outperforms buy and hold.	The interesting point is the collection of financial and economic variables and the use of an ANN with 24 years of data.
15	Lean Yu, Shouyang Wang, Kin Keung Lai	2005	A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates	Computers & Operations Research	(Yu, Wang, & Lai, 2005)	To obtain accurate prediction results and ameliorate forecasting performances.	Hybrid of generalized linear auto-regression (GLAR) with artificial neural networks (ANN)	Foreign exchange rate	Using generalized linear autoregressive (GLAR) and ANN models, and a novel ensemble technique, they outperform a single ANN and buy and hold.	A perfect combinations model for a hybrid. They offer two models for combined forecasting. The data are from a stock exchange, but the result is marvelous. The authors introduce equal weight and minimum error for comparisons.
16	Serdar Yümlü, Fikret Gürgen , Nesrin Okay	2005	A comparison of global, recurrent and smoothed-piecewise neural models for Istanbul stock exchange prediction	Pattern Recognition Letters	(Yümlü, Gürgen, & Okay, 2005)	To give a comparison of global feedback and smoothed-piecewise artificial neural prediction models for financial time series forecasting problems.	Recurrent and smoothed-piecewise neural models	Istanbul stock exchange	The results show that the smoothed-piecewise neural technique is advantageous for extracting volatility in index return series, improving on the conventional EGARCH volatility technique and the global and feedback neural method.	This study predicts integral squared error (ISE) with various architectures of ANNs, multilayer perceptrons, and a smoothed piecewise model and is compared to an exponential generalized autoregressive conditional heteroscedasticity (EGARCH) volatility model. The accuracy of the results are measured by hit rate, mean squared error (MSE), and mean absolute error (MAE).

17	Pai, Lin	2005	A hybrid ARIMA and SVM model in stock price forecasting	Omega	(Pai & Lin, 2005)	To see whether a novel neural network model can be successfully applied in order to solve nonlinear regression estimation issues.	ARIMA and SVM	Stock price	They claim their data are promising and that the hybrid model of ARIMA and an SVM outperforms single models.	They use daily stock data in a very simple study. They use ARIMA and SVM separately at first and then together as a hybrid on stocks' daily closing prices.
18	Heping Pan, Chandima Tilakaratne, John Yearwood	2005	Predicting Australian stock market index using NNs exploiting dynamical swings and intermarket influences	Journal of Research and Practice in Information Technology	(Pan et al., 2005)	To discover an effective NN or a set of adaptive neural networks for predicting the Australian stock market index, employing multilayer FNN from the time series data of AORD and numerous interrelated markets.	Feedforward NN	Australian stock market index	There are two results in this study: a time signature is applied because additional input provides applicable information, and a 6-day cycle is seen in the Australian stock market.	An interesting study because of the use of interrelated markets in Australia, such as the co-moving market. The open, high, low, close index level is used together with the traded volume for the trading period of time *t*. From the autocorrelation and partial autocorrelation tables (Tables 1 and 2), it is clear that there is a 6-day cycle in the working week in the Australian stock market, not a 5-day cycle.
19	David Enke, Suraphan Thawornwong	2005	The use of data mining and NNs for forecasting stock market returns	Expert Systems with Applications	(Enke & Thawornwong, 2005)	To introduce an information gain method used in machine learning for data mining to assess the predictive relationships of various variables.	ANN	Market returns	The outcomes show that the trading strategies led by the classification models make higher risk-adjusted gains than the buy-and-hold theory, those guided by linear regression models, and the level-estimation based forecasts of the neural network.	Very useful data-mining technique (with its literature and its reasons) for selecting the best and the most effective variables for stock forecasting. Important variables have also been used in this study.

20	Antonio Conejoa, Javier Contrerasa, Rosa Espınola, Miguel Plazas	2005	Forecasting electricity prices for a day-ahead pool-based electric energy market	International Journal of Forecasting	(Conejo et al., 2005)	To predict the 24 market-clearing prices of a day-ahead electric energy market.	NNs and a wavelet	Electricity prices	They consider and test all the techniques and claim that the combination of a neural network and wavelet outperforms the other methods.	They use a neural network and wavelet, and compare the results to ARIMA, a transfer function, and dynamic regression.
21	Chen, Wun-Hua Shih, Jen-Ying Wu, Soushan	2006	Comparison of SVMs and BPNNs in forecasting the six major Asian stock markets	International Journal of Electronic Finance	(Chen, Shih, & Wu, 2006)	To compare the forecasting accuracy of an SVM and BPNN.	SVM vs BPNN	Six major Asian stock markets	Their results indicate the superiority of SVM and BPNN models, compared with earlier studies and conventional methods.	They use technical indicators as input and SVM and BPNN models for comparison.
22	Aminghafari, Mina, Cheze, Nathalie, Poggi, Jean-Michel	2006	Multivariate denoising using wavelets and PCA	Computational Statistics & Data Analysis	(Aminghafari et al., 2006)	To propose a multivariate signal denoising process.	Wavelet transform and PCA		The results show that denoising multivariate signals by wavelet PCA (WPCA) performs better than univariate wavelet denoising on each factor separately, whereas WPCA extracts the same noise at different frequencies from elements of a multivariate signal.	They propose a multivariate signal denoising process using wavelet and principal component analysis. They conclude that this multivariate denoising outperforms univariate denoising process. The current study uses their denoising process with multivariate signal of OHLC, to capture the same noise in the signal and achieve a more accurate underlying or original form of signal (a futures indices' signal).
23	Kin Keung Lai, Kaijian He, Jerome Yen[3]	2007	Modeling VaR in crude oil market: a multi scale nonlinear ensemble approach incorporating wavelet analysis and ANN	International Conference on Computational Science, 2007	(Lai et al., 2007)	To adopt a VaR approach to calculate risks and offer multi-scale non-linear combined approaches to simulate the risk evolutions in the WTI crude oil market.	Wavelet analysis and ANN	Value at risk for crude oil market	The hybrid outcome shows promising results for decomposing useful information from a market with a wavelet and neural network, in order to forecast trends.	This study uses wavelet analysis to decompose useful information in order to feed to a neural network. The study shows another successful result from a wavelet and neural network, and demonstrates a simple but effective combination for a wavelet with an ANN.

24	Huang, Wei Lai, Kin Keung Nakamori, Yoshiteru Wang, Shouyang Yu, Lean	2007	NNs in finance and economics forecasting	International Journal of Information Technology & Decision Making	(Huang et al., 2007)	To discuss the input variables, types of neural network technique, and performance comparisons for the forecasting of economic growth and foreign exchanges.	Different types of ANN	Economic growth, foreign exchange rates, stock market index	They suggest that the forecasting performance of neural networks can be increased with integration with other intelligent technologies. Nonlinear combination prediction with neural networks also generates encouraging results.	A survey about neural networks and their functions in finance and economics that focuses on different variables, various models of ANN, performance measuring, and forecasting target in finance.
25	George Atsalakis, Kimon Valavanis	2009	Surveying stock market forecasting techniques–part II: soft computing methods	Expert Systems with Applications	(Atsalakis & Valavanis, 2009b)	To review forecasting techniques and variables for stock markets.	Variety of artificial intelligent techniques	A review study	They suggest hybrid and single forms of artificial intelligent models with various technical indicators. These astonishingly outperform traditional methods.	A review of different AI forecasting techniques with various variables and indicators used in the reviewed studies. The current study uses the common and most successful technical indicators such as MACD, OHLC, RSI, stochastics, ultimate oscillator, and volume.
26	Pei-Chann Chang, Chin-Yuan Fan	2009	A hybrid system integrating a wavelet and TSK fuzzy rules for stock price forecasting	IEEE Transactions on Systems, Man, and Cybernetics - TSMC	(Atsalakis & Valavanis, 2009b)	To propose a novel hybrid approach for predicting the Taiwan stock exchange index.	Wavelet and Takagi–Sugeno–Kang (TSK)-fuzzy-rule-based systems	Stock price	Simulation results show that the hybrid model with wavelet denoising successfully forecast the price variation for stocks, with an accuracy of up to 99.1%, for the Taiwan stock exchange index.	They propose a hybrid form of wavelet transform and fuzzy logic for forecasting stock prices.. Encouraging results for the use of wavelet analysis are presented in the study.
27	Camillo Lento	2008	A combined signal approach to technical analysis on the S&P 500	Available at SSRN 1113622 (2008)	(Lento, 2008)	To examine the effectiveness of nine technical trading rules on the S&P 500.	Combined signal approach (CSA) and buy-and-hold model	S&P 500 Index	They report that excess returns were made by employing a combined signal approach (CSA) on separated trading rules. Statistical significance was confirmed as robust through a subperiod technique and bootstrap simulations.	A good literature review and successful results for the S&P 500 from January 1950 to March 2008 with 14,646 items of daily data.

28	Ohanissian, Arek Russell, Jeffrey Tsay, Ruey	2008	True or spurious long memory? A new test	Journal of Business & Economic Statistics	(Ohanissian et al., 2008)	To differentiate between true long-term memory and spurious long-term memory.	Geweke Porter-Hudak	Foreign exchange rate	The results show that the long-term memory property in foreign exchange rate volatility is caused by a true long-term memory procedure.	This study proves that the long-term memory process in selected foreign exchange markets is truly generated and is not spurious. Structural breaks and short-term processes do not have any effect on the long-term memory process.
29	George Atsalakis, Kimon Valavanis	2009	Forecasting stock market short-term trends using a neuro-fuzzy based methodology	Expert Systems with Applications	(Atsalakis & Valavanis, 2009a)	To discover the prediction power of a proposed neuro-fuzzy model and control the stock market process.	Adaptive neuro-fuzzy inference system (ANFIS)	Athens and New York stock exchanges	The results show that the authors' method outperforms the buy-and-hold strategy. They compare their technique with the weak form of the efficient market hypothesis. They assess their method with root-mean-square error (RMSE) and the profitability of trading.	A neuro-fuzzy model is established to control the stock market process. The authors compare their method with the weak form of the EMH. The results show that their method outperforms the buy-and-hold strategy.
30	Vanstone, Finnie	2009	An empirical methodology for developing stock market trading systems using aANNs	Expert Systems with Applications	(Vanstone & Finnie, 2009)	To present a methodology for designing robust mechanical trading systems, employing soft computing technologies such as ANNs.	ANN, expert trading system	Stock market index	They report encouraging results by making a complete and robust trading system with an ANN that outperforms conventional trading strategies.	A complete expert system study, with a comparison of fundamental and technical indicators. The authors finally choose only technical indicators. They also introduce trading skills and trading methods to build a complete trading machine with ANN methodology.
31	Ito, Mikio, Sugiyama, Shunsuke	2009	Measuring the degree of time varying market inefficiency	Economics Letters	(Ito & Sugiyama, 2009)	To estimate a time-varying autocorrelation of stock market returns as a measure of market inefficiency.	Moving window method	US stock market	They investigate a time-varying autocorrelation of stock market returns as a measure of market inefficiency. The dependent inefficiency of the US stock market varies from 1955 to 2006.	They find that autocorrelation of US stock market returns varies over time, which is evidence of changing efficiency over time. This finding strengthens the theory of the adaptive nature of the market, which is consistent with the AMH.

32	Arash Bahrammirzaee	2010	A comparative survey of artificial intelligence applications in financial ANNs, expert systems, and HISs	Neural Computing and Applications	(Bahrammirzaee, 2010)	To review studies using ANN and its other hybrid forms in different fields of finance.		A review study	The results show that hybrid ANN forms perform better than a single form of AIs and all reported traditional strategies.	In this review, the hybrid ANN form is shown to perform the best. The hybrid ANN form is used in the current study in accordance with the strong and reliable literature, partly drawn from this study.
33	Boyacioglu, Melek Acar Avci, Derya	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	(Boyacioglu & Avci, 2010)	To investigate the predictability of stock market return with ANFIS.	ANFIS	Stock market returns of the Istanbul stock exchange	They report accuracy results of 98.3%. For reprocessing, they use a wavelet and then an adaptive neuro-fuzzy inference system (ANFIS).	They use some fundamental and technical variables, which are particularly interesting. They focus on the Istanbul market index.
34	Lu, Chi-Jie	2010	Integrating independent component analysis-based denoising scheme with an NN for stock price prediction	Expert Systems with Applications	(Lu, 2010)	To forecast stock prices that are inherently noisy and non-stationary data.	BPNN and ICA	NIKKEI 225 and TAIEX indices	The study's results show that the proposed model performs better than the BPN technique with non-filtered prediction variables and a random walk model.	This study introduces a very promising model to remove noise as much as possible: an integrated independent component analysis (ICA). Following this, an NN is used for forecasting.
35	Tsai, Chih-Fong, Hsiao, Yu-Chieh	2010	Combining multiple feature selection methods for stock prediction: union, intersection, and multi-intersection approaches	Decision Support Systems	(Tsai & Hsiao, 2010)	To predict effectively stock prices for traders and investors.	PCA, genetic algorithm, and decision trees	Stock market index	They use a combination of fundamental and technical variables in an ANN to forecast prices. Their results' accuracy is approximately 79%.	A reliable study for confirming the selection of variables. The authors use three models to choose their indicators through many fundamental variables (accounting ratios) and macroeconomic variables.

36	Cheng, Jao-Hong, Chen, Huei-Ping, Lin, Yi-Min	2010	A hybrid forecast marketing timing model based on a probabilistic NN, rough set, and C4.5	Expert systems with Applications	(Cheng et al., 2010)	To integrate a supply chain with finance to control marketing timing.	A probabilistic neural network, rough sets, and a C4.5 decision tree	Market indices	They use a combination of technical and fundamental variables with an ANN and a decision tree. The results show that the combination of technical and fundamental variables is encouraging.	A combination of technical and fundamental variables, with an ANN and a decision tree. A helpful methodology, especially in terms of classification. A decision tree is used to select technical and fundamental variables.
37	Lukas Menkhoff	2010	The use of technical analysis by fund managers: international evidence	Journal of Banking & Finance	(Menkhoff, 2010)	To show the importance of using technical indicators.	Survey	Technical and fundamental indicators	The results indicate that find managers rely on technical indicators.	This study is a survey of approximately 600 fund managers of five countries. It concludes that fund managers mostly trust and use technical indicators rather than fundamental indicators.
38	Tsai, Chih-Fong, Lin, Yuah-Chiao, Yen, David C, Chen, Yan-Min	2011	Predicting stock returns by classifier ensembles	Applied Soft Computing	(Tsai, Lin, Yen, & Chen, 2011)	To establish a better model for forecasting stock returns effectively and efficiently and investigate forecasting performance that utilizes the classifier ensemble model to analyze stock returns.	Neural networks, decision trees, and logistic regression	Stock returns	Their results show that multiple classifiers outperform single classifiers in terms of forecasting accuracy and investment returns.	They offer two types of ensembles: “homogeneous” classifier ensembles and “heterogeneous” classifier ensembles.

39	Rodríguez-González, Alejandro, García-Crespo, Ángel, Colomo-Palacios, Ricardo, Iglesias, Fernando Guldrís, Gómez-Berbís, Juan Miguel	2011	CAST: Using NNs to improve trading systems based on technical analysis by means of the RSI financial indicator	Expert Systems with Applications	(Rodríguez-González, García-Crespo, Colomo-Palacios, Iglesias, & Gómez-Berbís, 2011)	To check the prediction power of an offered model to improve trading systems.	ANN	Stock market index and market prices	They create and use an intelligent RSI with an ANN to forecast a market index and some market prices. They show that a modified RSI with an ANN outperforms RSI and conventional trading strategies.	They use an ANN to modify an RSI and create an intelligent RSI (iRSI) for better predictions. Then they use the model to forecast a market index and some market prices.
40	Hsieh, Hsiao, Yeh	2011	Forecasting stock markets using wavelet transforms and RNNs: an integrated system based on an artificial bee colony algorithm	Applied Soft Computing	(Hsieh et al., 2011)	To check the prediction accuracy of a proposed hybrid artificial intelligent model based on an artificial bee colony algorithm.	Wavelet transforms and recurrent neural network (RNN) based on an artificial bee colony (abc) algorithm (called ABC-RNN)	Stock market indices: DJIA, FTSE, NIKKEI, and TAIEX	They test their hybrid model in many markets. The results sound promising. They measure forecasting accuracy with various scales such as RMSE, MAE, MAPE, MSE, and the returns from trading.	This study uses a wavelet, a recurrent NN, stepwise regression-correlation selection (SRCS), and an artificial bee colony algorithm. The authors test their results in many markets. The results sound promising. The most important aspect of the study is the method that the authors use for the wavelet and recurrent NN. The results can also be compared with those of the current study.
41	Pang, Ye Xu, Wei Yu, Lean Ma, Jian Lai, Kin Keung Wang, Shouyang Xu, Shanying	2011	Forecasting the crude oil spot price with WNNs, using OECD petroleum inventory levels	New Mathematics and Natural Computation	(Pang et al., 2011)	To estimate the crude oil price with a WNN model and the use of an industrial petroleum inventory level.	Wavelet neural network	Crude oil spot price	They use a WNN to forecast crude oil prices and claim that their results are promising and that the WNN is successful. This WNN is slightly different from the current study's methodology, but shows inspiring results	The WNN used here differs slightly from the current study's methodology but shows promise.

									for a wavelet/ANN hybrid.	
42	Jian-Zhou Wang, Ju-Jie Wang, Zhe-George Zhang, Shu-Po Guo	2011	Forecasting stock indices with a BPNN	Expert Systems with Applications	(Wang et al., 2011)	To check the forecasting accuracy of a novel wavelet neural network model for the prediction of stock indices.	BPNN	Stock indices	They compare an ANN with a wavelet denoised ANN. They claim that the WNN outperforms a simple ANN; thus, the denoising process is successful.	They offer a new hybrid of wavelet denoising and ANN forecasting.
43	Luis Ortega	2012	A neuro-wavelet method for the forecasting of financial time series	Proceedings of the World Congress on Engineering and Computer Science	(Ortega, 2012)	To examine the effectiveness of nine technical trading rules on the S&P 500.	Neuro-wavelet	Stock returns	The results show that encouraging forecasting accuracy for one, three, and five step-ahead periods were achieved by the Jordan neural network wavelet model. A hybrid form of a wavelet and the Jordan NN model outperforms other models in the study.	This is a neuro-wavelet study to forecast Apple stock. A Jordan net, which is applied as an input variable to a NN, and the coefficients calculated from a Haar wavelet transform decomposition of the stock prices, display consistently superior predictive and modeling performance over other methods.
44	Rania Jammazi, Chaker Aloui	2012	Crude oil price forecasting: experimental evidence from wavelet decomposition and NN modeling	Energy Economics	(Jammazi & Aloui, 2012)	To achieve prominent forecasting of the crude oil price despite the intrinsic complexity of the oil market structure.	Wavelet decomposition and NN modeling	Crude oil prices	They claim their methodology and results are successful, and that a hybrid of wavelet decomposition and a NN outperform other models.	They use a Haar and Trous wavelet for decomposition and denoising, then a BPNN for predictions of crude oil prices.

45	Busse, Sebastian Helmholz, Patrick Weinmann, Markus	2012	Forecasting day ahead spot price movements of natural gas–an analysis of potential influence factors on basis of a NARX-NN	Multikonferenz Wirtschaftsinfo rmatik	(Busse et al., 2012)	To forecast day-ahead spot prices.	NARX-NN	Natural gas prices	The result show that the application of a NARX-NN model for the forecasting of natural gas prices could be successful.	They apply a NARX-NN model for natural gas price forecasting and achieve promising results.
46	Charles, Amélie Darné, Olivier, Kim, Jae	2012	Exchange-rate return predictability and the adaptive markets hypothesis: evidence from major foreign exchange rates	Journal of International Money and Finance	(Charles et al., 2012)	To examine return predictability.	Martingale difference hypothesis	Foreign exchange rate	When exchange rate returns are shown to be unpredictable most of the time, they identify a number of periods of statistically significant return predictability.	This results shows that the return predictability of foreign exchange rates occurs in different episodes depending on changing market circumstances. This is consistent with the explanation of the adaptive market hypothesis.
47	Arouri, Mohamed El Hedi Hammoudeh, Shawkat Lahiani, Amine Nguyen, Duc Khuong	2012	Long-term memory and structural breaks in modeling the return and volatility dynamics of precious metals	The Quarterly Review of Economics and Finance	(Arouri et al., 2012)	To examine the potential of structural changes and long-term memory attributes in returns and volatility.	ARFIMA-FIGRACH	Metal commoditie s	The empirical results show that the conditional volatility of the studied metal commodities is far better explained by long-term memory than by structural breaks.	They report valuable evidence of the existence of true long-term memory in the precious metals markets. Moreover, the conditional volatility of the studied metal commodities is far better explained by long-term memory than structural breaks.
48	Ling-Jing Kao, Chih-Chou Chiu , Chi-Jie Lu, Chih-Hsiang Chang	2013	A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting	Decision Support Systems	(Kao, Chiu, Lu, & Chang, 2013)	To forecast stock prices accurately.	Wavelet, multivariate adaptive regression splines (MARS), and support vector regression (SVR)	Stock market index	The empirical results show that the proposed approach of wavelet-SVR-MARS not only solves the problem of wavelet sub-series selection but also outperforms competing models.	They study uses wavelets in a hybrid model with MARS and SVR, and compares the results of the combinations of all three models. The authors suggest the use of this model by combining it with an NN model or "grey system theory." The hybrid model with a wavelet outperforms other methods in the study.

49	Nemes, Magdalena Daniela Butoi, Alexandru	2013	Data mining on Romanian stock market using NNs for price prediction	Informatica Economica	(Nemes & Butoi, 2013)	To forecast and data-mine the Romanian stock market.	NARX-NN	Stock price	They report promising results by applying the NARX-NN model in the Romanian stock market; however, MLP NN also has significant results, although the NARX model outperforms the MLP and other tested models.	They mine a good source of information and conclude that the NARX-NN model outperforms the MLP-NN and other tested models. This is another successful experience with the NARX-NN model and stock markets, and confirms the use of technical indicators.
50	Christopher Neely, David Rapach Jun Tu and Guofu Zhou	2014	Forecasting the equity risk premium: the role of technical indicators	Management Science	(Neely, Rapach, Tu, & Zhou, 2014)	To compare the predictive ability of technical indicators with that of macroeconomic variables in order to forecast US equity risk premium.	Convention al models	US equity risk premium	They show that substantial countercyclical ups and downs in equity risk premiums seem to be appropriately extracted by the combined information in macroeconomic variables and technical indicators.	A very interesting study. They compared technical indicators with macroeconomic variables. In this regard, they used them separately and then together to see the results. They do not use intelligent models. However, the results are still very interesting. With regard to fundamental variables, they chose the year 2014. They use data spanning 1950:12 to 2011:12 for 14 well-known macroeconomic variables from the literature and 14 common technical indicators.
51	Deepika Chandwani	2014	Stock direction forecasting techniques: an empirical study combining machine learning system with market indicators in the Indian context	International Journal of Computer Applications	(Chand wani & Saluja, 2014)	To encapsulate market indicators with artificial intelligence techniques to create useful extracts to improve decisions under uncertain conditions.	SVM, ANN, GA-SVM, and GA-ANN	Stock price movement	They conclude that GA significantly increases forecasting performance and that technical indicators, compared with fundamental variables, are superior in order to model a market.	This study considers ANN and SVM models with genetic algorithms to forecast an Indian stock index. They use technical indicators and fundamental variables.

52	Kaijian He, Lijun Wang, Yingchao Zou, Kin Keung Lai	2014	Exchange rate forecasting using entropy optimized multivariate wavelet denoising model	Mathematical Problems in Engineering	(He, Wang, Zou, & Lai, 2014)	To check the prediction power of a proposed hybrid model in an exchange rate market.	Entropy optimized multivariate wavelet denoising model	Exchange rate	Empirical outcomes in the Chinese and European stock markets are used as guidelines to show a significant performance increase when the proposed method is tested against conventional models.	They represent multivariate denoising with a wavelet and not an ANN. Moreover, they propose a heterogeneous market hypothesis based on exchange rate modeling methodology to predict a micro market structure. This study shows the power of wavelet analysis confronting noise in a market.
53	Nafiseh Behradmehr1, Mehdi Ahrari2	2015	Forecasting crude oil prices: a hybrid model based on wavelet transforms and an NN	The International Journal of Humanities	(Behradmehr & Ahrari, 2015)	To smooth and minimize the noise presented in crude oil prices and then investigate the effect of wavelet smoothing on oil price forecasting while using the GMDH neural network as the forecasting model.	Wavelet transforms and a neural network	Crude oil prices	They forecast crude oil prices with the usage of the noises effect (applying a wavelet and neural network) to obtain more accurate results, for which they claim a 40% increase in accuracy.	They present wavelet transforms and neural networks for the prediction of crude oil prices.
54	Kristoufek, Ladislav, Vosvrda, Miloslav	2014	Measuring capital market efficiency: long-term memory, fractal dimension, and approximate entropy	The European Physical Journal B	(Kristoufek & Vosvrda, 2014)	To measure capital market efficiency.	Long-term memory, fractal dimension, and approximate entropy	38 stock market indices	They conclude that the most efficient markets are situated in European countries such as France, Germany, and the Netherlands, and that the least efficient markets are in Latin American countries such as Venezuela and Chile.	They propose three different efficiency measurements for capital markets, including long-term memory with a Hurst exponent calculation and approximate entropy. The measurements indicate different efficiency rates among capital markets. Some markets are more efficient, such as the European markets, and some are less efficient, such as the Latin American markets. This result is consistent with the AMH.

55	Mensi, Wali, Beljid, Makra, Managi, Shunsuke	2014	Structural breaks and the time-varying levels of weak-form efficiency in crude oil markets: evidence from the Hurst exponent and Shannon entropy methods	International Economics	(Mensi et al., 2014)	To investigate the time dependency levels of the weak form of efficiency and the existence of structural breaks.	The Hurst exponent and the Shannon entropy approach	Two worldwide crude oil benchmarks	They show that the Hurst exponent performs better than the Shannon entropy model.	Regarding short-term memory and structural breaks, the Hurst exponent is more effective than the Shannon entropy in indicating financial crises and crashes as well as structural breaks and extreme events. The authors suggest using the Hurst exponent for more effective results and achieving true and spurious results.
56	Zhiqiang Guo , Huaiqing Wang, Jie Yang, David Miller	2015	A stock market forecasting model combining two-directional two-dimensional PCA and a radial basis function neural network (RBFNN)	PLoS One	(Guo, Wang, Yang, & Miller, 2015)	To test the predictability of an offered hybrid model in terms of stock market behavior.	Two-directional two-dimensional principal component analysis and RBFNN	Stock price movement	The empirical studies show that the proposed model of (2D2D)PCA-RBFN outperforms the PCA-based approach, as well as alternative methods based on ICA and on the multilayer perceptron approach.	They propose a two-directional two-dimensional PCA and radial basis function NN. They compare this with PCA-ANN and ICA-ANN. They choose 36 stock market technical indicators as the input features. This empirical study is a successful attempt at combining a PCA and ANN.
57	Babić, Jovana Božić,Đorđe	2015	EUR/RSD exchange rate forecasting using a hybrid wavelet-neural model: a case study	Computer Science and Information Systems, ISI	(Babić)	To examine and discuss modeling and forecasting the results of exchange rates.	WNN model	EUR/RSD exchange rate	They report that an increase in the level of resolution in wavelet decomposition does not increase prediction accuracy. The analysis of achieved outcomes indicates that the offered method sufficiently satisfies the characteristics of a financial time series predictor.	This study conerns a wavelet and a neural network; moreover, it represents statistical methods. The authors also make a forecasting program from their study in MATLAB. They use Daubiches and Haar for the wavelet. They claim that the combination of wavelet decomposition and an ANN has promising results and outperforms other models tested in the study.

58	Zhang Chengzhao, Pan Heiping, Zhou Ke	2015	Comparison of BPNNs and EMD-based NNs in forecasting the three major Asian stock markets	Journal of Applied Sciences	(Chengzhao et al., 2015)	To compare the forecasting accuracy of BPNNs and empirical mode decomposition (EMD), based on neural networks.	BPNN and EMD-NN	Stock indices: three major Asian stock markets	They claim that the combination of ANNs and EMD with parallel data input is superior in results for the NIKKEI, KOSPI, and Hang Seng markets.	This study is about ANN and EMD-ANN models and a comparison of the two. It tests the offered models in the NIKKEI, KOSPI, and Hang Seng markets. They use decomposition into intrinsic mode functions, an approach that is interesting to analyze as a methodology..
59	Junghwan Jin, Jinsoo Kim	2015	Forecasting natural gas prices using wavelets, time series, and ANNs	PLoS One	(Jin & Kim, 2015)	To predict natural gas prices more accurately than prior methods.	Wavelet and ANN	Natural gas prices	They show a wavelet as a successful method of denoising in financial time series. They compare their results with ARIMA and GARCH. The results of their proposed method with a wavelet and ANN outperform other models.	They use a wavelet for decomposition and then a neural network. They represent the wavelet as a successful method of denoising in financial time series.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Chapter Overview

This chapter describes the methodology used for the study. First, a brief introduction regarding the modeling of financial markets is given (Section 3.2). Sample data, analytical techniques for the sample data, and relevant results are then presented (Section 3.3). The next sections introduce the most suitable approaches to investigate the long-term memory and entropy of the markets, in terms of predictability indices, in order to address H1 (Sections 3.4 and 3.5). In the following section, the buy-and-hold strategy and the technical trading rules that are employed in this study are presented in order to address H2 and H3 (Section 3.6). After this, the new hybrid method for denoising and forecasting financial markets is presented in detail, with all fundamental, mathematical, and historical support (Section 3.7). Then, the sample design and structure of inputs are expressed as a research framework (Section 3.7.3). This is followed by a comprehensive presentation of all three models of prediction (pure NN, WNN, and WPCA-NN) in the form of the models' architecture, in order to address H4a and H4b (Section 3.7.4). In the next two subsections, the methods to evaluate the performance and profitability of modern trading strategies are described (Sections 3.7.5 and 3.7.6). After this, the hypotheses are reviewed in accordance with the research objectives, literature review, and data analysis (Section 3.8). Finally, a summary of the entire research methodology is given (Section 3.9). Based on the foregoing structure, this study is able to examine long-term memory, approximate entropy, predictability, forecasting performance, trading profitability, and the agreement or contradiction of the final outcome with RWH or AMH in all given futures markets. Table 3.1 presents an overview of the research methodology with the hypotheses and the relevant methods that address them.

Table 3.1: Research overview

	Objectives	Hypotheses	Model/Technique	Sections
1	To investigate whether the futures markets exhibit significant changes in predictability over time, considering long-term memory and approximate entropy.	H1: Futures markets exhibit significant changes in predictability over time, considering long-term memory and approximate entropy.	Hurst exponent Approximate entropy	3.4 and 3.5
2	To determine if the hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the RWH (a passive buy-and-hold strategy) for selected financial markets.	H2: The hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the RWH (a passive buy-and-hold strategy) for selected financial markets, as shown by significant mean differences and significant positive alphas in profit regressions.	Multivariate wavelet denoising using PCA plus forecasting with a NARX-NN (WPCA-NN) model Buy-and-hold strategy benchmark	3.6.1, 3.7.1.2, 3.7.2, and 3.7.4
3	To investigate, test, and establish whether the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other technical analysis indicators (such as optimized moving average).	H3: Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other technical analysis indicators (3a. MACD, 3b. RSI, 3c. stochastics, and 3d. OMA).	Multivariate wavelet denoising using PCA plus forecasting with a NARX-NN (WPCA-NN) model Technical trading rules: MACD, RSI, stochastics, OMA	3.6, 3.7.1.2, 3.7.2, and 3.7.4
4a	To investigate, test, and establish whether the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other pure NN methods of forecasting.	H4a: Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other pure ANN methods of forecasting.	Multivariate wavelet denoising using PCA plus forecasting with a NARX-NN (WPCA-NN) model Nonlinear autoregression with a (NARX-NN) model	3.7.1.2, 3.7.2, and 3.7.4
4b	To investigate, test, and establish whether the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets compared with the WNN.	H4b: Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets compared with the WNN model.	Multivariate wavelet denoising using PCA plus forecasting with a NARX-NN (WPCA-NN) model Univariate wavelet denoising plus forecasting with a NARX-NN (WNN) model	3.7.1.1, 3.7.1.2, 3.7.2, and 3.7.4

3.2 Modeling the Financial Market

Although a significant review of financial market modeling is presented in the prior chapter, a brief introduction is presented here in order to review the modeling context and discuss this study's methodology. Stock market modeling is of significant practical interest. Generally, there are three schools of thought regarding the stock market. The first school believes that there is no hope of achieving more than average trading benefits based on prior and current market data. There are well-known theories supporting this approach such as the RWH and the weak form of EMH. The RWH indicates that stock prices change without any reference to prior prices and move in a purely random and unforeseen manner. However, there is a convincing evidence to reject the RWH and thus encourages researchers to study better market forecasting (Taylor, 1986). For example, the empirical results of the foregoing study provide a contrast to RWH for short-run trading in some emerging and developed markets. According to the literature, a prominent evidence shown that stock returns are not volatile and retain certain elements of predictability, although there is an absence of any strong alternative theoretical clarifications to EMH. However, by making use of an evolutionary approach to economic interaction, it has been shown that the AMH can coexist with the EMH in a logically stable way (Lo, 2004). The AMH accommodates market frictions and states that markets advance over time, a situation that is not the case with the EMH, which considers frictionless markets. Given this perspective, which supports the other two schools of thought, this current study aims to examine whether the AMH is capable of providing a better explanation of futures' stock indices.

The second school of thought involves a fundamental analysis that addresses the financial status of a specific firm and the economic conditions of a stock market. Fundamental data, such as information related to accounting, competition, and management, determine a stock's value. Likewise, stock market indices are established by analyzing the fundamental and economic information associated with markets. Fundamental analysis provides the intrinsic value of a stock or market index and creates a signal to buy if the current value of the stock is less than its intrinsic value. Fundamental analysis of the market is determined by the following ratio: 10% by physiological factors and 90% by logical factors (Chavan & Patil, 2013). The main point about fundamental analysis is that it is useful for long-term trading (Soni, 2011).

The third school of thought involves technical analysis and states that a stock market can be predicted by its trends, which can be captured from historical data. Technical analysts believe in frequent patterns in a stock market that make it predictable. There are tools such as technical indicators, charting trends, and expert techniques. These tools include Fibonacci series, harmonic patterns, and head and shoulders analysis, all of which are used to find patterns. Indicators are built from open, highest, lowest, and close prices or index levels, and sometimes trading volumes. Traders monitor indicator changes and stock index numbers to decide their order.

Many different conventional and intelligent techniques have been used to model stock markets more effectively and achieve greater profits from trading. Traders prefer to apply various techniques in order to obtain more reliable signals. NNs as machine-learning techniques are often applied in order to model markets with technical and fundamental indicators. Moreover, ANNs in single and hybrid forms have been successful in many cases of stock price and stock market index forecasting. We apply a technical approach that uses historical data and technical indicators with a multivariate Wavelet-PCA denoising technique and NN system.

In the next section, we introduce the sample data and present a descriptive analysis of futures market data in order to check the predictability of the time series with statistical tools. After investigating the predictability of the time series of the futures markets, this study will consider and present hypotheses (Table 3.1).

3.3 Sample Data

This study's sample data are collected from Bloomberg and consist of 13 years of historical data from 2002 to 2014, 3,224 daily data items for the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500 and TAIEX futures markets. However, for all tested traditional and modern trading models, the trading period is from 2005 to 2014 (the data framework of the modern techniques are comprehensively explained in Section 3.8). The daily OHLC index and volume, as well as the technical indicators of the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500 and TAIEX futures markets, are collected from January 2, 2005 to December 31, 2014. These markets are analyzed in several studies (Necula, 2009; Huang et al., 2009; Yao et al., 1999; Kim et al., 1998; Lee et al., 2010; Leung et al., 2000; Necula, 2009; Quah & Srinivasan, 1999; Chiang & Doong, 2001; Enke & Thawornwong, 2005) and are the combination of developing and developed country with different economical circumstances, which eventually will give more general and robust results. Moreover, all these markets have been affected by the 2008 financial crisis. Futures markets indices have been chosen for this study because they are traded via most valid brokers and the data is updated and broadly accessible. The daily OHLC data, spot price, and volume, alongside their technical indicators such as RSI, MACD, MACD signal, stochastic fast %K, stochastic slow %K, stochastic %D, and ultimate oscillator, are used as inputs to the intelligent models. Spot price, ultimate oscillator and volume of trading diminished the forecasting performance and were eliminated from the study through a backward technique. These indicators are employed in several studies (Kuo et al., 1996; Chen et al., 2006; Olson & Mossman, 2003; Atsalakis & Valavanis, 2009b; Vanstone & Finnie, 2009; Tsai & Hsiao, 2010; Menkhoff, 2010; Neely et al., 2014; Chandwani & Saluja, 2014; Guo et al., 2015) and tested to be useful in forecasting. In addition to that, these variables are proven to be applicable and useful in our study with backward elimination technique. Since the data must be prepared and checked statistically before any further step, the next section introduces data analysis techniques and interprets their results for futures markets.

3.3.1 Data Analysis Techniques

In this section, we introduce data analysis techniques in order to discuss the descriptive analysis of the sample data. Since this study considers the forecasting of financial time series, we establish descriptive analytical tests suggested by the main studies in this field (Fama, 1991, 1965; Kuo et al., 1996; Makridakis, 1993) in order to confirm the predictability of the future movement of the sample data through the use of historical data.

3.3.1.1 Unit Root Test

In statistics, a unit root test examines whether time series data are stationary or non-stationary by applying an autoregressive method. A generally applied test that is effective in large models is the augmented Dickey–Fuller (ADF) test. Optimum limited sample tests for a unit root in autoregressive models have also been developed (Sargan & Bhargava, 1983). In experimental time series, Sargan–Bhargava statistics analyze the null hypothesis of the unit root in first command autoregressive models in contradiction of one-sided alternatives; for example, if the procedure is stationary or explosive in the alternative hypothesis. Another known test for unit root is the Phillips–Perron (PP) test.

In a stationary time series with a deterministic trend, the effect of shocks is transitory. However, in a time series with a stochastic trend or a unit root, the effect of shocks is permanent. Thus, these are the matters that make a unit root test important: to know whether the effect of shocks is permanent or transitory and whether the time series has an attractor.

Applying a non-stationary time series sample in financial models creates false and unreliable results and leads to misunderstandings and incorrect predictions. The solution to the issue is to convert the time series data so that it becomes stationary. If the non-stationary method is a random walk, it is converted to a stationary procedure by differencing. Additionally, if a time series sample shows a deterministic trend, the false outcomes can be avoided by de-trending. Occasionally, a non-stationary time series may combine a deterministic and a stochastic trend, simultaneously. In order to avoid misleading outcomes, both differencing and de-trending should be used because differencing omits the trend in the variance and de-trending omits the deterministic trend.

A time series Y_t $(t = 1,2...)$ is assumed to be stationary if its statistical parameters do not differ with time (variance, autocorrelation, and expectation). White noise is a sample of a stationary time series, with, for instance, the circumstance where Y_t obeys a normal distribution independent of time. A non-stationary time series can, for instance, be stationary in difference, log, or rate form, also known as integrated to the of order 1, where Y_t is non-stationary, but the $Y_t - Y_{t-1}$ first difference form is considered stationary (Sargan & Bhargava, 1983). A time series can also be stationary in its trend, which is the case for random walk.

Stationarity analyses enable researchers to confirm whether a time series is stationary or not. There are two different methods. Stationarity analyses such as the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test assume as the null hypothesis H0 that the time series is stationary. Unit root tests, such as the PP test or the ADF test, have an opposite null hypothesis such that the time series has a unit root and is thus non-stationary.

In econometrics and statistics, an ADF test analyzes a unit root in time series data. An augmented version of the Dickey–Fuller analysis is also considered for a more complex and larger collection of time series processes. The ADF numbers applied in the test are negative. The more negative a number is, the greater the rejection of the null hypothesis that a unit root exists at some level of confidence.

In addition to the foregoing, the conventional techniques for time series forecasting need a stationary time series because most real-time series are non-stationary. After neural networks had been introduced, traders and researchers could use original time series as prediction targets (Diaconescu, 2008) without any stationarity check. This is because the algorithm behind NNs is constructed on difference forms of data. Moreover, a survey mentions that there is no requirement for any data assumptions or data type for ANNs (Bahrammirzaee, 2010). Although there is no need to make the data stationary in order to feed to an NN (Diaconescu, 2008), we examine predictability and stationarity in the first difference form of the futures markets with a unit root and ADF test (Fama, 1991; Sargan, 1964). Table 3.2 presents the results.

Table 3.2: Results of the Unit Root (ADF) test

Futures markets	Test in the level form	The first difference form
Hang Seng	-2.59191	**-58.37271***
KLCI	-2.303409	**-60.82484***
KOSPI 200	-2.491356	**-55.64235***
NIKKEI 225	-1.312561	**-58.40276***
SiMSCI	-1.889653	**-60.65505***
S&P 500	-1.158461	**-62.90317***
TAIEX	-2.76804	**-53.85111***

*Significance at the 5% level.

The intercept value should be strongly negative to reject the hypothesis of the unit root. The descriptive analysis shows that all seven financial time series are non-stationary in the level form, but stationary in the first differenced form (Table 3.2). As the results of the unit root test in the first difference form are strongly negative, it can be concluded that all the data of futures markets for this study are stationary in the first difference form;

thus, historical data can be used to forecast the future movements of the selected futures markets. Consequently, because the ANN algorithm with the first step of delay $f(P_t, P_{t-1})$ makes the data stationary, we can use the original actual data as input variables without any changes (Diaconescu, 2008).

3.3.1.2 Serial Correlation Test

Serial correlation is the cross-correlation of a time series with itself at different points in time and is also known as autocorrelation or cross-autocorrelation (Zovko, 2008). In other words, Serial correlation is the similarity among observations as a function of the time delay between time series. Thus, it is a mathematical instrument for discovering iterating patterns, such as the presence of a periodic time series complicated by noise, or for recognizing the missing underlying frequency in a time series implied by its compatible frequencies. Serial correlation is regularly applied in signal processing for analyzing series of values or functions, such as time domain signals.

Hence, autocorrelation is the association between a particular variable and itself over different time intervals. Serial correlations or autocorrelations are frequently discovered in iterating patterns when the level of a parameter affects an upcoming level. In finance, technical analysts and chartists apply autocorrelation in order to determine how well the historical price of a stock forecasts the future price. Because technical analysis is founded completely on the movement of a stock's price and the related volume, rather than a company's fundamentals, seeking and evaluating profitable patterns is a necessary factor of the success in applying models.

In statistics and econometrics, the Breusch–Godfrey test (Breusch, 1978; Godfrey, 1978) is used to evaluate the credit of some of the modeling expectations embedded in applying regression-like methods to observed time series. Particularly, it examines for the existence of serial dependence that has not been contained in an offered model configuration and which, if present, would mean that spurious decisions would be drawn from other analyses, or that suboptimal estimates of model configurations are achieved if it is not considered. The regression techniques to which the analysis can be used include cases where delayed values of the dependent variables are applied as independent variables in the techniques' representations of future observations. This type of structure is usual in econometric processes. A similar valuation can also be performed with the application of the Durbin–Watson test. Since the test is founded on the basis of Lagrange multiplier testing, it is occasionally referred to as the LM test for serial correlation. The results of the test used here are in Table 3.3.

Table 3.3: Results of the Breusch–Godfrey serial correlation LM test

Futures markets	Obs R-squared	Prob. Chi-Square (5)
Hang Seng	8.050268*	0.0535
KLCI	2.907285*	0.0713
KOSPI 200	6.049925*	0.0304
NIKKEI 225	3.00968*	0.0695
SiMSCI	8.702567*	0.0125
S&P 500	6.388968*	0.0472
TAIEX	11.51784*	0.0542

*Significance at the 10% level.

In finance, technical analysts use serial correlation in order to distinguish how well the historical data of a security predicts future price movements. Descriptive analysis shows that serial correlation exists between prior and current prices in all selected futures markets at a confidence level of 10% (see Table 3.3). Since autocorrelation exists in futures markets, price movements are correlated to historical data; in other words, they can predict futures movements. These results are consistent with stationary results and confirm the predictability of futures markets. Since autocorrelation is the similarity among observations as a function of the time delay between time series, it can be concluded that all tested futures' time series are predictable by lag in their own time series, which suggests predictability with a similar delay feature as in the ANN algorithm.

Now that market predictability is confirmed through the use of historical data, it is time to investigate whether price movement is predictable through the use of technical indicators. Since we are going to examine the forecasting of a time series (future prices) with another time series (a technical indicator), the Granger causality test (Granger, 2004) is introduced.

3.3.1.3 Granger Causality Test

Although we introduced a test to examine whether the past prices of futures markets predict future prices, we also need to investigate the predictability of future prices with a time series of the relevant technical indicators. In statistics and econometrics, the Granger causality test is a theoretical test for specifying whether one time series is suitable for predicting another. The test was first suggested in 1969 (Granger, 1969). Generally, regressions expose broad correlations, but Granger (1969) debated that causality in econometrics could be analyzed by calculating the capability to forecast the upcoming values of a time series through the application of the prior values of another time series. Since the problem of true causality is genuinely philosophical, and because of the logical fallacy assumption that one factor preceding another can be applied as evidence of causation, econometricians emphasize that the Granger test discovers only predictive causality (Diebold, 1998).

A time series C is supposed to Granger-cause E if it can be indicated, generally within a series of F-tests and t-tests on delayed amounts of C (and with delayed amounts of E also involved), that the C amounts deliver statistically meaningful knowledge regarding future amounts of E (Granger, 2004).

Additionally, Granger stressed that some researches using Granger causality that analyze in fields outside economics come to absurd conclusions (Granger, 2004). However, the causality test remains in widespread use in time series because of its computational easiness (Eichler, 2012; Seth, 2007). The unique meaning of the Granger causality test is not considered in terms of hidden confounding impacts and does not have instantaneous and non-linear causal associations, although many extensions have been suggested to address these problems (Eichler, 2012).

Once a time series C Granger-causes time series E, the trends in C are almost iterated in E after some time delay (Granger, 2004). Hence, historical amounts of C can be used for the forecasting of upcoming amounts of E. Further, a variable C that changes over time Granger-causes another changing variable E if forecasts of the value of E founded on its own historical amounts and on the historical amounts of C are superior to forecasts of E based only on its own historical amounts. Thus, we employ the Granger causality test for the relationship between time series and their relevant technical indicators to investigate whether such indicators Granger-cause the relevant futures markets. In another word, we examine whether futures markets are predictable through the use of the selected technical indicators as inputs of the models. Table 3.4 presents the results of the Granger causality test.

A time series, x, is said to Granger-cause another time series, y, when past values of x assist the prediction of a recent level of y, given all other proper information. From the results of the Granger causality test (Table 3.4), all selected variables are confirmed to predict the future prices of closing indices through the use of their historical data. All results are significant at the 5% level. Thus, the examined technical indicators, RSI, MACD, signal, fast %K, fast %D, slow %D, and ultimate oscillator are able to forecast the price movement of the relevant futures markets. This finding confirms the usage of selected technical indicators as input variables into the forecasting models.

Table 3.4: Results of the Granger Causality Test for all markets and all variables

Variables	Hang Seng	KLCI	KSOPI	NIKKEI	SiMSCI	S&P 500	TAIEX
RSI	1.46579	6.27937	2.3908	6.7632	2.8292	1.07515	3.4124
MACD	6.02289	2.42921	3.421	3.55999	3.76423	3.85602	1.30403
Signal	4.51849	1.13104	4.87946	2.7193	1.79195	2.62161	1.39064
Fast %K	4.2437	15.4101	1.98852	1.86626	1.99723	1.91742	2.24604
Fast %D	3.68374	11.2376	1.51615	1.90635	1.63864	1.37864	2.10343
Slow %D	2.83582	6.7786	3.75203	1.09156	2.27504	5.87188	1.588
Ult. Osc.	0.56277	3.3664	1.51023	2.20122	1.26772	2.69577	1.03061

Note: Significance is at the 5% level.

3.3.2 Data Analysis

According to the unit root test, the intercept values of the first difference forms for the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets are respectively -58.37271, -60.82484, -55.64235, -58.40276, -60.65505, -62.90317, and -53.85111. These are significantly (at a 5% confidence level) and strongly negative. Thus, all data of the futures markets are stationary in the first difference form. Consequently, historical data can be employed to forecast the price movements of the selected futures markets. Moreover, we can employ the original actual data as input variables without any modification (Diaconescu, 2008) into ANN, since the ANN algorithm with the first step of delay $f(P_t, P_{t-1})$ is able to make the data stationary. Additionally, the Breusch–Godfrey serial correlation test shows that autocorrelation exists between the prior and current prices in the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets with R-squared values of 8.050268, 2.907285, 6.049925, 3.00968, 8.702567, 6.388968, and 11.51784 respectively (at a confidence level of 10%). In other words, historical data can predict futures' movements in the analyzed futures markets. These outcomes are consistent with the stationary results and confirm the predictability of futures markets. Hence, all tested time series of futures markets are predictable by lag in their own time series, which offers predictability with a similar delay feature as that in the ANN algorithm. According to Table 3.4 and the results of the Granger causality test, RSI, MACD, signal, fast %K, fast %D, slow %D, and ultimate oscillator are confirmed to predict the future prices of closing indices through the use of their historical data. This confirms the usage of selected technical indicators as input variables into the forecasting models. In addition to these tests, a cointegration test and error correction model are also used for further analysis of the time series of futures markets. These are presented in Appendix A.

In conclusion, according to the descriptive analysis and predictability of the sample data, all tests and models considered in the hypotheses are valid for use with the sample data. Moreover, the variables or technical indicators are selected properly and are able to predict futures' movements in the markets. The following sections describe the analyses and techniques used to examine the hypotheses.

3.4 Long-Term Memory

Investigating the presence of long-term memory is a suitable indicator of the predictability of financial time series (Mandelbrot & Wallis, 1969a). The Hurst exponent is applied as a calculation of the long-term memory of time series. It is associated with the serial correlations of the time series and the speed at which these fall as the delays between pairs of values rise. Studies that included the Hurst exponent were initially developed in hydrology for the practical problem of deriving optimal dam sizing for the Nile river's noisy rain and dryness conditions, which had been observed over a long period (Hurst, 1951). The term "Hurst coefficient," or "Hurst exponent," comes from Harold Edwin Hurst, who was the main scholar involved with these findings. The use of the notation "H" for the exponent is also associated with his name.

In fractal geometry, the universal Hurst coefficient has been designated by H or H_q in honor of both Ludwig Otto Hölder and Harold Edwin Hurst (Mandelbrot & Wallis, 1968).

H is associated with fractal dimension, D, and is a calculation of a time series' mild or wild uncertainty (Benoit & Hudson, 2004).

The Hurst coefficient is interpreted as the indicator of long-term dependence or index of dependence. It computes the relative tendency of a data series either to regress powerfully to a cluster or to a mean (Kleinow, 2002). An amount of H in the range 0.5–1 shows a data series with long-term positive autocorrelation, indicating that a large amount in the time series will possibly be followed by another large amount and that the amounts will also likely be large a long time into the future. An amount in the range 0–0.5 shows a data series with long-term change between low and high amounts in neighboring pairs, indicating that a single high amount will possibly be followed by a low amount and that the amount after that will likely be large, with this tendency to change between low and high amounts remaining for a long period into the future. An amount of H = 0.5 can represent an entirely uncorrelated time series; however, it is the usable amount for a time series for which the serial correlations or autocorrelations at short time delays can be negative or positive, but where the absolute amounts of the autocorrelations decline exponentially and rapidly to zero. This contradicts the normal power-law decline for the $0 < H < 0.5$ and $0.5 < H < 1$ cases.

The Hurst coefficient, H, is identified in terms of the asymptotic performance of the rescaled range as a function of the time period of a time series as follows (Rasheed & Qian, 2004).

$$E\left[\frac{R(n)}{S(n)}\right] = Cn^{H} \; as \; n \rightarrow \infty, \tag{4}$$

where

$E[x]$ is the expected value,

$R(n)$ is the range of the initial n values,

$S(n)$ is their standard deviation,

C is a constant, and

n is the number of data points in a data series.

Although there is enough support (Arouri et al., 2012; Kirkulak & Lkhamazhapov, 2014; Mensi et al., 2014; Ohanissian et al., 2008) in the literature (Section 2.4) for true long-term memory and the Hurst exponent measurement, we investigate structural breaks during the analyzed period of the futures markets. We study each year of the futures markets separately and examine whether periods that only include structural breaks show long-term memory, or periods without any structural breaks also display long-term memory. Using this approach, it is possible to distinguish the effect of short memory or structural breaks on the long-term memory of the entire market. Thus, the Bai–Perron tests are applied, using eviews software, for L globally optimized breaks, in contrast with the null hypothesis of no structural breaks.

The long-term memory approach makes it possible to observe changes in stock market predictability and efficiency. Since the data for the long-term memory process should be stationary, the first difference form of price is used in this test. The Hurst exponent is measured for all the given futures markets once, annually, from 2005 to 2014 as a predictability index (Kristoufek & Vosvrda, 2014) and once for the entire period of 2005 to 2014. This approach provides the predictability or inefficiency of the futures markets, together with the variation of the index (if any) and consequently its predictability from 2005 to 2014. When H is higher than 0.5 and less than 1, and a variation of the Hurst exponent is present over time in any market, it is possible to observe predictability and time-varying efficiency (Kristoufek, 2012; Kristoufek & Vosvrda, 2014; Mandelbrot & Hudson, 2005). This suggests that the market is supported by the AMH (Lo, 2004; Noda, 2012; Urquhart & Hudson, 2013; Urquhart & McGroarty, 2014). This support is a required basis in order to use prediction tools such as an ANN in the financial markets. The outcome and conclusions based on the Hurst exponent and approximate entropy (described in the following Section) form hypothesis H1.

3.5 Approximate Entropy

In addition to long-term memory, entropy is introduced as another index for market predictability (Kristoufek & Vosvrda, 2014; Pincus & Kalman, 2004). Entropy can be assumed as a calculation of a structure's complexity. Structures with high entropy can be considered as having no information and are consequently random; however, structures with low entropy can be considered deterministic (Pincus & Kalman, 2004). An efficient market is then characterized as a structure with maximum entropy. Further, the higher the entropy, the more efficient the market. For the purpose of a predictability measure, we require an entropy calculation that is bounded. Hence, I employ approximate entropy (Pincus, 1991). This approximate entropy is used by several studies (Kristoufek & Vosvrda, 2013, 2014; Zunino et al., 2010).

For each i in $1 \leq i \leq T - m + 1$, there is the following definition.

$$C_i^m(r) = \frac{\sum_{j=1}^{T-m+1} I_{d[i,j] \leq r}}{T-m+1}, \tag{5}$$

where $I_{d[i,j] \leq r}$ is equal to 1 and a binary indicator function if the condition in $I_{d[i,j] \leq r}$ is satisfied; otherwise, it is equal to 0 where

$$d[i,j] = \max_{k=1,2,\dots,m}\left(\left|x_{i+k-1} - u_{j+k-1}\right|\right). \tag{6}$$

$C_i^m(r)$ can thus be considered an amount of autocorrelation because it is based on a maximum gap between delayed series. Averaging $C_i^m(r)$ across *i* returns gives

$$C^m(r) = \frac{1}{T-m+1} \sum_{i=1}^{T-m+1} C_i^m(r), \tag{7}$$

which is associated with the correlation dimension

$$\beta_m = \lim_{r \to 0} \lim_{T \to +\infty} \frac{\log C^m(r)}{\log r}, \tag{8}$$

which in turn is assumed to be a calculation of complexity and entropy of the series (Pincus, 1991). β_m ranges between 0 and 1, which demonstrate completely deterministic ($\beta_m = 0$) and completely random ($\beta_m = 1$), respectively. Not only can the futures markets for market efficiency be ranked based on approximate entropy; it is also possible to capture changes in the markets' efficiency over time. This approach makes it possible to observe changes in stock market predictability and efficiency. The same data for the long-term memory process, which is the first difference form of price, is used in this analysis. We calculate the approximate entropy for all given futures markets from 2005 to 2014 as a predictability index. First, we establish the predictability or inefficiency of the futures markets together with the variation of the index from 2005 to 2014. If entropy is higher than 0 and lower than 1, and variation of approximate entropy exists over time in any market, there are statements of predictability and time-varying efficiency (Darbellay & Wuertz, 2000; Kristoufek & Vosvrda, 2014; Matesanz & Ortega, 2008; Mensi et al., 2014; Pincus, 1991). These observations show that price movements in futures market are supported by the AMH (Lo, 2004; Noda, 2012; Urquhart & Hudson, 2013; Urquhart & McGroarty, 2014). If the markets follow the AMH, then they can be predicted. Technical analysis is one of the schools of thought used to predict the markets' future prices. In the next section, the technical trading strategies that are used in this study are presented.

3.6 Technical Trading Rules

In this section, the buy-and-hold strategy and the chosen technical trading strategies that are based on the literature review and the best-performing strategies in the literature

(Chong et al., 2010; Chong & Ng, 2008; Fernández-Blanco et al., 2008; Rosillo et al., 2013) are introduced. The way in which the strategies are applied in this study to trading in the futures markets from 2005 to 2014 is also described. The results of the buy-and-hold strategy and the other chosen technical trading strategies, RSI, MACD, stochastics, and OMA, are compared with the novel hybrid model in the context of H2 and H3.

3.6.1 The Buy-and-Hold Benchmark

As mentioned previously, the buy-and-hold strategy is a passive trading or investment strategy that is considered the criterion of all trading rules in the market. The buy-and-hold strategy involves buying futures contracts at the beginning of a period and holding them until the end of the period. The return of a buy-and-hold strategy during a given period is calculated as follows:

$$Return = \frac{P_{end} - P_1}{P_1} \quad . \tag{9}$$

3.6.2 The RSI Trading Rule

When the RSI indicator falls below a value of 30, it indicates an oversold situation. A buy signal is triggered when the indicator line exceeds 30. An RSI value higher than 70 indicates an overbought situation. A sell signal is triggered when the indicator line falls below 70. The calculation for RSI is as follows.

$$RSI = 100 - \frac{100}{1 + \frac{\sum Gains over the past\ 14\ periods}{\sum Losses over the past\ 14\ periods}} \quad . \tag{10}$$

3.6.3 The MACD Trading Rule

When an MACD signal crosses over an MACD line, this is a signal to buy long; and when an MACD line crosses over an MACD signal, this is a signal to buy short. The calculation for an MACD is as follows:

$$MACD = (Averageofthepast\ 12\ periods - Averageofthepast\ 26\ periods), \quad (11)$$

$$\text{MACD Signal} = \text{Average of the MACD of the past 9 periods.} \quad (12)$$

3.6.4 The Stochastic Trading Rule

Stochastic K and D indicators that fall below a value of 20 indicate an oversold situation. A buy signal is triggered when one of the stochastic lines exceeds 20. Similarly, a stochastic value greater than 80 indicates an overbought situation. A sell signal is triggered when one of the stochastic lines falls below 80. The calculation for a stochastic indicator is as follows.

$$Stochastic\ Fast\ \%K = 100 * \frac{C_t - Thelowestlowofthepast\ 14\ periods}{ThehighestHofthepast\ 14\ periods - ThelowestLofthepast\ 14\ periods}, \quad (13)$$

$$Stochastic\ \%D = AverageoftheSlow\ \%Kofthepast\ 3\ periods,$$

where,

$$H = Highestindexintheperiod,$$

$$L = Lowestindexintheperiod,$$

$$C = Closingindexintheperiod.$$

3.6.5 The OMA Trading Rule

When an OMA line crosses under the price line, this is the signal to buy long, and when an OMA line crosses over the price line, this is the signal to buy short. The best parameter for an SMA for last three years, in-sample data, will be used for next three months, out-sample data, this method is so-called OMA. The SMA is suggested by other research (Jorion, 1995). The calculation for an SMA is as follows.

$$SMA_{t+1} = \left(\frac{1}{n}\right)(C_t + C_{t-1} + \cdots C_{t-n+1}). \quad (14)$$

3.6.6 The Return of Trading Rules

In order to employ the introduced technical indicators, a simple buy-and-hold or sell-and-hold trading strategy is performed. With regard to the selected technical trading rules, a long and hold approach is followed when the signal is to buy, a short and hold approach is taken when the signal is to sell. For each futures market, this study runs the strategy annually and calculates the return. The return of a trading strategy is widely considered a model's profitability performance (Enke & Thawornwong, 2005; Leitch & Tanner, 1991; Yao et al., 1999). Hence, the summed return of this trading rule can be calculated by the following equation and used as a comparison scale among the models and markets.

$$Return\ (\%) = 100 \times \left(\sum_{t=1}^{b}\left(\frac{y_{t+1}-y_t}{y_t}\right) + \sum_{t=1}^{s}\left(\frac{y_t-y_{t+1}}{y_t}\right)\right), \tag{15}$$

where b denotes the total number of days for buying futures and s represents the total number of days for selling futures. The trading strategies for all trading rules and for all futures markets are undertaken in two ways: without transaction costs and with transaction costs. The given procedure is considered without transaction costs. In accordance with a trading broker chosen for this study, Interactive Brokers LLC (LLC, 2015), this study runs live trading strategies on futures with offered fixed-commission rates; Hang Seng (0.08%), KOSPI 200 (0.08%), KLCI (0.08), NIKKEI 225 (0.05%), S&P 500 (0.02%), SiMSCI (0.08%) and TAIEX (0.08%). The initial capital is 1 million USD. The trading analyses are undertaken with Bloomberg Terminal. The results with and without transaction costs are displayed in Chapter 4. The return outcome of the buy-and-hold and other aforementioned technical trading strategies are compared with the novel WPCA-NN model in the form of hypotheses H2 and H3.

3.7 Developing a Modern Forecasting Technique

So far, this study has considered all the traditional and conventional methods that are used for the research. In this section, all the components and methods used to create the novel hybrid model of WPCA-NN are introduced. This section also covers the methodology of the other two modern models of pure NN and WNN, which are then compared with the WPCA-NN in the form of hypotheses H4a and H4b.

According to the literature (Section 2.9), ANNs and wavelet transforms have been successful in many cases of financial market forecasting in both single and hybrid forms (Bahrammirzaee, 2010). This current study first denoises the historical data of OHLC and the technical indicators by using a multivariate wavelet-PCA denoising technique. It then applies the resultant series in a NARX-NN. Thus, the actual data is first denoised by a wavelet PCA transform, then, a denoised signal is fed to an ANN. Consequently, we display the denoising process first and the ANN procedure afterward. The introduced model is a WPCA-NN, which is considered a novel hybrid model and is compared with the forecasting power of its constructive elements. The first element is a NARX-NN, which is the forecasting core of the hybrid without any denoising process. This element is the so-called pure NN model. The second element is a univariate wavelet denoising process, which is used with the NARX-NN for forecasting. This element is the so-called WNN model. The difference between the second element and the novel model (WPCA-NN) lies in the employment of PCA, which causes a search for the same noise in the OHLC signal. A comparison between the WPCA-NN model with the pure NN model addresses hypothesis H4b. A comparison between the WPCA-NN model and the WNN model addresses hypothesis H4a. The following sections introduce these models and their backgrounds.

3.7.1 Wavelet Principal Component Analysis Denoising

The fundamental goal of denoising is to remove noise while maintaining the main features of data. In recent years, the wavelet denoising technique has outperformed many traditional methods such as the exponential smoothing filter, the moving average filter, simple nonlinear noise reduction, and linear Fourier smoothing. The reason is that the wavelet denoising technique does not consider homogenous error structures and generates more accurate information in the denoised time series with respect to the original signal compared with other signal analyses (He et al., 2014). Hence, wavelet denoising algorithms have become a highly popular technique for single-dimensional signal filtering and mining.

The PCA technique is extensively applied in statistics, signal processing, and neural computing (Karhunen & Joutsensalo, 1995) as a competent feature-extraction tool. The basic concept of PCA is to discover the components that describe the maximum value of variance obtainable from a data vector with L dimensions and P linearly transformed components, using the mathematical technique of eigen analysis. The essential goal of PCA is to reduce the dimensions of the data. It can be demonstrated that the treatment given by PCA is an optimal method for decreasing linear dimensionality in a mean-square evaluation (Karhunen & Joutsensalo, 1995). Such a diminution in dimension has significant benefits. First, the computation required in further processing is decreased. Second, noise can be deducted and the significant underlying function identified. PCA can also simplify multi-scale signals by tracing new factors obtained from the main features of data (Bakshi, 1999). Univariate, multiple one-dimensional, and multivariate wavelet denoising are described in the following paragraphs (Aminghafari et al., 2006). The relevant procedures are also provided in the MATLAB library (Appendix B) as a coded function named “wmulden.” (MathWorks, 2014c).

The simplest classical univariate wavelet denoising model has the following form.

$$X(t) = f(t) + \varepsilon(t), t = 1, \dots, n, \tag{16}$$

where,

$\left(X(t)\right)_{1 \le t \le n}$ is the observed signal,

$\left(\varepsilon(t)\right)_{1 \le t \le n}$ is the centered Gaussian white noise of unknown variance σ^2, and

$f \in L^2$ is the unknown function to be recovered from the observation in accordance with a given orthogonal wavelet transform $((\phi_{j,k})_{k \in Z}, (\psi_{j,k})_{1 \le j \le J, k \in Z})$, where ϕ is the associated scaling function, ψ is a mother wavelet, and *J* is an appropriately selected decomposition level and where *j*, *k* and *x* are in a condition which $g_{j,k}(x) = 2^{-\frac{j}{2}} g\left(2^{-j}x - k\right)$, wavelet denoising is applied in the following three stages.

Stage 1. Decompose the observed signal by wavelet up to level *J*.

Stage 2. Create suitable thresholds for the wavelet detail coefficients.

Stage 3. Rebuild a denoised form of the initial signal, from the thresholded detail coefficients and the estimated coefficients, and then apply the inverse form of wavelet transform.

3.7.1.1 Multiple Univariate Wavelet Denoising

The first denoising procedure of this study is a direct generalization of the one-dimensional technique. The technique uses a modification of the procedure followed by a standard one-dimensional soft-thresholding approach. In this regard, consider the following *p*-dimensional model.

$$X(t) = f(t) + \varepsilon(t), t = 1, \dots, n, \tag{17}$$

where $X(t)$, $f(t)$, and $\varepsilon(t)$ are as previously defined in formula number 16, and of size $1 \times p$. In this equation, $\varepsilon(t)$ is a centered Gaussian white noise function with an unknown covariance matrix, $E\left(\varepsilon(t)^T \varepsilon(t)\right) = \Sigma_\varepsilon$. Each element of $X(t)$ is of the previously mentioned form (17). Moreover,

$$X^i(t) = f^i(t) + \varepsilon^i(t), t = 1, \ldots, n, \tag{18}$$

where $1 \leq i \leq p$ and f^i is a functional space such as L^2. Then, Σ_ε, as a covariance matrix that is assumed to be positive and definite, obtains the stochastic relationships among the elements of $X(t)$.

The following stages describe multiple one-dimensional denoising, with p original signals (the column of $X(t)$) with *n* dimensions presented as an $n \times p$ matrix, *X*.

Stage 1: Execute the wavelet decomposition at level *J* per signal as each column of X.

Stage 2: State $\widehat{\Sigma_\varepsilon}$ as an estimator of Σ_ε and then calculate a matrix, V, such that $\widehat{\Sigma_\varepsilon} = V \Lambda V^T$, where $\Lambda = diag(\lambda_i, 1 \leq i \leq p)$. Change the basis $D_j V, 1 \leq j \leq J$, and then apply the p univariate threshold strategies using the threshold $t_i = \sqrt{2\lambda_i \log(n)})$ for the i-th column of $D_j V$ for each detail. In addition, a list of strategies can be applied as thresholds at this stage (Donoho, 1995).

Stage 3: Rebuild the denoised matrix $\check{X}$ by inverting the wavelet transform from the estimation matrices and the simplified details.

Thus, this direct generalization and parallelization of the univariate wavelet denoising over time and space (a change of basis) acts first to change the basis in order to reduce the correlations among the p signals and, second, to employ p univariate wavelet denoising.

3.7.1.2 Multivariate Denoising Using Wavelet and Principal Component Analysis (WPCA)

A prior study (Aminghafari et al., 2006) employs multi-scale PCA denoising (Bakshi, 1999) in order to develop a generalized multivariate wavelet denoising approach. The introduction of a PCA stage can take advantage of the deterministic links among the signals, offering an extra layer of denoising by omitting insignificant principal components. The multiple univariate denoising previously discussed can be generalized by focusing on the deterministic links among the p signals.

A natural way to take the deterministic links among the p signals into account is first to use a threshold strategy, including a change of basis, employing V for the details. Then a PCA is applied by choosing the appropriate number of elements for the approximation. More accurately, the following is the general procedure for multivariate denoising.

Stage 1: Apply the level J wavelet decomposition of each column of X.

Stage 2: State $\widehat{\Sigma_\varepsilon}$, the estimator of Σ_ε as the noise covariance matrix. The figure is equivalent to the minimum covariance determinant (MCD) estimator (Rousseeuw, 1984) applied to D_1. Then calculate matrix V in such a way that $\widehat{\Sigma_\varepsilon} = V\Lambda V^T$, where $\Lambda = diag\ (\lambda_i, 1 \leq i \leq p)$. Change the basis $D_jV, 1 \leq j \leq J$, and then perform the p univariate thresholding strategies, employing a threshold such as $t_i = \sqrt{2\lambda_i \log(n)}$ to each element of the i-th column of D_jV. Moreover, a list of strategies can be applied as thresholds at this stage as follows (Donoho, 1995).

Stage 3: Apply the PCA of the matrix A_J and then choose the convenient number p_{J+1} of the principal components.

Stage 4. Rebuild the denoised matrix $\check{X}$ by inverting the wavelet transform from the estimation matrices and the simplified details.

In order to select those components with matching eigen values larger than the mean of all the eigen values, the Kaiser criterion can be employed (Karlis et al., 2003). Further, some variables and settings are required for the wavelet and PCA techniques, such as wavelet type, level of denoising, thresholding strategies, and the selection of the number of principal components. These variables and settings are discussed in Section 3.7.4.

3.7.2 A Nonlinear Autoregressive Neural Network with Exogenous Inputs (NARX-NN)

NNs are employed because of their advantages, such as their numeric nature, the absence of data distribution assumptions, the ability to insert new data and update inputs into a trained network, and their free model estimator nature (Bahrammirzaee, 2010; Chen & Leung, 2004; Garliauskas, 1999; Hassan et al., 2007; Hsieh et al., 2011; Ni & Yin, 2009).

An ANN is a set of interconnected simple processing factors. Each connection of the neural network has a weight attached to it. The FBNN algorithm appears to be one of the most broadly used machine-learning techniques for multilayer networks (McClelland et al., 1986). The standard FBNN generally contains an input layer, several hidden layers, and an output layer, as displayed in Figure 3.1. The elements in the network are linked in a feedforward style. The weights of the links have been set as initial values. The error term between the actual value and the predicted output value is backpropagated across the training network in order to revise the weights and consequently minimize the error between the predicted output and the actual value.

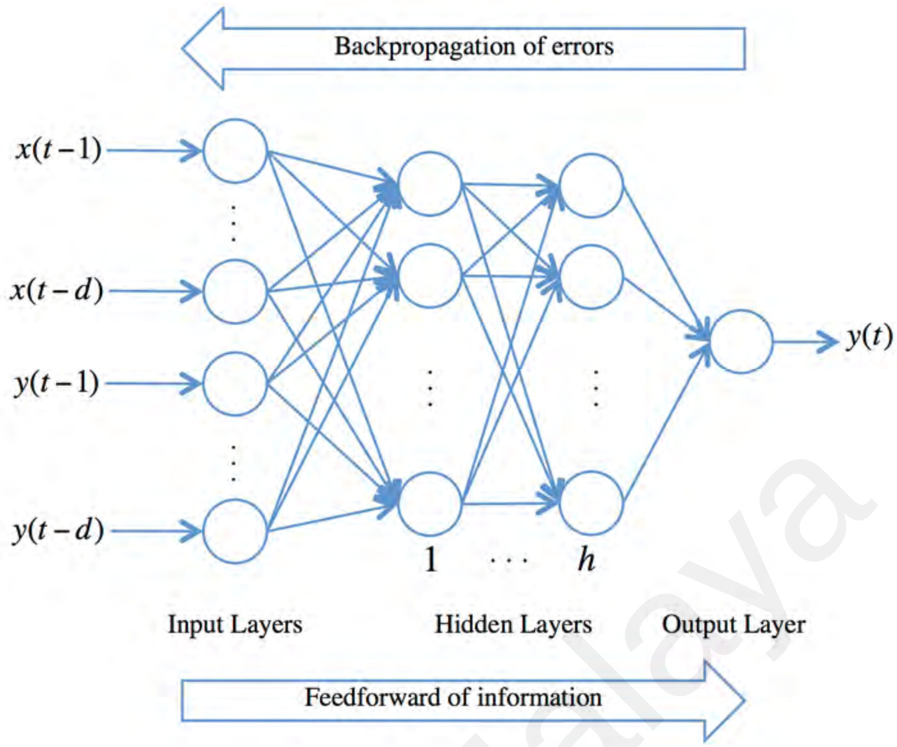


Figure 3.1: Feedforward backpropagation neural network architecture

A NARX-NN is a type of recurrent dynamic neural network with feedback links connecting some layers of the network (Siegelmann et al., 1997). The NARX model is generally applied in time series modeling and is built on the linear auto regressive exogenous (ARX) method.

The fundamental equation for the NARX model is as follows.

$$y(t) = f\big(y(t-1), y(t-2), \ldots, y(t-d), x(t-1), x(t-2), \ldots, x(t-d)\big), \quad (19)$$

where the obtained value of the dependent output signal $y(t)$ is regressed on d former values of the target signal $y(t)$ and d prior values of exogenous (independent) input signals $x(t)$. One can use the NARX-NN model by applying an FBNN to estimate the function f. Moreover, weights and biases in an FBNN are adjusted continuously to minimize the error term between output ($y^*(t+1)$) and target value ($y\ (t+1)$) in order to achieve the lowest mean of the error terms.

The NARX network has many applications, one of the more important being the modeling of nonlinear dynamic systems. An NN that is considered a learning-machine system applies input series and output series of *d* prior data. Such an NN predicts the next output and trains the network by making a comparison between the predicted output and the actual data of the time in question. This procedure is continuously performed, step by step and along the time series, in order to achieve the lowest mean error between the network output and the target.

In this study, a clear and efficiently coded tool in MATLAB, named "ntstool" (MathWorks, 2014b), is used to establish a one-step-ahead prediction model (Appendix B). The architecture of a NARX network includes the number of hidden layers, the number of delays (the amount of past data of that network that accounts for training), and portions of training, validation, and testing. NARX networks divide the data into three subsets: training, validation, and testing. These subsets are spread randomly along the time series, with a configured percentage for each of them. In this study, training is 70%, validation is 15%, and testing is 15%. Although the best architecture to apply depends on the type of problem to be solved by the network, there is no guidance for selecting the number of hidden layers and delays (Pesaran & Timmermann, 1994; Zurada, 1992). In this study, Levenberg–Marquardt optimization is used as the training algorithm. This optimization is a built-in algorithm in MATLAB (MathWorks, 2014a). Levenberg–Marquardt optimization is applied to solve non-linear least squares problems.

3.7.3 Research Framework

This study attempts to predict futures prices on the basis of daily historical prices along with their technical indicators: RSI, MACD, MACD signal, stochastic fast %K, stochastic slow %K, stochastic %D, and ultimate oscillator. The main aspect of this study is to examine the performance of a novel model named the WPCA-NN in the context of the futures markets of Hong Kong's Hang Seng index, Malaysia's KLCI futures, Japan's index (NIKKEI 225), Singapore's index (SiMSCI), South Korea's index (KOSPI 200), the S&P 500, and Taiwan's index (TAIEX). The results of the WPCA-NN model are evaluated with those of NN and WNN models and a threshold, passive, buy-and-hold strategy across these East Asian futures markets and the S&P 500 in order to check signal accuracy and trading profitability performance. The contribution of this study is the extraction of the best combination settings within the WPCA-NN model in each of these futures markets for further trading purposes.

3.7.3.1 Data Framework

The sample data for the study of the WPCA-NN, pure NN, and WNN models consist of 13 years of historical data and seven technical indicators (3,224 daily data items for each market). The daily OHLC, spot prices, and volumes of the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets are collected from January 2, 2005 to December 31, 2014 via Bloomberg terminal, which will be discussed more in the Section 3.8.1. The daily OHLC data and spot prices alongside their technical indicators, RSI, MACD, MACD signal, stochastic fast %K, stochastic slow %K, stochastic %D, and ultimate oscillator, are used as inputs for the models.

The research method uses the past three years' worth of data (in-sample data) of daily prices and technical indicators to forecast the following three months' (out-sample data) daily closing prices, as illustrated in Figure 3.2. This period of three years includes 70% training, 15% validation, and 15% testing. More years of training, validation, and testing, and a higher proportion of training, may cause overfitting problems—the memorizing of patterns by networks—thus reducing the generalizability of the models (Hsieh et al., 2011). The best-performing trained network settings of the in-sample data (the past three years) apply to the following out-sample data (three months or a quarter). The following period of three months (the out-sample) is considered the evaluation period, where each of the three techniques' (NN, WNN, and WPCA-NN) performances in terms of the mean absolute percentage error (MAPE) and profitability of trading strategies (Return) are measured each quarter and compared against the buy-and-hold strategy, selected technical trading rules, and each other. In accordance with popular quarterly portfolio management practice (Bahrammirzaee, 2010; Campbell et al. 2009; Jegadeesh & Titman, 1993; Pesaran & Timmermann, 1994; Sridharan, 2015), quarterly evaluation is performed in this study. This process continues for 10 years with each quarter in order to find the performance of the mean absolute percentage error (MAPE) and the returns of the proposed techniques. Thus, 40 quarters of seven future markets from 2005 to 2014 are studied in order to measure performance and compare robustness, as well as to ensure the generalizability and practicality of the method.

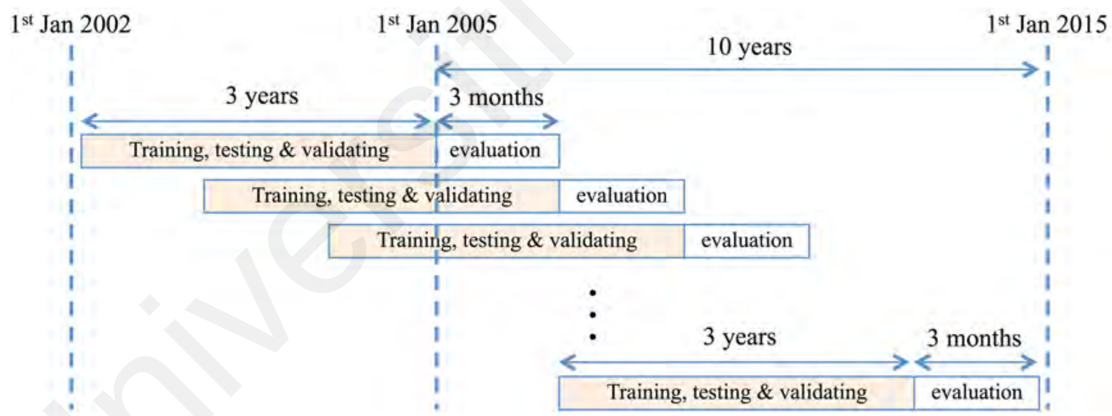


Figure 3.2. The arrangement of the continuous datasets for training and evaluation, 2002–2014

In the next figure, Figure 3.3, the arrangement of the continuous datasets for training and evaluation is illustrated for the Hang Seng futures market as an example of the data framework. This arrangement is exactly what is proposed in Figure 3.2 and is specifically displayed for Hang Seng futures, because these are the first futures to be studied. The figure starts with the chart of the first in-sample period (January 2002 to December 2004) and out-sample period (the first quarter in 2005, January, February, and March) and ends with the chart of the evaluation period of the last quarter in 2014 (October, November, and December). Figure 3.3 consists of 40 quarters, which are for evaluation, and the three years prior to each quarter, which are for training, validation, and testing. The duration of all charts are the same, from 2002 to 2014, and the in-sample and out-sample frames only move toward the end in order to display the research data framework. Using the approach of separating in-sample data and out-sample data, we eliminate the chance of data mining in the analysis (Neely & Weller, 2011).

The evaluation period of this study is three months because the trading investment community and fund managers usually evaluate their portfolio performances "quarterly" and then change their trading parameters for the next quarter. This study follows the same process. Some other studies also suggest quarterly evaluation and changing the trading settings and parameters quarterly (Jegadeesh & Titman, 1993; Sridharan, 2015; Campbell et al., 2009; Pesaran & Timmermann, 1994).

This study also uses a method to address missing data (Wang & Gupta, 2013). When data are missing for some days in the original time series because of public holidays, an average of the past five days is employed for the missing data point. The calculation is as follows.

$$x_t = \frac{x_{t-1}+x_{t-2}+x_{t-3}+x_{t-4}+x_{t-5}}{5}. \qquad (20)$$

Figure 3.3: The continuous datasets of training and evaluation for the Hang Seng futures market, from 2002 to 2014

3.7.3.2 Model Inputs

A prior survey of forecasting approaches indicates that technical analysts typically use indicators to forecast future prices (Atsalakis & Valavanis, 2009b). According to the authors' study and literature review, the key types of technical indicator used to forecast financial time series are spot price, RSI, MACD, MACD signal (MACDSig), MACD histogram (MACDHis), stochastics (fast %K, slow %K, and % D), ultimate oscillator, and volume. These indicators are derived from the OHLC and trading volumes of the futures' prices as follows.

$$RSI = 100 - \frac{100}{1+\frac{\sum Gainsoverthepast\ 14\ periods}{\sum Lossesoverthepast\ 14\ periods}}, \tag{21}$$

$$MACD\ = (Average of the past\ 12\ periods - Average of the past\ 26\ periods), \tag{22}$$

$$\text{MACD Signal} = \text{Average of the MACD of the past 8 periods}, \tag{23}$$

$$MACD\ Histogram = MACD - Signal, \tag{24}$$

$$Stochastic\ Fast\ \%K = 100 * \frac{C_t - The lowest low of the past\ 14\ periods}{The highest H of the past\ 14\ periods - The lowest L of the past\ 14\ periods}, \tag{25}$$

$$Stochastic\ Slow\ \%K = Average of the Fast\ \%K of the past\ 3\ periods, \tag{26}$$

$$Stochastic\ \%D = Average of the Slow\ \%K of the past\ 3\ periods, \tag{27}$$

$$UltimateOscillator = 100 * (4 * \left(\frac{BP(7)}{TR(7)}\right) + 2 * \left(\frac{BP(14)}{TR(14)}\right) + \frac{BP(28)}{TR(28)})/7, \tag{28}$$

where

$H\ = Highestindexintheperiod,$

$L\ = Lowestindexintheperiod,$

$C\ = Closingindexintheperiod,$

$BP(i) = BuyingPressure(i) = \sum_{n=t-i}^{t} C_n - Min(C_{n-1}, L_n),$

$$TR(i) = TrueRange(i) = \sum_{n=t-i}^{t} Max(C_{n-1}, H_n) - Min(C_{n-1}, L_n).$$

The same nonlinear analysis of a NARX NN and backward elimination technique are used to select the best set of technical indicators. The trading volume and the technical indicators are used as inputs to train the network in order to measure its performance in terms of MAPE over 10 years of each selected futures market. Thus, a MAPE for each futures market for the trained network with selected indicators is obtained. Each time, as a backward elimination technique, one of the indicators is omitted and the network is retrained to check whether the performance (MAPE) increases or falls. All possible network performances in all selected markets are compared to see which ones accurately determine their trends. The results show that the presence of OHLC, RSI, MACD, MACD signal, stochastic fast %K, stochastic slow %K, stochastic %D, and ultimate oscillator are significant as inputs to the models, whereas MACD histogram, spot price, and volume are not at all relevant for the purpose of achieving the best performances for the proposed models.

The valid inputs, namely OHLC, RSI, MACD, MACD signal, stochastic fast %K, stochastic slow %K, stochastic %D, and ultimate oscillator are then trained in NNs to estimate future market values. Different modern models, NN, WNN, and WPCA-NN, are examined to find the best model architecture to forecast the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets. Each of these models is evaluated quarterly over 10 years (January 1, 2005 to December 31, 2014), a total of 40 datasets per market. Each dataset consists of the past three years for training and the current quarter for predictive accuracy and profitability performance.

3.7.4 Architecture of the Models

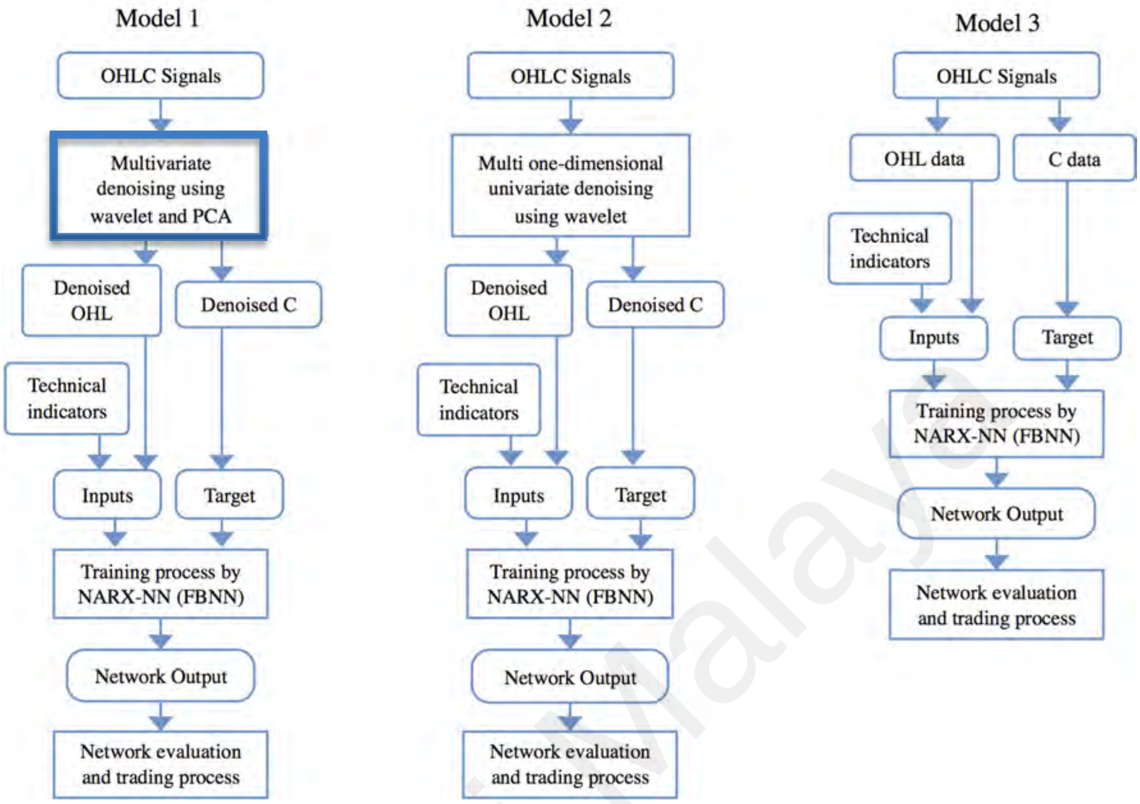
The architectures of the three intelligent models in this study are discussed here. First, the WPCA-NN model is considered a novel hybrid model and is compared with the forecasting power of its constructive elements. Second, a univariate wavelet denoising process, which is used with a NARX-NN model for forecasting, is the so-called WNN model. Third, a NARX NN, which is the forecasting core of the hybrid without any denoising process, is the so-called pure NN model. Comparisons of the WPCA-NN model (Model 1) with the WNN model (Model 2) and the pure NN model (Model 3) address hypotheses H4b and H4a respectively.

Figure 3.4 illustrates the proposed models' architectures with a flowchart. Model 1, the WPCA-NN model, proposes multivariate denoising by WPCA on the first part of each dataset of the OHLC signals in order to gain denoised OHLC signals. This process consists of various settings and variables as follows: wavelet type, level of denoising, thresholding strategy, and choosing the number of principal components. These settings and variables are discussed in the next section. Then, reconstructed denoised open, high, and low (OHL) index signals, with the selected technical indicators as inputs and a denoised close index (C signal) as targets, are introduced to the NARX-NN model with the Levenberg–Marquardt optimization algorithm in order to train the network. This procedure requires some variables and settings as follows: the number of delays, the number of hidden layers and portions of training, validation, and testing. These variables and settings are discussed in the next section. The trained networks from the prior three years are used to forecast the second part, the current quarter, with a one-step-ahead prediction technique. In this technique, the data of the first day of the second part are added to the data of Part 1; the denoising process is then applied again. After this, the currently trained network forecasts the next day of Part 2, based on the newly entered data. The predicted value (output) is compared with the actual data (target value) so as to calculate the forecasting error and obtain signals for buying or selling. This process continues daily over the next three months in order to evaluate the predictive performance of the dataset. The procedure is repeated over all the markets.

Model 2, the wavelet and NARX-NN model, or WNN, differs from Model 1 in terms of the denoising process only. Denoising in this model is performed as multiple univariate denoising by wavelet separately on each component of the OHLC signal. This procedure requires some variables and settings, including wavelet type, level of denoising, and thresholding strategy. The rest of the process follows that of Model 1.

Model 3 is a pure NARX-NN model with no denoising for the preprocessing of data, neither in the training nor in the evaluation steps. OHLC signals and the technical indicators are directly introduced to the NARX-NN model and the procedure continues in the same way as for the rest of the models in terms of training and evaluations. In conclusion, the ensemble and single models require different variables and settings in their structures. This issue is discussed in the next section.

Figure 3.4: Feedforward backpropagation neural network architecture



The key feature of a successful hybrid model with wavelet analysis is the settings of its structural elements, which are wavelet type, level of denoising, thresholding strategies, and the selection of the number of principal components. In order to perform a wavelet transform in the preprocessing stage, it is necessary to select the proper wavelet type from the range of strategies and various sequences available. These are Haar, Coiflets, Dmeyer, Daubechies, and Symlets (Antoniadis et al., 2001; Johnstone, 2011). Table 3.5 shows the wavelet families and their subsets used in this study to decompose the original OHLC signals. This study compares the performances of the various wavelets of each futures contract in order to obtain the best-performing settings. Since each futures contract has different characteristics, each requires different decomposition techniques to be preprocessed adequately. Although Daubechies and Symlets use future data for transformations, causing boundary problems, the effect is overcome here by using time adjustments in the evaluation (Johnstone, 2011). Thus, in order to overcome the boundary problem of the wavelet families, the data is inputted from the start to the n_i^{th} data point into the model in order to forecast the n_{i+1}^{th} index.

Table 3.5. Wavelet families and subsets

Wavelet name	Subsets*
Haar	Haar
Daubechies	db2 db3 db4 db5 db6 db7 db8 db9 db10
Symlets	sym2 sym3 sym4 sym5 sym6 sym7 sym8
Coiflets	coif1 coif2 coif3 coif4 coif5
Dmeyer	Dmeyer

* Subsets are presented in their shortened form as in the wmulden function (MathWorks, 2014c).

The next variable to select in preprocessing the wavelet transform is the level of decomposition, which is the number of times that the original signal is decomposed by wavelet transform. A prior study proposes the maximum number of decompositions, eight levels, and a rule to select the best level for decomposition (Aminghafari et al., 2006). Although with more decomposition it may be possible to remove more noise and obtain the underlying trend of the time series, fluctuations that carry market characteristics may also be removed. Hence, this research experiments with one to eight levels of decomposition in order to find the optimal denoising level for each of the futures markets studied.

After using each wavelet transform on the original signals, denoising parameters need to be selected. The first step is to select a thresholding method. The basic idea is based upon the base information that the wavelet coefficients propose. Intuitively, small wavelet coefficients are combined with noise, while large wavelet coefficients contain more signal information than noise (Wang & Gupta, 2013). In this situation, it is rational to attain a suitable denoising of a given signal if two basic processes are executed: remove those components with small coefficients and reduce the influence of components with large coefficients (Wang & Gupta, 2013). Generally, this study is thresholding or shrinking the absolute value of wavelet coefficients in accordance with a suitable method or rule such as fixed form threshold, rigorous Stein's unbiased risk estimate (SURE), heuristic SURE, minimax, penalized high, penalized medium, and penalized low (Donoho, 1995; Johnstone, 2011). The aforementioned thresholding rules are applied to each set of decomposed signals in order to estimate the noise covariance matrix.

After the wavelet decomposes each of the original OHLC signals separately, PCA then analyzes these four signals simultaneously in order to extract the similar noise contained in them and obtain the principal signals (denoised OHLC). The final step in this preprocessing stage is to select the appropriate number of useful principal components. PCA is the common name for a method that uses complex fundamental mathematical principles to convert a number of feasibly correlated variables into a lower number of linearly uncorrelated variables known as principal components. Although PCA has a variety of other usages, it is generally applied in multivariate data analysis (Aminghafari et al., 2006; Bakshi, 1999). This transformation is expressed in such a way that the initial principal component possesses the highest probability variance and each following component in turn possesses the largest variance possible, with the limitation that it is orthogonal to the earlier components. As suggested by a prior study (Aminghafari et al., 2006), the Kaiser criterion can be employed to select those components with matching eigenvalues that are larger than the mean of all the eigenvalues (Karlis et al., 2003). Implementing the Kaiser criterion, implanted as a module in the coded function of wmulden in MATLAB (MathWorks, 2014c), results in one principal component in model 1's analysis of all markets. Since OHLC consists of four signals, if this study sets four principal components, this is equivalent to not applying the PCA technique, an approach that is represented by model 2.

Figure 3.5 shows actual close data, generalized univariate wavelet denoising of close data, and multivariate denoising of close data using wavelet-PCA on the Hang Seng futures market for 2002, as an example. The wavelet settings used for denoising in Figure 3.5 is based on the best-performing settings as shown in the next section. According to the figures, although the levels of decomposition and thresholding strategy for both univariate and multivariate denoising are the same, wavelet-PCA seems to extract more noise than the univariate wavelet and to achieve a greater smoothed version of the original data. Hence, better forecasting results may be obtained from the underlying functions derived by wavelet-PCA.

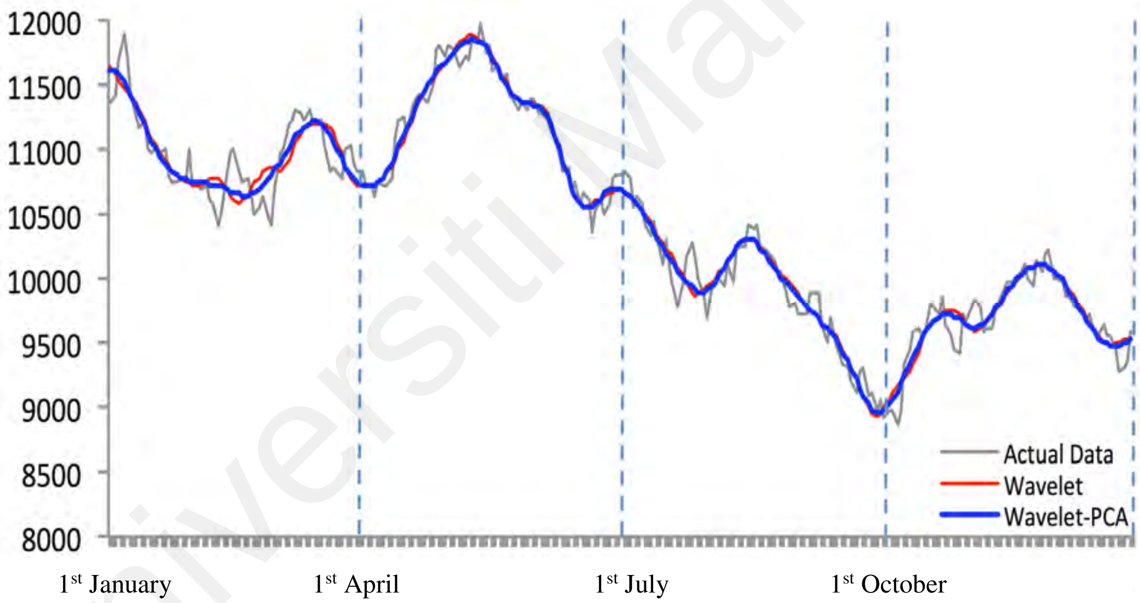


Figure 3.5: Univariate (wavelet) and multivariate (wavelet-PCA) denoising of Hang Seng futures, 2002

A loop of the wavelet-PCA denoising method is almost performed in order to achieve a set of all denoised signals. The aim is then to apply them as part of the input variables for the ANN model. Figure 3.6 illustrates this loop methodology for denoising and training the raw data with all the aforementioned settings of Model 1, the WPCA-NN model, and Model 2, the WNN model. This study uses 23 different wavelet settings (Table

3.5), one to eight levels of decomposition, seven different thresholding strategies, and two different PCA settings (one principal component for WPCA and four principal components for the WNN model) in order to achieve a total of 2,576 sets of denoised signals (denoised open, denoised high, denoised low, and denoised close).

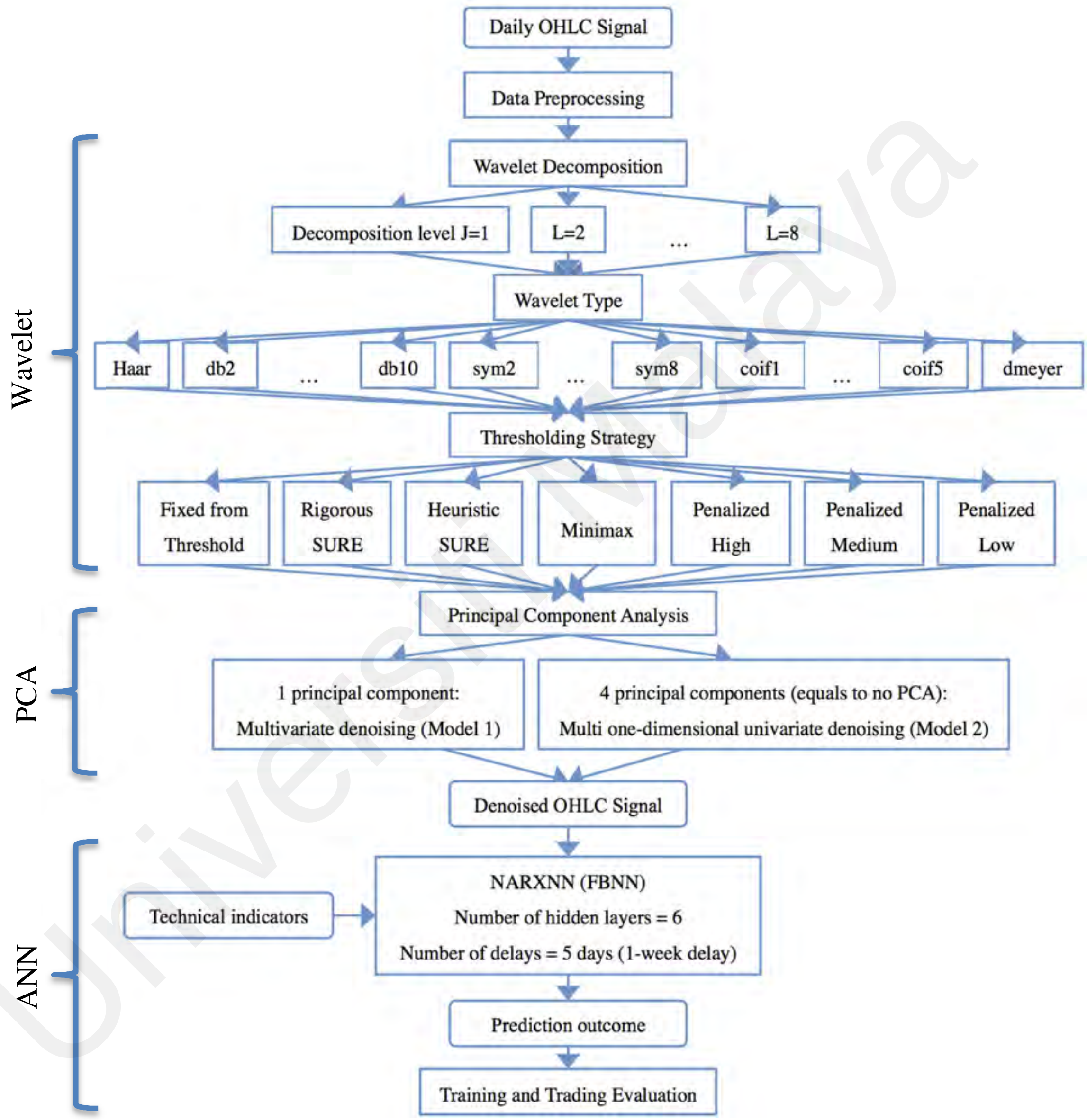


Figure 3.6: Research framework

Training an NN requires two main configurations: the numbers of hidden layers h and delays d. Although some techniques to choose suitable numbers for h and d exist, there is no flawless guidance. Thus, it is best to apply a backward elimination technique in order to achieve the optimum generalization result (Wang & Fu, 2006). Backward elimination may be computationally expensive; however, it is reliable. With the range of results derived from a backward elimination technique, networks constructed with five days' delay and six hidden nodes repeatedly achieve satisfactory results in all selected markets. There may be other combinations of h and d that perform better in a specific market, but since the main objective of this study is to investigate the performance of different denoising models, the research is undertaken repeatedly with one successful setting of the ANN among all markets.

In this study, the input series $x(t)$ in the NARX-NN model is denoised OHL signals together with the technical indicators, OHLC, RSI, MACD, MACD signal, stochastic fast %K, stochastic slow %K, stochastic %D, and ultimate oscillator, calculated by using the original OHLC signals. Further, $y(t)$ is the denoised close of the futures' time series, which are considered the target to be predicted. The prediction procedure is implemented with various settings of wavelet-PCA and the NARX-NN model, and predictive performance is examined by NN error terms over the evaluation periods.

An example of a time series response to training is illustrated in Figure 3.7, Figure 3.8, and Figure 3.9 for the intelligent models of pure ANN, WNN, and WPCA-NN. Since the first in-sample data is from 2002 to 2004 (Figure 3.2 and Figure 3.3) and the first market studied is the Hang Seng futures market, these three figures are presented as examples for 2002–2004 and the Hang Seng futures market.

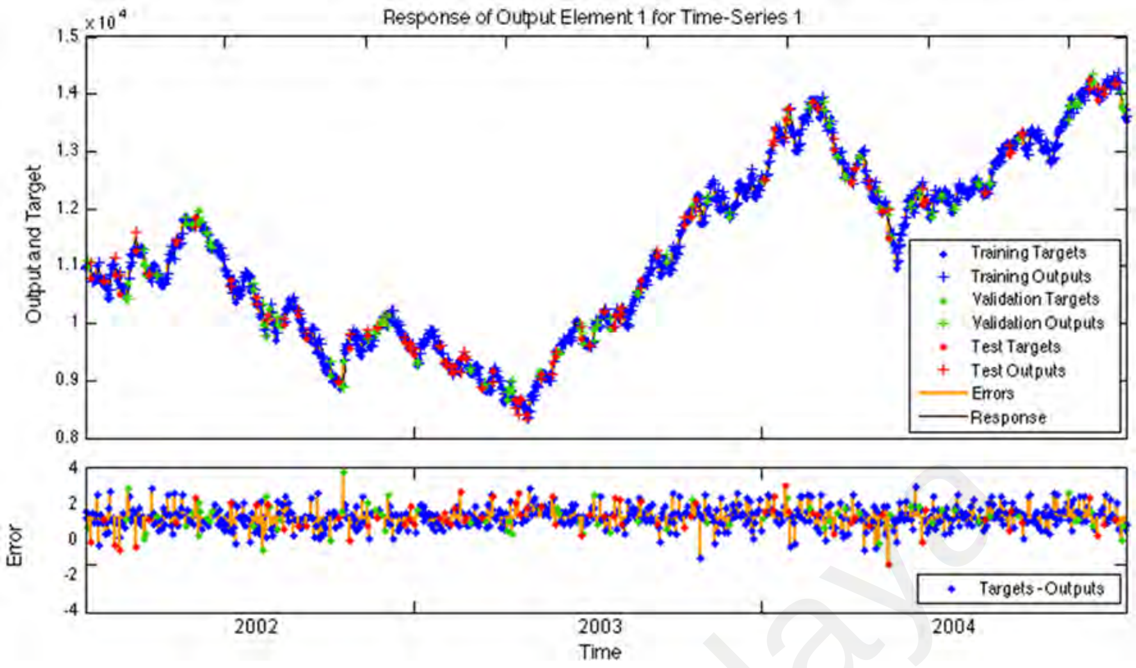


Figure 3.7: Time series response to training with the pure ANN model, the Hang Seng futures market, 2002–2004

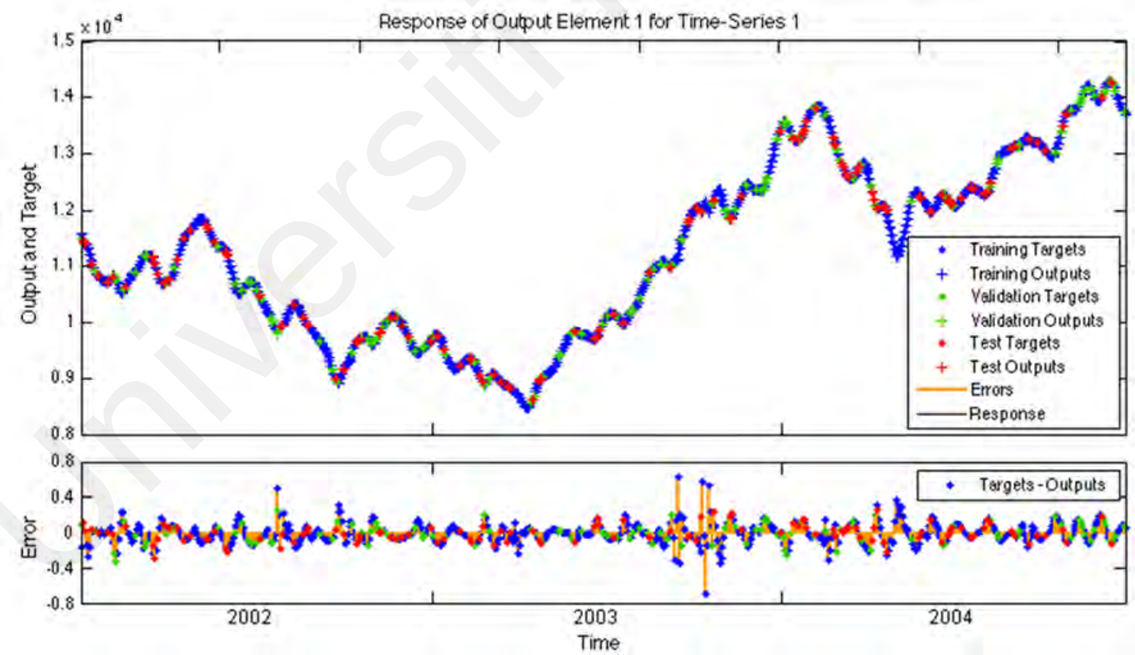


Figure 3.8: Time series response to training with the WNN model, the Hang Seng futures market, 2002–2004

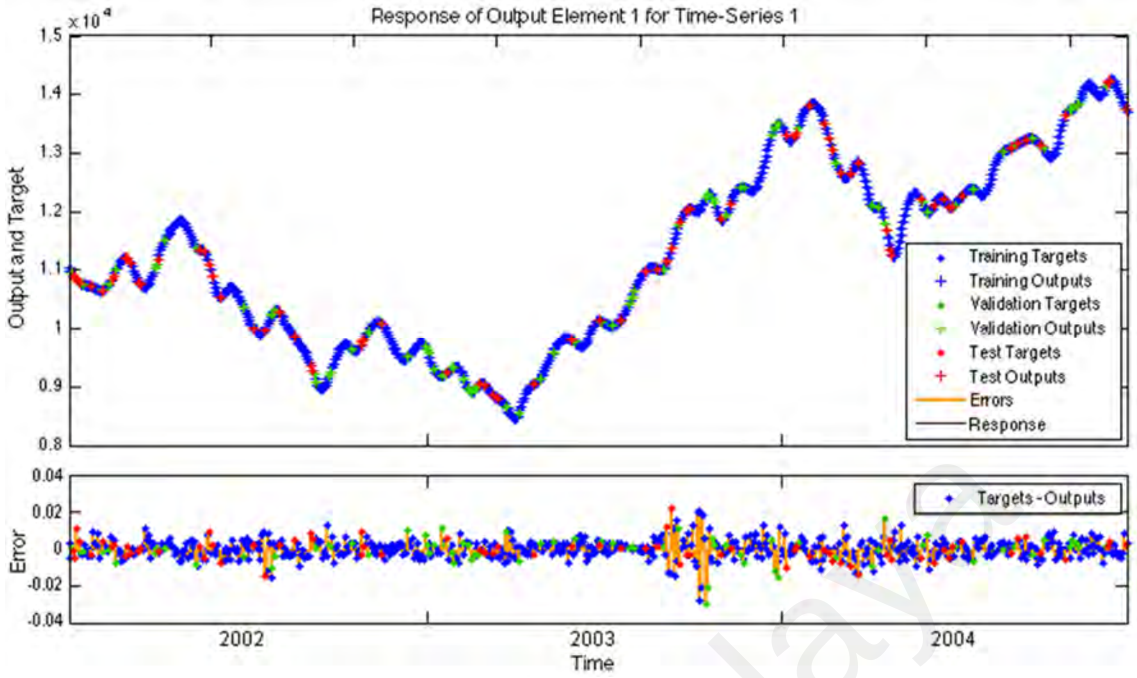


Figure 3.9: Time series response to training with the WPCA-NN model, the Hang Seng futures market, 2002–2004

According to the research framework, the in-sample part of the time series is divided into three: training (70%), validation (15%), and testing (15%). These are randomly distributed among the in-sample data. The training, validation, and testing portions are specified and presented in different pointers in Figure 3.7, Figure 3.8, and Figure 3.9. The first plot in all three figures displays how training, validation, and testing outcomes are located compared with target values and actual data. The closer the output of the simulation (training, validation, and testing) to the target, the better fitted the outcome is to the actual data. The second plot in the figures shows the difference between each figure's output and target values with a MAPE ratio as the error term for each day, from 2002–2004. According to these figures, the absolute amount of forecasting performance for the ANN model is between 0 and 4%; for the WNN model, it is between 0 to 0.8%; and for the WPCA-NN model, it is between 0 to 0.04%. Not only is the higher performance of the WPCA-NN model compared with the others a good sign in terms of its high forecasting power (0.04% < 0.8% < 4%), the specifically higher performance of the WPCA-NN model compared with the WNN model (0.04% < 0.8%) is also a significant suggestion that the enhancement to the denoising process by using multivariate wavelet-PCA is likely to result in a suitable outcome. Apart from investigating the performance of the WPCA-NN model compared with other traditional and modern methods, the examination of the enhancement of multivariate denoising in the WPCA-NN model, compared with the univariate denoising process of the WNN model, is one of the key objectives of this study and addresses hypothesis H4b. In the next section, the MAPE ratio, as the means of forecasting performance in this study, is introduced.

3.7.5 Forecasting Performance

MAPE is applied as a measure of predictive accuracy to select the model that performs the best (Makridakis, 1993). In this regard, it can evaluate and compare the predictive power of the models. The definition of MAPE is

$$MAPE = \frac{\sum_{t=1}^{n}\left|\frac{y_t - y_t^*}{y_t}\right|}{n}, \tag{29}$$

where y_t is the target value and y_t^* is the predicted value. A lower MAPE value indicates a network with higher performance. However, it is unlikely that the value will be close to zero because financial markets are volatile and fluctuate significantly. MAPE is selected here from other measures of predictive accuracy because its result is measured as a relative percentage, unlike other criteria such as root mean squared error and mean absolute error, which are biased to the scale of the index (Hsieh et al., 2011). Hence, the accuracy of the offered models is comparable across the seven futures markets, even with their different scales of indices.

3.7.6 Profitability (Return) Performance

In order to check a model's profitability in addition to its predictive accuracy, this study uses a buy-and-sell trading rule strategy. Such a strategy is widely used for profitability performance (Yao et al., 1999). The strategy works by buying when the next period's predicted value (target) is larger than the current market close and by selling when the next period's predicted value is smaller than the current market close. Thus,

$$y^*(t+1) > y(t): buy, \tag{30}$$

$$y^*(t+1) < y(t): sell, \tag{31}$$

where $y(t)$ is the current market close and $y^*(t+1)$ is the predicted value for the following market day. The net gain or loss is calculated every quarter during the second part of the evaluation period. The return of a trading strategy is widely considered the profitability performance of a model (Enke & Thawornwong, 2005; Leitch & Tanner, 1991; Yao et al., 1999). Hence, the summed return of this trading rule can be calculated by the following equation and used as a comparison scale among the models and markets.

$$Return\ (\%) = 100 \times \left(\sum_{t=1}^{b}\left(\frac{y_{t+1}-y_t}{y_t}\right) + \sum_{t=1}^{s}\left(\frac{y_t-y_{t+1}}{y_t}\right)\right), \tag{32}$$

where b denotes the total number of days for buying futures and s represents the total number of days for selling futures. The trading strategies for all trading rules and over all futures markets are undertaken with two approaches: without transaction cost and with transaction cost. The foregoing procedure is considered without transaction cost. In accordance with a trading broker chosen for this study, Interactive Brokers LLC (LLC, 2015), this study runs live trading strategies on futures with offered fixed-commission rates as follows: the Hang Seng, 0.08%; KOSPI 200, 0.08%; KLCI, 0.08%; NIKKEI 225, 0.05%; S&P 500, 0.02%; SiMSCI 0.08%; and TAIEX, 0.08%. The initial capital is 1 million USD.

3.8 The Methodology of Wavelet-PCA Neural Network

All the components and methods used to create the novel hybrid model of WPCA-NN were introduced according to the background of the study in the previous section. In this section the method of wavelet-PCA neural network will be introduced briefly and empirically.

3.8.1 Data Collection

First step is collecting the data from Bloomberg terminal, which is located in the lab of the faculty of Business and Accountancy, Universiti Malaya. In order to find the futures indices of the willing markets you can search to find them or may need to email the support team to get the proper advice. In addition to that, the following address is a way to find the willing futures indices in the Bloomberg terminal (https://www.bloomberg.com/markets/stocks/futures). The index of the selected futures markets in Bloomberg terminal are respectively, IK1:IND for KLCI, FT1:IND for TAIEX, KM1:IND for KOSPI 200, NK1:IND for Nikkei225, SD1:IND for SiMSCI, HI1:IND for Hang Seng and S&P 500 for ES1:IND. Not only opening, highest, lowest and closing indices can be downloaded from Bloomberg Terminal, but also technical indicators required for this study can be exported, Figure 3.10. Other than Bloomberg terminal, Yahoo finance and Google finance can be reliable sources of data to input for such studies. Moreover, cleaning the data is one of the most important things before using, as it is already discussed in Section 3.3.

Figure 3.10: A screenshot of Bloomberg terminal with RSI indicator

In order to show the methodology of WPCA-NN, we explain the prediction of the first quarter of the KLCI futures contract in the following, and the illustration with detailed screenshots from Matlab are displayed in Appendix B. The in-sample data for this illustration is the data from January 1st 2002 to December 31st 2004. Moreover, the out-sample data is the first quarter of 2005 which is January 1st 2005 to March 31st 2005. The next step is to import open, high, low and close price of the in-sample data (2002-2004) into the Matlab. Three years of KLCA data (2002-2004) is imported into the Matlab. According to the Section 3.7.3.1, three years of data is required to forecast the next quarter. Therefore, we apply the data of 2002-2004 as in-sample data, in order to forecast the first quarter of 2005-2014, which is the first period of out-sample data.

3.8.2 Denoising OHLC Signal Using WPCA

When the data is imported, according to the Sections 3.7.1.2 and 3.7.4, we are required to denoise the signal, OHLC data. In order to denoise the signal we use Wavelet toolbox in Matlab. The Wavelet-PCA is already a coded tool in Matlab toolbox. In order to open the Wavelet toolbox, we need to type Wavemenu in the Command Window. The wavelet tool main menu will be appeared, and then we select Multivariate Denoising under Multiple 1-D section.

Once the Multivariate denoising toolbox is opened, we need to load the input signal from the File menu on top. Then we need to open the file, which is already imported to the Matlab and includes the open, high, low and close signal. When the OHLC signal is loaded the wavelet family and the number of the level of decomposition should be chosen from the list. For instance, we have selected Symlet 4 as it is shown in the Appendix B. By clicking on the Decompose and Diagonalize, the OHLC signal will be decomposed and shown based on the given parameters. Now that the signal is decomposed, the denoising parameters needed to be set. The required thresholding method will be selected from the list in the section of select denoising parameters. For instance, we have selected Penalize High as it is illustrated in the Appendix B. There is a checkbox for PCA, which should be checked. There are two parameters of Nb Of PC for APP and Nb Of PC for Final PCA, which both should be set to 1 in order to extract the same noise within all open, high, low and close signals. By clicking on Denoise button, the input signals will be denoised and the original signals based on the wavelet and PCA parameters will be driven. The denoised signal or original signal will be shown along with the initial signal in the same chart if the Show Denoised Signals is checked in the toolbox frame. The final step in the denoising toolbox is to save the denoised signal by clicking Save Denoised Signals Only in the File tab in the menu, so we can load it into the Matlab workspace for future steps.

In this illustration the parameters of the Symlet 4, 3 levels of decomposition and Penalize High are used for denoising. However, there are other parameters for WPCA, which are completely described in the Section 3.7.1. Also, it is discussed that we will test the combination of all these parameters and settings for denoising of each futures indices. In order to do that, a loop of denoising command by WPCA is required for Matlab software. Figure 3.11 displays the Matlab code for calling the WPCA denoising technique. Wmulden is a function that is used for WPCA denoising technique, Section 3.7.1. Ohlc_signal is the signal of open, high, low and close, which will be denoised with Wmulden function. Level is the level of decomposition. Wname is the wavelet family name. The parameters of npc-app and npc-fin are the dimensions of the principle component matrix, which both are 1 in WPCA case. Tptr is the thresholding strategy, e.g. Penalize High. Sorh is the mode of denoising, which is set Soft for WPCA technique. Den_signal is the output of the WPCA technique, which is the denoised version of OHLC signal.

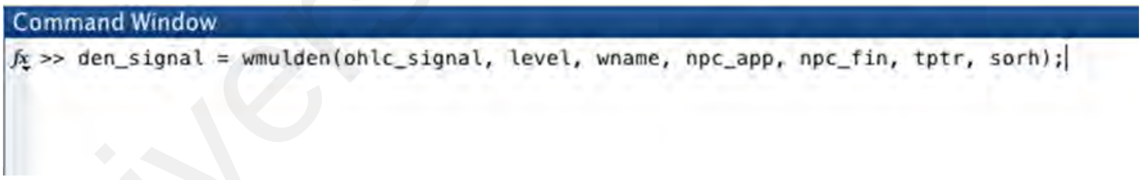


Figure 3.11: WPCA Denoising Code for Matlab

3.8.3 Preparing Denoised Data to Feed to Neural Network

So far, the OHLC signal is denoised and we are required to prepare the inputs and targets for the neural network toolbox. The denoised open, high and low signals will be separated from denoised OHLC signal and will be combined with the concurrent technical indicators, also called input series. The technical indicators have already been downloaded among with data from Bloomberg terminal. Therefore, there will be a series of three years data including, denoised OHL signal plus three years of concurrent technical indicators, and related to e.g. KLCI futures from January 2002 to December 2004. Moreover, the denoised close index will be stored in a separated series in order to be applied as target series for the neural network toolbox.

3.8.4 Training Network by Neural Network Toolbox

The next step is calling the NARX-NN toolbox in the Matlab. In order to that, we type the Ntstool in the command window. Neural time series toolbox will be opened, and then we select Nonlinear Autoregressive with External (Exogenous) Input (NARX) radiobutton. After clicking on the Next button, we need to select Inputs and Targets series from the prepared files. The series of inputs and targets are well discussed in the previous paragraph. After that, we need to select the suitable time series format depends on the physical orientation and format of the data in the series. If you have imported data from an excel file, you may need to change the format to Matrix Row. According to the Section 3.7.3, validation and testing percentages will be set to 15%. In the next page number of hidden neurons and number of delays are set to 6 and 5 respectively, regarding to the Section 3.7.2. In the next page, we need to choose Levenberg-Marquardt as a trading algorithm, according to the Section 3.7.2. By clicking on the Train button, the neural network training will open the training box and begin the progress. The Time-series response, Training performance, Number of iterations and some other parameters are the outcomes of the training, which can be saved and analyzed. Regarding to the Section 3.7.2, NARX-NN uses different biases and weights every time it starts the training. Therefore, each time we train the same inputs and targets with the same parameters, we achieve different trained network with different number of iterations and performance. There is no rule of thumbs for the number of iterations in order to qualify an outcome. However, according to the experience the larger the number of iterations to reach the trained network,the better. The performance value is the error term and the difference between targets and outputs (Section 3.7.5). Therefore, the lesser value of the performance the better network. We may retrain the network until we achieve a well performed and most fitted forecasting output to the target series.

3.8.5 Selecting The Best-Performing Network in The In-Sample Data

Each period of the in-sample data (past three years) in this study has been trained 30 times, and the 10 best-performed networks are selected and saved. The final page of the neural time series toolbox is the Save results. By checking the save network to Matlab network object and save outputs to Matlab matrix checkboxes, and clicking on Save Results button the trained network and the outputs need to be saved for further use.

So far 10 trained network is saved and ready to apply to the out-sample data. However, we need to select one of them to apply to the out-sample data. According to the Section 3.7.6, the profitability of these 10 networks during in-sample data is calculated and then the network which generate the highest profitability will be selected in order to forecast the out-sample data.

3.8.6 Preparing The Out-sample Data for Evaluation

So far the best-performing network based on forecasting performance and profitability in in-sample data is selected. The saved network, out-sample inputs and out-sample targets are required to forecast the out-sample data using the trained neural network. In order to prepare the inputs and targets series, the OHLC signal of the out-sample data is required to be denoised by the same parameters that is used to denoise in-sample data, at the beginning of the current section. The inputs series of the out-sample data includes denoised OHL signal plus the downloaded indicators. In addition, the target series of the out-sample data is the denoised close signal. When the inputs and targets series of the out-sample data is ready, they will be fed to the saved neural network to achieve the forecasting results.

3.8.7 Application of The Trained Network For Out-Sample Data

In order to apply the saved neural network with new inputs and targets series, we need to use Matlab codes in the command window. Figure 3.12 illustrate the Matlab codes required to prepare the series and call the saved neural network in order to forecast the out-sample data. The variables in the codes are as follow. The input is a series including denoised open, highest, lowest and selected technical indicators. The target is the denoised closing price. The “net” is the already trained and saved net by neural network toolbox (ntstool). The *y* is the output of the network, when the new data is fed and the network is applied. The commands of tonndata and preparenets are the predefined functions in Matlab to prepare the data and network for training. After running these codes in the command window, the output of the prediction also called *y* can be saved and used for calculating forecasting performance and profitability. According to the formula and methodology introduced in Section 3.7.5, the *T* as target series and *Y* as output series are used to calculate forecasting performance of the selected neural network in the out-sample data. Moreover, the profitability of the saved network is calculated via the trading strategy introduced in the Section 3.7.6.

```
Command Window
fx >> X = tonndata(inputs,false,false);
   T = tonndata(target,false,false);
   [x,xi,ai,t] = preparets(net,X,{},T);
   y = net(x,xi,ai);
```

Figure 3.12: Using a saved neural network in Matlab

3.8.8 Conclusion of the Methodology

So far, the best-performing network in in-sample data (e.g. data from 1st January 2002 to 31st December 2004 for KLCI) is selected, and then it is applied to out-sample data. Moreover, the forecasting performance and the trading profitability of the selected network is measured in out-sample data (e.g. data of 1st January 2005 to 31st March 2005 for KLCI). The results of the forecasting performance as MAPE percentage, and forecasting profitability as return percentage are recorded for analysis of the results.

Thus far, the first quarter of the out-sample data of one of the futures indices (e.g. 1st January 2005 to 31st March 2005 of KLCI futures) is evaluated using our novel hybrid model WPCA-NN. In addition, this study analysis 10 years data of 7 selected futures markets (Section 3.3), which consists of 40 quarters for each futures market. Therefore we are required to repeat this method 280 times in total. Ultimately, all the forecasting results will be prepared for data analysis in Chapter 4.

3.9 Hypothesis Review

Based on the research framework presented in this chapter and the literature review discussed in the prior chapter, the following hypotheses are proposed and examined in order to answer the research questions. All hypotheses are indicated in their alternative forms.

H1: All the futures markets in this study exhibit long-term memory and significant entropy.

$$Hurst\ exponent\ >\ 0.5 \qquad (0.5 < H < 1)$$

and

$$Approximate\ entropy\ <\ 1 \qquad (0 < E < 1).$$

H2: The hybrid WPCA-NN model generates higher returns than the RWH (a passive buy-and-hold strategy) for selected financial markets.

$$MAPE_{WPCA-NN} - MAPE_{Buy\&Hold} > 0 \tag{33}$$

and

$$R_{WPCA-NN} - R_{Buy\&Hold} > 0. \tag{34}$$

H3: The hybrid WPCA-NN model consistently generates more accurate forecasts and higher returns than:

a. MACD strategy for the selected financial markets.
b. RSI strategy for the selected financial markets.
c. Stochastics strategy for the selected financial markets.
d. OMA strategy for the selected financial markets.

H4a: The hybrid WPCA-NN model consistently generates more accurate forecasts and higher returns than an NN model for selected financial markets.

$$MAPE_{WPCA-NN} - MAPE_{NN} > 0 \tag{35}$$

and

$$R_{WPCA-NN} - R_{NN} > 0. \tag{36}$$

H4b: The hybrid WPCA-NN model consistently generates more accurate forecasts and higher returns than a WNN model for selected financial markets.

$$MAPE_{WPCA-NN} - MAPE_{WNN} > 0 \tag{37}$$

and

$$R_{WPCA-NN} - R_{WNN} > 0. \tag{38}$$

3.10 Summary

Based on the data analyses and predictability of the sample data by unit root, serial correlation, and Granger causality test, all analyses and models considered in the hypotheses are valid for use with the sample data. Moreover, the variables and technical indicators are appropriate and able to predict the movement of futures in the markets. The sample data is from the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets from 2002 to 2014. However, the evaluation period, or out-sample data, applies only from 2005 to 2014.

The buy-and-hold strategy and other technical trading strategies, such as RSI, MACD, stochastics and OMA, are tested using Bloomberg Terminal for the period from 2005 to 2014, with and without transaction costs. Moreover, the intelligent techniques of the WPCA-NN, WNN, and pure NN models are performed in a moving time framework of in-sample data and out-sample data. Out-sample data is each quarter for 10 years, starting from the first quarter of 2005 and ending with last quarter of 2014. Thus, there are 40 quarters in total for the out-sample data. The past three years prior to each item of out-sample data is in-sample data. The three intelligent models of this study are separately trained for in-sample data and the best-performing network of each intelligent model applies to trading with the out-sample data. This approach reflects the method that portfolio managers could use (Bahrammirzaee, 2010; Campbell et al., 2009; Jegadeesh & Titman, 1993; Pesaran & Timmermann, 1994; Sridharan, 2015). Moreover, with this approach, the chance of data mining in the analyses is eliminated (Neely & Weller, 2011).

According to the description of the hypotheses in the prior section, hypothesis H1 aims to answer whether the selected markets are predictable though long-term memory and entropy indices. The Hurst exponent and approximate entropy are measured for this hypothesis. A Hurst exponent value between 0.5 and 1 is required for a predictable market, while an approximate entropy value between 0 and 1 is required for an inefficient and predictable market. Further, an H that is equivalent to 0.5 and an E that is equivalent to 1 represent a random market. Hypothesis H2 is posited to answer whether the proposed hybrid model (WPCA-NN) outperforms a buy and-hold strategy in order to show that the futures markets do not follow the RWH. In this regard, the WPCA-NN model must generate statistically significant and higher returns than the buy-and-hold benchmark. Hypothesis H3 tests whether the novel hybrid model performs better than commonly used and best-performing technical trading strategies such as H3a) MACD, H3b) RSI, H3c) stochastics, and H3d) OMA. In such a context, the WPCA-NN model must achieve significantly higher returns than these best-performing technical trading rules. Hypothesis H4a aims to test whether the denoising process of the novel hybrid model adds to the forecasting performance of the pure NN model. The WPCA-NN and pure NN models are both tested for the same period and the same futures markets. The WPCA-NN model must forecast the markets with significantly higher performance and returns. Hypothesis H4b tests whether the hybrid multivariate denoising process of the new proposed model is better than the simple univariate denoising process of the WNN model. Here, the WPCA-NN model must show significantly higher performance and profitability compared with the WNN model. In addition, the outcome of hypothesis H4b will demonstrate whether the adjustment and enhancement of the denoising process by multivariate wavelet-PCA is a promising tool and can be employed in future studies. The results of the analyses for all these hypotheses, and the interpretation of the results, are presented in chapter 4.

CHAPTER 4: RESULTS AND RESEARCH ANALYSES

4.1 Introduction

This chapter presents the results of all proposed models and methods to address the hypotheses. It starts with the results of hypothesis H1 (Sections 4.2 and 4.3). An overview of the forecasting results is presented in the next Section 4.4. After this, the significance of the results with the relevant descriptive analysis and discussion of reliability are presented (Section 4.5). In Section 4.6, comprehensive discussion around the results of each futures market for H2, H3, and H4 are presented. Section 4.7 discusses risk, tests, and comparisons in the context of the results. The following Section 4.8 presents the research results, and Section 4.9 offers a summary of the discussion.

4.2 Hurst Exponent Index

The Hurst exponent (H) is a statistical amount applied to manage time series. It is also known as the rescales range or R/S. If $H = 0.5$, all autocorrelations tend promptly to 0 and the time series indicate a random walk. The further away the amount of R/S or H is from 0.5, the more powerful the memory impact in the time series. Small R/S or H ($H < 0.5$) suggest a mean reversion or anti-persistence in the time series, while large H ($H > 0.5$) suggests a mean aversion or persistence in the time series. This study considers the application of the Hurst exponent index to check the predictability and presence of long-term memory in selected time series. The Hurst exponent or R/S measure is calculated for all selected financial time series from 2005 to 2014 in accordance with the introduced measurement technique described in Chapter 3. The results are given in Table 4.1 and Figure 4.1. They show that for all markets, the Hurst exponents are larger than 0.5 and demonstrate long-range dependence and predictability of the current data in terms of their past data.

Table 4.1: Hurst exponent index results for all markets from 2002 to 2014

	Hang Seng	KLCI	KOSPI	NIKKEI	SiMSCI	S&P 500	TAIEX
2005	0.5801	0.5527	0.5898	0.5589	0.5928	0.5320	0.5344
2006	**0.6332**	**0.5481**	0.5860	0.5216	**0.6308**	0.5432	**0.5608**
2007	0.6247	0.5524	0.5920	0.5095	0.6153	0.5278	0.5389
2008	**0.6402**	**0.5706**	**0.6122**	**0.5509**	**0.6214**	**0.5303**	**0.5763**
2009	0.6315	0.5558	0.6016	0.5300	0.6303	0.5340	0.5686
2010	0.6157	**0.5452**	**0.5978**	0.5471	0.6183	0.5432	**0.5510**
2011	0.6045	0.5281	0.5851	0.5504	0.5781	0.5019	0.5391
2012	0.6423	0.5173	0.5999	0.5535	0.6107	0.5120	0.5518
2013	0.6319	**0.5318**	0.5763	**0.5593**	0.5971	**0.5463**	0.5439
2014	0.6247	0.5432	0.6006	0.5481	0.6175	0.5193	0.5658
2005–2014	0.5973	0.5507	0.5237	0.5109	0.5840	0.5128	0.5383
Ranking	7	5	3	2	6	1	4

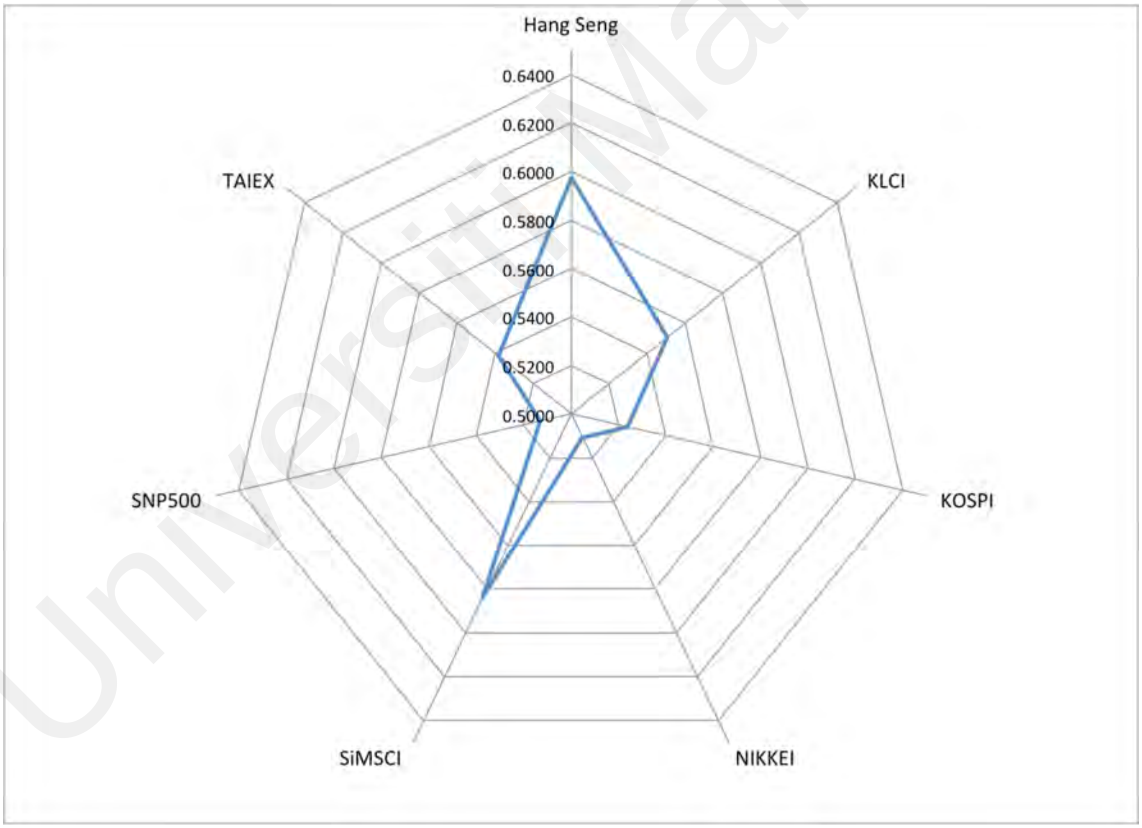


Figure 4.1: Hurst exponent for all futures markets from 2005 to 2014

According to the results displayed in Table 4.1 and Figure 4.1, the Hang Seng (0.5973), SiMSCI (0.5840), and KLCI (0.5507) futures markets are more inefficient and predictable than other futures markets. Further, the NIKKEI 225 (0.5109), S&P 500 (0.5128), and KOSPI 200 (0.5237) futures markets are considered to be more efficient and less predictable than other futures markets. Appendix C displays radar charts of the annual Hurst exponent values for each futures market. Taking into account these annual values, it is inferable that persistence changes annually. In other words, the market predictability and inefficiency of these futures markets vary over time. These results are consistent with other studies (Kristoufek & Vosvrda, 2014; Hull & McGroarty, 2014). Long-term memory has previously been studied in relation to 38 stock indices (Kristoufek & Vosvrda, 2014). The results show Hurst exponents for the same markets as used here; the Hang Seng (0.5945), KLCI (0.5489), KOSPI 200 (0.5135), NIKKEI 225 (0.5063), SiMSCI (0.5937), and S&P 500 (0.5026) from 2000 to 2011. These figures are in the same range as in this current study's results. Moreover, in accordance with the Hurst exponents, the prior figures confirm the ranking of the markets (S&P 500 – 1st, NIKKEI 225 – 2nd, KOSPI 200 – 3rd, KLCI – 4th, SiMSCI - 5th, and Hang Seng – 6th). The prior results also conclude that, based on the Hurst exponent, the studied markets have predictability and inefficiency, approximate entropy, and fractal dimension. Approximate entropy (Kristoufek & Vosvrda, 2014; Pincus & Kalman, 2004; Zunino et al., 2010) is also studied in this current research (Section 4.3), in accordance with Section 3.5.

Although several studies have proven the validity of the results of long-term memory using the Hurst index (Kristoufek & Vosvrda, 2014; Mensi et al., 2014; Willinger et al., 1999), this study considers structural breaks in accordance with Bai–Perron tests of L globally optimized breaks, which are in contrast with the null hypothesis of no structural breaks. The results for structural breaks are illustrated in (Appendix D) and shown in Table 4.1 in bold. 2008 is also added into the analysis as a known structural break. The values of the Hurst exponents in those years with structural breaks are not the highest or the lowest. Moreover, the values of the Hurst exponents in such years are not necessarily higher or lower than a year before or after. All values of the Hurst exponents, including those years with or without structural breaks, are almost in the same range for each futures market and almost higher than the overall value of the Hurst exponent. In their entirety, these outcomes for long-term memory investigation do not suggest any spurious results caused by structural breaks or the financial crisis.

4.3 Approximate Entropy

Another index for market predictability in this study is approximate entropy, which ranges between 0 (completely deterministic) and 1 (completely random). Not only can the futures markets be ranked for market predictability and efficiency according to approximate entropy, it is also possible to observe changes in market efficiency over the time. This method enables statements to be made regarding changes in stock market's predictability and inefficiency. Thus, the approximate entropy for all the futures markets from 2005 to 2014 can be computed in terms of a predictability index. First, the predictability or inefficiency of the futures markets are studied. Second, the variation of the index over time from 2005 to 2014 is considered. If entropy is higher than 0 and lower than 1, and approximate entropy shows variation over time in any market, there are statements of predictability and time-varying efficiency (Darbellay & Wuertz, 2000; Kristoufek & Vosvrda, 2014; Matesanz & Ortega, 2008; Mensi et al., 2014; Pincus, 1991), which demonstrate that price movements in futures markets are supported by the AMH (Lo, 2004; Noda, 2012; Urquhart & Hudson, 2013; Urquhart & McGroarty, 2014). Table 4.2 and Figure 4.2 display approximate entropy for all futures markets.

Table 4.2: Approximate entropy for all futures markets from 2005 to 2014

	Hang Seng	KLCI	KOSPI	NIKKEI	SiMSCI	S&P 500	TAIEX
2005	0.3049	0.0426	0.1419	0.1734	0.0325	0.1597	0.2211
2006	**0.1875**	**0.0437**	0.2478	0.1208	**0.0749**	0.1503	**0.2412**
2007	0.0996	0.2453	0.4391	0.0643	0.3276	0.505	0.0071
2008	**0.0217**	**0.3894**	**0.4978**	**0.036**	**0.2787**	**0.3829**	**0.0206**
2009	0.077	0.1737	0.4339	0.0637	0.236	0.381	0.0206
2010	0.0205	**0.2144**	**0.2548**	0.1881	0.0914	0.3559	**0.0077**
2011	0.036	0.2485	0.5554	0.3055	0.1407	0.4694	0.036
2012	0.0925	0.0969	0.3363	0.2772	0.1058	0.3641	0.0925
2013	0.1716	**0.2386**	0.3085	**0.1197**	0.1051	**0.3301**	0.194
2014	0.0915	0.1465	0.1647	0.1433	0.0571	0.3806	0.1586
2005–2014	0.3308	0.2712	0.4429	0.4376	0.2107	0.4221	0.3676
Ranking	5	6	1	2	7	3	4

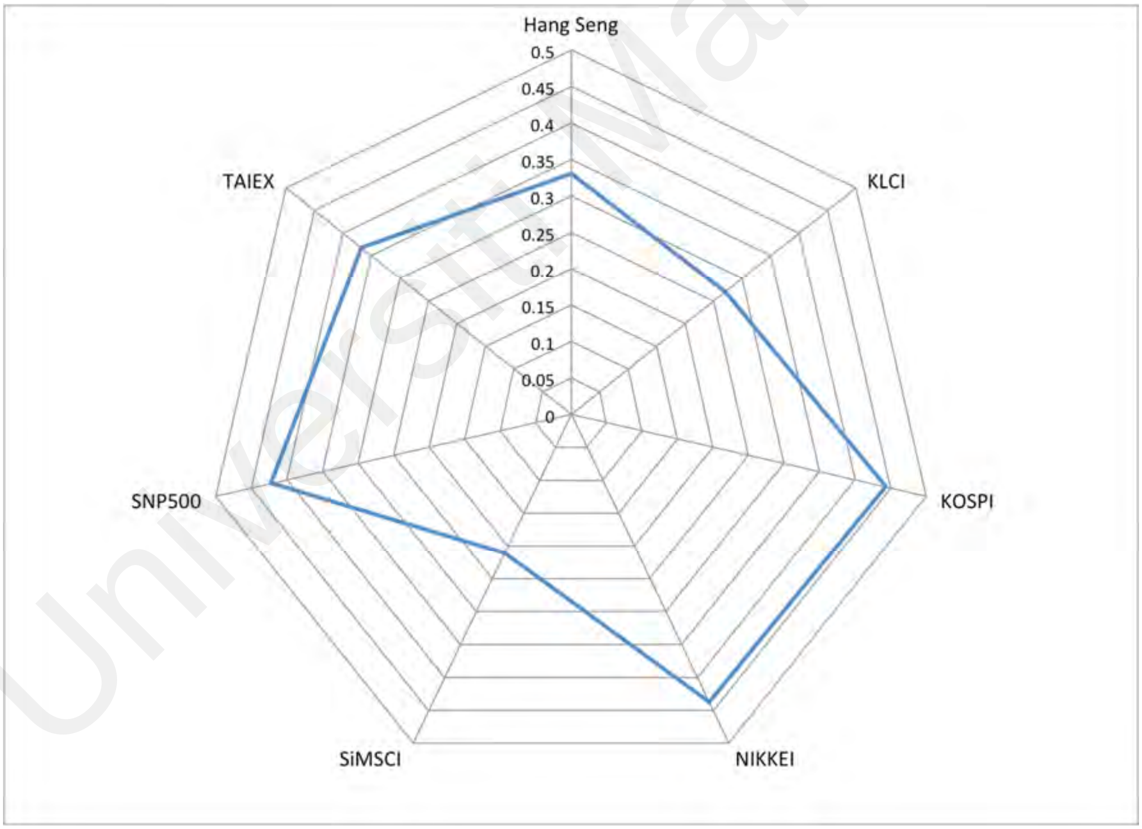


Figure 4.2: Approximate entropy for all futures markets from 2005 to 2014

Approximate entropy is calculated once for the 10 year period and once annually. All results are between 0 and 1, and none of them shows a completely random market. The results are: Hang Seng (0.3308), KLCI (0.2712), KOSPI 200 (0.4429), NIKKEI 225 (0.4376), SiMSCI (0.2107), S&P 500 (0.4221), and TAIEX (0.3676). The higher the value of approximate entropy, the more efficient the market. The futures markets for market efficiency can be ranked based on entropy as follows: KOSPI 200 – 1st, NIKKEI 225 – 2nd, S&P 500 – 3rd, TAIEX – 4th, Hang Seng – 5th, KLCI – 6th, and SiMSCI - 7th. Most importantly, entropy changes over time from 2005 to 2014 and presents changing market efficiency and predictability over this period. These results are consistent with the outcomes of a prior study (Kristoufek & Vosvrda, 2014). The results of this other study measure approximate entropy as follows: Hang Seng (0.3033), KLCI (0.1773), KOSPI 200 (0.4473), NIKKEI 225 (0.4285), SiMSCI (0.2027), and S&P 500 (0.3405) from 2000 to 2011. In this regard, the ranking of market predictability and efficiency is the same as that of the current study. Moreover, the entropy values for the futures markets are almost in the same range. Appendix E shows radar charts of annual approximate entropy for each futures market.

The years with structural breaks, and the year of the financial crisis, are given in Table 4.2 in bold. The financial crisis of 2008 is shown as a known structural break in the analysis. The values of approximate entropy in those years with structural breaks are not the highest or the lowest overall. In addition, the values of entropy in years with structural breaks are not necessarily higher or lower than a year before or after. Ultimately, these outcomes for entropy investigation do not suggest any spurious results caused by structural breaks or the financial crisis.

The results of approximate entropy show variation over time and confirm the results of the Hurst exponent regarding this variation. Thus, both approximate entropy and the

Hurst exponent suggest that these markets are predictable and that their predictability and inefficiency varies over time. These findings are compatible with the AMH, which suggests a time-varying attitude for market efficiency.

4.4 Results of Forecasting Performance: MAPE

So far, the results for long-term memory and approximate entropy are reported. In this section, an overview of the forecast results is presented. The overview addresses hypotheses H2, H3, and H4. This study investigates forecasting ability based on MAPE and the return profitability of the WPCA-NN model and compares the results to existing methods such as selected technical trading rules (Lukac et al., 1988), an NN model (Roman & Jameel, 1996; Siegelmann et al., 1997), and a WNN model (Aminghafari et al., 2006; Wang & Gupta, 2013), as well as to the threshold buy-and-hold strategy (Fama, 1965). Further, in order to show that the WPCA-NN model is sufficiently robust, the trading system has been applied to seven futures markets; namely the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets for ten years (2005–2014). Table 4.3 presents the results of the forecasting performance.

Table 4.3: Results of the forecasting performance, using a MAPE ratio, from 2005 to 2014

Futures Markets	WPCA-NN	WNN	NN
HANG SENG	4.065	15.443	579.4
KLCI	2.166	2.501	17.27
KOSPI 200	2.243	2.805	41.99
Nikkei 225	4.05	10.884	451.89
SiMSCI	1.142	1.792	7.26
S&P 500	2.335	2.354	29.53
TAIEX	2.713	5.005	143.77

Note: Performances in the tables are in accordance with a MAPE ratio (%).

According to Table 4.3, the trading results indicate that the WPCA-NN trading system with multivariate denoising using wavelet-PCA preprocessing and a NARX-NN model outperforms other intelligent models, such as pure NN and WNN in all seven futures markets. For instance, for the Hang Seng futures market, the MAPE ratio of the WPCA-NN model (4.065) is lower than those of the WNN model (15.443) and the NN model (579.4). This result shows lower error, and consequently greater fitted output, for the target value. In addition, multivariate denoising of the WPCA-NN model on OHLC enhances the denoising process and results in more accurate forecasting (4.065 < 15.443).

4.5 Profitability Results

Regarding Table 4.4 and Table 4.5, the profitability outcomes show that the WPCA-NN model has a greater return than other models, including the buy-and-hold and other chosen technical trading strategies, in all the examined futures markets from 2005 to 2014. For instance, for the Hang Seng futures market, the average annual trading profitability of the WPCA-NN model (34.8%) is greater than those of the WNN model (27.9%), the NN model (12.6%), the buy-and-hold (B & H) strategy (7.4%), MACD (2.8%), RSI (-3.1%), stochastics (13.3), and OMA (1.2%). Moreover, multivariate denoising of the WPCA-NN model on OHLC enhances the denoising process and results in a more profitable trading system (34.8% > 27.9%).

Table 4.4: Results of average annual returns from 2005 to 2014

Models	B & H	MACD	RSI	Stochastics	OMA	NN	WNN	WPCA-NN
Hang Seng futures	7.4	2.8	-3.1	13.3	1.2	12.6	27.9	34.8
KLCI futures	6.2	1.3	1.1	-6.9	3	21.7	40.2	47.2
KOSPI 200 futures	17.5	9.9	-13.8	-12.5	-8.3	17.8	40.8	44.8
NIKKEI 225 futures	4.9	0.7	3.7	-5.6	-6.4	13.3	34.9	42.1
SiMSCI futures	10.7	11.5	1.5	-10.6	20.4	19.9	40.6	54.4
S&P 500 futures	6.1	-1.1	4.7	5.3	-7.2	20.5	41.7	48.2
TAIEX futures	8.5	-0.2	2.9	4.6	2.7	8.6	35.3	45.3

Note: Average annual returns are in percentages.

Table 4.5: Results of average annual returns with transaction costs from 2005 to 2014

Models	B & H	MACD	RSI	Stochastics	OMA	NN	WNN	WPCA-NN
Hang Seng futures	9.3	0.5	-2.7	7.3	-0.7	7.8	17.2	21.7
KLCI futures	9.9	4	-0.3	-8.4	6.4	15.3	29.0	35.5
KOSPI 200 futures	29.3	4.5	-26.2	-22.7	-10.6	11.7	25.7	32.2
NIKKEI 225 futures	4.8	-1.5	3.5	-6.8	-9.1	7.9	22.6	29.9
SiMSCI futures	10.5	8.3	0.9	-12.2	16	13.2	25.8	39.8
S&P 500 futures	6.1	-1.9	4.5	4.7	-8.2	13.4	26.8	31.4
TAIEX futures	8.4	-3	2.3	2.7	-1.6	1.5	23.5	33.1

Note: Average annual returns are in percentages.

Further, the average returns using the WPCA-NN model are significantly higher than the outcomes from prior studies (Cheng et al., 2010; Chiang & Doong, 2001; Hsieh et al., 2011; Huang et al., 2009; Kim et al., 1998; Lee et al., 2010; Leung et al., 2000; Necula, 2009; Quah & Srinivasan, 1999).

The current studies outcomes are completely consistent with other similar studies (Booth et al., 2014; Hu et al., 2015; Xiao et al., 2014; Zhang et al., 2014) on machine-learning that use technical analysis indicators, where frequent accurate forecasting leads to excess returns over the buy-and-hold strategy.

Performances of the different models are measured and evaluated on the basis of the lowest MAPE and the highest return values. After 2,576 simulations on the seven futures markets, Table 4.6 displays the most common settings for the best-performing networks.

Table 4.6: Settings of the best-performing networks

Settings	Hang Seng	KLCI	KOSPI	NIKKEI	SiMSCI	S&P 500	TAIEX
Wavelet Name	coif5	db9	sym6 or db9	coif5	db7	db9	db9
Decomposition Level	3	3	3 or 4	2	4	3	2
Threshold	Penalized high	Penalized high	Penalized high	Penalized high	Penalized high	Penalized high	Penalized high
Principal Components	1	1	1	1	1	1	1
NARX-NN Delays	5	5	5	5	5	5	5
NARX-NN Hidden Layers	6	6	6	6	6	6	6
NARX-NN Training Period	3 years	3 years	3 years	3 years	3 years	3 years	3 years

4.6 Analysis

The return results of the offered novel model of trading need to be significant in comparison with other models. This means that the superiority of the offered model's results should be meaningful and reliable. In order to analyze the significance of the return results, the following tests and analyses are offered and explained.

Skewness, or skew, can be statistically defined as the average cubed deviation from the mean divided by the standard deviation cubed. If the outcome of the calculation is more than zero, the distribution is positively skewed. If the outcome is less than zero, the distribution is negatively skewed. If the outcome is equivalent to zero, this indicates that the distribution is symmetric. In terms of an explanation and analysis, the emphasis is on downside risk. Negatively skewed distributions have what mathematicians call a long left tail, which for researchers can mean a greater probability of largely negative results (Gencay, 1996). Positive skew would mean repeated minor negative results, in which case, largely negative scenarios will not happen most probably.

A skewed or nonsymmetrical distribution occurs when one side of the distribution plot does not mirror the other side. Used in investment profitability, nonsymmetrical distributions are usually explained as being either negatively skewed (representing repeated small gains and a few great losses) or positively skewed (representing repeated small losses and a few great gains).

For distributions with positive skewness, the mode (which is the point at the top of the bell curve) should be less than the median (which is the point lying at the midpoint of the frequency distribution), which should be less than the mathematical mean (the total observations divided by the number of observations). The contradictory rules are regarding negatively skewed distribution, whereby the arithmetic mean is less than the median, which is less than the mode (Zhang et al., 2014). Thus, Negative: Mode > Median > Mean

Positive: Mode < Median < Mean

Kurtosis indicates the peak degree in a distribution. In a leptokurtic situation, more peaks than usual denote that a distribution histogram possesses fatter tails; thus, there are fewer probabilities of excessive results compared with a normal distribution (Zhang et al., 2014). The kurtosis formula calculates the peak degree of a distribution histogram. Kurtosis should be three for a normal distribution; extra kurtosis indicates that kurtosis is below or above three.

For a measured skew number, which is calculated as average cubed deviations divided by the cubed standard deviation, it is necessary to consider the sign of the skewness in order to assess whether a return histogram as a distribution possesses positive skewness (skew > 0), negative skewness (skew < 0), or zero skewness, the last of which denotes symmetric distribution. Because the formula of a kurtosis value is average deviations to the fourth power divided by the standard deviation to the fourth power, the value is assessed in association with the normal distribution, for which kurtosis equals three (Gencay, 1996). Since excess kurtosis equals kurtosis minus three, any positive value for additional kurtosis would indicate that (Gencay, 1996) the distribution is leptokurtic (representing a lower risk of extreme results and fatter tails in the histogram). Table 4.7 presents the trading returns' descriptive analysis.

Table 4.7: Descriptive analysis of trading returns from 2005-2014

Futures Markets	Mean	Median	Maximum	Minimum	Std. Dev.	**Skewness**	**Kurtosis**
Hang Seng	**9.28**	8.43	33.04	-8.04	8.57	**0.48**	**3.33**
KLCI	**10.05**	8.96	32.11	-3.83	7.19	**0.74**	**3.59**
KOSPI 200	**11.26**	10.26	21.35	2.92	4.86	**0.25**	**2.32**
NIKKEI 225	**10.52**	9.81	26.59	-6.35	6.51	**0.09**	**3.81**
SiMSCI	**13.60**	11.91	29.33	3.49	7.09	**0.73**	**2.60**
S&P 500	**12.05**	10.51	39.57	0.73	8.43	**1.56**	**5.36**
TAIEX	**12.91**	11.72	26.53	-2.29	5.55	**0.19**	**3.93**

The skewness and the manner of return distribution of all futures markets are shown in Figure 4.3 to Figure 4.9.

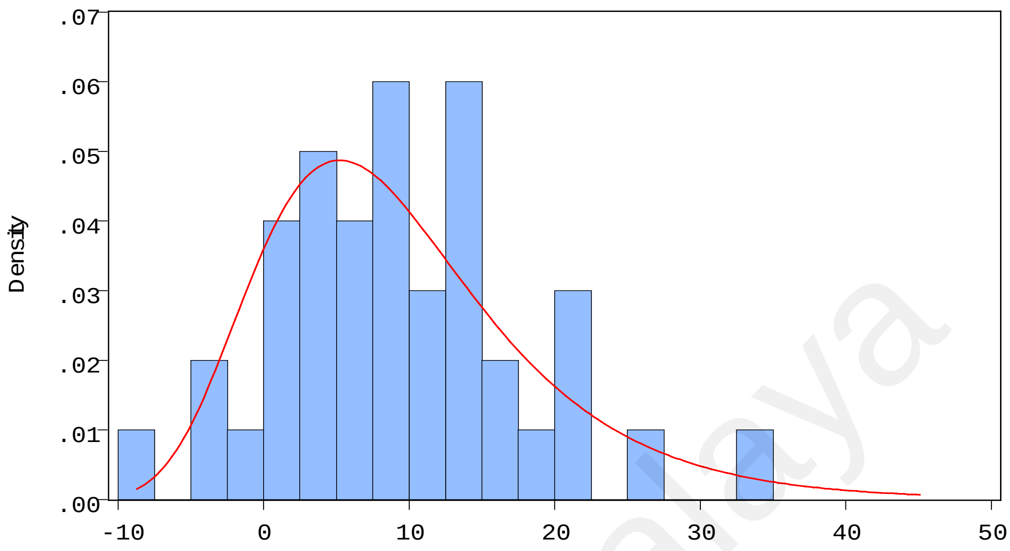


Figure 4.3: Histogram of return distribution, Hang Seng futures market

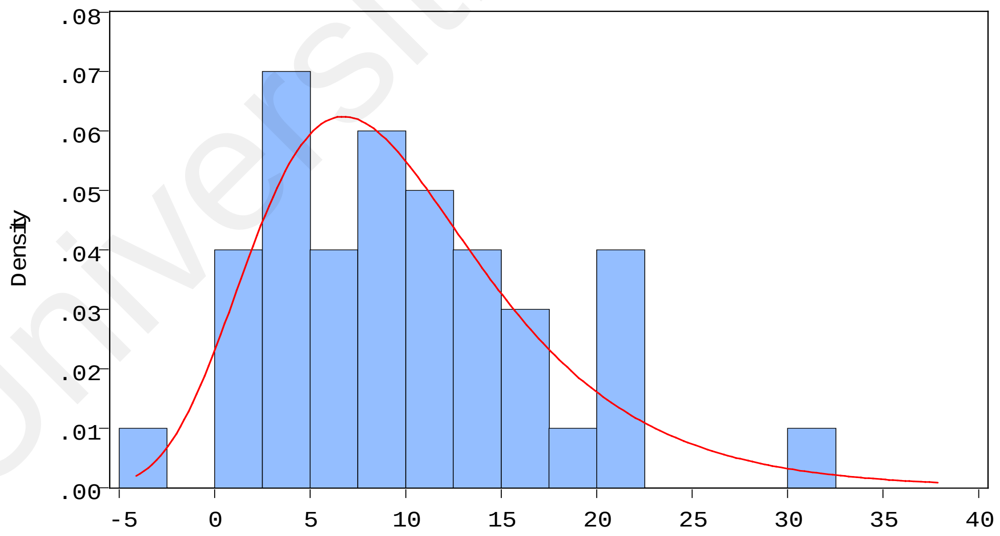


Figure 4.4: Histogram of return distribution, KLCI futures market

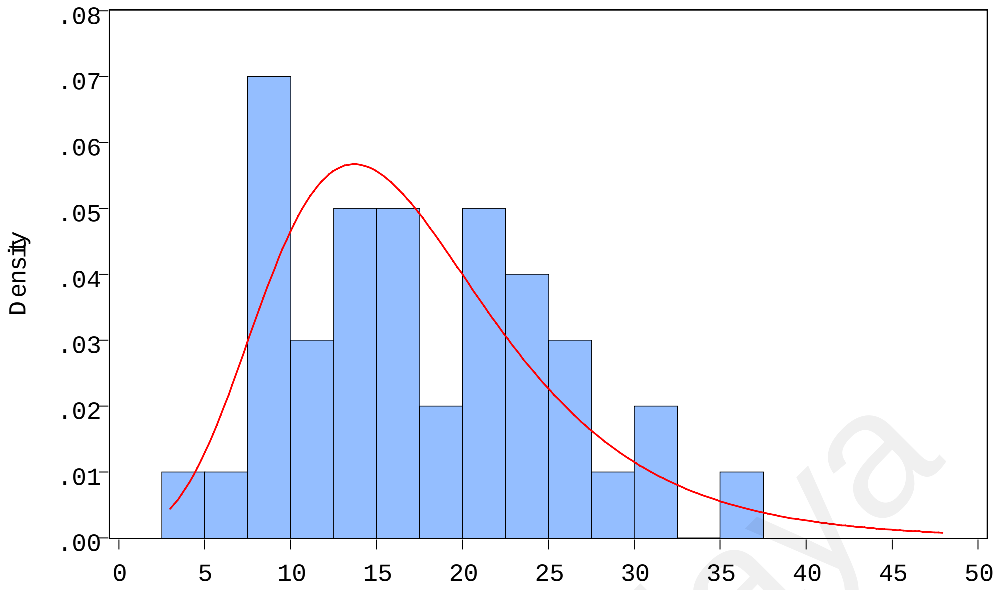


Figure 4.5: Histogram of return distribution, KOSPI 200 futures market

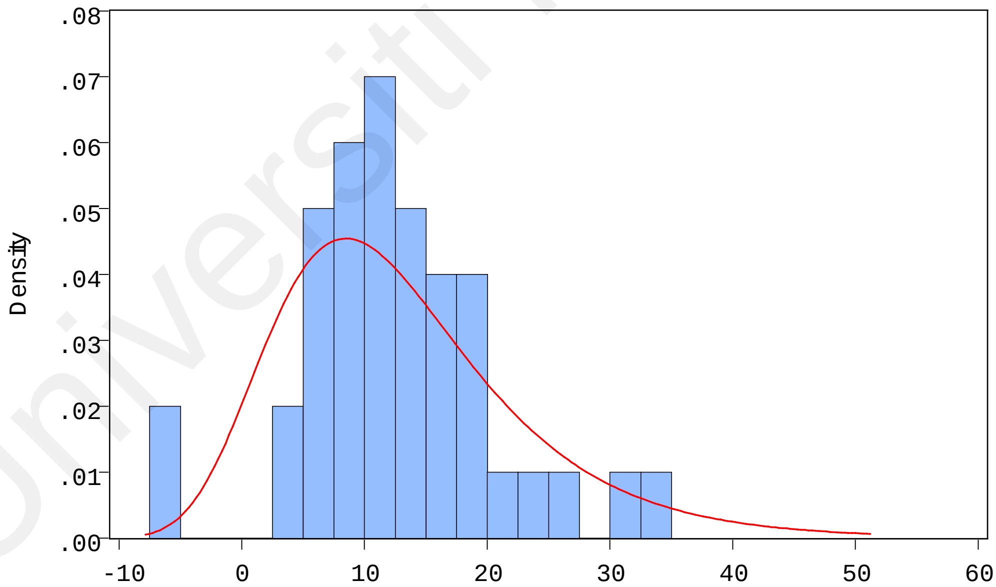


Figure 4.6: Histogram of return distribution, NIKKEI 225 futures market

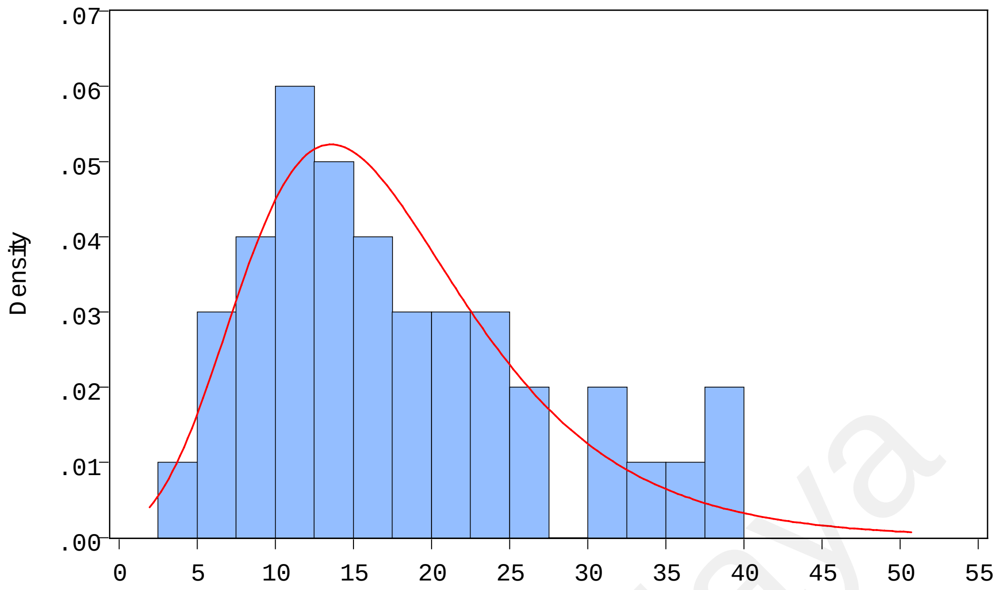


Figure 4.7: Histogram of return distribution, SiMSCI futures market

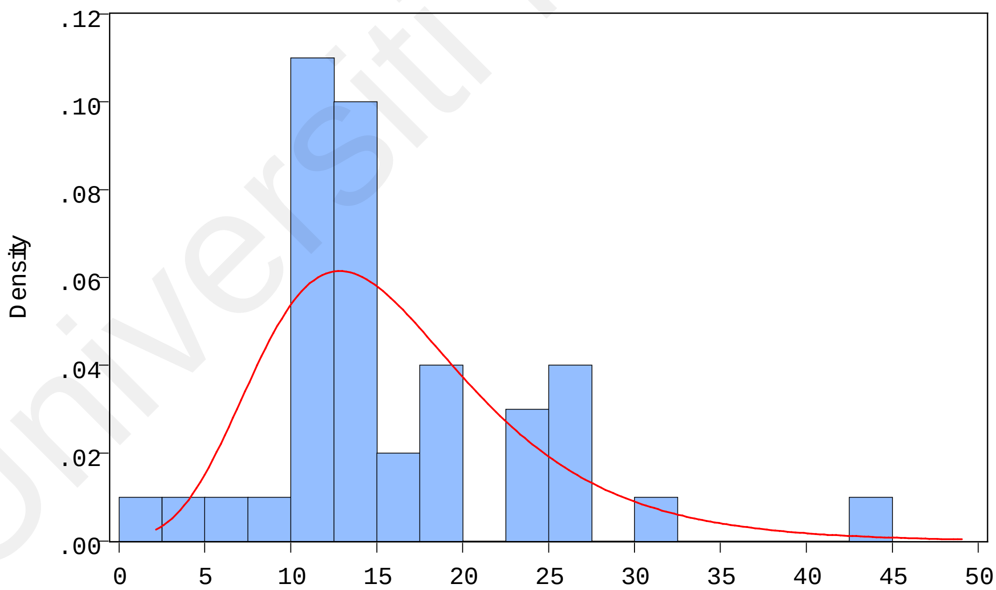


Figure 4.8: Histogram of return distribution, S&P 500 futures market

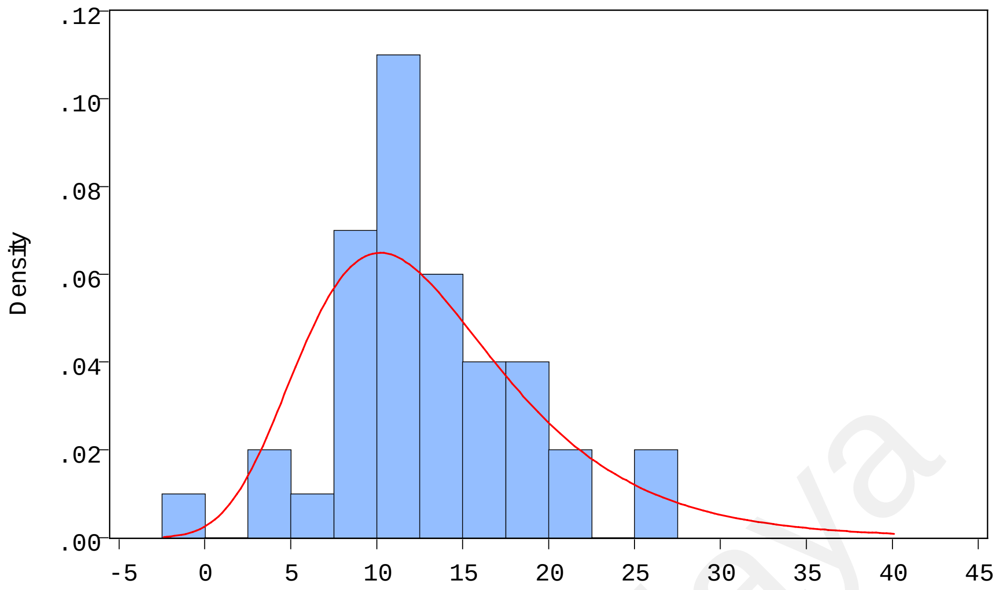


Figure 4.9: Histogram of return distribution, TAIEX futures market

If the outcome of the calculation is more than zero, the distribution is positively skewed. If the amount of the result is less than zero, the distribution is negatively skewed. If the outcome is equivalent to zero, this indicates that the distribution is symmetric. In terms of an explanation and analysis, the emphasis is on downside risk. Negatively skewed distributions have what mathematicians call a long left tail, which for researchers can mean a greater probability of largely negative results (Gencay, 1996). Positive skew would mean repeated minor negative results, and largely negative scenarios are not as probable. According to Table 4.7, the trading results of all futures markets are positively skewed, which means frequent small gains and negative results when the WPCA-NN model is applied. Moreover, extremely negative scenarios are not likely to happen based on the skewness results.

A kurtosis number is estimated in association with the normal distribution, for which kurtosis equals three. Since excess kurtosis equals kurtosis minus three, any positive value for excess kurtosis would mean fatter tails and lower risk of an extreme result. Regarding the kurtosis results presented in Table 4.7, all futures markets, except the KOSPI 200 (2.32) and SiMSCI (2.60), are leptokurtic, which means fatter tails and lower risk of an extreme outcome (Hang Seng (3.33); KLCI (3.59); NIKKEI 225 (3.81); S&P 500 (5.36); and TAIEX (3.93)). The return results for the KOSPI 200 (2.32) and SiMSCI (2.60) show kurtosis lower than three, which means thinner tails than other markets and higher risks of extreme returns from the WPCA-NN trading strategy.

Ultimately, the analysis of distribution for the trading returns of the WPCA-NN model displays a reliable distribution with repeated gains, small losses, and low probability for extremely negative scenarios for all futures markets. Regarding standard deviation and its effect on outcomes, the reliability of the results is tested with a Sharpe ratio in the next section.

4.7 Discussion of Results

4.7.1 Risks in Forecasting Results

When performing an algorithmic trading strategy, it is attractive to assume the annual return as the most suitable performance scale. However, there are many flaws with using the annual return separately. The computation of returns for definite strategies is not totally straightforward. This approach is particularly accurate for strategies that are not directional, such as strategies that make use of leverage or market-neutral variants; however, the parameters make it difficult to compare two different trading strategies based uniquely on their returns (Yao et al., 1999).

Moreover, if two trading strategies with equal returns are presented, how is it possible distinguish which one comprises greater risk? In addition, what is meant by "greater risk"? In finance, scholars and practitioners are regularly concerned with periods of drawdown and the volatility of returns. Thus, if one of these strategies possesses a meaningfully higher volatility of returns, it is probably less substantial, assuming that its historical data of returns are alike if they are not unique (Yao et al., 1999).

These issues of trading strategy comparison and risk valuation encourage the use of the Sharpe ratio (Mandelbrot et al., 2006). The Sharpe ratio is an amount for measuring risk-adjusted return developed by the Nobel laureate William Sharpe. It has become the standard calculation for such measurements in the industry and is the average return gained in addition to the risk-free rate for each unit of total risk or volatility (Yao et al., 1999). Deducting the risk-free rate from the average return, a performance related with risk-taking actions can be undertaken separately. One insight of this measurement is that a portfolio dealing with zero-risk investment, such as the purchase of United States treasury bills (with the expected return of a risk-free rate), possesses a Sharpe ratio of precisely zero. Commonly, when the value of a Sharpe ratio is high, the risk-adjusted return is more attractive (Mandelbrot et al., 2006).

The Sharpe ratio has become the most commonly used technique for computing risk-adjusted return. However, it may be imprecise when applied to assets or portfolios that do not possess a normal distribution of expected returns. Many assets possess negative skewness or a high degree of kurtosis that makes their distribution plots have fat tails.

Modern portfolio theory clarifies that increasing assets to a well-diversified portfolio that possesses correlations with each other that are lower than one reduces the risk of the portfolio without reducing the return. Such diversification can help to increase a portfolio's Sharpe ratio (Damodaran, 2016). Thus,

$$Sharpe\ ratio = \frac{Mean\ portfolio\ return - Risk\ free\ rate}{Standard\ deviation\ of\ portfolio\ return} = \frac{\bar{r}_p - r_f}{\sigma_p}, \tag{39}$$

where

$\bar{r}_p = Expected\ portfolio\ return,$

$r_f = Risk\ free\ rate,$

$\sigma_p = Portfolio\ standard\ deviation.$

The difference between ex post and ex ante Sharpe ratio formulas is that the ex post Sharpe ratio uses realized returns whereas the ex-ante Sharpe ratio applies expected returns.

The Sharpe ratio is regularly applied to monitor the change in a portfolio's total risk-return features once an asset class or new asset is attached to the portfolio. Moreover, the Sharpe ratio can help to clarify whether the excess returns of a portfolio are because of smart investment decisions or the outcome of high risk. Although one fund or portfolio can gain higher returns than its peers, it is simply a suitable investment if the aforementioned higher returns do not bring an increase in extra risk. When a portfolio's Sharpe ratio is higher, its risk-adjusted performance is better. A negative Sharpe ratio shows that an asset without risk achieves more than the security that was analyzed (Damodaran, 2016). Table 4.8 presents the results of the Sharpe ratios of the seven futures markets in the context of the trading strategies and models.

Table 4.8: Sharpe ratios with trading strategies and models, including the WPCA-NN model

Models	B & H	MACD	RSI	Stochastics	OMA	NN	WNN	WPCA-NN
Hang Seng futures	-0.10	-0.23	-0.40	0.08	-0.28	0.06	0.50	0.77*
KLCI futures	-0.33	-0.50	-0.51	-0.79	-0.44	0.21	0.85*	0.85*
KOSPI 200 futures	0.00	-0.39	-1.61	-1.54	-1.32	0.02	1.20	1.42*
NIKKEI 225 futures	0.01	-0.15	-0.04	-0.40	-0.43	0.33	1.16	1.44*
SiMSCI futures	0.03	0.06	-0.29	-0.72	0.37	0.36	1.09	1.57*
S&P 500 futures	-0.21	-0.42	-0.25	-0.23	-0.60	0.22	0.85	1.04*
TAIEX futures	0.07	-0.32	-0.18	-0.10	-0.19	0.08	1.28	2.02*

* The highest Sharpe ratio among the models.

Hence, the Sharpe ratio is applied in this study to evaluate the performance of the trading strategies in accordance with the risks and returns of their outcomes (seeTable 4.8). The Sharpe ratio represents how well a trading strategy performs in comparison with the return rate on a risk-free investment (Damodaran, 2016). The results are based on quarterly return results of the WPCA-NN model, with related standard deviation as the risk, and the risk-free rate of return. The higher the Sharpe ratio, the better the risk-adjusted performance and, consequently, the better the performance of the trading strategy on the specific futures market. According to the Sharpe ratio results, the performance of the WPCA-NN model is highly reliable and acceptable in most futures markets. The results show that the WPCA-NN model performs with higher risk in the Hang Seng (0.77) and KLCI (0.85) futures markets compared with the others because their Sharpe ratios are less than 1. However, the trading results of the Hang Seng and KLCI futures markets are still considered reliable because they are positive and close to 1. The WPCA-NN model achieves considerably reliable returns for the other futures markets according to their Sharpe ratios: KOSPI 200 (1.42), NIKKEI 225 (1.44), SiMSCI (1.57), S&P 500 (1.04), and TAIEX (2.02). Moreover, the Sharpe ratio results indicate that the cause of the returns achieved by the WPCA-NN model is the reliable and successful technique. In other words, the returns are not the outcomes of high risk.

Ultimately, the Sharpe ratios of the WPCA-NN model for the futures markets are considerably high, thus, the risk-adjusted performance of the WPCA-NN model is quite suitable. In order to compare the results of all models, we compare the means of the WPCA-NN model with those of the other models. This analysis is undertaken with a t-test. The results are reported in the next section.

4.7.2 T-Test and Comparison of Two Means

The differences between the return means of the WPCA-NN model and the other models are required to be meaningful and significant in order to report robust results. This section presents a t-test analysis that examines the differences between the return means.

Researchers may be interested in collecting data about two populations of information in order to compare them. In terms of the statistical implication for one population factor, confidence tests and intervals of significance are applicable to the statistical parameters in order to compare two population factors.

A confidence gap for the difference between two means provides a series of amounts along which this difference of the two populations of data may stand. The confidence interval for the difference between two means consists of all the amounts of $(\mu_1 - \mu_2)$ that would be accepted in the two-sided hypothesis analysis of H_0: $\mu_1 = \mu_2$against H_a: $\mu_1 \neq \mu_2$, namely, H_0: $\mu_1 - \mu_2 = 0$ against H_a:$\mu_1 - \mu_2 \neq 0$.

When the confidence interval contains 0, it is possible to conclude that there is no significant difference between the means of the two populations of data at an assumed level of confidence.

When there is a significant and meaningful difference between the means of two different return results from any two different trading models, it can be concluded that the model with the higher mean of return is superior.

A confidence interval for the difference between the means of the return of two different models indicates the superiority of one of them over the other because it is significant in terms of its p-value. The quarterly return mean differences between the WPCA-NN model and the other models are calculated with a paired t-test. The results are shown in Table 4.9.

Table 4.9: Satterthwaite-Welch t-test results for the differences in means

WPCA-NN vs.		HS futures	KLCI futures	KOSPI 200	NIKKEI 225	SiMSCI	S&P 500	TAIEX
WNN	T-Test	13.8	17.7	17.82	13.73	17.98	17.99	14.78
	Prob	0.0354	0.0332	0.0451	0.0336	0.0427	0.0334	0.0287
NN	T-Test	17.98	16.77	15.52	10.91	15.96	15.87	17.35
	Prob	0.0485	0.0031	0.0001	0.0062	0.0001	0.0002	0.0000
B & H	T-Test	17.67	17.6	10.20	13.30	13.70	14.68	11.83
	Prob	0.0279	0.0003	0.0279	0.0000	0.0017	0.0000	0.0062
MACD	T-Test	11.38	13.13	9.91	12.36	17.93	17.46	14.99
	Prob	0.002	0.0000	0.0359	0.0000	0.0000	0.0000	0.0000
RSI	T-Test	17.73	17.66	9.99	11.52	16.93	13.92	10.98
	Prob	0.0037	0.0000	0.009	0.0003	0.0000	0.0000	0.0077
Stochastics	T-Test	17.39	15.52	11.25	10.44	16.51	14.63	15.16
	Prob	0.0414	0.0000	0.0007	0.0005	0.0000	0.0000	0.0001
OMA	T-Test	12.84	17.99	11.03	9.91	17.11	16.13	17.94
	Prob	0.0014	0.0001	0.0017	0.0026	0.0009	0.0000	0.0000

* All results are significant at a 5% level of confidence.

According to the t-test results, it can be concluded that:

1) There are significant differences between the return means of the WPCA-NN model and those of the buy-and-hold strategy (the RWH outcome) (hypothesis H2).
2) There are significant differences between the return means of the WPCA-NN model and those of the MACD, RSI, stochastics, and OMA technical trading rules (hypotheses H3a, H3b, H3c, and H3d).
3) There are significant differences between the return means of the WPCA-NN and pure NN models.
4) There are significant differences between the return means of the WPCA-NN and WNN models.

The findings indicate that in all futures markets, the trading strategy using the WPCA-NN model generates a greater return than the RWH, the nonlinear autoregressive NN model, and the WNN model. Thus, the novel hybrid model does not support the RWH in the selected markets; moreover, its denoising procedure is superior to univariate wavelet denoising.

4.7.3 The Significant Alpha above Market Return

In addition to the t-test, an analysis of the significant alpha in a linear regression form of the WPCA-NN model's returns and the returns of the other models would contribute to the reliability of the results. Alpha is an amount of the active return regarding an investment when the investment's performance is compared with an appropriate market index. An alpha of 1 indicates that the return on an investment over a designated time is 1% better than the market for the same period. An alpha of -1 indicates that an investment underperformed compared with the market. In modern portfolio management theory, alpha is one of the five significant measures, the others being beta, R-squared, the Sharpe ratio, and standard deviation.

In the current financial markets, where indices are broadly accessible for purchasing, alpha is regularly used to evaluate the performance of similar investments and mutual funds. Since these funds contain various fees generally stated in percentage terms, each fund should have an alpha larger than its fees in order to deliver positive achievements compared with an index fund. Generally, significant numbers of traditional funds have possessed negative alphas, which have meant that capital transfers to non-traditional hedge funds and index funds.

In order to investigate a portfolio of investments and measure theoretical performance, it is also useful to apply the capital asset pricing model (CAPM) regularly. Returns on the said portfolio could be compared with theoretical returns, a comparison that is known as Jensen's alpha. This alpha is valuable for highly focused or non-traditional funds, while a sole stock index may not be representative of the holdings of an investment. The coefficient of alpha (α_i) in the CAPM is considered a significant factor. The alpha is the intercept of the security characteristic line (SCL), which is the coefficient of the constant in a market model regression form as follows.

$$SCL: \ R_{i,t} - R_f = \alpha_i + \beta_i\left(R_{M,t} - R_f\right) + \epsilon_{i,t}. \tag{40}$$

where

α_i is called the asset's alpha (abnormal return)

β_i is a nondiversifiable or systematic risk

$\epsilon_{i,t}$ is the non-systematic or diversifiable, non-market or idiosyncratic risk

$R_{M,t}$ is a market risk

R_f is a risk-free rate

The expected value of the alpha coefficient in the CAPM for an efficient market is zero. Hence, the alpha coefficient represents how an investment has performed after measuring for risk, and the alpha ranges are as follow:

$\alpha_i < 0$ where an investment has gained comparably little compared with the risk (or an investment is significantly risky in terms of the return).

$\alpha_i = 0$ where an investment has gained a satisfactory return compared with the risk that was taken.

$\alpha_i > 0$ where an investment has a return in excess of the reward for the presumed risk.

For example, although a return of 20% may seem appropriate, the investment may still possess a negative alpha if it is in an extremely risky position.

In modern financial markets, where futures are widely accessible for purchasing, alpha is commonly used to assess the performance of financial trading and similar investments.

This study establishes a simple linear regression of the returns of the WPCA-NN model against the returns of the buy-and-hold strategy. The alpha coefficient as an intercept of this regression should be positive and significant in order to show the superiority of the WPCA-NN model's result compared with the outcome of the buy-and-hold strategy. Thus, a regression of

$$R_{WPCA\_NN} = \alpha + \beta \times R_{Buy\&Hold} + \xi \qquad (41)$$

where

α is the trading's alpha (abnormal return)

β is the systematic risk

ξ is the non-systematic or diversifiable risk

R_{WPCA_NN} is the return of the WPCA-NN trading strategy

$R_{Buy\&Hold}$ is the return of the buy and hold trading strategy

is performed to check whether the alpha is positive and significant. Table 4.10 presents the regression results.

Table 4.10: Regression results for the WPCA-NN model and the buy-and-hold strategy

Futures Markets	Alpha Coefficients (α)	Prob.
Hang Seng	33.13	0.0043
KLCI	47.84	0.0001
KOSPI 200	42.56	0.0001
NIKKEI 225	43.63	0.0001
SiMSCI	54.26	0.0001
S&P 500	49.39	0.0001
TAIEX	43.57	0.0001

Note: All results are significant at the 5% level.

In order to check whether the WPCA-NN model outperforms the MACD, RSI, stochastics, and RSI trading strategies, four regressions are performed. Therefore, the following regressions (42, 43, 44 & 45) are performed to check whether the alpha is positive and significant. The results are presented in Table 4.11 to Table 4.14.

$$R_{WPCA\_NN} = \alpha + \beta \times R_{MACD} + \xi \tag{42}$$

$$R_{WPCA\_NN} = \alpha + \beta \times R_{RSI} + \xi \tag{43}$$

$$R_{WPCA\_NN} = \alpha + \beta \times R_{Stochastics} + \xi \tag{44}$$

$$R_{WPCA\_NN} = \alpha + \beta \times R_{SMA20} + \xi \tag{45}$$

where

α is the trading's alpha (abnormal return)

β is the systematic risk

ξ is the non-systematic or diversifiable risk

R_{WPCA_NN} is the return of the WPCA-NN trading strategy

R_{MACD} is the return of the MACD trading strategy

R_{RSI} is the return of the RSI trading strategy

$R_{Stochastics}$ is the return of the stochastics trading strategy

R_{OMA} is the return of the OMA trading strategy

Table 4.11: Regression results for the WPCA-NN model and the MACD trading strategy

Futures Markets	Alpha Coefficients (α)	Prob.
Hang Seng	34.22	0.0014
KLCI	47.72	0.0014
KOSPI 200	44.33	0.0000
NIKKEI 225	42.18	0.0000
SiMSCI	55.32	0.0002
S&P 500	48.30	0.001
TAIEX	45.39	0.0001

Note: All results are significant at the 5% level.

Table 4.12: Regression results for the WPCA-NN model and the RSI trading strategy

Futures Markets	Alpha Coefficients (α)	Prob.
Hang Seng	35.24	0.0017
KLCI	47.56	0.0001
KOSPI 200	43.32	0.0001
NIKKEI 225	42.48	0.0000
SiMSCI	55.04	0.0001
S&P 500	49.02	0.0000
TAIEX	45.63	0.0000

Note: All results are significant at the 5% level.

Table 4.13: Regression results for the WPCA-NN model and the stochastics trading strategy

Futures Markets	Alpha Coefficients (α)	Prob.
Hang Seng	42.54	0.0001
KLCI	40.13	0.0000
KOSPI 200	44.82	0.0001
NIKKEI 225	41.47	0.0001
SiMSCI	49.45	0.0000
S&P 500	48.28	0.0012
TAIEX	44.82	0.0000

Note: All results are significant at the 5% level.

Table 4.14: Regression results for the WPCA-NN model and the OMA trading strategy

Futures Markets	Alpha Coefficients (α)	Prob.
Hang Seng	34.62	0.0003
KLCI	46.26	0.0001
KOSPI 200	44.10	0.0001
NIKKEI 225	41.80	0.0001
SiMSCI	51.48	0.0001
S&P 500	47.89	0.0001
TAIEX	44.15	0.0001

Note: All results are significant at the 5% level.

The following two regressions are performed to check whether the WPCA-NN model outperforms the pure NN and WNN models. Table 4.15 and Table 4.16 display the results of the mentioned regressions and comparison between the models.

$$R_{WPCA\_NN} = \alpha + \beta \times R_{NN} + \xi \tag{46}$$

$$R_{WPCA\_NN} = \alpha + \beta \times R_{WNN} + \xi \tag{47}$$

where

α is the trading's alpha (abnormal return)

β is the systematic risk

ξ is the non-systematic or diversifiable risk

R_{WPCA_NN} is the return of the WPCA-NN trading strategy

R_{NN} is the return of the ANN trading strategy

R_{WNN} is the return of the WNN trading strategy

Table 4.15: Regression results for the WPCA-NN model and the pure NN model

Futures Markets	Alpha Coefficients (α)	Prob.
Hang Seng	36.60	0.0045
KLCI	27.95	0.0165
KOSPI 200	48.01	0.0044
NIKKEI 225	41.25	0.0001
SiMSCI	57.63	0.0001
S&P 500	49.42	0.0012
TAIEX	41.73	0.0001

Note: All results are significant at the 5% level.

Table 4.16: Regression results for the WPCA-NN model and the WNN model

Futures Markets	Alpha Coefficients (α)	Prob.
Hang Seng	27.78	0.0204
KLCI	14.87	0.0185
KOSPI 200	19.85	0.0307
NIKKEI 225	32.95	0.0012
SiMSCI	26.43	0.0045
S&P 500	31.32	0.0066
TAIEX	29.95	0.0008

Note: All results are significant at the 5% level.

According to Table 4.10 to Table 4.16, the alpha coefficients are significant and positive for all future markets. Thus, the WPCA-NN model generates significantly higher returns than the buy-and-hold benchmark (Table 4.10) regarding the positive and significant alphas for all futures markets ($\alpha_{HangSeng} = 33.13 > 0$, $\alpha_{KLCI} = 47.84 > 0$, $\alpha_{KOSPI} = 42.56 > 0$, $\alpha_{Nikkei} = 43.63 > 0$, $\alpha_{SiMSCI} = 54.26 > 0$, $\alpha_{S\&P500} = 49.39 > 0$, and $\alpha_{TAIEX} = 43.57 > 0$). According to these results, the markets do not follow random walk from 2005 to 2014. Moreover, the WPCA-NN model achieves significantly higher returns than the MACD, RSI, stochastics, and OMA trading strategies, according to the significant and positive alphas in their regression analyses (Table 4.11, Table 4.12, Table 4.13, and Table 4.14). In addition, the WPCA-NN model generates significantly higher returns than the pure NN and WNN models for all futures markets, according to the positive and significant alphas (Table 4.15 and Table 4.16). This superiority of the WPCA-NN model over the pure NN and WNN models is meaningful and significant, based on the results. Thus, the offered denoising process significantly and extremely increases the performance of forecasting because the regressions between the WPCA-NN model's returns and those of the NN model achieves positive and significant alphas. Additionally, the WPCA-NN model, with its multivariate denoising process, outperforms the WNN model, with its univariate denoising process, for all seven selected futures markets because of the positive and significant alphas reported for the regressions between the WPCA-NN model's returns and the WNN model's returns.

The results from Table 4.10 show that for all futures markets, the alphas are positive, extremely large, and significant. This means that in all futures markets, the WPCA-NN model generates significantly positive excess returns compared with the RWH. These results also suggest that it is possible to gain positive returns compared with a buy-and-hold strategy in futures markets and opens an opportunity to support the AMH. So far, the overall results and descriptive analyses of the outcomes are presented. The next section provides a detailed report of all outcomes for each futures market separately.

4.8 Research Results

In this section, the results of all models are reported for all seven futures markets in details. The forecasting results are discussed in the case of the forecasting performance and profitability. Moreover, the results are discussed around 2008 financial crisis.

4.8.1 Hang Seng Futures Market

Based on the results for the Hang Seng futures market presented in Table 4.17, the trained networks for all models are valid because their MAPE values are quite low and acceptable. The WPCA-NN model outperforms the WNN and NN models compared with the whole out-sample period (4.06 < 15.44 < 579.4). Indeed, the lower the MAPE ratio, the better. Moreover, the WPCA-NN model has the highest return (34.8), without transaction costs, among the buy-and-hold strategy (7.4); the MACD (2.8), RSI (-3.1), stochastics (13.3), and OMA (1.2) strategies, and the NN (12.6) and WNN (27.9) models, based on the results shown in Table 4.18. The NN model performs very poorly, with an error term of 579.4% compared with the other models. However, the results of the WPCA-NN and WNN models are more precise, with much lower error terms, from 2005 to 2014.

Table 4.17: Performance of the models and strategies measured by the MAPE ratios of the evaluation results for the Hang Seng futures market

Models	Period	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
WPCA-NN	In-sample	0.033*	0.037*	0.11*	0.039*	0.035*	0.033*	0.034*	0.027*	0.023*	0.023*	0.039*
	Out-sample	3.3*	3.72*	11*	3.929*	3.5*	3.342*	3.354*	2.711*	2.296*	3.495*	**4.065***
WNN	In-sample	0.034	0.056	0.512	0.416	0.116	0.083	0.164	0.057	0.051	0.038	0.153
	Out-sample	3.386	5.588	51.25	41.552	11.646	8.257	16.391	5.729	5.095	5.533	**15.443**
NN	In-sample	1.138	1.025	1.279	4.913	8.523	8.712	7.38	3.877	186.42	2.715	22.598
	Out-sample	99.19	245.13	1539.3	1598.01	621.58	379.93	615.35	214.39	230.25	250.88	**579.4**

*The best performance among the models.

Table 4.18: Returns of the models and strategies without transaction costs (in percentages) for the Hang Seng futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
B & H	5.2	31.4	36.6	-46.2	43*	4.1	-21.0	20.4	-0.2	1.4	**7.4**
MACD	-4.4	10.0	9.6	17.2	-6.9	8.4	-4.7	8.1	-2.1	-6.7	**2.8**
RSI	0.1	-25.7	47.6	-40.2	-38.4	-0.2	-8.1	9.6	19.2	5.3	**-3.1**
Stochastics	16.9	-17.9	35.2	-13.3	30.1	11.7	41*	-1.8	13.8	17.4	**13.3**
OMA	-9.4	2.8	11.0	6.8	1.9	7.8	-20.8	18.8	-2.0	-5.3	**1.2**
NN	-8.3	-0.8	-28.1	50.7*	3.0	12.4*	11.0	40.0	24.1	21.7	**12.6**
WNN	51.1	52.2	131.6*	-26.3	15.4	-8.1	-1.5	25.5	31.7	7.1	**27.9**
WPCA-NN	55.4*	76.1*	35.6	23.0	10.6	3.1	10.4	41.1*	54.5*	38.8*	**34.8***

*The highest return among the models.

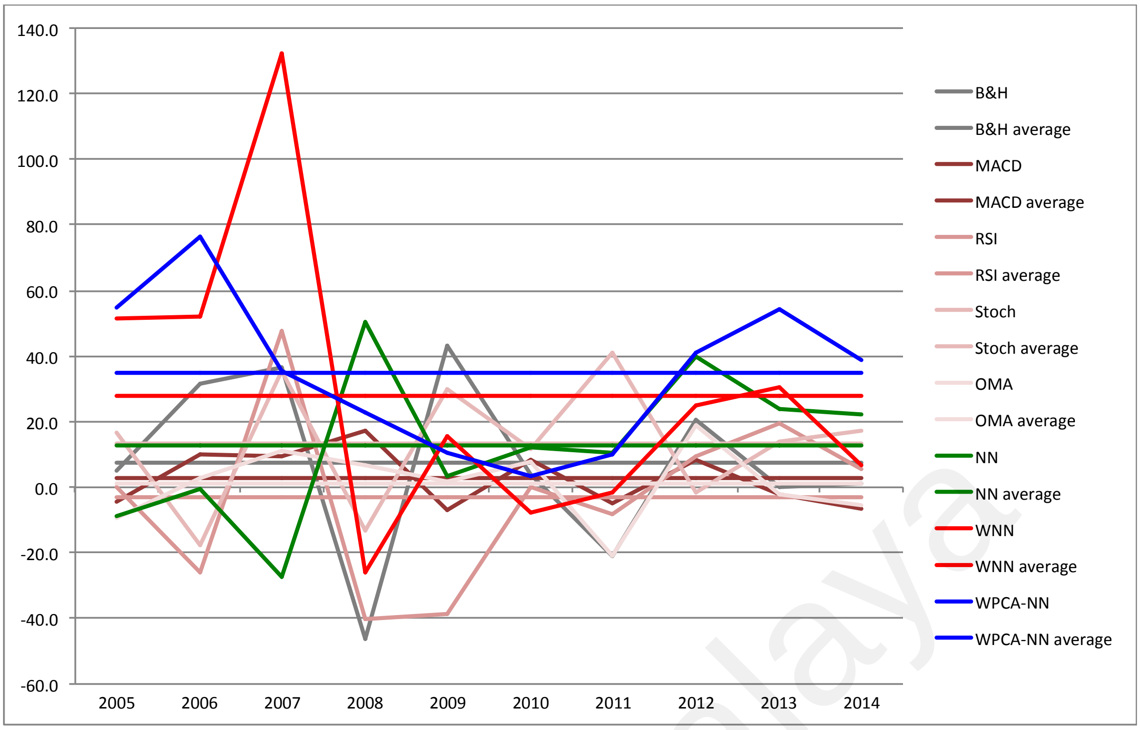


Figure 4.10: Returns of the models without transaction costs: the results for the Hang Seng futures market

Table 4.19 presents the returns of the models and strategies with transaction costs. The results confirm the same ranking for the models. Figure 4.10 and Figure 4.11 illustrate the return results of all the models and strategies without transaction costs and with transaction costs respectively. According to the best-performing networks in the Hang Seng futures market (Table 4.6), three levels of decomposition, a coif5 wavelet, and a penalized high thresholding strategy achieve significantly high performance and excess returns with the WPCA-NN model.

Table 4.19: Returns of the models and strategies with transaction costs: the results for the Hang Seng futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	**Average**
B & H	5.1	31.2	36.4	-46.2	42.8*	3.9	-0.9	20.2	-0.3	1.2	**9.3**
MACD	-7.3	7.1	7.0	13.7	-10.2	5.5	-0.2	4.9	-4.9	-10.1	**0.5**
RSI	-0.2	-25.9	46.4	-40.4	-38.6	-0.5	-0.3	9.0	18.6	4.8	**-2.7**
Stochastics	14.1	-19.7	32.2	-14.3	26.9	9.9*	2.2	-3.9	11.3	14.6	**7.3**
OMA	-13.7	-1.9	5.9	3.1	-2.6	3.3	-1.1	14.5	-6.1	-9.0	**-0.7**
NN	-14.9	-2.9	-33.8	45.2*	-5.5	9.4	7.8*	35.2*	18.2	19.5	**7.8**
WNN	40.7	37.3	122.2*	-36.1	5.8	-17.0	-11.8	13.0	21.7	-4.2	**17.2**
WPCA-NN	40.8*	64.2*	20.5	10.8	1.5	-9.7	-5.0	28.1	41.5*	24*	**21.7***

* The highest return among the models (the returns are in percentages).

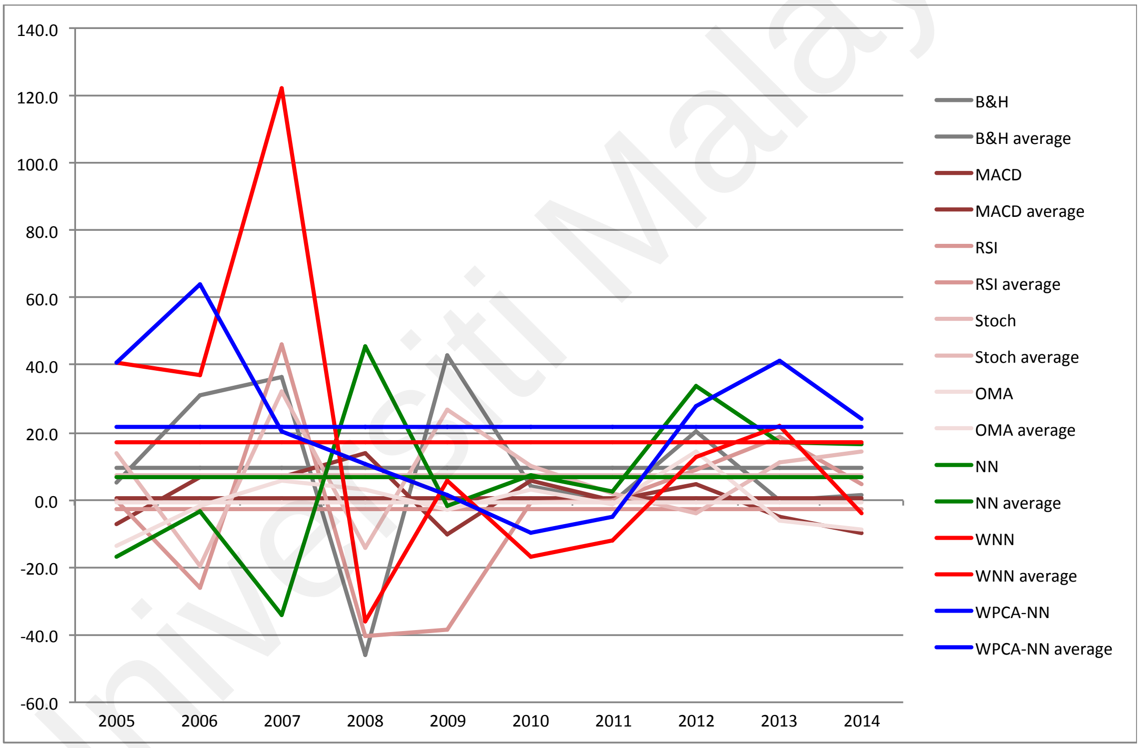


Figure 4.11: Returns of the models and strategies with transaction costs: the results for the Hang Seng futures market

Figure 4.12 displays the forecasting results of all intelligent models for 2005-2014. A prior study applied a generalized hyperbolic distribution in order to forecast the Hang Seng index and achieved a 5.3% annual return (Necula, 2009). Another study performed a hierarchical coevolutionary fuzzy predictive model and gained a 14.25% return (Huang et al., 2009).

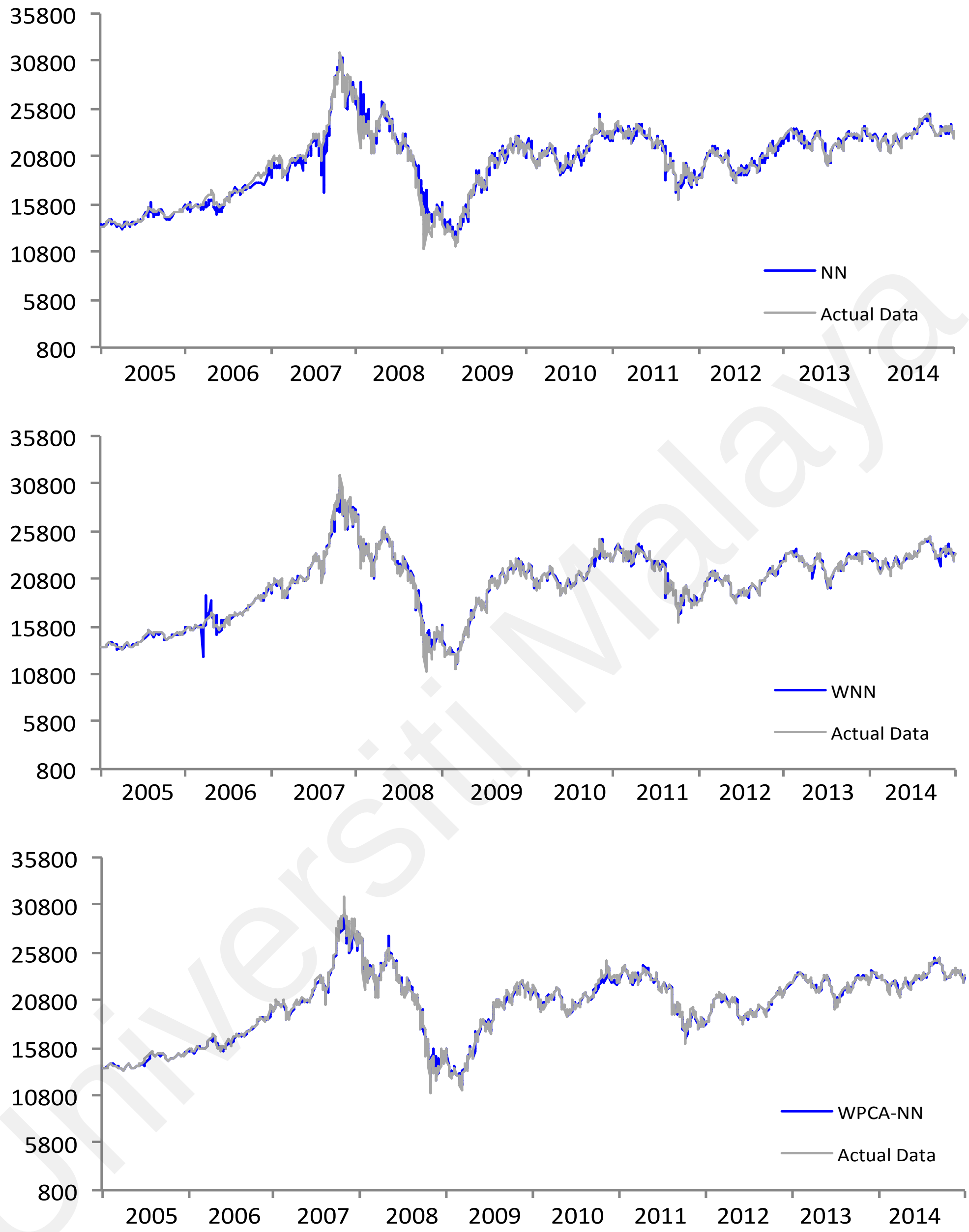


Figure 4.12: Forecasting results of NN, WNN and WPCA-NN models for the Hang Seng futures market in out-sample data

4.8.2 KLCI Futures Market

Based on the results of the KLCI futures market presented in Table 4.20 and Table 4.21, the trained networks for all models are valid because their MAPE values are less than 5%. Moreover, the WPCA-NN model outperforms the WNN and NN models (2.17 < 2.5 < 17.27) according to the MAPE evaluation, and achieves the highest return (47.2) among the buy-and-hold strategy (6.2), the MACD (1.3), RSI (1.1), stochastics (-6.9), and OMA (3) trading strategies; and the NN (21.7) and WNN (40.2) models, according to the annual returns.

Table 4.20: Performance of the models and strategies measured by the MAPE ratios of the evaluation results for the KLCI futures market

Models	Period	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
WPCA-NN	In-sample	0.006*	0.004*	0.048	0.028	0.013	0.026*	0.023*	0.015*	0.022*	0.021	0.021*
	Out-sample	0.589*	0.427*	4.815	2.755	1.312	2.595*	2.253*	1.502*	2.237*	3.17	**2.166***
WNN	In-sample	0.018	0.028	0.031*	0.022*	0.013*	0.047	0.025	0.02	0.028	0.012*	0.024
	Out-sample	1.847	2.752	3.11*	2.222*	1.26*	4.688	2.545	1.967	2.825	1.793*	**2.501**
NN	In-sample	0.037	0.028	0.038	0.099	0.123	0.172	0.144	0.099	0.057	0.046	0.084
	Out-sample	3.15	16.5	52.66	25.57	7.44	32.79	11.8	5.86	11.38	5.58	**17.27**

* The best performance among the models.

Table 4.21: Returns of the models and strategies without transaction costs: the results for the KLCI futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
B & H	0.6	-2.8	29.2	-38.6	40.2	18.7	-2.0	9.8	11.2	-4.8	**6.2**
MACD	10.6	4.8	-12.3	8.9	-13.7	6.0	8.4	3.8	4.7	-8.5	**1.3**
RSI	5.2	1.1	12.5	-21.8	-32.1	9.8	15.4	18.2	1.9	0.7	**1.1**
Stochastics	-3.8	3.7	-15.0	-24.4	-28.2	-8.7	-2.8	0.4	9.2	0.3	**-6.9**
OMA	8.3	-20.6	-18.8	47.3	2.8	0.4	9.3	9.0	-4.7	-2.7	**3.0**
NN	21.9	4.0	17.2	48.8	32.5	5.3	37.8	16.7	14.8	17.9	**21.7**
WNN	12.0	35.4	66.5	79.5*	54.9	29.3*	40.9	20.0	38.8*	24.3	**40.2**
WPCA-NN	33.6*	43.5*	75.3*	76.5	57*	27.3	59.7*	31.3*	37.2	30.7*	**47.2***

* The highest return among the models.

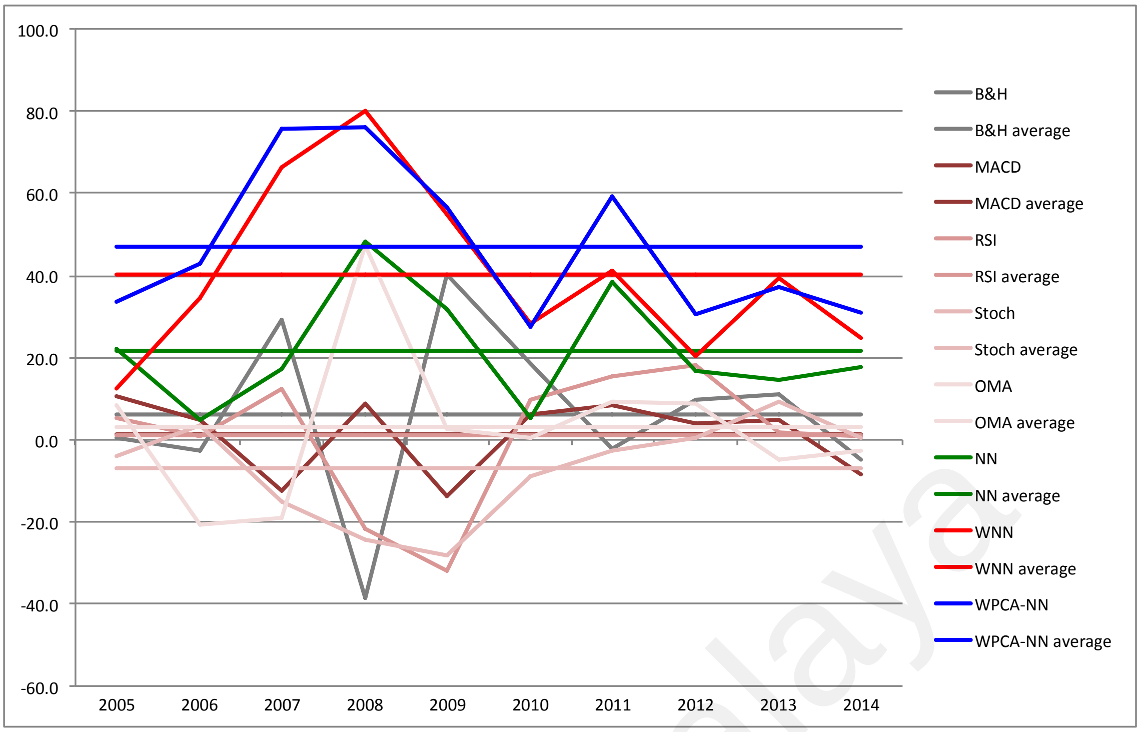


Figure 4.13: Returns of the models and strategies without transaction costs: the results for the KLCI futures market

In addition, Table 4.22 presents the returns of the models and strategies with transaction costs. The findings confirm the same ranking of the models and the superiority of the WPCA-NN model to the others. Figure 4.13 and Figure 4.14 illustrate the return results of all models and strategies without and with transaction costs respectively.

Table 4.22: Returns of the models and strategies with transaction costs: the results for the KLCI futures market

Models and strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	**Average**
B & H	0.5	31.2	37.5	-41.7	42.3	30.3*	-5.2	12.6	2.6	-10.8	**9.9**
MACD	7.6	14.9	-19.8	19.2	5.3	12.8	-16.7	-0.4	17.2	-0.3	**4.0**
RSI	4.9	-10.3	11.4	-25.5	-35.1	30.1	2.5	11.9	10.4	-2.9	**-0.3**
Stochastics	-5.2	-21.1	-9.1	-29.0	-20.1	-16.0	2.3	-4.1	20.0	-1.9	**-8.4**
OMA	3.2	23.0	-19.2	34.0	20.0	11.0	-15.7	-3.7	4.7	6.6	**6.4**
NN	16.4	-3.5	11.0	43.3	25.1	1.7	30.2	8.4	7.7	12.8	**15.3**
WNN	2.5	25.3	53.8	68.6*	46.3*	14.1	31.7	10.6	27.8*	9.6	**29.0**
WPCA-NN	18.6*	34.2*	64.8*	66.6	43.6	19.1	48.2*	19.9*	22.6	17.3*	**35.5***

* The highest return among the models.

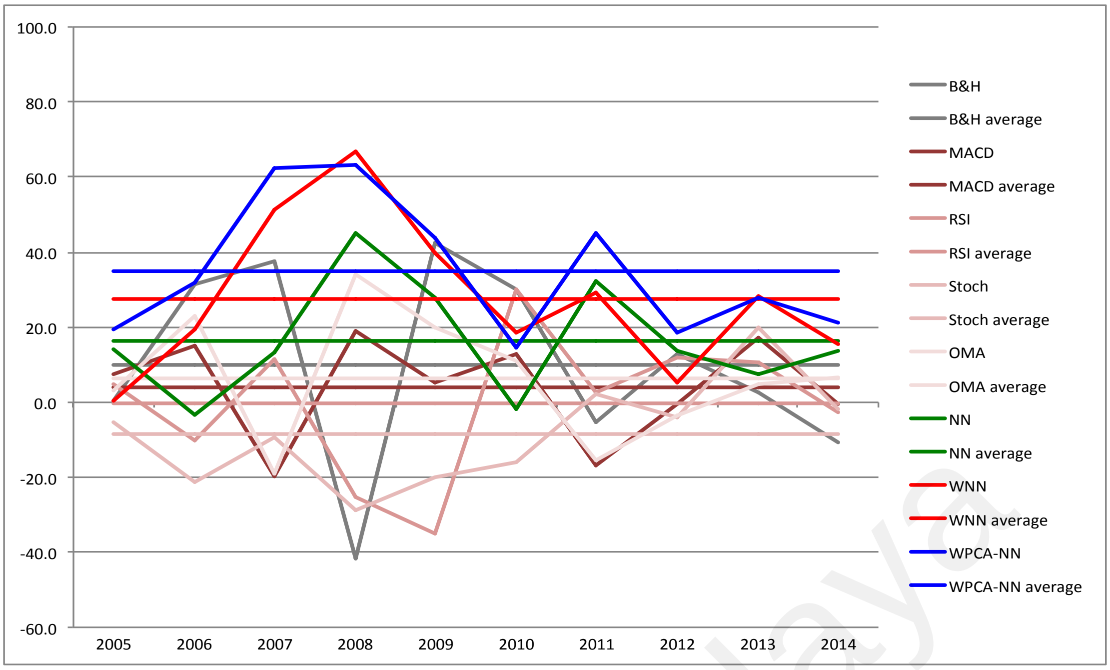


Figure 4.14: Returns of the models and strategies with transaction costs: the results for the KLCI futures market

According to the best-performing networks in the KLCI futures market (Table 4.6), three levels of decomposition, a db9 wavelet, a penalized high thresholding strategy, and a delay of five days achieve considerably high performance and returns with the WPCA-NN model. Although these results are robust and valid in different evaluation subsets and years, they relate specifically to the characteristics of the KLCI futures market and may change in the future. However, the foregoing combination is currently the most appropriate for KLCI forecasting; further, any other combination of these parameters performs comparably better than others because the parameters represent common characteristics of the market across 10 years. Figure 4.15 displays the forecasting results of all intelligent models for 2005-2014. In addition, a prior study applied an ANN and ARIMA to the KLCI futures market based on technical indicators and achieved average annual returns of 20.8% and 15.29% respectively (Yao et al., 1999).

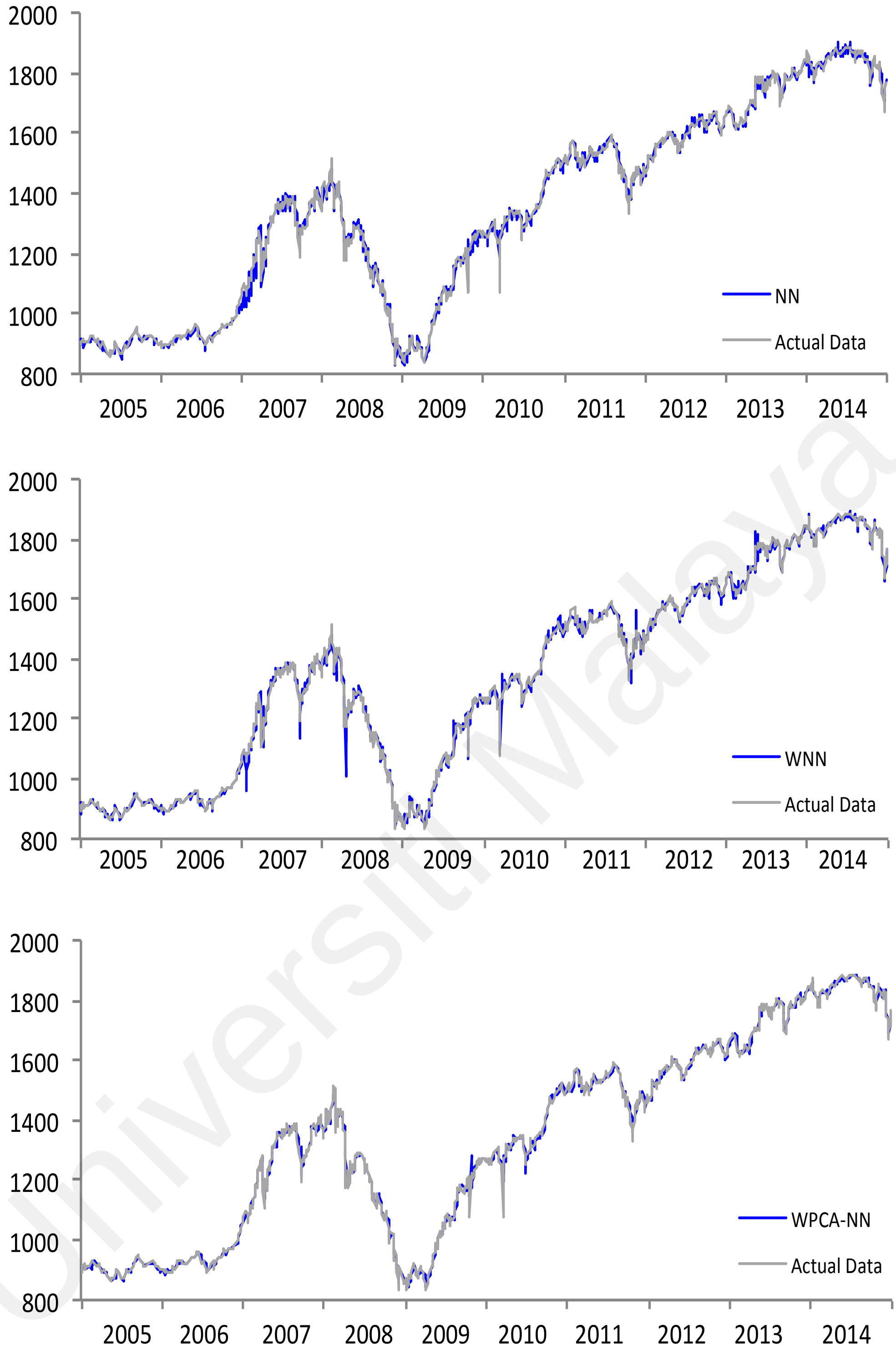


Figure 4.15: Forecasting results of NN, WNN and WPCA-NN models for the KLCI futures market in out-sample data

4.8.3 KOSPI 200 Futures Market

According to the MAPE ratio evaluation in Table 4.23, the WPCA-NN model outperforms the WNN and NN models (2.24 < 2.8 < 41.99) for the KOSPI 200 futures market. Moreover, the profitability of the WPCA-NN model (44.8) is higher than the returns of the buy-and-hold strategy (17.5); the MACD (9.9), RSI (-13.8), stochastics (-12.5), and OMA (-8.3) trading strategies; and the NN (17.5) and WNN (38) models, as shown in Table 4.24. Moreover, the results of the networks of the WPCA-NN and WNN models are valid and reliable for the period considered in the study.

Table 4.23: Performance of the models and strategies measured by the MAPE ratios of evaluation results for the KOSPI 200 futures market

Models	Period	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
WPCA-NN	In-sample	0.023	0.021*	0.026*	0.03*	0.023*	0.021*	0.029*	0.022	0.017	0.01*	0.022*
	Out-sample	2.348	2.111*	2.621*	3.027*	2.27*	2.059*	2.876*	2.231	1.687	1.2*	**2.243***
WNN	In-sample	0.023*	0.024	0.036	0.034	0.03	0.032	0.045	0.022*	0.017*	0.011	0.028
	Out-sample	2.338*	2.429	3.637	3.374	3.019	3.157	4.542	2.22*	1.677*	1.652	**2.805**
NN	In-sample	0.162	0.137	0.142	0.248	0.382	0.39	0.321	4.352	0.249	0.206	0.659
	Out-sample	33.07	30.43	93.36	68.6	31.02	27.42	89.56	23.73	14.37	8.3	**41.99**

* The best performance among the models.

Table 4.24: Returns of the models and strategies without transaction costs: results for the KOSPI 200 futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
B & H	100*	0.0	25.0	-60.0	100*	25*	-20.0	25.0	0.0	-20.0	**17.5**
MACD	33.3	66.7*	5.0	100*	-50.0	-55.0	-28.0	87.5*	-40.0	-20.0	**9.9**
RSI	-100.0	66.7*	5.0	-60.0	-100.0	25*	20.0	25.0	0.0	-20.0	**-13.8**
Stochastics	0.0	-50.0	47.0	-40.0	-66.7	25*	20.0	-40.0	0.0	-20.0	**-12.5**
OMA	-85.2	0.0	16.7	20.0	0.0	-25.0	-20.0	12.5	50*	-52.0	**-8.3**
NN	26.7	23.6	-3.7	18.7	26.8	15.5	14.4	16.3	16.0	20.9	**17.5**
WNN	49.6	49.7	57.0	22.6	57.3	23.2	27.9	35.3	23.5	33.9	**38.0**
WPCA-NN	52.0	40.9	61.8*	52.2	71.6	20.6	43.6*	47.6	21.5	35.9*	**44.8***

* The highest return among the models.

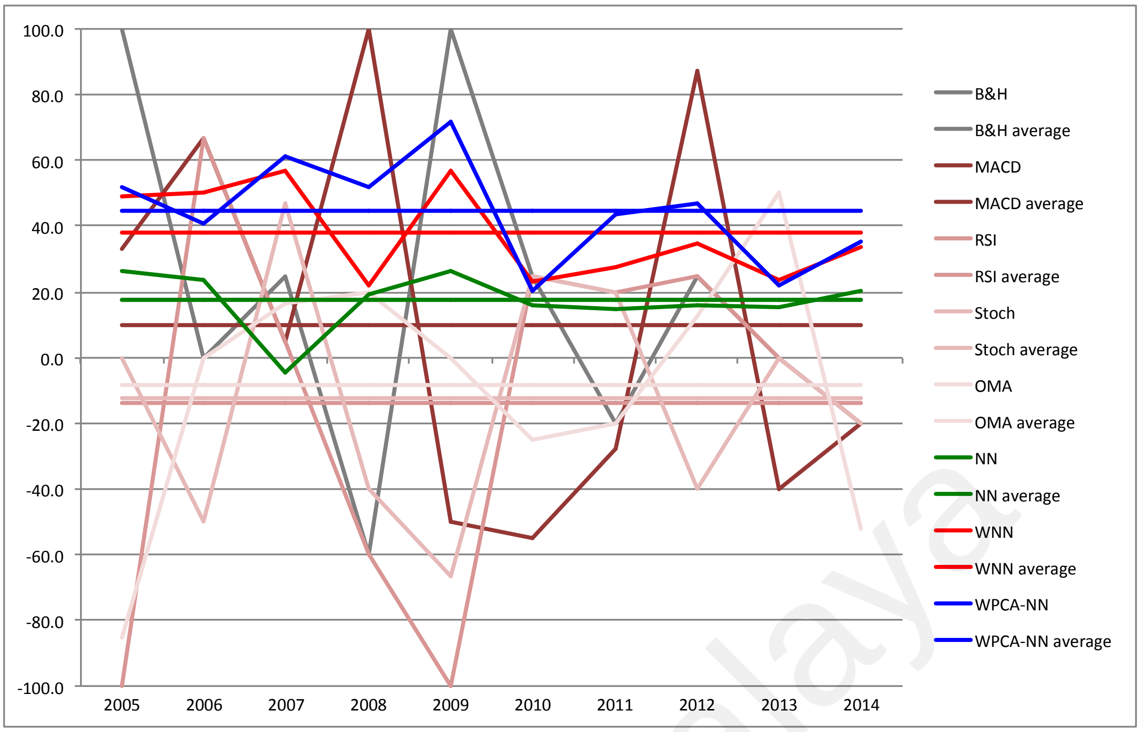


Figure 4.16: Returns of the models and strategies without transaction costs: the results for the KOSPI 200 futures market

In addition, Table 4.25 presents the returns of the models and strategies with transaction costs. The findings confirm the same ranking of the models and also confirm the higher return of the WPCA-NN model compared with the buy-and-hold strategy and other models. Figure 4.16 and Figure 4.17 show the return results of all models and strategies without and with transaction costs respectively. Based on the best-performing networks in the KOSPI 200 futures market (Table 4.6), three or four levels of decomposition, sym6 or db9 wavelets (For more information refer to Table 3.5), and a penalized high thresholding strategy achieve significantly high performance and excess returns with the WPCA-NN model.

Table 4.25: Returns of the models and strategies with transaction costs: the results for the KOSPI 200 futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	**Average**
B & H	99.7*	-0.2	24.8	-60.1	99.7*	24.8*	99.7*	24.8	-0.2	-20.1	**29.3**
MACD	29.4	61.5	0.9	95.1*	-51.7	-55.9	-51.7	81.9*	-42.1	-22.5	**4.5**
RSI	-100.2	65.9*	4.3	-60.1	-100.2	24.4	-100.2	24.6	-0.5	-20.4	**-26.2**
Stochastics	-1.5	-51.5	44.2	-40.9	-67.5	23.0	-67.5	-41.2	-1.9	-21.9	**-22.7**
OMA	-85.7	-4.1	12.4	15.3	-5.7	-28.3	-5.7	6.7	43.9*	-54.6	**-10.6**
NN	19.4	20.5	-7.4	11.9	20.0	11.9	10.3	12.0	8.8	16.7	**12.4**
WNN	36.2	35.4	45.4	10.8	43.7	10.1	17.3	26.3	11.2	20.8	**25.7**
WPCA-NN	39.4	31.6	50.8*	39.3	61.4	5.1	32.2	32.8	7.5	21.5*	**32.2***

* The highest return among the models.

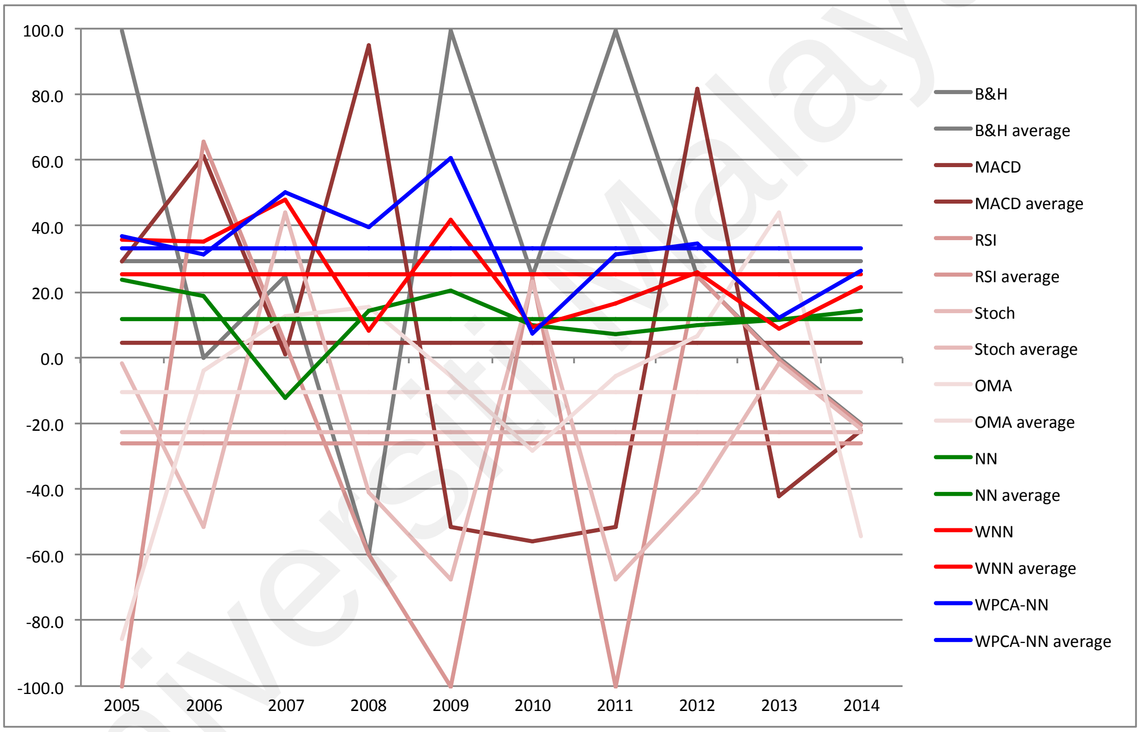


Figure 4.17: Returns of the models and strategies with transaction costs: the results for the KOSPI 200 futures market

Figure 4.18 displays the forecasting results of all intelligent models for 2005-2014. A prior study applied an ANN and case-based reasoning, and achieved an annual return of 40.9% on the KOSPI 200 index (Kim et al., 1998). A further study achieved 28.57% using a real-time rule-based trading system (Lee et al., 2010).

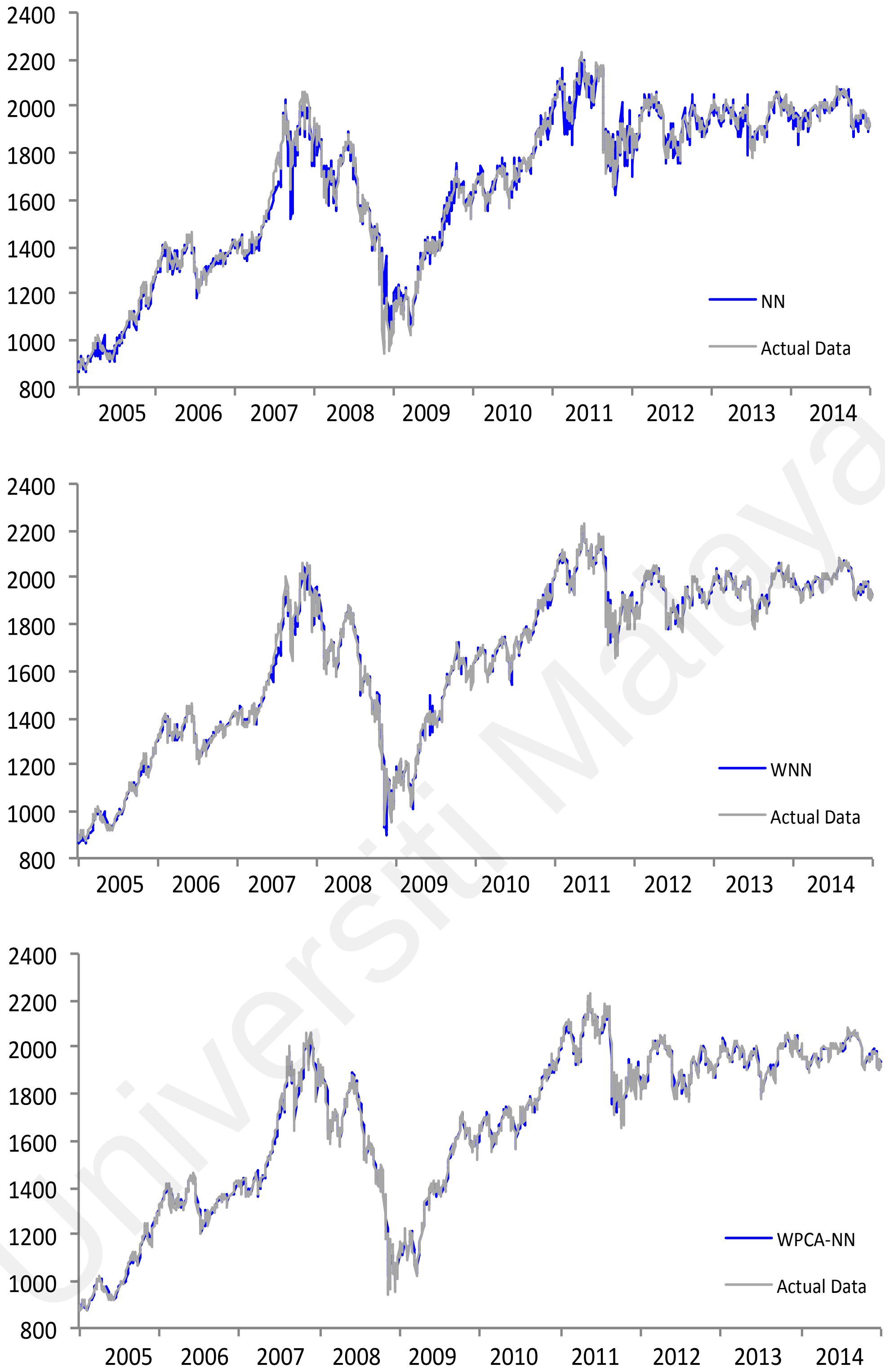


Figure 4.18: Forecasting results of NN, WNN and WPCA-NN models for the KOSPI 200 futures market in out-sample data

4.8.4 NIKKEI 225 Futures Market

Based on the outcomes presented in Table 4.26 and Table 4.27, the NIKKEI 225 futures market also confirms the superiority of the WPCA-NN model over the WNN and NN models, the buy-and-hold strategy, and other technical trading systems. The WPCA-NN model's evaluation performance is better than those of the WNN and NN models (4.05 < 10.88 < 451.89). Further, the annual return (42.1) is the highest among the buy-and-hold strategy (4.9); the MACD (0.7), RSI (3.7), stochastics (-5.6), and OMA (-6.4) trading strategies; and the NN (13.3) and WNN (34.9) models. Moreover, the WPCA-NN and WNN models are valid in the training and evaluation periods and achieve considerably higher returns than the NN model, the buy-and-hold strategy, and other technical trading strategies.

Table 4.26: Performance of the models and strategies measured by the MAPE ratio of evaluation results for the NIKKEI 225 futures market

Models	Period	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
WPCA-NN	In-sample	0.027*	0.032*	0.033*	0.084*	0.028*	0.027*	0.031*	0.032*	0.079*	0.018*	0.039*
	Out-sample	2.741*	3.233*	3.283*	8.439*	2.782*	2.724*	3.13*	3.187*	7.856*	3.124*	**4.05***
WNN	In-sample	0.115	0.077	0.063	0.29	0.06	0.034	0.064	0.035	0.26	0.089	0.109
	Out-sample	11.512	7.673	6.31	29.03	6.036	3.362	6.389	3.463	26.013	9.05	**10.884**
NN	In-sample	1.591	1.433	1.443	2.442	3.546	11.726	3.517	1.676	1.615	2.022	3.101
	Out-sample	395.82	391.55	320.73	1378.62	228.21	211.1	277.77	128.94	814.43	371.72	**451.89**

* The best performance among the models.

Table 4.27: Returns of the models and strategies without transaction costs: the results for the NIKKEI 225 futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
B & H	27.3	7.1	-6.7	-23.1	10.0	0.0	-8.3	9.1	33.3	0.0	**4.9**
MACD	-27.7	-11.1	-15.0	-13.6	27.3	-3.3	15.2	7.4	-3.0	30.6	**0.7**
RSI	-20.0	23.6	6.7	-23.1	12.5	9.1	36.4	18.2	-33.3	7.1	**3.7**
Stochastics	-27.0	61.7*	-14.2	-34.9	-26.7	-17.9	-11.9	-10.5	-5.7	30.6	**-5.6**
OMA	-30.3	-22.0	-34.9	-59.7	52.8*	-3.0	-37.2	7.4	13.5	49.3*	**-6.4**
NN	-36.5	45.4	28.0	5.1	8.9	25.6	-15.6	34.2	0.1	37.4	**13.3**
WNN	33.4*	18.9	46.7	51.1	46.8	40.0	26.4	33.1	50.5*	1.8	**34.9**
WPCA-NN	31.1	30.2	56.3*	53.1*	39.3	46.4*	39.2*	40.5*	42.9	41.9	**42.1***

* The highest return among the models.

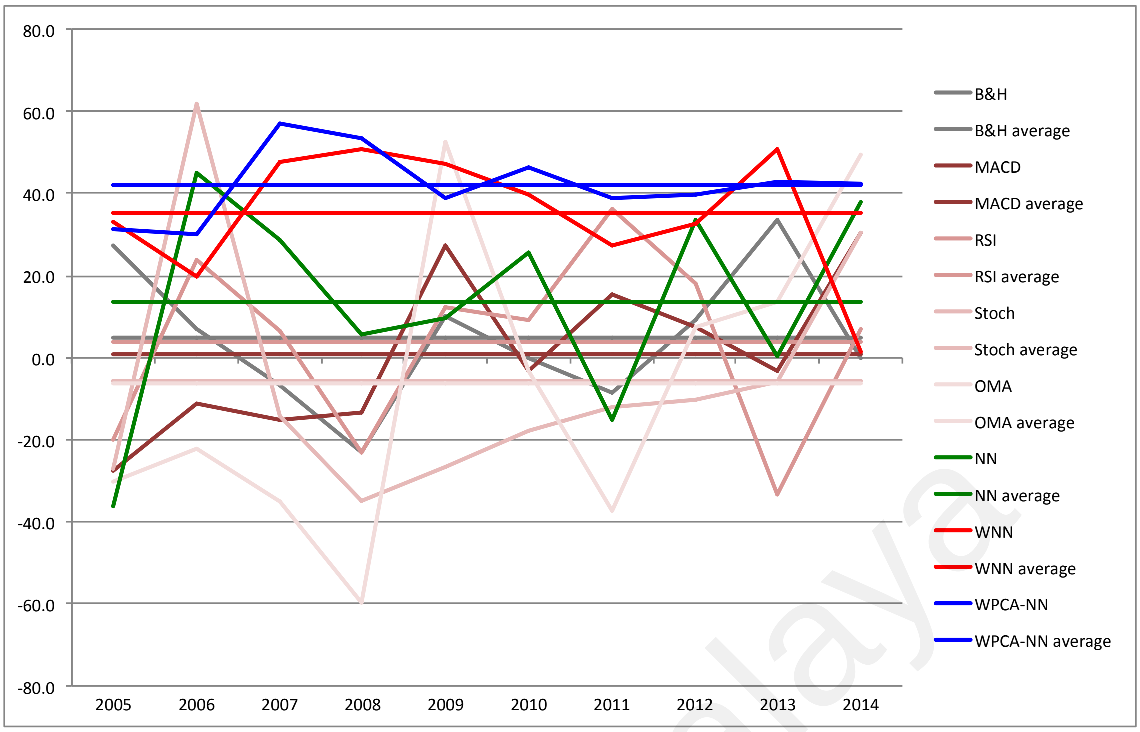


Figure 4.19: Returns of the models and strategies without transaction costs: the results for the NIKKEI 225 futures market

In addition, Table 4.28 presents the returns of the models and strategies with transaction costs. The findings confirm the same ranking of the models. Figure 4.19 and Figure 4.20 display the return results of all models and strategies without and with transaction costs respectively. According to the best-performing networks in the NIKKEI 225 futures market (Table 4.6), two levels of decomposition, a coif5 wavelet, and a penalized high thresholding strategy achieve considerably high performance and returns with the WPCA-NN model.

Table 4.28: Returns of the models and strategies with transaction costs: the results for the NIKKEI 225 futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	**Average**
B & H	27.1	7.0	-6.8	-23.2	9.9	-0.1	-8.4	9.0	33.2	-0.1	**4.8**
MACD	-29.3	-12.7	-17.6	-15.5	25.0	-5.4	13.1	5.5	-5.1	27.3	**-1.5**
RSI	-20.2	23.4	6.6	-23.2	12.3	9.0	36*	17.8	-33.4	7.0	**3.5**
Stochastics	-27.8	59.4*	-15.5	-35.4	-27.2	-19.2	-12.9	-11.5	-6.8	28.7	**-6.8**
OMA	-32.0	-23.7	-37.4	-60.9	50.4*	-4.9	-39.5	5.1	9.3	42.3*	**-9.1**
NN	-40.0	37.6	25.0	2.5	1.6	19.3	-23.8	30.2	-7.3	33.7	**7.9**
WNN	22.9	4.0	32.9	39.8	36.7	25.5	18.0	19.5	39.4*	-12.8	**22.6**
WPCA-NN	22.0	16.1	46.7*	42.2*	26.2	34.2*	25.1	28.5	31.2	26.9	**29.9***

* The highest return among the models.

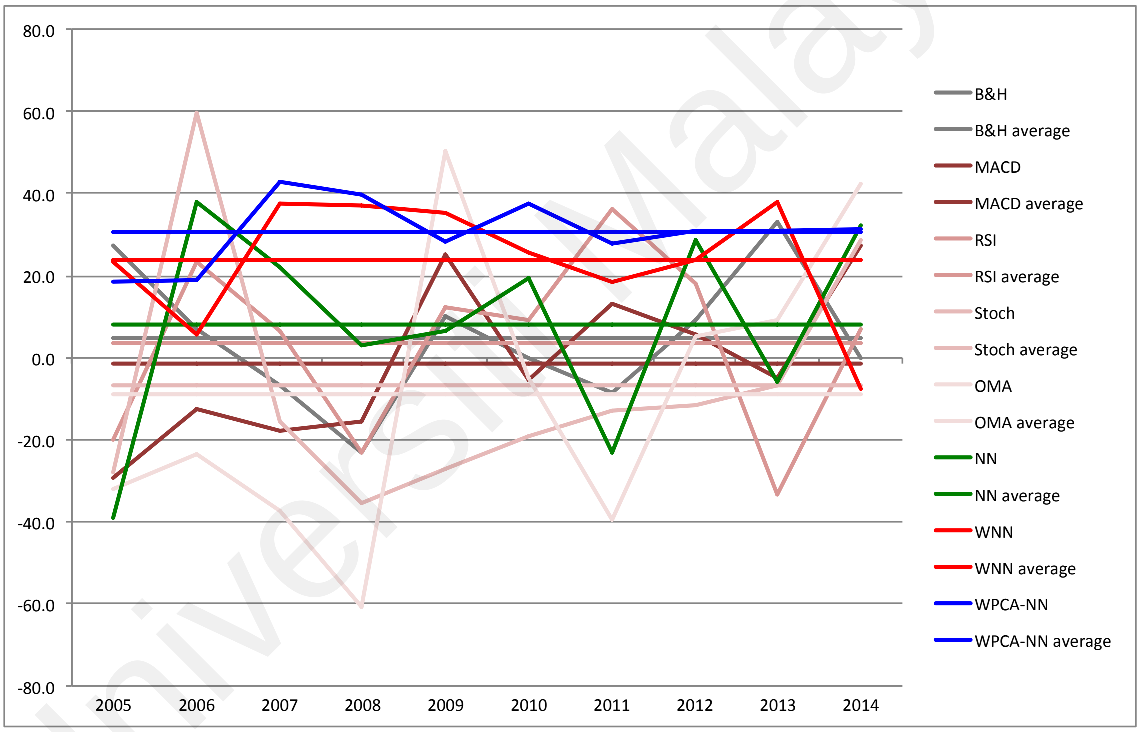


Figure 4.20: Returns of the models and strategies with transaction costs: the results for the NIKKEI 225 futures market

Figure 4.21 displays the forecasting results of all intelligent models for 2005-2014. In addition, a prior study on forecasting the Nikkei index achieved a 17.2% annual return with discriminant analysis and a 13.78% annual return with a multilayered feedforward neural network (Leung et al., 2000). A further study applied a generalized hyperbolic distribution in order to model the Nikkei 225 futures market and achieved an 8.6% annual return (Necula, 2009).

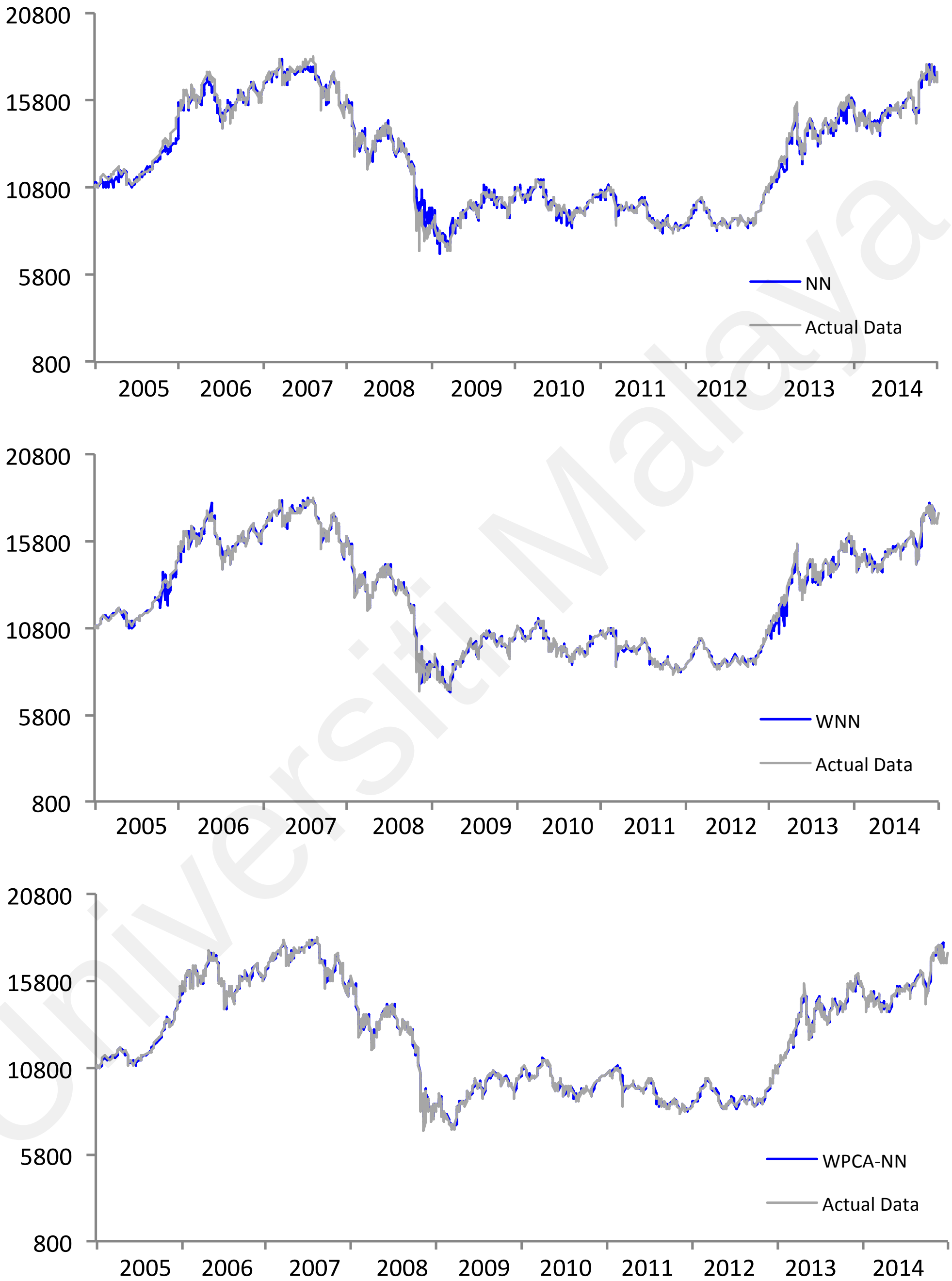


Figure 4.21: Forecasting results of NN, WNN and WPCA-NN models for the NIKKEI 225 futures market in out-sample data

4.8.5 SiMSCI Futures Market

According to the results for the SiMSCI futures market, shown in Table 4.29 and Table 4.30, the WPCA-NN model performs better than the WNN and NN models (1.14 < 1.79 < 7.26). Moreover, the WPCA-NN model (54.4) achieves the highest return among the buy-and-hold strategy (10.7); the MACD (11.5), RSI (1.5), stochastics (-10.6), and OMA (20.4) trading strategies; and the NN (19.9) and WNN (40.6) models. Although the NN model performs more poorly than the other two intelligent models, all of them achieve greater returns than a buy-and-hold strategy.

Table 4.29: Performance of the models and strategies measured by the MAPE ratio of evaluation results for the SiMSCI futures market

Models	Period	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
WPCA-NN	In-sample	0.009*	0.01	0.026	0.022*	0.01*	0.003*	0.006*	0.012*	0.005*	0.011	0.011*
	Out-sample	0.886*	0.986	2.585	2.173*	0.959*	0.32*	0.551*	1.202*	0.508*	1.254*	**1.142***
WNN	In-sample	0.012	0.008*	0.022*	0.025	0.028	0.024	0.012	0.021	0.012	0.007*	0.017
	Out-sample	1.21	0.84*	2.164*	2.497	2.775	2.428	1.24	2.109	1.193	1.467	**1.792**
NN	In-sample	0.023	0.02	0.026	0.062	0.092	0.09	0.064	0.035	0.03	0.022	0.046
	Out-sample	3.05	6.7	21.64	17.59	6.24	3.4	6.49	3.26	2.37	1.87	**7.26**

* The best performance among the models.

Table 4.30. Returns of the models and strategies without transaction costs: the results for the SiMSCI futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
B & H	11.7	40.9	23.8	-48.4	62.2	16.8	-21.7	24.4	-2.5	-0.6	**10.7**
MACD	0.7	-4.9	-10.8	41.2	10.4	18.2	2.8	26.2	20.8	10.4	**11.5**
RSI	10.4	-9.3	7.4	-44.1	-8.6	38.4	-4.4	20.1	2.9	1.9	**1.5**
Stochastics	21.6	-36.9	-13.9	-51.7	-25.9	-1.0	-4.5	-11.2	14.0	4.1	**-10.6**
OMA	-6.3	36.1	-2.7	55.3	33.3	32.0	0.3	37.4*	8.2	10.5	**20.4**
NN	14.0	15.2	0.9	27.9	36.0	31.5	16.3	23.8	6.0	27.7	**19.9**
WNN	36.6	56.3*	65.6	35.9	55.5	45.5*	47.1	25.4	17.8	20.4	**40.6**
WPCA-NN	45.4*	53.9	78.5*	76.2*	75.1*	42.0	55.5*	37.3	48.1*	32.3*	**54.4***

* The highest return among the models.

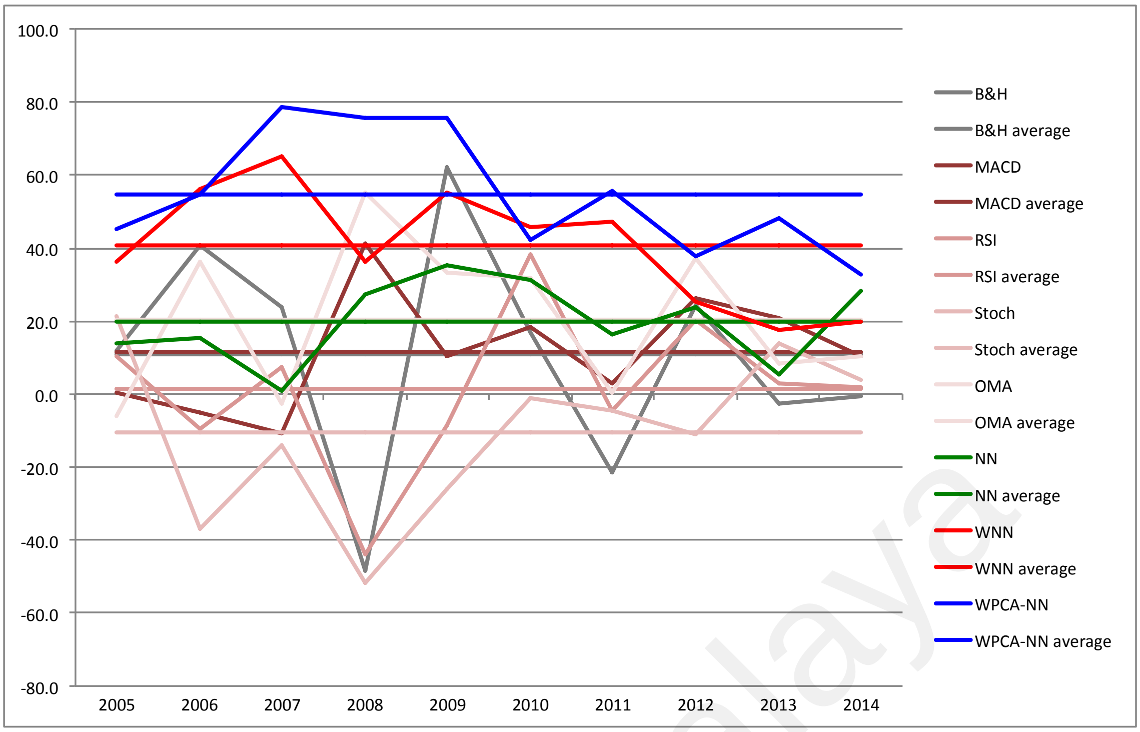


Figure 4.22: Returns of the models and strategies without transaction costs: the results for the SiMSCI futures market

In addition, Table 4.31 presents the returns of the models and strategies with transaction costs. The findings confirm the same ranking of the models and that the WPCA-NN model has the highest profitability compared with the other models. Figure 4.22 and Figure 4.23 illustrate the return results of all models and strategies without and with transaction costs respectively. Based on the best-performing networks in the SiMSCI futures market (Table 4.6), four levels of decomposition, db7 wavelets, and a penalized high thresholding strategy achieve significantly high performance and returns with the WPCA-NN model.

Table 4.31: Returns of the models and strategies with transaction costs: the results for the SiMSCI futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	**Average**
B & H	11.6	40.7	23.6	-48.5	62.0	16.6	-21.8	24.2	-2.7	-0.7	**10.5**
MACD	-2.5	-9.3	-14.2	38.6	7.4	14.9	0.2	23.6	16.8	7.6	**8.3**
RSI	10.0	-9.8	6.7	-44.2	-9.0	37.3*	-4.8	19.2	2.2	1.5	**0.9**
Stochastics	18.9	-37.7	-15.9	-52.3	-27.4	-2.3	-7.2	-13.1	12.2	2.4	**-12.2**
OMA	-11.5	33.5	-7.5	50.5	28.2	28.7	-3.0	32.4*	3.1	5.0	**16.0**
NN	6.5	11.0	-5.6	19.0	26.2	21.7	12.2	19.4	0.9	20.6*	**13.2**
WNN	25.2	42.2*	48.9	24.6	40.3	28.2	32.2	9.8	5.4	1.5	**25.8**
WPCA-NN	27.8*	40.6	62.4*	58.5*	64.3*	27.3	41.9*	23.4	30.6*	21.4	**39.8***

* The highest return among the models.

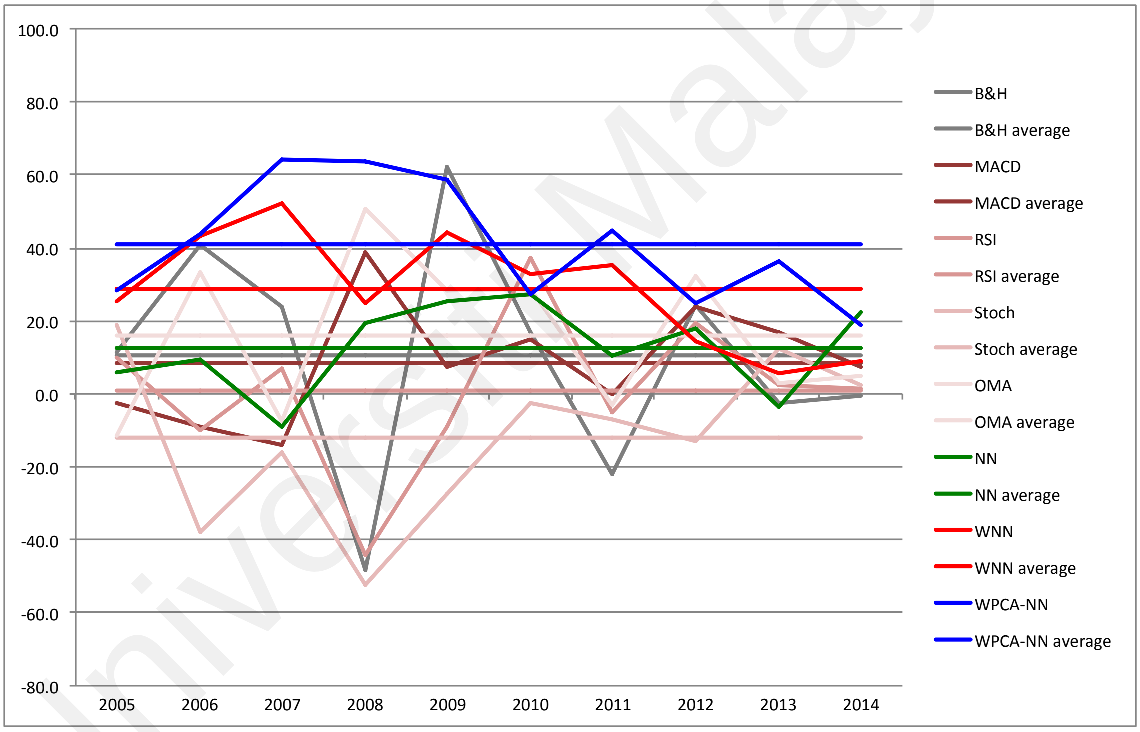


Figure 4.23: Returns of the models and strategies with transaction costs: the results for the SiMSCI futures market

Figure 4.24 displays the forecasting results of all intelligent models for 2005-2014. A prior study used an NN model to forecast the Singapore index and achieved a 25.64% annual return (Quah & Srinivasan, 1999). Another study achieved an approximately 15%

annual return performing generalized autoregressive conditional heteroscedasticity on the same index (Chiang & Doong, 2001).

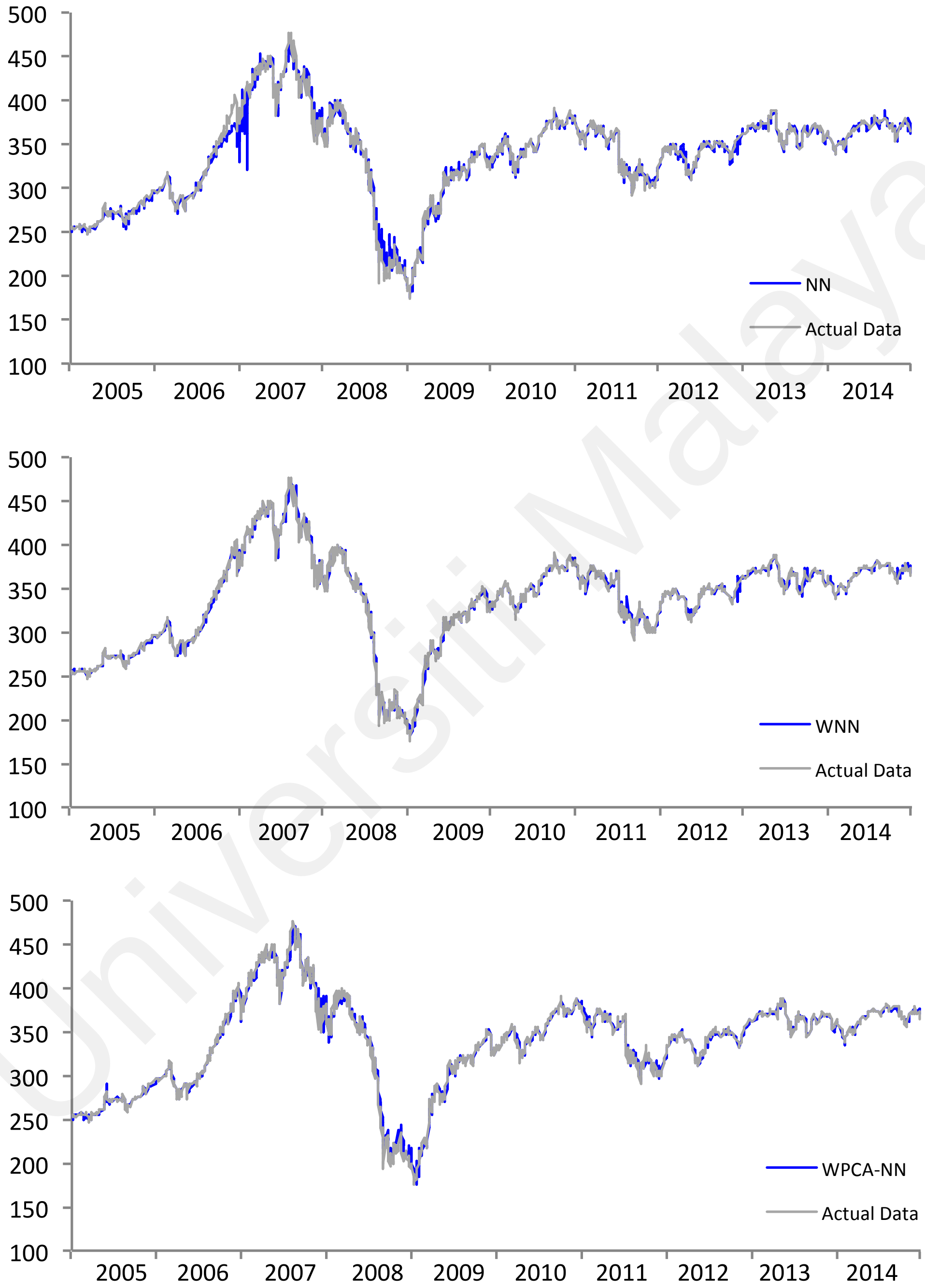


Figure 4.24: Forecasting results of NN, WNN and WPCA-NN models for the SiMSCI futures market in out-sample data

4.8.6 S&P 500 Futures Market

Based on the results for the S&P 500 futures market presented in Table 4.32, the trained networks for all models are valid because their MAPE values are quite low and acceptable. The WPCA-NN model outperforms the WNN and NN models over the entire testing period (2.33 < 2.35 < 29.55). In this regard, the lower the MAPE ratio, the better. Moreover, the WPCA-NN model (44.8) achieves the highest return among the buy-and-hold strategy (6.1); the MACD (-1.1), RSI (4.7), stochastics (5.3), and OMA (-7.2) trading strategies; and the NN (20.5) and WNN (37.8) models, based on the results shown in Table 4.33.

Table 4.32: Performance of the models and strategies measured by the MAPE ratio of evaluation results for the S&P 500 futures market

Models	Period	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
WPCA-NN	In-sample	0.006	0.028	0.025	0.026*	0.013*	0.024	0.05	0.02*	0.017*	0.014*	0.022*
	Out-sample	1.475*	2.804	2.511	2.646*	1.348*	2.384	5.007	2.028*	1.736*	1.413*	**2.335***
WNN	In-sample	0.006*	0.025*	0.022*	0.033	0.022	0.023*	0.028*	0.023	0.018	0.027	0.023
	Out-sample	1.485	2.541*	2.175*	3.293	2.162	2.296*	2.794*	2.342	1.759	2.693	**2.354**
NN	In-sample	0.107	0.117	0.163	0.285	0.332	0.25	0.105	0.05	0.053	0.105	0.157
	Out-sample	12.59	16.82	13.99	40.15	14.64	28.77	131.94	21.86	6.18	8.36	**29.53**

* The best performance among the models.

Table 4.33: Returns of the models and strategies without transaction costs: the results for the S&P 500 futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
B & H	4.0	12.1	2.7	-38.3	20.1	11.0	-1.0	11.6	26.4	12.3	**6.1**
MACD	-0.9	-17.8	3.0	27.4	-9.5	3.4	-4.2	-14.3	-1.7	4.0	**-1.1**
RSI	9.1	-1.3	6.1	-32.9	9.2	17.4	14.9	18.2	-23.2	29.0	**4.7**
Stochastics	5.8	-8.1	2.4	-30.8	30.0	11.5	19.1	-9.7	19.3	13.0	**5.3**
OMA	-2.2	-5.6	-7.7	-32.1	6.4	14.7	-30.8	-13.8	-1.0	0.1	**-7.2**
NN	18.1	-5.2	-1.8	19.3	10.9	37.4	38.3	32.0	30.0	25.9	**20.5**
WNN	24.8	55.7*	29.7	28.6	34.4	49.4	45.9	47.4*	36.2	25.8	**37.8**
WPCA-NN	36.3*	49.6	34*	57.8*	60.7*	57.6*	49.7*	28.4	42.1*	32.1*	**44.8***

* The highest return among the models.

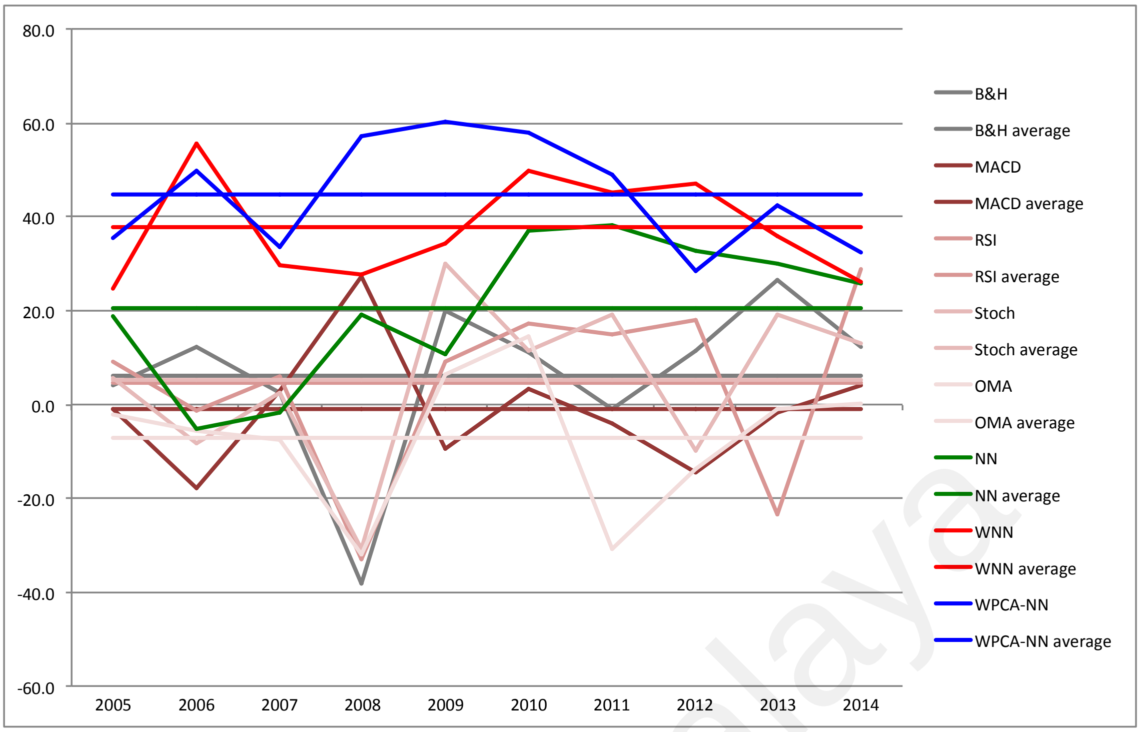


Figure 4.25: Returns of the models and strategies without transaction costs: the results for the S&P 500 futures market

In addition, Table 4.34 presents the returns of the models and strategies with transaction costs. The findings confirm the same ranking of the models. The WPCA-NN model achieves higher returns (39.6) than the other models and strategies, including the buy-and-hold strategy (6.1). Figure 4.25 and Figure 4.26 illustrate the return results of all models and strategies without and with transaction costs respectively. Based on the MAPE ratio, all forecasting results of the WPCA-NN, WNN, and NN models are valid from 2005 to 2014. According to the best-performing networks in the S&P 500 futures market (Table 4.6), three levels of decomposition, a db9 wavelet, and a penalized high thresholding strategy achieve significantly high performance and excess returns with the WPCA-NN model.

Table 4.34: Returns of the models and strategies with transaction costs: the results for the S&P 500 futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	**Average**
B & H	4.0	12.0	2.7	-38.3	20.1	11.0	-1.0	11.6	26.4	12.3	**6.1**
MACD	-1.6	-19.0	2.2	26.6	-10.3	2.5	-5.0	-15.0	-2.5	3.1	**-1.9**
RSI	9.0	-1.4	6.0	-32.9	9.1	17.3	14.9	18.0	-23.2	28.9*	**4.5**
Stochastics	5.3	-8.6	1.9	-31.1	29.3	10.9	18.4	-10.1	18.7	12.5	**4.7**
OMA	-3.4	-6.7	-8.7	-33.0	5.3	13.9	-31.8	-14.8	-2.2	-0.9	**-8.2**
NN	11.7	-10.0	-6.1	14.3	6.1	32.4	33.9	29.3	23.6	19.4	**15.4**
WNN	12.0	44.3*	18.7	18.4	24.0	39.0	33.9	35.9*	26.4	15.8	**26.8**
WPCA-NN	22.9*	39.0	21.5*	41.9*	44.9*	45.9*	37.2*	15.5	28.9*	16.3	**31.4***

* The highest return among models

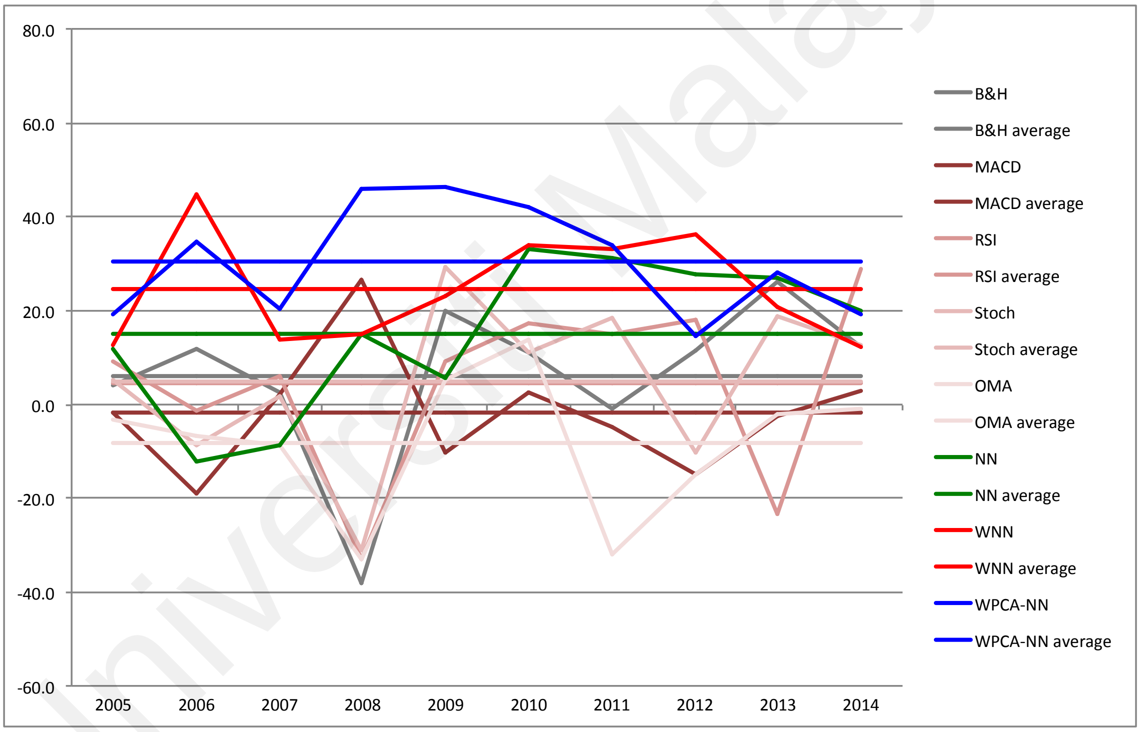


Figure 4.26: Returns of the models and strategies with transaction costs: the results for the S&P 500 futures market

Figure 4.27 displays the forecasting results of all intelligent models for 2005-2014. A prior study shows an 18.96% annual return with the application of a generalized regression NN on the S&P 500 futures market (Enke & Thawornwong, 2005).

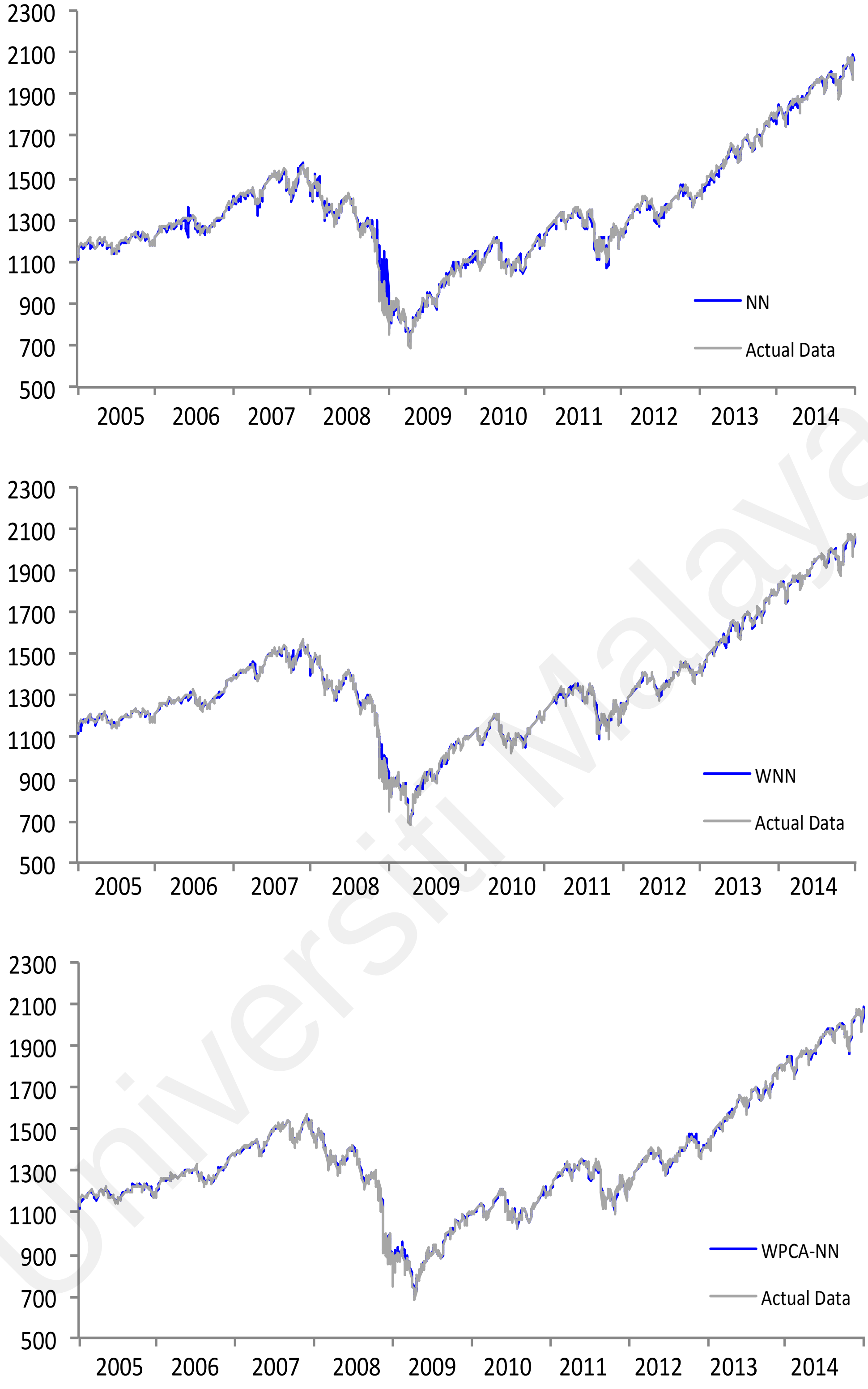


Figure 4.27: Forecasting results of NN, WNN and WPCA-NN models for the S&P 500 futures market in out-sample data

4.8.7 TAIEX Futures

Based on the results for the TAIEX futures market presented in Table 4.35 and Table 4.36, the trained networks for all models are valid because their MAPE values are less than 5%. Moreover, the WPCA-NN model outperforms the WNN and NN models (2.71 < 5.00 < 143.77), according to the MAPE evaluation. The WPCA-NN model also achieves a greater return (45.3) than the buy-and-hold strategy (8.5); the MACD (-0.2), RSI (2.9), stochastics (4.6), and OMA (2.7) trading strategies; and the NN (8.6) and WNN (35.3) models, according to the annual returns.

Table 4.35: Performance of the models and strategies measured by the MAPE ratio of evaluation results for the TAIEX futures market

Models	Period	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
WPCA-NN	In-sample	0.018*	0.019*	0.027*	0.028*	0.023*	0.024*	0.045*	0.04*	0.028*	0.016*	0.027*
	Out-sample	1.921*	1.915*	2.689*	2.845*	2.325*	2.437*	4.537*	3.99*	2.826*	1.641*	**2.713***
WNN	In-sample	0.024	0.033	0.031	0.094	0.035	0.037	0.079	0.091	0.036	0.028	0.049
	Out-sample	3.598	3.328	3.141	9.358	3.459	3.745	7.911	9.1	3.629	2.778	**5.005**
NN	In-sample	0.754	0.831	0.969	1.153	1.689	1.514	1.198	0.643	0.66	0.96	1.037
	Out-sample	55.92	43.44	80.63	212.44	97.32	130.28	397.79	259.53	108.03	52.32	**143.77**

* The best performance among the models.

Table 4.36: Returns of the models and strategies without transaction costs: the results for the TAIEX futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
B & H	4.1	21.6	6.1	-45.1	80.3*	19.4	-25.2	13.8	7.8	2.4	**8.5**
MACD	16.1	1.8	31.2	-35.1	-26.8	22.9	-16.4	1.2	-3.2	5.9	**-0.2**
RSI	9.8	18.2	66.5*	-46.4	-73.0	12.8	-6.6	35.0	9.7	3.3	**2.9**
Stochastics	-1.4	-9.2	-3.3	-2.7	6.4	59.1*	-14.2	-6.5	4.3	13.9	**4.6**
OMA	19.3	8.5	3.9	-22.5	12.9	22.8	-4.5	-11.3	-0.4	-2.0	**2.7**
NN	0.0	-0.4	-3.6	-9.3	15.2	37.5	-7.4	19.7	27.7	6.4	**8.6**
WNN	53.8*	49*	-8.0	6.5	42.4	56.7	23.3	48.8	45.5	35.5	**35.3**
WPCA-NN	49.3	43.8	24.6	22.9*	60.5	55.5	58.9*	51.5*	47.5*	38.6*	**45.3***

* The highest return among the models.

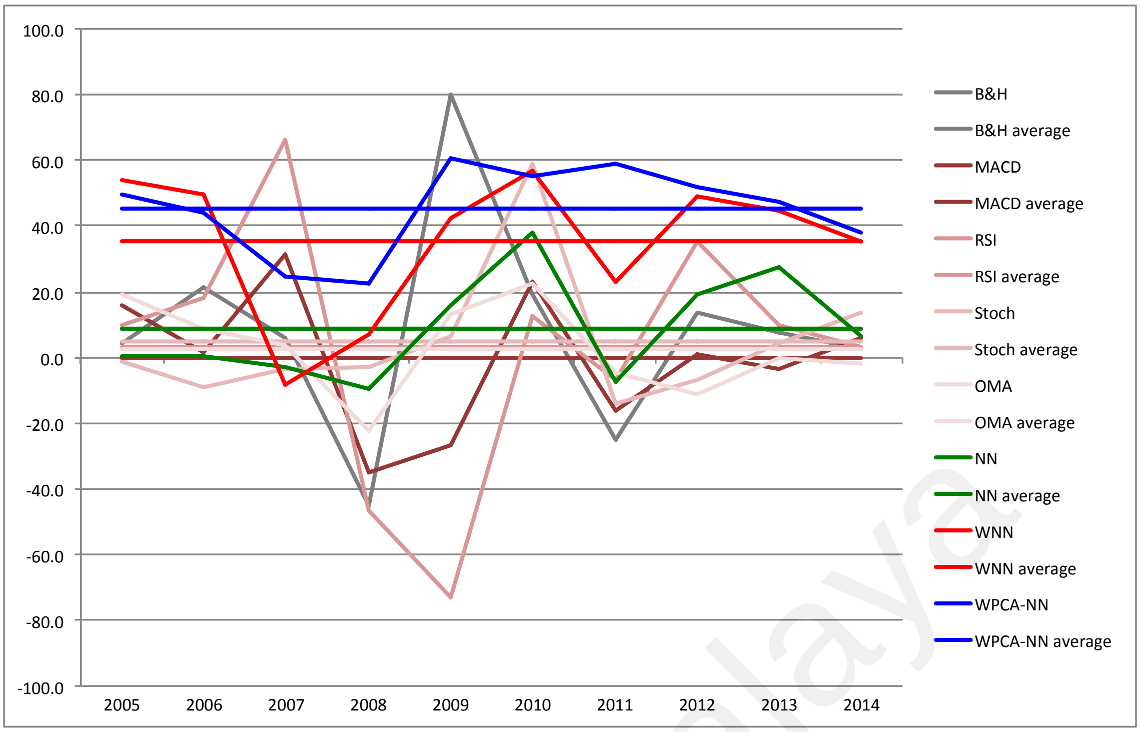


Figure 4.28: Returns of the models and strategies without transaction costs: the results for the TAIEX futures market

In addition, Table 4.37 presents the returns of the models and strategies with transaction costs. These confirm the same ranking of the models. The returns also show that the WPCA-NN model has the highest return (37.6) compared with the other models. Figure 4.28 and Figure 4.29 display the return results of all models and strategies without and with transaction costs respectively. According to the best-performing networks in the TAIEX futures market (Table 4.6), three levels of decomposition, a db9 wavelet, a penalized high thresholding strategy, and a delay of one week achieve considerably high performance and returns with the WPCA-NN model. Although these results are robust and valid in different evaluation subsets and years, they relate to the characteristics of the market and may change in the future. However, this combination is the most appropriate for forecasting of the TAIEX futures market. Further, any other combination of these parameters performs comparably better than the others because the parameters may represent common characteristics of the TAIEX futures market across 10 years.

Table 4.37: Returns of the models and strategies with transaction costs: the results for the TAIEX futures market

Models and Strategies	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	**Average**
B & H	4.0	21.4	5.9	-45.1	80*	19.2	-25.3	13.6	7.6	2.3	**8.4**
MACD	12.8	-1.0	27.5	-37.2	-29.8	20.2	-18.8	-1.2	-6.4	3.3	**-3.0**
RSI	8.9	17.7	65*	-46.6	-73.1	11.8	-6.9	34.0	9.2	2.8	**2.3**
Stochastics	-2.9	-10.8	-4.8	-4.5	4.5	56.1*	-16.0	-8.7	2.4	11.4	**2.7**
OMA	15.1	3.8	-0.5	-25.3	7.9	18.9	-9.2	-14.6	-5.4	-6.6	**-1.6**
NN	-8.5	-4.8	-12.4	-17.5	11.3	33.4	-13.4	12.3	17.3	-3.1	**1.5**
WNN	44.4*	39.3*	-18.9	-8.0	29.5	40.5	13.3	37.4	35.7*	21.5	**23.5**
WPCA-NN	40.3	32.2	14.9	6.4*	46.8	44.9	43.6*	40.9*	32.8	28.6*	**33.1***

* The highest return among the models.

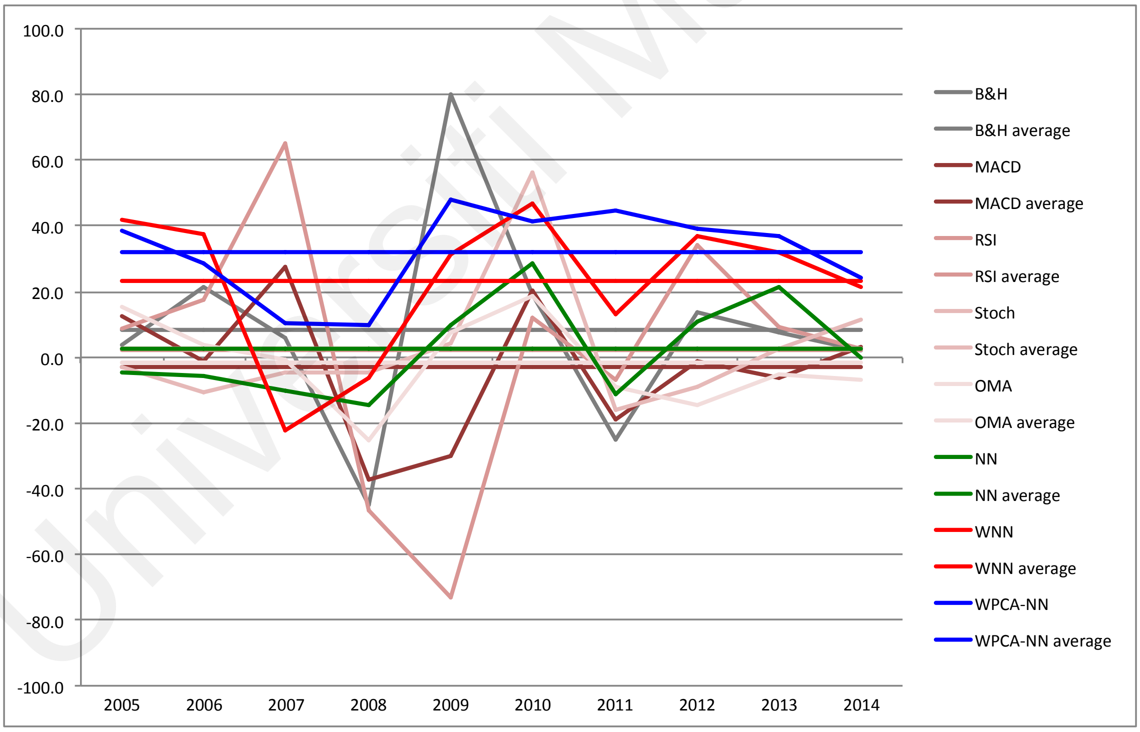


Figure 4.29: Returns of the models and strategies with transaction costs: the results for the TAIEX futures market

Figure 4.30 displays the forecasting results of all three intelligent models for 2005-2014. A prior study shows 11.42% return using artificial bee colony neural network on TAIEX (Hsieh et al., 2011). Another study achieve 17.27% average return using a hybrid model based on rough set theory and genetic algorithm (Cheng et al., 2010).

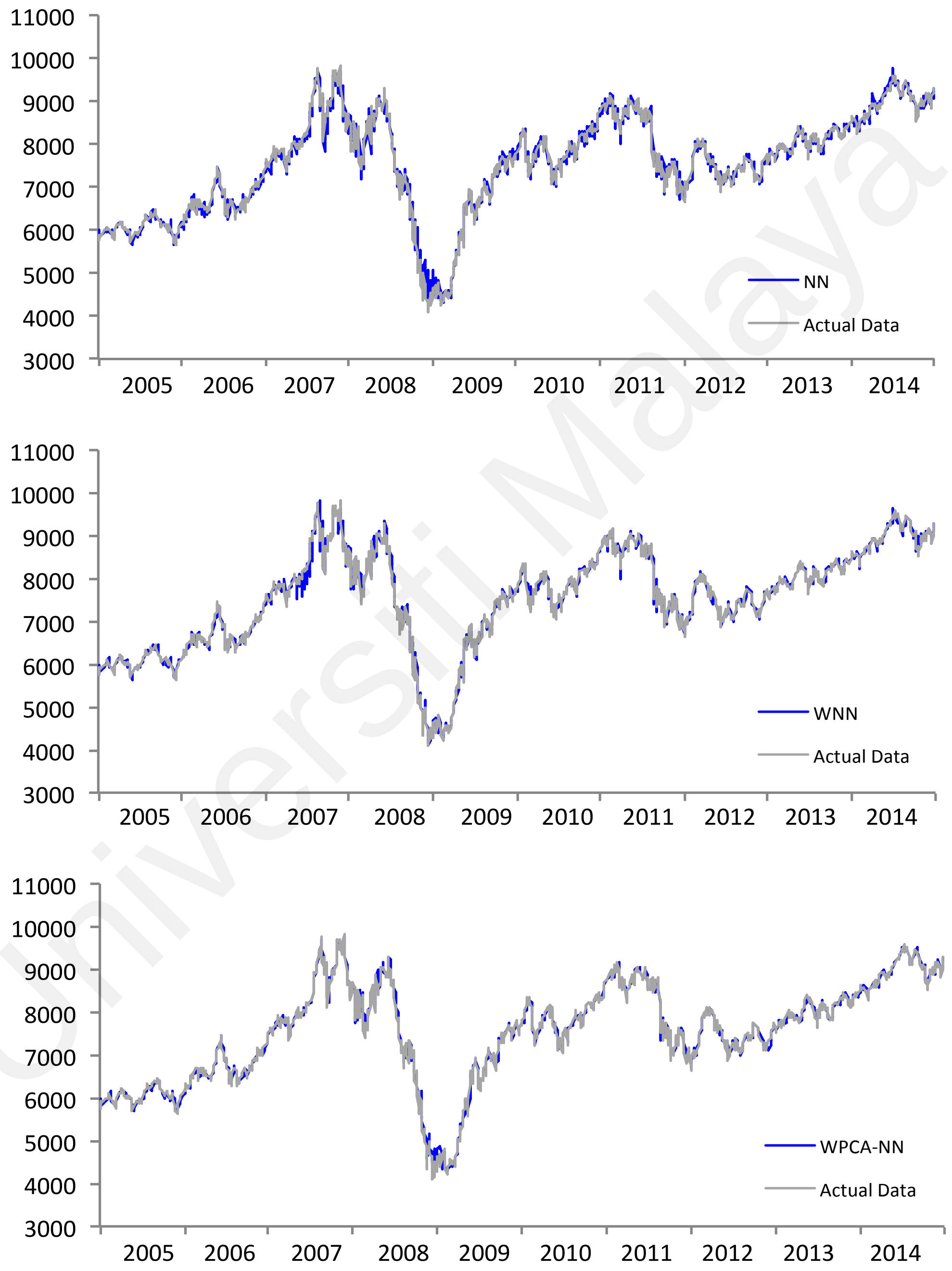


Figure 4.30: Forecasting results of NN, WNN and WPCA-NN models for the TAIEX futures market in out-sample data

4.9 Summary and Discussion of Results

This section is a detailed and comprehensive report of the study's tests and analyses. It starts with the sample data and the descriptive analysis. After this, a thorough report of the forecasting performance and trading results of all models and strategies for all the studied futures markets is presented. The significance and descriptive analysis of the results are also discussed in order to demonstrate the reliability of the results. In addition, the results that are applicable to the time around the financial crisis are investigated. All results are presented in a way that addresses the hypotheses and objectives of this study.

The sample data of the study is collected from Bloomberg L.P. and constitutes 13 years of historical data from 2002 to 2014 (3,224 daily data items for each market). However, for all investigated conventional and modern trading models, the trading period or out-sample data is from 2005 to 2014. The daily OHLC and volume are collected from January 2, 2002 to December 31, 2014 together with the RSI, MACD, MACD signal, stochastic Fast %K, stochastic slow %K, stochastic %D, and ultimate oscillator data of the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets. The daily OHLC data, spot prices, and volumes of the technical indicators are used as inputs for the intelligent models. The data are prepared and checked statistically.

According to the unit root test, the intercept value of the first difference forms for the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets are, respectively, -58.37271, -60.82484, -55.64235, -58.40276, -60.65505, -62.90317, and -53.85111. These are significantly (to a 5% confidence level) and strongly negative. Hence, it can be concluded that all futures market data are stationary in the first difference form; thus, historical data can be used to forecast the price movements of the selected futures markets. Further, the original actual data can be employed as input variables without any modification (Diaconescu, 2008) for an ANN because the ANN technique with the first step of delay $f(P_t, P_{t-1})$ is already able to make the data stationary. In addition to stationary analysis, a Breusch–Godfrey serial correlation test shows that autocorrelation exists between prior prices and current prices in the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets with R-squared values of 8.050268, 2.907285, 6.049925, 3.00968, 8.702567, 6.388968, and 11.51784 respectively (at a confidence level of 10%). In other words, historical data can be used to predict futures movements in tested futures markets. These findings are consistent with stationary results and confirm the predictability of futures markets. Thus, all tested futures time series are predictable by lag in their own time series, which suggests predictability with a similar delay feature in the ANN technique. According to the results of a Granger causality test, the RSI, MACD, MACD signal, stochastic fast %K, stochastic fast %D, stochastic slow %D, and ultimate oscillator trading strategies are confirmed to predict the future prices of futures market. These findings are confirmatory evidence for the usage of selected technical indicators as input variables for the forecasting models. Ultimately, according to descriptive analyses and predictability of the sample data, all tests and models considered in the hypotheses are valid for use with the sample data. In addition, the variables and technical indicators have been selected correctly and will be able to forecast futures' price movements.

According to the sample data's descriptive analysis, the tests and analyses offered in the hypotheses are valid. Hypothesis H1 aims to answer whether the selected markets are predictable in terms of long-term memory and entropy indices. The Hurst exponent and approximate entropy are measured for this objective. Hypothesis H2 is posited in order to answer whether the proposed hybrid model (WPCA-NN) outperforms the buy-and-hold strategy, the purpose being to show that the futures markets do not follow the RWH. In this regard, the WPCA-NN model must generate statistically significant higher returns than the buy-and-hold benchmark. Hypothesis H3 is proposed to show whether the novel hybrid model performs better than commonly used and best-performing technical trading strategies, these being MACD (H3a), RSI (H3b), stochastics (H3c), and OMA (H3d). In this context, the WPCA-NN model must achieve significantly higher returns than the selected technical trading strategies. Hypothesis H4a aims to test whether the denoising process of the novel hybrid model adds to the forecasting performance of a pure NN model. The WPCA-NN and pure NN models are tested for the same period and the same futures markets. The WPCA-NN model must forecast the markets with significantly higher performance and returns. Hypothesis H4b tests whether the hybrid multivariate denoising process of the new proposed model is better than the simple univariate denoising process in the WNN model. This requires significantly higher performance and returns from the WPCA-NN model compared with the WNN model. In addition, the outcome of H4b should demonstrate that the adjustment and enhancement of the denoising process by a multivariate wavelet-PCA is a promising tool and can be employed in future studies. The following findings address this study's hypotheses and objectives.

The Hurst exponent for all futures markets is calculated annually (from 2005 to 2014) as a long-term memory approach that makes it possible to observe changes in stock market predictability and efficiency (Kristoufek & Vosvrda, 2014). The variation of this index also provides predictability and inefficiency over time. The results show that in all markets, Hurst exponents are larger than 0.5 and demonstrate the long-range dependence and predictability of the current data through their past data (Hang Seng (0.5973), KLCI (0.5507), KOSPI 200 (0.5237), NIKKEI 225 (0.5109), SiMSCI (0.5840), S&P 500 (0.5128) and TAIEX (0.5383)). According to the Hurst exponents, the Hang Seng (0.5973), SiMSCI (0.5840), and KLCI (0.5507) futures markets are more inefficient and predictable than the other futures markets. However, based on the Hurst exponent results, the NIKKEI 225 (0.5109), S&P 500 (0.5128), and KOSPI 200 (0.5237) futures markets are considered to be more efficient and less predictable than other futures markets. Taking into account the annual value of the Hurst exponent for each futures market, it is inferable that persistence changes annually. In other words, the market predictability and inefficiency of these futures markets varies over time. Since the Hurst exponents are higher than 0.5 and less than 1, and the Hurst exponents vary over time in the studied futures markets (e.g., for the Hang Seng futures market, 2005 (0.5801), 2006 (0.6332), 2007 (0.6247), 2008 (0.6402), 2009 (0.6315), 2010 (0.6157), 2011 (0.6045), 2012 (0.6423), 2013 (0.6319) and 2014 (0.6247)), it is possible to observe predictability and time-varying efficiency (Kristoufek, 2012; Kristoufek & Vosvrda, 2014; Mandelbrot & Hudson, 2005). This suggests that the markets are supported by the AMH (Lo, 2004; Noda, 2012; Urquhart & Hudson, 2013; Urquhart & McGroarty, 2014). These results are consistent with other studies (Hull & McGroarty, 2014; Kristoufek & Vosvrda, 2014). Indeed, the latter of these studies considers long-term memory regarding 38 stock indices (Kristoufek & Vosvrda, 2014). The authors present Hurst exponents for the same markets as considered in the current research as follows; Hang Seng (0.5945), KLCI (0.5489), KOSPI 200 (0.5135), NIKKEI 225 (0.5063), SiMSCI (0.5937), and S&P 500 (0.5026) from 2000 to 2011. These results are consistent with the results of the current study and confirm the same ranking for the markets according to the Hurst exponent (S&P 500 (1^{st}), NIKKEI 225 (2^{nd}), KOSPI 200 (3^{rd}), KLCI (4^{th}), SiMSCI (5^{th}) and Hang Seng (6^{th})). In addition, the authors conclude that there is predictability and inefficiency for the studied markets based on the Hurst exponents and approximate entropy. Approximate entropy (Kristoufek & Vosvrda, 2014; Pincus & Kalman, 2004; Zunino et al., 2010) is also considered in the current study. Together, the Hurst exponent and approximate entropy form hypothesis H1. Thus, the futures markets show significant changes in predictability over time, taking into account long-term memory.

Several studies support verifiable results of long-term memory using the Hurst index (Kristoufek & Vosvrda, 2014; Mensi et al., 2014; Willinger et al., 1999). However, this study considers structural breaks by using Bai–Perron tests of L globally optimized breaks in contrast with the null hypothesis of no structural breaks. Moreover, this study adds the 2008 financial crisis as a known structural break into the analysis. The values of Hurst exponents in years with structural breaks are not the highest or the lowest in general. In addition, the values of Hurst exponents in years with structural breaks are not essentially higher or lower than a year before or after them. All values of Hurst exponents, including those years with or without structural breaks, are almost in the same range for each futures market and almost higher than the overall values of the Hurst exponents. Ultimately, these findings for long-term memory analysis do not suggest any spurious results caused by structural breaks or the financial crisis.

Like long-term memory, approximate entropy is calculated here once for 10 years and once annually for all futures markets. All the results are between 0 and 1, and none of them shows a completely random market (Entropy = 1). The results of approximate entropy for 10 years are as follow; Hang Seng (0.3308), KLCI futures (0.2712), KOSPI 200 (0.4429), NIKKEI 225 (0.4376), SiMSCI (0.2107), S&P 500 (0.4221), and TAIEX (0.3676). The higher the value of approximate entropy, the more efficient the market. The futures markets can be ranked in terms of market efficiency based on entropy as follows; KOSPI 200(1st), NIKKEI 225 (2nd), S&P 500 (3rd), TAIEX (4th), Hang Seng (5th), KLCI (6th) and SiMSCI (7th). Most importantly, entropy varies over time from 2005 to 2014, a finding that represents changing market efficiency and predictability over this period. These results are consistent with the outcomes of a prior study (Kristoufek & Vosvrda, 2014). These authors measure approximate entropy for the Hang Seng (0.3033), KLCI (0.1773), KOSPI 200 (0.4473), NIKKEI 225 (0.4285), SiMSCI (0.2027), and S&P 500 (0.3405) futures markets from 2000 to 2011. The ranking of market predictability or efficiency in their study is the same as the ranking in the current study. Moreover, the entropy values for the futures markets are almost in the same range. The values of approximate entropy in years with structural breaks are not the highest or the lowest in general. Moreover, the values of entropy in years with structural breaks are not essentially higher or lower than a year before or after them. Ultimately, these outcomes for entropy investigation do not suggest any fake results caused by structural breaks or the financial crisis.

The results of approximate entropy show variation over time and confirm the results of the Hurst exponents regarding this variation. Thus, both approximate entropy and the Hurst exponents suggest that these markets are predictable and that their predictability and inefficiency vary over time. These findings are compatible with the AMH, which suggests a time-varying attitude for market efficiency. Consequently, futures markets exhibit significant changes in predictability over time, taking into account long-term memory and approximate entropy (hypothesis H1).

The return results of the offered novel model of trading (WPCA-NN) need to be significant in comparison with other models. This requirement means that the superiority of the WPCA-NN model's results should be meaningful and reliable. In order to analyze the significance of the return results, skewness, kurtosis, Sharpe ratios, and significant alphas are examined.

The trading results of all futures markets are positively skewed (Hang Seng (0.48), KLCI (0.74), KOSPI 200 (0.25), NIKKEI 225 (0.09), SiMSCI (0.73), S&P 500 (1.56) and TAIEX (0.19)), which means that frequent small gains or losses are achieved with the WPCA-NN model. Further, extremely negative scenarios are not likely to happen based on the skewness results. A kurtosis value is estimated in relation to normal distribution, on which kurtosis equals three. Based on the value of the excess kurtosis (kurtosis minus three), any positive value for excess kurtosis would mean fatter tails for a distribution plot and consequently a lower risk of an extreme result. Regarding the kurtosis results, all futures markets, except the KOSPI 200 (2.32) and SiMSCI (2.60), are leptokurtic, which means fatter tails and a lower risk of an extreme outcome (Hang Seng (3.33), KLCI (3.59), NIKKEI 225 (3.81), S&P 500 (5.36) and TAIEX (3.93)). The return results for the KOSPI 200 (2.32) and SiMSCI (2.60) futures markets show kurtosis smaller than three, which means thinner tails than other markets and a higher risk of extreme return results from the WPCA-NN model. Finally, the analysis of distribution for the trading returns of the WPCA-NN model shows a reliable distribution with repeated small gains, small losses, and a low probability for extremely negative scenarios to occur in all futures markets. Regarding standard deviation and its impact on the trading outcomes, the reliability of the results is tested with Sharpe ratios.

Based on the Sharpe ratio results, the performance of the WPCA-NN model is highly reliable and acceptable in all futures markets. The higher the Sharpe ratio, the better the risk-adjusted performance and, accordingly, the better the performance of the model or trading strategy on the futures market. The results show that the WPCA-NN model performs with higher risk in the Hang Seng (0.77) and KLCI (0.85) futures markets compare with the other futures markets. The reason is that the Sharpe ratio values of the Hang Seng and KLCI futures markets are both less than one. However, the trading results of the Hang Seng and KLCI futures markets are still considered reliable because they are positive and close to one. The WPCA-NN model achieves considerably reliable returns for other futures markets according to their Sharpe ratios. The results of Sharpe ratio for WPCA-NN trading strategy for all markets are KOSPI 200 (1.42), NIKKEI 225 (1.44), SiMSCI (1.57), S&P 500 (1.04), and TAIEX (2.02). In addition, the Sharpe ratio results show that the returns achieved by the WPCA-NN model are because the model is reliable and successful, and are not the result of high risk. Finally, the Sharpe ratios of the WPCA-NN model for the futures markets are considerably high; thus, the risk-adjusted performance of the WPCA-NN model is suitable. In order to compare the outcomes of all models and strategies, the means of the WPCA-NN model are compared with those of the other models and strategies through t-test analysis.

According to all the futures markets' results, the trained networks of the WPCA-NN model generate positive and significantly higher returns compared with the buy-and-hold strategy. Thus, in the context of no transaction costs:

$(R_{WPCA-NN,HSI}(34.8) > R_{B\&H,HSI}(7.4)$,

$R_{WPCA-NN,KLCI}(47.2) > R_{B\&H,KLCI}(6.2)$,

$R_{WPCA-NN,KOSPI}(44.8) > R_{B\&H,KOSPI}(17.5)$,

$R_{WPCA-NN,NIKKEI}(42.1) > R_{B\&H,NIKKEI}(4.9)$,

$R_{WPCA-NN,SiMSCI}(54.1) > R_{B\&H,SiMSCI}(10.7)$,

$R_{WPCA-NN,S\&P500}(48.2) > R_{B\&H,S\&P500}(6.1)$,

$R_{WPCA-NN,TAIEX}(45.3) > R_{B\&H,TAIEX}(8.5))$.

In the context of transaction costs:

$(R_{WPCA-NN,HSI}(29.5) > R_{B\&H,HSI}(9.3)$,

$R_{WPCA-NN,KLCI}(41.4) > R_{B\&H,KLCI}(9.9)$,

$R_{WPCA-NN,KOSPI}(40) > R_{B\&H,KOSPI}(29.3)$,

$R_{WPCA-NN,NIKKEI}(35.9) > R_{B\&H,NIKKEI}(4.8)$,

$R_{WPCA-NN,SiMSCI}(47.2) > R_{B\&H,SiMSCI}(10.5)$,

$R_{WPCA-NN,S\&P500}(39.8) > R_{B\&H,S\&P500}(6.1)$,

$R_{WPCA-NN,TAIEX}(37.6) > R_{B\&H,TAIEX}(8.4))$.

This difference between the trading returns of the WPCA-NN model and the buy-and-hold strategy should be meaningful and significant. Moreover, the alpha coefficients of the regression models between the return results of these two trading systems should be significant and positive.

When there is a significant and meaningful difference between the mean of two different return results from two different trading models, it can be concluded that the model with the higher mean of return is superior to the other. The quarterly return mean difference between the WPCA-NN model and the buy and hold strategy is calculated with a paired t-test. Since all Satterthwaite–Welch t-test values are positive and significant at the 5% level of confidence, the following conclusion can be drawn: There is a significant difference between the return mean of the WPCA-NN model and the buy and hold strategy (the RWH outcome).

The alpha coefficients of the regression models for the trading results of the WPCA-NN model and the buy and hold strategy show that the WPCA-NN model generates significantly higher returns than the buy-and-hold benchmark because of the positive and significant alphas for all futures markets ($\alpha_{HangSeng} = 33.13 > 0$; $\alpha_{KLCI} = 47.84 > 0$; $\alpha_{KOSPI} = 42.56 > 0$; $\alpha_{Nikkei} = 43.63 > 0$; $\alpha_{SiMSCI} = 54.26 > 0$; $\alpha_{S\&P500} = 49.39 > 0$; and $\alpha_{TAIEX} = 43.57 > 0$). Based on the results, these futures markets do not follow random walk from 2005 to 2014. Because the results of the regression models suggest that it is possible to achieve better positive returns than the buy-and-hold strategy in futures markets, the AMH is supported. Overall, the hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the RWH (a passive buy-and-hold strategy) for selected financial markets as shown by the significant mean differences and significant positive alphas in the profit regressions.

According to all futures markets' results, the trained networks of the WPCA-NN model generate higher returns than the best-performing technical indicators as follows, without transaction costs.

MACD:

$(R_{WPCA-NN,HSI}(34.8) > R_{MACD,HSI}(2.8)$,

$R_{WPCA-NN,KLCI}(47.2) > R_{MACD,KLCI}(1.3)$,

$R_{WPCA-NN,KOSPI}(44.8) > R_{MACD,KOSPI}(9.9)$,

$R_{WPCA-NN,NIKKEI}(42.1) > R_{MACD,NIKKEI}(0.7)$,

$R_{WPCA-NN,SiMSCI}(54.4) > R_{MACD,SiMSCI}(11.5)$,

$R_{WPCA-NN,S\&P500}(48.2) > R_{MACD,S\&P500}(-1.1)$, and

$R_{WPCA-NN,TAIEX}(45.3) > R_{MACD,TAIEX}(-0.2)$).

RSI:

($R_{WPCA-NN,HSI}(34.8) > R_{RSI,HSI}(-3.1)$,

$R_{WPCA-NN,KLCI}(47.2) > R_{RSI,KLCI}(1.1)$,

$R_{WPCA-NN,KOSPI}(44.8) > R_{RSI,KOSPI}(-13.8)$,

$R_{WPCA-NN,NIKKEI}(42.1) > R_{RSI,NIKKEI}(3.7)$,

$R_{WPCA-NN,SiMSCI}(54.4) > R_{RSI,SiMSCI}(1.5)$,

$R_{WPCA-NN,S\&P500}(48.2) > R_{RSI,S\&P500}(4.7)$, and

$R_{WPCA-NN,TAIEX}(45.3) > R_{RSI,TAIEX}(2.9)$).

Stochastics:

($R_{WPCA-NN,HSI}(34.8) > R_{Stochastics,HSI}(13.3)$,

$R_{WPCA-NN,KLCI}(47.2) > R_{Stochastics,KLCI}(-6.9)$,

$R_{WPCA-NN,KOSPI}(44.8) > R_{Stochastics,KOSPI}(-12.5)$,

$R_{WPCA-NN,NIKKEI}(42.1) > R_{Stochastics,NIKKEI}(-5.6)$,

$R_{WPCA-NN,SiMSCI}(54.4) > R_{Stochastics,SiMSCI}(-10.6)$,

$R_{WPCA-NN,S\&P500}(48.2) > R_{Stochastics,S\&P500}(5.3)$, and

$R_{WPCA-NN,TAIEX}(45.3) > R_{Stochastics,TAIEX}(4.6)$.

OMA:

($R_{WPCA-NN,HSI}(34.8) > R_{SMA,HSI}(1.2)$,

$R_{WPCA-NN,KLCI}(47.2) > R_{SMA,KLCI}(3)$,

$R_{WPCA-NN,KOSPI}(44.8) > R_{SMA,KOSPI}(-8.3)$,

$R_{WPCA-NN,NIKKEI}(42.1) > R_{SMA,NIKKEI}(-6.4)$,

$R_{WPCA-NN,SiMSCI}(54.4) > R_{SMA,SiMSCI}(20.4)$,

$R_{WPCA-NN,S\&P500}(48.2) > R_{SMA,S\&P500}(-7.2)$, and

$R_{WPCA-NN,TAIEX}(45.3) > R_{SMA,TAIEX}(2.7)$)

Taking into account trading strategies with transaction costs, the WPCA-NN model achieves higher returns, as follows.

MACD:

($R_{WPCA-NN,HSI}(21.7) > R_{MACD,HSI}(0.5)$,

$R_{WPCA-NN,KLCI}(35.5) > R_{MACD,KLCI}(4)$,

$R_{WPCA-NN,KOSPI}(32.2) > R_{MACD,KOSPI}(4.5)$,

$R_{WPCA-NN,NIKKEI}(29.9) > R_{MACD,NIKKEI}(-1.5)$,

$R_{WPCA-NN,SiMSCI}(39.8) > R_{MACD,SiMSCI}(8.3)$,

$R_{WPCA-NN,S\&P500}(31.4) > R_{MACD,S\&P500}(-1.9)$, and

$R_{WPCA-NN,TAIEX}(33.1) > R_{MACD,TAIEX}(-3)$).

RSI:

($R_{WPCA-NN,HSI}(21.7) > R_{RSI,HSI}(-2.7)$,

$R_{WPCA-NN,KLCI}(35.5) > R_{RSI,KLCI}(-0.3)$,

$R_{WPCA-NN,KOSPI}(32.2) > R_{RSI,KOSPI}(-26.2)$,

$R_{WPCA-NN,NIKKEI}(29.9) > R_{RSI,NIKKEI}(3.5)$,

$R_{WPCA-NN,SiMSCI}(39.8) > R_{RSI,SiMSCI}(0.9)$,

$R_{WPCA-NN,S\&P500}(31.4) > R_{RSI,S\&P500}(4.5)$, and

$R_{WPCA-NN,TAIEX}(33.1) > R_{RSI,TAIEX}(2.3)$).

Stochastics:

($R_{WPCA-NN,HSI}(21.7) > R_{Stochastics,HSI}(7.3)$,

$R_{WPCA-NN,KLCI}(35.5) > R_{Stochastics,KLCI}(-8.4)$,

$R_{WPCA-NN,KOSPI}(32.2) > R_{Stochastics,KOSPI}(-22.7)$,

$R_{WPCA-NN,NIKKEI}(29.9) > R_{Stochastics,NIKKEI}(-6.8)$,

$R_{WPCA-NN,SiMSCI}(39.8) > R_{Stochastics,SiMSCI}(-12.2)$,

$R_{WPCA-NN,S\&P500}(31.4) > R_{Stochastics,S\&P500}(4.7)$, and

$R_{WPCA-NN,TAIEX}(33.1) > R_{Stochastics,TAIEX}(2.7)$).

OMA:

($R_{WPCA-NN,HSI}(21.7) > R_{SMA,HSI}(-0.7)$,

$R_{WPCA-NN,KLCI}(35.5) > R_{SMA,KLCI}(6.4)$,

$R_{WPCA-NN,KOSPI}(32.2) > R_{SMA,KOSPI}(-10.6)$,

$R_{WPCA-NN,NIKKEI}(29.9) > R_{SMA,NIKKEI}(-9.1)$,

$R_{WPCA-NN,SiMSCI}(39.8) > R_{SMA,SiMSCI}(16)$,

$R_{WPCA-NN,S\&P500}(31.4) > R_{SMA,S\&P500}(-8.2)$, and

$R_{WPCA-NN,TAIEX}(33.1) > R_{SMA,TAIEX}(-1.6)$).

According to the t-test, there are significant differences between the return means of the WPCA-NN model and the tested technical trading rules, MACD, RSI, stochastics, and OMA. Based on the significant and positive alphas in the regression analyses of the WPCA-NN model, the WPCA-NN model achieves significantly higher returns than the best-performing trading strategies in the literature (Chong et al., 2010; Chong & Ng, 2008; Fernández-Blanco et al., 2008; Rosillo et al., 2013) as follows; MACD ($\alpha_{HangSeng} = 34.22 > 0$, $\alpha_{KLCI} = 47.72 > 0$, $\alpha_{KOSPI} = 44.33 > 0$, $\alpha_{Nikkei} = 42.18 > 0$, $\alpha_{SiMSCI} = 55.32 > 0$, $\alpha_{S\&P500} = 48.3 > 0$, and $\alpha_{TAIEX} = 45.39 > 0$); RSI ($\alpha_{HangSeng} = 35.24 > 0$, $\alpha_{KLCI} = 47.56 > 0$, $\alpha_{KOSPI} = 43.32 > 0$, $\alpha_{Nikkei} = 42.48 > 0$, $\alpha_{SiMSCI} = 55.04 > 0$, $\alpha_{S\&P500} = 49.02 > 0$, and $\alpha_{TAIEX} = 45.63 > 0$); stochastics ($\alpha_{HangSeng} = 42.54 > 0$, $\alpha_{KLCI} = 40.13 > 0$, $\alpha_{KOSPI} = 44.82 > 0$, $\alpha_{Nikkei} = 41.47 > 0$, $\alpha_{SiMSCI} =$

49.45 > 0, $\alpha_{S\&P500} = 48.28 > 0$, and $\alpha_{TAIEX} = 44.82 > 0$); and OMA ($\alpha_{HangSeng} = 33.62 > 0$, $\alpha_{KLCI} = 46.26 > 0$, $\alpha_{KOSPI} = 44.1 > 0$, $\alpha_{Nikkei} = 41.8 > 0$, $\alpha_{SiMSCI} = 51.48 > 0$, $\alpha_{S\&P500} = 47.89 > 0$, and $\alpha_{TAIEX} = 44.15 > 0$). Significant mean differences and significant positive alphas in the profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other technical trading rules (i.e. MACD, RSI, stochastics, and OMA).

According to all futures markets' results, the trained networks for all three intelligent models (NN, WNN, and WPCA-NN) are valid because their MAPE values are low and acceptable. The WPCA-NN model outperforms the NN model over the entire out-sample period, whereby the lower the MAPE ratio the better, as follows:

($MAPE_{WPCA-NN,HSI}(4.06) > MAPE_{NN,HSI}(579.4)$,

$MAPE_{WPCA-NN,KLCI}(2.17) > MAPE_{NN,KLCI}(17.27)$,

$MAPE_{WPCA-NN,KOSPI}(2.24) > MAPE_{NN,KOSPI}(41.99)$,

$MAPE_{WPCA-NN,NIKKEI}(4.05) > MAPE_{NN,NIKKEI}(451.89)$,

$MAPE_{WPCA-NN,SiMSCI}(1.14) > MAPE_{NN,SiMSCI}(7.26)$,

$MAPE_{WPCA-NN,S\&P500}(2.34) > MAPE_{NN,S\&P500}(29.53)$, and

$MAPE_{WPCA-NN,TAIEX}(2.71) > MAPE_{NN,TAIEX}(143.7)$).

Further, the WPCA-NN model outperforms the WNN model over the whole out-sample period, whereby the lower the MAPE ratio the better, as follows:

($MAPE_{WPCA-NN,HSI}(4.06) > MAPE_{WNN,HSI}(15.44)$,

$MAPE_{WPCA-NN,KLCI}(2.17) > MAPE_{WNN,KLCI}(2.5)$,

$MAPE_{WPCA-NN,KOSPI}(2.24) > MAPE_{WNN,KOSPI}(2.8)$,

$MAPE_{WPCA-NN,NIKKEI}(4.05) > MAPE_{WNN,NIKKEI}(10.88)$,

$MAPE_{WPCA-NN,SiMSCI}(1.14) > MAPE_{WNN,SiMSCI}(1.79)$,

$MAPE_{WPCA-NN,S\&P500}(2.34) > MAPE_{WNN,S\&P500}(2.35)$, and

$MAPE_{WPCA-NN,TAIEX}(2.71) > MAPE_{WNN,TAIEX}(5)$).

Moreover, the WPCA-NN model achieves higher returns than other intelligent techniques without transaction costs, as follows:

NN:

($R_{WPCA-NN,HSI}(34.8) > R_{NN,HSI}(12.6)$,

$R_{WPCA-NN,KLCI}(47.2) > R_{NN,KLCI}(21.7)$,

$R_{WPCA-NN,KOSPI}(44.8) > R_{NN,KOSPI}(17.8)$,

$R_{WPCA-NN,NIKKEI}(42.1) > R_{NN,NIKKEI}(13.9)$,

$R_{WPCA-NN,SiMSCI}(54.4) > R_{NN,SiMSCI}(19.9)$,

$R_{WPCA-NN,S\&P500}(48.2) > R_{NN,S\&P500}(20.5)$, and

$R_{WPCA-NN,TAIEX}(45.3) > R_{NN,TAIEX}(8.6)$).

WNN:

($R_{WPCA-NN,HSI}(34.8) > R_{WNN,HSI}(27.9)$,

$R_{WPCA-NN,KLCI}(47.2) > R_{WNN,KLCI}(40.2)$,

$R_{WPCA-NN,KOSPI}(44.8) > R_{WNN,KOSPI}(40.8)$,

$R_{WPCA-NN,NIKKEI}(42.1) > R_{WNN,NIKKEI}(34.9)$,

$R_{WPCA-NN,SiMSCI}(54.4) > R_{WNN,SiMSCI}(40.6)$,

$R_{WPCA-NN,S\&P500}(48.2) > R_{WNN,S\&P500}(41.7)$, and

$R_{WPCA-NN,TAIEX}(45.3) > R_{WNN,TAIEX}(35.3)$).

In addition, when all trading is tested with transaction costs, the WPCA-NN model still achieves higher returns than the NN model, as follows:

($R_{WPCA-NN,HSI}(21.7) > R_{NN,HSI}(7.8)$,

$R_{WPCA-NN,KLCI}(35.5) > R_{NN,KLCI}(15.3)$,

$R_{WPCA-NN,KOSPI}(32.2) > R_{NN,KOSPI}(11.7)$,

$R_{WPCA-NN,NIKKEI}(29.9) > R_{NN,NIKKEI}(7.9)$,

$R_{WPCA-NN,SiMSCI}(39.8) > R_{NN,SiMSCI}(13.2)$,

$R_{WPCA-NN,S\&P500}(31.4) > R_{NN,S\&P500}(13.4)$, and

$R_{WPCA-NN,TAIEX}(33.1) > R_{NN,TAIEX}(1.5)$).

Further, when all trading is tested with transaction costs, the WPCA-NN model still achieves higher returns than the WNN model, as follows:

($R_{WPCA-NN,HSI}(21.7) > R_{WNN,HSI}(17.2)$,

$R_{WPCA-NN,KLCI}(35.5) > R_{WNN,KLCI}(29.0)$,

$R_{WPCA-NN,KOSPI}(32.2) > R_{WNN,KOSPI}(25.7)$,

$R_{WPCA-NN,NIKKEI}(29.9) > R_{WNN,NIKKEI}(22.6)$,

$R_{WPCA-NN,SiMSCI}(39.8) > R_{WNN,SiMSCI}(22.8)$,

$R_{WPCA-NN,S\&P500}(31.4) > R_{WNN,S\&P500}(26.8)$, and

$R_{WPCA-NN,TAIEX}(33.1) > R_{WNN,TAIEX}(23.5)$).

The NN model performs very poorly compared with the other models. The results of the WPCA-NN and WNN models are more precise, with much lower error terms, from 2005 to 2014.

According to the t-test, there are significant differences between the return means of the WPCA-NN model and other intelligent techniques such as the NN and WNN models. In addition, according to the significant and positive alphas in the regression analyses of the WPCA-NN model and the two other intelligent techniques, NN ($\alpha_{HangSeng} = 36.60 > 0$, $\alpha_{KLCI} = 27.95 > 0$, $\alpha_{KOSPI} = 48.01 > 0$, $\alpha_{Nikkei} = 41.25 > 0$, $\alpha_{SiMSCI} = 57.63 > 0$, $\alpha_{S\&P500} = 49.42 > 0$, and $\alpha_{TAIEX} = 41.73 > 0$) and WNN ($\alpha_{HangSeng} = 27.78 > 0$, $\alpha_{KLCI} = 14.87 > 0$, $\alpha_{KOSPI} = 19.85 > 0$, $\alpha_{Nikkei} = 32.95 > 0$, $\alpha_{SiMSCI} = 26.43 > 0$, $\alpha_{S\&P500} = 31.32 > 0$, and $\alpha_{TAIEX} = 29.95 > 0$), the WPCA-NN model generates significantly higher returns than the pure NN and WNN models for all futures markets. This superiority of the WPCA-NN model over the pure NN and WNN models is meaningful and significant, with a 5% level of confidence. Thus, it is concluded that the proposed denoising process significantly and extremely increases trading performance because the regression between the WPCA-NN model's returns and the NN model's returns achieves positive and significant alphas. Moreover, the WPCA-NN model, with its multivariate denoising process, outperforms the WNN model, with its univariate denoising process, in all seven futures markets, according to the positive and significant alphas reported for the regression between the WPCA-NN returns and the WNN returns. The significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for all studied futures markets than the pure NN and WNN models' methods of forecasting.

Further, the denoising processes of multivariate wavelet-PCA and univariate wavelet both perform better than the NARX-NN (the so-called pure NN) model in terms of forecasting performance and trading profitability. Hence, not only is it useful to employ a denoising process before the NN model, wavelet transform analysis also proves to be a successful tool for denoising purposes. However, the multivariate denoising process embedded in the WPCA-NN model enhances the denoising procedure and results in more accurate forecasting and a more profitable trading system than the univariate denoising process.

Regarding this study's intelligent models, the NN model is the most sensitive to data fluctuations. However, the WPCA-NN and WNN models appear to be less sensitive to noise because of the application of wavelet transforms. Moreover, not only do the forecasting results of the WPCA-NN and WNN models fit better to the actual data (lower

MAPE values), they also achieve higher excess returns. Further, the WNN model is more sensitive to data fluctuations than the WPCA-NN model. Thus, the WPCA-NN model appears to result in a smoother version of the original signal. For some days, the forecasting results of the WNN model show large differences and incorrect market direction compared with the actual data. However, on the same days, the WPCA-NN model recognizes the changing direction of the market much more accurately and achieves predictions that fit better and higher returns compared with the WNN model.

Moreover, the average returns using the WPCA-NN model are significantly higher than the outcomes from prior studies (Cheng et al., 2010; Chiang & Doong, 2001; Hsieh et al., 2011; Huang et al., 2009; Kim et al., 1998; Lee et al., 2010; Leung et al., 2000; Necula, 2009; Quah & Srinivasan, 1999). This study's outcomes are completely consistent with other similar studies (Booth et al., 2014; Hu et al., 2015; Xiao et al., 2014; Zhang et al., 2014) on machine-learning using technical analysis indicators, where frequent accurate forecasting leads to excess returns over the buy-and-hold strategy.

The most notable crisis during the period of this study is the 2008 financial crisis. The trading returns of the WPCA-NN model during this financial crisis are higher than the returns of all other tested models for the NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets (Appendix F). However, during the financial crisis, the NN and WNN models and the MACD trading strategy achieve higher returns than the WPCA-NN model for the Hang Seng, KLCI, and KOSPI 200 futures markets respectively. Nonetheless, the WPCA-NN model achieves the second-highest gains in 2008 for these markets. Moreover, the best trading performance, or the lowest MAPE ratio, during the crisis is that of the WPCA-NN model compared with the other intelligent models. In addition, the WPCA-NN model generates higher returns than average for the KLCI (R_{2008}: 76.5 > $R_{average}$: 48.2), KOSPI 200 (R_{2008}: 52.2 > $R_{average}$: 44.8), NIKKEI 225

(R_{2008}: 53.1 > $R_{average}$: 43.1), SiMSCI (R_{2008}: 76.2 > $R_{average}$: 54.4), and S&P 500 (R_{2008}: 57.8 > $R_{average}$: 44.8) futures markets during the financial crisis. However, the forecasting performances during the crisis for all futures markets except the Hang Seng ($MAPE_{2008}$: 3.929 < $MAPE_{average}$: 4.065) are lower than the average. Ultimately, the proposed hybrid model performs better than the other intelligent models and achieves higher returns than the other trading methods during the financial crisis. In addition, the return results of the model are promising and reliable during this period. Comparisons of the years before and after the financial crisis for selected futures markets are as follow:

Hang Seng (R_{07}: 35.6, R_{08}: 23 *and* R_{09}: 10.6),

KLCI (R_{07}: 75.3, R_{08}: 76.5 *and* R_{09}: 57),

KOSPI 200 (R_{07}: 61.8, R_{08}: 52.2 *and* R_{09}: 71.6),

NIKKEI 225 (R_{07}: 56.3, R_{08}: 53.1 *and* R_{09}: 39.3),

SiMSCI (R_{07}: 78.5, R_{08}: 76.2 *and* R_{09}: 75.1),

S&P 500 (R_{07}: 34, R_{08}: 57.8 *and* R_{09}: 60.7), and

TAIEX (R_{07}: 24.6, R_{08}: 22.9 *and* R_{09}: 60.5).

Results show that the returns of the WPCA-NN model are not only not extremely higher or lower than the year before and the year after the crisis, they are also not necessarily higher or lower. Thus, the analysis of the WPCA-NN model's returns during the 2008 financial crisis does not suggest any spike, or any spurious impact on profitability, of the WPCA-NN model. Instead, there is evidence of reliable and robust profitability compared with the returns of the year before, the returns of the year after, and the average returns for all futures markets.

The following review is presented in order to compare the results of the current study to other similar studies in the seven selected futures markets. In one prior study, a generalized hyperbolic distribution was applied to forecast the Hang Seng index. This method achieved a 5.3% annual return (Necula, 2009). A further study performed a hierarchical coevolutionary fuzzy predictive model and gained a 14.25% return (Huang et al., 2009). In addition, an ANN and ARIMA were applied to the KLCI futures market based on technical indicators. These methods achieved average annual returns of 20.8% and 15.29% respectively (Yao et al., 1999). An ANN and case-based reasoning were applied to the KOSPI 200 index and achieved an annual return of 40.9% (Kim et al., 1998), while a real-time rule-based trading system achieved 28.57% (Lee et al., 2010). Moreover, a 17.2% annual return by the use of discriminant analysis and 13.78% by the use of a multilayered feedforward neural network were achieved for the NIKKEI 225 index (Leung et al., 2000). A generalized hyperbolic distribution was applied to model the NIKKEI 225 futures market and achieved a 8.6% annual return (Necula, 2009). Moreover, an NN model was used to forecast the Singapore index and achieved a 25.64% annual return (Quah & Srinivasan, 1999), while an approximately 15% annual return was achieved by using generalized autoregressive conditional heteroscedasticity (Chiang & Doong, 2001). In addition, an 18.96% annual return was realized with the application of a generalized regression NN on the S&P 500 futures market (Enke & Thawornwong, 2005).

In the following, the parameters of the best-performing networks in each of the futures markets are presented. This study not only suggests that the introduced combination of parameters is the most appropriate setting for forecasting in each particular futures markets, it also makes the point that any other combination of these parameters performs comparably better than others because the parameters represent common characteristics of the market across 10 years. Although these findings are robust and valid in different evaluation subsets and periods, they relate to the characteristics of each specific futures market and may change in the future. In addition, based on the forecasting response in evaluation periods, the WPCA-NN model generates the best-fitted output of the target series of the NIKKEI 225 futures market compared with the NN and WNN models.

According to the best-performing networks in the Hang Seng futures market, three levels of decomposition, a coif5 wavelet, and a penalized high thresholding strategy achieve continuously high performance and excess returns with the WPCA-NN model. Based on the best-performing networks in the evaluation results of the KLCI futures market, three levels of decomposition, a db9 wavelet, a penalized high thresholding strategy, and a delay of five days achieve considerably high performance and returns with the WPCA-NN model. The best-performing networks in the KOSPI 200 futures market employ three or four levels of decomposition, sym6 or db9 wavelets, and a penalized high thresholding strategy. These networks repeatedly achieve significantly high performance and excess returns over the buy-and-hold benchmark with the WPCA-NN model. According to the best-performing networks in the NIKKEI 225 futures market, two levels of decomposition, a coif5 wavelet, and a penalized high thresholding strategy achieve considerably high performance and excess returns with the WPCA-NN model. The best-performing networks in the SiMSCI futures market use four levels of decomposition, a db7 wavelet, and a penalized high thresholding strategy to achieve significantly high performance and returns with the WPCA-NN model. According to the best-performing networks in the S&P 500 futures market, three levels of decomposition, a db9 wavelet, and a penalized high thresholding strategy achieve significantly high performance and excess returns with the WPCA-NN model. The best-performing networks in the TAIEX futures market use three levels of decomposition, a db9 wavelet, a penalized high thresholding strategy, and a delay of one week to achieve considerably high performance and returns repeatedly with the WPCA-NN model. Although these results are robust and valid in different evaluation subsets and years, they relate to the characteristics of the markets and may vary in the future.

Table 4.38: Summary of the results

	Objectives	Hypotheses	Model/Technique	Results
1	To investigate whether the futures markets exhibit significant changes in predictability over time, considering long-term memory and approximate entropy.	H1: Futures markets exhibit significant changes in predictability over time, considering long-term memory and approximate entropy.	Hurst exponent. Approximate entropy.	All futures markets show long-term memory with a Hurst exponent greater than 0.5 and less than 1. Considering structural breaks, the Hurst exponent achieves true results. Considering long-term memory as a predictability or efficiency index of a market, all futures markets have long-term memory variation over time. All studied futures markets display predictability with approximate entropy as another measure for efficiency and predictability. The inefficiency and predictability of all futures markets varies over time.
2	To determine whether the hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the RWH (a passive buy-and-hold strategy) for selected financial markets.	H2: The hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the RWH (a passive buy-and-hold strategy) for selected financial markets as shown by significant mean differences and significant positive alphas in profit regressions.	Multivariate wavelet denoising using PCA plus forecasting with the NARX-NN model (WPCA-NN). Buy-and-hold strategy benchmark.	The WPCA-NN model consistently generates significantly higher returns than forecasted by the RWH for all futures markets. The differences between their means are significant for all markets. The alphas in the profit regression model between the returns of the WPCA-NN model and the buy-and-hold strategy are positive and significant for all markets. Selected futures markets do not follow random walk from 2005–2014.
3	To investigate, test, and find whether the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other technical analysis indicators (such as optimized moving average).	H3: Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other technical trading rules (MACD, RSI, stochastics, and OMA).	Multivariate wavelet denoising using PCA plus forecasting with the NARX-NN model (WPCA-NN). Technical trading rules: MACD, RSI, Stochastics, and OMA.	The WPCA-NN model consistently generates significantly higher returns than predicted by MACD, RSI, stochastics, and OMA trading strategies for all futures markets. The differences between the means of the WPCA-NN model and other technical trading strategies are significant for all markets. The alphas in the profit regression model between the returns of the WPCA-NN model and other technical trading strategies are positive and significant for all markets.
4a	To investigate, test, and find whether the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than the pure NN method of forecasting.	H4a: Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than the pure NN method of forecasting.	Multivariate wavelet denoising using PCA plus forecasting with the NARX-NN model (WPCA-NN). Nonlinear autoregressive with exogenous input neural network (NARX-NN).	The WPCA-NN model consistently generates significantly higher returns than predicted by the NN model for all futures markets. The differences between their means are significant for all markets. The alphas in the profit regression model between the returns of the WPCA-NN model and the NN model are positive and significant for all markets. The WPCA-NN model's results fit better to the target time series or actual data compared with the NN model. The denoising process with wavelet transform highly assists the prediction procedure with the ANN technique.
4b	To investigate, test, and find whether the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than the WNN method of forecasting.	H4b: Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than the WNN method of forecasting	Multivariate wavelet denoising using PCA plus forecasting with the NARX-NN model (WPCA-NN). Univariate wavelet denoising plus forecasting with the NARX-NN model (WNN).	The WPCA-NN model consistently generates significantly higher returns than forecasted by the WNN model for all futures markets. The differences between their means are significant for all markets. The alphas in the profit regression model between the returns of the WPCA-NN model and the WNN model are positive and significant for all markets. The WPCA-NN model's results fit better to the actual data compared with the WNN model. The enhancement of the denoising process by multivariate denoising using a wavelet and PCA results in a promising technique.

CHAPTER 5: CONCLUSION

5.1 Summary

The essential need to perform human tasks with the least expense and fastest speed, along with the requirement to process large amounts of data, supports the growth of computational intelligent methods in numerous scientific fields, including finance. In addition, the fact that financial prediction is inherently associated with the extreme degree of uncertainty controlling the modern world, such adaptive models can be capable alternatives to conventional models. Further, to succeed, financial academicians must put a great deal of effort into accurate computational science-specific calibration. The scope of this study is to generalize the implementation of these models by combining the virtues of some computational intelligence techniques into superior hybrid systems.

The predictability and inefficiency of the selected markets are analysed with unit root and serial correlation tests, long-term memory and approximate entropy. According to these analyses and their significant results, all selected futures markets are concluded to be inefficient and predictable during 2005-2014 (Hypothesis H1). Therefore, they follow the adaptive market hypothesis and are predictable. According to the literature, conventional and traditional forecasting models have reached to their limitations, and artificial intelligent methods proved to be significantly successful in forecasting. ANNs are one of the best-performing artificial intelligent methods, but have deficiencies to noises in financial time series. Hence, wavelet analysis is entered to the preprocessing part of the method to denoise the financial time series. The combination of wavelet denoising and ANN forms the forecasting model, namely WNN. In addition to that, an enhancement has been done to the denoising part by including PCA in order to analyse the OHLC as a multivariate signal. This enhancement formed a novel combination with ANN, namely WPCA-NN. ANN, WNN and WPCA-NN are examined with each other and some traditional technical strategies such as MACD, RSI, Stochastics, OMA and passive buy-and-hold. The results represent the superiority of WPCA-NN over other models and strategies. The next Section (5.2) begins with the questions and objectives of the study, and briefly explains all the conclusions taken from the results. Summary of the conclusion is presented in Table 5.1.

Table 5.1: Summary of conclusion

	Hypotheses	Model/Tests	Results	Conclusions
1	H1: Futures markets exhibit significant changes in predictability over time, considering long-term memory and approximate entropy.	Hurst exponent. Approximate entropy.	All futures markets show long-term memory with a Hurst exponent greater than 0.5 and less than 1. Considering structural breaks, the Hurst exponent achieves true results. Considering long-term memory as a predictability or efficiency index of a market, all futures markets have long-term memory variation over time. All studied futures markets display predictability with approximate entropy as another measure for efficiency and predictability. The inefficiency and predictability of all futures markets varies over time.	The long memory results exhibit serial dependence and aperiodic long cycles in these futures markets. The results of approximate entropy also confirm the predictability and inefficiency of the futures markets from 2005 to 2014. Moreover, the results of statistical tests approve the inefficiency of the futures markets during the study. The findings from each of the analyses propose a strong presence of nonlinear dependence in the movement of futures markets throughout the studied period, suggesting possible predictability of price movement and consequent excess returns. Further, the studied futures markets are still evolving because they have not gone through a particular period of efficiency. Hence, the futures markets do not follow random walk from 2005 to 2014. However, the time-varying index of long memory and approximate entropy over time in all the studied futures markets supports adaptive market hypothesis, a finding that is consistent with similar studies (Kristoufek, 2012; Kristoufek & Vosvrda, 2014; Lo, 2005).
2	H2: The hybrid WPCA-NN model consistently generates significantly higher returns than predicted by the RWH (a passive buy-and-hold strategy) for selected financial markets as shown by significant mean differences and significant positive alphas in profit regressions.	Multivariate wavelet denoising using PCA plus forecasting with the NARX-NN model (WPCA-NN). Buy-and-hold strategy benchmark.	The WPCA-NN model consistently generates significantly higher returns than forecasted by the RWH for all futures markets. The differences between their means are significant for all markets. The alphas in the profit regression model between the returns of the WPCA-NN model and the buy-and-hold strategy are positive and significant for all markets. Selected futures markets do not follow random walk from 2005–2014.	These outcomes are significant in all studied markets and during the whole testing period, so, they confirm the inefficiency and predictability of these markets during the period of the study and also confirm the results of long-term memory, approximate entropy, and the adaptive market hypothesis.
3	H3: Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than other technical trading rules (MACD, RSI, stochastics, and OMA).	Multivariate wavelet denoising using PCA plus forecasting with the NARX-NN model (WPCA-NN). Technical trading rules: MACD, RSI, Stochastics, and OMA.	The WPCA-NN model consistently generates significantly higher returns than predicted by MACD, RSI, stochastics, and OMA trading strategies for all futures markets. The differences between the means of the WPCA-NN model and other technical trading strategies are significant for all markets. The alphas in the profit regression model between the returns of the WPCA-NN model and other technical trading strategies are positive and significant for all markets.	These findings will add to the existing literature of comparison between traditional and intelligent techniques, showing the hybrid intelligent systems significantly outperform conventional technical trading rules.

4a	H4a: Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than the pure NN method of forecasting.	Multivariate wavelet denoising using PCA plus forecasting with the NARX-NN model (WPCA-NN). Nonlinear autoregressive with exogenous input neural network (NARX-NN).	The WPCA-NN model consistently generates significantly higher returns than predicted by the NN model for all futures markets. The differences between their means are significant for all markets. The alphas in the profit regression model between the returns of the WPCA-NN model and the NN model are positive and significant for all markets. The WPCA-NN model's results fit better to the target time series or actual data compared with the NN model. The denoising process with wavelet transform highly assists the prediction procedure with the ANN technique.	The superiority of WPCA-NN and WNN over NN in both forecasting performance and trading profitability indicates the significant denoising work undertaken by wavelet transform, since both WPCA-NN and WNN include the wavelet denoising process. The ability of wavelet transforms to analyze a signal in two domains of frequency and time, as well as signal denoising using threshold estimators, helps the hybrid models to significantly overcome noise in financial time series. This finding adds to the existing literature and confirms not only the limitation of the existence of noise in ANNs in financial time series, but also the strength of wavelet analysis in denoising financial time series to identify the series' underlying functions.
4b	H4b: Significant mean differences and significant positive alphas in profit regressions show that the WPCA-NN model consistently generates significantly higher returns and achieves higher forecasting performance for futures markets than the WNN method of forecasting	Multivariate wavelet denoising using PCA plus forecasting with the NARX-NN model (WPCA-NN). Univariate wavelet denoising plus forecasting with the NARX-NN model (WNN).	The WPCA-NN model consistently generates significantly higher returns than forecasted by the WNN model for all futures markets. The differences between their means are significant for all markets. The alphas in the profit regression model between the returns of the WPCA-NN model and the WNN model are positive and significant for all markets. The WPCA-NN model's results fit better to the actual data compared with the WNN model. The enhancement of the denoising process by multivariate denoising using a wavelet and PCA results in a promising technique.	However, the superiority of WPCA-NN over WNN in both forecasting performance and trading profitability suggests that multivariate denoising of WPCA on OHLC enhances the denoising process and results in more accurate forecasting and a more profitable trading system. Open, high, low, and close signals are successfully considered as a multivariate signal. Further, multivariate denoising, namely WPCA, of OHLC as a multivariate signal increases the precision of the estimation of the main features and consequently enhances the denoising process, which finally results in more accurate forecasting and a more profitable trading system. In other words, the denoising process estimated the original signals better by eliminating the same noise from all open, high, low, and close signals. This finding will add to the existing literature of wavelet analysis and multivariate signals. Moreover, considering open, high, low, and close signals as a multivariate signal helps the denoising process and, consequently, forecasting ability. These findings not only will add to financial prediction and planning, but also will be helpful in other finance fields where hybrid intelligent systems are needed, such as credit valuation and portfolio management.

5.2 Conclusion

Due to its integral role within the fields of economics and finance, forecasting financial markets has been the primary focus of this thesis. The literature review identified numerous fundamental areas linked to fluctuations in the financial markets, such as efficiency (Fama, 1991, 1965; Taylor, 1986; Thaler, 1985; Lo 1989), forecasting ability (Gencay, 1996; Lo, 1989, 2004, 2005; Mandelbrot, 1971; Mandelbrot & Hudson, 2005), noise (Chang et al., 2004; Ramsey & Zhang, 1997; Aminghafari et al., 2006; Donoho, 1995; Hsieh et al., 2011), and profitability (Gencay, 1999; Jin & Kim, 2015; Lai et al., 2007; Lotrič, 2004; Ortega, 2012; Wang & Gupta, 2013) of various prediction models.

The four empirical questions related to these areas within this topic laid out for analysis are:

Question 1: Do futures markets, particularly market index futures, have long-term memory and approximate entropy?

Question 2: If so, can the proposed novel model, namely WPCA-NN, consistently generate abnormal higher returns greater than those advocated by the RWH for selected futures markets?

Question 3: Does the WPCA-NN model consistently generate significantly higher returns for futures markets than the best performing technical analysis indicators?

Question 4: Does the WPCA-NN model consistently generate significantly higher returns and achieve higher forecasting performance for futures markets than other intelligent forecasting techniques such as: a) pure NN and b) WNN?

This study investigates the predictability and efficiency of particular futures markets. To perform the analysis in different economic conditions, futures markets of some developing and developed countries are selected. The predictability and efficiency of the futures markets are examined with long-term memory and approximate entropy in addition to statistical tests, such as unit root and serial correlation tests. The studied futures markets are proven to have long-term memory based on Hurst exponent results (Question 1), which exhibit serial dependence and aperiodic long cycles in these futures markets (Cheung & Lai, 1995). The results of approximate entropy also confirm the predictability and inefficiency of the futures markets from 2005 to 2014. Moreover, the results of statistical tests prove the inefficiency of the futures markets during the study. The findings from each of the analyses propose a strong presence of nonlinear dependence in the movement of futures markets throughout the studied period, suggesting possible predictability of price movement and consequent excess returns. Further, the studied futures markets are still evolving because they have not gone through a specific period of efficiency. Hence, the futures markets do not follow a random walk from 2005 to 2014. However, the time-varying index of long memory and approximate entropy over time in all the studied futures markets supports the adaptive market hypothesis, a finding that is consistent with similar studies (Kristoufek, 2012; Kristoufek & Vosvrda, 2014; Lo, 2005). Since the adaptive market hypothesis supports the selected futures markets in the period of the study, they can be predicted using forecasting tools. Based on the literature review, and the noisy nature of financial time series, success of modern intelligent techniques, and proven predictability of the selected futures markets using AMH, long-term memory and approximate entropy, this study offers a hybrid intelligent technique to predict future movements of these futures markets.

According to the literature review, an artificial neural network, one of the well-known mathematical prediction techniques, is used for the core element of the hybrid system. The novel hybrid trading system, namely WPCA-NN, consists of multivariate denoising using Wavelet-PCA preprocessing and a nonlinear autoregressive neural network. WPCA-NN is developed in four stages; (1) data preprocessing using wavelet analysis and PCA as a multivariate denoising technique, which is applied to decompose the futures' price time series to eliminate the same noise from the OHLC signals, (2) the use of some technical indicators and denoised OHLC signals to construct the input series selected via the backward elimination technique, (3) the application of a nonlinear autoregressive neural network with exogenous inputs, and (4) the use of a simple trading strategy to provide more empirical results. This novel technique is required to address the rest of the suggested research questions.

The proposed trading system, WPCA-NN, is compared with the random walk hypothesis (Fama, 1965), and proved to generate significantly higher returns than the buy-and-hold strategy in all selected futures markets to address Question 2. These findings are significant in all studied markets and during the whole testing period, therefore, they show the inefficiency and predictability of these markets during the period of the study and confirm the results of long-term memory, approximate entropy, and the adaptive market hypothesis.

To answer the third question, WPCA-NN is compared to MACD, RSI, Stochastics, and OMA, which shows our novel model generates higher returns than traditional technical indicators. These findings will add to the existing literature of comparison between traditional and intelligent techniques, showing the hybrid intelligent systems significantly outperform conventional technical trading rules.

The proposed trading system, WPCA-NN, is also compared with NARX-NN (Abdulkadir & Yong, 2014; Busse et al., 2012; Diaconescu, 2008; Shahbazi et al., 2016; Siegelmann et al., 1997), and WNN (Aminghafari et al., 2006; Wang & Gupta, 2013) to address the fourth question. NARX-NN is considered the pure neural network and WNN is the generalized form of univariate denoising in multiple one-dimensional signals combined with a neural network. To show that WPCA-NN is sufficiently robust, this trading system has been applied to seven futures markets, the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures, over 10 years (2005–2014). Simulation results indicate that a WPCA-NN trading system outperforms other intelligent models, such as NN and WNN. Moreover, WPCA-NN earns significantly greater returns than NN (Question 4a) and WNN (Question 4b). Therefore, adding a preprocessing feature of wavelet denoising to the NN increases the forecasting ability and profitability of the pure NN models.

The superiority of WPCA-NN and WNN over NN in both forecasting performance and trading profitability indicates the significant denoising work undertaken by wavelet transform (Question 4a), since both WPCA-NN and WNN include the wavelet denoising process. The ability of wavelet transforms to analyze a signal in two domains of frequency and time, as well as signal denoising using threshold estimators, helps the hybrid models to significantly overcome noise in financial time series. This finding adds to the existing literature and confirms not only the limitation of the existence of noise in ANNs in financial time series, but also the strength of wavelet analysis in denoising financial time series to identify the series' underlying functions.

However, the superiority of WPCA-NN over WNN in both forecasting performance and trading profitability suggests that multivariate denoising of WPCA on OHLC enhances the denoising process and results in more accurate forecasting and a more profitable trading system (Question 4b). Open, high, low, and close signals are successfully considered as a multivariate signal. Further, multivariate denoising, namely WPCA, of OHLC as a multivariate signal increases the precision of the estimation of the main features and consequently enhances the denoising process, which finally results in more accurate forecasting and a more profitable trading system. In other words, the denoising process estimated the original signals better by eliminating the same noise from all open, high, low, and close signals. This finding will add to the existing literature of wavelet analysis and multivariate signals. Moreover, considering open, high, low, and close signals as a multivariate signal helps the denoising process and, consequently, forecasting ability. These findings not only will add to financial prediction and planning, but also will be helpful in other finance fields where hybrid intelligent systems are needed, such as credit valuation and portfolio management.

Thus, this study successfully proposed a hybrid model to overcome traditional trading strategies, the best-performing technical strategies, such as MACD, RSI, stochastics, and OMA, pure NN, and WNN. The superiority of WPCA-NN contributes to the literature of neural networks, wavelet analysis, denoising procedures, the RWH, and traditional technical trading strategies as another successful hybrid of ANNs. Further, the average returns using WPCA-NN are significantly higher than results from prior studies (Cheng et al., 2010; Chiang & Doong, 2001; Hsieh et al., 2011; Huang et al., 2009; Kim et al., 1998; Lee et al., 2010; Leung et al., 2000; Necula, 2009; Quah & Srinivasan, 1999).

The results are entirely consistent with similar studies (Booth et al., 2014; Hu et al., 2015; Xiao et al., 2014; Zhang et al., 2014) on machine learning using technical analysis indicators, whereby frequent accurate predictions lead to excess returns. The findings of this research indicate that, in view of the growing efficiency of rapidly growing financial markets that have moved beyond traditional technical analysis tools (Coronel-Brizio et al., 2007; Olson, 2004), machine-learning trading systems may be a novel profitable strategy for trading futures contracts. It can be concluded that a significant implication arising from the findings of this study is that for at least three months ahead, traders in Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures could use the best possible settings obtained by repeated simulations of the prior three years for a promising new trading system, WPCA-NN, to combat the volatile nature of these rapidly changing markets. Moreover, this study's model can be applied in other fields that include multivariate signals and require forecasting tools to expand their boundaries.

5.3 Significance of the Study

Overall, this thesis addresses problems and provides extensions to the knowledge of the field of finance. Although universal estimations have never been found in scientific research, earlier chapters have strong implications for decision-making. Moreover, more light is shed on the demanding issue of achieving statistical and trading profitability in the futures markets using computational intelligent techniques. Hedge fund managers and traders should experiment beyond the imitations of traditional and conventional models. Their trading decisions should be based on modern expectations from techniques and strategies that are improved using a hybrid trading and statistical approach. Government, financial institutions, and central banking policies can also be affected in the same context. All these entities have endless concerns related to monitoring large numbers of indicators that could affect inflation or unemployment, and, consequently, the economy. The models proposed could be found particularly worthwhile in these monitoring procedures. However, there are still many roads to be taken in the exploration of efficient adjustment of computational intelligent techniques for financial forecasting tasks.

For at least the next three months, traders in the Hang Seng, KLCI, KOSPI 200, NIKKEI 225, SiMSCI, S&P 500, and TAIEX futures markets could use the best settings obtained by repeated simulations of the prior three years in a promising new trading system, WPCA-NN, to combat the volatile nature of the rapidly changing markets. In addition, traders in other financial markets can use this novel hybrid technique.

This study contributes to the existing literature of efficiency of the selected markets, long-term memory and investigating spurious results, approximate entropy, changes in inefficiency and predictability of the markets, adaptive market hypothesis, artificial intelligent techniques, NN's hybrid, wavelets, denoising and finally forecasting financial time series. This research offers a successful hybrid forecasting machine, which can be used not only in financial fields, such as credit evaluation, portfolio management, and market prediction, but also in other fields such as analysis of population, physical systems, face and voice recognition, weather forecasting, cancer pattern recognition, and so on. Another significance of this study is the enhancement has been done in denoising procedure. Considering open, high, low and close as a multivariate signal and denoising them using the approach of wavelet-PCA results in promising results for this enhancement. This approach can be used not only in financial forecasting, but also in other fields of science, which may require denoising in preprocessing. The approach would help because the results show that complicated and noisy time series with a multivariate signals nature can be modeled and identified with the complex combination of a wavelet, PCA, and an artificial neural network.

5.4 Suggestions for Future Studies

Although the proposed forecasting machine represents promising outcomes, it still possesses some limitations. The selected futures markets were chosen because, first, this research is investigating a novel approach with a hybrid system; thus, it is necessary to test the system with a range of emerging markets and developed markets for generalizability and stability. Second, they are typical markets that an Asian fund manager in the Asian time zone would trade in. Although, there are lots of other regions and markets that would make this study more interesting and generalized, adding markets to the research is computationally expensive and time consuming. Implementing the proposed hybrid model to other futures indices and markets such as stock, commodity and Forex would be interesting to scholars and traders, which is highly suggested for futures study.

In live trading, other than transaction costs, slippage bothers algorithm trading and may lower the performance of the proposed forecasting machine. In the moments of illiquidity in the market, slippages are high and may affect the trading results. However, it totally depends on the volume of trading in each futures market. A live trading test of the proposed model is recommended in order to investigate the actual trading results. However, it is highly time consuming.

Although the profitability results are promising and better than some academic research, there are approximately the same or better profitability outcomes in literature and in the real world. For example, a hybrid model using an approach based on a firefly algorithm, built on an established multi-output support vector regression, could be used to forecast financial time series (Xiong et al., 2014). The authors achieve promising forecasting performance and average annualized returns for the S&P 500 (58.90%), FTSE 100 (50.01%), and Nikkei 225 (51.44%). Another study achieves an average annual return of approximately 40% for the S&P 500, with the application of a support vector machine (SVM) and reinforcement learning (Shen et al., 2012). However, there are no reports regarding the reliability of the strategies and risk adjustments. Other authors have performed a study on out-of-sample and post-publication profitability and predictability using 82 models that are proposed in published academic research (McLean & Pontiff, 2016). The average post-publication decay is about 35%, which the authors attribute to price pressure from aware investors and statistical bias. Nonetheless, these types of achievement and promising results from a literature review provide significant motivation for academicians to continue spending time on this field of study.

Although the proposed hybrid system achieves promising prediction results, it still possesses some deficiencies, which means that the approach could be investigated further. Other than the suggested approaches for future studies in the previous section (Section 5.4), a different intelligent ensemble model, such as a support vector machine (SVM) or adaptive neuro-fuzzy inference system (ANFIS) with different algorithms, such as genetic or bee colony algorithms, could be employed together with wavelet-PCA to address the forecasting problems of financial time series. In addition, the selection of best-performing networks in this current study has two different destinies via statistical performance (e.g., MAPE value) and trading strategies (return). Thus, suggested topics for future studies are the analysis of statistical and trading performance, and the establishment of a more precise method to choose the best network and setting from the results.

REFERENCES

Abdulkadir, S. J., & Yong, S.-P. (2014). *Empirical analysis of parallel-NARX recurrent network for long-term chaotic financial forecasting.* Paper presented at the International Conference on Computer and Information Sciences (ICCOINS), 2014.

Abdulkadir, S. J., Yong, S.-P., Marimuthu, M., & Lai, F.-W. (2014). Hybridization of ensemble Kalman filter and non-linear auto-regressive neural network for financial forecasting *Mining Intelligence and Knowledge Exploration* (pp. 72-81): Springer.

Abreu, D., & Brunnermeier, M. K. (2003). Bubbles and crashes. *Econometrica, 71*(1), 173-204.

Adebiyi, A., Ayo, C., Adebiyi, M., & Otokiti, S. (2012). An improved stock price prediction using hybrid market indicators. *African Journal of Computing & ICT, 5*(5), 124-135.

Adebiyi, A., Ayo, C., Adebiyi, M. O., & Otokiti, S. (2012). Stock price prediction using neural network with hybridized market indicators. *Journal of Emerging Trends in Computing and Information Sciences, 3*(1), 1-9.

Aggarwal, S., & Krishna, V. (2011). *CS698o: Project Report Stock Price Direction Prediction, Indian Institute of Technology Kanpur*.

Alfaro, E., García, N., Gámez, M., & Elizondo, D. (2008). Bankruptcy forecasting: An empirical comparison of AdaBoost and neural networks. *Decision Support Systems, 45*(1), 110-122.

Alvarez-Ramirez, J., Rodriguez, E., & Espinosa-Paredes, G. (2012). Is the US stock market becoming weakly efficient over time? Evidence from 80-year-long data. *Physica A: Statistical Mechanics and Its Applications, 391*(22), 5643-5647.

Aminghafari, M., Cheze, N., & Poggi, J.-M. (2006). Multivariate denoising using wavelets and principal component analysis. *Computational Statistics & Data Analysis, 50*(9), 2381-2398.

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

Anderson, C. W., Devulapalli, S. V., & Stolz, E. A. (1995). Determining mental state from EEG signals using parallel implementations of neural networks. *Scientific Programming, 4*(3), 171-183.

Antoniadis, A., Bigot, J., & Sapatinas, T. (2001). Wavelet estimators in nonparametric regression: a comparative simulation study. *Journal of Statistical Software, 6*, 1-83.

Arouri, M. E. H., Hammoudeh, S., Lahiani, A., & Nguyen, D. K. (2012). Long memory and structural breaks in modeling the return and volatility dynamics of precious metals. *The Quarterly Review of Economics and Finance, 52*(2), 207-218.

Atsalakis, G. S., & Valavanis, K. P. (2009a). Forecasting stock market short-term trends using a neuro-fuzzy based methodology. *Expert Systems with Applications, 36*(7), 10696-10707.

Atsalakis, G. S., & Valavanis, K. P. (2009b). Surveying stock market forecasting techniques–Part II: Soft computing methods. *Expert Systems with Applications, 36*(3), 5932-5941.

Aydogan, K., & Booth, G. G. (1988). Are there long cycles in common stock returns? *Southern Economic Journal*, 141-149.

Babić, J. B. EUR/RSD Exchange Rate Forecasting Using Hybrid Wavelet-Neural Model: A CASE STUDY. *Computer Science and Information Systems, 12*(2), 487-508.

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.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics, 73*(1), 5-59.

Bajgrowicz, P., & Scaillet, O. (2012). Technical trading revisited: False discoveries, persistence tests, and transaction costs. *Journal of Financial Economics, 106*(3), 473-491.

Bakshi, B. R. (1999). Multiscale analysis and modeling using wavelets. *Journal of Chemometrics, 13*(3-4), 415-434.

Barber, B. M., Lehavy, R., McNichols, M., & Trueman, B. (2006). Buys, holds, and sells: The distribution of investment banks' stock ratings and the implications for the profitability of analysts' recommendations. *Journal of Accounting and Economics, 41*(1), 87-117.

Barkoulas, J. T., & Baum, C. F. (1996). Long-term dependence in stock returns. *Economics Letters, 53*(3), 253-259.

Barkoulas, J. T., Baum, C. F., & Travlos, N. (2000). Long memory in the Greek stock market. *Applied Financial Economics, 10*(2), 177-184.

Basheer, I., & Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods, 43*(1), 3-31.

Batten, J. A., Ellis, C. A., & Hogan, W. P. (2005). Decomposing intraday dependence in currency markets: evidence from the AUD/USD spot market. *Physica A: Statistical Mechanics and Its Applications, 352*(2), 558-572.

Benoit, M., & Wallis, J. (1968). Noah Joseph And Operational Hydrology. *Water Resources Research, 4*(5), 909-924.

Behradmehr, N., & Ahrari, M. (2015). Forecasting Crude Oil Prices: A Hybrid Model Based on Wavelet Transforms and Neural Networks. *The International Journal of Humanities, 21*(3), 131-150.

Benoit, M., & Hudson Richard, L. (2004). The (Mis) behavior of markets. *Physica A: Statistical mechanics and Its applications,* 121-137.

Bentes, S. R., Menezes, R., & Mendes, D. A. (2008). Long memory and volatility clustering: Is the empirical evidence consistent across stock markets? *Physica A: Statistical Mechanics and Its Applications, 387*(15), 3826-3830.

Bessembinder, H., & Chan, K. (1995). The profitability of technical trading rules in the Asian stock markets. *Pacific-Basin Finance Journal, 3*(2), 257-284.

Bezdek, J. C. (1994). What is computational intelligence? *Expert Systems with Applications, 37*(1), 2-11.

Bildirici, M., & Ersin, Ö. Ö. (2009). Improving forecasts of GARCH family models with the artificial neural networks: An application to the daily returns in Istanbul Stock Exchange. *Expert Systems with Applications, 36*(4), 7355-7362.

Boboc, I.-A., & Dinică, M.-C. (2013). An algorithm for testing the efficient market hypothesis. *PloS One, 8*(10), e78234.

Booth, A., Gerding, E., & McGroarty, F. (2014). Automated trading with performance weighted random forests and seasonality. *Expert Systems with Applications, 41*(8), 3651-3661.

Boyacioglu, M. A., & 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.

Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers, 17*(31), 334-355.

Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance, 47*(5), 1731-1764.

Busse, S., Helmholz, P., & Weinmann, M. (2012). Forecasting day ahead spot price movements of natural gas–An analysis of potential influence factors on basis of a NARX neural network. *Multikonferenz Wirtschaftsinformatik 2012, Tagungsband Der MKWI 2012*.

Cajueiro, D. O., & Tabak, B. M. (2004a). Evidence of long range dependence in Asian equity markets: the role of liquidity and market restrictions. *Physica A: Statistical Mechanics and Its Applications, 342*(3), 656-664.

Cajueiro, D. O., & Tabak, B. M. (2004b). The Hurst exponent over time: testing the assertion that emerging markets are becoming more efficient. *Physica A: Statistical Mechanics and Its Applications, 336*(3), 521-537.

Cajueiro, D. O., & Tabak, B. M. (2004c). Ranking efficiency for emerging markets. *Chaos, Solitons & Fractals, 22*(2), 349-352.

Cajueiro, D. O., & Tabak, B. M. (2005). Ranking efficiency for emerging equity markets II. *Chaos, Solitons & Fractals, 23*(2), 671-675.

Campbell, S. D., Davis, M. A., Gallin, J., & Martin, R. F. (2009). What moves housing markets: A variance decomposition of the rent–price ratio. *Journal of Urban Economics, 66*(2), 90-102.

Chandwani, D., & Saluja, M. S. (2014). Stock direction forecasting techniques: an empirical study combining machine learning system with market indicators in the Indian context. *International Journal of Computer Applications, 92*(11).

Chang, J., Jung, Y., Yeon, K., Jun, J., Shin, D., & Kim, H. (1996). Technical indicators and analysis methods. *Seoul: Jinritamgu Publishing*.

Chang, P.-C., & Fan, C.-Y. (2008). A hybrid system integrating a wavelet and TSK fuzzy rules for stock price forecasting. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 38*(6), 802-815.

Chang, P.-C., Wang, Y.-W., & Yang, W.-N. (2004). An investigation of the hybrid forecasting models for stock price variation in Taiwan. *Journal of the Chinese Institute of Industrial Engineers, 21*(4), 358-368.

Charbonneau, L., & Kharma, N. (2009). *Evolutionary inference of rule-based trading agents from real-world stock price histories and their use in forecasting*. Paper presented at the Proceedings of the 11th Annual conference on Genetic and evolutionary computation.

Charles, A., Darné, O., & Kim, J. H. (2012). Exchange-rate return predictability and the adaptive markets hypothesis: Evidence from major foreign exchange rates. *Journal of International Money and Finance, 31*(6), 1607-1626.

Chavan, P. S., & Patil, S. T. (2013). Parameters for Stock Market Prediction. *International Journal of Computer Technology and Applications, 4*(2), 337.

Chavarnakul, T., & Enke, D. (2008). Intelligent technical analysis based equivolume charting for stock trading using neural networks. *Expert Systems with Applications, 34*(2), 1004-1017.

Chen, A.-S., & Leung, M. T. (2004). Regression neural network for error correction in foreign exchange forecasting and trading. *Computers & Operations Research, 31*(7), 1049-1068.

Chen, A.-S., Leung, M. T., & Daouk, H. (2003). Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index. *Computers & Operations Research, 30*(6), 901-923.

Chen, H.-J., Huang, S. Y., & Lin, C.-S. (2009). Alternative diagnosis of corporate bankruptcy: A neuro fuzzy approach. *Expert Systems with Applications, 36*(4), 7710-7720.

Chen, K., Lin, H.-Y., Yu, C., & Chen, Y.-C. (2009). *The prediction of Taiwan government bond yield by neural networks.* Paper presented at the Proceedings of the 13th WSEAS international conference on Systems.

Chen, W.-H., Shih, J.-Y., & Wu, S. (2006). Comparison of support-vector machines and back propagation neural networks in forecasting the six major Asian stock markets. *International Journal of Electronic Finance, 1*(1), 49-67.

Cheng, C.-H., Chen, T.-L., & Wei, L.-Y. (2010). A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting. *Information Sciences, 180*(9), 1610-1629.

Cheng, J.-H., Chen, H.-P., & Lin, Y.-M. (2010). A hybrid forecast marketing timing model based on probabilistic neural network, rough set and C4. *Expert Systems with Applications, 37*(3), 1814-1820.

Chengzhao, Z., Heiping, P., & Ke, Z. (2015). Comparison of Back Propagation Neural Networks and EMD-Based Neural Networks in Forecasting the Three Major Asian Stock Markets. *Journal of Applied Sciences, 15*(1), 90-117.

Cheung, Y.-W., & Lai, K. S. (1995). A search for long memory in international stock market returns. *Journal of International Money and Finance, 14*(4), 597-615.

Cheung, Y. W., Chinn, M. D., & Marsh, I. W. (2004). How do UK ‐ based foreign exchange dealers think their market operates? *International Journal of Finance & Economics, 9*(4), 289-306.

Chiang, T. C., & Doong, S.-C. (2001). Empirical analysis of stock returns and volatility: Evidence from seven Asian stock markets based on TAR-GARCH model. *Review of Quantitative Finance and Accounting, 17*(3), 301-318.

Chiarella, C., He, X.-Z., & Hommes, C. (2006). A dynamic analysis of moving average rules. *Journal of Economic Dynamics and Control, 30*(9), 1729-1753.

Chong, T. T.-L., Cheng, S. H.-S., & Wong, E. N.-Y. (2010). A comparison of stock market efficiency of the BRIC countries. *Technology and Investment, 1*(04), 235-251.

Chong, T. T.-L., & Ng, W.-K. (2008). Technical analysis and the London stock exchange: testing the MACD and RSI rules using the FT30. *Applied Economics Letters, 15*(14), 1111-1114.

Chow, K. V., Denning, K. C., Ferris, S., & Noronha, G. (1995). Long-term and short-term price memory in the stock market. *Economics letters, 49*(3), 287-293.

Conejo, A. J., Contreras, J., Espinola, R., & Plazas, M. A. (2005). Forecasting electricity prices for a day-ahead pool-based electric energy market. *International Journal of Forecasting, 21*(3), 435-462.

Cont, R. (2007). Volatility clustering in financial markets: empirical facts and agent-based models. *Long Memory in Economics* (289-309): Springer.

Coronel-Brizio, H., Hernández-Montoya, A., Huerta-Quintanilla, R., & Rodriguez-Achach, M. (2007). Evidence of increment of efficiency of the Mexican Stock Market through the analysis of its variations. *Physica A: Statistical Mechanics and Its Applications, 380*, 391-398.

Corrado, C. J., & Lee, S. H. (1992). Filter rule tests of the economic significance of serial dependencies in daily stock returns. *Journal of Financial Research, 15*(4), 369-387.

Costa, R. L., & Vasconcelos, G. (2003). Long-range correlations and nonstationarity in the Brazilian stock market. *Physica A: Statistical Mechanics and Its applications, 329*(1), 231-248.

Dacorogna, M., Gencay, R., Muller, U., Olsen, R., & Pictet, O. (2001). An Introduction to High Frequency Finance Academic Press. *San Diego Academic Press*, 383-384.

Dai, W., Wu, J.-Y., & Lu, C.-J. (2012). Combining nonlinear independent component analysis and neural network for the prediction of Asian stock market indexes. *Expert Systems with Applications, 39*(4), 4444-4452.

Damodaran, A. (2016). *Damodaran on valuation: security analysis for investment and corporate finance* (Vol. 324): John Wiley & Sons.

Darbellay, G. A., & Wuertz, D. (2000). The entropy as a tool for analysing statistical dependences in financial time series. *Physica A: Statistical mechanics and Its applications, 287*(3), 429-439.

De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting, 22*(3), 443-473.

Dempster, M., & Jones, C. (2001). A real-time adaptive trading system using genetic programming. *Quantitative Finance, 1*(4), 397-413.

Diaconescu, E. (2008). The use of NARX neural networks to predict chaotic time series. *WSEAS Transactions on Computer Research, 3*(3), 182-191.

Diebold, F. X. (1998). Elements of forecasting. *South-Western College Pub*, 170-176.

Dionisio, A., Menezes, R., & Mendes, D. A. (2006). An econophysics approach to analyse uncertainty in financial markets: an application to the Portuguese stock market. *The European Physical Journal B-Condensed Matter and Complex Systems, 50*(1-2), 161-164.

Donoho, D. L. (1995). De-noising by soft-thresholding. *Information Theory, IEEE Transactions on, 41*(3), 613-627.

Dunis, C. L., & Huang, X. (2002). Forecasting and trading currency volatility: An application of recurrent neural regression and model combination. *Journal of Forecasting, 21*(5), 317-354.

Dunis, C. L., Laws, J., & Sermpinis, G. (2011). Higher order and recurrent neural architectures for trading the EUR/USD exchange rate. *Quantitative Finance, 11*(4), 615-629.

Dunis, C. L., & Shannon, G. (2005). Emerging markets of south-east and central Asia: Do they still offer a diversification benefit? *Journal of Asset Management, 6*(3), 168-190.

Edwards, R. D., & Magee, J. (1997). Technical Analysis of Stock Trends (Amacom, New York).

Eichler, M. (2012). Causal inference in time series analysis. *Causality: Statistical perspectives and applications*, 327-354.

Eigner, P., & Umlauft, T. S. (2015). The Great Depression (s) of 1929-1933 and 2007-2009? Parallels, Differences and Policy Lessons. *Parallels, Differences and Policy Lessons (July 1, 2015). Hungarian Academy of Science MTA-ELTE Crisis History Working Paper* (2).

Enke, D., & Thawornwong, S. (2005). The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with Applications, 29*(4), 927-940.

Errunza, V. R. (1994). Emerging markets: some new concepts. *The Journal of Portfolio Management, 20*(3), 82-87.

Faller, W. E., & Schreck, S. J. (1995). Real-time prediction of unsteady aerodynamics: Application for aircraft control and manoeuvrability enhancement. *IEEE Transactions on Neural Networks, 6*(6), 1461-1468.

Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal, 51*(1), 75-80.

Fama, E. F. (1991). Efficient capital markets: II. *The Journal of Finance, 46*(5), 1575-1617.

Fama, E. F., & Blume, M. E. (1966). Filter rules and stock-market trading. *The Journal of Business, 39*(1), 226-241.

Fatima, S., & Hussain, G. (2008). Statistical models of KSE100 index using hybrid financial systems. *Neurocomputing, 71*(13), 2742-2746.

Fernández-Blanco, P., Bodas-Sagi, D. J., Soltero, F. J., & Hidalgo, J. I. (2008). *Technical market indicators optimization using evolutionary algorithms.* Paper presented at the Proceedings of the 10th annual conference companion on Genetic and evolutionary computation.

Fernandez-Rodrıguez, F., Gonzalez-Martel, C., & 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.

Fong, W. M., & Yong, L. H. (2005). Chasing trends: recursive moving average trading rules and internet stocks. *Journal of Empirical Finance, 12*(1), 43-76.

Frisch, R., & Cassel, G. (1933). Propagation problems and impulse problems in dynamic economics. *The Journal of Portfolio Management,* 76-95.

Garliauskas, A. (1999). *Neural network chaos and computational algorithms of forecast in finance.* Paper presented at the Systems, Man, and Cybernetics, 1999. IEEE International Conference.

Gencay, R. (1996). Non-linear prediction of security returns with moving average rules. *Journal of Forecasting, 15*(3), 165-174.

Gencay, R. (1998a). Optimization of technical trading strategies and the profitability in security markets. *Economics Letters, 59*(2), 249-254.

Gencay, R. (1998b). The predictability of security returns with simple technical trading rules. *Journal of Empirical Finance, 5*(4), 347-359.

Gencay, R. (1999). Linear, non-linear and essential foreign exchange rate prediction with simple technical trading rules. *Journal of International Economics, 47*(1), 91-107.

Geweke, J., & Porter‐Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis, 4*(4), 221-238.

Ghazali, R. (2007). *Higher order neural networks for financial time series prediction.* Liverpool John Moores University.

Ghazani, M. M., & Araghi, M. K. (2014). Evaluation of the adaptive market hypothesis as an evolutionary perspective on market efficiency: Evidence from the Tehran stock exchange. *Research in International Business and Finance, 32*, 50-59.

Ghosh, J., & Shin, Y. (1992). Efficient higher-order neural networks for classification and function approximation. *International Journal of Neural Systems, 3*(04), 323-350.

Gifford, E. (1995). *Investor's guide to technical analysis: predicting price action in the markets*: *Financial Times/Pitman Publishing Series. 45-140.*

Giglio, R., Matsushita, R., Figueiredo, A., Gleria, I., & Da Silva, S. (2008). Algorithmic complexity theory and the relative efficiency of financial markets. *EPL Europhysics Letters, 84*(4), 48005.

Godfrey, L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica: Journal of the Econometric Society*, 1293-1301.

Goonatilake, S., & Khebbal, S. (1995). Intelligent hybrid systems: issues, classifications and future directions. *Intelligent Hybrid Systems, John Wiley & Sons*, 1-20.

Gradojevic, N., & Gençay, R. (2013). Fuzzy logic, trading uncertainty and technical trading. *Journal of Banking & Finance, 37*(2), 578-586.

Granger, C., & Newbold, P. (2001). Spurious regressions in econometrics. *Essays in econometrics: Collected papers of Clive WJ Granger, 2*, 109.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424-438.

Granger, C. W. (2004). Time series analysis, cointegration, and applications. *American Economic Review*, 421-425.

Grau-Carles, P. (2005). Tests of long memory: a bootstrap approach. *Computational Economics, 25*(1-2), 103-113.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review, 70*(3), 393-408.

Gunasekarage, A., & Power, D. M. (2001). The profitability of moving average trading rules in South Asian stock markets. *Emerging Markets Review, 2*(1), 17-33.

Guo, Z., Wang, H., Yang, J., & Miller, D. J. (2015). A Stock Market Forecasting Model Combining Two-Directional Two-Dimensional Principal Component Analysis and Radial Basis Function Neural Network. *PloS One, 10*(4), e0122385.

Handelman, D., Lane, S. H., & Gelfand, J. J. (1990). Integrating neural networks and knowledge-based systems for intelligent robotic control. *IEEE Control Systems Magazine, 10*(3), 77-87.

Harrald, P. G., & Kamstra, M. (1997). Evolving artificial neural networks to combine financial forecasts. *IEEE Transactions on Evolutionary Computation, 1*(1), 40-52.

Harris, L. (2013). What to do about high-frequency trading. *Financial Analysts Journal, 69*(2), 6-9.

Hassan, M. R., Nath, B., & Kirley, M. (2007). A fusion model of HMM, ANN and GA for stock market forecasting. *Expert Systems with Applications, 33*(1), 171-180.

Hassler, U., & Olivares, M. (2008). Long memory and structural change: New evidence from german stock market returns. *Goethe University Frankfurt Discussion Paper*.

Hatanaka, M. (1996). *Time-series-based econometrics: unit roots and co-integrations*: OUP Catalogue Oxford University Press, 67-83.

He, K., Wang, L., Zou, Y., & Lai, K. K. (2014). Exchange Rate Forecasting Using Entropy Optimized Multivariate Wavelet Denoising Model. *Mathematical Problems in Engineering, 2014*.

Henry, O. T. (2002). Long memory in stock returns: some international evidence. *Applied Financial Economics, 12*(10), 725-729.

Hiremath, G. S., & Kumari, J. (2014). Stock returns predictability and the adaptive market hypothesis in emerging markets: evidence from India. *SpringerPlus, 3*(1), 1-17.

Hsieh, T.-J., Hsiao, H.-F., & Yeh, W.-C. (2011). Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm. *Applied Soft Computing, 11*(2), 2510-2525.

Hsu, P.-H., & Kuan, C.-M. (2005). Reexamining the profitability of technical analysis with data snooping checks. *Journal of Financial Econometrics, 3*(4), 606-628.

Hsu, S.-H., Hsieh, J. P.-A., Chih, T.-C., & Hsu, K.-C. (2009). A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression. *Expert Systems with Applications, 36*(4), 7947-7951.

Hu, Y., Feng, B., Zhang, X., Ngai, E., & Liu, M. (2015). Stock trading rule discovery with an evolutionary trend following model. *Expert Systems with Applications, 42*(1), 212-222.

Hua, Z., Wang, Y., Xu, X., Zhang, B., & Liang, L. (2007). Predicting corporate financial distress based on integration of support vector machine and logistic regression. *Expert Systems with Applications, 33*(2), 434-440.

Huang, H., Pasquier, M., & Quek, C. (2009). Financial market trading system with a hierarchical coevolutionary fuzzy predictive model. *Evolutionary Computation, IEEE Transactions on, 13*(1), 56-70.

Huang, W., Lai, K. K., Nakamori, Y., Wang, S., & Yu, L. (2007). Neural networks in finance and economics forecasting. *International Journal of Information Technology & Decision Making, 6*(01), 113-140.

Huang, W., Nakamori, Y., & Wang, S.-Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & Operations Research, 32*(10), 2513-2522.

Hudson, R., Dempsey, M., & 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.

Hull, M., & McGroarty, F. (2014). Do emerging markets become more efficient as they develop? Long memory persistence in equity indices. *Emerging Markets Review, 18*, 45-61.

Hurst, H. E. (1951). {Long-term storage capacity of reservoirs}. *Trans. Amer. Soc. Civil Eng., 116*, 770-808.

Hussain, A. J., Ghazali, R., Al-Jumeily, D., & Merabti, M. (2006, November). Dynamic ridge polynomial neural network for financial time series prediction. *Innovations in Information Technology, 2(1),* 1-5.

Ince, H., & Trafalis, T. B. (2006). A hybrid model for exchange rate prediction. *Decision Support Systems, 42*(2), 1054-1062.

Interactive Brokers (2015). Interactivebrokers.com. Retrieved from https://www.i_nteractivebrokers.com/en/home.php

Ito, M., & Sugiyama, S. (2009). Measuring the degree of time varying market inefficiency. *Economics Letters, 103*(1), 62-64.

Jammazi, R., & Aloui, C. (2012). Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling. *Energy Economics, 34*(3), 828-841.

Jaruszewicz, M., & Mańdziuk, J. (2004). *One day prediction of NIKKEI index considering information from other stock markets.* Paper presented at the International Conference on Artificial Intelligence and Soft Computing.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65-91.

Jensen, M. C. (1978). Some anomalous evidence regarding market efficiency. *Journal of Financial Economics, 6*(2/3), 95-101.

Jensen, M. C., & Benington, G. A. (1970). Random walks and technical theories: Some additional evidence. *The Journal of Finance, 25*(2), 469-482.

Jin, J., & Kim, J. (2015). Forecasting Natural Gas Prices Using Wavelets, Time Series, and Artificial Neural Networks. *PloS One, 10*(11), e0142064.

Johnstone, I. M. (2011). Gaussian estimation: Sequence and wavelet models. *Applied Soft Computing, 5*(2), 439-472.

Jolliffe, I. (2002). *Principal component analysis*: *Intelligent Hybrid Systems, John Wiley & Sons*, 21-32.

Jorion, P. (1995). Predicting volatility in the foreign exchange market. *The Journal of Finance, 50*(2), 507-528.

Kamo, T., & Dagli, C. (2009). Hybrid approach to the Japanese candlestick method for financial forecasting. *Expert Systems with Applications, 36*(3), 5023-5030.

Kanas, A., & Yannopoulos, A. (2001). Comparing linear and nonlinear forecasts for stock returns. *International Review of Economics & Finance, 10*(4), 383-398.

Kao, L.-J., Chiu, C.-C., Lu, C.-J., & Chang, (2013). A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting. *Decision Support Systems, 54*(3), 1228-1244.

Karhunen, J., & Joutsensalo, J. (1995). Generalizations of principal component analysis, optimization problems, and neural networks. *Neural Networks, 8*(4), 549-562.

Karlis, D., Saporta, G., & Spinakis, A. (2003). A simple rule for the selection of principal components. *Communications in Statistics-Theory and Methods, 32*(3), 643-666.

Kestner, L. (2003). *Quantitative trading strategies: harnessing the power of quantitative techniques to create a winning trading program*: McGraw-Hill Professional*, 34-51.*

Kim, H.J., & Shin, K.S. (2007). A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets. *Applied Soft Computing, 7*(2), 569-576.

Kim, J. H., Shamsuddin, A., & Lim, K.-P. (2011). Stock return predictability and the adaptive markets hypothesis: Evidence from century-long US data. *Journal of Empirical Finance, 18*(5), 868-879.

Kim, K., Han, I., & Chandler, J. S. (1998). Extracting trading rules from the multiple classifiers and technical indicators in stock market.

Kim, K.J. (2003). Financial time series forecasting using support vector machines. *Neurocomputing, 55*(1), 307-319.

Kim, K.J. (2006). Artificial neural networks with evolutionary instance selection for financial forecasting. *Expert Systems with Applications, 30*(3), 519-526.

Kim, K.J., & Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems with Applications, 19*(2), 125-132.

Kirkulak Uludag, B., & Lkhamazhapov, Z. (2014). Long memory and structural breaks in the returns and volatility of gold: evidence from Turkey. *Applied Economics, 46*(31), 3777-3787.

Kleinow, H. D.-M. T. (2002). *Testing continuous time models in financial markets.* Humboldt Universität zu Berlin. *11*(2), 56-74.

Kristoufek, L. (2012). How are rescaled range analyses affected by different memory and distributional properties? A Monte Carlo study. *Physica A: Statistical Mechanics and Its Applications, 391*(17), 4252-4260.

Kristoufek, L., & Vosvrda, M. (2013). Measuring capital market efficiency: Global and local correlations structure. *Physica A: Statistical Mechanics and Its Applications, 392*(1), 184-193.

Kristoufek, L., & Vosvrda, M. (2014). Measuring capital market efficiency: long-term memory, fractal dimension and approximate entropy. *The European Physical Journal B, 87*(7), 1-9.

Kuo, R. J., Lee, L. C., & Lee, C. F. (1996, October). Integration of artificial neural networks and fuzzy delphi for stock market forecasting. In *Systems, Man, and Cybernetics, 1996., IEEE International Conference on* (Vol. 2. 1073-1078). IEEE.

Kwon, Y. S., Han, I., & Lee, K. C. (1997). Ordinal pairwise partitioning (OPP) approach to neural networks training in bond rating. *International Journal of Intelligent Systems in Accounting Finance and Management, 6*, 23-40.

Lai, K. K., He, K., & Yen, J. (2007). *Modeling VaR in crude oil market: A multi scale nonlinear ensemble approach incorporating wavelet analysis and ANN.* Paper presented at the International Conference on Computational Science.

Lam, M. (2004). Neural network techniques for financial performance prediction: integrating fundamental and technical analysis. *Decision Support Systems, 37*(4), 567-581.

Lan, B. L., & Tan, Y. O. (2007). Statistical properties of stock market indices of different economies. *Physica A: Statistical Mechanics and Its Applications, 375*(2), 605-611.

LeBaron, B. (1999). Technical trading rule profitability and foreign exchange intervention. *Journal of International Economics, 49*(1), 125-143.

LeBaron, B. (2000). The stability of moving average technical trading rules on the Dow Jones Index. *Derivatives Use, Trading and Regulation, 5*(4), 324-338.

Lee, M.-C. (2009). Using support vector machine with a hybrid feature selection method to the stock trend prediction. *Expert Systems with Applications, 36*(8), 10896-10904.

Lee, S. J., Ahn, J. J., Oh, K. J., & Kim, T. Y. (2010). Using rough set to support investment strategies of real-time trading in futures market. *Applied Intelligence, 32*(3), 364-377.

Leitch, G., & Tanner, J. E. (1991). Economic forecast evaluation: profits versus the conventional error measures. *The American Economic Review*, 580-590.

Lento, C. (2008). A combined signal approach to technical analysis on the S&P 500. *Lakehead University Publication, 28-45*.

Lento, C., Gradojevic, N., & Wright, C. (2007). Investment information content in Bollinger Bands? *Applied Financial Economics Letters, 3*(4), 263-267.

Lertpalangsunti, N. (1999). *An implemented framework for the construction of hybrid intelligent forecasting systems*. Artificial Intelligence in Engineering 13(1), 21-42.

Leung, J. M.-J., & Chong, T. T.-L. (2003). An empirical comparison of moving average envelopes and Bollinger Bands. *Applied Economics Letters, 10*(6), 339-341.

Leung, M. T., Daouk, H., & Chen, A.-S. (2000). Forecasting stock indices: a comparison of classification and level estimation models. *International Journal of Forecasting, 16*(2), 173-190.

Levis, M., & Liodakis, M. (1999). The profitability of style rotation strategies in the United Kingdom. *The Journal of Portfolio Management, 26*(1), 73-86.

Li, K., Sarkar, A., & Wang, Z. (2003). Diversification benefits of emerging markets subject to portfolio constraints. *Journal of Empirical Finance, 10*(1), 57-80.

Li, S.-T., & Kuo, S.-C. (2008). Knowledge discovery in financial investment for forecasting and trading strategy through wavelet-based SOM networks. *Expert Systems with Applications, 34*(2), 935-951.

Lillo, F., & Farmer, J. D. (2004). The long memory of the efficient market. *Studies in Nonlinear Dynamics & Econometrics, 8*(3).

Lim, K.-P. (2007). Ranking market efficiency for stock markets: A nonlinear perspective. *Physica A: Statistical Mechanics and Its Applications, 376*, 445-454.

Lim, K.-P., & Brooks, R. D. (2006). The evolving and relative efficiencies of stock markets: Empirical evidence from rolling bicorrelation test statistics. *Applied Economics Letters,* 24, 10, (668).

Lim, K.-P., Luo, W., & Kim, J. H. (2013). Are US stock index returns predictable? Evidence from automatic autocorrelation-based tests. *Applied Economics, 45*(8), 953-962.

Lim, K. P., & Brooks, R. (2011). The evolution of stock market efficiency over time: a survey of the empirical literature. *Journal of Economic Surveys, 25*(1), 69-108.

Ling, S.-H., Leung, F. H.-F., Lam, H., Lee, Y.-S., & Tam, P. K.-S. (2003). A novel genetic-algorithm-based neural network for short-term load forecasting. *IEEE Transactions on Industrial Electronics, 50*(4), 793-799.

Liu, M. (2000). Modeling long memory in stock market volatility. *Journal of Econometrics, 99*(1), 139-171.

Lo, A. W. (1989). *Long-term memory in stock market prices*. International Journal Of Finance And Economics, 59 – 67.

Lo, A. W. (2004). The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management, 2(3), 48-55.*

Lo, A. W. (2005). Reconciling efficient markets with behavioral finance: the adaptive markets hypothesis. *Journal of Investment Consulting, 7*(2), 21-44.

Lo, A. W., & MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies, 1*(1), 41-66.

Los, C. A. (2003), Financial Market Risk: Measurement & Analysis, Routledge International Studies in Money and Banking, Vol. 24, Taylor & Francis Books Ltd, London, UK.

Lotrič, U. (2004). Wavelet based denoising integrated into multilayered perceptron. *Neurocomputing, 62*, 179-196.

Lotric, U., & Dobnikar, A. (2005). Predicting time series using neural networks with wavelet-based denoising layers. *Neural Computing & Applications, 14*(1), 11-17.

Lu, C.-J. (2010). Integrating independent component analysis-based denoising scheme with neural network for stock price prediction. *Expert Systems with Applications, 37*(10), 7056-7064.

Lukac, L. P., Brorsen, B. W., & Irwin, S. H. (1988). A test of futures market disequilibrium using twelve different technical trading systems. *Applied Economics, 20*(5), 623-639.

Luther, R. K. (1998). An artificial neural network approach to predicting the outcome of Chapter 11 bankruptcy. *The Journal of Business and Economic Studies, 4*(1), 57-68.

Maddala, G. S., & Kim, I.-M. (1998). *Unit roots, cointegration, and structural change*: Cambridge University Press. 2(3), 121-129.

Makridakis, S. (1993). Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting, 9*(4), 527-529.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., . . . Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting, 1*(2), 111-153.

Malkiel, B. G. (2007). *A random walk down Wall Street: the time-tested strategy for successful investing*: WW Norton & Company. 137-156.

Malkiel, B. G., & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383-417.

Manahov, V., & Hudson, R. (2014). A note on the relationship between market efficiency and adaptability–New evidence from artificial stock markets. *Expert Systems with Applications, 41*(16), 7436-7454.

Mandelbrot, B. B. (1971). When can price be arbitraged efficiently? A limit to the validity of the random walk and martingale models. *The Review of Economics and Statistics*, 225-236.

Mandelbrot, B. B. (1975). Limit theorems on the self-normalized range for weakly and strongly dependent processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 31*(4), 271-285.

Mandelbrot, B. B., & Hudson, R. L. (2004). The (mis) behavior of markets: a fractal view of risk, ruin, and reward. *Basic Books Publication*. 10-246.

Mandelbrot, B. B., & Wallis, J. R. (1969a). Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence. *Water Resources Research, 5*(5), 967-988.

Mandelbrot, B. B., & Wallis, J. R. (1969b). Some long - run properties of geophysical records. *Water Resources Research, 5*(2), 321-340.

Marshall, B. R., Cahan, R. H., & Cahan, J. M. (2008). Can commodity futures be profitably traded with quantitative market timing strategies? *Journal of Banking & Finance, 32*(9), 1810-1819.

Matesanz, D., & Ortega, G. J. (2008). A (econophysics) note on volatility in exchange rate time series: entropy as a ranking criterion. *International Journal of Modern Physics C, 19*(07), 1095-1103.

MathWorks. (2014a). Levenberg-Marquadrt Function for use with MATLAB. Retrieved from http://www.mathworks.com/help/nnet/ref/trainlm.html.

MathWorks. (2014b). NARX Toolbox for use with MATLAB. Retrieved from http://www.mathworks.com/help/nnet/ref/narxnet.html.

MathWorks. (2014c). WMULDEN Function for use with MATLAB. Retrieved from http://www.mathworks.com/help/wavelet/ref/wmulden.html.

Matia, K., Pal, M., Salunkay, H., & Stanley, H. (2004). Scale-dependent price fluctuations for the Indian stock market. *Europhysics Letters, 66*(6), 909.

McClelland, J. L., Group, P. d. p. r., & Rumelhart, D. E. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Foundations*: MIT Press. 4(2), 162-175.

McLean, R. D., & Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance, 71*(1), 5-32.

Medsker, L. R. (2012). *Hybrid neural network and expert systems*: Springer Science & Business Media. 1(1), 145-171.

Menkhoff, L. (2010). The use of technical analysis by fund managers: International evidence. *Journal of Banking & Finance, 34*(11), 2573-2586.

Menkhoff, L., & Schmidt, U. (2005). The use of trading strategies by fund managers: some first survey evidence. *Applied Economics, 37*(15), 1719-1730.

Mensi, W., Beljid, M., & Managi, S. (2014). Structural breaks and the time-varying levels of weak-form efficiency in crude oil markets: Evidence from the Hurst exponent and Shannon entropy methods. *International Economics, 140*, 89-106.

Metaxiotis, K., & Psarras, J. (2003). Expert systems in business: applications and future directions for the operations researcher. *Industrial Management & Data Systems, 103*(5), 361-368.

Metghalchi, M., Chang, Y.-H., & Garza-Gomez, X. (2012). Technical analysis of the Taiwanese stock market. *International Journal of Economics and Finance, 4*(1), 90-108.

Milionis, A. E., & Papanagiotou, E. (2011). A test of significance of the predictive power of the moving average trading rule of technical analysis based on sensitivity analysis: application to the NYSE, the Athens Stock Exchange and the Vienna

Stock Exchange. Implications for weak-form market efficiency testing. *Applied Financial Economics, 21*(6), 421-436.

Mizuno, H., Kosaka, M., Yajima, H., & Komoda, N. (1998). Application of neural network to technical analysis of stock market prediction. *Studies in Informatic and Control, 7*(3), 111-120.

Moazzami, M., Khodabakhshian, A., & Hooshmand, R. (2013). A new hybrid day-ahead peak load forecasting method for Iran's National Grid. *Applied Energy, 101*, 489-501.

Motiwalla, L., & Wahab, M. (2000). Predictable variation and profitable trading of US equities: a trading simulation using neural networks. *Computers & Operations Research, 27*(11), 1111-1129.

Mukherjee, I., Sen, C., & Sarkar, A. (2011). Long memory in stock returns: Insights from the Indian market. *The International Journal of Applied Economics and Finance, 5*(1), 62-74.

Mundra, P., & Rajapakse, J. (2007). SVM-RFE with relevancy and redundancy criteria for gene selection. *Pattern Recognition in Bioinformatics*, 242-252.

Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*: Penguin Publication. 32-56

Necula, C. (2009). Modeling Heavy Tailed Stock Index Returns Using the Generalized Hyperbolic Distribution. *Romanian Journal of Economic Forecasting, 10*(2), 118-131.

Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management Science, 60*(7), 1772-1791.

Neely, C. J., & Weller, P. A. (2011). Technical analysis in the foreign exchange market. *Federal Reserve Bank of St. Louis Review*, (Sep), 23-38.

Nelson, C. R., & Plosser, C. R. (1982). Trends and random walks in macroeconmic time series: some evidence and implications. *Journal of Monetary Economics, 10*(2), 139-162.

Nemes, M. D., & Butoi, A. (2013). Data mining on romanian stock market using neural networks for price prediction. *Informatica Economica, 17*(3), 125-141.

Ni, H., & Yin, H. (2009). Exchange rate prediction using hybrid neural networks and trading indicators. *Neurocomputing, 72*(13), 2815-2823.

Nikolaos, L. (2008). Radial basis functions networks to hybrid neuro-genetic RBFNs in financial evaluation of corporations. *International Journal of Computers, 2*(2), 176-183.

Noda, A. (2012). A Test of the Adaptive Market Hypothesis using a Time-Varying AR Model in Japan. *Journal of Forecasting, 2*(1), 43-59.

O'Neill, M., Brabazon, A., Ryan, C., & Collins, J. (2001). *Evolving market index trading rules using grammatical evolution.* Paper presented at the Workshops on Applications of Evolutionary Computation.

Oh, G., Kim, S., & Eom, C. (2007). Market efficiency in foreign exchange markets. *Physica A: Statistical Mechanics and Its Applications, 382*(1), 209-212.

Oh, G., Um, C.-J., & Kim, S. (2006). Statistical properties of the returns of stock prices of international markets. *Journal of Forecasting, 1*(3), 121-137.

Ohanissian, A., Russell, J. R., & Tsay, R. S. (2008). True or spurious long memory? A new test. *Journal of Business & Economic Statistics, 26*(2), 161-175.

Olson, D. (2004). Have trading rule profits in the currency markets declined over time? *Journal of Banking & Finance, 28*(1), 85-105.

Olson, D., & Mossman, C. (2003). Neural network forecasts of Canadian stock returns using accounting ratios. *International Journal of Forecasting, 19*(3), 453-465.

Ortega, L. F. (2012). A neuro-wavelet method for the forecasting of financial time series. *Proceedings of the World Congress on Engineering and Computer Science* 1(1), 24-26.

Pai, P.-F., & Lin, C.-S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega, 33*(6), 497-505.

Pan, H., Tilakaratne, C., & Yearwood, J. (2005). Predicting Australian stock market index using neural networks exploiting dynamical swings and intermarket influences. *Journal Of Research And Practice In Information Technology, 37*(1), 43-56.

Panas, E. (2001). Estimating fractal dimension using stable distributions and exploring long memory through ARFIMA models in Athens Stock Exchange. *Applied Financial Economics, 11*(4), 395-402.

Panda, C., & Narasimhan, V. (2007). Forecasting exchange rate better with artificial neural network. *Journal of Policy Modeling, 29*(2), 227-236.

Pang, Y., Xu, W., Yu, L., Ma, J., Lai, K. K., Wang, S., & Xu, S. (2011). Forecasting the crude oil spot price by wavelet neural networks using OECD petroleum inventory levels. *New Mathematics and Natural Computation, 7*(02), 281-297.

Park, C. H., & Irwin, S. H. (2007). What do we know about the profitability of technical analysis? *Journal of Economic Surveys, 21*(4), 786-826.

Pérez-Rodríguez, J. V., Torra, S., & Andrada-Félix, J. (2005). STAR and ANN models: forecasting performance on the Spanish Ibex-35 stock index. *Journal of Empirical Finance, 12*(3), 490-509.

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics, 16*(3), 289-326.

Pesaran, M. H., & Timmermann, A. (1994). Forecasting stock returns An examination of stock market trading in the presence of transaction costs. *Journal of Forecasting, 13*(4), 335-367.

Peters, E. E. (1994). *Fractal market analysis: applying chaos theory to investment and economics* (Vol. 24): John Wiley & Sons.

Phillips, P. C. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics, 33*(3), 311-340.

Pincus, S., & Kalman, R. E. (2004). Irregularity, volatility, risk, and financial market time series. *Proceedings of the National Academy of Sciences of the United States of America, 101*(38), 13709-13714.

Pincus, S. M. (1991). *Approximate entropy as a measure of system complexity.* Paper presented at the National Academy of Sciences.

Pindoriya, N., Singh, S., & Singh, S. (2008). An adaptive wavelet neural network-based energy price forecasting in electricity markets. *Power Systems, IEEE Transactions on, 23*(3), 1423-1432.

Popović, S., Mugoša, A., & Đurović, A. (2013). Adaptive Markets Hypothesis: Empirical Evidence from Montenegro Equity Market. *Economic Research-Ekonomska Istraživanja, 26*(3), 31-46.

Poterba, J. M., & Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics, 22*(1), 27-59.

Prechelt, L. (2012). Early stopping—but when? *Neural Networks: Tricks of the Trade*, 2(1), 53-67.

Preen, R. (2009). An XCS approach to forecasting financial time series. *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference: Late Breaking Papers*, 2625-2632.

Pukthuanthong, K., Levich, R. M., & Thomas, L. R. (2006). Do foreign exchange markets still trend? *International Journal of Forecasting, 17*(1), 371-384.

Quah, T.-S., & Srinivasan, B. (1999). Improving returns on stock investment through neural network selection. *Expert Systems with Applications, 17*(4), 295-301.

Ramsey, J. B., & Zhang, Z. (1997). The analysis of foreign exchange data using waveform dictionaries. *Journal of Empirical Finance, 4*(4), 341-372.

Rasheed, K., & Qian, B. (2004). *Hurst exponent and financial market predictability.* Paper presented at the IASTED conference on Financial Engineering and Applications (FEA 2004).

Refenes, A. N., Zapranis, A., & Francis, G. (1994). Stock performance modeling using neural networks: a comparative study with regression models. *Neural Networks, 7*(2), 375-388.

Rege, S., & Martín, S. G. (2011). Portuguese Stock Market: A Long-Memory Process? *Business: Theory and Practice/Verslas: Teorija ir Praktika, 12*(1), 75-84.

Regnault, J. (1863). *Calcul des chances et philosophie de la bourse*: Mallet-Bachelier.

Risso, W. A. (2008). The informational efficiency and the financial crashes. *Research in International Business and Finance, 22*(3), 396-408.

Rodríguez-González, A., García-Crespo, Á., Colomo-Palacios, R., Iglesias, F. G., & Gómez-Berbís, J. M. (2011). CAST: Using neural networks to improve trading systems based on technical analysis by means of the RSI financial indicator. *Expert Systems with Applications, 38*(9), 11489-11500.

Roman, J., & Jameel, A. (1996, January). Backpropagation and recurrent neural networks in financial analysis of multiple stock market returns. *Proceedings of the Twenty-Ninth Hawaii International Conference,* 2, 454-460.

Rosillo, R., De la Fuente, D., & Brugos, J. A. L. (2013). Technical analysis and the Spanish stock exchange: testing the RSI, MACD, momentum and stochastic rules using Spanish market companies. *Applied Economics, 45*(12), 1541-1550.

Rosso, O., Larrondo, H., Martin, M., Plastino, A., & Fuentes, M. (2007). Distinguishing noise from chaos. *Physical Review Letters, 99*(15), 154102.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American Statistical Association, 79*(388), 871-880.

Sadique, S., & Silvapulle, P. (2001). Long ‐ term memory in stock market returns: international evidence. *International Journal of Finance & Economics, 6*(1), 59-67.

sargan, J. D. (1964). Wages and prices in the United Kingdom: A study in econometric methodology. *Econometric Analysis for National Economic Planning*, 25-63.

Sargan, J. D., & Bhargava, A. (1983). Testing residuals from least squares regression for being generated by the Gaussian random walk. *Econometrica: Journal of the Econometric Society*, 153-174.

Sermpinis, G., Dunis, C., Laws, J., & Stasinakis, C. (2012). Forecasting and trading the EUR/USD exchange rate with stochastic Neural Network combination and time-varying leverage. *Decision Support Systems, 54*(1), 316-329.

Sermpinis, G., Theofilatos, K., Karathanasopoulos, A., Georgopoulos, E. F., & Dunis, C. (2013). Forecasting foreign exchange rates with adaptive neural networks using radial-basis functions and particle swarm optimization. *European Journal of Operational Research, 225*(3), 528-540.

Seth, A. (2007). Granger causality. *Scholarpedia, 2*(7), 1667-1684.

Shafie-Khah, M., Moghaddam, M. P., & Sheikh-El-Eslami, M. (2011). Price forecasting of day-ahead electricity markets using a hybrid forecast method. *Energy Conversion and Management, 52*(5), 2165-2169.

Shah, J. R., & Murtaza, M. B. (2000). A neural network based clustering procedure for bankruptcy prediction. *American Business Review, 18*(2), 80-99.

Shahbazi, N., Memarzadeh, M., & Gryz, J. (2016). *Forex Market Prediction Using NARX Neural Network with Bagging.* Paper presented at the MATEC Web of Conferences.

Shen, L., & Loh, H. T. (2004). Applying rough sets to market timing decisions. *Decision Support Systems, 37*(4), 583-597.

Shen, S., Jiang, H., & Zhang, T. (2012). Stock market forecasting using machine learning algorithms. *Journal of Research And Practice In Information Technology, 28* (3), 59-72.

Shiller, R. J. (1980). Do stock prices move too much to be justified by subsequent changes in dividends? *National Bureau of Economic Research Cambridge, Mass. 12(2), 47-69.*

Shiller, R. J., Fischer, S., & Friedman, B. M. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity, 1984*(2), 457-510.

Shin, Y., & Ghosh, J. (1991). The pi-sigma network: An efficient higher-order neural network for pattern classification and function approximation. *Proceedings of the Neural Networks, Seattle International Joint Conference*, 13-18.

Siegelmann, H. T., Horne, B. G., & Giles, C. L. (1997). Computational capabilities of recurrent NARX neural networks. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 27*(2), 208-215.

Smith, G. (2012). The changing and relative efficiency of European emerging stock markets. *The European Journal of Finance, 18*(8), 689-708.

Soni, S. (2011). Applications of ANNs in stock market prediction: a survey. *International Journal of Computer Science & Engineering Technology, 2*(3), 71-83.

Sridharan, S. A. (2015). Volatility forecasting using financial statement information. *The Accounting Review, 90*(5), 2079-2106.

Szafarz, A. (2012). Financial crises in efficient markets: How fundamentalists fuel volatility. *Journal of Banking & Finance, 36*(1), 105-111.

Taffese, W. Z. (2007). Case-based reasoning and neural networks for real estate valuation. *Artificial Intelligence and Applications*, 14(2), 98-104.

Taha, I., & Ghosh, J. (1997). Hybrid intelligent architecture and its application to water reservoir control. *International Journal Of Smart Engineering System Design, 1*(1), 59-75.

Taha, I. A. E. (1998). A hybrid intelligent architecture and revising domain knowledge. *International Journal of Smart Engineering Systems, vol. 1, pp. 59-75.*

Taylor, S. (1986). Modelling Financial Time Series, John Wiley & Sons. *Great Britain*.

Temin, P. (2010). The Great Recession & the Great Depression. *Daedalus, 139*(4), 115-124.

Tenti, P. (1996). Forecasting foreign exchange rates using recurrent neural networks. *Applied Artificial Intelligence, 10*(6), 567-582.

Teoh, H. J., Chen, T.-L., Cheng, C.-H., & Chu, H.-H. (2009). A hybrid multi-order fuzzy time series for forecasting stock markets. *Expert Systems with Applications, 36*(4), 7888-7897.

Teräsvirta, T., Van Dijk, D., & Medeiros, M. C. (2005). Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting, 21*(4), 755-774.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science, 4*(3), 199-214.

Thawornwong, S., & Enke, D. (2004). The adaptive selection of financial and economic variables for use with artificial neural networks. *Neurocomputing, 56*, 205-232.

Todea, A., Ulici, M., & Silaghi, S. (2009). Adaptive markets hypothesis: Evidence from Asia-Pacific financial markets. *The Review of Finance and Banking, 1*(1).

Tolvi, J. (2003). Long memory and outliers in stock market returns. *Applied Financial Economics, 13*(7), 495-502.

Touzet, C. F. (1997). Neural reinforcement learning for behaviour synthesis. *Robotics and Autonomous Systems, 22*(3), 251-281.

Tsai, C.-F., & Chiou, Y.-J. (2009). Earnings management prediction: A pilot study of combining neural networks and decision trees. *Expert Systems with Applications, 36*(3), 7183-7191.

Tsai, C.-F., & Hsiao, Y.-C. (2010). Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. *Decision Support Systems, 50*(1), 258-269.

Tsai, C.-F., Lin, Y.-C., Yen, D. C., & Chen, Y.-M. (2011). Predicting stock returns by classifier ensembles. *Applied Soft Computing, 11*(2), 2452-2459.

Tsaih, R., Hsu, Y., & Lai, C. C. (1998). Forecasting S&P 500 stock index futures with a hybrid AI system. *Decision Support Systems, 23*(2), 161-174.

Tsakonas, A., Dounias, G., Doumpos, M., & Zopounidis, C. (2006). Bankruptcy prediction with neural logic networks by means of grammar-guided genetic programming. *Expert Systems with Applications, 30*(3), 449-461.

Urquhart, A., & Hudson, R. (2013). Efficient or adaptive markets? Evidence from major stock markets using very long run historic data. *International Review of Financial Analysis, 28*, 130-142.

Urquhart, A., & McGroarty, F. (2014). Calendar effects, market conditions and the Adaptive Market Hypothesis: Evidence from long-run US data. *International Review of Financial Analysis, 35*, 154-166.

Vanstone, B., & Finnie, G. (2009). An empirical methodology for developing stockmarket trading systems using artificial neural networks. *Expert Systems with Applications, 36*(3), 6668-6680.

Varutbangkul, E. (2013). Integrating Traditional Stock Trading Systems: A Data Mining Approach. *Expert Systems with Applications, 42*(2), 404-417.

Wang, J.-Z., Wang, J.-J., Zhang, Z.-G., & Guo, S.-P. (2011). Forecasting stock indices with back propagation neural network. *Expert Systems with Applications, 38*(11), 14346-14355.

Wang, L., & Fu, X. (2006). *Data mining with computational intelligence*: Springer Science & Business Media. 22(3), 142-156.

Wang, L., & Gupta, S. (2013). Neural networks and wavelet de-noising for stock trading and prediction. *Time Series Analysis, Modeling and Applications*. 229-247

Wang, L., Lin, J., & Metzler, D. (2010). Learning to efficiently rank. *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, 138-145.

Wang, Y.-H. (2009). Nonlinear neural network forecasting model for stock index option price: Hybrid GJR–GARCH approach. *Expert Systems with Applications, 36*(1), 564-570.

Wen, Q., Yang, Z., & Song, Y. (2009). *Hybrid approaches for stock price prediction. Computers & Operations Research.* 38(3), 708-722.

Weron, R. (2002). Estimating long-range dependence: finite sample properties and confidence intervals. *Physica A: Statistical Mechanics and Its Applications, 312*(1), 285-299.

Weron, R. (2011/Sep). Hurst: MATLAB function to compute the Hurst exponent using R/S Analysis [COMP]: Hugo Steinhaus Center, Wroclaw University of Technology. Retrieved from https://ideas.repec.org/c/wuu/hscode/m11003.html

White, H. (1988). Economic prediction using neural networks: The case of IBM daily stock returns. *Proceedings of the Second Annual IEEE Conference on Neural Networks*, 451-458

Williams, C. J. (2012). Euro crisis imperils recovering global economy, OECD warns. *Retrieved From Los Angeles Times*. 22

Willinger, W., Taqqu, M. S., & Teverovsky, V. (1999). Stock market prices and long-range dependence. *Finance and Stochastics, 3*(1), 1-13.

Wu, X., Fung, M., & Flitman, A. (2001). *Forecasting stock market performance using hybrid intelligent system.* Paper presented at the International Conference on Computational Science, Springer, Berlin, Heidelberg.

Xiao, Y., Xiao, J., Lu, F., & Wang, S. (2014). Ensemble ANNs-PSO-GA approach for day-ahead stock e-exchange prices forecasting. *International Journal of Computational Intelligence Systems, 7*(2), 272-290.

Xiong, T., Bao, Y., & Hu, Z. (2014). Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting. *Knowledge-Based Systems, 55*, 87-100.

Yalama, A., & Celik, S. (2013). Real or spurious long memory characteristics of volatility: Empirical evidence from an emerging market. *Economic Modelling, 30*, 67-72.

Yao, J., Tan, C. L., & Poh, H.-L. (1999). Neural networks for technical analysis: a study on KLCI. *International Journal of Theoretical and Applied Finance, 2*(2), 221-241.

Ye, C., & Huang, J. (2008). Non-classical oscillator model for persistent fluctuations in stock markets. *Physica A: Statistical Mechanics and Its Applications, 387*(5), 1255-1263.

Yoon, Y., Swales Jr, G., & Margavio, T. M. (1993). A comparison of discriminant analysis versus artificial neural networks. *Journal of the Operational Research Society*, 51-60.

Yu, L., Wang, S., & Lai, K. K. (2005). A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates. *Computers & Operations Research, 32*(10), 2523-2541.

Yu, T. H.-K., & Huarng, K.-H. (2008). A bivariate fuzzy time series model to forecast the TAIEX. *Expert Systems with Applications, 34*(4), 2945-2952.

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between Time-Series? a study in sampling and the nature of time-series. *Journal of The Royal Statistical Society, 89*(1), 1-63.

Yümlü, S., Gürgen, F. S., & Okay, N. (2005). A comparison of global, recurrent and smoothed-piecewise neural models for Istanbul stock exchange (ISE) prediction. *Pattern Recognition Letters, 26*(13), 2093-2103.

Zadeh, L. A. (1994). The role of fuzzy logic in modeling, identification and control. *Modeling Identification and Control, 15*(3), 191-212.

Zhang, B.-L., Coggins, R., Jabri, M. A., Dersch, D., & Flower, B. (2001). Multiresolution forecasting for futures trading using wavelet decompositions. *IEEE Transactions on Neural Networks, 12*(4), 765-775.

Zhang, B.-L., & Dong, Z.-Y. (2001). An adaptive neural-wavelet model for short term load forecasting. *Electric Power Systems Research, 59*(2), 121-129.

Zhang, X., Hu, Y., Xie, K., Wang, S., Ngai, E., & Liu, M. (2014). A causal feature selection algorithm for stock prediction modeling. *Neurocomputing, 142*, 48-59.

Zhang, Y., & Wu, L. (2009). Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. *Expert Systems with Applications, 36*(5), 8849-8854.

Zhang, Y.-C. (1999). Toward a theory of marginally efficient markets. *Physica A: Statistical Mechanics and Its Applications, 269*(1), 30-44.

Zhu, Y., & Zhou, G. (2009). Technical analysis: An asset allocation perspective on the use of moving averages. *Journal of Financial Economics, 92*(3), 519-544.

Zhuang, Y., Green, C. J., & Maggioni, P. (2000). The great rebound, the great crash, and persistence in British stock prices. *Economic Research Paper, Loughborough University*.

Zovko, I. I. (2008). *Topics in market microstructure*: Amsterdam University Press. 23(2), 68-80.

Zunino, L., Zanin, M., Tabak, B. M., Pérez, D. G., & Rosso, O. A. (2010). Complexity-entropy causality plane: A useful approach to quantify the stock market inefficiency. *Physica A: Statistical Mechanics and Its Applications, 389*(9), 1891-1901.

Zurada, J. M. (1992). *Introduction to artificial neural systems*. *IEEE Transactions on Evolutionary Computation, 2*(1), 61-75.