

# HYBRID APPROACH FOR FINANCIAL FORECASTING WITH SUPPORT VECTOR MACHINES

Withanage Don Sameera Roshan

(118006X)



University of Moratuwa, Sri Lanka.  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)  
Degree of Master of Science

Department of Mechanical Engineering

University of Moratuwa  
Sri Lanka

June 2014

# **HYBRID APPROACH FOR FINANCIAL FORECASTING WITH SUPPORT VECTOR MACHINES**

Withanage Don Sameera Roshan

(118006X)



University of Moratuwa, Sri Lanka.  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)

Thesis submitted in partial fulfillment of the requirements for  
the degree Master of Science

Department of Mechanical Engineering

University of Moratuwa  
Sri Lanka

June 2014

## Declaration

I declare that this is my own work and this thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis/dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Mr. W.D.S. Roshan



Signature:

University of Moratuwa, Sri Lanka.  
Date:  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)

The above candidate has carried out research for the Master's thesis under our supervision.

Dr. R.A.R.C Gopura  
Department of Mechanical Engineering,  
University of Moratuwa

Signature of the supervisor:

Date:

Dr. A.G.B.P Jayasekara  
Department of Electrical Engineering,  
University of Moratuwa

Signature of the supervisor:

Date:

*Dedicated*  
To my parents




University of Moratuwa, Sri Lanka.  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)

## **Acknowledgements**

I would like to thank my supervisors Dr. A.G.B.P Jayasekara and Dr. R.A.R.C. Gopura for their continuous support given to me throughout the master's program. Thank you for your immense support for thesis writing guidelines, reference materials, motivation and vision provided by Dr. A.G.B.P Jayasekara. It has been a great pleasure to explore the new knowledge with the friendly environment created by Dr. R.A.R.C. Gopura. I really appreciate the trust and confidence you gave me throughout the whole period of time. Thank you both for the great support given.

I would like to thank the members of the progress review committee, Dr. K.T.M.U. Hemapala and Dr.V.P.C. Dassanayake for giving their ideas, directions, positives and cons by overseeing the progress of the work.

Withanage Don Sameera Roshan  
sameeraroshan@gmail.com  
 University of Moratuwa, Sri Lanka.  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)

## Abstract

Financial markets are the biggest business platforms in the world. Therefore, financial forecasting is getting a lot of attention in today's economic context. Accurate forecast is beneficial to broker firms, governments, individuals etc.

Vast range of forecasting methods, models have introduced by the research community. However, the risk involved with trading on those markets are very high. Such complexity makes a difficulty of making consistent profit. Building an accurate forecasting model is still an active and interesting research area for the academic community.

Recently, nonlinear statistical models such as neural network, support vector machine have shown greater capability to forecast financial markets over conventional methods. This dissertation proposed a hybrid support vector machine model which consists of wavelet transform and k-means clustering for foreign exchange market forecasting. The proposed model analyzes the trends and makes a forecast by entirely depending on the past exchange data. Wavelet transform is used to remove the noise of the time series. K-means clustering cluster the input space according to the similarities of the input vectors and finally support vector models make a forecast for the relevant cluster.

The proposed hybrid forecasting system was tested on real market environment to check the forecasting capability. Auto trading algorithm developed on 'metatrader4' platform used the forecast of the model to trade on the real conditions. Results confirmed that the proposed model can forecast price movements with greater accuracy that leads to profitable trades on foreign exchange market.



University of Moratuwa, Sri Lanka.  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)

## Table of content

Declaration .....	i
Dedication .....	ii
Acknowledgements .....	iii
Abstract .....	iv
Table of content .....	v
List of figures .....	vii
List of tables .....	x
List of abbreviations.....	xi
1. INTRODUCTION.....	1
1.1 Financial Markets .....	1
1.2 Prediction Approaches.....	2
1.3 Motivation .....	3
1.4 Overview .....	4
2. CURRENT APPROACHES FOR FINANCIAL FORECASTING .....	5
2.1 Data Preparation .....	7
2.1.1 Input selection .....	7
2.1.2 Input data preprocessing.....	10
2.2 Model Design .....	11
2.2.1 Learning algorithm .....	11
2.2.2 Architectures.....	12
2.2.3 Performance measure .....	12
3. DEVELOPMENT OF FORECASTING MODEL .....	13
3.1 Removal of Noise and Non-Stationary Characteristics of Input Data .....	13
3.1.1 Wavelet based denoising issue .....	25
3.2 Learning Algorithm Selection .....	34
3.3 Cluster Input Space.....	39
3.4 Proposed Hybrid Model .....	41

4. REAL TIME IMPLEMENTATION AND ISSUES.....	43
4.1 Wavelet Based Denoising Issue and Errors at the Edge.....	43
4.2 Forecasting Methodology and Performance Measure.....	47
5. MODEL TESTING.....	50
5.1 Meta Trader Expert Advisor.....	50
5.2 Meta Trader Back Testing.....	52
6. SIMULATIONS AND RESULTS.....	53
6.1 Forecast for EUR/USD 1 Hr.....	53
6.2 Forecast for EUR/USD 5min.....	54
6.3 Test Results for Real Market Conditions .....	55
6.4 Results for EUR/USD back testing .....	58
6.5 Results for USD/JPY back testing.....	61
7. CONCLUSION AND FUTURE DIRECTIONS.....	65
References.....	67
Appendix A.....	72
Appendix B.....	75



University of Moratuwa, Sri Lanka.  
 Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)



## List of figures

Figure 3.1- Debauchees 1 wavelet .....	17
Figure 3.2- Debauchees 2 wavelet .....	17
Figure 3.3-Debauchees 6 wavelet .....	18
Figure 3.4- Debauchees 10 wavelet .....	18
Figure 3.5- Coiflet 1 wavelet .....	19
Figure 3.6- Coiflet 2 wavelet .....	19
Figure 3.7-Coiflet 5 wavelet .....	20
Figure 3.8 - Debauchee's wavelet transform for EUR/USD. Db1 and Db2 wavelets combined with SWT and DWT denoising is shown in the figure. Soft and had thresholding have been used for both SWT and DWT for comparison.....	21
Figure 3.9 - Debauchee's wavelet transform for EUR/USD. Db6 and Db10 wavelets combined with SWT and DWT denoising is shown in the figure. Soft and had thresholding have been used for both SWT and DWT for comparison.....	22
Figure 3.10 - Coiflet wavelet transform for EUR/USD. Coiflet 1 and coiflet 2 wavelets combined with SWT and DWT denoising is shown in the figure. Soft and had thresholding have been used for both SWT and DWT for comparison.....	23
Figure 3.11 - Coiflet wavelet transform for EUR/USD. Coiflet 5 wavelets combined with SWT and DWT denoising is shown in the figure. Soft and had thresholding have been used for both SWT and DWT for comparison.....	24
Figure 3.12- Comparison of the accuracy of Haar wavelet in SWT denoising. Figure 3.12 (A) shows the denoised results of two timeframes superimposed. Figure 3.12 (B) zoomed into selected area of Figure 3.12 (A). Yellow and Green lines of Figure 3.12 (B) shows the wrong trend direction generated from the Haar based SWT soft and hard thresholding.....	27
Figure 3.13 - Comparison of the accuracy of Haar wavelet in DWT denoising. Figure 3.13 (A) shows the denoised results of two timeframes superimposed. Figure 3.13 (B) zoomed into selected area of Figure 3.13 (A). All the lines of Figure 3.13 (B) super imposed with one another indicate that the noise removal at the edge for Haar based denoising is accurate.....	28
Figure 3.14 Comparison of the accuracy of Coiflet 3 in SWT denoising Figure 3.14 (A) shows the denoised results of two timeframes superimposed. Figure 3.14 (B) zoomed into selected area of Figure 3.14 (A). Yellow and Green lines of Figure 3.14 (B) shows the wrong trend direction generated from the Coiflet 3 based SWT soft and hard thresholding .....	29

Figure 3.15- Comparison of the accuracy of Coiflet3 in DWT denoising. Figure 3.15 (A) shows the denoised results of two timeframes superimposed. Figure 3.15 (B) zoomed into selected area of Figure 3.15 (A). Yellow and Green lines of Figure 3.15 (B) shows the correct trend direction generated from the Coiflet3 based SWT soft and hard thresholding. Accuracy is less than Haar based DWT denoising ..... 30

Figure 3.16- Comparison of the accuracy of Db6 wavelet in SWT denoising. Figure 3.16 (A) shows the denoised results of two timeframes superimposed. Figure 3.16 (B) zoomed into selected area of Figure 3.16 (A). Yellow and Green lines of Figure 3.16 (B) shows the wrong trend direction generated from the Db6 based SWT soft and hard thresholding ..... 31

Figure 3.17- Comparison of the accuracy of Db6 wavelet in DWT denoising. Figure 3.17 (A) shows the denoised results of two timeframes superimposed. Figure 3.17 (B) zoomed into selected area of Figure 3.17 (A). Yellow and Green lines of Figure 3.17 (B) shows the correct trend direction generated from the Db6 based SWT soft and hard thresholding which is more accurate..... 32

Figure 3.18- Optimal hyper-plane for set of data points ..... 36

Figure 3.19 - Proposed hybrid model combining wavelet transform, K-means clustering and support vector machine ..... 42

Figure 4.1- EUR/USD closing price denoised with discrete wavelet soft thresholding .. 44

Figure 4.2- Discrete wavelet denoising with soft thresholding is done for EUR/USD time series. Original time series is denoised with 128 data and 256 data. Denoising is accurate for past data. However, denoised data is not always correct for current time frame. .... 45

Figure 4.3- Zoomed into current time frame of Figure 4.2 ..... 46

Figure 4.4 - Low mean squared error (MSE) and high directional accuracy (D%) is lead to profit. Higher mean squared error with lower directional accuracy is also leaded to profit..... 47

Figure 4.5 – EMA (30) and ideal zig-zag forecast on the same graph..... 48

Figure 5.1: Flow diagram of the overall process..... 51

Figure 6.1- Model forecast for 400 one-hour data points ..... 53

Figure 6.2- Steady forecast for uptrend with noisy environment..... 54

Figure 6.3- Steady forecast for down trend with noisy environment..... 54

Figure 6.4- Model forecast for 1000 five minute data points. .... 55

Figure 6.5- Zoomed area for evaluate model robustness under noisy environment. .... 55

Figure 6.6 - Profit chart for the period of 2009 January to 2010 January..... 58

Figure 6.7 - Back testing results for the period of 2010 January to 2011 January..... 59

Figure 6.8 - Profit for the period of 2011 January to 2012 January ..... 60

Figure 6.9- Profit for the period of 2009 January to 2010 January.....61  
Figure 6.10 - Profit for the period of 2010 January to 2011 January.....62  
Figure 6.11- Profit for the period of 2011 January to 2012 January.....63



University of Moratuwa, Sri Lanka.  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)

## List of tables

Table 2.1-Winners of the automated trading championship (2007-2011).....	6
Table 2.2 - Types of inputs used for the forecasting model.....	9
Table 3.1- Error comparison with level 2 decomposition.....	26
Table 3.2- Debauchee's denoising for DWT with soft thresholding( MSE- Mean Squared error, D% - Directional Statistics).....	33
Table 3.3- Error with levels of triple exponential moving average .....	35
Table 4.1 - The comparison of two graphs on Figure 4.4.....	47
Table 4.2- Comparison of two graphs on Figure 4.5 .....	48
Table 6.1- Back testing results for the period of 2009 January to 2010 January .....	58
Table 6.2- Back testing results for the period of 2010 January to 2011 January .....	59
Table 6.3 - Back testing results for the period of 2011 January to 2012 January.....	60
Table 6.4 - Back testing results for the period of 2009 January to 2010 January.....	61
Table 6.5- Back testing results for the period of 2010 January to 2011 January .....	62
Table 6.6- Back testing results for the period of 2011 January to 2012 January .....	63
Table 6.7 – Summary of the results.....	64
Table A.1 – Summary of the reviewed articles.....	72
Table B.1-First 100 orders of EUR/USD trading for year 2009.....	75

## List of abbreviations

ABC - Artificial Bee Colony Algorithm  
APE - Absolute Percentage Error  
ARIMA - Auto Regression Integrated Moving Average  
BP - Back Propagation  
CD - Correct Down Trend  
DS - Directional Symmetry  
DWT - Discrete Wavelet Transform  
EA - Expert Advisor  
EKF - Extended Kalman Filter  
EMA - Exponential Moving Average  
EUR - Euro  
FFT - Fast Fourier Transform  
FLNN - Functional Link Neural Network  
FOREX – Foreign Exchange  
GA - Genetic Algorithm  
GARCH - Generalized Autoregressive Conditional Heteroskedasticity  
GLAR - Generalized Auto Regression  
IBCO - Improved Bacterial Chemotaxis Optimization  
ICA - Independent Component Analysis  
ICA - Independent Component Analysis  
JPY – Japan Yen  
MAD – Mean Absolute Deviation  
MAE - Mean Absolute Error  
MLP - Multilayer Perceptron  
MSE - Mean Squared Error  
NMSE - Normalized Mean Squared Error  
PCA - Principal Component Analysis  
PCA - Principal Component Analysis

PSNN - Pi-Sigma Neural Network  
PSO - Particle-Swarm Optimization  
RBF - Radial Basis Functions  
RMSE - Root Mean Square Error  
RNN - Recurrent Neural Network  
RPNN - Ridge Polynomial Neural Network  
RW - Random Walk  
SOM - Self-Organizing Maps  
SVM - Support Vector Machine  
SVR - Support Vector Regression  
SWT - Stationary Wavelet Transforms  
TAIEX - Taiwan Capitalization Weighted Stock Index  
TEMA - Triple Exponential Moving Average  
USD - US Dollar  
VC - Vapnik-Chervonenkis  
WT - Wavelet Transform



University of Moratuwa, Sri Lanka.  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)