Informative Correlation Extraction from and for Forex Market Analysis

Lei Song

A thesis submitted to Auckland University of Technology
in fulfillment of the requirements
for the degree of Master of Computer and Information Sciences

May, 2010

School of Computing and Mathematical Sciences

Primary Supervisor: Dr. Shaoning Pang
Secondary Supervisor: Prof. Nikola Kasabov

Abstract

The forex market is a complex, evolving, and a non-linear dynamical system, and its forecast is difficult due to high data intensity, noise/outliers, unstructured data and high degree of uncertainty. However, the exchange rate of a currency is often found surprisingly similar to the history or the variation of an alternative currency, which implies that correlation knowledge is valuable for forex market trend analysis.

In this research, we propose a computational correlation analysis for the intelligent correlation extraction from all available economic data. The proposed correlation is a synthesis of channel and weighted Pearson's correlation, where the channel correlation traces the trend similarity of time series, and the weighted Pearson's correlation filters noise in correlation extraction. In the forex market analysis, we consider 3 particular aspects of correlation knowledge: (1) historical correlation, correlation to previous market data; (2) cross-currency correlation, correlation to relevant currencies, and (3) macro correlation, correlation to macroeconomic variables.

While evaluating the validity of extracted correlation knowledge, we conduct a comparison of Support Vector Regression (SVR) against the correlation aided SVR (cSVR) for forex time series prediction, where correlation in addition to the observed forex time series data is used for the training of SVR. The experiments are carried out on 5 futures contracts (NZD/AUD, NZD/EUD, NZD/GBP, NZD/JPY and NZD/USD) within the period from January 2007 to December 2008. The comparison results show that the proposed correlation is computationally significant for forex market analysis in that the cSVR is performing consistently better than purely SVR on all 5 contracts exchange rate prediction, in terms of error functions MSE, RMSE, NMSE, MAE and MAPE.

However, the cSVR prediction is found occasionally differing significantly from the actual price, which suggests that despite the significance of the proposed correlation,

how to use correlation knowledge for market trend analysis remains a very challenging difficulty that prevents in practice further understanding of the forex market. In addition, the selection of macroeconomic factors and the determination of time period for analysis are two computationally essential points worth addressing further for future forex market correlation analysis.

Acknowledgment

I would like to thank all people who have helped and inspired me during my master study.

I especially want to thank my advisors, Dr. Paul S. Pang and Prof. Nik Kasabov, for their guidance during my research and study at Auckland University of Technology. Their perpetual energy and enthusiasm for research have motivated all their students, including me. In addition, Dr. Paul was always accessible and willing to help his students with their research. As a result, research life became smooth and rewarding for me. Many thanks also go in particular to Joyce D'Mello, for being supportive and helpful whatever the occasion.

All my lab buddies at the KEDRI made it a convivial place to work. In particular, I would like to thank Gary Chen and Kshitij Dhoble for their friendship and help during my thesis. All other folks, including Harya Widiputra and Yingjie Hu, had inspired me in research and life through our interactions during the long hours in the lab. Thanks.

My deepest gratitude goes to my family for their unflagging love and support throughout my life; this thesis is simply impossible without them. I am indebted to my father, Degong Song, for his care and love. As a typical father in a Chinese family, he worked industriously to support the family and spared no effort to provide the best possible environment for me to grow up and attend school. He had never complained in spite of all the hardships in his life. I cannot ask for more from my mother, Shuqing Hou, as she is simply perfect. I have no suitable word that can fully describe her everlasting love for me. I remember her constant support when I encountered difficulties and I remember, most of all, her delicious dishes

Last but not least, thanks be to God for my life through all tests in the past years. You have made my life more bountiful. May your name be exalted, honoured, and glorified.

Contents

Chapter 1

Introduction

1.1 Background

The forex market is a complex, evolving, and a non-linear dynamical system. Financial forecasting is challenging due to the data intensity, noise/outliers, unstructured data, high degree of uncertainty, and hidden relationships in previous time series data. It is noticeable that influential occurrences such as political events, general economic conditions, and traders' extractions, seriously affect the variation of market, and sometimes even determine the market trend. Thus for financial market analysis, information beyond the market is often observed in addition to the historical data from the market investigated in terms of trend prediction. For example, while we observe the exchange rate of New Zealand Dollars (NZD) and US Dollars (USD), we have to pay attention to the related data of historical pass-through, Official Cash Rate (OCR), federal fund interest rate, employment, Gross Domestic Product (GDP), Purchasing Power Parity (PPP), monetary policy, etc. This is because that correlation exists widely in the forex market affecting the variation of the market, which makes financial time series forecasting extremely complicated.

The application of correlation to forex market analysis has been investigated in previous researches. The significance of correlation is verified in the forecasting of future market direction (Kirkpatrick & Dahlquist, 2006). Historical correlation is prevalently used in the literature. However, after the results have been less analysis data and inaccurate prediction. Walter and Lopez (2000) test historical correlation in the USD/DEM/JPY currency trio. The finding shows that the correlation extracted from historical data is not significant enough to ensure accurate market prices prediction. Also, Yao and Tan (2000) examine historical correlation in the AUD/USD, AUD/USD, CHF/USD, DEM/USD, GBP/USD, and JPY/USD for prediction of the period from Nov 1993 to July 1995 using a neural networks model. The prediction with historical correlation is found not performing as well as that incorporating other rules. In addition, a 1996 study (Bruno, Cyril & Yann, 1996) examines cross-currency correlation among 8 currencies exchange rate. The results indicate that strong correlation exists in international forex market.

In market analysis technical analysis and fundamental analysis are jointly used. Technical analysis utilize models and trading rules based on price and volume transformations, such as the relative strength index, moving averages, regressions, inter-market and intra-market price correlations, cycles or, classically, through recognition of chart patterns. On the other hand, fundamental analysis of a business involves analyzing its financial statements and health, its management and competitive advantages, and its competitors and markets. When applied to futures and forex, it focuses on the overall state of the economy, interest rates, production, earnings, and management (Benjamin & David, 2008). For the usages of technical and fundamental analysis on market correlation analysis, the Bank of England did a survey in 1992 among chief foreign exchange dealers based in London (M. P. Taylor & Allen, 1992). The results revealed that at least 90% of respondents perfered to use technical analysis to conduct correlation analysis for forex market when they were forming views one or more time horizons. In 2002, the Bank of Canada carried out an evaluation of fundamental analysis (DSouza, 2002), and identified that the correlations from fundamental analysis provided strong evidence for forex market trend variation. The Bank suggested that such correlation must be considered by forex traders.

Although the analysis of forex market in terms of correlations between technical and fundamental analysis is powerful and popular, using the two methods in conjunction is practically difficult, which is impossible even for a finance professional (Neely, 1997). Technical analysis only suits correlations analysis within shorter time horizons. When applying it to longer time horizons, obtained correlations show merely a varied meaning for different horizon, which often turns out to be inaccurate for ongoing analysis.

Fundamental analysis provides correlative data to many macroeconomic domains for obtaining long term trend. However, for short term trading it is difficult to make decisions just by general correlations to macroeconomic domains. Furthermore, if the financial conditions (i.e. macroeconomic) undergo changes over a longer period of time, such fundamental analysis has to be performed again. Meanwhile, when traders observe a currency, some related currency also need to be considered. For example, when we trade New Zealand dollars, Australia dollar is also observed. However, neither technical analysis nor fundamental analysis could produce this kind of data.

1.2 Research Objectives

This research aims to come up with a correlation extraction method for the collection of reliable and efficient correlation information for forex market analysis. The method concerns both technical analysis and fundamental analysis, and addresses three types of correlation knowledge extraction: 1) correlation to previous market performance for the same currency, called historical correlation or temporal correlation; 2) correlation to alternative currencies, called cross-currency correlation or spatial correlation; 3) correlation to domestic macroeconomic factors, called micro-correlation. The core of the proposed method is a synthesis of correlation extracted with the use of two correlation approximation: the channel correlation, approximating the trend similarity of time series by a graphical channel; and the weighted Pearson's correlation, estimating distance similarity of time series, meanwhile filtering out noise from the correlation extraction. The extracted correlation knowledge is examined in support vector regression (SVR) Forex time series prediction.

1.3 Research Contributions

The contributions from the thesis are:

1. A computational correlation extraction method is proposed that incorporates historical correlation, cross-currency correlation and mico-correlation for forex market analysis.

2. The proposed computational correlation analysis is demonstrated to be capable of dealing with a periodic trend in forex market analysis. Dealing with 'zig zag' type trend is often difficult in forex market analysis. The proposed method is able to extract correlation for such zig zag path time series by synthesizing channel method and weighted Pearson's correlation.

3. A correlation-aided Support Vector Regression (cSVR) is proposed for forex time series prediction, through which the correlation is able to be exploited by SVR for time series prediction.

1.4 Thesis Structure

Chapter 2 gives a review of previous studies on correlation analysis and motivations for the presented research. In the review, the traditional forex analysis methods-technical analysis and fundamental analysis are investigated first, followed by a discussion on statistical linear and nonlinear correlation analysis.

Chapter 3 describes the proposed correlation extraction method in the context of forex market analysis. The chapter begins with an introduction to the channel method and the weighted Pearson's correlation, which are two core components of the proposed correlation synthesis method. For the purpose of forex market analysis, the proposed correlation is employed for deriving three types of informative correlation data.

Chapter 4 describes the utilization and evaluation of correlation information that is extracted with the use of the proposed correlation analysis method in terms of time series prediction. To this end, typical Support Vector Regression (SVR) is extended for correlation aided time series prediction using the correlation data in addition to the observed time series data.

Chapter 5 commences with a brief introduction of machine learning methods for time series prediction. Next, the chapter discuss to the experiments and analyses the five currencies trading that are used for the evaluation of extracted correlation knowledge.

Chapter 6 presents the conclusions of the thesis and directions for future work.

Chapter 2

Literature Review and Motivations

The chapter reviews technical analysis, fundamental analysis, and statistical correlation analysis methods, followed by motivations of the presented research.

Technical analysis and fundamental analysis are the most popular methods used to visualize and analyze the behavior of forex trading time series. Technical analysis falls into the security analysis discipline. It is used for forecasting the future direction of prices through the study of past market data, primarily price and volume (Kirkpatrick & Dahlquist, 2006). Technical analysis deals with presenting forecasts or trading advice that is largely based on the assessment of past prices, regardless of any underlying economic factors. Numerous technical analysis methods have been developed for financial market forecasting through different graphical representations. Moving Averages, Japanese Candlesticks, Oscillator and Miscellaneous Patterns are the most commonly used technical analysis methods. Moving Averages shows a movement of the average values of a security's price over certain periods. It is often used to measure the momentum and define areas of possible support and resistance. Japanese Candlesticks are formed by using the opening, highest, lowest and closing price on each trading day, where a series of candlesticks indicate the movement of a currency. Oscillator is banded between two extreme values and built with the results from a trend indicator, which helps to determine whether a market is in an overbought or oversold condition. Miscellaneous Patterns are comparisons of current market movement with previous patterns. It aims to predict the movement of price in the near future. Since the technical analysis does not consider any economic and financial factors of the observed organization or country, it can not provide any long-term forecasting. A 2002 study (Saacke, 2002) tested technical analysis on exchange rate of DEM/USD from 2 Jan, 1979 to 25 July, 1994 by moving average trading rules. Their study confirmed that the moving average trading rules are highly profitable in short term analysis, but they can not always predict exchange rate which seems warranted by the use of rules of the exchange rate.

Fundamental analysis is a method for evaluating the security (i.e. a note, stock, preferred shares, bonds, debentures, future, swap, rights, warrants, or virtually any other financial asset) by measuring its intrinsic value and examining the related economic, financial and other qualitative and quantitative factors (Abarbanell & Bushee, 1997). Fundamental analysis provides a forecasting on a currency by evaluating the macroeconomic factors and individual specific factors of a country, where macroeconomic factors include the overall country's economic conditions and business management of the government. For instance, when a currency is under observation, the economic growth and business/economy cycle become two important macroeconomic factors that need to be evaluated. Economic growth is an increase in the amount of the goods and services, produced by an economy over time. Business cycle refers to the fluctuations of economic activity that occur over several months or years. These fluctuations are often measured by using the growth rate of the real gross domestic product. In practice, if the investors are confident in the trend of economic growth or business/economy cycle, then they will invest into that currency. Fundamental analysis is more reliable than technical analysis, it is an appropriate solution to forecasting a long term trend of a currency. However, it cannot be interpreted by charting (which is a form of technical analysis), as this provides just a trend, rather than a specific value prediction. Hilde and Hvard (Hilde & Hvard, 2006) investigated the relation among purchasing power parity (PPP), interest rate and exchange rate. Their findings indicate that PPP and interest rate influence the exchange rate, but sometimes prove to be insignificant.

2.1 Technical Analysis

Normally, people analyze forex market by using technical analysis. Technical analysis is based on graphical representation and is also called charting. Many types of technical analysis have been created. Some of the commonly used technical analyses are Moving Averages, Japanese Candlesticks, Oscillator and Miscellaneous Patterns.

2.1.1 Moving Averages

Among smaller investors, Moving Averages are one of the most popular technical analysis tools. The calculation is based on time periods' closing prices of a share or index. The moving average (MA5) represents the calculation that uses the previous 5 days moving average at a given time (T) as shown in Eq.(2.1).

$$MA(T) = \frac{1}{w} \sum_{i=T-w-1}^{T-1} y_i, \tag{2.1}$$

where w is the window size and y_i is the closing price of a day(i). In this case, the current moving average at time (T) is calculated by using the mean of previous 5 days moving average. Similarly, the $MA(T-1)$ represents the number of the moving average of previous day. (Neely, 1997)

In stock market analysis, the moving average calculations are normally divided into 5 days (MA5), 10 days (MA10), 20 days (MA20), 60 days (MA60), 120 days (MA120) and 250 days (MA250). If the moving average of MA_c is greater than MA_t (where $t > c$), then the price of stock or index is in a higher expectation price. Also, if the moving average of MA_c is lower than MA_t (where $t > c$), then the price of stock or index is in a lower expectation price. It is noticeable that t is always greater than c. The decision of buying and selling is based on the graphical intersection created

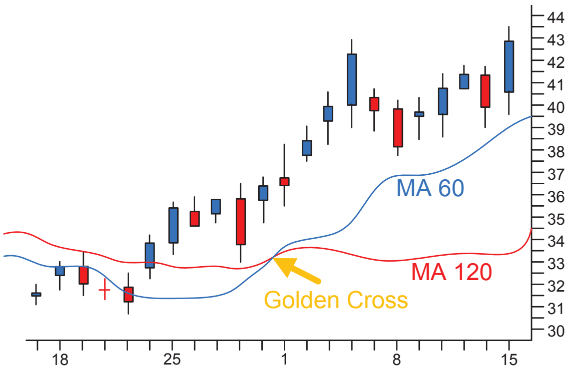


Figure 2.1: Moving Average

by MA_c and MA_t, where $t > c$. If the graphic line of MA_c intersects MA_t from below, then it indicates the pip is going up. And vice versa if the MA_c intersects MA_t from above, then it indicates the pip is going down. Similarly, if MA_c intersects multiple MA_{t_i}s' (where $t_i > c$) from below then it indicates a strong trend of the pip rising. They are shown in Figure.2.1. Therefore, depending on the indication of the pip behavior, people can buy and sell when the pip rises and drops respectively.

2.1.2 Japanese Candlesticks

Japanese Candlesticks is a widely used stock analysis tool. It was developed by Munehisa Homma for rice market trading in 1700's in Japan (Pring, 2002). The candlesticks show the highest price, lowest price, opening price and closing price for each day's trading. The vertical narrow line shows the price range of the day, the coloured rectangular box shows the relationship between opening price and closing price. The colours adopted for the rectangular boxes varies depending on the software vendor. If closing price is above the opening price, then the box is not filled with any colour (e.g. green or blue). On the other hand, if the opening price is above the closing price then the box is filled (black or red). The candles are shown as Figure.2.2. A single candlestick that includes the vertical narrow line and the rectangular box cannot show the trend of stock. The trend can only be observed or predicted if ten or more days of data are available.

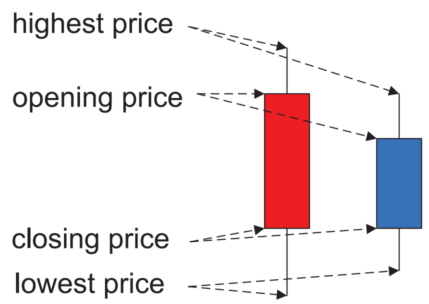


Figure 2.2: Japanese Candlesticks. The left is a negative candle, the right is a positive candle

There are three observed patterns which are important for investors. These candlesticks patterns are bullish reversals, bearish reversals and continuation patterns. Bullish reversals happen when a downward trend is going to stop and reverse to upward trend. Bearish reversals is opposite of Bullish reversal where the upward moving trend stops and reverses is direction to downward trend. In the following descriptions, only strong Reversals patterns have been explained.

Bullish Reversals

1. Abandoned body candlestick

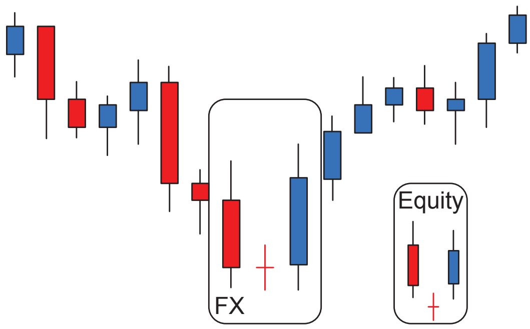


Figure 2.3: Abandoned body candlestick

The graph in Figure.2.3 depicts an abandoned body which is a rare bullish reversal pattern. It is identified by a large downward move (i.e. red candlestick) followed by a small red candle (cross), and a third candle (i.e. blue candlestick) moving in the opposite direction. As shown from the hierarchy Figure.2.3, the abandoned body candlestick falls under the highly reliable bullish reversal pattern. Above graph represents an abandoned body candlestick with a Bullish direction. The prediction based on this technique is highly reliable. A red coloured candlestick represents a bearish trend and is also called as "a red day". Whereas, a blue candlestick is called "a blue day".

The first red candle in the graph represents a continuation of the bear market. After that the small candle which reflects a trading in small range, suggests uncertainty in trend. Since each candle represents a trading day, therefore the above graph shows a moderate strength bullish pattern till the second day. The gapping caused by the abandoned body is common in less efficient market. Since currency market offers 24 hour trading, abandoned body is only seen to some extent after weekends.

2. Morning Doji Star

The morning doji star (Figure.2.4) is represented by a long red day followed by a doji and a blue day. The doji is the small blue cross in the second

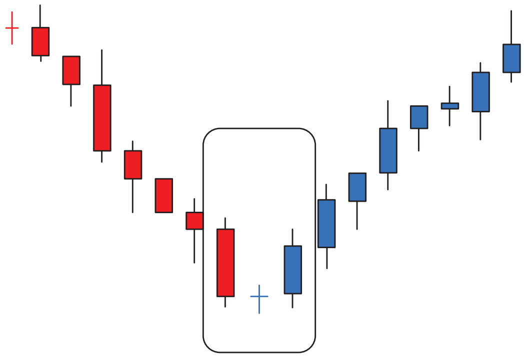


Figure 2.4: Morning Doji Star

trading day. It happens very rarely but it is a strong bullish reversal signal in the forex market. The doji indicates the continuation of a bearish trend. It reflects indecisive movement of trend. After the day of indecisiveness, the confirmation of the trend reversal has been made by a strong upward move on the third day.

3. Three Inside Up

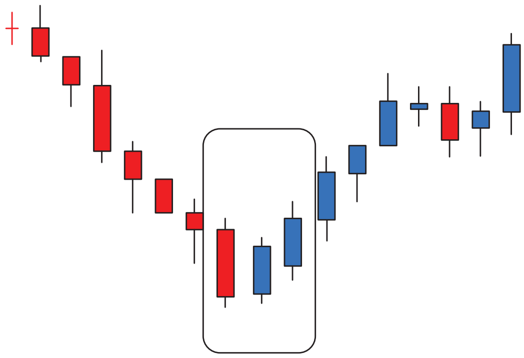


Figure 2.5: Three Inside Up

The 'Three Inside Up' (Figure.2.5) is represented by a long red day followed by a blue day that trades up to the midpoint of the first day, and a third blue day carrying the price above the first bear candle. The second days pattern is referred as a Bullish Harami pattern. Haramis show a clear-cut formation reflecting buyers overtaking the strength of the moving down trend. It often means a continuing rally in price. The 'Three Inside Up' is an additional confirmation that the price will enter a long lasting upward trend.

4. Three Outside Up

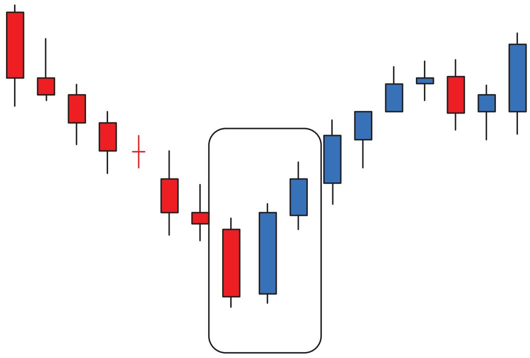


Figure 2.6: Three Outside Up

The 'Three Outside Up' (Figure.2.6) pattern occurs after a downward trend where the closing price of the second day is above the opening price of the previous day. Consecutively, the closing price of the third day is higher than the closing price of the second day. It reflects buyers overtaking selling strength and often leads to a continuous rally in price.

5. Three White Soldiers

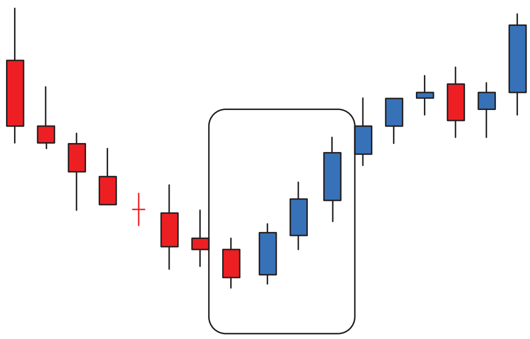


Figure 2.7: Three White Soldiers

The 'Three White Soldiers'(Figure.2.7) pattern occurs after a moving down trend. Three consecutive long blue days show up where the closing price of each day is higher than the previous day. This is a strong indicator of the trend moving upwards. But if the candles are overextended the analysts worry that it will cause the market to be overbought, consequently resulting in a pause.

Bearish Reversals

In bearish reversals, abandoned body, three inside down and three outside down are exactly the opposite to the bullish reversals. The three different patterns are explained below.

1. Dark Cloud Cover

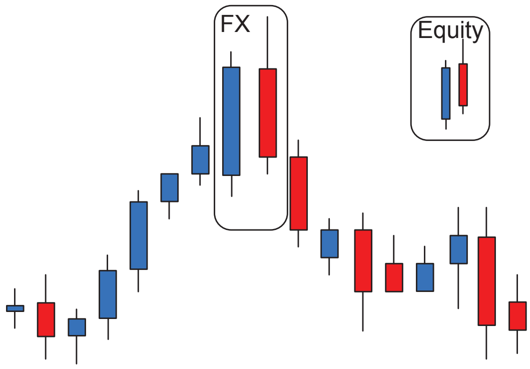


Figure 2.8: Dark Cloud Cover

The dark cloud cover (Figure.2.8) represents a long blue day follows with a red candle. The red candle closes below the middle of the body of previous day. In Forex trading, it is considered to be a good sign if the second day opens very high, since it shows that more sellers will be able to drive the price back down. A dark cloud cover confirms sellers have controlled the market and this long upward trend loses its bullish momentum. A declining trend follows a dark cloud cover. Based on this pattern the sellers confidently keep on selling out the stocks until a clear stop signal is given at the highest price on the second day. If the second candle does not reach below the middle of the first candle, traders normally feel safe and wait for confirmation on the third day.

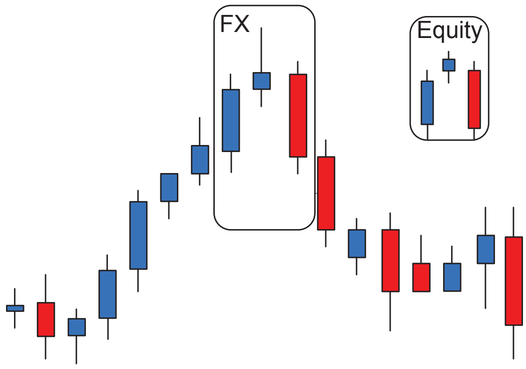


Figure 2.9: Evening Doji Star

2. Evening Doji Star

   Following a bullish movement, an evening doji star (Figure.2.9) shows up. The second candle shows a trend continually moving up. Therefore a sell-off keeps the opening and closing prices very close or same. However, the first two days blue candle indicates that the upward trend has lost momentum. The first two days show a pattern similar to the Bearish Shooting Star which is a weak-to-moderate strength reversal pattern. The Shooting Star can be of any colour and follows an additional sell-off on the third day. If the Shooting Star indicator is low, there should be a reversal confirmation on the next day lead by a red candle. Traders should sell off the stock based on this confirmation on the next day.

Continuation Patterns

All patterns except strong Reversal Patterns can be Continuation Patterns. However, the Continuation Patterns cannot be marked since they depend on the trends movement, therefore it can be very difficult and risky to mark this as a Continuation Pattern.

2.1.3 Miscellaneous patterns

Miscellaneous Patterns are artistic technical analysis. It has been proved to be very useful for predicting movement of price in future by experiences of practitioners.

1. Head and Shoulders

Head and shoulders (Figure.2.10) often occur in stock or forex market. It is formed by two shoulders and a head rising above between the two shoulders. As the diagram shows, the price after a head and shoulders follows a down trend. This pattern is one of the highly reliable signals of predictors. It also can be seen as an upside down formation.

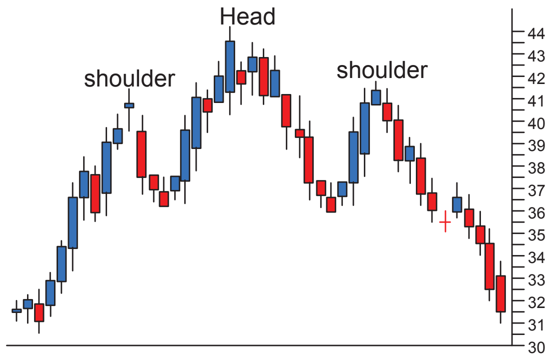


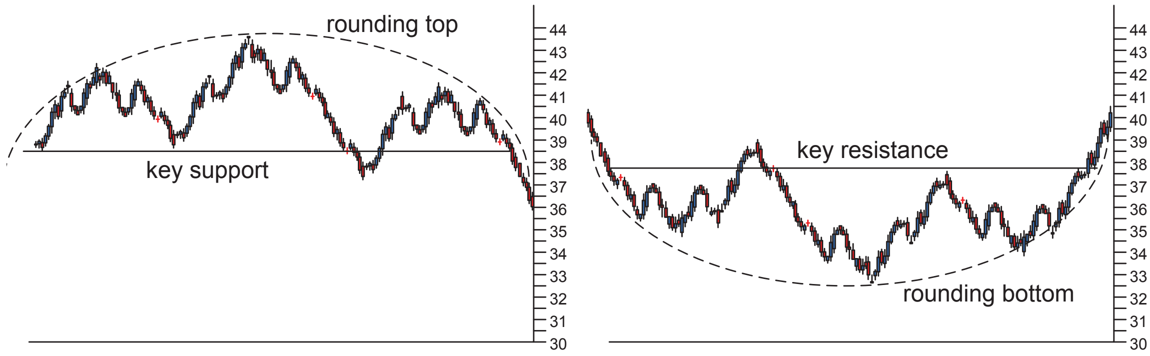
Figure 2.10: Head and Shoulders

2. Saucer tops and bottoms

   Saucer tops (Figure.2.11a) and bottoms (Figure.2.11b) are also called rounding tops and bottoms. They are made by a gradual change in supply and demand of the market. Normally, the shape is fairly symmetrical as the prices go upwards and downwards. These patterns are very rare. They do not contain any predictive information about the new trends direction.

3. Double and triple tops and bottoms

   Double (Figure.2.12a and Figure.2.12b) and triple (Figure.2.12c) tops and bottoms are very rare patterns. The triple occurs more rarely than the double. A double top shows an“M”-shaped pattern and a double bottom shows a“W”-shaped pattern. The triple top and bottom are similar but there are three peaks or troughs. A triple top means the pattern has three peaks and in-between them the trend fall below the support line. It places a stop-loss just above the last peak. In this case, price normally goes back to the support line which then acts as a resistance level. It is also a reversal signal and therefore places a stop-loss just over the resistance level. The triple bottom is an opposite of triple top where the trend fall till the resistance line and bounces back to



(a) rounding top (b) rounding bottom

Figure 2.11: Saucer tops and bottoms

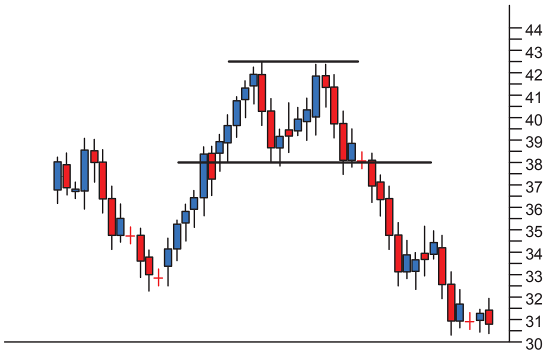
an upward trend. It also places a stop-loss just below the last trough. In fact, the price often goes back to the resistance line which than acts as a support level. It goes short on a reversal signal and places a stop-loss just below the support level.

2.1.4 Oscillator

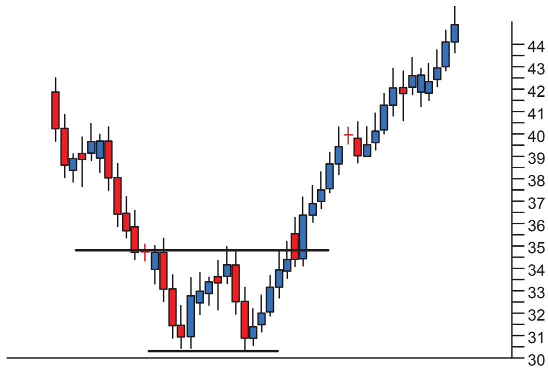
Oscillator pattern displays a fluctuating pattern occurring above and below the average line or between set levels whose value changes over time. It shows extreme levels of upward and downward trends for extended periods. However this trend does not last for a continued period. When the upward trend goes above the upper set level then this is a strong indication for sellers market. Similarly, when the downward trend falls below the lower set level then it is an indication for buyers market.

1. Moving Average Convergence/Divergence (MACD)

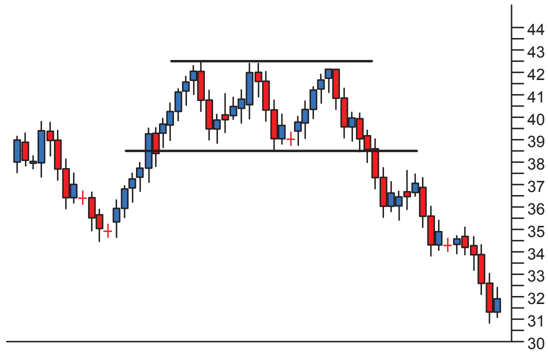
   MACD (Figure.2.13) provides the difference between a fast and slow Exponential Moving Average (EMA) (Hunter, 1990) of closing prices. The most standard periods were designed by Gerald Appel in the 1960s (Appel, 2005) for 12 and 26 days. EMA considers weighting factors to exponentially reduce the closing price. The weighting for each previous data point decreases or increases exponentially, giving much more important information to recent



(a) Double top



(b) Double bottom



(c) Triple top

Figure 2.12: Double and triple tops and bottoms

observations from previous observations. It is calculated as Eq.(2.2).

$$EMA = \alpha \times (p_1 + (1-\alpha) \times p_2 + (1-\alpha)^2 \times p_3 + (1-\alpha)^3 \times p_4 ... (1-\alpha)^{(i-1)} \times p_i) \quad (2.2)$$

Smoothing factor α is the degree of weighing decrease expressed. $\alpha = \frac{2}{i+1}$, therefore α is only between 0 and 1. p_i is the data point. The MACD is calculated as Eq.(2.3).

$$MACD = EMA_{12} \; of \; price - EMA_{26} \; of \; price \quad (2.3)$$

In general, there are three trading signals given by MACD. The signal line (Figure.2.13) in MACD method represents a trading rule.

Trading signal 1: If the MACD crosses the signal line from top direction, it

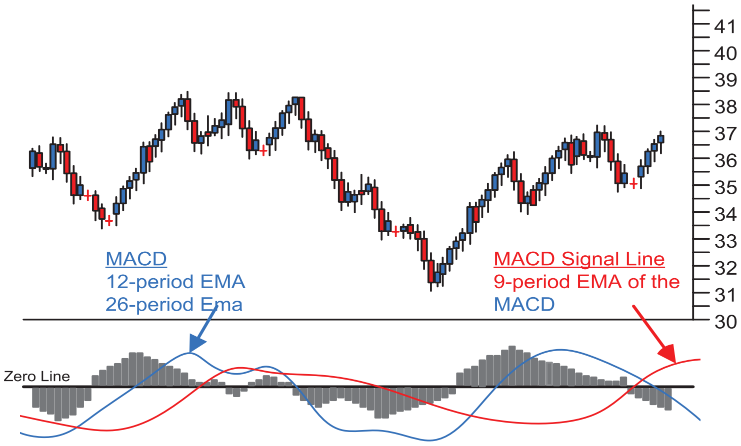


Figure 2.13: Moving Average Convergence/Divergence (MACD)

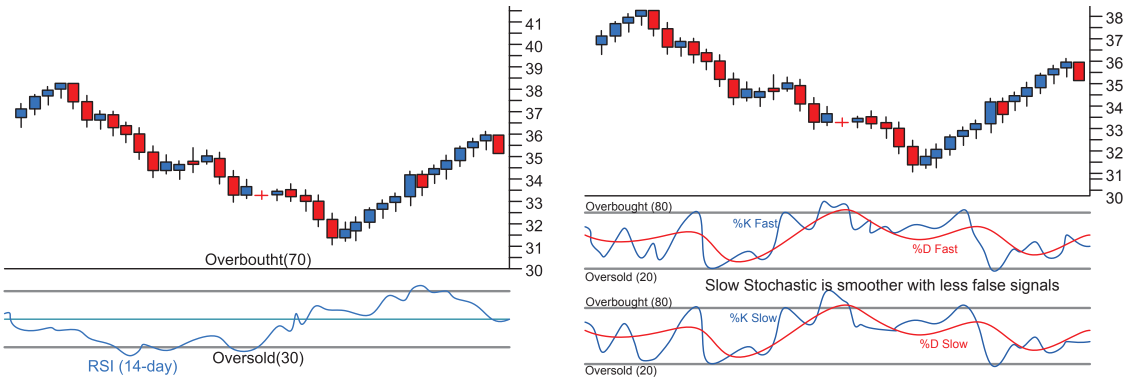
is a signal for selling. Conversely, if MACD line crosses the signal line from below, it is a signal for buying.

Trading signal 2: If the MACD crosses the zero line from top direction, it signifies a bearish trend. Conversely, if MACD line crosses the zero line from below, it indicates a bullish trend.

Trading signal 3: A divergence occurs between price and histogram or between MACD line and price. A positive divergence occurs when MACD forms consecutive higher 'lows' or when the second low is higher than the previous 'low'. This pattern is a strong indication that the upcoming trend will be bullish. In this scenario there is a possibility of bullish reversal.

Similarly, for negative divergence, MACD forms consecutive higher 'high' or when it forms the second 'high' which is lower than the previous 'high'. This pattern is a strong indication that the upcoming trend will be bearish. In this scenario there is a possibility of bearish reversal.

2. Relative Strength Index (RSI) and the Stochastic Oscillator (STO)



(a) Relative Strength Index (RSI) (b) Stochastic Oscillator (STO) Fast & Slow

Figure 2.14: Relative Strength Index (RSI) and the Stochastic Oscillator (STO)

$$\begin{aligned} U = &\quad close_{today} - close_{yesterday} \\ D = &\quad 0 \ or \\ D = &\quad close_{today} - close_{yesterday} \\ U = &\quad 0 \end{aligned} \tag{2.4}$$

Eq.(2.4) shows, U is upward and D is downward. If yesterdays and todays closing price are same, than both U and D equal to zero. The average of U is calculated with an exponential moving average by means of the given (i) days smoothing factor, and same for D. The Relative Strength is calculated through the ratio of those averages (Wilder, 1978; Park & Irwin, 2004). Eq.(2.5) describes the form of RSI.

$$\begin{aligned} RS = &\quad \frac{EMA(i)ofU}{EMA(i)ofD} \\ RSI = &\quad 100 - 100 \times \frac{1}{1+RS} \\ RSI \in &\quad \{0, 100\} \end{aligned} \tag{2.5}$$

STO (Murphy, 1999; Schwager, 1996) has two indicators: K (fast) and D (slow). They are used to determine the future variations in price. The difference in latest closing price and the lowest price over the previous N number of days is used for calculating K (fast)(Eq.(2.6)).

$$K = \frac{ClossingPrice_{today} - Low_{lowestNdays}}{High_{highestNdays} - Low_{lowestNdays}} \times 100 \tag{2.6}$$

And D (slow) is calculated with the simple moving average (SMA) (Chou, 1975) of the Stoch K statistic across s periods (Eq.(2.7) and Eq.(2.8)).

$$SMA = \frac{\sum_{N=1} P}{N} \tag{2.7}$$

$$D = \frac{SMA_s of(ClossingPrice_{today} - Low_{lowestNdays})}{SMA_s of(High_{highestNdays} - Low_{lowestNdays})} \times 100 \tag{2.8}$$

The previous n day's closing prices unweighted average value is also technically known as simple moving average (SMA).

The bands in RSI (Wilder, 1978; Park & Irwin, 2004) for overbought and oversold are usually set at 70 and 30. If the RSI value is greater than 70, the market is in an overbought situation. If the RSI value is less than 30, the market is in an oversold situation. Similarly, in STO the overbought and oversold are set at 80 and 20. Both RSI and STO have an overbought and oversold setting value ranging from 0 to 100.

2.2 Fundamental Analysis

Fundamental analysis is based on macroeconomics factors. Some of the factors have been proven to affect exchange rates. The commonly used macroeconomics factors are: interest rate, Purchase Power Parity (PPP), Gross Domestic Product (GDP) and monetary policy.

2.2.1 The Relationship Between Changes in Interest Rates and Exchange Rates

Interest rate is defined as the percentage that is charged, or paid, for the use of money. Here, the interest rate is paid by central bank when money is deposited.

The interest rate influences the demand and the supply of currencies on the forex market. The speculative purposes of forex trading is moving funds from one currency to another, in order to take advantage of price movements or to take advantage of better returns in another country. For example, if the federal reserve interest rate in the U.S was 0.25% and the Official Cash Rate (OCR) in New Zealand was 2.5%, there are advantages gained from moving money from US dollars based securities to NZ dollars, because NZ banks are paying interest that is ten times higher than US banks. In this case, a move towards selling US dollars on forex market and buying NZ dollars is expected, which results in increasing demand of NZ dollars. Therefore, the NZ dollar would get a pressure to push its value up against US dollars.

The previous studies have proved that interest rate influence exchange rate. According to the survey in 1988(Goodhart, 1988) , they tested the interest rate against UK pound. Their results show that the relationship between interest rate and exchange rate is positive. Fleming and Remolona (1999) examine if the exchange rate is influenced by interest rate on US dollars to other currencies from 23 Aug, 1993 to 19 Aug, 1994. The results also shows the positive correlation between them. A working paper (Snchez, 2005) in European Central Bank defines relationship between the interest rate and exchange rate as shown in Eq.(2.9), given time point t,

$$r_t = R_t - E_t\pi_{t+1}, \tag{2.9}$$

where r is the real interest rate; E is the real exchange rate; R is interpreted as a risk premium term and π is a simple aggregate supply schedule which states that prices (t+1) are determined by the last period expectations of the current (t) price level. It gives more evidence of the positive relationship between them.

2.2.2 The Relationship Between Changes in Purchase Power Parity (PPP) and Exchange Rates

The PPP uses two countries long-term equilibrium exchange rates in order to equalize their purchasing power (Cassel, 1918). It states that identical goods should have only one price in ideally efficient markets. Bases on the theory of PPP, if a country has a relatively high inflation rate, then the value of its currency will decrease. For example, lets consider two fictional countries: A and B. The price of everything was the same in 2006, e.g. can of coke cost 1.5 dollars in both countries. If PPP holds, 1 dollar in country A must be worth 1 dollar in country B, otherwise there will be a risk-free profit buying a can of coke in country A and selling it in country B. So PPP here requires a 1 for 1 exchange rate. Suppose inflation rate in country A was 50% and inflation rate in country B was zero in 2008. If the inflation in country A impacts all products equally, then the price of a can of coke would be 2.25 dollars in 2008. Since there is no inflation in country B, the price of a can of coke would still be 1.5 dollars in 2008. If PPP holds, there is no profit from buying coke in country B and selling it in country A, then 2.25 dollars in country A would cost 1.5 dollars in country B at that time. If 2.25 dollars in country A equals to 1.5 dollars in country B, then 1.5 dollars in country A must equal 1 dollar in country B. Thus, it will cost 1.5 dollars in country A to purchases 1 dollar in country B on foreign exchange markets. If there are differing rates in both countries, the relative prices of products in the two countries will change e.g. the price of coke. The relative price of products is linked to the exchange rate through the PPP theory.

Previous studies have tested performance of PPP influencing exchange rates. Frankel and Rose (1996) examined the relationship between PPP and real exchange rates using a panel of 150 countries in the previous 45 years. Their results show a strong evidence that PPP movement is similar to long term exchange rate trend. The same evidence between PPP and exchange rates is also shown in a study reported by Abuaf and Jorion (1990) study. They re-examine the evidence on Purchasing Power Parity (PPP) in 10 European countries and their currencies from Jan, 1973 to Dec, 1987. A recent study (Lothian & Taylor, 2000) examines exchange rate between the British Sterling and US dollar and how it influenced by PPPs in UK and USA from 1792 to 1990. In this long-term run, the exchange rate between the two countries is slowly adjusted by their PPPs.

2.2.3 The Relationship Between Changes in Gross Domestic Product (GDP) and Exchange Rates

The performance evaluation for the economic is done through the country's gross domestic product (GDP). The countries productions correlation with the standard of living is usually considered for calculating GDP. There are three ways in which GDP can be defined:

1. equals all final products and services total expenditure of a country annually.

2. equals every stage of productions total cost utilised by all the industries in a country, including the untaxed subsidies on products annually.

3. equals the overall generated income sum through production within a country, including employees' compensation, production taxes and gross operating surplus (or profits).

Normally, the exchange rate increases when GDP grows. A 2003 study (Broda, 2004) examines that the GDP influences exchange rate in 75 developing countries. This result shows that there is a strong positive correlation between GDP and exchange rate. Another study (Calvo, Leiderman & Reinhart, 1993) tested the factors that affect exchange rates between countries in Latin America and US. GDP also shows a strong influence to those exchange rates. Lane and Milesi-Ferretti (2005) review the relationship between GDP and exchange rate in their research, their study empirically explores some of the inter-connections between financial factors and exchange rate adjustment. GDP is a very important factor on evaluating a currency.

2.2.4 The Relationship Between Changes in Monetary Policy and Exchange Rates

Monetary policy is the process controlled by the government, the central bank, or a monetary authority of a country. It controls the following items:

1. money supply;

2. money availability;

3. and interest rates.

The goal is to align its objectives with the economy's growth and stability. Monetary policy can be either an expansionary policy, or a contractionary policy. Expansionary policy is intended to augment the total money supply in the economy for reasons such as countermeasures against unemployment during depression. This countermeasure allows lowering of the interest rate. Alternatively, contractionary policy is intended to reduce the total supply of money and raises interest rates as a countermeasure against inflation. Distinct from fiscal policy, monetary policy refers to a government borrowing, spending and taxation.

The reason for monetary policy influencing exchange rate is that monetary policy controls inflation in a country. A high inflation rate leads to a decrease of a country's currency price. J. B. Taylor (2001) has reviewed that the most national central banks' control setting new monetary policy for dealing with inflation increases and interest rate, the exchange rate therefore floats and drifts follows the monetary policy. Gali and Monacelli (2005) review how three alternative monetary policy regimes for the small open economy to control the exchange rate in a long term run. Devereux and Engel (2003) investigate the implications of monetary policies for exchange-rate flexibility by reviewing many previous studies. Their findings are that optimal monetary policy results in a fixed exchange rate regardless of country–specific shocks.

2.3 Correlation Extraction Methods

Correlation in statistics indicates the strength and direction of a relationship between two random variables (Rodgers & Nicewander, 1988). Depending on correlation distributions, correlation can be categorized into two main types: Pearson's Correlation (positive and/or negative linear correlation) and non-parametric correlation. The most popular correlation extraction method for forex market analysis is Pearson's correlation.

2.3.1 Linear Correlation

Pearson's correlation (Pearson, 1897) is briefed as follows. Given time series $X = \{x_1, x_2, \ldots, x_N\}$ and $Y = \{y_1, y_2, \ldots, y_N\}$, the Pearson product-moment correlation coefficient ($\rho_{X,Y}$) is calculated as:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E((X-\mu_X)(Y-\mu_Y))}{\sigma_X \sigma_Y}, \tag{2.10}$$

where *cov* is the covariance; σ_X and σ_Y are standard deviations; μ_X and μ_Y are the expected value; and E is the expected value operator. Practically, except $\rho_{X,Y}$, Pearson's correlation returns a probability p-value (p). p-value in statistical hypothesis testing is the probability of obtaining a test statistic at least as extreme as the one that was actually observed (Y to X), assuming that the null hypothesis is true. Null hypotheses are typically statements of no difference or effect. The p-values are crucial for their correct interpretation as they are based on this hypothesis. Therefore, a lower p-value or assumption of the null hypothesis can be thought of as the production of a statistically significant result. p is calculated as:

$$\mathrm{p} = \frac{1}{N-1}\sum_{i=1}^{N-1}\mathrm{p}_i \tag{2.11}$$

where,

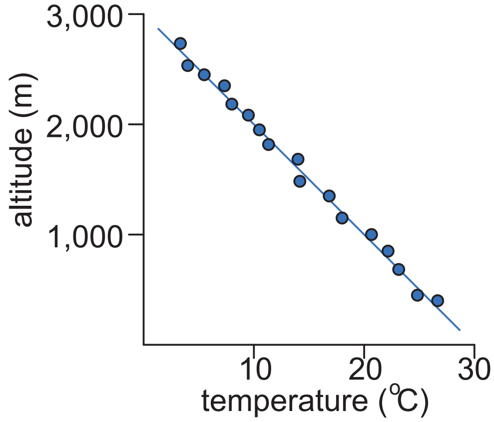
$$\mathrm{p}_i = \begin{cases} 0 & \text{if } \Delta x_i > 0 \textit{ and } \Delta y_i > 0 \\ 1 & \text{if } \Delta x_i < 0 \textit{ and } \Delta y_i > 0 \\ 1 & \text{if } \Delta x_i > 0 \textit{ and } \Delta y_i < 0 \end{cases} \tag{2.12}$$

Consider $\sigma_X^2 = E[(X - E(X))^2] = E(X^2) - E^2(X)$ Due to $\mu_X = E(X)$ and likewise for Y. Also, $E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$. Eq.(2.10) is often formulated with p as:

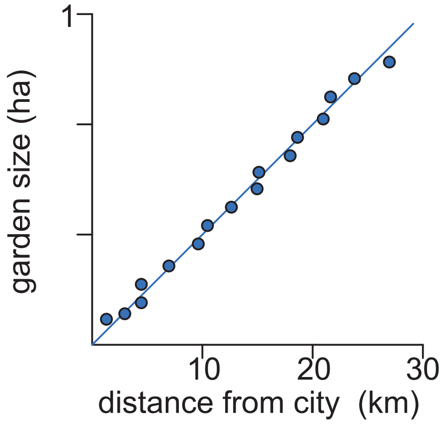
$$\begin{aligned} \rho_{X,Y} = & \quad \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}} \\ \textit{subject to}: & \quad \mathrm{p} < 0.05, \end{aligned} \tag{2.13}$$

$\rho_{X,Y}$ is ranged from +1 to -1, which follows that Pearson's correlation includes positive correlation and negative correlation. A positive correlation ($\rho_{X,Y} \rightarrow 1$) means that, as one variable/time series (X) becomes large, the other (Y) also becomes large, and vice versa. $\rho_{X,Y} \rightarrow +1$ means a perfect positive linear relationship between X and Y. In case of negative correlation($\rho_{X,Y} \rightarrow -1$), as one variable ($X$) increases the other ($Y$) decreases, and vice versa. Figure.2.15, explains the case of negative, positive, and no Pearson's correlation, respectively. Note that Pearson's correlation $\rho_{X,Y}$ is statistically significant, only if p is less than 0.05.

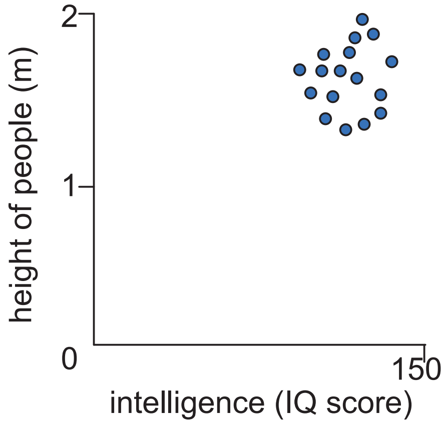
The advantage of using Pearson's correlation is that more accurate prediction can be made when a strong correlation exists amongst variables/time series patterns. The suitability of Pearson's correlation for financial market forecasting has been demonstrated by Kondratenko and Kuperin (2003). They used Pearson's correlation to aid neural networks (NN) to forecast the exchange rates between American Dollar to four other major currencies: Japanese Yen, Swiss Frank, British Pound and EURO. The results show that the NN gets better performance with Pearson's correlation extraction information than without it. Also, a recent study (Kwapien,



(a) negative correlation



(b) positive correlation



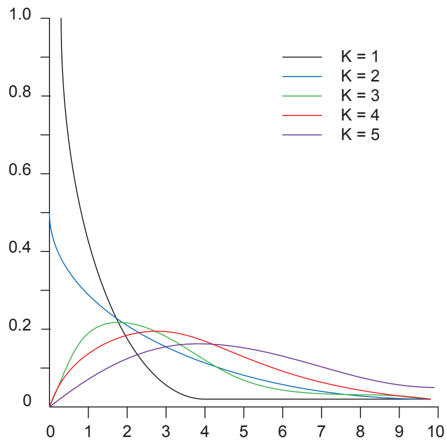
(c) no correlation

Figure 2.15: Linear Correlation. Temperature decreases when altitude increases. The garden size out of city is often bigger than inside of city. There is no correlation between height of people and their IQ.

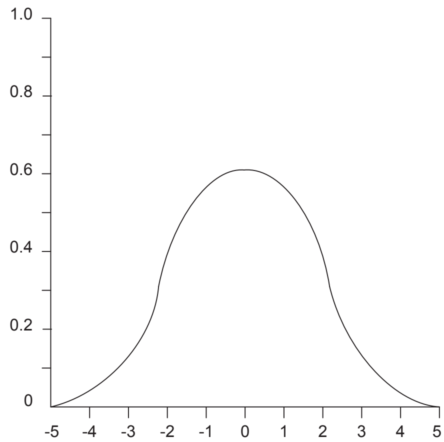
Gworek & Drozdz, 2009) tested NN model work with Pearson's correlation and found out that it results in better average internode distance on ten exchange rates when comparing to other correlation methods. However, both articles report that their Pearson's correlation aided time series prediction is not reliable.

2.3.2 Non-parametric Correlation

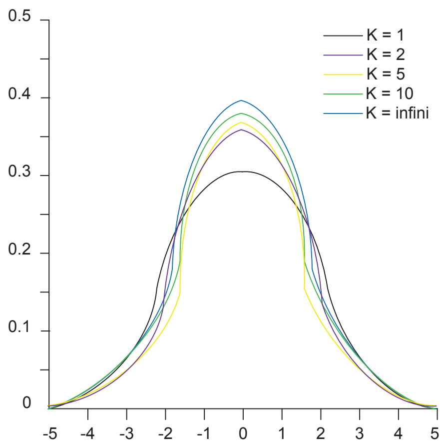
In contrast to Pearson's correlation influenced by outliers, unequal variances, non-normality, non-parametric correlation is calculated by implementing the Pearson's correlation formula to the ranks of the data, instead of the actual data values themselves. In doing so, several distortions present in the Pearson's correlation are reduced significantly. In the literature, Chi-square correlation (Plackett, 1983), Point biserial correlation (Tate, 1954), Spearman's correlation (Myers & Well, 2003), and Kendall's correlation (DETSKY et al., 1987) are some of the well known non-parametric correlation methods. Depending on the type of non-parametric correlation, they represent the correlated data distribution differently as curves shown in Figure.2.16



(a) Chi-square distributions



(b) Point biserial correlation (normal distributions)



(c) Spearman's ρ and Kendall's τ (student's τ-distribution)

Figure 2.16: Data distributions on different Non-parametric correlations

It is worth noting that the efficiency of a particular non-parametric correlation method depends on the type of probability distribution inherent in the data. Thus, different non-parametric correlations in practice have their characteristic applications. Chi-square correlation works well for age-adjusted death rates, life-table analysis (Mantel, 1963), lung cancer analysis (Paez et al., 2004) and cardiac resynchronization therapy (CRT) in heart failure (HF) (Yu et al., 2005). Point biserial correlation is used in the analysis of children reading attainment (Hewison & Tizard, 2004), schizophrenia research (Akdede, Alptekin, Kitis, Arkar & Akvardar, 2005) and academic achievement prediction (Deberard, Spielmans & Julka, 2004). Spearman's correlation usually performs well on psoriasis disease analysis (Gelfand et al., 2004), analysis of lung inflammation in asthma (Sutherland et al., 2004) and gaucher disease prediction (Boot et al., 2004). Kendall's correlation is unique on the analysis of drugs composition (Panackal et al., 2006), network coupled motions (Wong, Selzer, Benkovic & Hammes-Schiffer, 2005) and information ordering evaluation (Lapata, 2006).

2.4 Motivations for the Presented Research

Previous forex market analysis has confirmed that correlation information/knowledge to have an unique, sometime even deterministic role on market trend analysis and forecast, despite the chaotic variation of forex market. Thus from a technical viewpoint, correlation data is believed to be essential for any computational market analysis, in addition to the original data from the observed market. This holds especially when insufficient market data is available, or the observed market data gives little indication on future direction of market.

Aiming to extract correlation data significant for forex market trend analysis, in this work we developed a new correlation computing and synthesis approach, in which correlation knowledge is derived from historical data from the observed currency pair, relevant currency pairs, as well as important domestic/international microeconomics variables. Based on computational analysis of all available market data, the proposed correlation extraction would enable an ordinary trader to conduct expert market trend analysis the same way as a financial professional would do with his/her years of experience in traditional technical and fundamental analysis.

In statistics, standard correlation analysis method calculates a correlation coefficient based on a certain type of distance-base covariance calculation over every time point of the time line. Its worth noting that standard correlation counts just one type of distance similarity. Applying to market analysis, significant correlation knowledge on trend similarity might be lost because the time point mismatches happen to most financial time series. For example, given two periods time series in a similar increasing trend and varied zig zag paths, the obvious correlation on trend similarity is easily ignored as standard correlation calculation often gives a rather low coefficient due to the mismatches between the two zig zag paths.

As a solution to the problem, we take trend similarity between two time series as an important correlation for financial data analysis, and approximate the correlation graphically by a channel method followed by weighted Pearson's correlation method to extract the most similar and correlative patterns in the observed time series. Hence, we utilize technical analysis, fundamental analysis and correlation distribution theory in conjunction, to obtain efficient correlation data for learning, assisting computational inference models such as SVR for enhanced forex market forecast.

Chapter 3

The Proposed Correlation Analysis Method

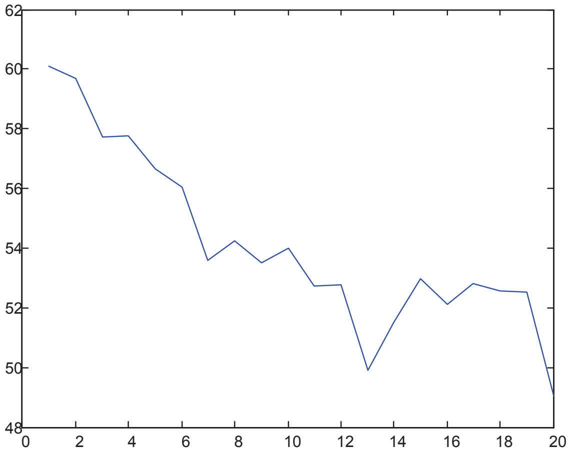
The proposed correlation extraction method is a hybrid method based on the channel model and weighted Pearson's correlation analysis. The channel model is used to model a concrete arc, approximating the time series for trend prediction and the weighted Pearson's correlation analysis adopts a method based on Person's correlation algorithm. The chapter discusses the theories of channel model and weighted Pearson's correlation analysis and illustrates them with real world examples.

Statistical correlation which measures the strength and the direction of a relationship between two time series by calculating a distance-based covariance for every point along the time line.

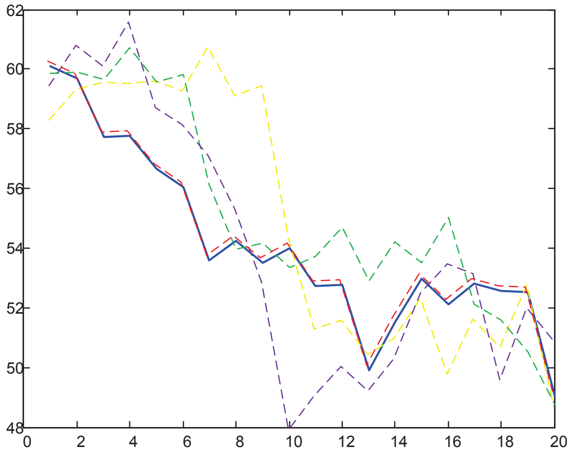
Unlike statistical correlation, the proposed correlation analysis achieves a balancing trade-off between trend similarity and distance similarity, which evaluates the correlation on trend similarity straightforwardly by a graphical channel approach, and the correlation on distance similarity through a weighted Pearson's Correlation analysis.

3.1 The Channel Correlation

Market movement varies over time, dynamically and evolutionarily, making it difficult to establish any particular rule that the market follows. However, it is noticeable that similar market variation often occurs in a pattern in the historical data. For example, on forex market, NZD to JPY in the period between 04 Nov, 2008 and 01 Dec, 2008, had turned on a falling-down as shown in Figure.3.1a. On searching historical data within 01 Dec, 2006 to 01 Dec, 2008, we find 43 very similar downward trends, and have 4 of them plotted with normalization in Figure.3.1b. Surprisingly, similar patterns are also found from other currencies, especially from AUD/JPY trading data. This could be normally explained by the fact that the trend of New Zealand's economy had been the same as that of Australian economy in previous years.



(a) the closing prices during 04 Nov, 2008 to 01 Dec, 2008



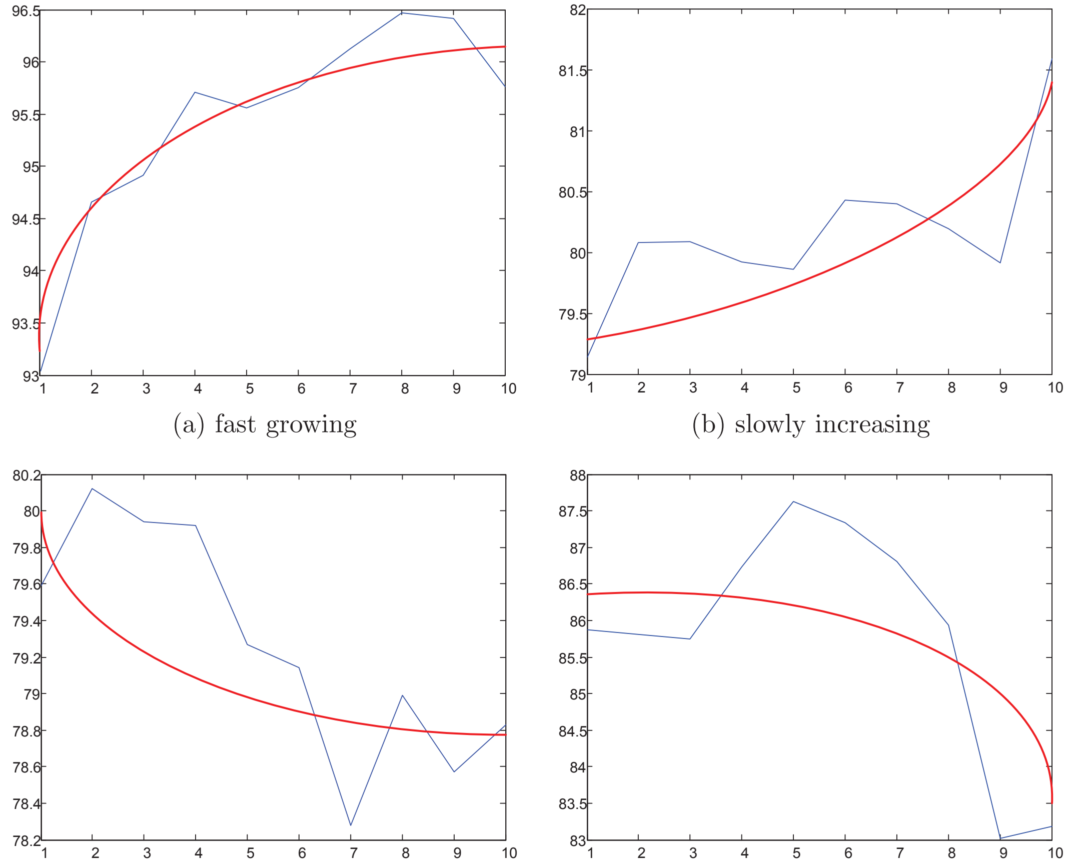
(b) 4 of 43 trends similar to (a)

Figure 3.1: An example of trend pattern for NZD/JPY forex occurred during 2006 to 2008. Solid line: observed time series; Dash line: similar patterns.

3.1.1 Channel Correlation Extraction

The channel method for correlation extraction is used to model a concrete arc, approximating graphically the trend similarity between two time series. Figure.3.2 depicts the diagrams of 4 typical trend patterns: fast growing, slowly increasing, fast dropping and slowly decreasing.

Straightforwardly, each of the above trend patterns can be described graphically by one piece of arc with its function formulated as a sub-circle shown in Figure3.3. In this way, we have the following 4 arc functions describing the 4 trend patterns,



(a) fast growing (b) slowly increasing

(c) fast dropping (d) slowly decreasing

Figure 3.2: Four trend patterns used for channel approximation

respectively.

$$(x-x_0)^2+(y-y_0)^2=R^2 \left| \begin{array}{l} x_0=0, y_0=R \\ x\in[0, sin\alpha \cdot R\sqrt{2(1-cos2\alpha)}] \\ \text{see Figure.3.3a} \end{array} \right. \tag{3.1}$$

$$(x-x_0)^2+(y-y_0)^2=R^2 \left| \begin{array}{l} x_0=R, y_0=0 \\ x\in[0, sin\alpha \cdot R\sqrt{2(1-cos(\pi-2\alpha))}] \\ \text{see Figure.3.3b} \end{array} \right. \tag{3.2}$$

$$(x-x_0)^2+(y-y_0)^2=R^2 \left| \begin{array}{l} x_0=0, y_0=0 \\ x \in [0,(1-cos\alpha)\cdot R\sqrt{2(1-cos(\pi-2\alpha))}] \\ \text{see Figure.3.3c} \end{array} \right. \tag{3.3}$$

$$(x-x_0)^2+(y-y_0)^2=R^2 \left| \begin{array}{l} x_0=R, y_0=R \\ x \in [0,(1-cos\alpha)\cdot R\sqrt{2(1-cos2\alpha)}] \\ \text{see Figure.3.3d} \end{array} \right. \tag{3.4}$$

where $\alpha \in (0, \pi/4)$. $\angle\alpha$ is the parameters reflecting the speed of increasing or decreasing trend. Radius R determines the length of the trend pattern corresponding to the time period of observation. In practice, a discrete arc will be produced according to the length of time series for channel approximation.

Given the observed time series X with N data points and another time series Y with T points, $N \leq T$. Applying Eq.(3.1) - Eq.(3.4) to X, respectively, one of 4 types arc (i.e. functions) called 'channel pattern' is selected with its parameter α tuned to best suit the time series under observation,

$$p^* = \arg\min_{\alpha, i\in[1,4]} \frac{\sum_{t=1}^{N}\|p_t^i - x_t\|}{N}. \tag{3.5}$$

For discovering a correlation of Y to X, an Euclidean mean distance from the observed time series X to the channel pattern p^* is estimated at every time point t:

$$d_t = \frac{\sum_{t=1}^{N}\|p_t^* - y_t\|}{N}. \tag{3.6}$$

Then, correlation extraction is carried out through a shifting distance comparison:

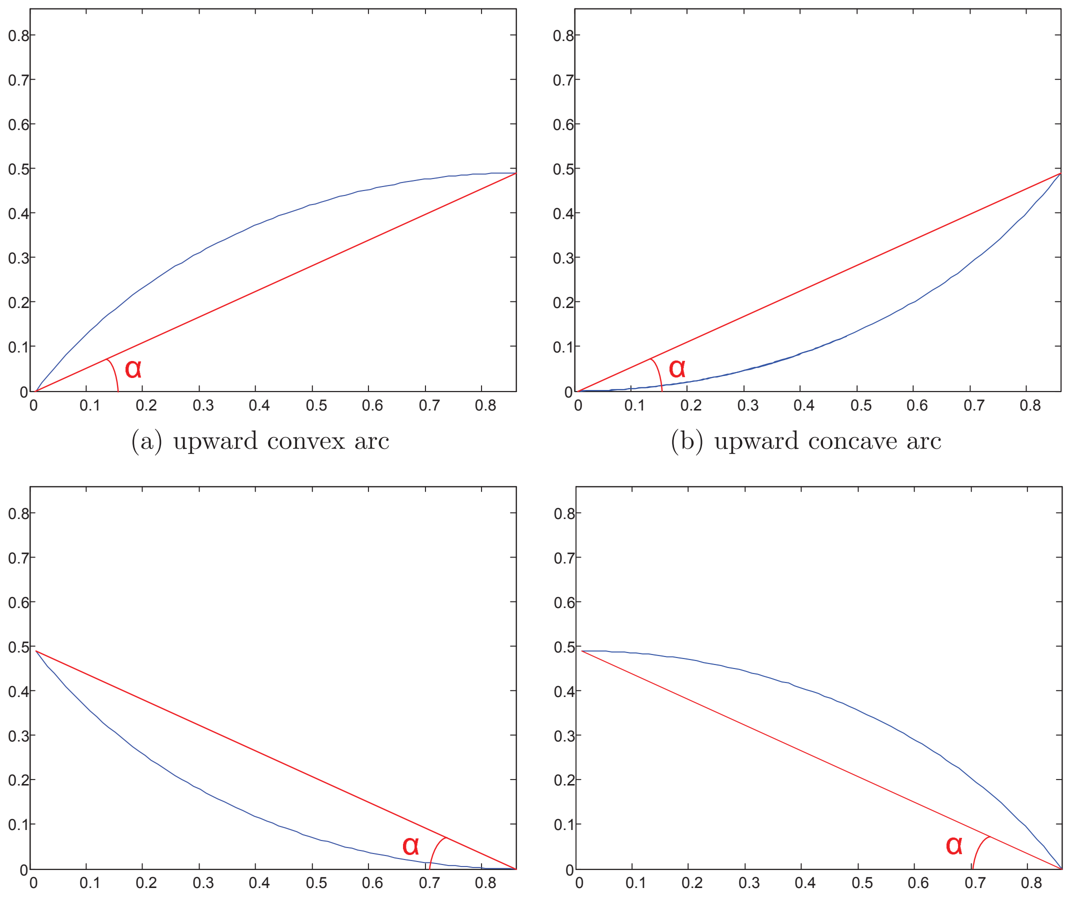
$$\begin{aligned} &\mathcal{C}_c(X,Y) = \{y_t, y_{t+1}, \ldots, y_{t+N}\} \\ &.subject\ to: d_t < \xi, t = 1, \ldots, T \end{aligned} \tag{3.7}$$

A subperiod time series of Y is judged correlated to X, only if its distance to the channel pattern p^* is less than the distance threshold ξ. ξ is often fixed based on the average distance between the selected channel pattern p^* and the observed time series X,

$$\xi = \frac{\sum_{t=1}^{N} \|p_t^* - x_t\|}{N}. \tag{3.8}$$

Alternatively, the correlation information is also extractable through a ransack searching within period T using the minimum distance as the distance threshold.

$$\begin{aligned} \mathcal{C}_c(X,Y) &= \{y_t, y_{t+1}, \ldots, y_{t+N}\} \\ subject\ to: d_t &\leq \min_{t\in[1,T]} \frac{\sum_{t=1}^{N} \|p_t^* - y_t\|}{N}. \end{aligned} \tag{3.9}$$



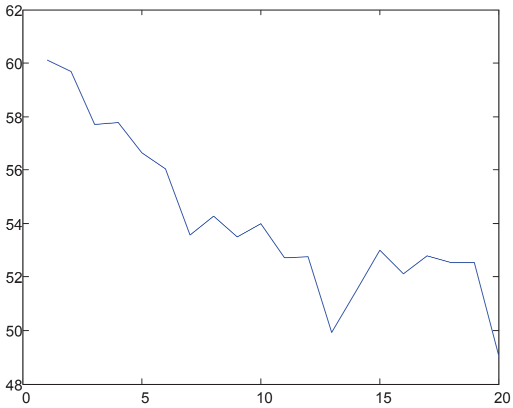
(a) upward convex arc (b) upward concave arc

(c) downward concave arc (d) downward convex arc

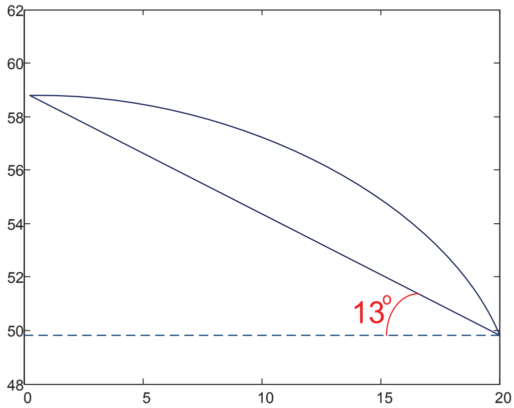
Figure 3.3: 4 types arc ruler corresponding to 4 trend patterns shown in Figure.3.2

3.1.2 An Example of Channel Correlation Analysis

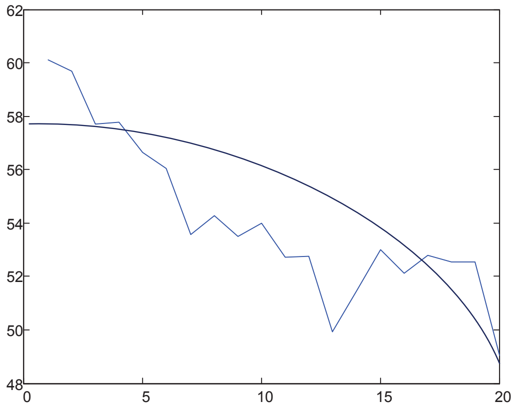
As an example, a real time series dataset is selected from the NZD to JPY forex market for 20 trading days between 04 Nov, 2008 and 01 Dec,2008, as shown in Figure.3.4a. By Eq.(3.5), an upward convex arc with $\angle 13^o$ and $\xi_t = 0.0321$ (Figure.3.4b) are foamed out. We therefore use the arc as a ruler to measure the same length (period) of data in the whole range of historical time series. As a result, we obtain a set periods of time series data, which are judged as the correlation knowledge to the observed time series. Figure.3.5 shows 4 periods of time series as the example of extracted correlation knowledge.



(a) Closing price from 04 Nov, 2008 to 01 Dec, 2008



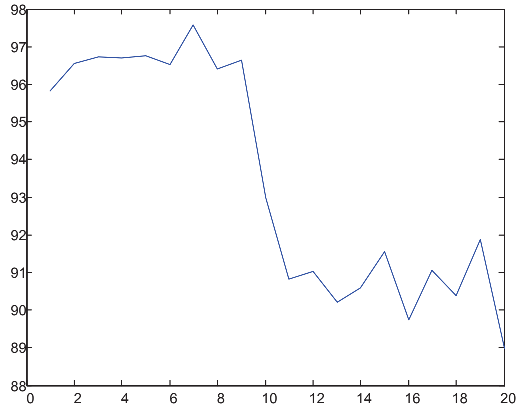
(b) The modeled arc ruler on Figure.3.4a time series



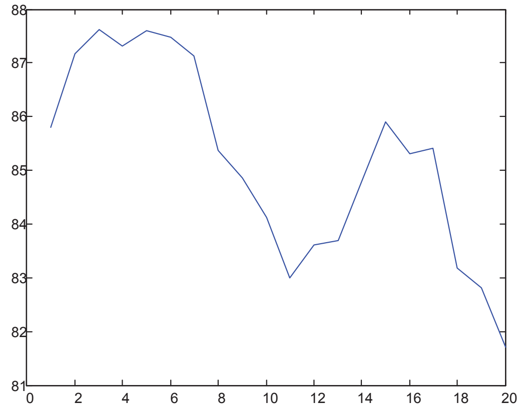
(c) Measuring correlation by the ruler in Figure. 3.4b

Figure 3.4: The procedure of the channel method implementation for correlation analysis.

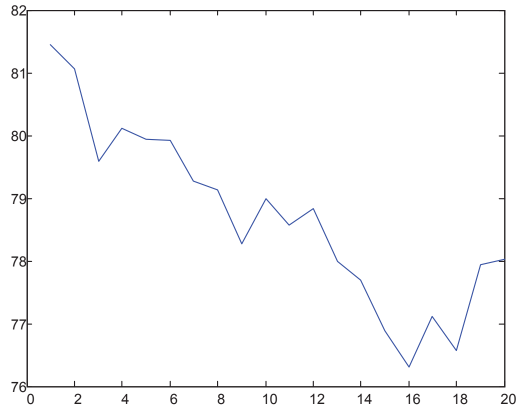
The channel method approximates the trend of time series by a graphical channel,



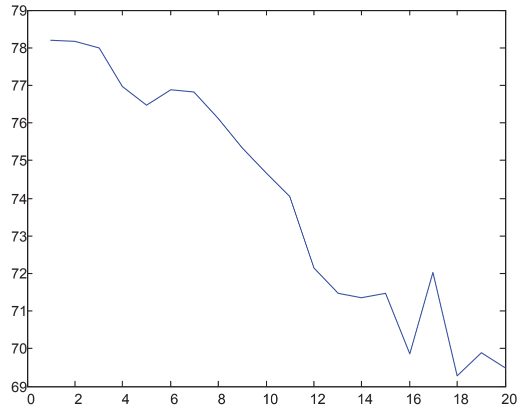
(a) Closing price from 13 Jul, 2007 to 09 Aug, 2007



(b) Closing price from 20 Dec, 2007 to 17 Jan, 2008



(c) Closing price from 22 Jul, 2008 to 18 Aug, 2008



(d) Closing price from 20 Aug, 2008 to 17 Sep, 2008

Figure 3.5: 4 periods of time series, the example of correlation knowledge extracted using the proposed channel method.

and evaluates the trend similarity between two time series by calculating the distance to the graphical channel, rather than the distance to the observed time series. This makes the channel method suitable for the zig zag path market data analysis, because it extracts correlations from time series with similar general variation trends, in spite of their different zig-zag shapes.

However, it is worth noting that the threshold ξ_t varies for different data analysis. Single channel correlation extraction confronts difficulties in some case, such as: if $\xi_t \approx 0$, Eq.(3.9) gives often no correlation output; if ξ_t is set to a large value, then Eq.(3.9) is likely to present correlation that includes noise.

3.2 The Weighted Pearson's Correlation

To overcome the drawbacks of the channel method, we develop here a weight Person's correlation analysis, extending standard Person's correlation for correlation knowledge extraction with minimized noise and in the mean time minimized useful information lost.

3.2.1 Weighted Person's Correlation Extraction

According to Eq.(2.13), the implementation correlation extraction by Pearson's correlation actually is subjected to the p condition. For forex time series analysis, two similar time series gives often have a high p value because of the time point mismatches between two variables. From this, it follows that a high correlation degree is normally associated with high p value. This implies that significant correlation information is likely to be missed due to the high p value, and therefore the Pearson's correlation is ineffective for extracting useful information for forex market analysis.

For a feasible and effective correlation extraction through Pearson's correlation analysis, the following weighted Pearson's correlation is proposed. The method sets a hyperplane on both sides of the perfect positive correlation ($Y = X$), so that a similarity margin a is allocated to exclude noisy correlation data, and Eq.(2.10) can be presented as,

$$\begin{aligned} r = & \quad \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{(\sum X^2 - \frac{\sum X^2}{N})(\sum Y^2 - \frac{\sum Y^2}{N})}} \\ subject\ to: & \quad Y - X - a < 0, and\, Y - X + a > 0 \end{aligned} \tag{3.10}$$

where a is the weight identifying the relatedness to the target analysis $Y = f(X)$. Figure.3.6 gives an illustration of weighted Pearson's correlation analysis.

To discover correlation from Y, the distance from point (x_t, y_t) to the perfect Pearson's correlation line $Y = X$ is estimated as in Figure.3.7 for every time point t:

$$d_t = \frac{\sum_{t \in N} \frac{|x_t + y_t|}{\sqrt{2}}}{N}. \tag{3.11}$$

Then, similar to the previously described channel method, correlation data is ex-

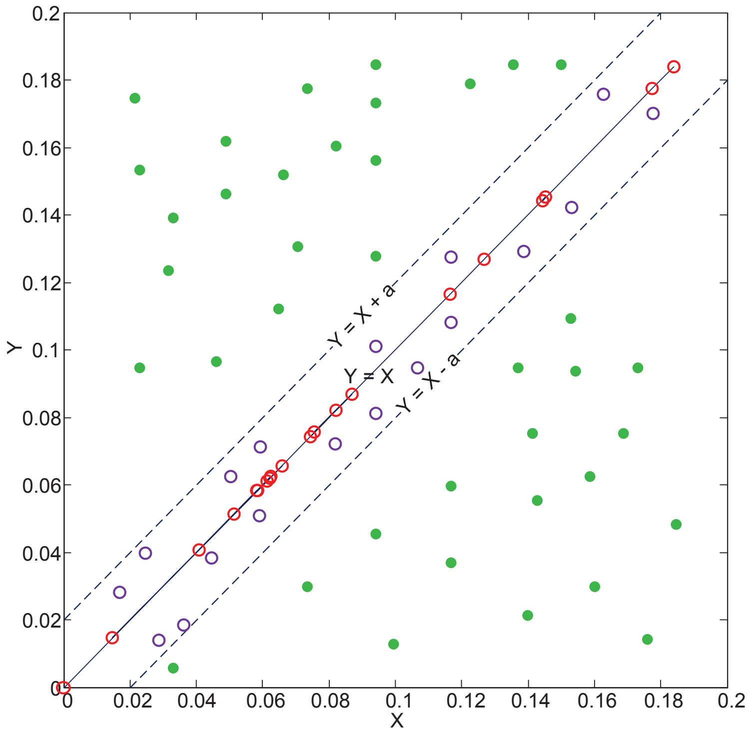


Figure 3.6: The illustration of weighted Pearson's correlation. A perfect positive correlation distributed on Person's correlation theory. The data used is the closing price from 04 Nov, 2008 to 01 Dec, 2008 and o is represents the closing price on each trading day.

tracted from Y through a shifting distance comparison as,

$$\mathcal{C}_p(X,Y) = \{y_t, y_{t+1}, \ldots, y_N\}$$
$$, subject\ to : d_t < a, t = 1, \ldots, T \tag{3.12}$$

where a is the weight identifying the width of correlation margin.

3.2.2 An Example of Weighted Pearson's Correlation Analysis

As an example, a real time series dataset is selected from NZD to JPY forex market for 20 trading days from 04 Nov, 2008 to 01 Dec, 2008, as shown in Figure.3.8a. The weighted Pearson's correlation brings it to a perfect positive linear correlation axis with $a = 0.06$ in Figure.3.8b. We use the model for correlation data extraction from

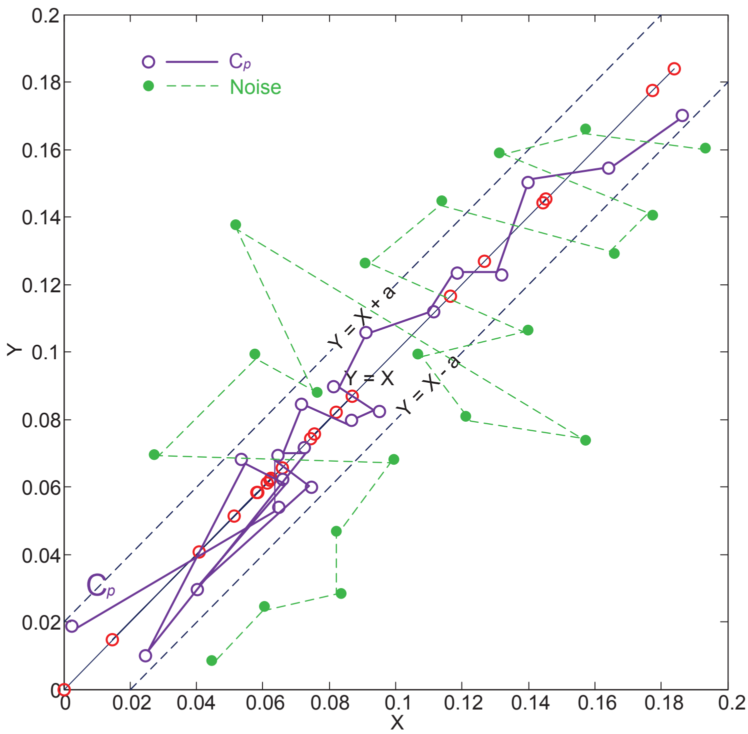


Figure 3.7: The illustration of weighted Pearson's correlation extraction

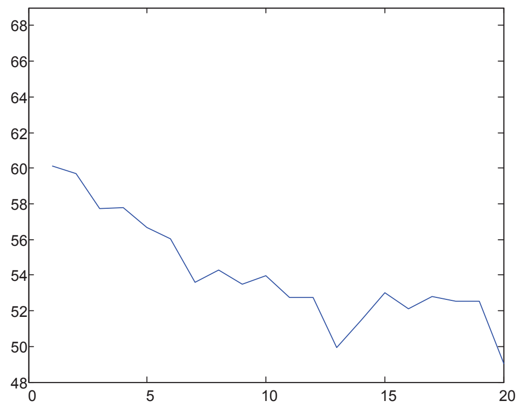
historical data, and obtain a set of time series matching this model (Figure.3.8c). Figure.3.9 presents four examples of correlated time series extracted by the weighted Pearson's correlation analysis.

The difficulty of the weighted Pearson's correlation when analyzing such zig zag path time series is presented here. If $a \approx 0$ a, Eq.(3.12) gives no correlation data out; and if a is set with to a large value, then Eq.(3.12) is likely to present correlation that includes noise.

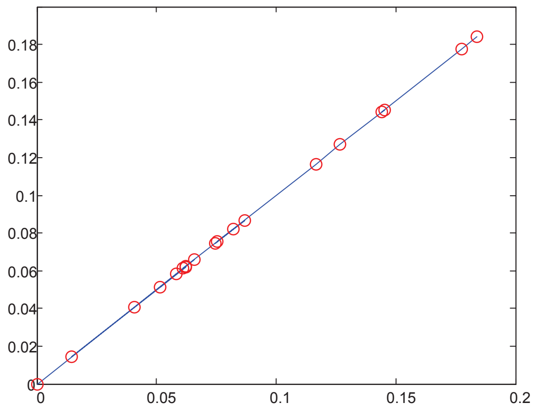
3.3 Correlation Synthesis for Forex Market Analysis

As discussed above, both the channel method and the weighted Pearson's correlation method have certain limitations. However, the combination of channel and weighted Pearson's analysis provide an optimal correlation extraction.

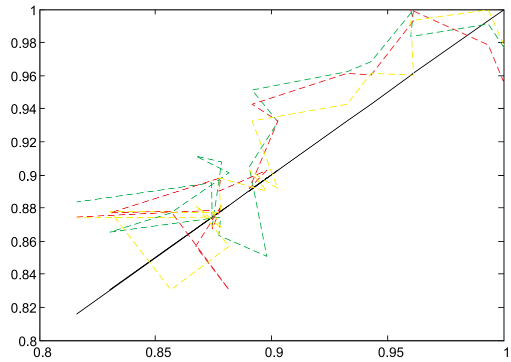
Technically, the channel correlation has the threshold ξ_t in Eq.(3.8) determined



(a) Closing price from 04 Nov, 2008 to 01 Dec, 2008



(b) The weighted Pearson's correlation model on the data of Figure. 3.8(a)

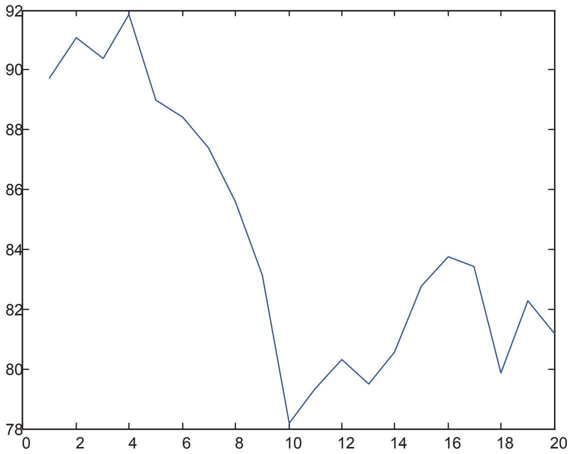


(c) The definition in weighted Pearson's correlation

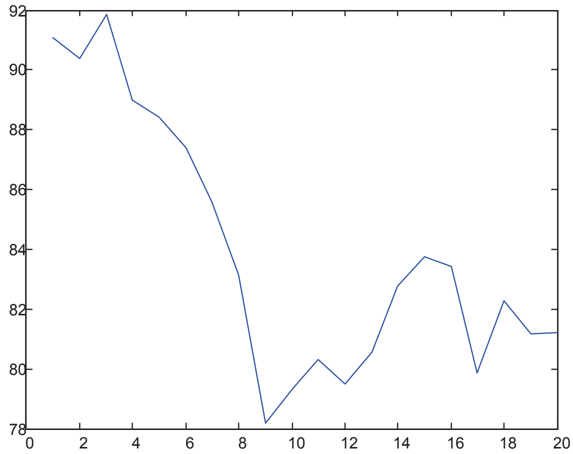
Figure 3.8: The procedure of the proposed weighted Pearson's correlation analysis.

by the average distance from the arc to the observed time series. A very small ξ_t often is given when the arc has a good match with the observed time series, which causes no correlation output from the channel method. In this case, the weighted Pearson method is always able to extract correlation within a proper correlation margin a. Also, when the observed time series is shaped as a zig zag path, no correlation output the weighted Pearson's method does not produce correlation due to the big mismatches caused by zig zag path. In this case, the channel method is able to trace trends similarity, as Eq.(3.8) produces surely a big ξ_t value on the zig zag path.

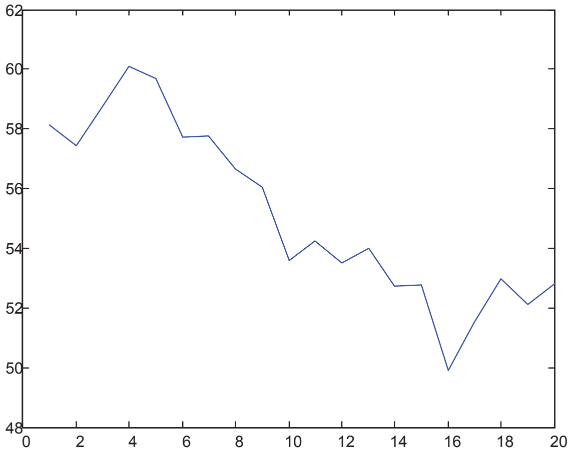
On the other hand, the combination of channel and weighted Pearson's correlation methods takes into account the balancing tradeoff between trend similarity and



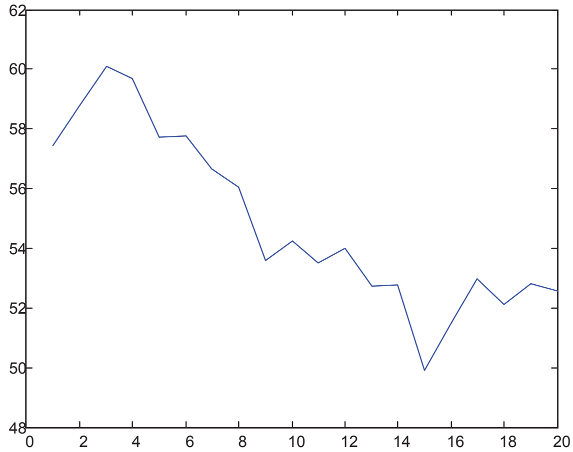
(a) Closing price from 03 Aug, 2007 to 30 Aug, 2007



(b) Closing price from 06 Aug, 2007 to 31 Aug, 2007



(c) Closing price from 31 Oct, 2008 to 27 Nov, 2008



(d) Closing price from 31 Oct, 2008 to 27 Nov, 2008

Figure 3.9: Four examples of correlation time series extracted by the weighted Pearson's correlation analysis

distance similarity for correlation knowledge extraction. The obtained correlation data is expected to have more weightage than the data from any one of the two methods. Thus, significant correlation knowledge is composed as,

$$\mathcal{C}(X,Y) = \mathcal{C}_c(X,Y) \cup \mathcal{C}_p(X,Y). \tag{3.13}$$

In forex market analysis, we consider 3 aspects of correlation knowledge: (1) historical correlation of the observed time series to previous market data, called historical correlation; (2) the correlation to relevant currencies, called cross-currency correlation; and (3) the correlation to macroeconomic variables, called macro-correlation. which is,

$$\mathcal{C}^* = \{\mathcal{C}(X, Y_i^{<h>})\} \bigcup \{\mathcal{C}(X, Y_j^{<c>})\} \bigcup \{\mathcal{C}(X, Y_k^{<m>})\}, \tag{3.14}$$

where Y_i^h, Y_j^c, Y_k^m is an individual time series from historical market, correlated currency exchange rates, and microeconomic variables, respectively.

Chapter 4

Correlation Knowledge Verification

Once correlation information and knowledge have been extracted, they have to be evaluated. In this chapter, we study machine learning technologies for correlation knowledge verification.

The evaluation is based on the theory of time series prediction. In machine learning, artificial neural networks and support vector machine regression are the most popular tools. The chapter introduces both methods and explains the reason why we choose support vector machine regression.

4.1 Time Series Prediction

To inspect the validity of extracted correlation knowledge, a straightforward approach is to use the obtained correlation knowledge directly for market trend analysis, as valuable correlation information is expected to contribute positively to the enhancement of forex time series prediction.

A forex time series prediction is modeled based on current and past market data to predict the future value (Sapankevych & Sankar, 2009) as: $\hat{x}(t + \Delta_t) = f(x(t - a), x(t - b), x(t - c), ...)$, where $\hat{x}$ is the predicted value of a discrete time series x; $f(x)$ is the perdition function which predicts an unbiased and consistent value of x at a future time point $t + \Delta_t$.

4.1.1 Artificial Neural Networks

Artificial Neural Network (ANN) are designed after the biological neurons and are also known as "Neural Network" (NN). They can be said as the mathematical or computational model that simulates the biological neurons functional aspects in the neural networks.

There are interconnected groups of artificial neurons and process information, which use the connectionist approach for computation inside an ANN model. Also, ANN can be seen as an adaptive system since it undergoes structural changes based on incoming information that traverses through the network during the learning phase. Due to the introduction of activation / transfer function, it can be seen as a non-linear data modeling tool and can be used to represent complex relationships amongst the input and output signals (information) or to find particular patterns or special events in a dataset (Mitchell, 1999).

Artificial Neural Networks Time Series Prediction

The structure of an ANN as shown in Figure.4.1, is an interconnected group of nodes. ANN time series prediction uses a group of interconnected functions to calculate $\hat{x}(t+\Delta_t)$ by analyzing x within t period. Suppose an ANN has n composition

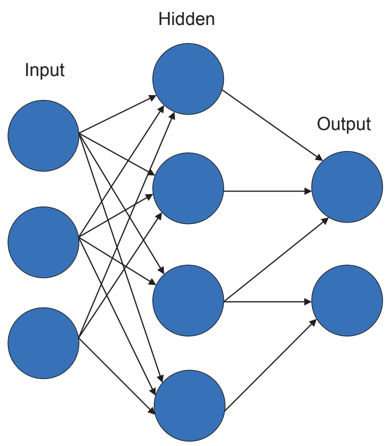


Figure 4.1: A neural network is an interconnected group of nodes

functions $((g_1(x), g_2(x), ...g_n(x)))$. The ANN function $f(x)$ is defined over a number of functions $f(x) = (g_1(x), g_2(x), ...g_n(x))$. The commonly used type of composition is the nonlinear weighted sum shown in Eq.(4.1)

$$f(x) = K\left(\sum_i w_i g_i(x)\right), \tag{4.1}$$

where K denotes a predefined function, for example a hyperbolic tangent function. For the sake of convenience, the set of functions g_i can be considered as a vector $g = (g_1, g_2, \ldots, g_n)$. Therefore, an ANN can be described as a graph composed by a set of 2-dimensional vectors and 3-dimensional vectors as Figure.4.2.

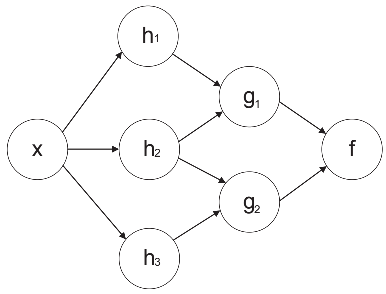


Figure 4.2: The input x is transformed into a 3-dimensional vector h, which is then transformed into a 2-dimensional vector g, which is finally transformed into f

For ANN optimization, ANN learning phases are required to use a set of observations to find $f^* \in F$, based on which the ANN can produce some meaningful results. F is a group of functions. One of the significant schemes used in machine learning is the concept of *Cost function* $C : F \rightarrow \mathbb{R}$, where the set of functions F should reach the minimum risk value $\mathbb{R}$. Through the optimization, the learning method minimizes the risk value. The cost function C shows how far away it is from a particular solution. Since the risk value $\mathbb{R}$ should be as less as possible, the learning algorithm explores the solution space to achieve the least possible cost. To achieve the optimal solution f^*, cost function can bee calculated as $C(f^*) \leq C(f) \forall f \in F$. However, in real practice the real obtained solution never reaches the optimal solution cost, but is only able to find a solution that falls closest to the optimal solutions cost.

To train a ANN model, historical data is often used for cost function C estimation. For example, given data $\mathcal{D}$, let the data pairs derived from it be (x, y), here the problem lies in building a model f, such that it minimizes $C = E\left[(f(x) - y)^2\right]$. However in practice, the least minimization of $C = E\left[(f(x) - y)^2\right]$ can be reached due to the availability of only N samples obtained from $\mathcal{D}$. Therefore, minimization can be carried out simply on limited data samples instead of the entire data set.

The ANN learning can be categorized into three major learning paradigms, namely supervised learning, unsupervised learning and reinforcement learning. Each suits a particular type of learning task. In the following, a brief explanation is given on the three learning paradigms.

Supervised learning: Consider a set of pairs (x, y), $x \in X$, $y \in Y$, supervised learning is to find a $f : X \to Y$ function that matches the given examples(Shubhabrata & Malay, 2004). In a nutshell, the mapping needs to be obtained from the given data. Since the mapping is based on prior knowledge concerning the problem domain, cost function C is utilized to find the difference between our mapping and the data, for example, Mean Squared Error (MSE) and Multi-Layer Perceptrons (MLP) are two popular cost functions for supervised learning neural network construction. MSE minimizes the average squared error between ANN output $f(x)$ and target value y for the observed data samples (x, y), $x \in X$, $y \in Y$; and MLP uses gradient descent for MSE minimizing.

Supervised learning is used for reoccurring patterns. It can be used for pattern recognition task such as classification and regression. It is also employed for sequential data such as speech and gesture recognition.

Unsupervised learning: Different to supervised learning, unsupervised learning performs learning based on priori assumptions (Agatonovic-Kustrin & Beresford, 2000), thus does not require target data information y. This leads to that the minimization of cost function is task and priori assumptions dependent. For instance, suppose that a is the output of $f(x)$ and C is calculated as $C = E[(x - f(x))^2]$ from priori assumptions, then the minimized C is found when a equals to the mean of the data. However, the cost function could be associated with the mutual information or posterior probability for some applications. In these cases, the cost function will be maximized instead of minimized by learning just the priori assumptions.

Unsupervised learning is applicable to tasks involving clustering, statistical distributions estimation, compression and filtering.

Reinforcement learning: Reinforcement learning (RF) is a category of machine learning with the minimization of cost function dynamic over the time. RF corrects input/output pairs and optimal actions at each time point t, exploiting trade-off between reward and punishment (Kaelbling, Littman & Moore, 1996). The cost C is calculated by the data x_t and y_t at each time point t. During a long term learning, the dynamic cost C_t for each optimal action can be approximated by cumulation. As dealing with some complicated dataset, x sometime is not given, reinforcement learning is capable of generating a new observation x_t via the optimal action on minimizing dynamic cost C_t.

A popular RF modeling is based on Markov Decision Process (MDP) having states $s_1, ..., s_n \in S$ and actions $a_1, ..., a_m \in A$. A MDP includes probabilities of instantaneous cost distributions $P(c_t|s_t)$, observation distributions $P(x_t|s_t)$ and transition $P(s_t + 1|s_t, a_t)$. MDP produces a number of Markov Chains (MC) to connect each function in RF learning. The action policy of a given observation is discovered and the cost function is minimized by conducting MCs.

Reinforcement learning is often used in economics, game theory, control problems and other sequential decision making problems.

ANN Applications

Artificial Neural Network (ANN) have been popularly employed in forex market prediction for the past two decades and are still under development. A case study on Australian foreign exchange by Kamruzzaman and Sarker (2003) compares the performances of three ANN prediction models: standard backpropagation, scaled conjugate gradient and backpropagation with Baysian regularization. Auto-Regressive Integrated Moving Average technique (ARIMA) has been used in the study for predicting six different currencies against Australian dollar. The results are evaluated by Normalized Mean Square Error (NMSE), Mean Absolute Error (MAE), Directional Symmetry (DS), Correct Up trend (CU) and Correct Down trend (CD).

For both testing periods (35 weeks and 65 weeks), ANN model shows better performances than ARIMA. Following the development of ANN, feed forward neural networks were recently considered for flexible non-linear modeling of censored survival data through the generalization of both discrete and continuous time models. A 1998 study (Elia, Patrizia, Luigi & Ettore, 1998) reviews feed forward neural networks in theory, and shows that it is a more efficient prediction technology for forex market than other time series prediction method. The study reported by (Emam,

2008) tested an optimal ANN technology to predict the foreign exchange rate between Japanese Yen and US dollar from 20 Aug, 2006 to 20 Sep, 2006. The chosen models are Moving Average(MA 10, MA 20, MA50) and RSI. The results are evaluated by Mean Square Error (MSE) and show that the optimal ANN technology performs better than a previously suggested ANN model (feed forward neural networks).

4.1.2 Support Vector Machine

Support Vector Regression (SVR) is the application of Support Vector Machines (SVM) (Vapnik, 1999; Drucker, Burges, Kaufman, Smola & Vapnik, 1997; Scholkopf, Burges & Smola, 1999) to general regression analysis. The SVR departs from more traditional time series prediction methodologies in the strict sense where there is no "model" to make the prediction and depends only on the data from one domain.

Support Vector Regression Time Series Prediction

Given a forex time series $x(t)$ where t represents the time point. Suppose the present time point is N, then a prediction x for $t > N$ is computed over the training data $\mathcal{X}(t) = \{x(1), x(2), ..., x(N)\}$. Thus, the goal is to find a function $f(x)$ that matches the actually obtained targets $x(t)$ of next time point for all the training data.

According to (Vapnik, 1999), a non-linear estimation of $f(x)$ is computed in Eq.(4.2)

$$f(x) = (w \cdot \phi(x)) + b, \tag{4.2}$$

where "$\cdot$" means a dot product (Takeshi, 2005) and $\phi(x)$ refers to the kernel function $k(x, x^{'}) = \left\langle \Phi(x), \Phi(x^{'}) \right\rangle$, which enables performing a linear regression in higher dimensional feature space.

To find an optimal set of parameters: weight w and threshold b. Firstly, the weights is flatted by the Euclidean norm ($\|w\|^2$). Secondly, the empirical risk (error) is generated by the estimation process of the value. Thus, the overall goal is the minimization of the regularized risk $R_{reg}(f)$,

$$R_{reg}(f) = R_{emp}(f) + \frac{\lambda}{2}\|w\|^2, \tag{4.3}$$

where $R_{emp}(f)$ is the empirical risk

$$R_{emp}(f) = \frac{1}{N}\sum_{i=0}^{N-1} L(x(i), y(i), f(x(i), w)), \tag{4.4}$$

where i is an index to discrete time points $t = \{0, 1, 2, ...N-1\}$ and $y(i)$ is the predicted value being sought. $L(.)$ is a "loss function" to be defined. λ is the capacity control factor, a scale factor regard as regularization constant which reduces "over-fitting" of data and minimizes negative effects of generation.

A quadratic programming problem is formed to solve the optimal weights and minimize the regularized risk using the general ϵ-insensitive loss function:

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} L(y(i), f(x(i), w)), \tag{4.5}$$

where

$$L(y(i), f(x(i), w)) = \begin{cases} 0 & \text{if } |y(i) - f(x(i), w)| \le \epsilon \\ |y(i) - f(x(i), w)| - \epsilon & \text{otherwise.} \end{cases} \tag{4.6}$$

$\mathcal{C}$ is a positive constant which includes the $(1/N)$ summation normalization factor and ϵ refers to the precision by which the function is to be approximated. They are both user defined constants and can be typically determined by cross validation tests.

Solving Eq.(4.6) is an exercise in convex optimization, thus it is easy to use Lagrange multipliers and form the dual optimization problem as:

$$\begin{array}{ll} \text{Maximize:} & -\frac{1}{2}\sum_{i,j=1}^{N}(a_i - a_i^*)(a_j - a_j^*)\langle x(i), x(j)\rangle \\ & -\epsilon\sum_{i=1}^{N}(a_i - a_i^*) + \sum_{i=1}^{N} y(i)(a_i - a_i^*) \\ \text{Subject to:} & \sum_{i-1}^{N}(a_i - a_i^*) = 0 : a_i, a_i^* \in [0, C] \end{array} \tag{4.7}$$

In this way, $f(x)$ is approximated as the sum of the optimal weights times the dot products between the data points as:

$$f(x) = \sum_{i=1}^{N}(a_i - a_i^*)\langle x, x(i)\rangle + b, \tag{4.8}$$

where those data points on or outside the ϵ tube with non-zero Lagrange multipliers

a are defined as support vectors.

To figure out the non-linear SVR regression, it is necessary to map the input space $x(i)$ into a (possibly) higher dimension feature space $\Phi(x(i))$. The solution of the SVR does not rely on the input data, a kernel function that satisfies Mercer's conditions can be written as:

$$k(x, x') = \left\langle \Phi(x), \Phi(x') \right\rangle, \tag{4.9}$$

which can be put back into Eq.(4.8) and the optimal weights w can be calculated in feature space in exactly the same fashion.

There are some well known kernel functions listed below (Table 4.1).

SVM Kernel	
Polynomial (homogeneous):	$k(x, \acute{x}) = (x \cdot \acute{x})^d$
Polynomial (inhomogeneous):	$k(x, \acute{x}) = (x \cdot \acute{x} + 1)^d$
Radial Basis Function (RBF):	$k(x, \acute{x}) = exp(-\gamma \|x - \acute{x})\|^2)$, for $\gamma > 0$
Gaussian Radial basis function:	$k(x, \acute{x}) = exp(-\frac{\|x-\acute{x})\|^2}{2\sigma^2})$
Sigmoid:	$k(x, \acute{x}) = tanh(\kappa x \cdot \acute{x} + c)$ for some (not every) $\kappa > 0$ and $c < 0$

Table 4.1: SVM Kernels

SVR Applications

SVM is a popular financial time series prediction after used instead of ANN currently. Over fitting problem is a main problem in Neural Networks. SVR can easily fix over fitting problem by RBF kernel (Trafalis & Ince, 2000; Tay & Cao, 2001a). Some researchers (Ajith Abraham, 2004; Ongsritrakul & Soonthornphisaj, 2003; Tay & Cao, 2001b) also compare SVM with ANN, Difference Boosting Neural Network (DBNN), decision trees and multi-layer-perception (MLP). SVM retrieves the best results on Root Mean Squared Error (RMSE), Mean Absolute deviation (MAD) and Mean Absolute Percentage Error (MAPE).

SVR has been widely used on financial data time series prediction, for instance, electricity price forecasting, credit rating analysis, auto instance market prediction, financial analysis and production value prediction of the Taiwanese machinery industry. Sansom, Downs and Saha (2003) employed SVR to predict Australian national Electricity Market Prices 7 days ahead. They gained a better result than when using ANN. Huang, Chen, Hsu, Chen and Wu (2004) reported corporate credit rating prediction and Hur and Lim (2005) predicted customer "churn" ratio for auto insurance market by SVM. Their result was much better than when using ANN. Bose and Raktim (2005) analyzed the fate of failed Dotcoms with SVM. It revealed that SVM was easier to classify a surveyed dotcom company than a failed one. Pai and Lin (2005) used SVR Gaussian RBF kernel to predict the one-step ahead production values of the Taiwanese machinery industry. The MAE, MAPE, RMSE and NMSE of their result are all better than using general regression neural network.

4.2 Correlation Aided SVR Time Series Prediction

To evaluate the validity of correlation knowledge, we consider a comparison of original SVR time series prediction against a correlation-aided SVR (cSVR) time series prediction. The cSVR employs correlation data $\mathcal{C}$ in addition to the observed time series data X for regression. Let $\mathcal{C}$ be the correlation data to the observed time series X, then Eq.(4.2) is extended for cSVR as:

$$f(x) = (w \cdot \phi([x\,\mathcal{C}])) + b. \tag{4.10}$$

Chapter 5

Experiments on Forex Time Series Prediction

In our experiment, a number of tests have been performed with SVR. Five exchange rates have been individually tested to evaluate the effectiveness of the extracted correlation. This chapter reports the forex data used for experiments, the steps of time series prediction experiment, as well as the comparison experimental results.

5.1 Forex data

For testing the utilization of $\mathcal{C}$ on real forex time series forecasting, we examined the correlation data $\mathcal{C}$ with a RBF kernelled SVR on five real futures contracts. The five real futures contracts are collected from the exchange rate NZD - AUD, NZD - EUD, NZD - GBP, NZD - JPY and NZD - USD. Their corresponding time periods used are listed in Table 5.1, and the daily closing prices are used as the data sets. The macroeconomic data is also employed into this study. Due to the shortage of macroeconomic reports on daily basis from most countries, the presented study had five sets of stock market data from each observed country as assistant analysis data. Table 5.2 gives the time periods for these five sets of stock market data.

5.2 Experimental setup

The proposed CSM is implemented in MATLAB version 7.6.0, on a 1.86Hz Intel Core 2 machine with 2GB RAM. In this experiment, we use correlation information

Names:	Time periods
NZD/AUD	01/01/2007 - 31/12/2008
NZD/EUD	01/01/2007 - 31/12/2008
NZD/GBP	01/01/2007 - 31/12/2008
NZD/JPY	01/01/2007 - 31/12/2008
NZD/USD	01/01/2007 - 31/12/2008

Table 5.1: Five futures contracts

Names:	Time periods
NZX 50	01/01/2007 - 31/12/2008
S_P_ASX 200	01/01/2007 - 31/12/2008
ftse100	01/01/2007 - 31/12/2008
nikkei255	01/01/2007 - 31/12/2008
NYSE	01/01/2007 - 31/12/2008

Table 5.2: Five assistant analysis data

extracted from CSM and SVR RBF kernel. Channel method's parameter ξ_t is automatically selected; weighted Pearson's correlation's parameter a is set to 0.07, and the SVR RBF parameter γ is set to 250. The regression period of time series N is generally determined by traders' experience. In our experiment, N is fixed to 20 by a cross-validation prediction tests on NZD/AUD within 2006 (Figure.5.1).

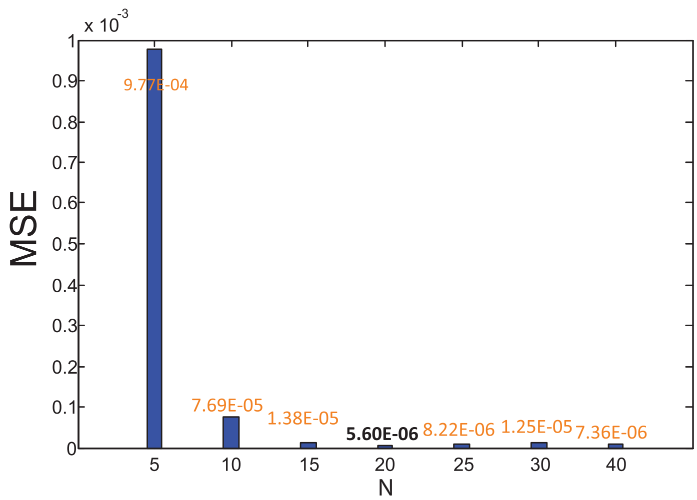


Figure 5.1: The MSE from prediction for NZD/AUD in 2006 by different length of time series N. The best result is when $N = 20$.

To exhibit the advantages of our method, we set a reliable prediction performance evaluation by means of the directional symmetry (DS), mean squared error(MSE), root mean squared error (RMSE), normalized mean square error (NMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE).

5.3 Experimental Results

Table 5.3, 5.4, 5.5, 5.6 and 5.7 show the results of forex time series prediction from 2 Jan, 2007 to 31 Dec, 2008 for five currency pairs. As seen from the tables, the cSVR in clearly shows a more advanced capability than SVR on the forex time series prediction in terms of MSE, RMSE, NMSE, MAE and MAPE. On DS, although the cSVR does not outperform SVR, there is particularly difference between the DS of SVR and cSVR.

Among the 5 currency pairs, it is worth noting that the most obvious evidence on cSVR is shown in NZD/JPY prediction. As can be observed in table 5.6, the MSE produced by cSVR is 3 times smaller than that produced by SVR in both 2007 and 2008. RMSE in cSVR prediction for 2007 is about 8 times smaller than that of SVR. The NMSE of cSVR is forty times smaller in 2007, four times in 2008 smaller than that of SVR prediction. Also for MAE and MAPE, the cSVR is giving significantly smaller errors than those from SVR.

NZD/AUD						
	DS	MSE	RMSE	NMSE	MAE	MAPE
SVR						
2007	**55.02%**	3.1903e-005	0.0056	2.0218e-006	0.0037	2.3705e-005
2008	45.28%	5.4424e-005	0.0074	3.4149e-006	0.0050	6.9396e-006
cSVR						
2007	53.41%	**1.5189e-005**	**0.0039**	**9.6063e-007**	**7.1164e-004**	**1.9383e-006**
2008	**46.40%**	**6.5802e-006**	**0.0026**	**4.1534e-007**	**4.8927e-004**	**5.5700e-007**

Table 5.3: *Training* 4 Jan, 1999 - 30 Dec, 2005, *Validation* 3 Jan, 2006 - 29 Dec, 2006 in NZD/AUD, *Testing* 2 Jan, 2007 - 31 Dec, 2008.

For daily exchange rates forecast using SVR, Figure.5.2 and Figure.5.3 depict the diagrams showing the differences between the predicted and the actual time series of five contracts exchange rates for the period of 2007 and 2008, respectively. As seen from the diagrams, the fitness between the predicted prices and the actual prices is mismatched in the five future contracts prediction. Obvious gaps exist between the

NZD/EUD						
	DS	MSE	RMSE	NMSE	MAE	MAPE
SVR						
2007	52.61%	1.9504e-005	0.0044	1.2360e-006	0.0032	2.1510e-005
2008	45.28%	2.6656e-005	0.0052	1.6726e-006	0.0037	6.6451e-005
cSVR						
2007	**53.82%**	**9.1822e-008**	**3.0302e-004**	**5.8073e-009**	**1.8954e-004**	**1.1167e-006**
2008	**47.60%**	**1.2423e-006**	**0.0011**	**7.8413e-008**	**3.0918e-004**	**1.5854e-006**

Table 5.4: *Training* 4 Jan, 1999 - 30 Dec, 2005, *Validation* 3 Jan, 2006 - 29 Dec, 2006 in NZD/EUD, *Testing* 2 Jan, 2007 - 31 Dec, 2008.

NZD/GBP						
	DS	MSE	RMSE	NMSE	MAE	MAPE
SVR						
2007	**54.62%**	1.2368e-005	0.0035	7.8380e-007	0.0024	1.9437e-005
2008	**51.57%**	2.1833e-005	0.0047	1.3699e-006	0.0032	4.3566e-005
cSVR						
2007	53.82%	**2.4687e-006**	**0.0016**	**1.5614e-007**	**3.0980e-004**	**4.4691e-006**
2008	50.00%	**3.5398e-006**	**0.0019**	**2.2343e-007**	**2.8227e-004**	**1.4314e-006**

Table 5.5: *Training* 4 Jan, 1999 - 30 Dec, 2005, *Validation* 3 Jan, 2006 - 29 Dec, 2006 in NZD/GBP, *Testing* 2 Jan, 2007 - 31 Dec, 2008.

NZD/JPY						
	DS	MSE	RMSE	NMSE	MAE	MAPE
SVR						
2007	**57.83%**	1.3234	1.1504	0.0839	0.8159	6.4106e-006
2008	**46.46%**	2.0997	1.4490	0.1317	0.9733	9.2072e-005
cSVR						
2007	56.22%	**0.0334**	**0.1828**	**0.0021**	**0.0744**	**3.0934e-006**
2008	46.00%	**0.6052**	**0.7779**	**0.0382**	**0.2241**	**7.7199e-006**

Table 5.6: *Training* 4 Jan, 1999 - 30 Dec, 2005, *Validation* 3 Jan, 2006 - 29 Dec, 2006 in NZD/JPY, *Testing* 2 Jan, 2007 - 31 Dec, 2008.

two curves, indicating that the errors on MSE, RMSE, NMSE, MAE and MAPE are all at high level.

As a comparison, Figure.5.4 and 5.5 present the daily exchange rate forecast results from the cSVR. As seen, the prediction from cSVR is consistently better than the prediction from SVR for NZD/AUD in 2007, NZD/GBP in 2007, NZD/JPY in both 2007 and 2008, and NZD/USD in 2007. It is noticeable that those gaps occurring in SVR prediction either disappeared or are mostly reduced in the cSVR predic-

NZD/USD						
	DS	MSE	RMSE	NMSE	MAE	MAPE
SVR						
2007	55.42%	7.3222e-005	0.0086	4.6403e-006	0.0057	1.2783e-005
2008	**47.64%**	9.1267e-005	0.0096	5.7266e-006	0.0069	4.0427e-005
cSVR						
2007	**56.22%**	**3.0260e-005**	**0.0055**	**1.9138e-006**	**0.0012**	**1.7798e-006**
2008	46.00%	**7.7186e-006**	**0.0028**	**4.8719e-007**	**7.9949e-004**	**2.4130e-004**

Table 5.7: *Training* 4 Jan, 1999 - 30 Dec, 2005, *Validation* 3 Jan, 2006 - 29 Dec, 2006 in NZD/USD, *Testing* 2 Jan, 2007 - 31 Dec, 2008.

tion. A few downward/upward overfittings occur in the cSVR prediction, which leads to cSVR not performing as well as SVR for these time points prediction. For example, for the prediction during 07 to 11 Jun, 2007, cSVR in Figure.5.4a is seen suddenly losing accuracy, performing even worse than the SVR in Figure.5.2. The explanation for this could be that the correlation data might pose a trend different/conflicted to the state indicated in the observed time series, which eventually causes the overfitting of cSVR training. Nevertheless, the contribution of the extracted correlation knowledge to the forex market trend prediction is confirmed according to the statistics for the predictions within the whole 2007 and 2008 period.

5.4 Summary

In the experiment, cSVR has better performance than SVR on its own. Except DS, cSVR has got much smaller MSE, RMSE, NMSE, MAE and MAPE (Table.5.3 to 5.7). The results has also been displayed on Figure.5.2 to 5.5. The predicted curves in Figure.5.4 and 5.5 are much closer to actual curves than they are in Figure.5.2 and 5.3. However, the correlation data might pose a trend different/conflicted to the state indicated in the observed time series, which eventually causes the overfitting of cSVR training.

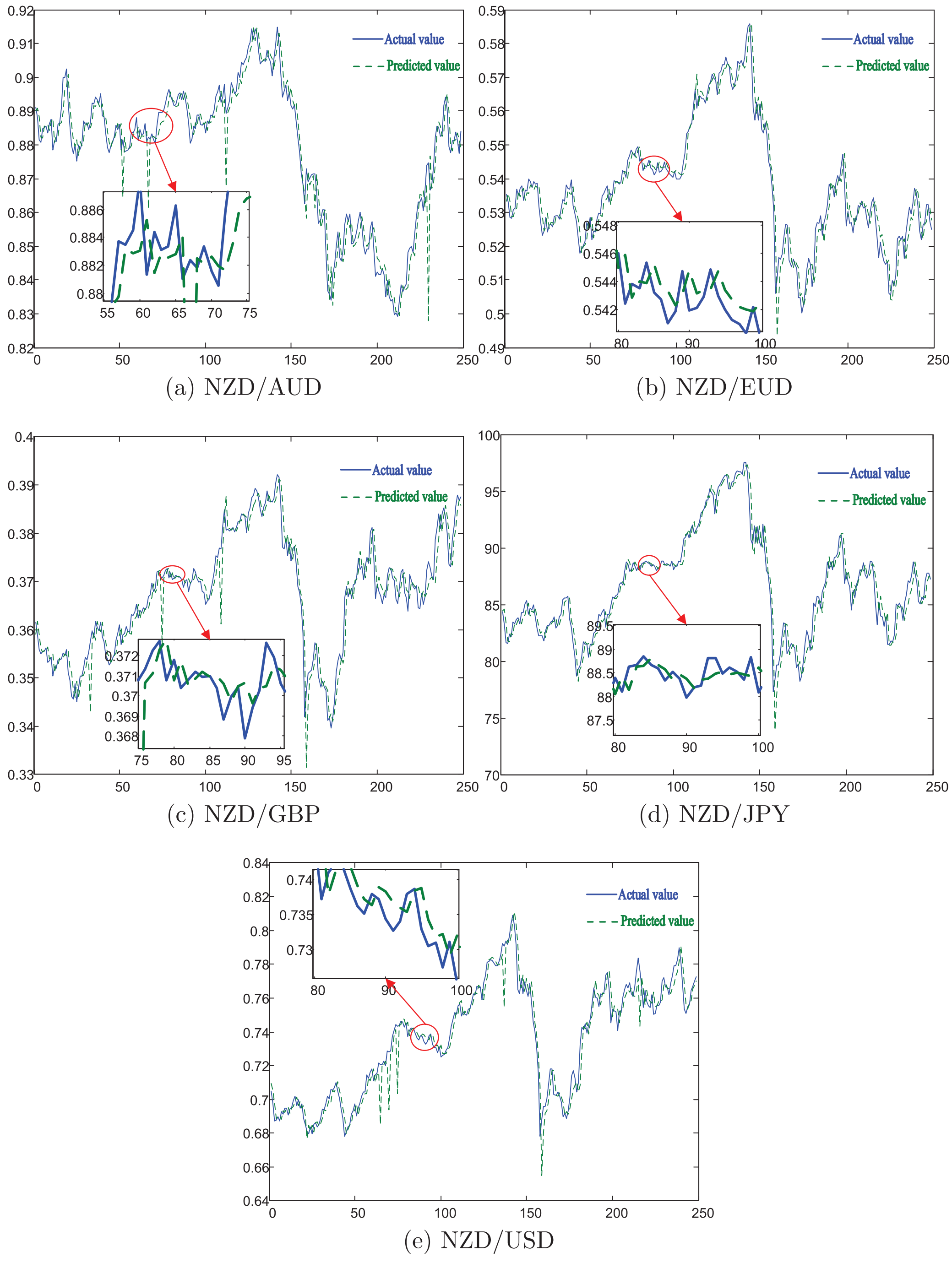
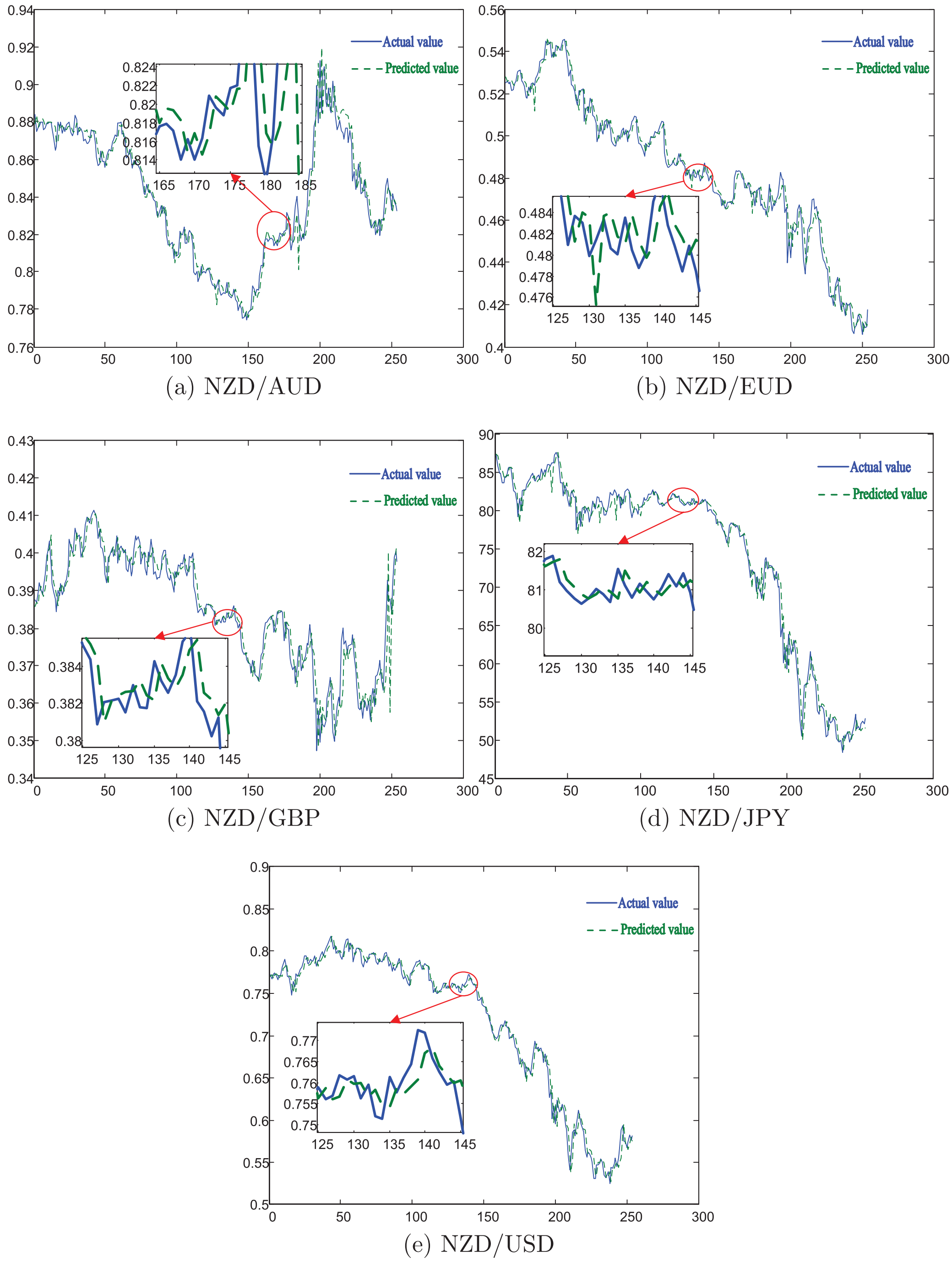
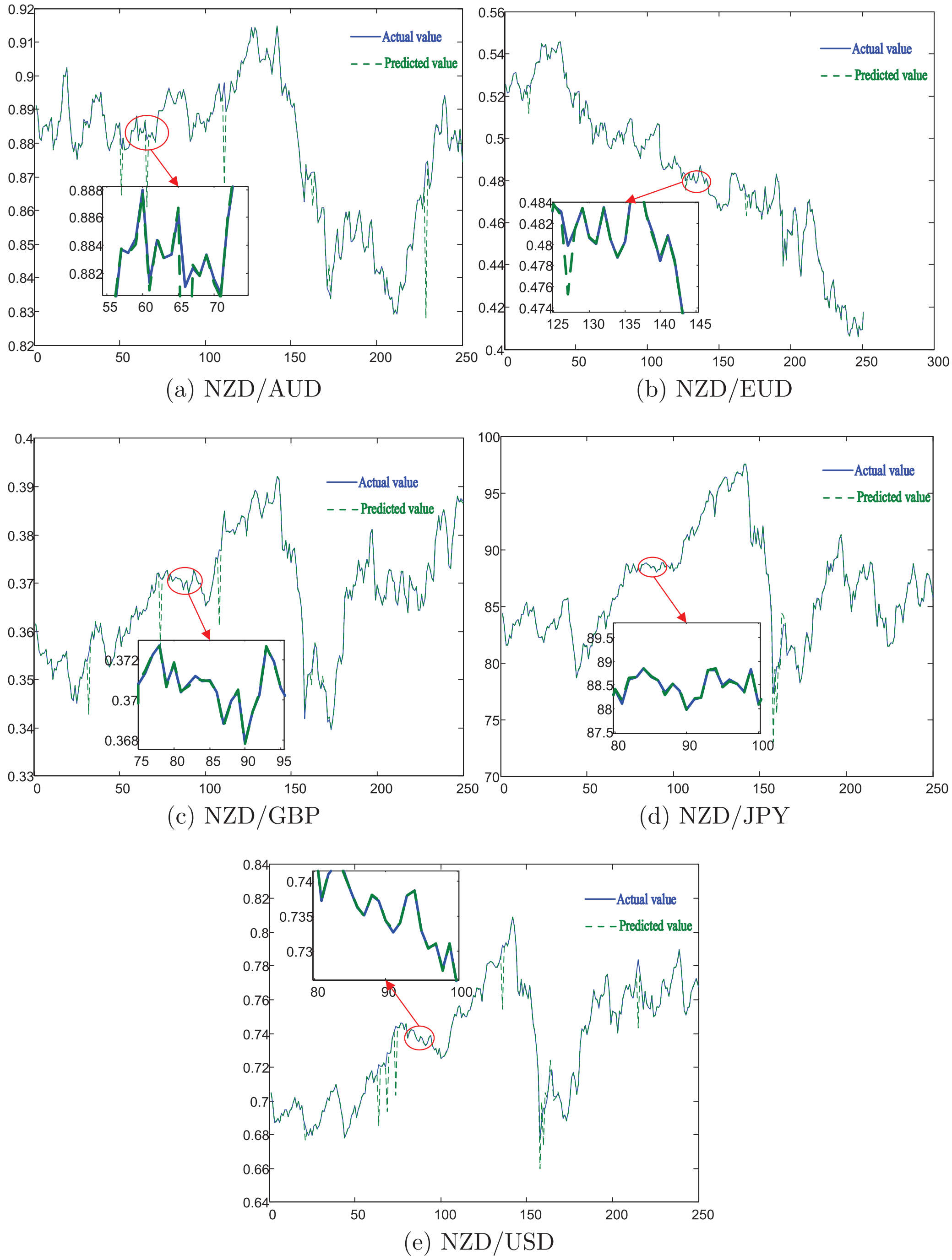


Figure 5.2: SVR prediction on daily exchange rate, January 2 007 to December 2007. Zoomed areas are 20 trading days randomly selected from this year.



(a) NZD/AUD (b) NZD/EUD (c) NZD/GBP (d) NZD/JPY (e) NZD/USD

Figure 5.3: SVR prediction on daily exchange rate, January 2008 to December 2008. Zoomed areas are 20 trading days randomly selected from this year.

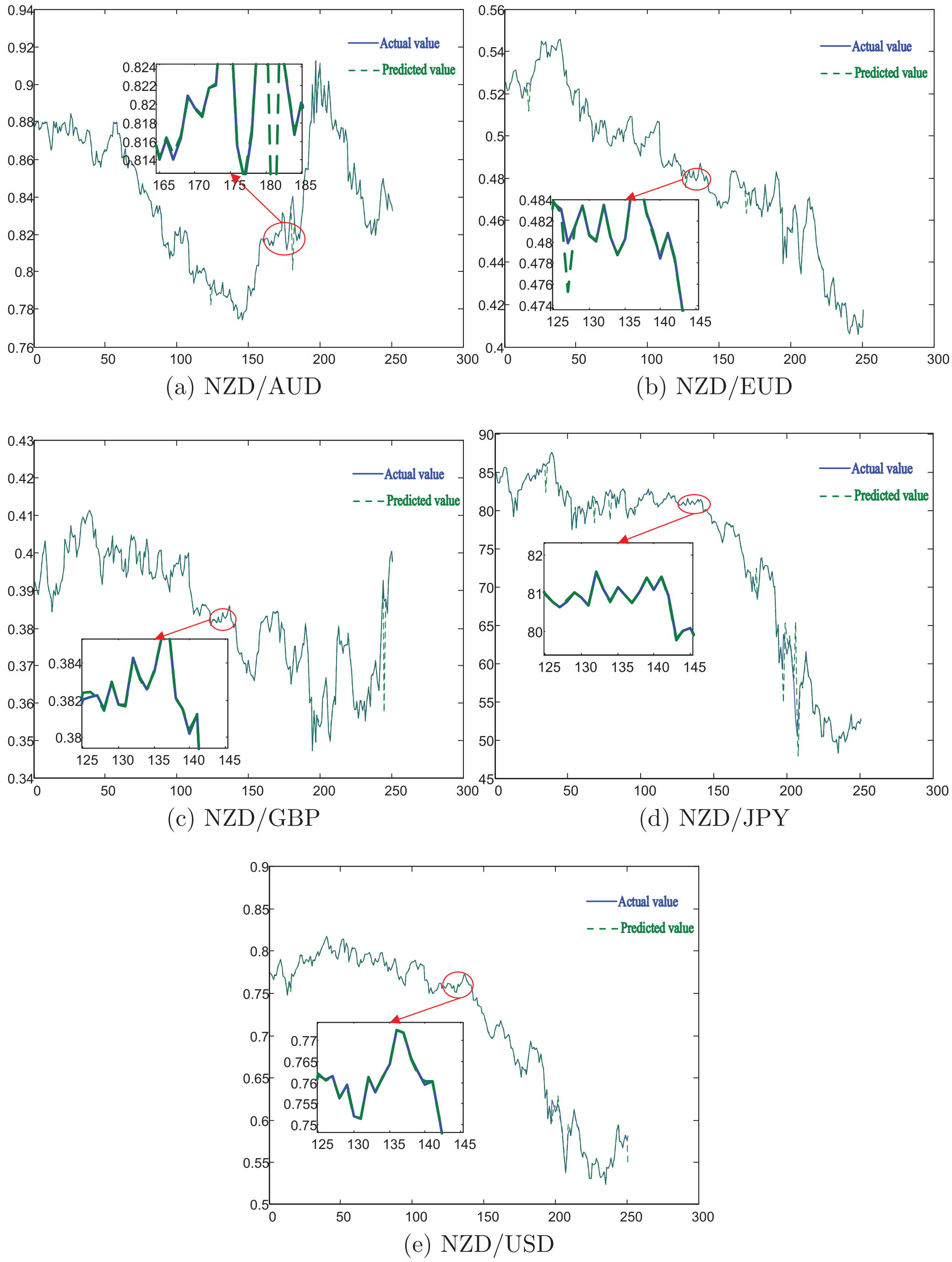


(a) NZD/AUD (b) NZD/EUD

(c) NZD/GBP (d) NZD/JPY

(e) NZD/USD

Figure 5.4: cSVR prediction on daily exchange rate, January 2007 to December 2007. Zoomed areas are 20 trading days randomly selected from this year.



(a) NZD/AUD (b) NZD/EUD

(c) NZD/GBP (d) NZD/JPY

(e) NZD/USD

Figure 5.5: cSVR prediction on daily exchange rate, January 2008 to December 2008. Zoomed areas are 20 trading days randomly selected from this year.

Chapter 6

Conclusions and Directions for Future Research

For forex market analysis, correlation is widely sought as an indicator of market variation. However, to discover comprehensive correlation knowledge from an observed market is practically very difficult for traditional technical and fundamental analysis approaches, because the variation of the market is influenced by diverse factors from domestic economic, historical market states, as well as international economic background.

This thesis proposes a computational correlation analysis for the automatic correlation extraction from available market and economic data. The proposed correlation is a synthesis of channel and weighted Pearson's correlation, where the channel correlation traces trend similarity even for a zig zag path time series, and the weighted Pearson's correlation filters noise in correlation extraction.

For correlation validity evaluation, correlation data is employed directly for aiding SVR time series prediction on 5 futures contracts (NZD/AUD, NZD/EUD, NZD/GBP, NZD/JPY and NZD/USD) within the period from January 2007 to December 2008. The results of comparison between cSVR and SVR show that the proposed correlation has been demonstrated to be significant for forex market analysis, as cSVR is performing consistently better than on all 5 contracts exchange rate prediction in terms of error functions such as MSE, RMSE, NMSE, MAE and MAPE.

The cSVR prediction is found sometime suffering unexpectedly far away from the truth value, which implies that despite the significance of the proposed correlation, how to use and fuse correlation into the present market data remains a challenge preventing us from enhancing further market understanding through computational analysis. In addition, the selection of macroeconomic factors and the determination of time period N for analysis are two computationally essential points worth addressing further for future forex market correlation analysis.

References

Abarbanell, J. S. & Bushee, B. J. (1997). Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research*, *35*(1), 1–24.

Abuaf, N. & Jorion, P. (1990). Purchasing power parity in the long run. *The Journal of Finance*, *45*(1), 157–174.

Agatonovic-Kustrin, S. & Beresford, R. (2000). Basic concepts of artificial neural network (ann) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*, *22*(5), 717 - 727.

Ajith Abraham, P. S., Ninan Sajith Philip. (2004). Modeling chaotic behavior of stock indices using intelligent paradigms. *CoRR*, *cs.AI/0405018*.

Akdede, B. B. K., Alptekin, K., Kitis, A., Arkar, H. & Akvardar, Y. (2005). Effects of quetiapine on cognitive functions in schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *29*(2), 233 - 238.

Appel, G. (2005). *Technical analysis:power tools for active investors*. FT Press.

Benjamin, G. & David, D. (2008). In *Security analysis: Principles and technique sixth edition* (p. 41-64). The McGraw.Hill Companies.

Boot, R. G., Verhoek, M., Fost, M. d., Hollak, C. E. M., Maas, M., Bleijlevens, B. et al. (2004). Marked elevation of the chemokine ccl18/parc in gaucher disease: a novel surrogate marker for assessing therapeutic intervention. *Blood*, *103*(1), 33-39.

Bose, I. & Raktim, P. (2005). Using support vector machines to evaluate financial fate of dotcoms. *PACIS 2005 Proceedings. Paper 42.*

Broda, C. (2004). Terms of trade and exchange rate regimes in developing countries. *Journal of International Economics*, *63*(1), 31 - 58.

Bruno, S., Cyril, B. & Yann, L. F. (1996). International market correlation and volatility. *Financial Analysts Journal*, *52*(5), 17-34.

Calvo, G. A., Leiderman, L. & Reinhart, C. M. (1993). Capital inflows and real exchange rate appreciation in latin america. *IMF Staff Papers*, *40*(1).

Cassel, G. (1918). Abnormal deviations in international exchanges. *The Economic Journal*, *28*(112), 413–415.

Chou, Y.-l. (1975). *Statistical analysis*. Holt International.

Deberard, M. S., Spielmans, G. I. & Julka, D. L. (2004). Predictors of academic achievement and retention among college freshmen: A longitudinal study. *College Student Journal*, *38*.

DETSKY, A. S., MCLAUGHLIN, J. R., BAKER, J. P., JOHNSTON, N., WHITTAKER, S., MENDELSON, R. A. et al. (1987). What is subjective global assessment of nutritional status? *parenteral and enteral nutrition*, *11*(1), 8-13.

Devereux, M. B. & Engel, C. (2003). Monetary policy in the open economy revisited: Price setting and exchange-rate flexibility. *The Review of Economic Studies*, *70*(4), 765–783.

Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A. & Vapnik, V. (1997). Support vector regression machines. *Advances in neural information processing systems*(9), 115-161.

DSouza, C. (2002). A market microstructure analysis of foreign exchange intervention in canada. *Bank of Canada Working Paper 2002-16*, *1192-5434*.

Elia, B., Patrizia, B., Luigi, M. & Ettore, M. (1998). Feed forward neural networks for the analysis of censored survival data: a partial logistic regression approach. *Statistics in Medicine*, *17*(10), 1169-1186. (10.1002/(SICI)1097-0258(19980530)17:10¡1169::AID-SIM796¿3.0.CO;2-D)

Emam, A. (2008). Optimal artificial neural network topology for foreign exchange forecasting. In *Acm-se 46: Proceedings of the 46th annual southeast regional conference on xx* (pp. 63–68). New York, NY, USA: ACM.

Fleming, M. J. & Remolona, E. M. (1999). Price formation and liquidity in the u.s. treasury market: The response to public information. *The Journal of Finance*, *54*(5), 1901–1915.

Frankel, J. A. & Rose, A. K. (1996). A panel project on purchasing power parity: Mean reversion within and between countries. *Journal of International Economics*, *40*(1-2), 209 - 224.

Gali, J. & Monacelli, T. (2005). Monetary policy and exchange rate volatility in a small open economy. *Review of Economic Studies*, *72*(252), 707-734.

Gelfand, J., Feldman, S., Stern, R., Thomas, J., Rolstad, T. & Margolis, D. (2004). Determinants of quality of life in patients with psoriasis: A study from the us population. *American Academy of Dermatology*, *51*(5), 704-708.

Goodhart, C. (1988). The foreign exchange market: A random walk with a dragging anchor. *Economica*, *55*(220), 437–460.

Hewison, J. & Tizard, J. (2004). *Major themes in education*. David Wray.

Hilde, C. B. & Hvard, H. (2006). The importance of interest rates for forecasting the exchange rate. *Journal of Forecasting*, *25*(3), 209-221. (10.1002/for.983)

Huang, Z., Chen, H., Hsu, C.-J., Chen, W.-H. & Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: a market comparative study. *Decision Support Systems*, *37*(4), 543 - 558. (Data mining for financial decision making)

Hunter, J. S. (1990). [exponentially weighted moving average control schemes: Properties and enhancements]: Discussion. *Technometrics*, *32*, 21-22.

Hur, Y. & Lim, S. (2005). Customer churning prediction using support vector machines in online auto insurance service. In *Advances in neural networks c isnn 2005* (Vol. 3497/2005, p. 928-933). Springer Berlin / Heidelberg.

Kaelbling, L. P., Littman, M. L. & Moore, A. W. (1996). Reinforcement learning: A survey. *Articial IntelligenceResearch*(4), 237-285.

Kamruzzaman, J. & Sarker, R. A. (2003). Forecasting of currency exchange rates using ann: a case study. In *Neural networks & signal processing (icnnsp03)* (p. 793-797). Proc. of IEEE Int.

Kirkpatrick, C. D. & Dahlquist, J. R. (2006). *Technical analysis:the complete resource for financial market technicians* (Vol. 704). Financial Times.

Kondratenko, V. V. & Kuperin, Y. A. (2003). *Using recurrent neural networks to forecasting of forex.*

Kwapien, J., Gworek, S. & Drozdz, S. (2009). Structure and evolution of the foreign exchange networks. *ACTA PHYS.POL.B*, *40*, 175.

Lane, P. R. & Milesi-Ferretti, G. M. (2005). Financial globalization and exchange rates. *CEPR Discussion Paper*(4745).

Lapata, M. (2006). Automatic evaluation of information ordering: Kendall's tau. *MIT Press*, *32*(4), 471-484.

Lothian, J. R. & Taylor, M. P. (2000). Purchasing power parity over two centuries: strengthening the case for real exchange rate stability: A reply to cuddington and liang. *Journal of International Money and Finance*, *19*(5), 759 - 764.

Mantel, N. (1963). Chi-square tests with one degree of freedom; extensions of the mantel- haenszel procedure. *Journal of the American Statistical Association*, *58*(303), 690–700.

Mitchell, T. M. (1999). Machine learning and data mining. *Commun. ACM*, *42*(11), 30–36.

Murphy, J. J. (1999). Technical analysis of the financial markets. *New York Institute of Finance*, 264.

Myers, J. L. & Well, A. (2003). *Research design and statistical analysis* (second edition ed.). Lawrence Erlbaum.

Neely, C. J. (1997). Technical analysis in the foreign exchange market: a layman's guide. *Review*(Sep), 23-38.

Ongsritrakul, P. & Soonthornphisaj, N. (2003, July). Apply decision tree and support vector regression to predict the gold price. In (Vol. 4, p. 2488-2492 vol.4).

Paez, J. G., J?nne, P. A., Lee, J. C., Tracy, S., Greulich, H., Gabriel, S. et al. (2004). Egfr mutations in lung cancer: Correlation with clinical response to gefitinib therapy. *Science Express*, *304*(5676), 1497 - 1500.

Pai, P.-F. & Lin, C.-S. (2005). Using support vector machines to forecast the production values of the machinery industry in taiwan. *The International Journal of Advanced Manufacturing Technology*, *27*(1-2), 205-210.

Panackal, A. A., Gribskov, J. L., Staab, J. F., Kirby, K. A., Rinaldi, M. & Marr, K. A. (2006). Clinical significance of azole antifungal drug cross-resistance in candida glabrata. *Clinical Microbiology*, *44*(5), 1740-1743.

Park, C. & Irwin, S. (2004). *The profitability of technical analysis: A review* (Tech. Rep.). AgMAS.

Pearson, K. (1897). Mathematical contributions to the theory of evolution.–on a form of spurious correlation which may arise when indices are used in the measurement of organs. *The Royal Society*, 489-498.

Plackett, R. L. (1983). Karl pearson and the chi-squared test. *International Statistical Review*, *51*(1), 59C72.

Pring, M. J. (2002). *Candlesticks explained* (illustrated ed.). McGraw-Hill Professional.

Rodgers, J. L. & Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, *42*(1), 59–66.

Saacke, P. (2002). Technical analysis and the effectiveness of central bank intervention. *Journal of International Money and Finance*, *21*(4), 459 - 479.

Sansom, D. C., Downs, T. & Saha, T. K. (2003). Evaluation of support vector machine based forecasting tool in electricity price forecasting for australian national electricity market participants. *Electrical and Electronics Engineering,Australia*, *22*(3), 227-234.

Sapankevych, N. I. & Sankar, R. (2009). Time series prediction using support vector machines: A survey. *IEEE Computational Intelligence*, *4*(2), 25 - 37.

Scholkopf, B., Burges, C. J. C. & Smola, A. J. (1999). *Advances in kernel methods - support vector learning.* The MIT Press.

Schwager, J. D. (1996). Technical analysis. *John Wiley and Sons*, 545.

Shubhabrata, D. & Malay, K. B. (2004). Optimizing parameters of supervised learning techniques (ann) for precise mapping of the input-output relationship in tmcp steels. *Scandanavian Journal of Metallurgy*, *33*(6), 310-315.

Snchez, M. (2005). *The link between interest rates and exchange rates do contractionary depreciations make a difference?* (Tech. Rep.). Social Science Research Network electronic library.

Sutherland, E., Martin, R., Bowler, R., Zhang, Y., Rex, M. & Kraft, M. (2004). Physiologic correlates of distal lung inflammation in asthma. *Allergy and Clinical Immunology*, *113*(6), 1046-1050.

Takeshi, O. (2005). Accurate sum and dot product. *SIAM journal on scientific computing*, *26*(6), 1955.

Tate, R. F. (1954). Correlation between a discrete and a continuous variable. point-biserial correlation. *The Annals of Mathematical Statistics*, *25*(3), 603–607.

Tay, F. E. H. & Cao, L. J. (2001a). Application of support vector machines in financial time series forecasting. *Omega*, *29*, 209 - 317.

Tay, F. E. H. & Cao, L. J. (2001b). Improved financial time series forecasting by combining support vector machines with self-organizing feature map. *Intell. Data Anal.*, *5*(4), 339–354.

Taylor, J. B. (2001). The role of the exchange rate in monetary-policy rules. *The American Economic Review*, *91*(2), 263–267.

Taylor, M. P. & Allen, H. (1992). The use of technical analysis in the foreign exchange market. *Journal of International Money and Finance*, *11*(3), 304 - 314.

Trafalis, T. B. & Ince, H. (2000). *Support vector machine for regression and applications to financial forecasting* (No. 6). Neural Networks, 2000. IJCNN 2000.

Vapnik, V. N. (1999). An overview of statistical learning theory. *Neural Networks, IEEE Transactions on*, *10*(5), 988-999.

Walter, C. A. & Lopez, J. A. (2000). Is implied correlation worth calculating? evidence from foreign exchange options. *The Journal of Derivatives*, *7*(3), 65-81.

Wilder, J. W. (1978). *New concepts in technical trading systems.* Trend Research Greensboro, NC.

Wong, K. F., Selzer, T., Benkovic, S. J. & Hammes-Schiffer, S. (2005). Impact of distal mutations on the network of coupled motions correlated to hydride transfer in dihydrofolate reductase. *The National Academy of Sciences*.

Yao, J. & Tan, C. L. (2000). A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing (Amsterdam)*, *34*(1), 79.

Yu, C., Chan, Y., Zhang, Q., Yip, G., Chan, C., Kum, L. et al. (2005). Benefits of cardiac resynchronization therapy for heart failure patients with narrow qrs complexes and coexisting systolic asynchrony by echocardiography. *American College of Cardiology*, *48*(11), 2251-2257.