Prediction of International Stock Market Movement Using Technical Analysis Methods and TSK

Mahammad Abdulrazzaq Thanoon

Submitted to the
Institute of Graduate Studies and Research
in partial fulfilment of the requirements for the Degree of

Master of Science
in
Computer Engineering

Eastern Mediterranean University
April 2014
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

____

Prof. Dr. Elvan Yılmaz
Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Master of Science in Computer Engineering.

Prof. Dr. Işık Aybay
Chair, Department of Computer Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Computer Engineering.

Asst. Prof. Dr. Mehmet Bodur
Supervisor

Examining Committee

1. Asst. Prof. Dr. Adnan Acan
2. Asst. Prof. Dr. Mehmet Bodur
3. Asst. Prof. Dr. Ahmet Ünveren

ABSTRACT

This research aimed to propose a method to improve forecasting accuracy of the technical analysis of future closing price using Takagi-Sugeno-Kang (TSK) fuzzy model to merge the forecasting of three technical prediction methods. The historical data available for London Stock Market is employed in this study to verify the performance of the proposed model compared to technical predictions.

Fuzzy data modelling emerges as an advanced technique in predicting future closing prices. In this study, the predictions of three technical analysis methods were modelled by Fuzzy Methods to enhance the predicted closing price. The fuzzy rules were extracted by using Fuzzy-C-Means (FCM) algorithm.

Data set from year 2008 to 2012 is dividing in two parts for training and verification purpose. The Fuzzy C-Means clustering (FCM) is applied on the six days Moving Average (SDMA), the Moving Average Convergence Divergence (MACD), and the Relative Strength Index (RSI) technical analysis to predict the future price, which is, target variable of the TSK fuzzy model. A prediction accuracy close to 94.7%, is achieved in predicting two days ahead closing prices of London Stock Market. The results are very encouraging and easy to implement in real-time trading system.

Keywords: Technical Forecasting, Fuzzy Modelling, TSK, FCM, clustering, moving average.

ÖZ

Bu araştırma Takagi Sugeno Kang (TSK) modeli kullanarak üç kapanış fiyatı teknik analiz yönteminin tahminlerini geliştirmeyi amaçlamıştır. Londra Hisse Senedi piyasası verileri önerilen yöntemin performansını sınamak üzere kullanılmıştır.

Bulanık mantıklı veri modellemesi piyasaların gelecekteki kapanış fiyatını tahminde başarılı bir yöntem olarak ortaya çıkmıştır. Bu çalışmada üç standart teknik analiz metodunun tahmin güçleri, kuralları FCM algoritması ile elde edilen bulanık modelleme yöntemleri ile birleştirilerek arttırılmıştır.

Kullanılan 2008 ile 2012 yılları arasındaki menkul değer piyasa kapanış fiyatları model oluşturma ve model sınama amaçlı iki bölüme ayrılmıştır. Altı-günlük ortalama, kayar ortalamalı yakınsama ıraksama, ve göreceli dayanım indeksi olmak üzere üç teknik analiz yöntemi model oluşturma verisi ile fiyat tahmini için kullanılmış, çıkan tahminler FCM ile gelecekteki fiyat hedeflenerek değerlendirilmiştir. Sınama verisini kullanarak yapılan karşılaştırmada Londra menkul değer piyasalarında 94.7% civarında başarıyla tahmin gerçekleşmiştir. Yöntemin uygulanmasının kolaylığı nedeniyle gerçek zamanlı yatırım uygulamalarında kullanımı açısından cesaret vericidir.

Anahtar Kelimeler: Teknik tahmin, Bulanık model, TSK, FCM, öbekleme, kayar ortalama.

DEDICATION

To my family

My Father & Mother

My wife

My siblings

My son

ACKNOWLEDGMENT

With all the respect and gratitude, I would like to thank my supervisor Asst. Prof. Dr. Mehmet Bodur for his endless assistance and support in carrying out this thesis (God bless him as a symbol for knowledge and learning), and without him I would never see the light of the future upon my life.

I would also like to thanks my great university Eastern Mediterranean University (EMU) and especially the Department of Computer Engineering for valuable support and suggestions. My thankful goes to my precious inveterate university (University of Mosul), which I am honoured to be part in stiff family, and absolutely goes to my wonderful College (College of Electronic Engineering) deanery and doctors.

My special thanks to all my friends, especially Dr. Hashem Al hendi, who help or advised to finish this study. Ask Allah to give them all the best in their life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
AR	Autoregressive
ARIMA	Autoregressive Moving Integrated Average
P	Closing Price
EMA	Exponential Moving Average.
FCM	Fuzzy C-Means.
GBP	British Pound
LD	London stock market
MA	Moving Average.
MACD	Moving Average Convergence Divergence.
Matlab	A software for matrix operations, Math Works, Inc., R2014a.
MF	Membership Function.
NN	Neural Networks.
RMSE	Root Mean Square Error.
RSI	Relative Strength Index.
SDMA	Six Days Moving Average
SMPM	Stock Market Prediction Model
TSK	Takagi-Sugeno-Kang.

LIST OF SYMBOLS

y	The target point
x_a	The previous point
$\bar{x}$	Sample mean
x_b	The next point
y_b	Next point value
y_a	Previous point value
$\in$	Set membership to the interval [0, 1].
A(x)	The membership function (MF) for the fuzzy set.
A, B	A fuzzy set in X.
A_1	Corresponding fuzzy set in TSK model
A_2	Corresponding fuzzy set in TSK model
b_0, b_1, b_2	Linear consequent parameters in TSK model
c	Cluster number.
z^*_i	Constant output value for each rule in zero-order Sugeno fuzzy model .
E	The prediction error
f	Function
x^*_k	Forecast value.
m	Fuzzification power of FCM.
n	The number of data points in x i.e. the time period.
N	Size of time series observations
P_k	Closing price observation at time $_k$

$P_{S,k+1}$	$SDMA_k$ next closing price.
v_i	The centre of cluster.
$v_{i,j}$	The degree to which element x_i belongs to cluster.
X	A collection of objects denoted normally by x.
x	Input variable
z	A crisp function in the consequence for the fuzzy set.
σ	Standard deviation.
α	Linear regression coefficient vector.
$SMA(U,n)$	Average of up closed days.
$SMA(D,n)$	Average of down closed days.
$E_{RMS,NT}$	Estimation error for training.
$E_{RMS,NV}$	Estimation error for verification.
X_T	Matrix of training inputs and outputs.
X_V	Matrix of verification inputs and outputs.

Chapter 1

INTRODUCTION

1.1 Time Series Data Set

A time series is a collection of observations collected consecutive in time, such as a particulate pollution measurements and temperature information. Most information in economic science and finance are time series data sets. Time series models in science of economics have mostly fixed intervals between the observations, such as days, weeks, months, etc., [1]. A time series is a set of N observations with observation periods T in historical ordering

$$Y = \{y_1, y_2, ..., y_k, ..., y_N\},$$

where k points the time of observations $t = kT$.

Many forecasting methods cannot use a time series if the statistical characteristic (mean, autocorrelation, and variance) changes over the time. A time series dataset with an almost constant mean and variance is called *stationary*. Transforming a non-stationary data set to difference data or removing its slope from the data set converts it to a stationary data set [2].

1.2 Financial time series estimation

Estimation of the future values for a time series is an essential task for business and decisions on financial investments. There are many analysis strategies for the estimation of the future value of a time series, such as Six day moving average (SDMA), Relative Strength Index (RSI) and 26 days Moving Average Convergence

Divergence (MACD) techniques, which are applied in this thesis. Other than these three techniques there are commonly used methods based on Artificial Neural Nets (ANN), Auto Regressive Integrated Moving Average (ARIMA), and Fuzzy Logic Modelling of time series data set. ARIMA discovers the dynamic behaviour of the system from the data set, and estimates the future effect of input error by the dynamic behaviour of the system. That means, if the estimated value diverges from the actual value, the future value is calculated according to the auto regressive moving average effect of the prediction error and the dynamic behaviour of the system. Stock market expectation is a very significant financial matter that has concerned researchers' consideration for many decades. Approaches of stock market prediction may involve assumptions of some relationships between the stock return and several variables of the observations that build the market data [3]. The columns of data set may include some variables such as, interest rates, exchange rates, growth rates, client value, financial gain statements and dividend yields. *Fundamental analysis* focuses on the overall economic indices and the success of the industry groups related to a business. Fundamental analysis is a precise active approach to estimate economic situations, but not essentially actual market prices. Financial time series analysis aims to discover the patterns of movements, unexpected changes on a non-stationary transform of time series data set.

The *technical approach* relies on the idea that the price depends on the psychological state of a mass ('the crowd') in action. The approach tries to connect the future value to the psychological condition of people, negatively such as alarm, fear, and doubt, or positively sureness, extreme assurance, and greed [4]. Stock price prediction has always an important matter for numerous financiers and skill analysts. Even so, looking for the most effective time to buy or sell has continued a troublesome task as a result of there are a factors which it guidance a stock price.

Technical analysis, as one of method for analysis, assumes that the market activities, related news, and psychological observations may affect stock prices, and they must be considered in forecasting future prices and market trends. Among the techniques, that builds this class of analysis, MA, and recent artificial intelligence techniques are gained more attention. The MA has smoothing effect on a dataset to explain the dataset trends. It is called moving because for every time step the latest period is added and the ancient period is dropped. The MA is based on historical prices.

Clustering is the method of distributing dataset into groups. As a result, the items in one group are similar as probable, and items in unlike group are different as probable. The measurements that used in clustering contain distance, connectivity, and intensity. A commonly algorithm used fuzzy clustering is the Fuzzy C-Mean (FCM) algorithm [5].

Fuzzy Takagi-Sugeno-Kang (TSK) model provides advanced forecasting methods based on extracted fuzzy rules from a given data set. The rules provide knowledge about behaviour of the modelled system. The fuzzy rules may be extracted directly from the data set, as well as, from the results of some reliable methods that summarize the characteristics of the data set [6] [7] [8] [9] [10].

The basic indication of fuzzy logic is to let not only the values *0* and *1*, matching to *false* and *true* but the whole interval [0, 1] as marks of truth.

The aim of this thesis is to predict two-day-ahead future price of London Stock Market by TSK fuzzy model, using MACD, SDMA and RSI. The prediction performance of the proposed TSK based method is compared against the performance of linear regression based prediction models using these three indices. The coefficients of the linear expressions are obtained from the first half of five years stock prices (2008-2012), and verified using the last half of the time series observations. The proposed method attempts to improve the prediction accuracy of the three indices using TSK model. This proposed model also provides hints of conditions to expect a rising or falling stock markets. The structure of a fuzzy rule of a first order TSK model with two input features x_1 and x_2, is:

$$\text{Rule: If } x_1 \text{ is}_r\ A_1\text{, and } x_2 \text{ is}_r\ A_2\text{, then } y^* = b_0 + b_1 x_1 + b_2 x_2 \ , \qquad (1)$$

where x_1 and x_2 are scalar linguistic variables, A_1 and A_2 are linguistic terms described by fuzzy sets and b_0, b_1, b_2 are coefficients of a liner expression that forms the consequent part of the rule. The inference for an input x is obtained by aggregating the y^* values of the rules using membership values of x in the terms of the premise [11], [12].

The performance of the model is measured by comparing the inferred y' values to the actual y values using Root Mean Square Error of the predicted values over the test period.

1.3 Summary of the Proposed Method

The important motivation for trying to predict the stock market prices is financial advantage. The skill to find a mathematical model that can reliably predict the direction of the future stock prices would make the owner of the model actual wealthy. Thus, academics, investors and investment professionals are always trying to find a stock market model that would profit them with higher returns.

This thesis proposes application of Takagi–Sugeno–Kang (TSK) method to merge three indices of closing prices, SDMA, MACD, RSI, and the current closing prices into the next day's predicted closing price. The TSK fuzzy model uses these indices as input variables to discover the rules that explain the dependence of the change of closing price on the change of indices by using the training data set. The consequents of the TSK rule base is a linear combination of the input variables. Once the rule base is extracted, TSK inference method predicts the next day's closing price from the current closing prices and available indices with a smaller prediction error than prediction of closing prices using each index alone.

1.4 Organization of the Thesis

In this Thesis Chapter 1 introduces Stock markets, source of data set, technical prediction methods, TSK Fuzzy Modelling and explains the problem. Chapter 2 describes the data set and theoretical of methods that used in this work which are as follows: the SDMA, the MACD and, the RSI technical analysis; the FCM and the TSK fuzzy model. Chapter 3 consists of the methodology for prediction model that proposed in this work. The performance of prediction of the proposed method is discussed in Chapter 4. Chapter 5 concludes the advantages of the proposed method, and comments on the future work.

Chapter 2

DATA SET AND EXISTSING METHODS OF PREDICTION

2.1 The Data Sets and Pre-processing

This thesis proposes a time series prediction method, and demonstrates the proposed method on five years closing prices (P) of LD stock market from 1-01-2008 to 1-01-2013. The dataset is achieved from *www.finance.yahoo.com* website [13].

The original dataset taken from *www.finance.yahoo.com* had missing days because the stock markets were closed during the weekends and holidays. The missing values of the time series data has negative effect in prediction of future prices because the skipped time steps in the data set corresponds to economic consumption during that period. There are mainly three methods to fill the missing records for the off days: the simplest method is filling either the next day or the previous day prices into the missing day's closing price. The third method fills the missing price of k^{th} day, P_k, by linear interpolation using both the previous and next day prices, P_{k-1} and P_{k+1}. Table 2.1 shows an example for this method; the value for the fourth day does not exist.

Table 2. 1 Data with missing value

day k	value P_k
0	0
1	0.3
2	0.5
3	0.1
4	Missing
5	0.7

Previous-value method fills P_4 by P_3. Next-value method fills P_4 by P_5. Linear interpolation method calculates P_k using the available previous and the next prices P_{k-n} and P_{k+m} by [14]

$$P_k = P_{k-n} + (P_{k+m} - P_{k-n})n/(m+n) \tag{2.1}$$

i.e., in Table 2.1, interpolation gives $P_4 = P_3 + (P_5 - P_3)(4-3)/(5-3)$.

2.2 Closing Prices of London Stock Market

The arbitrary movement of the closing prices for London stock market is shown in the Figure 2.1. The closing prices in London stock market in the beginning of 2008 are about 6000 GBP (Figure 2.1). It doesn't change well totally after 4 years from the impact of the financial crisis. The average of the closing prices over 5 years is 5365 GBP.

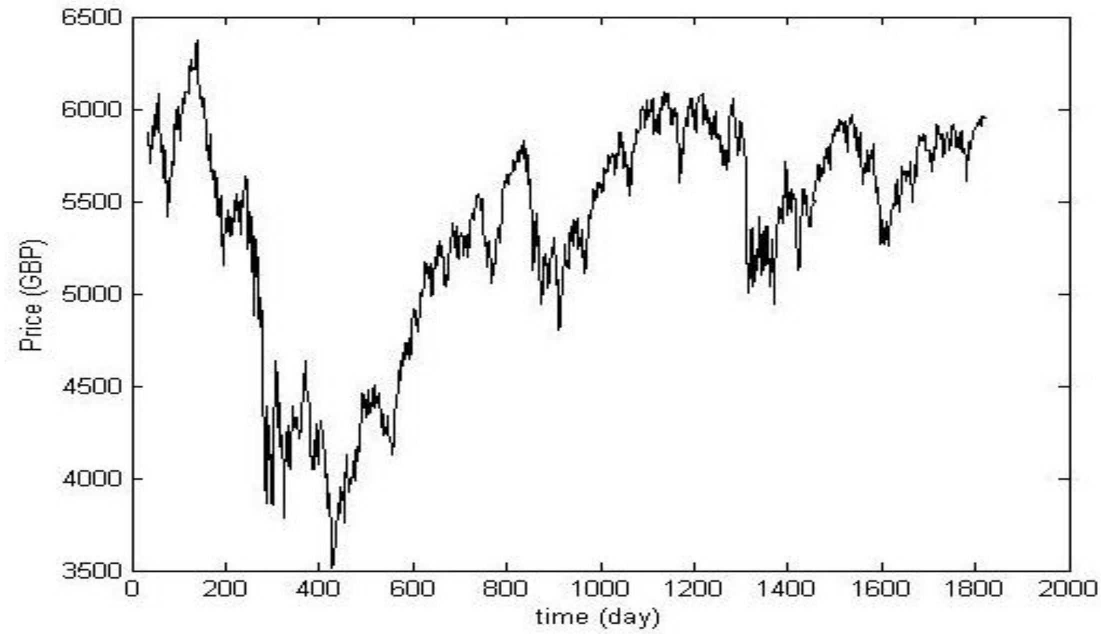


Figure 2.1: Closing price of London Stock Market

2.3 Moving Average Index of a Market

Moving Average (*MA*) smoothens out the noise on a data series, making it easy and reliable to compare to the latest value. It is named moving because for each time step the latest period is added into average, and the oldest period is dropped from the average value. In a stock market, it integrates the closing prices of the moving average period, and therefore, in technical analysis it has a lagging effect as an indicator. Simple Moving Average over *n* day is defined by

$$MA_k = \frac{1}{n}\sum_{i=0}^{n-1} P_{k-i} \tag{2.2}$$

In this thesis, 6-day Moving Average ($SDMA_k$) is used to indicate the short term closing prices with reduced effect of volatility based noise [15].

In general, $SDMA_k$ together with P_k may provide a prediction of next closing price P_{k+1} through a linear expression

$$P_{S,k+1} = \alpha_{S,1}\ P_k + \alpha_{S,2}\ SDMA_k \qquad (2.3)$$

The error in k^{th} prediction by *SDMA* is $E_{S,k} = P_k - P_{S,k}$. Over n observations the RMS error of the estimation by *SDMA* is calculated by

$$E_{RMS,S} = \sqrt{\frac{1}{n}\sum_{i=0}^{n-1}(P_{k-i} - P_{S,k-i})^2} \qquad (2.4)$$

MA method used to display when an investor can sell or buy in a particular financial market because it provides a measure of the momentum. Also, the investor can use the *MA* method to determine when the prices are likely to change its direction. Depending on historical commercialism ranges, the support and confrontation points are recognized where the price of a stock market reversed its upward or downward trend, within the past price, by this way the buy or sell decision are decides. The most common applications of *MA* are to identify the trend direction and to determine support and resistance levels.

2.4 Moving Average Convergence Divergence (MACD)

The difference between the long-term exponential moving average and the short-term exponential moving average is called Moving Average Convergence Divergence (*MACD*). It is a technical analysis indicator created by Gerald Appel in the late 1970s [16]. Thomas Aspray added a bar chart to the *MACD* in 1986, as a means to expect *MACD* crossovers, an indicator of necessary moves in the underlying security. It is used to spot changes within the strength, direction, momentum, and period of a trend during a stock's price. *MACD* is calculated in MatLab simply by the *macd*() function. *MACD* is calculated over the prices *Pk*, k=1...n using exponential moving average (*EMA*) [17]

$$EMA_{12,k} = 0.985 EMA_{12,k-1} + 0.015\ P_k \qquad (2.5)$$

and 26-day *EMA*

$$EMA_{26,k} = 0.926\, EMA_{26,k-1} + 0.074\, P_k \quad (2.6)$$

The difference of EMA_{12} and EMA_{26},

$$MACD_k = EMA_{12,k} - EMA_{26,k} \quad (2.7)$$

gives the short-term trend of the price. A 9-day *EMA* indicates the buying or selling day of the stock if the decision shall be based only on *MACD*.

The theory of *MACD* is that when two *MA*s cross, a major change of trend in the stock's price is more probably to occur. Like all indicators, the *MA* crossover has large uncertainty to consider it as an absolute truth in trading stocks [18].

In general, $MACD_k$ together with P_k may provide a prediction of next closing price P_{k+1} through a linear expression [19]

$$P_{M,k+1} = \alpha_{M,1}\ P_k + \alpha_{M,2}\, MACD_k \quad (2.8)$$

2.5 Relative Strength Index (RSI)

The relative strength index (RSI), one of the common technical indicators, RSI was first introduced by Welles Wilder [20]. A list of step-by-step commands to calculate and interpret the RSI is provided in Wilder's book, New Concepts in Technical Trading Systems. RSI is calculated in MatLab simply by the *RSindex()* function. It is based on the simple moving averages of the up steps ($SMA(U, n)$) and down steps ($SMA(D,n)$) in total n=14 day period.

$$RS = SMA(U,n)/SMA(D,n) \quad (2.9)$$

Consequently, *RS* is simply the ratio of average up closed days to the average of down closes days.

RSI is normalised value of *RS* to move between 0 and 100 by using the formula given below:

$$RSI_k = 100 - \frac{100}{(1 + RS_k)} \qquad (2.10)$$

RSI is a technical momentum indicator that compares the magnitude of recent gains to recent losses to detect overbought and oversold conditions of an asset. RSI may create false buy or sell signals. RSI is a valuable balance to other technical analysis methods.

Next closing price P_{k+1} is estimated using RSI_k and P_k through a linear expression

$$P_{R,k+1} = \alpha_{R,1} \; P_k + \alpha_{R,2} \; RSI_k \qquad (2.11)$$

2.6 Fuzzy System Modelling using an Observed Data Set

The input and output relation of a system can be obtained by many approaches. Zadeh's Fuzzy number and extension principle provides a solid ground to develop fuzzy modelling methods by a number of fuzzy rules that consists of two main parts: a premise and a consequent [21]. Several methods were proposed to transfer the experts' ideas or the characteristic relations in a set of observed inputs and consequent outputs into the fuzzy rules. Zadehs Singleton Method (SM), Kosko's Standard Additive Method (SAM), Mamdani's Center of Gravity Method (CoG), and Takagi-Sugeno's method (TS) are widely used and well known rule construction, representation, and inference methods [22].

2.6.1 Fuzzy Sets and Fuzzy Logic

In early seventies Zadeh introduced the fundamentals of fuzzy set and fuzzy logic [22]. Fuzzy logic provides several modelling methods with properties of a universal approximate, which means it can imprecise any continuous function in an interval to any degree of precision using the required number of fuzzy rules [24].

The idea behind Fuzzy logic is to use the whole interval [0, 1] as a measure of truth. In the 1920, J. Lukasiewicz introduced multivalued logic calculus, but its application was restricted until the introduction of computer technology in the late 1950 [21].

Fuzzy set extends the membership predicate "∈" to a membership value in the interval [0, 1]. This implies that a collection can contain points with a certain degree of membership. This degree of membership is often considered in several ways. Fuzzy set theory allows us to consider the uncertainty data attributes.

Let X is a set of objects denoted by x, a fuzzy set A in X is a set of well-ordered pairs:

$$A=\{(x, A(x) \mid x \in X, A(x) \in (0, 1) \}$$

where $A(x)$ is the membership function (MF) for the fuzzy set. The MF $A(x)$ maps each element of X to a membership value $A(x) \in (0, 1)$.

2.6.2 Fuzzy c-Means in Extracting Input-Output Relation

Bezdek's Fuzzy c-Means clustering method provides an easy approach to extract input-output relations from large observation data sets. Data clustering is the process of partitioning a data set into a number of classes or clusters. The aim of clustering is to have similar elements in the same class and dissimilar items in different classes. There are methods to cluster data according to distance, intensity, or property [25].

In fuzzy clustering, a data vector belongs to all clusters with some membership values in (0, 1). A vector belongs to one of the clusters dominantly if the membership value for that cluster is considerably higher than all other clusters. For the rule extraction purpose, FCM clustering algorithm is applied on input-output data vectors xy_k using a distance metrics $d(v_i, xy_k)$ which stands for the distance from xy_k to v_j,. Let the set of data contains N vectors. The iterative algorithm returns c cluster centres $\{v_1, v_2, ... v_c\}$ in a matrix form

$V=[v_1\ v_2\ ...\ v_c]^T$,

a partition matrix,

$U= [u_{j,k}]$, where $u_{j,k}$ is in the interval [0,1], for $j= 1,\dots,c$ and $k=1,\dots,N$.

where $u_{j,k}$ tells the degree to which element xy_k belongs to cluster of v_j. FCM aims to minimize the objective function for a data set $XY=[\ xy_1\ \ xy_2\ ...\ xy_k\ ...,xy_N]^T$:

$$J_m(U,V;XY) = \sum_{k=1}^{N} \sum_{j=1}^{c} u_{j,k}{}^{m}.d(xy_k, v_j\ ); \quad (2.12)$$

by calculating

$$u_{j,k} = \frac{1}{\sum_{i=1}^{c} \left(\frac{d(v_j, xy_k)}{d(v_i, xy_k)} \right)^{2/(m-1)}} \quad (2.13)$$

and

$$v_j = \frac{\sum_{k=1}^{N} u_{j,k}{}^{m} xy_k}{\sum_{k=1}^{N} u_{j,k}{}^{m}} \quad (2.14)$$

iteratively after each other.

The fuzzification constant m determines the level of cluster fuzziness. Any point xy has a set of coefficients giving the degree of being within the i^{th} cluster with the center v_i. With fuzzy c-means, the centre of a cluster is the fuzzy mean of all points, weighted by their degree of membership values in that cluster as expressed by (2.14)

The degree of belonging, $u_{j,k}$ depends inversely on the distance from xy to the cluster centre v_j and the fuzzification power m that controls how much weight is given to the closest centre [26].

2.6.3 TS and TSK Models

A *fuzzy model* is a mathematical model that uses fuzzy sets to describe the input output relationships in a data set. The models are based on fuzzy rules and inference methods. The fuzzy rules represent the relationship between the variables through linguistic terms.

The TS fuzzy model was proposed by Takagi and Sugeno as a novel method to express the input-output relation of a system by TS fuzzy rules that contains linear expression to compute the output as the consequent part of the rule. Takagi, Sugeno and Kang developed this approach using FCM to extract rules from a set of input-output vectors [26].

For two-input plus an output variable observations (x, z), where input x has two components (x_1, x_2), a fuzzy rule in a Sugeno fuzzy model has the following structure:

If (x_1 is_r A_1) and (x_2 is_r A_2) then $z = f(x_1, x_2)$,

where A_1 and A_2 are fuzzy sets in the antecedent, and $z = f(x_1, x_2)$ is an arithmetic expression in the consequence. $z = f(x_1, x_2)$ is a polynomial in the input variables x_1 and x_2. It can be any function as long as it approximates the input-output of the observations in a fuzzy region specified by the antecedent of the rule. When $f(x_1, x_2)$ is a first order polynomial, the resultant fuzzy inference system is named a first-order Sugeno fuzzy model [27].

For multiple fuzzy rules R_i , i= 1…c.,

R_i: if x_1 is $A_{i,1}$ and x_2 is $A_{i,2}$ then $z^*_I = f_i(x_1, x_2)$.

The model infers to z^*, the predicted value of z, by

$$z^*(x)= \sum_{i=1}^{c} \beta\left(A_{i,1}(x_1), A_{i,2}(x_2) \right) f_i(x_1,x_2) \Big/ \sum_{i=1}^{c} \beta\left(A_{i,1}(x_1), A_{i,2}(x_2) \right) \tag{2.15}$$

where $\beta(A_{i,1}(x_1), A_{i,2}(x_2))$ is the degree of fulfilment of input vector $x=(x_1, x_2)$ by the rule R_i. Zadeh's fuzzy singleton rule is formed from the Sugeno fuzzy model using zero-order z^*_i, a constant output value for each rule.

TS model originally proposed that experts shall set the premise part of the rules and the liner expressions in the consequents are then calculated using the available set of observations. TSK model introduced that the premise can be extracted by a clustering method that relates the inputs to the outputs. For each cluster obtained by FCM, the fuzzy sets of the premise of a fuzzy rule is obtained by processing the projections of the membership values of the observations, using convex points method that finds the envelope covering all projected membership values for a component of the inputs. The convex points are furthermore approximated by a simple membership function such as a triangular, trapezoidal, or Gaussian shape by determining the parameters of one of these shapes [11].

Chapter 3

PROPOSED PROCESS OF FORECASTING

This chapter presents the proposed forecasting method to build a Stock Market Prediction Model (SMPM), depending on TSK to predict the stock prices using well known forecasting methods in the input arguments. Closing prices of London Stock Market from 21/08/2009 to 2012, total 1200 days, were used as the time series data to verify the proposed method. A block diagram of all processes of proposed SMPM including its test processes is shown in Figure 3.1.where $E_{RMS,NT}$ = Estimation error for training, $E_{RMS,NV}$= Estimation error for verification.

3.1 Data Pre-processing

Data pre-processing is performed on the time series data in order to bridge the gap of the missing dates using the interpolation method. The comparison of the prediction error with and without missing data is carried on one-day and two-day ahead predictions using linear regression of the last two days to verify the effect of this process. After completing the missing values, the time series is divided in two parts: From almost 1800 days in time series data, the first 900 is used for training purpose in calculating the parameters of the prediction methods, and the last 900 is used for verification purpose to determine the performance of the prediction method.

3.2 Calculation of Stock Market Indices

Three technical indexes, SDMA, MADC, and RSI, are calculated on the pre-processed time series data using the procedures described in Chapter 2. From these indices, MADC and RSI show the trend of the prices, and SDMA gives an accurate short term delayed price status. The indices individually were used in predicting the one-day and two-day-ahead prices by linear regression to determine the level of information content in each of these indices by the following procedure.

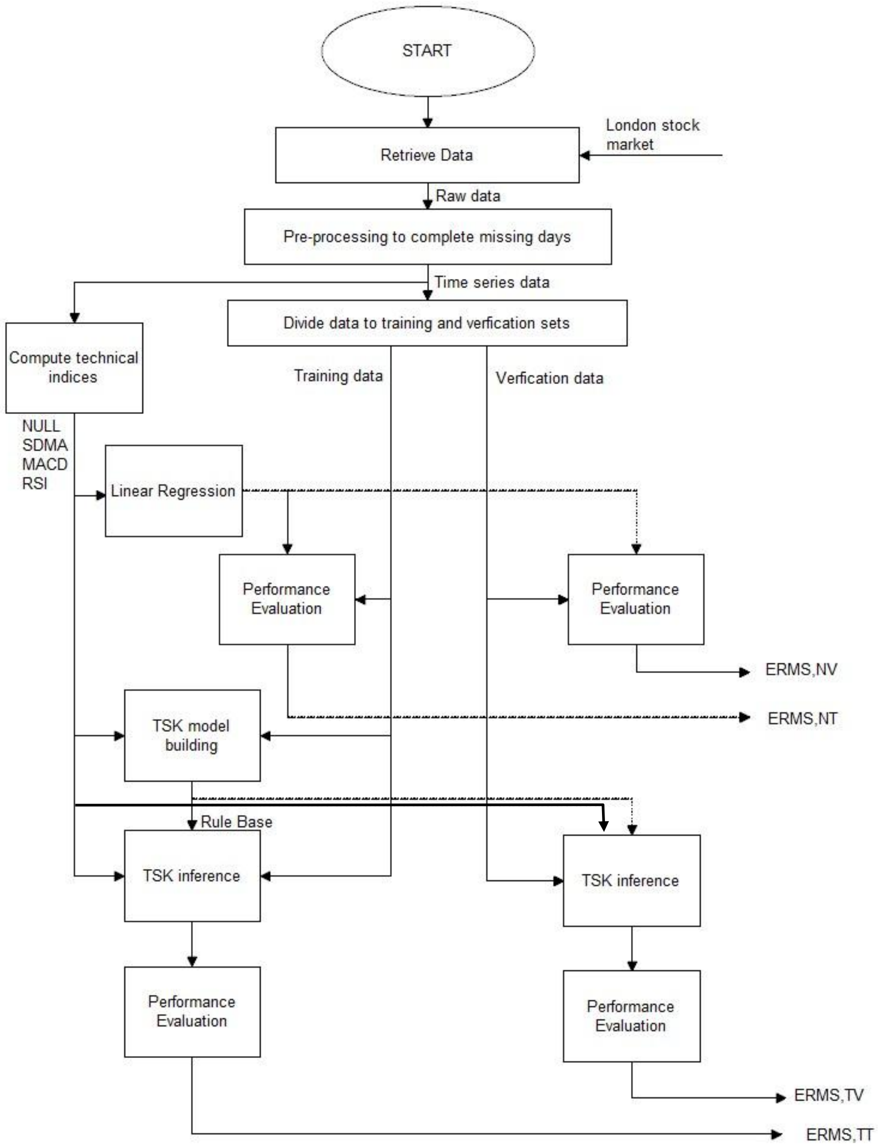


Figure 3.1 The structure of the proposed process including its testing.

3.3 Regression without Technical Indices

Regression without Technical Indices (Null Regression) is necessary to compare the prediction performance of an index in predicting the a-day-ahead future price $P_{N,k+a}$

with and without that index. A-day-ahead price regression is carried to determine the coefficients of the linear expression

$$P_{N,k+a} = \alpha_{N,1}\ P_k + \alpha_{N,2}\ P_{k-1} \tag{3.1}$$

The coefficients $\alpha_N = (\alpha_{N,1}, \alpha_{N,2})$ are obtained by forming matrices of training inputs and outputs

$$X_T = \begin{bmatrix} P_k & P_{k-1} \\ \dots & \dots \\ P_{N+k} & P_{N+k-1} \end{bmatrix};\ Y_T = \begin{bmatrix} P_{k+a} \\ \dots \\ P_{N+k+a} \end{bmatrix} \tag{3.2}$$

where a=1 provides 1-day ahead, and a=2 provides 2-day-ahead coefficients. Using X and Y the expression is written in the matrix form

$$Y = X\ \alpha_N , \tag{3.3}$$

and consequently α_N is calculated by

$$\alpha_N = (X^T X)^{-1} X^T\ Y . \tag{3.4}$$

Once α_N is determined, the RMS error for training and verification are

$$Y_{NT} = X_T\ \alpha_N , \text{ and } \quad Y_{NV} = X_V\ \alpha_N , \tag{3.5}$$

The estimation errors for training and verification are

$$E_{RMS,NT} = \sqrt{\frac{1}{N}(Y_{NT} - Y_T)^T (Y_{NT} - Y_T)} \tag{3.6}$$

$$E_{RMS,NV} = \sqrt{\frac{1}{N}(Y_{NV} - Y_V)^T (Y_{NV} - Y_V)} \tag{3.7}$$

3.4 SDMA index regression

SDMA is an indicator of the short term price with an approximate time lag of three days. A prediction of a-day ahead price by linear regression requires the last two days prices together with SDMA

$$SDMA_k = (P_k + P_{k-1} + \dots + P_{k-5})/6 \tag{3.8}$$

$$P_{S,k+a} = \alpha_{S,1} \; P_k + \alpha_{S,2} \; SDMA_k \tag{3.9}$$

The coefficients vector $\alpha_S = (\alpha_{S,1} + \alpha_{S,2})$ is obtained from regression using the training data set, and the training and verification errors $E_{RMS,ST}$ and $E_{RMS,SV}$ are computed by similar calculations given for the Null Regression.

3.5 MACD index regression

MACD is easily computed by MatLab function macd(), which calculates it as explained in Chapter 2. Once the $MACD_k$ value of the day is available, a linear expression predicts the future price

$$P_{M,k+a} = \alpha_{M,1} \; P_k + \alpha_{M,2} \; MACD_k \tag{3.10}$$

Similar to Null and SDMA cases, the parameters $\alpha_M = (\alpha_{M,1} \;\; \alpha_{M,2})$ are obtained by regression using training data, and the errors $E_{RMS,MT}$ and $E_{RMS,MV}$ are computed by similar calculations given for the Null Regression.

3.6 RSI index regression

RSI index needs counting the loss and gain days, as well as computing the total loss of loss days and the total gain of the gain days over the last 14 days period. The function RSindex() in MatLab calculates the index RSI_k as described in Chapter 2.

The RSI index is a score of trend, rather than the price value. A linear expression with RSI

$$P_{R,k+a} = \alpha_{R,1} \; P_k + \alpha_{R,2} \, RSI_k \tag{3.11}$$

gives the a-day-ahead price prediction after calculating the coefficients

$$\alpha_R = (\alpha_{R,1} \; \alpha_{R,2})$$

by regression using training data set.

The errors $E_{RMS,RT}$ and $E_{RMS,RV}$ are computed in similar way as described for Null Regression.

3.7 Linear Prediction of all Three Indices

The expression

$$P_{A,k+a} = \alpha_{A,1} P_k + \alpha_{A,2} SDMA_k + \alpha_{A,3} MACD_k + \alpha_{A,4} RSI_k \quad (3.12)$$

calculates the a-day-ahead future price from all three indices. The parameter vector

$$\alpha_A = (\alpha_{A,1} \ \ \alpha_{A,2} \ \ \alpha_{A,3} \ \ \alpha_{A,4}) \quad (3.13)$$

is obtained using training data set by linear regression method, while the future price is calculated by

$$P_{A,k+a} = (P_k \ \ SDMA_k \ \ MACD_k \ \ RSI_k)\, \alpha_A \quad (3.14)$$

The errors $E_{RMS,AT}$ and $E_{RMS,AV}$ are computed in similar way as described for Null Regression.

3.8 TSK modelling of future prices

The TSK fuzzy modelling has two main sections: the training section to build a model with sufficient fuzzy TS rules that describes the input-output relations in the data, and the inference section that infers the value of output for a given input vector. This thesis propose to use the stock market indices SDMA, MACD, RSI and the price movement in the last two days as the input vector, and the price movement from the present day to the future day two indexes are chosen through stepwise. The null regression arguments P_k, and P_{k-1} are also included to these selected indexes in form of price movements $P_k - P_{k-1}$ and $P_{k-1} - P_{k-2}$ to predict the a-day-ahead price difference $P_{k+a} – P_k$. An input vector of TSK prediction model consists of five features:

$$x_k = (P_k – P_{k-1} \ \ P_k – P_{k-2} \ SDMA_k \ MACD_k \ RSI_k) \quad (3.15)$$

3.8.1 Structural Parameters of the TSK Model

A TSK model is constructed for a number of fuzzy rules. The number of rules nr depends on the number of clusters c in FCM. Furthermore, the fuzzification power m of FCM plays an important role in clustering the input-output observation vectors by determining the extent of fuzziness of the clusters.

3.8.2 Obtaining the Fuzzy Sets of the Rule

Next phase in TSK modelling is curve-fitting on the membership value of each training observation after projecting the point on the plane of an input feature vs. membership value. At this phase, TSK modelling requires the choice for one of the possible membership functions such as triangular, trapezoidal, or Gaussian. This thesis prefers Gaussian MF for two major reasons. Gaussian MF is well defined and non-zero over the whole discourse of universe of the input feature, and it is a simple function that can be easily fit on the projected FCM membership values for each cluster. The Gaussian function fits on the projected membership values by

$$\sigma_{i,j} = \sum_{k=1}^{N} \frac{1}{N} \sqrt{-\frac{(x_{k,j} - v_{i,j})^2}{2\log(u_{i,k})}} \quad (3.16)$$

where $x_{k,j} \in R$ is the j^{th} feature of the k^{th} observation among N training observations; the $v_{i,j} \in R$ is the j^{th} feature of the i^{th} cluster center; $u_{i,k}$ is the FCM membership of k^{th} observation in i^{th} cluster. The computed $\sigma_{i,j}$ defines the Gaussian MF of the fuzzy set $A_{i,j}$, corresponding to the i^{th} rule, j^{th} input feature i on the plane of membership value vs. j^{th} feature of input vector, N is total number of observations.

3.9 Computing the Coefficients of Consequent Expressions

The linear output expression for the i^{th} rule consists of one constant and nx coefficients.

$$z^*_i = f_i(x_k) = b_{i,0} + b_{i,1}\, x_{k,1} + b_{i,2}\, x_{k,2} + \ldots + b_{i,nx}\, x_{k,nx}\ . \tag{3.17}$$

The constant and the coefficients can be obtained from the observations by forming the homogeneous input matrix Xi, output vector Z_i, and coefficient matrix B_i for the i^{th} rule using the i^{th} cluster membership values $u_{i,k}$ of the observation (x_k,z_k)

$$X_i=\begin{bmatrix} u_{i,1} & u_{i,1}x_1 \\ u_{i,2} & u_{i,2}x_2 \\ \ldots & \ldots \\ u_{i,nt} & u_{i,ny}x_k \end{bmatrix};\ Z_i=\begin{bmatrix} u_{i,1}z_1 \\ u_{i,2}z_2 \\ \ldots \\ u_{i,nt}z_{nt} \end{bmatrix};\text{ and } B_i=\begin{bmatrix} b_{i,0} \\ b_{i,1} \\ \ldots \\ b_{i,nx} \end{bmatrix}, \tag{3.18}$$

so that the observations are written

$$\mathrm{Z}=X\, B_i \tag{3.19}$$

Accordingly, the least squares error solution for the coefficients are obtained by

$$B_i = (X_i^{\mathrm{T}} X_i)^{-1} X_i^{\mathrm{T}}\, Z_i \quad i=1 \ldots \mathrm{c}\ . \tag{3.20}$$

3.10 Predicted Output by Inference of input vector

The inference of the TS model calculates the *degree of fulfilment* β_i of each rule i=1 … c for the input vector $\mathrm{x}=(\mathrm{x}_1, \mathrm{x}_2, \ldots \mathrm{x}_{nx})$ to be used for prediction.

$$\beta_i(x) = A_{i,1}(\mathrm{x}_1) \wedge A_{i,2}(\mathrm{x}_2) \wedge \ldots \wedge A_{i,2}(\mathrm{x}_{nx}) \tag{3.21}$$

The degree of fulfilments of the rules provides the predicted value of output z as a weighted average of the predictions of each rule.

$$z^*(x)=\frac{\sum_{i=1}^{c} \beta_i(x)\, z_i^*(x)}{\sum_{i=1}^{c} \beta_i(x)} \tag{3.22}$$

3.11 Evaluation of the Prediction Performance

The prediction performance of the model is evaluated by the RMS error of the verification data set, which is obtained using the last half of the time series data by

$$E_{\mathrm{RMS,FV}} = \sqrt{\frac{1}{nz}(Z^*_{FV} - Z_{\mathrm{FV}})^{\mathrm{T}}(Z^*_{FV} - Z_{\mathrm{FV}})}\ . \tag{3.23}$$

Where *nz* is number of verification vectors which are employed to get predictions Z_{FV}^{*}

3.12 Selection and Determination of Significance of Features

The input variables, which are called features of fuzzy model, are of prime importance for the success of the model to have sufficiently low RMS errors of predictions. The feature selection method used in this study is based on overall performance of the fuzzy model by testing one-missing and one-added features.

The proposed procedure in this thesis is based on testing all combinations of the input variables to determine the significance of the variables, as well as to simplify the model by eliminating the non-contributing inputs from the TSK model.

Chapter 4

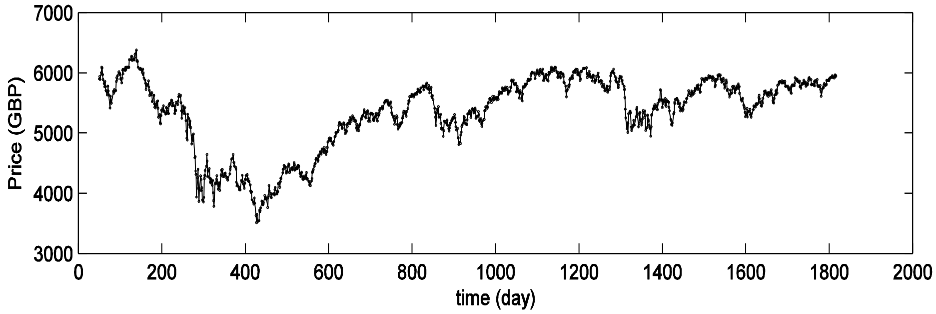
THE RESULTS OF FORECASTING USING THE PROPOSED MODEL

The proposed SMDM is applied on London Stock Market data set which is collected from finance section of *www.yahoo.com* web site and listed in Appendix. This Chapter focuses on i) performances of prediction by only stock market indices SDMA, MACD, and RSI, ii) performance improvement of pre-processing, iii) performance of fuzzy model compared to the stock market indices, iv) feature selection and feature significance determination, and finally v) the graphical representation of the model with the best RMS error.

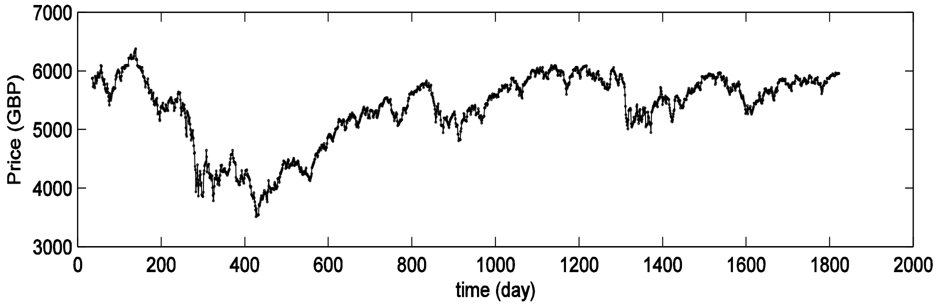
The training and verification time series data are over 900 points, which makes the point wise graphical representation impossible because of the large uncertainty in the data set. This thesis overcomes it by plotting a sample of training and test graphs instead of the full range plots. The sampled plots contain 100 days from 950 days training and 950 days verification data. Trading in a typical stock market through a bank takes almost one day, and therefore all predictions were carried over two-day-ahead prices instead of the usual one day ahead forecasting. Figure 4.1 a, b, c, d show the raw and pre-processed data sets, in complete range, and in the sampled period from the day-600 to the day-700. Table 4.1 (a) displays the parameters for linear prediction as a result of regression with training data set. The 5-year time series data has closing price properties (minimum=3512, maximum=6479, average=5365) GBP.

4.1 Forecasting by only Technical Indices SDMA, MACD, RSI

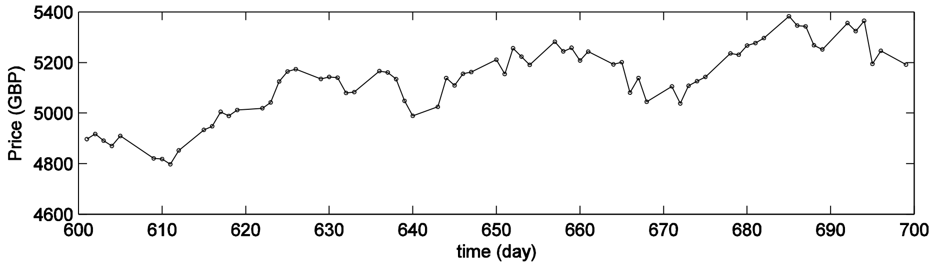
The errors of two-day-ahead forecasted prices using the indices SDMA, MACD, and RSI are shown in Table 4.1.



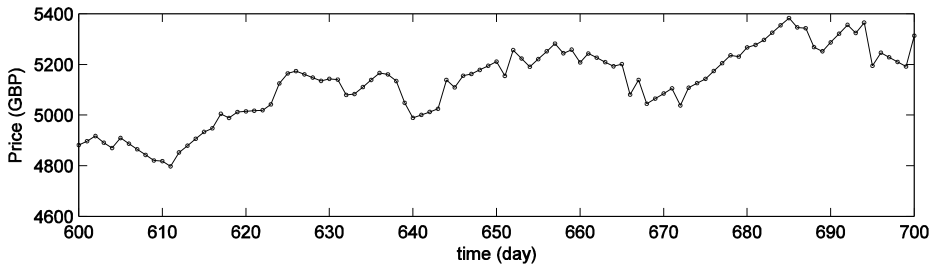
a)



b)



c)



d)

Figure 4.1 London Stock Market prices from 2-1-2008 to 29-12-2012
(a) raw data without completing the missing dates, (b) pre-processed data after completing the missing dates by interpolation. (c) Sample of raw data and from day-600 to day-700 (d) Sample of completed data from day-600 to day-700.

Table 4.1 RMS Errors of Linear Estimations using Technical Indices for raw data

	Indices	Regression Coefficients	RMSE training	RMSE verification
	None (by last two prices)	(0.8769 0.1224)	115.855	87.7402
	SDMA + last day price	(0.8727 0.1266)	115.676	**87.1570**
	MACD + last day price	(0.9993 -0.0075)	116.313	87.4836
	RSI + last day price	(1.0001 -0.0734)	116.301	87.4435
	All indices together+last day	(0.7746 0.2195 0.0143 0.4981)	**115.307**	87.5284

Table 4.2 RMS Errors using Technical Indices for pre-processed data

#	Indices	Regression Coefficients	RMSE training	RMSE verification
1	None (by last two prices)	(1.0710 -0.0715)	91.4520	**68.0317**
2	SDMA + last day price	(0.9615 0.0380)	91.5129	68.1659
3	MACD + last day price	(0.9995 -0.0045)	91.5625	68.1096
4	RSI + last day price	(0.9987 0.0715)	91.5391	68.2451
5	All indices together+last day	(0.8953 0.1014 -0.0102 0.2655)	**91.3384**	68.6755

4.2 Effect of the Missing Days on Prediction Performance

Table 4.1 and Table 4.2 displays an apparent benefit of the pre-processing. Accordingly, from this point on the computational efforts are focused on pre-processed data set. The prediction performance of linear model with all three indices after interpolating the missing days is improved from

$$\eta = \frac{5365 - 87.5284}{5365} = 98.37\ \% \qquad (4.1)$$

to

$$\eta = \frac{5365 - 68.6755}{5365} = 98.72\ \% \qquad (4.2)$$

with a percentage of 0.35%, which corresponds to 27.3% reduction in RMS error.

Table 4.2 also points that the last two day prices are as important as the technical indices in predicting the future prices. The over-fitting of model-5 to the crisis-period notch (days 200 to 600) in the training data set may be the reason of the considerably high verification RMS error.

Figure 4.2 and Figure 4.3 (b), (c) displays the calculated values of technical indices SDMA, MADC, and RSI along a sample of closing prices. SDMA (plot-a) has a smooth curve with a small amount of lag compared to the closing prices. MACD and RSI are indicators of trend rather than the price value.

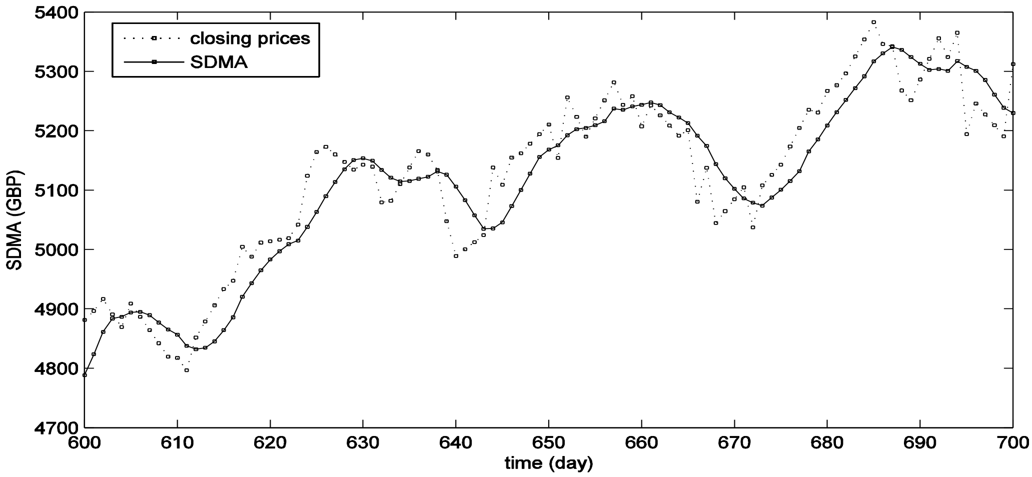
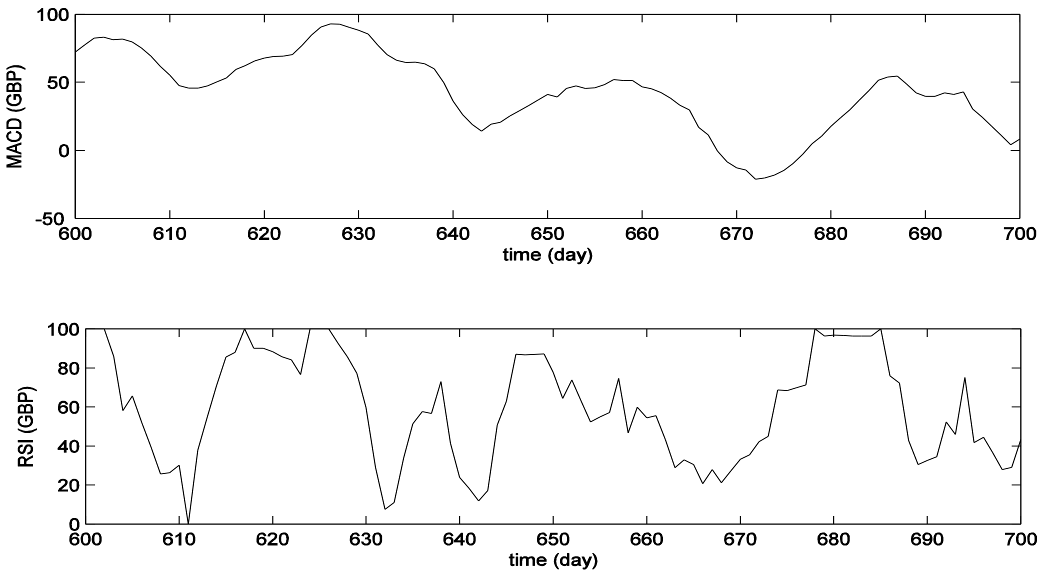


Figure 4.2 Plot of SDMA,



(a)

(b)

Figure 4.3 Plots of (a) MADC, (b) RSI values for a sample of data.

As seen in Figure 4.3, the Relative Strength Index is a percentage that changes between 0 and 100, and in usual practice an equity is preferred to buy when the index exceeds 80%.

4.3 Prediction Performance of TSK Model

This section compares RMS errors of prediction by TSK against the prediction error of technical indices (SDMA, MACD, RSI). TSK model is generated by using five input variables

$$x_k = [P_k – P_{k–1} \;\; P_k – P_{k–2} \;\; SDMA_k \;\; MACD_k \;\; RSI_k\,] \qquad (4.3)$$

and one output variable

$$\mathrm{yk} = P_{k+a} – P_k\,, \qquad (4.4)$$

where a=2 provides two-day-ahead prediction of TSK model.

Clustering a data set with the original ranges of observation variables may result in unexpected distribution of the cluster centres. Normalization of the input-output data vectors scales the vectors into a unit hyper-square, and similarity of the vectors are scored after this scaling. It makes the cluster centres distributed among all variables rather than along the input variable with the largest range, which is in this case the SDMA column.

For best performance, a set of TSK models with number of cluster from 2 to 12 were generated using MatLab Fuzzy Toolbox after correcting the rule-extraction section of the internal MatLab code. These models were built using the training partition of the 5-year time series data set with and without normalization.

RMS errors of TSK models without normalization of the data set are shown in Table 4.3. In the table, the six-rule TSK (c=6) gives the best model that has both training and verification error the lowest among the tested c values. The corresponding Fuzzy Rule Base for c=6 is plotted in Figure 4.4.

Table 4.3 RMSE values by TSK without normalization of features

Method	RMSE training	RMSE verification
All indices together	91.33847	68.67550
TS, c=2	90.888	68.8181
TS, c=3	91.0149	69.191
TS, c=4	90.9532	68.6699
TS, c=5	90.9881	68.0841
TS, c=6	**90.5881**	**68.257**
TS, c=7	90.7222	68.3209
TS, c=8	90.9628	68.6347
TS, c=9	91.0001	68.1932
TS, c=10	91.1865	68.311
TS, c=11	91.1867	68.4423
TS, c=12	91.479	68.5519

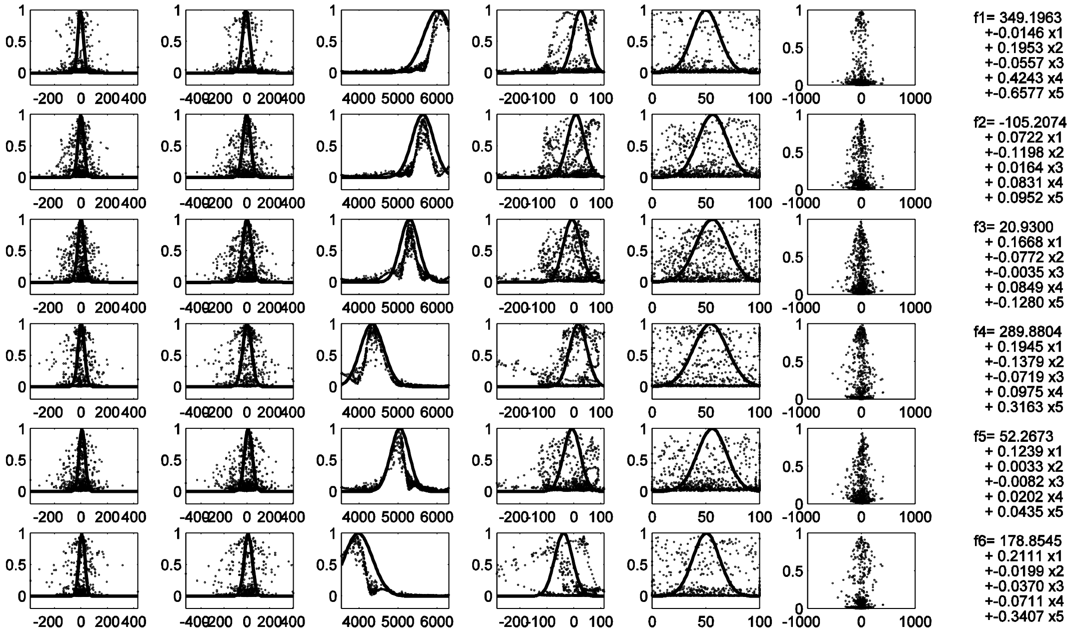


Figure 4.4 The rule base of TSK c=6 without normalization

Table 4. 4 TSK Rule Base Parameters of input fuzzy sets

rule:	1		2		3		4		5		6	
inp	s	c	s	c	s	c	s	c	s	c	s	c
1	20.11	-5.052	22.24	-2.176	24.45	-1.311	24.76	0.1829	22.30	6.399	22.25	5.463
2	30.57	-9.498	33.32	-3.085	37.05	-1.133	38.85	0.6142	34.34	10.42	34.33	9.300
3	363.0	6031.	276.9	5638.	233.4	5296.	300.5	4337.	235.9	5050.	364.1	3981.
4	28.61	24.74	31.38	7.14	31.94	-7.151	32.49	16.13	31.06	-7.474	32.99	-38.62
5	12.77	49.93	13.48	56.061	14.87	55.62	15.79	54.09	13.65	55.74	12.40	50.27

Table 4.5 TSK Rule Base Parameters of output expressions

Coeff: Rule# *i*	$b_{i,0}$	$b_{i,1}$	$b_{i,2}$	$b_{i,3}$	$b_{i,4}$	$b_{i,5}$
1	349.1963	-0.0146	0.1953	-0.0557	0.4243	-0.6577
2	-105.2074	0.0722	-0.1198	0.0164	0.0831	0.0952
3	20.93	0.1668	-0.0772	-0.0035	0.0849	-0.128
4	289.8804	0.1945	-0.1379	-0.0719	0.0975	0.3163
5	52.2673	0.1239	0.0033	-0.0082	0.0202	0.0435
6	178.8545	0.2111	-0.0199	-0.037	-0.0711	-0.3407

The normalization of data set is obtained by scaling and shifting each variable (x) using the minimum (x_{min}) and maximum (x_{max}) of that variable along all training observations.

$$x_n=(x-x_{min}) / (x_{max}-x_{min}). \qquad (4.5)$$

After the modeling, the predicted values are denormalized by

$$y= y_{min} + y_n (y_{max}-y_{min}). \qquad (4.6)$$

The training and verification RMS errors for the TSK Model with normalization are listed in Table 4.6. Considering that higher number of rules have higher risk of over fitting, the model with 6 rules appear to be the lowest error model with balanced training and verification error values.

Table 4.6 RMSE values in TSK with normalization

Method	RMSE train	RMSE verif
All indices together	91.33847	68.67550
TS, c=2	90.5448	**68.3445**
TS, c=3	90.4009	68.3849
TS, c=4	90.192	68.367
TS, c=5	90.1439	68.4931
TS, c=6	**90.1302**	**68.3439**
TS, c=7	90.1759	68.4035
TS, c=8	90.1289	68.4759
TS, c=9	90.0965	68.4681
TS, c=10	90.1063	68.5021
TS, c=11	**90.0579**	68.3967
TS, c=12	90.5206	68.4829

4.4 Significance of each input variable

Table 4.7 contains the training and verification RMS errors obtained for 6-cluster models with only one added feature, and with only one missing features. Tests with one-added-feature (1 2, 1 3, 1 4, 1 5) shows that, none of the technical indices dominate in predicting closing prices better than others, and the previous day price (input-2) is more informative than any individual of the three technical indices. However, the one-missing-feature tests (1345 1245 1235 1234) indicate that when all three technical indices are involved in prediction, the previous day price has disturbing effect on prediction.

Table 4.7 RMSE values in TSK with normalization

Method	RMSE train	RMSE verify
All indices together	91.33847	68.67550
TS, c=6, features 1 2 3 4 5	**90.5881**	**68.2570**
TS, c=6, features 1 2	93.4763	72.9736
TS, c=6, features 1 3	**90.7601**	**68.0068**
TS, c=6, features 1 4	95.9887	71.0612
TS, c=6, features 1 5	96.7880	73.5772
TS, c=6, features 1 2 3	90.5196	68.0668
TS, c=6, features 1 2 4	93.4646	71.3128
TS, c=6, features 1 2 5	93.7798	72.8904
TS, c=6, features 1 3 4	**91.0398**	**68.0678**
TS, c=6, features 1 3 5	90.8098	68.0882
TS, c=6, features 1 4 5	94.8400	71.9937
TS, c=6, features 1 2 3 4	90.7550	68.1503
TS, c=6, features 1 2 3 5	90.5549	68.4081
TS, c=6, features 1 2 4 5	92.9171	71.2290
TS, c=6, features 1 3 4 5	**91.0375**	**68.0801**

In conclusion, models with features (1, 3, 4, 5) and (1, 3, 4) provide highest reduction in verification error while keeping the training error low as well. RSI looks like the least significant among three technical indices, and SDMA is the most significant in reducing the prediction error.

Accordingly, the model with reduced features is obtained as shown in Figure 4.5, using the observation vectors and the RMS error of predicted prices

$$x_k = [P_k - P_{k-1} \;\; SDMA_k \;\; MACD_k \;] \qquad (4.7)$$

and one output variable

$$yk = P_{k+a} - P_k \,, \qquad (4.8)$$

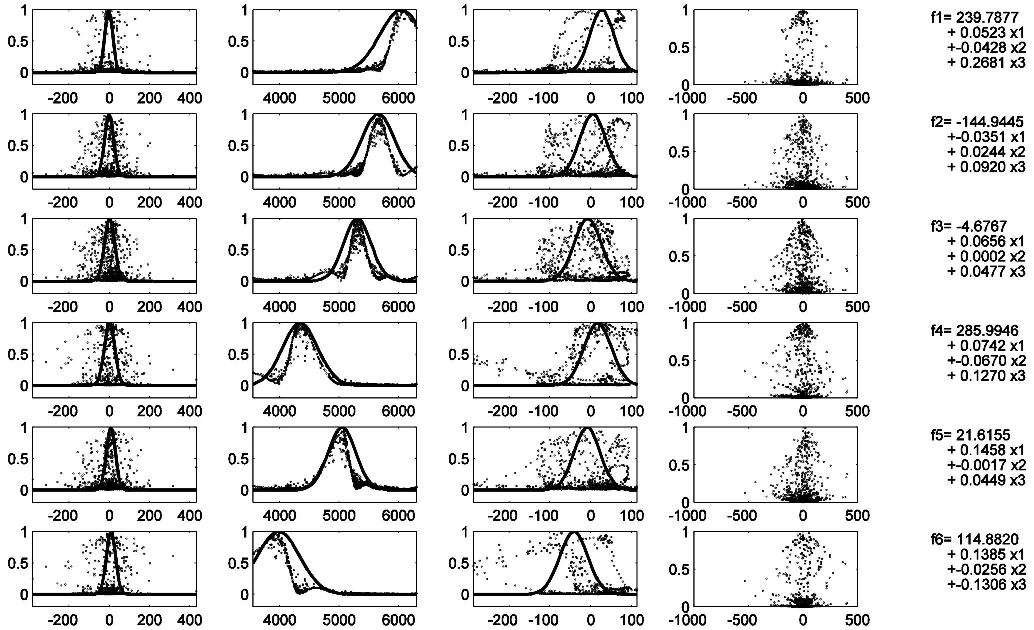


Figure 4.5 The rule base of TSK c=6 without normalization

The rules of the reduced model is shown in Figure 4.5, and the predicted 2-day-ahead prices for the training (from day-600 to day-700) and verification (from day-1500 to day-1600) sample periods are plotted in Figure 4.6 and Figure 4.7.

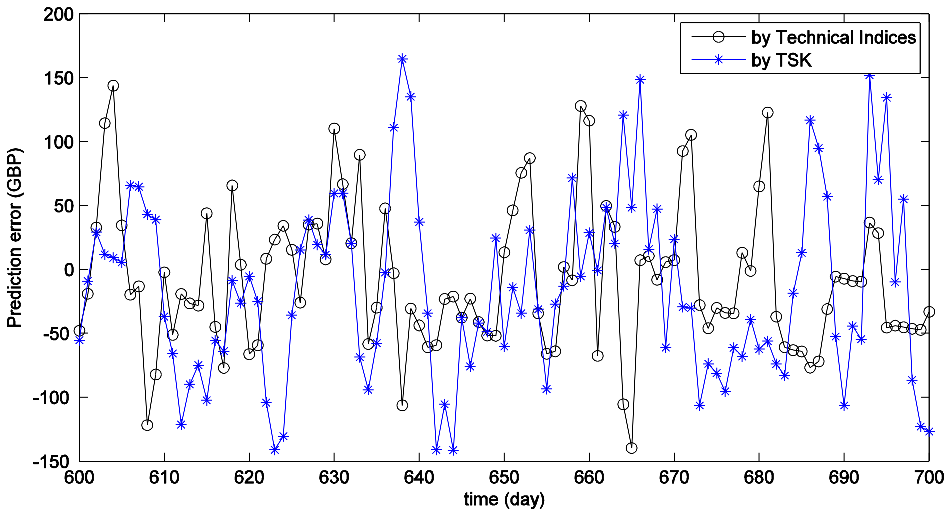


Figure 4.6 The prediction error in a sample of training data for reduced model

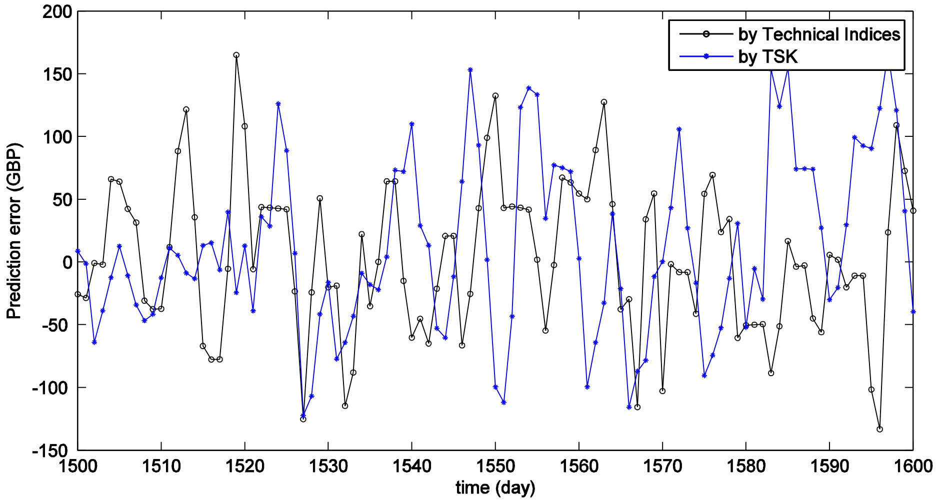


Figure 4.7 The prediction error in a sample of verification data for reduced model

The input membership function parameters, and the output expression coefficients of the rules are shown in Table 4.8 and 4.9.

Table 4.8 TSK Rule Base Parameters of input fuzzy sets of reduced model

rule:	1		2		3		4		5		6	
inp	s	c	s	c	s	c	s	c	s	c	s	c
1	20.00	-5.019	22.49	-3.001	25.52	-0.8958	26.14	-0.9429	22.48	5.988	22.90	6.308
3	357.8	6042.	273.3	5653	230.2	5303.	296.6	4339.	231.3	5044.	359.4	3974.
4	28.84	26.75	31.79	4.585	33.14	-7.777	33.68	16.70	31.66	-9.277	34.35	-41.84

Table 4.9 TSK Rule Base Parameters of output expressions of reduced model

Coeff: Rule# i	$b_{i,0}$	$b_{i,1}$	$b_{i,3}$	$b_{i,4}$
1	0.0523	-0.0428	0.2681	239.7877
2	-0.0351	0.0244	0.092	-144.945
3	0.0656	0.0002	0.0477	-4.6767
4	0.0742	-0.067	0.127	285.9946
5	0.1458	-0.0017	0.0449	21.6155
6	0.1385	-0.0256	-0.1306	114.882

The reduced model has verification RMS error 68.0678, which is 0.169 less than the verification error of linear regression with all technical indices (=68.257). It corresponds approximately to 0.2% reduction of the RMS error. With this reduction of the error, the approximate success of prediction becomes

$$\eta=(5385 - 68.0688)/5385= 98.74\% .$$

The improvement looks like small, but the positive contribution of the fuzzy modeling is apparent in Figure 4.5 as smaller error compared to the linear method.

Chapter 5

CONCLUSION

This study introduces a new method to improve forecasting accuracy of technical indices using a TSK fuzzy model. The proposed method is tested on London stock market time series data set from 2008 January to 2012 December.

According to the results of this research, the pre-processing of the time series data set to complete the closing prices of the missing dates significantly improved the prediction accuracy. The reduction of the RMSE error in verification data is around 25%.

Although the clustering of the normalized data set was expected to distribute the cluster centres along the ranges of all variables, the results indicated there is a very minor difference between the normalized and non-normalized models. In both cases, the lowest error figures were obtained at 6 clusters that gives a 6-rule prediction model.

Future work: This thesis accomplished an implementation of TSK fuzzy model, which can be expanded to include several additional features that may reduce the uncertainty. The techniques introduced here may be applied to other time series data sets. The stock market prices are almost randomly changing time series and prediction of this high uncertainty is beyond the capability of the TSK model. It is expected that markets with less uncertainty may give better performance increase by the TSK model, as well as other methods such as type-2 fuzzy may help to predict the variance of the future error together with the future price.

REFERENCES

[1] H. B. Nielsen, “Introduction To Time Series,” *Econometrics 2,* pp. 1-15, 2004.

[2] J. D. Hamilton , Time Series Analysis, Princeton University Press, 1994.

[3] Kolarik and G. Rudorfer, “Time series Forecasting Using Neural Networks, Department Of Applied,” *Vienna University Of Economics and Business Administration,* no. 1090, p. pp. 2–6, 1997.

[4] R. Balvers, . T. Cosimano and B. McDonald, “Predicting Stock Returns In An Efficient Market,” *Journal of Finance,* vol. 55, p. 1109–1128, 1990.

[5] J. Jang , C. Sun and . E. Mizutani, Neuro-Fuzzy and Soft Computing: A Computational Approach, NJ: Prentice-Hall, 1997.

[6] A. Esfahanipour and W. Aghamiri, “Adapted Neuro-Fuzzy Inference System on Indirect Approach TSK fuzzy rule base for stock market analysis,” *Expert Systems with Applications,* no. 37, p. 4742–4748, 2010.

[7] P.C. Chang and C.-H. Liu, “A TSK type fuzzy rule based system for stock price prediction,” *Expert Systems with Applications,* no. 34, pp. 135-144, 2008.

[8] C.-H. Cheng, H. J. Teoh and T.-L. Chen, “Forecasting Stock Price Index Using Fuzzy Time-Series Based on Rough Set,” *IEEE,* 2007.

[9] M. Männle, “Identifying Rule-Based TSK Fuzzy Models”,in Proc. EUFIT, Aachen, Germany,Sep.1999.

[10] V. Vaidehi, S. Monica and M. S. Safeer, “A Prediction System Based on Fuzzy Logic,” *World Congress on Engineering and Computer Science,* 2008.

[11] M. Bodur, A. Acan, T. Akyol, , “Fuzzy system modeling with the genetic and differential evolutionary optimization,” in *Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, vol. 1, pp. 432 - 438,,* Vien, 2005.

[12] Bodur M., Ahmederaghi B, “A comparison of fuzzy functions with lse and ts-fuzzy methods in modeling uncertain dataset,” in *Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control, ICSCCW 2009. 5th International Conference,* Famagusta, 2009.

[13] “YAHOO! Finance,” 10 1 2013. [Online]. Available: finance.yahoo.com/..

[14] B. Thierry , . T. Philippe and U. Michael , “Linear Interpolation Revitalized,” *IEEE Transactions On Image processing,* vol. 13, no. 5, pp. 710-719, 2004.

[15] J. Murphy, Technical Analysis of the Futures Market, Haarlem , 1991.

[16] G. Appel, Technical Analysis Power Tools for Active Investors, Financial Times Prentice Hall, 1999, p. 166.

[17] D. C. Montgomery, Introduction to Statistical Quality Control, New York: John Wiley and Sons, 1991.

[18] E. Seykota, ““MACD: Sweet Anticipation?,” *Technical Analysis of Commodities and Stocks,* 1991.

[19] I. Copsey, “The Principles of Technical Analysis,” *Dow Jones Telerate,* 1993.

[20] J. W. Wilder, New Concepts in Technical Trading Systems, 1st Ed., Business & Economics, 1978.

[21] L. A. Zadeh, “Fuzzy sets,” *Information and Control* , no. 8, pp. 338-353, 1965.

[22] L. A. Zadeh, “A rationale for Fuzzy Control,” *Journal of Dynamic Systems, Measurement, Control 94,* no. 6, pp. 3-4, 1972.

[23] J. J. Buckley, “Universal fuzzy controllers,” no. 28, pp. 1245-1248, 1992.

[24] R. Nock and . F. Nielsen, “On Weighting Clustering,” *IEEE Transactions On Pattern Analysis And Machine Intelligence*, vol. 28, no. 8, pp. 1-13, 2006.

[25] U. Kaymak and . M. Setnes, “Extended Fuzzy Clustering Algorithms,” *Erim Report Series Research In Management,* pp. 1-24, 2000.

[26] T. Takagi and M. Sugeno, “Fuzzy Identification Of Systems and its Applications To modeling and control,” *IEEE Trans. On Systems, Man and Cybernetics*, no. 15, pp. 116-132, 1985.

[27] Abu-Ghoush, M., M. Samhouri, E. Yaseen and T.. Herald, 2010. Fuzzy clustering modeling and evaluation of surface interactions and emulsions of selected wheat protein combined to Iota-Carrageenan and gum Arabic solutions.J. Food Proccess Eng., 33:133-141.

APPENDICES

Appendix 1: The thesis code

```
function A = M0504
clc; clear('all'); format('compact'); sP=@sprintf ;
close('all'); CompleteP=1;
% read raw data
DP = LondonSM; nP=size(DP,1); % Raw Prices
% complete missing days rand
DC=completemissing(DP,nP);
if(CompleteP), DR=DC; grts=600;grvs=1500; grr=100;
else DR=DP; grts=416;grvs=1070; grr=69; end
grt=[grts:grts+grr]; grv=[grvs:grvs+grr];
% DP=alldata;
Dve=size(DR,1)-4;  %2 % with missing days
Dts=35; Dte=floor(Dts+(Dve-Dts)*0.5);  Dvs=Dte+1;
nkT=Dte-Dts+1; nkV=Dve-Dvs;
kA=[1:Dve+4]; kAA=[Dts-4:Dve+4];
kT=[Dts:Dte]; kV=[Dvs:Dve]; kTV=[kT,kV];
%=================================================
days=2; normalization=0; % gr options: kTV,kV,kT,grt,grv
gr=grt; gra='  EF RF'; gri=6; fig=1;% PL+ EA MA RA
%gra options PL SA MA RA EA EF
%=================================================
% x = -pi:pi/20:pi;
% plot(x,cos(x),'-ro',x,sin(x),'-.b')
% h = legend('cos_x','sin_x',2);
% set(h,'Interpreter','none')
D=DR(kA,2);
if(strf(gra,' PL')>0),figure(fig); fig=fig+1;
plot(DR(gr,1),D(gr),'-ok','Markersize',2);
```

```
ylabel('Price (GBP)'); xlabel('time (day)');
h=legend('closing prices','SDMA');
set(h,'Interpreter','none');
%set(gcf,'Position', [20  50  800  200] );
pause(0.5);  end
% Calculation of SDMA (six-day-simple-moving-average)
SA(:,1)= filter(ones(1,6)/6,[1],D(kA)); ST=SA(kT); SV=SA(kV);
% Calculation of MACD (Moving Av. Convergence Divergence)
if(strf(gra,' SA')>0),figure(fig); fig=fig+1;
plot(DR(gr,1),D(gr),':sk','Markersize',2); hold on;
plot(DR(gr,1),SA(gr),'-sk','Markersize',2);
ylabel('SDMA (GBP)'); xlabel('time (day)');
set(gcf,'Position', [20  50  800  400] );
pause(0.5);  end
MA= macd(D(kA)); MA(1:25)=0; MT=MA(kT); MV=MA(kV);
% Calculate 14 days RSI index
if(strf(gra,' MA')>0),figure(fig); fig=fig+1;
plot(DR(gr,1),MA(gr),'-k');
ylabel('MACD (GBP)'); xlabel('time (day)');
set(gcf,'Position', [20  50  800  400] );
pause(0.5); figure(fig); end
RA = rsindex(D(kA),6); RT=RA(kT); RV=RA(kV);
if(strf(gra,' RA')>0),figure(fig); fig=fig+1;
plot(DR(gr,1),RA(gr),'-k');
ylabel('RSI (GBP)'); xlabel('time (day)');
set(gcf,'Position', [20  50  800  400] );
pause(0.5); end
[aN,aS,aM,aR,aA,EN,ES,EM,ER] ...
=trainverifNMSR(D,SA,MA,RA,kT,kV,kA,days);
```

```
% TS modeling
disp('TSK RMS errors for training and verification.');
kA2=kA(10:end); kTS=kT-Dts+1; kVS=kV-Dts+1;
PM=[zeros(10,2) ; [D(kA2) MA(kA2)]]*aM;
PS=[zeros(10,2) ; [D(kA2) SA(kA2)]]*aS;
PR=[zeros(10,2) ; [D(kA2) RA(kA2)]]*aR;
PA=[zeros(10,4) ; [D(kA2) SA(kA2) MA(kA2) RA(kA2)]]*aA;
DTV =[D(kTV)-D(kTV-1) D(kTV)-D(kTV-2) ...
SA(kTV) MA(kTV) RA(kTV) D(kTV+days)-D(kTV) ];
DTT =DTV(1:nkT,:); DTV=DTV(nkT+1:nkT+1+nkV,:);
DDX= [ 0 0 0 0 0 0; 1 1 1 1 1 1]; % for w/o normalization
% 3 2 4 1 5
% Normalization of data set
[NTPar]= normalparam([DDX]);
if(normalization),[NTPar]= normalparam([DTT]); end
DNT=normalize(DTT,NTPar); DNV=normalize(DTV,NTPar);
% Feature selection
FS=[1 2 3 4 5];nOut=6; OS=[nOut]; %
%FS=[ 3 2 4 1 5 ];nOut=6; OS=[nOut]; %
disp([ sprintf('completed=%d #days=%d norm=%d',
CompleteP, days,normalization)]);
disp([ 'FS:' sprintf(' %d ',FS)]);
% TS Modeling
disp('  clusters    test     verif');
for nc=2:12
FMT =genfisTS(DNT(:,FS),DNT(:,OS), ...
'sugeno',nc,[2,100,0,0]);
RandStream.setDefaultStream(...
RandStream('mcg16807', 'Seed',0));
```

```
[Vv, Uu] = fcm( [DNT(:,FS) DNT(:,OS)], ...
nc, [2,100,0,0]);
[nR,nV]=size(Vv);
% sort cluster centers ascending output order
[V,Iv] = sortrows(Vv,nV); U=Uu; % Uu=zeros(lenDt,numC);
for ir=1:nR, U(ir,:) = Uu(Iv(ir),:); end
%PLOTC plots a complete Fuzzy Rule-Base
if(nc==gri)&&(strf(gra,' RF'))
figure(fig); fig=fig+1;
plotc( FMT, [DNT(:,FS) DNT(:,OS)], U(:,:))
dispruleparam(FMT);end
% get predicted prices for normalized prices
PNT=evalfis(DNT(:,FS),FMT); %3
% denormalization of prices and getting error
PFT =denormalize(PNT,NTPar(:,nOut))+D(kT);
EFT=PFT-D(kT+days);
ERFT=  sqrt(mean(EFT.^2));
% get error for validation
PNV=evalfis(DNV(:,FS),FMT);
% denormalization of prices and getting error
PFV=denormalize(PNV,NTPar(:,nOut))+D(kV);
EFV=PFV-D(kV+days);
ERFV=  sqrt(mean(EFV.^2));
EF=[EFT;EFV];
if( (strf(gra,' EF')>0) && (gri==nc) ),
figure(fig); fig=fig+1;
plot(DC(gr,1),EF(gr),'-ok'); hold on;
plot(DC(gr,1),PA(gr)-D(gr+days),'-*b'); hold on;
ylabel('Prediction error (GBP)');
```

```
xlabel('time (day)');
legend('by Technical Indices','by TSK');
set(gcf,'Position', [20  50  800  400] );
pause(0.5); end;
disp([ sP('%10d %10.4f %10.4f',nc, ERFT, ERFV) ]);
end
return
%=============================================
%=============================================
function [PN]=normalparam(X)
PN(1,:)=min(X); PN(2,:)= max(X); PN(3,:)=  PN(2,:)-PN(1,:)  ;
return
function [XN]=normalize(X,PN)
n=size(X,1); XN= (X-ones(n,1)*PN(1,:))./(ones(n,1)*PN(3,:)) ;
return
function [X]=denormalize(XN,PN)
n=size(XN,1);X= ones(n,1)*PN(1,:) +
XN.*(ones(n,1)*PN(3,:)) ;
return
function [DC] = completemissing(D,nP)
k=1;
for i=1:nP
while(k<D(i,1)),
DC(k,:)=D(i-1,:)+(D(i,:)-D(i-1,:))...
    *(k-D(i-1,1))/(D(i,1)-D(i-1,1));
k=k+1;end
if(k==D(i,1)), DC(k,:)=D(i,:); k=k+1;end
end
return
```

```
function [aN,aS,aM,aR,aA,EN,ES,EM,ER]= ...
trainverifNMSR(D,SA,MA,RA,kT,kV,kA,days)
kTV=[kT,kV];
sP=@sprintf ;
disp('====  train      verif      coeffs.  ');
% Null Test
aN=[ D(kT) D(kT-1)]\D(kT+days);
EN=[D(kTV) D(kTV-1)]*aN -D(kTV+days);
ENT=[D(kT) D(kT-1)]*aN -D(kT+days);
ERNT= sqrt(mean(ENT.^2));
ENV=[D(kV) D(kV-1)]*aN -D(kV+days);
ERNV= sqrt(mean(ENV.^2));
disp(['RMSE.N=' sP(' %10.5f ', ERNT, ERNV) ...
'  aN=' sP(' %10.4f', aN )  ]);
% Calculation of Coefficients for
%Pmacd(k)=[P(k-1 P(k-2)MACD(k-1)]*[A1 A2 A3] ;
aM=[ D(kT) MA(kT) ]\D(kT+days);
EM=[D(kTV) MA(kTV)]*aM -D(kTV+days);
EMT=[D(kT) MA(kT)]*aM -D(kT+days);
ERMT= sqrt(mean(EMT.^2));
EMV=[D(kV) MA(kV)]*aM -D(kV+days);
ERMV= sqrt(mean(EMV.^2));
disp(['RMSE.M='sP(' %10.5f ', ERMT, ERMV)  ...
'  aM=' sP(' %10.4f', aM )  ]);
% Predicting Prices using SDMA by a linear expression
% Psdma(k) = [P(k-1) P(k-2) SDMA(k-1)][a1 a2 a3];
aS=[ D(kT) SA(kT)]\D(kT+days);
ES=[D(kTV) SA(kTV)]*aS -D(kTV+days);
EST= [D(kT) SA(kT)]*aS -D(kT+days);
ERST= sqrt(mean(EST.^2));
```

```
ESV= [D(kV) SA(kV)]*aS -D(kV+days);
ERSV= sqrt(mean(ESV.^2));
disp(['RMSE.S= ' sP(' %10.5f', ERST, ERSV) ...
'  aS=' sP(' %10.4f', aS )  ]);
% predicting Prices using RSI by a linear expression
% Prsi(k) = [P(k-1) P(k-2) RSI(k-1)][a1 a2 a3];
aR=[ D(kT) RA(kT)]\D(kT+days);
ER=[D(kTV) RA(kTV)]*aR -D(kTV+days);
ERT= [ D(kT) RA(kT)]*aR-D(kT+days);
ERRT= sqrt(mean(ERT.^2));
ERV= [ D(kV) RA(kV)]*aR-D(kV+days);
ERRV= sqrt(mean(ERV.^2));
disp(['RMSE.R=' sP(' %10.5f ', ERRT, ERRV) ...
'  aR=' sP(' %10.4f', aR )  ]);
%  Predicting Prices using ALL by a linear expression
% Pall(k) = a1*P(k-1)+a2*P(k-2)+a3*S(k-1)+a4*M(k-1)+a5*R(k-1);
aA=[ D(kT) SA(kT) MA(kT) RA(kT)]\D(kT+days);
EAT= [ D(kT) SA(kT) MA(kT) RA(kT)]*aA-D(kT+days);
ERAT= sqrt(mean(EAT.^2));
EAV= [ D(kV) SA(kV) MA(kV) RA(kV)]*aA-D(kV+days);
ERAV= sqrt(mean(EAV.^2));
disp(['RMSE.A= ' sP(' %10.5f ', ERAT, ERAV) ...
'  aA=' sP(' %10.4f', aA ) ]);
return
function fismat = genfisTS( ...
X, Y, fistype, nC, fcmoptions)
%GENFIS3 Generates a FIS using FCM clustering
%FIS = GENFIS3(XIN, XOUT,TYPE,CLUSTER_N, FCMOPTIONS)
%  allows you to specify options for the FCM algorithm.
%   For TSK type may be 'sugeno' or 'mamdani'
```

```
if nargin < 4,
disp('X, Y, fistype, nC required.'); end
if nargin < 5, fcmoptions = []; end
mftype = 'gaussmf'; % only option
% Check fistype
fistype = lower(fistype);
if ~isequal(fistype, 'mamdani') ...
&& ~isequal(fistype, 'sugeno')
disp('Unknown fistype specified.'); end
RandStream.setDefaultStream(...
RandStream('mcg16807', 'Seed',0));
[Vv, Uu] = fcm([X Y], nC, fcmoptions);
[nR,nV]=size(Vv); U=Uu;
% sort cluster centers in ascending order of the output column
[V,Iv] = sortrows(Vv,nV); % Uu=zeros(lenDt,numC);
for ir=1:nR, U(ir,:) = Uu(Iv(ir),:); end
% Check Xin, Xout
numX = size(X,2); numY = size(Y,2);
% Initialize a FIS
theStr = sprintf('%s%g%g',fistype,numX,numY);
fismat = newfis(theStr, fistype);
% Loop through and add inputs
for i = 1:1:numX
fismat = addvar(fismat,'input', ...
['in' num2str(i)],minmax(X(:,i)'));
% Loop through and add mf's
for j = 1:1:nC
params = computemfparams(mftype, ...
X(:,i), U(j,:)', V(j,i));
```

```
fismat = addmf(fismat,'input', i, ...
['in' num2str(i) 'cluster' num2str(j)], ...
mftype, params); end;  end
switch lower(fistype)
case 'sugeno'
% Loop through and add outputs
for i=1:1:numY
fismat = addvar(fismat,'output', ...
['out' num2str(i)],minmax(Y(:,i)'));
% Loop through and add mf's
for j = 1:1:nC
%MB correction
%MB params = computemfparams('linear', ...
%MB     [Xin Xout(:,i)] );
params = computemfparams ('linear', ...
[X Y(:,i)],U(j,:)');
fismat = addmf(fismat,'output', i, ...
['out' num2str(i) 'cluster' num2str(j)], ...
'linear', params); end;  end
case 'mamdani'
% Loop through and add outputs
for i = 1:1:numOutp
fismat = addvar(fismat,'output', ...
  ['out' num2str(i)],minmax(Y(:,i)'));
% Loop through and add mf's
for j = 1:1:cluster_n
params = computemfparams (mftype,...
X(:,i), U(j,:)', V(j,numInp+i));
fismat = addmf(fismat,'output', i, ...
```

```
['out' num2str(i) 'cluster' num2str(j)],...
mftype, params); end; end
otherwise
error('unknownfistype', ...
'Unknown fistype specified'); end
% Create rules
ruleList = ones(nC, numX+numY+2);
for i = 2:1:nC, ruleList(i,1:numX+numY) = i; end
fismat = addrule(fismat, ruleList);
% Set the input variable ranges
minX = min(X); maxX = max(X); ranges = [minX ; maxX]';
for i=1:numX, fismat.input(i).range = ranges(i,:); end
% Set the output variable ranges
minY = min(Y); maxY = max(Y); ranges = [minY ; maxY]';
for i=1:numY, fismat.output(i).range = ranges(i,:); end
return
function mfparams = computemfparams(mf,x,m,c)
switch lower(mf)
case 'gaussmf'
sigma = invgaussmf4sigma (x, m, c);
mfparams = [sigma, c];
case 'linear'
[N, dims] = size(x);
%MB correction   xin = [x(:,1:dims-1) ones(N,1)];
%MB correction   xout = x(:, dims);
xin = [x(:,1:dims-1) ones(N,1)].*(m*ones(1,dims));
xout = x(:, dims).*m; b = xin \ xout;
mfparams = b';
otherwise
disp('Unknown type of membership function');
```

```
end
return
function pr=minmax(p)
if iscell(p), [m,n] = size(p); pr = cell(m,1);
for i=1:m, pr{i} = minmax([p{i,:}]); end
elseif isa(p,'double'), pr = [min(p,[],2) max(p,[],2)];
else disp('Argument has illegal type.')
end
return
function  plotc( R, Fd, U)
%PLOTC plots a complete Fuzzy Rule-Base
%  plotc( R,Fd,U)
%   R Fuzzy model struct
%   Fd, Training Data
%   U,  FCM Membership values of Training Data
%   para, Nr of parameters of the mdfn
nR= size(R.rule,2);
nV=size(R.rule(1,1).antecedent,2);
hold off;
NXDiv = 100;  %=nr of divisions on X axis
for ir=1:nR
for iv=1:nV
maxX=R.input(1,iv).range(1,2);
minX=R.input(1,iv).range(1,1);
ivv=(ir-1)*(nV+2)+iv;
subplot(nR,nV+2,ivv);
% plot the data points
x1=Fd(:,iv); y1=U(ir,:);
plot(x1,y1,'ko','MarkerSize',1); hold on;
```

```
% plot the curve
for i=1:NXDiv
x11(i) = minX+(i-1)*(maxX-minX)/(NXDiv-1);
y11(i) = gaussmf(x11(i), ...
R.input(1,iv).mf(1,ir).params(1,:));
end
plot(x11,y11,'k-','LineWidth',2); hold off;
xlim([minX maxX]); ylim([-0.2 1.0]); end
% select the box
subplot(nR,nV+2,ivv+1);
% plot the data points
x1=Fd(:,iv+1); y1=U(ir,:);
plot(x1,y1,'ko','MarkerSize',1); hold on;
subplot(nR,nV+2,ivv+2);
text(0.1, 0.9,sprintf('f%i= %7.4f',ir,...
R.output.mf(1,ir).params(1,iv+1)));
for iv=1:nV,
text(0.2,0.9-iv*0.2,sprintf('+%7.4f x%i',...
R.output.mf(1,ir).params(1,iv),iv));end
set(gcf,'Position', [20  50  1200  650] );
axis off;
end
pause(0.5);
return
function a=strf(b,c);
a=-1; if(strfind(b,c)>0) a=strfind(b,c);end; return
function dispruleparam(R)
sP=@sprintf;
ic= size(R.input(1,1).mf,2); ix= size(R.input,2);
```

```
disp('Rule Base Parameters');
for i=1:ic % for each rule
for j=1:ix
disp([sP('%d %d',i, j) sP('%9.4f ', ...
R.input(1,j).mf(1,i).params(1,:))]);
end
disp(R.output(1,1).mf(1,i).params(1,:));
end
return
```

Appendix 2: The Raw Data Set

Training Data Set		Verification Data Set	
date	closing price	date	closing price
02/01/2008	6416.7	02/07/2010	4838.1
03/01/2008	6479.4	05/07/2010	4823.5
04/01/2008	6348.5	06/07/2010	4965
07/01/2008	6335.7	07/07/2010	5014.8
08/01/2008	6356.5	08/07/2010	5105.5
09/01/2008	6272.7	09/07/2010	5132.9
10/01/2008	6222.7	12/07/2010	5167
11/01/2008	6202	13/07/2010	5271
14/01/2008	6215.7	14/07/2010	5253.5
15/01/2008	6025.6	15/07/2010	5211.3
16/01/2008	5942.9	16/07/2010	5158.9
17/01/2008	5902.4	19/07/2010	5148.3
18/01/2008	5901.7	20/07/2010	5139.5
21/01/2008	5578.2	21/07/2010	5214.6
22/01/2008	5740.1	22/07/2010	5313.8
23/01/2008	5609.3	23/07/2010	5312.6
24/01/2008	5875.8	26/07/2010	5351.1
25/01/2008	5869	27/07/2010	5365.7
28/01/2008	5788.9	28/07/2010	5319.7
29/01/2008	5885.2	29/07/2010	5314
30/01/2008	5837.3	30/07/2010	5258
31/01/2008	5879.8	02/08/2010	5397.1
01/02/2008	6029.2	03/08/2010	5396.5
04/02/2008	6026.2	04/08/2010	5386.2
05/02/2008	5868	05/08/2010	5365.8
06/02/2008	5875.4	06/08/2010	5332.4
07/02/2008	5724.1	09/08/2010	5410.5
08/02/2008	5784	10/08/2010	5376.4
11/02/2008	5707.7	11/08/2010	5245.2
12/02/2008	5910	12/08/2010	5266.1
13/02/2008	5880.1	13/08/2010	5275.4
14/02/2008	5879.3	16/08/2010	5276.1
15/02/2008	5787.6	17/08/2010	5350.5
19/02/2008	5966.9	18/08/2010	5302.9
20/02/2008	5893.6	19/08/2010	5211.3
21/02/2008	5932.2	20/08/2010	5195.3
22/02/2008	5888.5	23/08/2010	5234.8
25/02/2008	5999.5	24/08/2010	5156
26/02/2008	6087.4	25/08/2010	5109.4
27/02/2008	6076.5	26/08/2010	5155.8
28/02/2008	5965.7	27/08/2010	5201.6
29/02/2008	5884.3	31/08/2010	5225.2
03/03/2008	5818.6	01/09/2010	5366.4
04/03/2008	5767.7	02/09/2010	5371

05/03/2008	5853.5	03/09/2010	5428.1
06/03/2008	5766.4	06/09/2010	5439.2
07/03/2008	5699.9	07/09/2010	5407.8
10/03/2008	5629.1	08/09/2010	5429.7
11/03/2008	5690.4	09/09/2010	5494.2
12/03/2008	5776.4	10/09/2010	5501.6
13/03/2008	5692.4	13/09/2010	5565.5
14/03/2008	5631.7	14/09/2010	5567.4
17/03/2008	5414.4	15/09/2010	5555.6
18/03/2008	5605.8	16/09/2010	5540.1
19/03/2008	5545.6	17/09/2010	5508.5
20/03/2008	5495.2	20/09/2010	5602.5
25/03/2008	5689.1	21/09/2010	5576.2
26/03/2008	5660.4	22/09/2010	5551.9
27/03/2008	5717.5	23/09/2010	5547.1
28/03/2008	5692.9	24/09/2010	5598.5
31/03/2008	5702.1	27/09/2010	5573.4
01/04/2008	5852.6	28/09/2010	5578.4
02/04/2008	5915.9	29/09/2010	5569.3
03/04/2008	5891.3	30/09/2010	5548.6
04/04/2008	5947.1	01/10/2010	5592.9
07/04/2008	6014.8	04/10/2010	5556
08/04/2008	5990.2	05/10/2010	5635.8
09/04/2008	5983.9	06/10/2010	5681.4
10/04/2008	5965.1	07/10/2010	5662.1
11/04/2008	5895.5	08/10/2010	5657.6
14/04/2008	5831.6	11/10/2010	5672.4
15/04/2008	5906.9	12/10/2010	5661.6
16/04/2008	6046.2	13/10/2010	5747.4
17/04/2008	5980.4	14/10/2010	5727.2
18/04/2008	6056.5	15/10/2010	5703.4
21/04/2008	6053	18/10/2010	5742.5
22/04/2008	6034.7	19/10/2010	5703.9
23/04/2008	6083.6	20/10/2010	5728.9
24/04/2008	6050.7	21/10/2010	5757.9
25/04/2008	6091.4	22/10/2010	5741.4
28/04/2008	6090.4	25/10/2010	5752
29/04/2008	6089.4	26/10/2010	5707.3
30/04/2008	6087.3	27/10/2010	5646
01/05/2008	6087.3	28/10/2010	5677.9
02/05/2008	6215.5	29/10/2010	5675.2
06/05/2008	6215.2	01/11/2010	5694.6
07/05/2008	6261	02/11/2010	5757.4
08/05/2008	6270.8	03/11/2010	5749
09/05/2008	6204.7	04/11/2010	5862.8
12/05/2008	6220.6	05/11/2010	5875.4
13/05/2008	6211.9	08/11/2010	5850
14/05/2008	6216	09/11/2010	5875.2
15/05/2008	6251.8	10/11/2010	5816.9
16/05/2008	6304.3	11/11/2010	5815.2

19/05/2008	6376.5	12/11/2010	5796.9
20/05/2008	6191.6	15/11/2010	5820.4
21/05/2008	6198.1	16/11/2010	5681.9
22/05/2008	6181.6	17/11/2010	5692.6
23/05/2008	6087.3	18/11/2010	5768.7
27/05/2008	6058.5	19/11/2010	5732.8
28/05/2008	6069.6	22/11/2010	5680.8
29/05/2008	6068.1	23/11/2010	5581.3
30/05/2008	6053.5	24/11/2010	5657.1
02/06/2008	6007.6	25/11/2010	5698.9
03/06/2008	6057.7	26/11/2010	5668.7
04/06/2008	5970.1	29/11/2010	5551
05/06/2008	5995.3	30/11/2010	5528.3
06/06/2008	5906.8	01/12/2010	5642.5
09/06/2008	5877.6	02/12/2010	5767.6
10/06/2008	5827.3	03/12/2010	5745.3
11/06/2008	5723.3	06/12/2010	5770.3
12/06/2008	5790.5	07/12/2010	5808.5
13/06/2008	5802.8	08/12/2010	5794.5
16/06/2008	5794.6	09/12/2010	5808
17/06/2008	5861.9	10/12/2010	5813
18/06/2008	5756.9	13/12/2010	5860.8
19/06/2008	5708.4	14/12/2010	5891.2
20/06/2008	5620.8	15/12/2010	5882.2
23/06/2008	5667.2	16/12/2010	5881.1
24/06/2008	5634.7	17/12/2010	5871.8
25/06/2008	5666.1	20/12/2010	5891.6
26/06/2008	5518.2	21/12/2010	5951.8
27/06/2008	5529.9	22/12/2010	5983.5
30/06/2008	5625.9	23/12/2010	5996.1
01/07/2008	5479.9	24/12/2010	6008.9
02/07/2008	5426.3	29/12/2010	5996.4
03/07/2008	5476.6	30/12/2010	5971
04/07/2008	5412.8	31/12/2010	5899.9
07/07/2008	5512.7	04/01/2011	6013.9
08/07/2008	5440.5	05/01/2011	6043.9
09/07/2008	5529.6	06/01/2011	6019.5
10/07/2008	5406.8	07/01/2011	5984.3
11/07/2008	5261.6	10/01/2011	5956.3
14/07/2008	5300.4	11/01/2011	6014
15/07/2008	5171.9	12/01/2011	6050.7
16/07/2008	5150.6	13/01/2011	6023.9
17/07/2008	5286.3	14/01/2011	6002.1
18/07/2008	5376.4	17/01/2011	5985.7
21/07/2008	5404.3	18/01/2011	6056.4
22/07/2008	5364.1	19/01/2011	5976.7
23/07/2008	5449.9	20/01/2011	5867.9
24/07/2008	5362.3	21/01/2011	5896.3
25/07/2008	5352.6	24/01/2011	5943.9
28/07/2008	5312.6	25/01/2011	5917.7

29/07/2008	5319.2	26/01/2011	5969.2
30/07/2008	5420.7	27/01/2011	5965.1
31/07/2008	5411.9	28/01/2011	5881.4
01/08/2008	5354.7	31/01/2011	5862.9
04/08/2008	5320.2	01/02/2011	5957.8
05/08/2008	5454.5	02/02/2011	6000.1
06/08/2008	5486.1	03/02/2011	5983.3
07/08/2008	5477.5	04/02/2011	5997.4
08/08/2008	5489.2	07/02/2011	6051
11/08/2008	5541.8	08/02/2011	6091.3
12/08/2008	5534.5	09/02/2011	6052.3
13/08/2008	5448.6	10/02/2011	6020
14/08/2008	5497.4	11/02/2011	6062.9
15/08/2008	5454.8	14/02/2011	6060.1
18/08/2008	5450.2	15/02/2011	6037.1
19/08/2008	5320.4	16/02/2011	6085.3
20/08/2008	5371.8	17/02/2011	6087.4
21/08/2008	5370.2	18/02/2011	6083
22/08/2008	5505.6	21/02/2011	6014.8
26/08/2008	5470.7	22/02/2011	5996.8
27/08/2008	5528.1	23/02/2011	5923.5
28/08/2008	5601.2	24/02/2011	5920
29/08/2008	5636.6	25/02/2011	6001.2
01/09/2008	5602.8	28/02/2011	5994
02/09/2008	5620.7	01/03/2011	5935.8
03/09/2008	5499.7	02/03/2011	5914.9
04/09/2008	5362.1	03/03/2011	6005.1
05/09/2008	5240.7	04/03/2011	5990.4
08/09/2008	5446.3	07/03/2011	5973.8
09/09/2008	5415.6	08/03/2011	5974.8
10/09/2008	5366.2	09/03/2011	5937.3
11/09/2008	5318.4	10/03/2011	5845.3
12/09/2008	5416.7	11/03/2011	5828.7
15/09/2008	5204.2	14/03/2011	5775.2
16/09/2008	5025.6	15/03/2011	5695.3
17/09/2008	4912.4	16/03/2011	5598.2
18/09/2008	4880	17/03/2011	5696.1
19/09/2008	5311.3	18/03/2011	5718.1
22/09/2008	5236.3	21/03/2011	5786.1
23/09/2008	5136.1	22/03/2011	5762.7
24/09/2008	5095.6	23/03/2011	5795.9
25/09/2008	5197	24/03/2011	5880.9
26/09/2008	5088.5	25/03/2011	5900.8
29/09/2008	4818.8	28/03/2011	5904.5
30/09/2008	4902.5	29/03/2011	5932.2
01/10/2008	4959.6	30/03/2011	5948.3
02/10/2008	4870.3	31/03/2011	5908.8
03/10/2008	4980.3	01/04/2011	6009.9
06/10/2008	4589.2	04/04/2011	6017
07/10/2008	4605.2	05/04/2011	6007.1

08/10/2008	4366.7	06/04/2011	6041.1
09/10/2008	4313.8	07/04/2011	6007.4
10/10/2008	3932.1	08/04/2011	6055.8
13/10/2008	4256.9	11/04/2011	6053.4
14/10/2008	4394.2	12/04/2011	5964.5
15/10/2008	4079.6	13/04/2011	6010.4
16/10/2008	3861.4	14/04/2011	5963.8
17/10/2008	4063	15/04/2011	5996
20/10/2008	4282.7	18/04/2011	5870.1
21/10/2008	4229.7	19/04/2011	5896.9
22/10/2008	4040.9	20/04/2011	6022.3
23/10/2008	4087.8	21/04/2011	6018.3
24/10/2008	3883.4	26/04/2011	6069.4
27/10/2008	3852.6	27/04/2011	6068.2
28/10/2008	3926.4	28/04/2011	6069.9
29/10/2008	4242.5	03/05/2011	6082.9
30/10/2008	4291.6	04/05/2011	5984.1
31/10/2008	4377.3	05/05/2011	5920
03/11/2008	4443.3	06/05/2011	5976.8
04/11/2008	4639.5	09/05/2011	5942.7
05/11/2008	4530.7	10/05/2011	6018.9
06/11/2008	4272.4	11/05/2011	5976
07/11/2008	4365	12/05/2011	5945
10/11/2008	4403.9	13/05/2011	5925.9
11/11/2008	4246.7	16/05/2011	5923.7
12/11/2008	4182	17/05/2011	5861
13/11/2008	4169.2	18/05/2011	5923.5
14/11/2008	4233	19/05/2011	5956
17/11/2008	4132.2	20/05/2011	5948.5
18/11/2008	4208.5	23/05/2011	5835.9
19/11/2008	4005.7	24/05/2011	5858.4
20/11/2008	3875	25/05/2011	5870.1
21/11/2008	3781	26/05/2011	5881
24/11/2008	4153	27/05/2011	5938.9
25/11/2008	4171.3	31/05/2011	5990
26/11/2008	4152.7	01/06/2011	5928.6
27/11/2008	4226.1	02/06/2011	5847.9
28/11/2008	4288	03/06/2011	5855
01/12/2008	4065.5	06/06/2011	5863.2
02/12/2008	4122.9	07/06/2011	5864.6
03/12/2008	4170	08/06/2011	5808.9
04/12/2008	4163.6	09/06/2011	5856.3
05/12/2008	4049.4	10/06/2011	5765.8
08/12/2008	4300.1	13/06/2011	5773.5
09/12/2008	4381.3	14/06/2011	5803.1
10/12/2008	4367.3	15/06/2011	5742.5
11/12/2008	4388.7	16/06/2011	5698.8
12/12/2008	4280.4	17/06/2011	5714.9
15/12/2008	4277.6	20/06/2011	5693.4
16/12/2008	4309.1	21/06/2011	5775.3

17/12/2008	4324.2	22/06/2011	5773
18/12/2008	4330.7	23/06/2011	5674.4
19/12/2008	4286.9	24/06/2011	5697.7
22/12/2008	4249.2	27/06/2011	5722.3
23/12/2008	4256	28/06/2011	5766.9
24/12/2008	4216.6	29/06/2011	5856
29/12/2008	4319.4	30/06/2011	5945.7
30/12/2008	4392.7	01/07/2011	5989.8
31/12/2008	4434.2	04/07/2011	6017.5
02/01/2009	4561.8	05/07/2011	6024
05/01/2009	4579.6	06/07/2011	6002.9
06/01/2009	4638.9	07/07/2011	6054.5
07/01/2009	4507.5	08/07/2011	5990.6
08/01/2009	4505.4	11/07/2011	5929.2
09/01/2009	4448.5	12/07/2011	5869
12/01/2009	4426.2	13/07/2011	5906.4
13/01/2009	4399.1	14/07/2011	5847
14/01/2009	4180.6	15/07/2011	5843.7
15/01/2009	4121.1	18/07/2011	5752.8
16/01/2009	4147.1	19/07/2011	5790
19/01/2009	4108.5	20/07/2011	5853.8
20/01/2009	4091.4	21/07/2011	5899.9
21/01/2009	4059.9	22/07/2011	5935
22/01/2009	4052.2	25/07/2011	5925.3
23/01/2009	4052.5	26/07/2011	5929.7
26/01/2009	4209	27/07/2011	5856.6
27/01/2009	4194.4	28/07/2011	5873.2
28/01/2009	4295.2	29/07/2011	5815.2
29/01/2009	4190.1	01/08/2011	5774.4
30/01/2009	4149.6	02/08/2011	5718.4
02/02/2009	4077.8	03/08/2011	5584.5
03/02/2009	4164.5	04/08/2011	5393.1
04/02/2009	4228.6	05/08/2011	5247
05/02/2009	4228.9	08/08/2011	5069
06/02/2009	4291.9	09/08/2011	5164.9
09/02/2009	4307.6	10/08/2011	5007.2
10/02/2009	4213.1	11/08/2011	5162.8
11/02/2009	4234.3	12/08/2011	5320
12/02/2009	4202.2	15/08/2011	5350.6
13/02/2009	4189.6	16/08/2011	5357.6
16/02/2009	4134.8	17/08/2011	5331.6
17/02/2009	4034.1	18/08/2011	5092.2
18/02/2009	4006.8	19/08/2011	5040.8
19/02/2009	4018.4	22/08/2011	5095.3
20/02/2009	3889.1	23/08/2011	5129.4
23/02/2009	3850.7	24/08/2011	5205.9
24/02/2009	3816.4	25/08/2011	5131.1
25/02/2009	3849	26/08/2011	5129.9
26/02/2009	3915.6	30/08/2011	5268.7
27/02/2009	3830.1	31/08/2011	5394.5

02/03/2009	3625.8	01/09/2011	5418.6
03/03/2009	3512.1	02/09/2011	5292
04/03/2009	3645.9	05/09/2011	5102.6
05/03/2009	3529.9	06/09/2011	5156.8
06/03/2009	3530.7	07/09/2011	5318.6
09/03/2009	3542.4	08/09/2011	5340.4
10/03/2009	3715.2	09/09/2011	5214.6
11/03/2009	3693.8	12/09/2011	5129.6
12/03/2009	3712.1	13/09/2011	5174.3
13/03/2009	3753.7	14/09/2011	5227
16/03/2009	3864	15/09/2011	5337.5
17/03/2009	3857.1	16/09/2011	5368.4
18/03/2009	3805	19/09/2011	5259.6
19/03/2009	3816.9	20/09/2011	5363.7
20/03/2009	3842.9	21/09/2011	5288.4
23/03/2009	3952.8	22/09/2011	5041.6
24/03/2009	3911.5	23/09/2011	5066.8
25/03/2009	3900.3	26/09/2011	5089.4
26/03/2009	3925.2	27/09/2011	5294
27/03/2009	3898.9	28/09/2011	5217.6
30/03/2009	3762.9	29/09/2011	5196.8
31/03/2009	3926.1	30/09/2011	5128.5
01/04/2009	3955.6	03/10/2011	5075.5
02/04/2009	4125	04/10/2011	4944.4
03/04/2009	4029.7	05/10/2011	5102.2
06/04/2009	3993.5	06/10/2011	5291.3
07/04/2009	3930.5	07/10/2011	5303.4
08/04/2009	3925.5	10/10/2011	5399
09/04/2009	3983.7	11/10/2011	5395.7
14/04/2009	3989	12/10/2011	5441.8
15/04/2009	3968.4	13/10/2011	5403.4
16/04/2009	4053	14/10/2011	5466.4
17/04/2009	4092.8	17/10/2011	5436.7
20/04/2009	3990.9	18/10/2011	5410.4
21/04/2009	3987.5	19/10/2011	5450.5
22/04/2009	4030.7	20/10/2011	5384.7
23/04/2009	4018.2	21/10/2011	5488.6
24/04/2009	4156	24/10/2011	5548.1
27/04/2009	4167	25/10/2011	5525.5
28/04/2009	4096.4	26/10/2011	5553.2
29/04/2009	4189.6	27/10/2011	5713.8
30/04/2009	4243.7	28/10/2011	5702.2
01/05/2009	4243.2	31/10/2011	5544.2
05/05/2009	4336.9	01/11/2011	5421.6
06/05/2009	4396.5	02/11/2011	5484.1
07/05/2009	4398.7	03/11/2011	5545.6
08/05/2009	4462.1	04/11/2011	5527.2
11/05/2009	4435.5	07/11/2011	5510.8
12/05/2009	4425.5	08/11/2011	5567.3
13/05/2009	4331.4	09/11/2011	5460.4

14/05/2009	4362.6	10/11/2011	5444.8
15/05/2009	4348.1	11/11/2011	5545.4
18/05/2009	4446.5	14/11/2011	5519
19/05/2009	4482.3	15/11/2011	5517.4
20/05/2009	4468.4	16/11/2011	5509
21/05/2009	4345.5	17/11/2011	5423.1
22/05/2009	4365.3	18/11/2011	5362.9
26/05/2009	4411.7	21/11/2011	5222.6
27/05/2009	4416.2	22/11/2011	5206.8
28/05/2009	4387.5	23/11/2011	5139.8
29/05/2009	4417.9	24/11/2011	5127.6
01/06/2009	4506.2	25/11/2011	5164.6
02/06/2009	4477	28/11/2011	5312.8
03/06/2009	4383.4	29/11/2011	5337
04/06/2009	4386.9	30/11/2011	5505.4
05/06/2009	4438.6	01/12/2011	5489.3
08/06/2009	4405.2	02/12/2011	5552.3
09/06/2009	4404.8	05/12/2011	5568
10/06/2009	4436.8	06/12/2011	5568.7
11/06/2009	4461.9	07/12/2011	5546.9
12/06/2009	4442	08/12/2011	5483.8
15/06/2009	4326	09/12/2011	5529.2
16/06/2009	4328.6	12/12/2011	5427.9
17/06/2009	4278.5	13/12/2011	5490.1
18/06/2009	4280.9	14/12/2011	5366.8
19/06/2009	4345.9	15/12/2011	5400.9
22/06/2009	4234	16/12/2011	5387.3
23/06/2009	4230	19/12/2011	5365
24/06/2009	4280	20/12/2011	5419.6
25/06/2009	4252.6	21/12/2011	5389.7
26/06/2009	4241	22/12/2011	5457
29/06/2009	4294	23/12/2011	5512.7
30/06/2009	4249.2	28/12/2011	5507.4
01/07/2009	4340.7	29/12/2011	5566.8
02/07/2009	4234.3	30/12/2011	5572.3
03/07/2009	4236.3	03/01/2012	5699.9
06/07/2009	4194.9	04/01/2012	5668.5
07/07/2009	4187	05/01/2012	5624.3
08/07/2009	4140.2	06/01/2012	5649.7
09/07/2009	4158.7	09/01/2012	5612.3
10/07/2009	4127.2	10/01/2012	5696.7
13/07/2009	4202.1	11/01/2012	5670.8
14/07/2009	4237.7	12/01/2012	5662.4
15/07/2009	4346.5	13/01/2012	5636.6
16/07/2009	4361.8	16/01/2012	5657.4
17/07/2009	4388.8	17/01/2012	5694
20/07/2009	4443.6	18/01/2012	5702.4
21/07/2009	4481.2	19/01/2012	5741.1
22/07/2009	4493.7	20/01/2012	5728.5
23/07/2009	4559.8	23/01/2012	5782.6

24/07/2009	4576.6	24/01/2012	5751.9
27/07/2009	4586.1	25/01/2012	5723
28/07/2009	4528.8	26/01/2012	5795.2
29/07/2009	4547.5	27/01/2012	5733.5
30/07/2009	4631.6	30/01/2012	5671.1
31/07/2009	4608.4	31/01/2012	5681.6
03/08/2009	4682.5	01/02/2012	5790.7
04/08/2009	4671.4	02/02/2012	5796.1
05/08/2009	4647.1	03/02/2012	5901.1
06/08/2009	4690.5	06/02/2012	5892.2
07/08/2009	4731.6	07/02/2012	5890.3
10/08/2009	4722.2	08/02/2012	5875.9
11/08/2009	4671.3	09/02/2012	5895.5
12/08/2009	4716.8	10/02/2012	5852.4
13/08/2009	4755.5	13/02/2012	5905.7
14/08/2009	4714	14/02/2012	5899.9
17/08/2009	4645	15/02/2012	5892.2
18/08/2009	4685.8	16/02/2012	5885.4
19/08/2009	4689.7	17/02/2012	5905.1
20/08/2009	4756.6	20/02/2012	5945.3
21/08/2009	4850.9	21/02/2012	5928.2
24/08/2009	4896.2	22/02/2012	5916.5
25/08/2009	4916.8	23/02/2012	5937.9
26/08/2009	4890.6	24/02/2012	5935.1
27/08/2009	4869.4	27/02/2012	5915.5
28/08/2009	4908.9	28/02/2012	5927.9
01/09/2009	4819.7	29/02/2012	5871.5
02/09/2009	4817.5	01/03/2012	5931.3
03/09/2009	4796.8	02/03/2012	5911.1
04/09/2009	4851.7	05/03/2012	5874.8
07/09/2009	4933.2	06/03/2012	5765.8
08/09/2009	4947.3	07/03/2012	5791.4
09/09/2009	5004.3	08/03/2012	5859.7
10/09/2009	4987.7	09/03/2012	5887.5
11/09/2009	5011.5	12/03/2012	5892.8
14/09/2009	5018.9	13/03/2012	5955.9
15/09/2009	5042.1	14/03/2012	5945.4
16/09/2009	5124.1	15/03/2012	5940.7
17/09/2009	5164	16/03/2012	5965.6
18/09/2009	5172.9	19/03/2012	5961.1
21/09/2009	5134.4	20/03/2012	5891.4
22/09/2009	5142.6	21/03/2012	5892
23/09/2009	5139.4	22/03/2012	5845.6
24/09/2009	5079.3	23/03/2012	5854.9
25/09/2009	5082.2	26/03/2012	5902.7
28/09/2009	5165.7	27/03/2012	5869.5
29/09/2009	5159.7	28/03/2012	5809
30/09/2009	5133.9	29/03/2012	5742
01/10/2009	5047.8	30/03/2012	5768.5
02/10/2009	4988.7	02/04/2012	5874.9

05/10/2009	5024.3	03/04/2012	5838.3
06/10/2009	5138	04/04/2012	5703.8
07/10/2009	5108.9	05/04/2012	5723.7
08/10/2009	5154.6	10/04/2012	5595.5
09/10/2009	5161.9	11/04/2012	5634.7
12/10/2009	5210.2	12/04/2012	5710.5
13/10/2009	5154.1	13/04/2012	5651.8
14/10/2009	5256.1	16/04/2012	5666.3
15/10/2009	5223	17/04/2012	5767
16/10/2009	5190.2	18/04/2012	5745.3
19/10/2009	5281.5	19/04/2012	5744.5
20/10/2009	5243.4	20/04/2012	5772.1
21/10/2009	5257.9	23/04/2012	5665.6
22/10/2009	5207.4	24/04/2012	5709.5
23/10/2009	5242.6	25/04/2012	5718.9
26/10/2009	5191.7	26/04/2012	5748.7
27/10/2009	5201	27/04/2012	5777.1
28/10/2009	5080.4	30/04/2012	5737.8
29/10/2009	5137.7	01/05/2012	5812.2
30/10/2009	5044.5	02/05/2012	5758.1
02/11/2009	5104.5	03/05/2012	5766.5
03/11/2009	5037.2	04/05/2012	5655.1
04/11/2009	5107.9	08/05/2012	5554.5
05/11/2009	5125.6	09/05/2012	5530
06/11/2009	5142.7	10/05/2012	5544
09/11/2009	5235.2	11/05/2012	5575.5
10/11/2009	5230.5	14/05/2012	5465.5
11/11/2009	5266.8	15/05/2012	5437.6
12/11/2009	5276.5	16/05/2012	5405.3
13/11/2009	5296.4	17/05/2012	5338.4
16/11/2009	5382.7	18/05/2012	5267.6
17/11/2009	5345.9	21/05/2012	5304.5
18/11/2009	5342.1	22/05/2012	5403.3
19/11/2009	5267.7	23/05/2012	5266.4
20/11/2009	5251.4	24/05/2012	5350
23/11/2009	5355.5	25/05/2012	5351.5
24/11/2009	5324	28/05/2012	5356.3
25/11/2009	5364.8	29/05/2012	5391.1
26/11/2009	5194.1	30/05/2012	5297.3
27/11/2009	5245.7	31/05/2012	5320.9
30/11/2009	5190.7	01/06/2012	5260.2
01/12/2009	5312.2	06/06/2012	5384.1
02/12/2009	5327.4	07/06/2012	5447.8
03/12/2009	5313	08/06/2012	5435.1
04/12/2009	5322.4	11/06/2012	5432.4
07/12/2009	5310.7	12/06/2012	5473.7
08/12/2009	5223.1	13/06/2012	5483.8
09/12/2009	5203.9	14/06/2012	5467
10/12/2009	5244.4	15/06/2012	5478.8
11/12/2009	5261.6	18/06/2012	5491.1

14/12/2009	5315.3	19/06/2012	5586.3
15/12/2009	5285.8	20/06/2012	5622.3
16/12/2009	5320.3	21/06/2012	5566.4
17/12/2009	5217.6	22/06/2012	5513.7
18/12/2009	5196.8	25/06/2012	5450.6
21/12/2009	5294	26/06/2012	5447
22/12/2009	5328.7	27/06/2012	5523.9
23/12/2009	5372.4	28/06/2012	5493.1
24/12/2009	5402.4	29/06/2012	5571.1
29/12/2009	5437.6	02/07/2012	5640.6
30/12/2009	5397.9	03/07/2012	5687.7
31/12/2009	5412.9	04/07/2012	5684.5
04/01/2010	5500.3	05/07/2012	5692.6
05/01/2010	5522.5	06/07/2012	5662.6
06/01/2010	5530	09/07/2012	5627.3
07/01/2010	5526.7	10/07/2012	5664.1
08/01/2010	5534.2	11/07/2012	5664.5
11/01/2010	5538.1	12/07/2012	5608.3
12/01/2010	5498.7	13/07/2012	5666.1
13/01/2010	5473.5	16/07/2012	5662.4
14/01/2010	5498.2	17/07/2012	5629.1
15/01/2010	5455.4	18/07/2012	5685.8
18/01/2010	5494.4	19/07/2012	5714.2
19/01/2010	5513.1	20/07/2012	5651.8
20/01/2010	5420.8	23/07/2012	5533.9
21/01/2010	5335.1	24/07/2012	5499.2
22/01/2010	5303	25/07/2012	5498.3
25/01/2010	5260.3	26/07/2012	5573.2
26/01/2010	5276.9	27/07/2012	5627.2
27/01/2010	5217.5	30/07/2012	5693.6
28/01/2010	5145.7	31/07/2012	5635.3
29/01/2010	5188.5	01/08/2012	5712.8
01/02/2010	5247.4	02/08/2012	5662.3
02/02/2010	5283.3	03/08/2012	5787.3
03/02/2010	5253.1	06/08/2012	5808.8
04/02/2010	5139.3	07/08/2012	5841.2
05/02/2010	5060.9	08/08/2012	5845.9
08/02/2010	5092.3	09/08/2012	5851.5
09/02/2010	5111.8	10/08/2012	5847.1
10/02/2010	5132	13/08/2012	5831.9
11/02/2010	5161.5	14/08/2012	5864.8
12/02/2010	5142.5	15/08/2012	5833
15/02/2010	5167.5	16/08/2012	5834.5
16/02/2010	5244.1	17/08/2012	5852.4
17/02/2010	5276.6	20/08/2012	5824.4
18/02/2010	5325.1	21/08/2012	5857.5
19/02/2010	5358.2	22/08/2012	5774.2
22/02/2010	5352.1	23/08/2012	5776.6
23/02/2010	5315.1	24/08/2012	5776.6
24/02/2010	5342.9	28/08/2012	5775.7

25/02/2010	5278.2	29/08/2012	5743.5
26/02/2010	5354.5	30/08/2012	5719.5
01/03/2010	5405.9	31/08/2012	5711.5
02/03/2010	5484.1	03/09/2012	5758.4
03/03/2010	5533.2	04/09/2012	5672
04/03/2010	5527.2	05/09/2012	5657.9
05/03/2010	5599.8	06/09/2012	5777.3
08/03/2010	5606.7	07/09/2012	5794.8
09/03/2010	5602.3	10/09/2012	5793.2
10/03/2010	5640.6	11/09/2012	5792.2
11/03/2010	5617.3	12/09/2012	5782.1
12/03/2010	5625.6	13/09/2012	5819.9
15/03/2010	5593.9	14/09/2012	5915.5
16/03/2010	5620.4	17/09/2012	5893.5
17/03/2010	5644.6	18/09/2012	5868.2
18/03/2010	5642.6	19/09/2012	5888.5
19/03/2010	5650.1	20/09/2012	5854.6
22/03/2010	5644.5	21/09/2012	5852.6
23/03/2010	5673.6	24/09/2012	5838.8
24/03/2010	5677.9	25/09/2012	5859.7
25/03/2010	5727.6	26/09/2012	5768.1
26/03/2010	5703	27/09/2012	5779.4
29/03/2010	5710.7	28/09/2012	5742.1
30/03/2010	5672.3	01/10/2012	5820.5
31/03/2010	5679.6	02/10/2012	5809.5
01/04/2010	5744.9	03/10/2012	5825.8
06/04/2010	5780.4	04/10/2012	5827.8
07/04/2010	5762.1	05/10/2012	5871
08/04/2010	5712.7	08/10/2012	5841.7
09/04/2010	5771	09/10/2012	5810.3
12/04/2010	5777.6	10/10/2012	5776.7
13/04/2010	5761.7	11/10/2012	5829.8
14/04/2010	5796.3	12/10/2012	5793.3
15/04/2010	5825	15/10/2012	5805.6
16/04/2010	5744	16/10/2012	5870.5
19/04/2010	5727.9	17/10/2012	5910.9
20/04/2010	5783.7	18/10/2012	5917
21/04/2010	5723.4	19/10/2012	5896.1
22/04/2010	5665.3	22/10/2012	5882.9
23/04/2010	5723.6	23/10/2012	5797.9
26/04/2010	5753.9	24/10/2012	5804.8
27/04/2010	5603.5	25/10/2012	5805
28/04/2010	5586.6	26/10/2012	5806.7
29/04/2010	5617.8	29/10/2012	5795.1
30/04/2010	5553.3	30/10/2012	5849.9
04/05/2010	5411.1	31/10/2012	5782.7
05/05/2010	5341.9	01/11/2012	5861.9
06/05/2010	5261	02/11/2012	5868.5
07/05/2010	5123	05/11/2012	5839.1
10/05/2010	5387.4	06/11/2012	5884.9

11/05/2010	5334.2	07/11/2012	5791.6
12/05/2010	5383.5	08/11/2012	5776
13/05/2010	5433.7	09/11/2012	5769.7
14/05/2010	5262.9	12/11/2012	5767.3
17/05/2010	5262.5	13/11/2012	5786.3
18/05/2010	5307.3	14/11/2012	5722
19/05/2010	5158.1	15/11/2012	5677.8
20/05/2010	5073.1	16/11/2012	5605.6
21/05/2010	5062.9	19/11/2012	5737.7
24/05/2010	5069.6	20/11/2012	5748.1
25/05/2010	4940.7	21/11/2012	5752
26/05/2010	5038.1	22/11/2012	5791
27/05/2010	5195.2	23/11/2012	5819.1
28/05/2010	5188.4	26/11/2012	5786.7
01/06/2010	5163.3	27/11/2012	5799.7
02/06/2010	5151.3	28/11/2012	5803.3
03/06/2010	5211.2	29/11/2012	5870.3
04/06/2010	5126	30/11/2012	5866.8
07/06/2010	5069.1	03/12/2012	5871.2
08/06/2010	5028.1	04/12/2012	5869
09/06/2010	5085.9	05/12/2012	5892.1
10/06/2010	5132.5	06/12/2012	5901.4
11/06/2010	5163.7	07/12/2012	5914.4
14/06/2010	5202.1	10/12/2012	5921.6
15/06/2010	5217.8	11/12/2012	5925
16/06/2010	5237.9	12/12/2012	5945.9
17/06/2010	5253.9	13/12/2012	5929.6
18/06/2010	5250.8	14/12/2012	5921.8
21/06/2010	5299.1	17/12/2012	5912.1
22/06/2010	5247	18/12/2012	5935.9
23/06/2010	5178.5	19/12/2012	5961.6
24/06/2010	5100.2	20/12/2012	5958.3
25/06/2010	5046.5	21/12/2012	5940
28/06/2010	5071.7	24/12/2012	5954.2
29/06/2010	4914.2	27/12/2012	5954.3
30/06/2010	4916.9	28/12/2012	5925.4
01/07/2010	4805.8	31/12/2012	5897.8