INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

A COMPARATIVE STUDY OF DIFFERENT METHODS OF PREDICTING TIME SERIES

Sutanuka Bhattacharya

A thesis

in

The Department

of

Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements

For the Degree of Master of Science

Concordia University

Montréal, Québec, Canada

March 1997

© Sutanuka Bhattacharya, 1997

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-25985-4

Canada

Concordia University

School of Graduate Studies

This is to certify that the thesis prepared

By: **Sutanuka Bhattacharya**

Entitled: **A Comparative Study of different methods of predicting time series**

and submitted in partial fulfillment of the requirements for the degree of

Master of Science

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining commitee:

______________________________ Chair

______________________________ Examiner

______________________________ Examiner

______________________________ Examiner

______________________________ Supervisor

Approved ______________________________

Chair of Department or Graduate Program Director

__________ 19 _____ ______________________________

Dean of Faculty

Abstract

A Comparative Study of different methods of predicting time series

Sutanuka Bhattacharya

This thesis work presents a comparative study of different methods for predicting future values of time series data and implement them to predict the currency exchange rates. The current thesis focuses mainly on two approaches in predicting a time series. One of them is the traditional statistical approach which involves building models based on certain assumptions and then applying them to do the predictions. The models considered in this thesis are multiple regression, exponential smoothing, double exponential smoothing, Box-Jenkins method, and Winter's method. The second approach is using the concept of training neural nets and pattern recognition. This involves in designing a neural network and training it using different learning methods. The learning algorithms used in the current work involves the backpropagation method, recurrent nets learning method, adaptively trained neural nets, and fuzzy learning methods. In addition to these, some methods for forecasting a chaotic time series and fractional differencing are also mentioned in the thesis. In order to compare the performances of different techniques of forecasting the future values of a time series, experiments were conducted using the exchange rates of different currencies with respect to the US dollar. These exchange rates exhibit a lot of randomness in their behaviour and hence it was very challenging to predict their future values. Different prediction zones were selected to conduct the experiments and analysis of the results have been presented towards the end of the thesis.

To my loving father,

the late Professor Ajit Kumar Bhattacharji

who gave me all the inspiration to start this work

but did not live to see it completed.

God took him away too soon!

Acknowledgments

I would like to take this opportunity to thank my thesis supervisor Prof. Tao Li for his guidance, enthusiasm and interest in this project. I am grateful to my parents and my sister Mrs. Paramita Sikdar for their patience and constant encouragement throughout the project. I am thankful to Mr. Steven Robbins for giving me unlimited access to his computer system which made it possible to carry out the experiments. His timely suggestions and comments on different programming techniques proved to be extremely helpful. I am thankful to Mr. Pankaj Kamthan for providing me some useful references. Thanks are due to Ms Nancy Gravenor for her help and cooperation, specially in times of need. Finally, I wish to thank all my Professors, colleagues and friends for their support and encouragement.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One of the most important aspects of any organization is *decision making.* Making a good decision depends largely on *predicting* the future events and conditions. This thesis work presents a

(a) Comparative study of different methods for predicting future values of time series data.

(b) Application of the methods in predicting currency exchange rates.

The basic assumption made when forecasting is that, there is always an underlying pattern which describes the event and conditions and it repeats in the future. A time series is a chronological sequence of observations on a particular variable.

Some variables in real life change with time. Hence we can use time series data, which are collected based on observations of the variable of interest (the event) over a period of time, as past data for forecasting. Chapter 2 gives a brief overview of different concepts of a time series and its applications in the real world.

Chapter 3 focuses on discussion of measuring different types of errors and their application to time series.

Chapter 4 deals with statistical time series modelling. In case of statistical methods of prediction, the general idea behind the prediction of the future values of a time series or, in other words, forecasting a time series, involves the following steps:

1. Analysis of past data in order to identify the pattern of behaviour.
2. The pattern is extrapolated or extended in order to prepare a forecast.
3. It is assumed that the pattern that has been identified will remain more or less unchanged over a given time period.

The last one is one of the vital points to be noted. The entire forecasting theory will fail if there is an abrupt change in the behaviour of the pattern. This is the situation where the problem becomes interesting and difficult. Hence the forecaster has to be aware of this situation and try to anticipate the sudden changes in pattern and make appropriate changes in the forecasting system. The methods used for forecasting statistical time series are multiple regression, exponential smoothing, double exponential smoothing, Box-Jenkins methodology and Winter's method. The derivations of the mathematical formulae accompanying these methods have been presented in the appendix.

Chapter 5 deals with forecasting time series by training neural nets by pattern recognition and fuzzy logic. A neural net is basically an interconnection of computational units called nodes or neurons. Typically, a neural net consists of a layer of neurons receiving the input pattern, a hidden layer which has no connection with the outer world, but is responsible for the performance of the network, and the output layer which returns the output of the network after each training session. The influence of each neuron on the performance of the network is measured in terms of the connection weights between the neurons. The basic concept of using neural nets for predicting future values of a time series involves the following steps:

1. Design a network and initialize all the components with random values.
2. Train the network with a set of input patterns chosen from a given time series and update the weights accordingly.

3. Once the weights are stabilized, use the network for predicting the future values of the time series from which the input pattern has been chosen.

Thus it can be noted that the method of prediction using neural nets is different from that of statistical methods. Here, the predictions are done based on training the network and no models are designed as in the case of statistical methods. The collective of neurons approximates the true model through the process of training. The performance of this approach depends on the learning algorithm and of course the structure of the network and its parameters. The configuration of a neural net depends largely on the learning method used and the complexity of the pattern (which in this case is the time series) itself. The training methods presented in this thesis include adaptively trained neural nets, backpropagation, recurrent nets, and fuzzy learning.

Chapter 6 includes a comparison of traditional statistical techniques and neural networks in forecasting the future values of a time series. This is done by analyzing their performance in predicting the exchange rates for different countries with the US dollar. These exchange rates exhibit complex patterns and hence it was challenging to predict their future values. Different prediction zones have been chosen to carry out the experiments.

In addition to these, some techniques for predicting white noises (which appears when deriving Brownian motions) and chaotic time series has been presented in some details in Chapter 7.

Chapter 8 contains a summary and conclusions of results from the experiments which where carried out during the course of the entire thesis work.

Chapter 2

Properties and Applications of Time Series

In order to understand the effectiveness of forecasting future values of a time series, it is important to first identify what a time series consists of and what are its main properties. This chapter gives a brief account of the component and properties of a time series. It is followed by a detailed description of the application of time series analysis in prediction of exchange rates. Finally, a discussion on the use of time series in physiology has been presented.

2.1 Components and Properties of a Time Series

A *time series* is defined as the a chronological sequence of observations on a particular variable.

The components of a time series are :

1. Trend
2. Cycle
3. Seasonal Variations
4. Irregular Fluctuations

The definitions of the above four components of a time series are given below:

1. Trend: refers to the upward or downward movement that characterizes a time series over a period of time. Thus, trend refers a long-run growth or decline in the time series.

2. Cycle: refers to recurring up and down movements around trend levels. These fluctuations can have a duration anywhere from 2 to 10 years or even longer measured from peak to peak or trough to trough.

3. Seasonal Variations: are periodic patterns in time series that complete themselves within a calendar year and are then repeated on a yearly basis.

4. Irregular Fluctuations: are erratic movements in a time series that follow no recognizable or regular pattern. This factor is the source of errors in forecasting.

It should be noted that a time series is a combination of some or all the above components, hence no single best forecasting technique exists. One of the most important problems in designing an efficient time series model is trying to match the appropriate forecasting technique to the pattern of the available time series data.

Once an appropriate technique has been chosen, the methodology involves analyzing the time series data in such a way that different components that are present can be estimated. The different estimates are then combined in order to produce a forecast.

2.1.1 Forecasting Methods

Forecasting methods can be divided into two basic types:

1. Qualitative Methods

2. Quantitative Methods

Qualitative Methods generally use the opinions of experts to subjectively predict future events. Such methods are used when historical data concerning the events to be predicted either are not available at all or are scarce. For example, consider a situation when a new product is being introduced in the market. No historical sales data for the product being available, a company has to rely on expert opinion (which can be supplied by members of its sales force and the market research team) to forecast future sales. This method is also used to predict changes in the historical data patterns.

Quantitative Methods involves the analysis of historical data in an attempt to predict future values of a variable of interest. Quantitative methods can be grouped into two types: *time series* and *causal*. The most common quantitative methods are the time series models. In such models, the historical data on the variable to be forecast is analyzed in an attempt to identify a data pattern.Thus the forecasting is based solely on extrapolating the historical data. Therefore time series modelling is quite effective under unchanged conditions. But it loses its effectiveness if there are sudden changes in the conditions.

The use of causal forecasting models involves the identification of other variables that are related to the variable to be predicted. Once these related variables has been identified, a statistical model that describes the relationship between these variables and the variable to be forecast is developed. The statistical relationship derived is then used to forecast the variable of interest. Causal methods are of interest in the business world because they allow the management to evaluate the impact of various alternative policies. For example, the management might want to predict how various price structures and levels of advertising expenditures will affect sales. A causal model relating these variables could be used here. However, causal models have its own disadvantages. First of all, developing the model is sometimes difficult. Secondly, the historical data for all the variables are necessary, i.e., for both the dependent (the variable to be forecasted) and the independent variables. And also, the ability of the forecaster to predict future values of the dependent variable depends on the accuracy of the prediction of the future values of the independent variables.

2.2 Applications of Time Series

It is of a matter of great interest to researchers to devise an efficient tool to forecast the future values of a time series. These tools (or methods) and their effectiveness in forecasting will be dealt in subsequent chapters. One of the major interests in forecasting future values of a time series lies in the field of finance. And one of the interesting areas is forecasting exchange rates of different currencies and study their changing patterns. This is what the current thesis deals with.

2.2.1 Application of Time Series in Predicting Exchange Rates

As mentioned before, prediction of exchange rates is of vital importance specially in the world trading market. These exchange rates exhibit a lot of fluctuations depending on the economy of the country and other factors which are not measurable. These rates have very complex patterns and are chaotic in nature. Hence, it is a big challenge to predict the future values.

In this thesis, the currency exchange rates of Australia, Great Britain, Canada, France, Germany, Japan and Switzerland with respect to the United States dollar have been used as experimental data. Inferences regarding the performances of the models discussed in this thesis are based on their prediction of these exchange rates.

The details of experiments have been discussed in the following chapters. To get a feeling of the complexity of the forecasting problem, the data used for the experiments are presented in the following graphs. Please note that the time series data has been normalized to the range of [0.1,0.8] so that the neural network methods perform well.

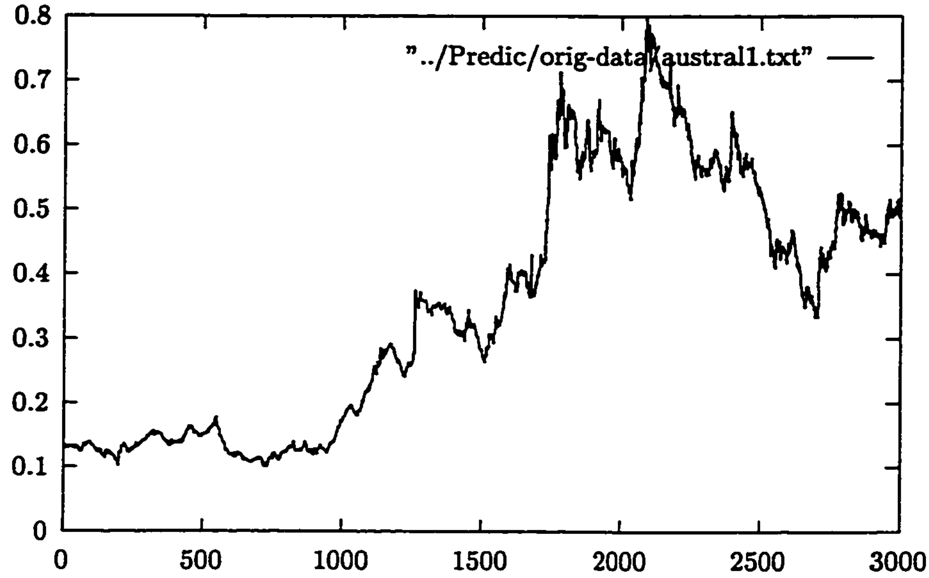


Figure 1: Conversion rate of the Australian dollar with respect to the United States dollar

Fig.1 describes the changes in exchange rates of Australian dollar with respect to United States dollar. We see that there is a general upward trend up to the data point 2300 where the graph reaches its peak and then there is a downward slope. In between there are some local maxima and minima and there is hardly any smoothness in the curves.

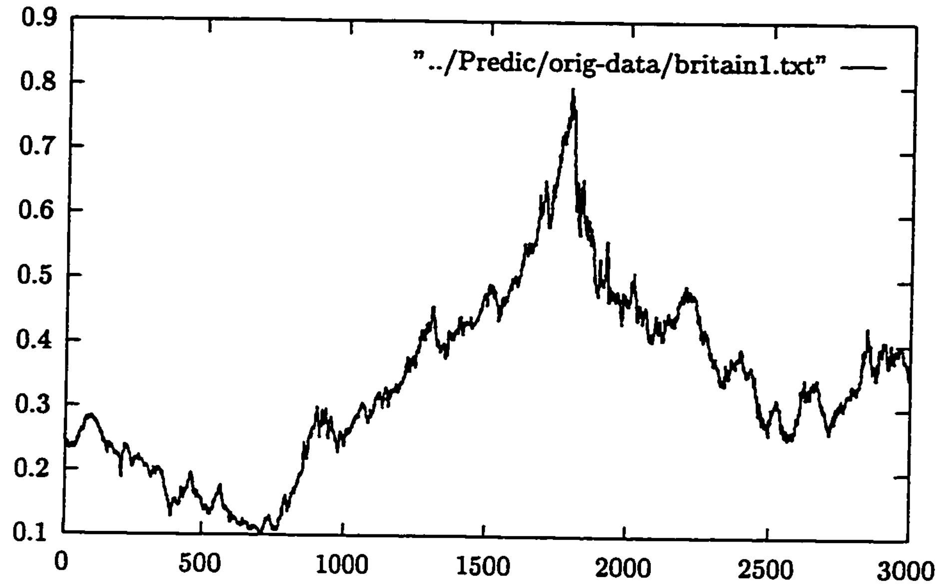


Figure 2: Conversion rate of the British Sterling with respect to the United States dollar

Fig.2 describes the changes in exchange rates of the English Sterling with respect to the United States dollar. Here the fluctuations are even more random and the pattern is more chaotic than the Australian exchange pattern as illustrated in Fig.1. However, we can loosely say that there is a general upward trend in and around 2000, the peak is reached around the data point 2300 and downward trend after. The minimum point is around 700.

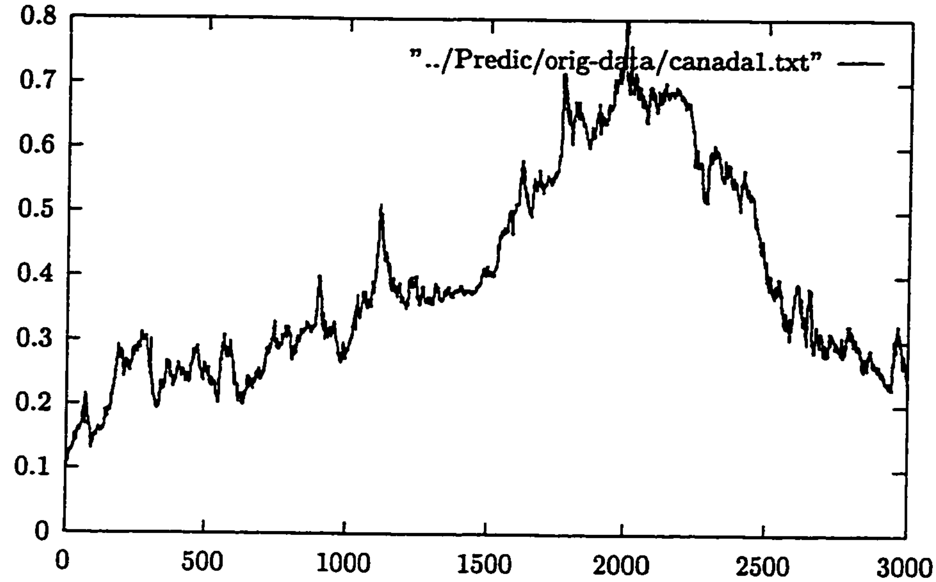


Figure 3: Conversion rate of the Canadian dollar with respect to the United States dollar

Fig.3 denotes the exchange rate variation of the Canadian dollar with respect to the United States dollar. Here the pattern is more complicated. There is a very gradual upward trend and the maximum is reached around 2000 and then there is a downward trend. There is no minimal dip as seen in Fig.2 for Sterling-U.S. dollar conversion.

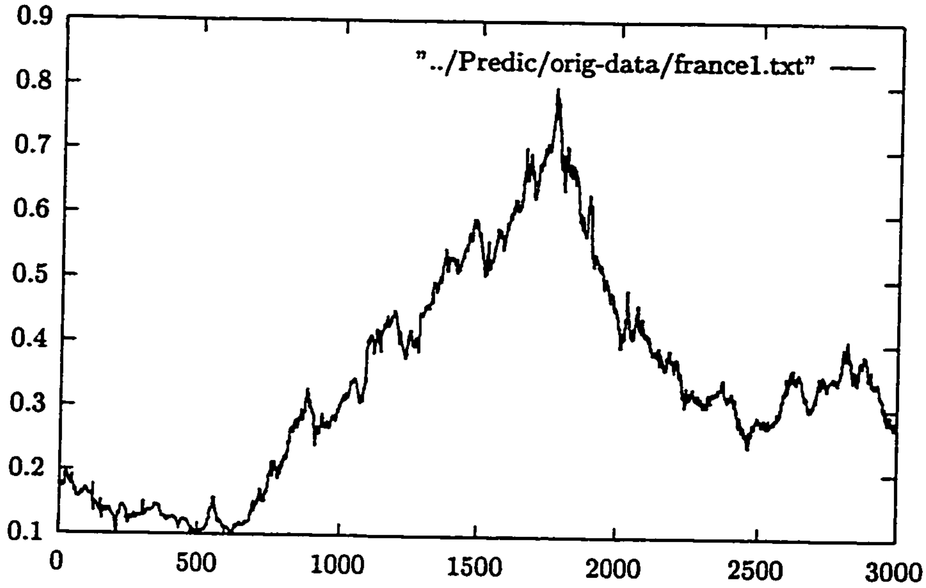


Figure 4: Conversion rate of the French Franc with respect to the United States dollar

It is interesting to note that the graphs illustrated in figures Fig. 2 and Fig. 4 have similar appearance. As seen in Fig. 2, this curve has an upward trend from about 700 to 1700, the maximum is reached around 1700 and the minimum around 600.

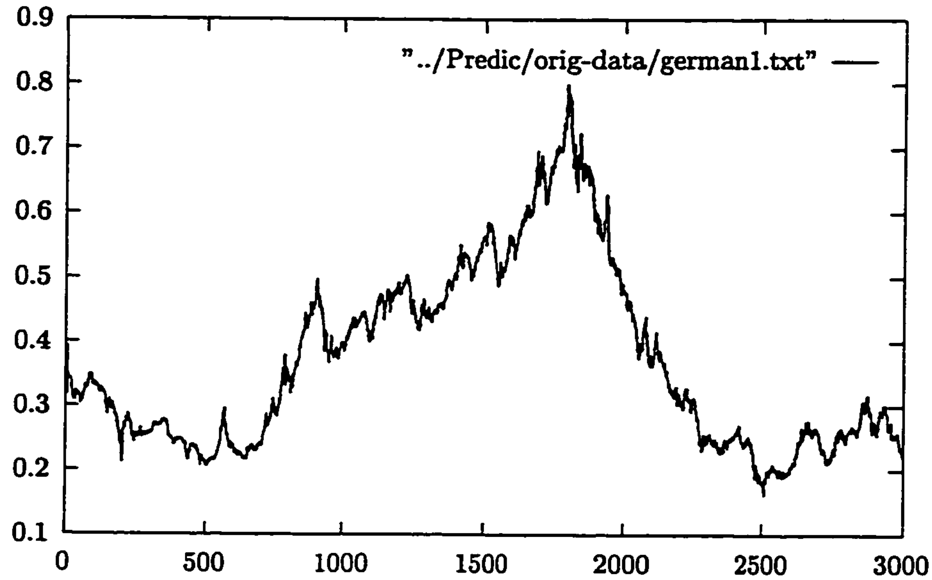


Figure 5: Conversion rate of the German Mark with respect to the United States dollar

The appearance can be compared to that in Fig.2. It has an increasing trend from 700 to 1700 and a steep slope downwards from 1700 to 2000. The minimum point is around 2000.

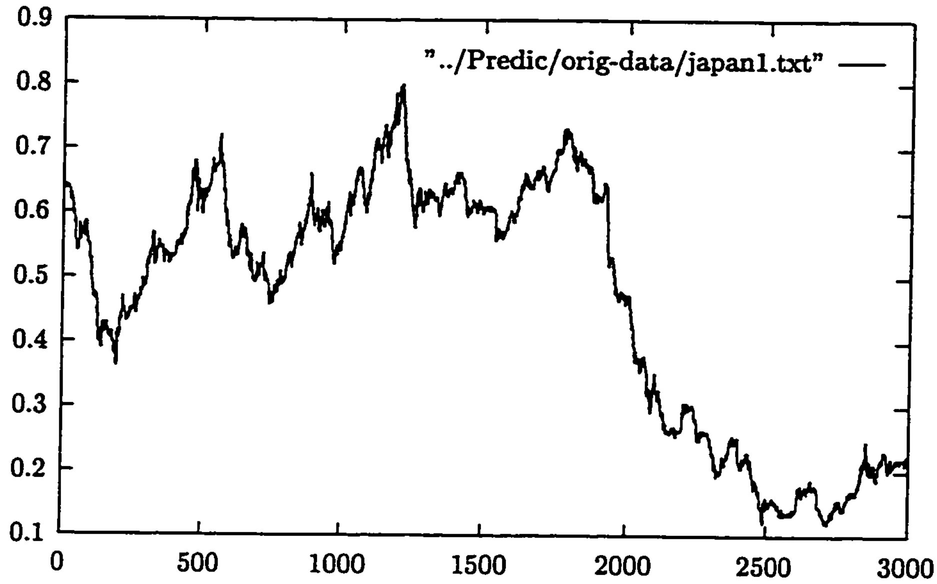


Figure 6: Conversion rate of the Japanese Yen with respect to the United States dollar

The graph illustrated in Fig.6 that describes the exchange rates of the Japanese Yen with respected to the U.S. dollar is totally different from the previous ones. Here we notice that the graph is quite flat up to 1700 and suddenly begins a downward trend after that point and reaches a minimum at 2500. Then it starts to go up very slowly.

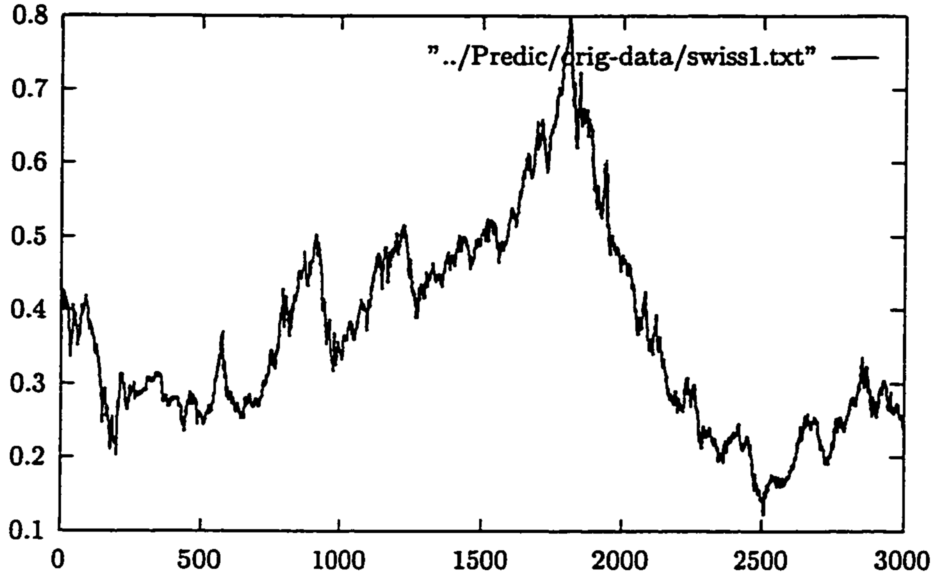


Figure 7: Conversion rate of the Swiss Franc with respect to the United States dollar

The changes in the conversion rates of the Swiss Franc with respect to the U.S. dollar is illustrated in Fig.7. We see that the pattern of the curve can be compared with that of the French Franc in Fig.4 and German Mark Fig.5.

So, in short, we may conclude that the behaviour of the European market in relation to the United States were quite comparable to each other. This is depicted by the graphs of the exchange rates.

2.2.2 Design of Experiments in Predicting Exchange Rates

Let the time series be defined by the points $x_0, x_1, \cdots$. Let x_i be the current point. Our aim is to predict from the known values $x_{i-k}, x_{i-k+1}, \cdots$ the value of x_i. $[x_{i-k}, x_{i-k+1}, \cdots, x_i]$ is called a window of size k.

The number of days (in advance) on which the prediction is done is called the prediction horizon.

In order to implement the methods of prediction discussed in the thesis, the following sequence of experiments were conducted.

The exchange rates of Australia, Britain, Canada, France, Germany, Japan and

Switzerland with the U.S. dollar were used as input time series data. Experiments were conducted using 3000 data points from each country and predictions were made in the following sequence:

1. Predicting the exchange rates of each country using window sizes of 3, 4, 5 and prediction horizon of 1 day, 2 days and 3 days.

2. Predicting the averages of the exchange rates of each country using window size of 3, 4 and 5 and prediction horizon of 3 and 5 days.

All the experiments were carried out using the C Programming Language on the computer. The results of each experiments are presented in subsequent chapters.

2.2.3 Application of Time Series in Physiology

Apart from scientific research and financial planning, time series is used in the areas of physiology and medicine. This section briefly deals with that aspect of time series. In real life we deal with events that are constantly changing with time. Even the living organisms are not constant in time, there are sub-cellular, cellular and super-cellular activities like cell divisions, respiration, blood pressure regulation etc. taking place at every instant of time. These activities generate complex patterns and exhibit what is known as *chaotic dynamics*. These are aperiodic rhythms sensitive to initial conditions. It has been theoretically derived that some biological systems exhibit chaos, nevertheless, it is a matter of interest to biologists to recognize the patterns of time series. There are some standard as well as highly mathematical analysis involved for studying these time series.

The most basic type of time series is carried out by the human eye (according to Glass [9]).. it is in fact, an excellent pattern recognition device and is capable of carrying out the sophisticated analysis needed to classify time series. For example, the interpretation of complex Electrocardiograms (ECGs) requires nothing more than some knowledge in cardiology combined with measurement of timing of occurrence of beats on comparatively short records. Interpretation of Electroencephalograms (EEGs) is carried out in a similar manner by skilled clinicians who have the knowledge of how to interpret the frequency, amplitude and morphology of recordings of electrical activity from different scalp positions. This is sufficient for identifying a great number of clinical disorders.

In research, most of the time series analysis takes place with the analysis of mean and standard deviation. These simple statistics can provide important physiological information. For example, the mean heart rate can be used to determine the level of exertion and a low standard deviation can be associated with pathology. The respiratory cycles can change with age.

The other methods uses the concept of non-linear dynamical systems. This has a strong impact on research in time series analysis in physiology and medicine. However, the discussion of this aspect is beyond the scope of the current thesis work. Prediction of chaotic time series is discussed briefly in Chapter 7.

Thus we have seen the importance of time series in our lives. The task of a forecaster is not complete until the forecasting errors are analyzed. The next chapter deals with errors in forecasting and how they influence the performance of a forecasting system.

Chapter 3

Measurement of Errors in Forecasting a Time Series

As stated in the previous chapters, time series play a very important role in our day-to-day lives. Hence, it is important to predict the future values of a time series with as much accuracy as possible.The closer the estimation, the better is the technique of prediction. Various methods of predicting a time series have been presented in details in the subsequent chapters.

Now, the question arises : "How do we measure the performance of a given prediction system ?". Well, the answer is not very simple as there are plenty of "performance metrics" available which can be used to measure the quality of prediction. Some important performance measures (or metrics) are being presented in this chapter. For further details, please refer to Armstrong and Callopy [1],Caldwell [5].

Before proceeding to discuss different performance measures, it is better to get an understanding of the characteristics of performance measures. This is being presented in the next section. Different performance measures and their significance are given in the subsequent sections.

3.1 Characteristics of Performance Measures

Armstrong and Callopy [1], gave some metrics to measure the performances of a time series prediction. These metrics were based typically in measuring the errors between the actual and the predicted values. The characteristics of these metrics are listed below :

1. Reliability: if the metric is consistent over and across different data series. One method of determining the reliability of a metric is to rank it against others using different data sets.

2. Validity: if the metric is measuring what it is supposed to measure. The way to go about it is to check if different metrics are measuring the the same thing for a given data set.

3. Expense: it is the cost of implementation of the metric (computing power needed to calculate it). In today's world of fast machines, the comparison of computing power of different metrics is insignificant. However, for iterative computations, like training process of a neural network, if one is interested to constantly monitor the performance, then the cost of the metric becomes worth consideration.

4. Understandability: if the metric is hard to understand, it would be hard to apply and interpret. Hence, most metrics are straightforward calculations and easily applicable. However, the interpretation of results of various metrics might differ depending upon their applications. Hence, studies on prediction systems should also include how results are interpreted.

5. Sensitivity: it is the response of the performance of the system (measured by the metric) to the variation of system parameters.

3.2 Some Performance Metrics

The analysis of prediction performance typically involves calculation of errors between desired (actual) and calculated results. Some of the traditional statistical forecasting error measures are listed below. These are used in measuring the performance of different forecasting techniques in predicting the future values of exchange rates of Australia, Britain, Canada, France, Germany, Japan and Switzerland with respect to the U.S. dollar. In addition to the traditional performance metrics, an additional measure called the directional symmetry *DS* has been presented at the end of the section. This metric is useful for measuring the performance of a neural network.

	p_i	be the	predicted value.
Let:	a_i	be the	actual value.
	n	be the	total number of data points.

1. *MSE : Mean Squared Error*

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(a_i - p_i)^2$$

MSE averages the total squared error over n points.

2. *RMSE : Root Mean Squared Error*

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(a_i - p_i)^2}$$

RMSE is the square root of MSE.

3. *MAE : Mean Absolute Error*

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|a_i - p_i|$$

MAE is the averaged absolute error value over n points.

4. $MAPE$: *Mean Absolute Percent Error*

$$MAPE = \frac{1}{n}100\sum_{i=1}^{n}\frac{|a_i - p_i|}{|a_i|}$$

MAPE represents the average error with respect to the true value over n points.

The 4 traditional metrics are not ideal for decision-making, and hence these may not be a matter of interest to the developers of financial forecasting techniques (Armstrong and Collopy [1]). $RMSE$ does describe the magnitude of the error and therefore is more useful to decision making than others. However, certain tests indicate that $RMSE$ may not be a reliable measure. The $MAPE$ is not very useful in terms of decision-making because the expression of error as a percentage does not give any direct implication of the performance.

In the case of decision-making, interest lies in the accuracy of predicting the direction of movement rather than the magnitude of error. Thus, in the interest of decision making, a measure to determine the direction is given below :

5. DS : *Directional Symmetry*

$$DS = 100\frac{1}{n}\sum_{i=1}^{n} d_i$$

where $d_i = \begin{cases} 1 & \text{if } (a_i - a_{i-1})(p_i - p_{i-1}) > 0 \\ 0 & \text{otherwise} \end{cases}$

Hence a $DS = 47\%$ means that the predicted direction was correct in 47% of the predictions. Or, in other words, the directional errors (DIR) was $100 - 47 = 53\%$. That is, the predicted direction was not correct in 53% of the cases. However, this does not give the magnitude of the movement. Hence, lower is the DIR value, better is the prediction of direction of movement of the data.

Please note that in order to measure the performance of a forecasting method, the directional errors DIR have been considered.

Due to biases that may occur during relatively long-term trends, it is advisable to first test the normalcy in the data.

All the above five performance metrics have been calculated and presented in tables for each prediction method.

Chapter 4

Forecasting Time Series using Statistical Methods

Statisticians, from very early times, have been interested and curious about time series because of their influence on every walks of life. Hence, they sought to develop models and theories to support their analysis to forecast the future values of a time series. In this chapter, some popular statistical forecasting models have been presented along with their effectiveness in forecasting future values of a time series. For experimentation, the exchange rates of Australia, Great Britain, Canada, France, Germany, Japan and Switzerland with respect to the United States dollar has been chosen as time series data. The statistical methods experimented include the following:

1. Multiple regression
2. Exponential Smoothing
3. Double Exponential Smoothing
4. Box-Jenkins Methodology

The underlying theory and derivations of the modelling equations have also been presented. Please refer to Bowerman and O'Connell [3] for further details. In addition to the above four methods a discussion of Winters' method which involves forecasting seasonal time series has been stated. Since the behaviour of the exchange data does not follow seasonal variations, it has not been considered suitable for experimentation.

For implementing the above mentioned statistical models, the following strategies were adopted to give best results. The predictions were made using inputs of window size. i.e., 3, 4, or 5 days. This enabled curve-fitting for a small range of values and hence better approximation. The prediction horizons were for 1, 2 and 3 days for the regular(raw) data of exchange rates. A prediction horizon of 3 and 5 days were used for predicting the averages of exchange rates.

4.1 Multiple Regression Analysis

In this section a quantitative forecasting technique has been presented, which is known as multiple regression. The forecasting method discussed below is interesting in the sense that it involves a causal multiple regression model to forecast the future values of a time series. Multiple regression models are widely used as causal forecasting models. Hence its application in time series makes the forecasting of future values very effective.

A general multiple regression model has the following form:

$$y_t = \beta_0 + \beta_1 x_{t1} + \cdots + \beta_k x_{tk} + \varepsilon_t \tag{4.1.1}$$

where y_t denotes the dependent variable in time t; k represents the number of independent variables in the model; $x_{t1}, x_{t2}, \cdots, x_{tk}$ represents the values of the k independent variables in time t; $\beta_1, \beta_2, \cdots, \beta_k$ are unknown parameters relating the dependent variable y_t to the k independent variables $x_{t1}, x_{t2}, \cdots, x_{tk}$; and ε_t is a random error component that describes the influence on y_t of all factors other than the k independent variables. It is assumed that the expected value of ε_t is 0. According to Bowerman and O'Connell [3], the general multiple regression model states that time series y_t can be represented by an average level that changes over the time according to the function on the right hand side of the above equation 4.1.1. The random fluctuations are responsible for the time series to deviate from this average level. Let this average level be represented by μ_t, i.e.,

$$\mu_t = \beta_0 + \beta_1 x_{t1} + \cdots + \beta_k x_{tk} \tag{4.1.2}$$

The value μ_t represents the average of all the values of the dependent variable y_t that could ever possibly be observed when the values of the independent variables are fixed at $x_{t1}, x_{t2}, \cdots, x_{tk}$. If $b_0, b_1, \cdots, b_k$ are the least squares estimates of $\beta_0, \beta_1, \beta_2, \cdots, \beta_k$, then the point estimate of μ_t is given by:

$$\hat{y}_t = b_0 + b_1 x_{t1} + b_2 x_{t2} + \cdots + b_k x_{tk}$$

Since we assume that the error component ε_t averages out to 0 in the long run, and that 0 is therefore a reasonable guess for any future value of ε_t, it follows that $\hat{y}_t$ is the point forecast of the actual time-series value

$$y_t = \mu_t + \varepsilon_t$$

In order to find how far $\hat{y}_t$ is from y_t and μ_t, we have to make certain assumptions on the error component ε_t. These are:

1. For each and every period t the random error follows a normal distribution.

2. The variance of y_t, which measures the spread of all the potential values of the dependent variable y_t around the average level μ_t, is the same for each and every value of t.

3. The time series values $y_1, y_2, \cdots$ in different periods are statistically independent of or not related to each other.

The above three assumptions are the so-called *inference assumptions.*

4.1.1 Building a Multiple Regression Model

In order to construct a multiple regression model to provide accurate forecasts of a given time series, the specification of the appropriate independent variables and their functional relationship with the dependent variable is very important. This functional relationship may be *linear* or *quadratic* of two or more variables multiplied together to form an *interaction variable.*

Let us consider the following regression model:

$$y_t = \mu_t + \varepsilon_t$$

where μ_t is the average level of the time series at time t and ε_t is the random error component. The dependent variable y_t is said to be related to a single independent variable x_{t1} in a *linear fashion* if

$$\mu_t = \beta_0 + \beta_1 x_{t1}$$

and in a *quadratic fashion* if

$$\mu_t = \beta_0 + \beta_1 x_{t1} + \beta_{t2} x_{t1}^2$$

This means that the average level is *randomly fluctuating* around an average level μ_t that changes in a linear or quadratic fashion as x_{t1} increases. Thus the average level μ_t of the time series is increasing/decreasing at an increasing or decreasing rate as x_{t1} increases.

For our present purposes, only linear regression models have been considered.

Hence the forecasted output for a linear regression model is given by:

$$y_t = b_0 + b_1 x_t \tag{4.1.3}$$

4.1.2 The Initial Estimates

The initial estimates of β_0 and β_1 denoted by b_0 and b_1 respectively are given by the least squares methods described by the following equations:

$$b_1 = \frac{n\sum_{t=1}^{n} t y_t - (\sum_{t=1}^{n} t)(\sum_{t=1}^{n} y_t)}{n\sum_{t=1}^{n} t^2 - (\sum_{t=1}^{n} t)^2} \tag{4.1.4}$$

$$b_0 = \frac{\sum_{t=1}^{n} y_t}{n} - b_1(\frac{\sum_{t=1}^{n} t}{n}) \tag{4.1.5}$$

A detailed description of the measures of contribution of the independent variables in prediction has been stated in appendix A.1.

4.1.3 The Multiple Regression Algorithm

Step 1: Choose an initial set of input data.

Step 2: Compute the initial estimates b_0 and b_1 using equations 4.1.4 and 4.1.5.

Step 3: Compute the output using equation 4.1.3.

Step 4: IF stopping condition is met, THEN stop ELSE go to Step 3.

Please note that the stopping condition here is to carry out the regression method for all 3000 data.

4.1.4 Experimental Results

The performance of multiple regression method is illustrated in the following graph: The overall *MSE* ranges from 0.00349 to 0.00409. The mean is around 0.0040. The

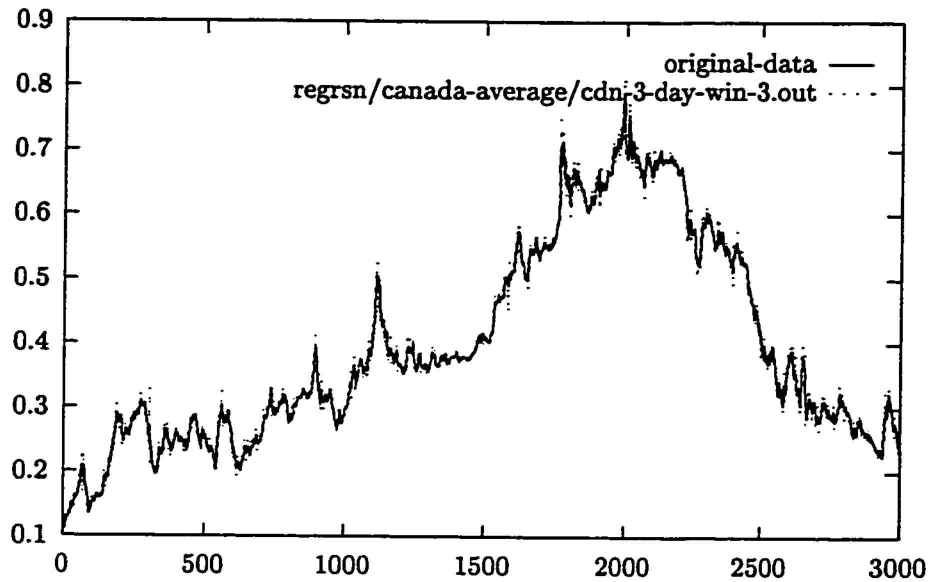


Figure 8: Predicting 3 days averages of Canadian-U.S. dollar exchange rates for window size 3 using multiple regression

range of *MSE* for the averages is also from 0.00349 to 0.00409, but the mean is about 0.0037. The range of *RMSE* is from 0.059 to 0.064 for exchange rates and from 0.0599 to 0.06334 for the averages of exchange rates. The mean of the former is around 0.063 and that for the latter is 0.062. *MAE* ranges from 0.00562 to 0.01051 for exchange rates and from 0.0057 to 0.01136 for the averages of exchange rates. The mean is about 0.007 for exchange rates and about 0.006 for the averages of exchange rates. *MAPE* ranges from 1.61 to 3.2601 for exchange rates and from 1.85 to 3.34 for the averages of exchange rates. The mean of the *MAPE* values is around 2.3 for the exchange rates and 2.15 for the averages of exchange rates. Finally, for the *DIR*, the value ranges from 0.484 to 0.5077 for exchange rates and from 0.4720 to 0.5147. The overall mean for the exchange rates is about 0.50 whereas that for the averages of exchange rates is about 0.48. The results are stated in the following tables.

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	1	0.00350	0.05919	0.00629	1.63987	0.4940
	3	2	0.00355	0.05959	0.00787	2.04784	0.4937
	3	3	0.00359	0.05991	0.00910	2.37831	0.4933
	4	1	0.00352	0.05934	0.00703	1.82297	0.4937
	4	2	0.00356	0.05969	0.00843	2.17705	0.4933
	4	3	0.00360	0.06004	0.00949	2.47147	0.4933
	5	1	0.00349	0.05907	0.00630	1.61768	0.4933
	5	2	0.00353	0.05944	0.00775	2.00245	0.4933
	5	3	0.00358	0.05984	0.00887	2.32526	0.4937
Britain	3	1	0.00371	0.06088	0.00639	1.95005	0.4840
	3	2	0.00374	0.06119	0.00784	2.40467	0.4840
	3	3	0.00380	0.06162	0.00913	2.80789	0.4843
	4	1	0.00372	0.06100	0.00695	2.12547	0.4840
	4	2	0.00377	0.06144	0.00836	2.56919	0.4843
	4	3	0.00381	0.06170	0.00956	2.93604	0.4843
	5	1	0.00372	0.06103	0.00630	1.93083	0.4843
	5	2	0.00376	0.06129	0.00780	2.39091	0.4843
	5	3	0.00380	0.06165	0.00910	2.79540	0.4843
Canada	3	1	0.00398	0.06311	0.00757	2.32648	0.4927
	3	2	0.00403	0.06345	0.00924	2.78404	0.4927
	3	3	0.00407	0.06383	0.01060	3.15433	0.4927
	4	1	0.00400	0.06321	0.00822	2.49081	0.4927
	4	2	0.00404	0.06360	0.00980	2.92544	0.4927
	4	3	0.00409	0.06398	0.01103	3.26093	0.4927
	5	1	0.00398	0.06305	0.00738	2.25170	0.4927
	5	2	0.00403	0.06345	0.00902	2.70462	0.4927
	5	3	0.00407	0.06378	0.01043	3.09846	0.4927
France	3	1	0.00385	0.06208	0.00592	1.93015	0.5007
	3	2	0.00387	0.06217	0.00706	2.27613	0.5010
	3	3	0.00389	0.06235	0.00804	2.58443	0.5010
	4	1	0.00385	0.06203	0.00626	2.02691	0.5010
	4	2	0.00387	0.06220	0.00736	2.36786	0.5010
	4	3	0.00391	0.06251	0.00833	2.68148	0.5007
	5	1	0.00383	0.06190	0.00562	1.81886	0.5010
	5	2	0.00387	0.06221	0.00689	2.22470	0.5007
	5	3	0.00390	0.06241	0.00801	2.57662	0.5007

Table 1: Predicting exchange rates using regression method

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	1	0.00399	0.06316	0.00624	1.95369	0.4897
	3	2	0.00403	0.06349	0.00759	2.33723	0.4897
	3	3	0.00408	0.06384	0.00874	2.67718	0.4893
	4	1	0.00401	0.06334	0.00667	2.09028	0.4897
	4	2	0.00405	0.06367	0.00791	2.44961	0.4893
	4	3	0.00409	0.06394	0.00905	2.76658	0.4897
	5	1	0.00401	0.06332	0.00600	1.91879	0.4893
	5	2	0.00404	0.06358	0.00742	2.31013	0.4897
	5	3	0.00409	0.06393	0.00871	2.68442	0.4897
Japan	3	1	0.00397	0.06305	0.00668	1.89479	0.5043
	3	2	0.00401	0.06332	0.00819	2.25459	0.5047
	3	3	0.00405	0.06362	0.00955	2.60456	0.5047
	4	1	0.00399	0.06314	0.00726	2.03426	0.5047
	4	2	0.00402	0.06342	0.00865	2.37822	0.5047
	4	3	0.00407	0.06380	0.00988	2.70636	0.5047
	5	1	0.00397	0.06302	0.00650	1.85115	0.5047
	5	2	0.00402	0.06339	0.00806	2.24530	0.5047
	5	3	0.00406	0.06370	0.00940	2.59280	0.5047
Switzerland	3	1	0.00396	0.06296	0.00707	2.21385	0.5070
	3	2	0.00402	0.06343	0.00867	2.70071	0.5073
	3	3	0.00407	0.06376	0.01008	3.11202	0.5073
	4	1	0.00400	0.06323	0.00765	2.40993	0.5073
	4	2	0.00404	0.06353	0.00912	2.84087	0.5073
	4	3	0.00406	0.06375	0.01051	3.22563	0.5077
	5	1	0.00398	0.06308	0.00687	2.18530	0.5073
	5	2	0.00400	0.06328	0.00855	2.65413	0.5077
	5	3	0.00405	0.06361	0.01011	3.10161	0.5073

Table 2: Predicting exchange rates using regression method (contd.)

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	3	0.00359	0.05990	0.00882	2.34411	0.4720
	3	5	0.00349	0.05907	0.00616	1.60257	0.4883
	4	3	0.00360	0.06004	0.00944	2.52852	0.4717
	4	5	0.00354	0.05946	0.00767	2.02263	0.4883
	5	3	0.00360	0.05997	0.00933	2.52314	0.4717
	5	5	0.00352	0.05931	0.00723	1.92855	0.4883
Britain	3	3	0.00375	0.06128	0.00708	2.18199	0.5067
	3	5	0.00371	0.06088	0.00626	1.91456	0.4867
	4	3	0.00380	0.06168	0.00971	2.98732	0.5067
	4	5	0.00374	0.06119	0.00782	2.39684	0.4867
	5	3	0.00380	0.06167	0.00973	2.99691	0.5067
	5	5	0.00374	0.06112	0.00753	2.30870	0.4867
Canada	3	3	0.00398	0.06309	0.00825	2.48536	0.4833
	3	5	0.00396	0.06296	0.00735	2.24189	0.4823
	4	3	0.00408	0.06389	0.01136	3.36056	0.4833
	4	5	0.00401	0.06334	0.00913	2.73634	0.4823
	5	3	0.00408	0.06386	0.01133	3.35549	0.4833
	5	5	0.00400	0.06322	0.00870	2.61636	0.4823
France	3	3	0.00387	0.06221	0.00641	2.07348	0.4807
	3	5	0.00385	0.06204	0.00570	1.84917	0.4860
	4	3	0.00393	0.06267	0.00868	2.78745	0.4807
	4	5	0.00388	0.06229	0.00705	2.27006	0.4860
	5	3	0.00393	0.06265	0.00868	2.79108	0.4807
	5	5	0.00387	0.06222	0.00674	2.17746	0.4860

Table 3: Predicting averages of exchange rates using regression method

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	3	0.00401	0.06336	0.00694	2.15152	0.5073
	3	5	0.00399	0.06315	0.00609	1.90925	0.5033
	4	3	0.00407	0.06380	0.00954	2.85554	0.5077
	4	5	0.00402	0.06341	0.00769	2.34901	0.5033
	5	3	0.00407	0.06379	0.00955	2.86534	0.5077
	5	5	0.00401	0.06334	0.00739	2.27253	0.5033
Japan	3	3	0.00407	0.06379	0.00967	2.62656	0.5003
	3	5	0.00397	0.06304	0.00655	1.85245	0.4913
	4	3	0.00409	0.06396	0.01030	2.78903	0.5010
	4	5	0.00403	0.06347	0.00828	2.29017	0.4900
	5	3	0.00409	0.06392	0.01035	2.79959	0.4997
	5	5	0.00402	0.06341	0.00799	2.22659	0.4900
Switzerland	3	3	0.00398	0.06306	0.00790	2.43777	0.5147
	3	5	0.00398	0.06311	0.00755	2.35098	0.4960
	4	3	0.00407	0.06381	0.01103	3.34480	0.5147
	4	5	0.00400	0.06327	0.00882	2.71241	0.4963
	5	3	0.00407	0.06382	0.01111	3.36901	0.5147
	5	5	0.00399	0.06320	0.00855	2.63371	0.4963

Table 4: Predicting averages of exchange rates using regression method (contd.)

4.2 The Exponential Smoothing Technique

In this section, we consider a time series with *no trend.* This technique involves a choice of an appropriate *smoothing constant*: it determines the influence of past observations of a time series on forecasting. A large smoothing constant gives a larger weight to the more recent observations in the time series and results in a more rapid response to changes in the time series.

4.2.1 The Modelling

Suppose at the end of time period $T-1$ we obtained a set of observations for the time series denoted by:

$$y_1,\ y_2,\ \cdots\ y_{T-1}$$

. Given these observations we wish to estimate β_0 which is the average level of the time series. Let this estimate be $b_0(T)$ for the time period T.

Let $b_0(T-1)$ given by:

$$b_0(T-1) = \bar{y} = \sum_{t=1}^{T-1} y_t/(T-1)$$

be the forecast for future values. The forecasting error for the time period T is given by:

$$e_T = y_T - b_0(T-1)$$

Thus it is the difference between the observed value in the period T and the forecast made for the period T in period $T-1$. The updated estimate is given by:

$$b_0(T) = b_0(T-1) + \alpha[y_T - b_0(T-1)] \qquad (4.2.1)$$

Thus the new estimate is dependent on the old estimate by some factor α known as the *smoothing constant.* If the old estimate forecasted a value for period T that was too low, then the new estimate is higher, and vice versa. The magnitude of this up and down adjustment is taken care by the smoothing constant α. Let us define $S_T = b_0(T)$. Then (4.2.1) can be rewritten as:

$$S_T = S_{T-1} + \alpha[y_T - S_T(T-1)] = \alpha y_T - (1-\alpha)S_{T-1} \qquad (4.2.2)$$

This equation 4.2.2 defines the updating procedure called the *Simple Exponential Smoothing.* We call S_T the *the smoothed estimate* or *smoothed statistic.*

4.2.2 The Initial Estimates

We can change the time origin so that the initial estimate of β_0 is assumed to be generated in time period zero. Let the initial estimate be denoted by S_0. S_0 can be calculated by taking the average of initial set of observations if they are available, or it is set equal to the first observed value of the time series.

Hence S_0 can be calculated as:

$$S_0 = \frac{\sum_{t=1}^{T} y_t}{T} \qquad (4.2.3)$$

4.2.3 Updating the Estimates

The equation 4.2.2 can be expressed as follows:

$$S_t = \alpha y_t - (1-\alpha)S_{t-1} \qquad (4.2.4)$$

Equation 4.2.4 can be used to update the estimates for each time period t from period 1 to period T , the present time period. The smoothed estimate for the current period T can be represented as a combination of the past observations by the following recursive scheme:

Putting $t = T$ in 4.2.4 we have:

$$S_{T-1} = \alpha y_{T-1} - (1-\alpha)S_{T-2} \tag{4.2.5}$$

Substituting S_{T-1} in 4.2.2 we have:

$$\begin{aligned} S_T &= \alpha y_T - (1-\alpha)[\alpha y_{T-1} - (1-\alpha)S_{T-2}] \\ &= \alpha y_T - \alpha(1-\alpha)y_{T-1} - (1-\alpha)^2 S_{T-2} \end{aligned} \tag{4.2.6}$$

4.2.4 Forecasting the Future Values

Suppose we are in the current period T. The current estimate of β_0 is $S_T = b_0(T)$. We wish to forecast the time series for a future period $T+\tau$ in period T. This can be done as follows:

Since the model is given by

$$y_t = \beta_0 + \epsilon_t$$

the forecast is given by:

$$\hat{y}_{T+\tau} = S_T \tag{4.2.7}$$

where S_T is the current estimate of β_0.

The impacts of the smoothing constant α in predicting future values are discussed in details in appendix A.2.

4.2.5 The Exponential Smoothing Algorithm

Step 1: Choose an initial estimate of β_0 defined by equation 4.2.3.

Step 2: Use the smoothing equation 4.2.2 to simulate all the historical data available for a particular α.

Step 4: Calculate the output using equation 4.2.7.

Step 5: IF all data has been used THEN stop ELSE go to Step 2.

4.2.6 Experimental Results

An example of the performance of exponential smoothing method is illustrated in the following graph: The overall MSE errors range from 0.0047 to 0.00552. The mean is

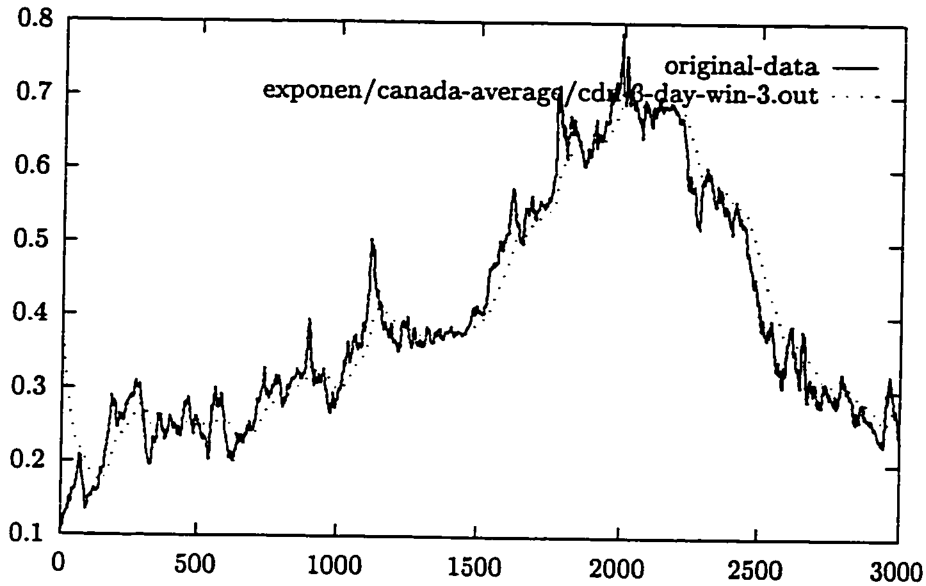


Figure 9: Predicting 3 days averages of Canadian-U.S. dollar exchange rates for window size 3 using exponential smoothing

around 0.0054. The range of MSE errors for the averages is from 0.00467 to 0.00552, the mean is about 0.005. The range of $RMSE$ is from 0.06856 to 0.0734 for exchange rates and from 0.0687 to 0.0732 for the averages of exchange rates. The mean of the former is around 0.073 and that for the latter is 0.070. MAE ranges from 0.02655 to 0.3235 for exchange rates and from 0.02649 to 0.03239 for the averages of exchange rates. The mean is about 0.028 for exchange rates and about 0.027 for the averages. $MAPE$ ranges from 7.56 to 9.42 for exchange rates and from 7.48 to 9.38 for the averages of exchange rates. The mean of the $MAPE$ values is around 8.3 for the exchange rates and 8.96 for the averages of exchange rates. Finally, for the DIR, the value ranges from 0.484 to 0.5077 for exchange rates and from 0.4720 to 0.5147. The overall mean for the exchange rates is about 0.50 whereas that for the averages of exchange rates is about 0.48. The results are summarized in the following tables.

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	1	0.00508	0.07127	0.02856	9.57696	0.4940
	3	2	0.00508	0.07125	0.02856	9.57738	0.4937
	3	3	0.00507	0.07119	0.02856	9.57462	0.4933
	4	1	0.00508	0.07125	0.02856	9.57738	0.4937
	4	2	0.00507	0.07119	0.02856	9.57462	0.4933
	4	3	0.00506	0.07116	0.02856	9.57318	0.4933
	5	1	0.00507	0.07119	0.02856	9.57462	0.4933
	5	2	0.00506	0.07116	0.02856	9.57318	0.4933
	5	3	0.00507	0.07121	0.02856	9.57505	0.4937
Britain	3	1	0.00470	0.06856	0.02656	8.32390	0.4840
	3	2	0.00470	0.06857	0.02655	8.32361	0.4840
	3	3	0.00472	0.06869	0.02655	8.32836	0.4843
	4	1	0.00470	0.06857	0.02655	8.32361	0.4840
	4	2	0.00472	0.06869	0.02655	8.32836	0.4843
	4	3	0.00471	0.06865	0.02655	8.32513	0.4843
	5	1	0.00472	0.06869	0.02655	8.32836	0.4843
	5	2	0.00471	0.06865	0.02655	8.32513	0.4843
	5	3	0.00472	0.06871	0.02655	8.32781	0.4843
Canada	3	1	0.00539	0.07344	0.02832	9.42555	0.4927
	3	2	0.00539	0.07340	0.02832	9.40732	0.4927
	3	3	0.00539	0.07339	0.02832	9.39131	0.4927
	4	1	0.00539	0.07340	0.02832	9.40732	0.4927
	4	2	0.00539	0.07339	0.02832	9.39131	0.4927
	4	3	0.00539	0.07340	0.02832	9.37172	0.4927
	5	1	0.00539	0.07339	0.02832	9.39131	0.4927
	5	2	0.00539	0.07340	0.02832	9.37172	0.4927
	5	3	0.00538	0.07336	0.02832	9.35014	0.4927
France	3	1	0.00491	0.07006	0.02711	9.00454	0.5007
	3	2	0.00490	0.06997	0.02712	8.99938	0.5010
	3	3	0.00489	0.06996	0.02712	8.99927	0.5010
	4	1	0.00490	0.06997	0.02712	8.99938	0.5010
	4	2	0.00489	0.06996	0.02712	8.99927	0.5010
	4	3	0.00491	0.07006	0.02713	9.01130	0.5007
	5	1	0.00489	0.06996	0.02712	8.99927	0.5010
	5	2	0.00491	0.07006	0.02713	9.01130	0.5007
	5	3	0.00491	0.07005	0.02713	9.01336	0.5007

Table 5: Predicting exchange rates using exponential smoothing method

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	1	0.00493	0.07019	0.02646	7.56779	0.4897
	3	2	0.00494	0.07029	0.02647	7.58676	0.4897
	3	3	0.00495	0.07038	0.02648	7.60720	0.4893
	4	1	0.00494	0.07029	0.02647	7.58676	0.4897
	4	2	0.00495	0.07038	0.02648	7.60720	0.4893
	4	3	0.00496	0.07040	0.02649	7.61726	0.4897
	5	1	0.00495	0.07038	0.02648	7.60720	0.4893
	5	2	0.00496	0.07040	0.02649	7.61726	0.4897
	5	3	0.00497	0.07050	0.02651	7.64026	0.4897
Japan	3	1	0.00552	0.07427	0.03238	8.33185	0.5043
	3	2	0.00551	0.07424	0.03235	8.32723	0.5047
	3	3	0.00551	0.07422	0.03233	8.32378	0.5047
	4	1	0.00551	0.07424	0.03235	8.32723	0.5047
	4	2	0.00551	0.07422	0.03233	8.32378	0.5047
	4	3	0.00552	0.07427	0.03230	8.33161	0.5047
	5	1	0.00551	0.07422	0.03233	8.32378	0.5047
	5	2	0.00552	0.07427	0.03230	8.33161	0.5047
	5	3	0.00551	0.07425	0.03228	8.32975	0.5047
Switzerland	3	1	0.00518	0.07197	0.02964	8.70733	0.5070
	3	2	0.00520	0.07210	0.02964	8.72693	0.5073
	3	3	0.00520	0.07208	0.02964	8.72718	0.5073
	4	1	0.00520	0.07210	0.02964	8.72693	0.5073
	4	2	0.00520	0.07208	0.02964	8.72718	0.5073
	4	3	0.00518	0.07196	0.02963	8.71268	0.5077
	5	1	0.00520	0.07208	0.02964	8.72718	0.5073
	5	2	0.00518	0.07196	0.02963	8.71268	0.5077
	5	3	0.00518	0.07195	0.02964	8.71158	0.5073

Table 6: Predicting exchange rates using exponential smoothing method (contd.)

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	3	0.00505	0.07105	0.02839	9.53248	0.4720
	3	5	0.00503	0.07094	0.02827	9.50143	0.4883
	4	3	0.00504	0.07101	0.02839	9.53125	0.4717
	4	5	0.00503	0.07094	0.02827	9.50143	0.4883
	5	3	0.00504	0.07101	0.02839	9.53125	0.4717
	5	5	0.00503	0.07094	0.02827	9.50143	0.4883
Britain	3	3	0.00472	0.06870	0.02649	8.31447	0.5067
	3	5	0.00467	0.06837	0.02636	8.25786	0.4867
	4	3	0.00469	0.06847	0.02645	8.29043	0.5067
	4	5	0.00467	0.06837	0.02636	8.25786	0.4867
	5	3	0.00469	0.06847	0.02645	8.29043	0.5067
	5	5	0.00467	0.06837	0.02636	8.25786	0.4867
Canada	3	3	0.00536	0.07320	0.02809	9.38083	0.4833
	3	5	0.00535	0.07312	0.02792	9.31501	0.4823
	4	3	0.00536	0.07320	0.02809	9.38083	0.4833
	4	5	0.00535	0.07312	0.02792	9.31501	0.4823
	5	3	0.00536	0.07320	0.02809	9.38083	0.4833
	5	5	0.00535	0.07312	0.02792	9.31501	0.4823
France	3	3	0.00490	0.07002	0.02696	8.96108	0.4807
	3	5	0.00489	0.06990	0.02690	8.93564	0.4860
	4	3	0.00490	0.07002	0.02696	8.96108	0.4807
	4	5	0.00489	0.06994	0.02689	8.93676	0.4860
	5	3	0.00490	0.07002	0.02696	8.96108	0.4807
	5	5	0.00489	0.06994	0.02689	8.93676	0.4860

Table 7: Predicting averages of exchange rates using exponential smoothing method

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	3	0.00492	0.07017	0.02631	7.53349	0.5073
	3	5	0.00490	0.07002	0.02620	7.48069	0.5033
	4	3	0.00491	0.07006	0.02627	7.49485	0.5077
	4	5	0.00490	0.07002	0.02620	7.48069	0.5033
	5	3	0.00491	0.07006	0.02627	7.49485	0.5077
	5	5	0.00490	0.07002	0.02620	7.48069	0.5033
Japan	3	3	0.00552	0.07430	0.03239	8.32797	0.5003
	3	5	0.00549	0.07410	0.03222	8.28219	0.4913
	4	3	0.00553	0.07439	0.03249	8.34474	0.5010
	4	5	0.00552	0.07429	0.03238	8.31682	0.4900
	5	3	0.00552	0.07428	0.03239	8.32320	0.4997
	5	5	0.00552	0.07429	0.03238	8.31682	0.4900
Switzerland	3	3	0.00515	0.07179	0.02950	8.64364	0.5147
	3	5	0.00515	0.07178	0.02935	8.61415	0.4960
	4	3	0.00515	0.07179	0.02950	8.64364	0.5147
	4	5	0.00514	0.07171	0.02935	8.59970	0.4963
	5	3	0.00515	0.07179	0.02950	8.64364	0.5147
	5	5	0.00514	0.07171	0.02935	8.59970	0.4963

Table 8: Predicting averages of exchange rates using exponential smoothing method (contd.)

4.3 The Double Exponential Smoothing Technique

The Double Exponential Smoothing technique is quite similar to exponential smoothing method. In this case there is a *linear trend* observed in the time series. The average level of the time series changes over time in a linear fashion.

4.3.1 The Modelling

The appropriate model for the time series can thus be defined as:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t \tag{4.3.1}$$

This implies that the time series can be described by the trend implied by the straight line with slope β_1 and intercept β_0 combined with the random fluctuations which cause the time series to deviate from the trend line. In double exponential smoothing technique, the updating of the estimates of β_0 and β_1 are done using the following equation:

$$b_1(T) = \frac{\alpha}{1-\alpha}(S_T - S_T^{[2]} \tag{4.3.2}$$

$$b_0(T) = 2S_T - S_T^{[2]} - Tb_1(T) \tag{4.3.3}$$

Here S_T is the *single smoothed statistic* which is found by using the smoothing equation 4.2.4 and $S_T^{[2]}$ is the *double smoothed statistic*. It is found by applying the smoothing operation to the output of the single smoothing equation. i.e.,

$$S_T^{[2]} = \alpha S_T + (1-\alpha)S_{T-1}^{[2]} \tag{4.3.4}$$

Where $0 \leq \alpha \leq 1$ is called the smoothing constant. Suppose the data up to and including period T is available. Then the forecast for a future time period $T + \tau$ is given by:

$$\hat{y}_{T+\tau} = b_0(T) + b_1(T)(T+\tau) \quad (4.3.5)$$

4.3.2 The Initial Estimates

The initial estimates can be found by using the following equations. These are derived by substituting $T = 0$ in equations 4.3.2 and 4.3.3 and solving for S_T and $S_t^{[2]}$ in terms of b_0 and b_1.

$$S_0 = b_0(0) - (\frac{1-\alpha}{\alpha})b_1(0) \quad (4.3.6)$$

$$S_0^{[2]} = b_0(0) - 2(\frac{1-\alpha}{\alpha})b_1(0) \quad (4.3.7)$$

The values of b_0 and b_1 can be found either by regression method (using the least squares estimates) given in equations 4.1.4 and 4.1.5, or by using the exponential smoothing approach. In the latter case, the initial estimates are given by the following equations:

$$b_1(T-1) = \frac{(T-1)\sum_{t=1}^{(T-1)} ty_t - (\sum_{t=1}^{(T-1)} t)(\sum_{t=1}^{(T-1)} y_t)}{(T-1)\sum_{t=1}^{(T-1)} t^2 - (\sum_{t=1}^{(T-1)} t)^2} \quad (4.3.8)$$

$$b_0(T-1) = \frac{\sum_{t=1}^{(T-1)} y_t}{(T-1)} - b_1(T-1)(\frac{\sum_{t=1}^{(T-1)} t}{(T-1)}) \quad (4.3.9)$$

In case the initial estimates b_0 or b_1 are not available, the first observation is used to assign S_0 and $S_0^{[2]}$.

4.3.3 Updating the Estimates

Given the initial estimates, the smoothing equations

$$S_T = \alpha y_T - (1-\alpha)S_{T-1} \quad (4.3.10)$$

and 4.3.4 are used to update the estimates of β_0 and β_1.

4.3.4 Forecasting the Future Values

The forecasting equation for time period T is given by:

$$\hat{y}_{T+\tau}(T) = (2 + \frac{\alpha\tau}{(1-\alpha)})S_T - (1 + \frac{\alpha\tau}{(1-\alpha)})S_T^{[2]} \qquad (4.3.11)$$

4.3.5 The Double Exponential Smoothing Algorithm

Step 1: Choose an initial estimate of β_0 and β_1 defined by equations 4.3.8 and 4.3.9 or equations 4.1.4 and 4.1.5, depending on the technique chosen.

Step 2: Use the smoothing equations 4.3.4 and 4.3.8 to simulate all the historical data available for a particular α.

Step 3: Calculate the output using equation 4.3.11.

Step 4: IF all data has been used THEN stop ELSE go to Step 2.

4.3.6 Experimental Results

An example of the performance of double exponential smoothing method is illustrated in the following graph: The overall MSE range from 0.00013 to 0.00035. The mean is around 0.00021. The range of MSE for the averages is from 0.00017 to 0.00031, the mean is about 0.00014. The range of $RMSE$ is from 0.015 to 0.011 for exchange rates and from 0.011 to 0.014 for the averages of exchange rates. The mean of the former is around 0.014 and that for the latter is 0.014. MAE ranges from 0.00562 to 0.01051 for exchange rates and from 0.02 to 0.0111 for the averages of exchange rates. The mean for the exchange rates is about 0.007 and that of the averages is about 0.008. $MAPE$ ranges from 3.24 to 3.2601 for exchange rates and from 1.85 to 3.34 for the averages of exchange rates. The mean of the $MAPE$ values is around 2.3

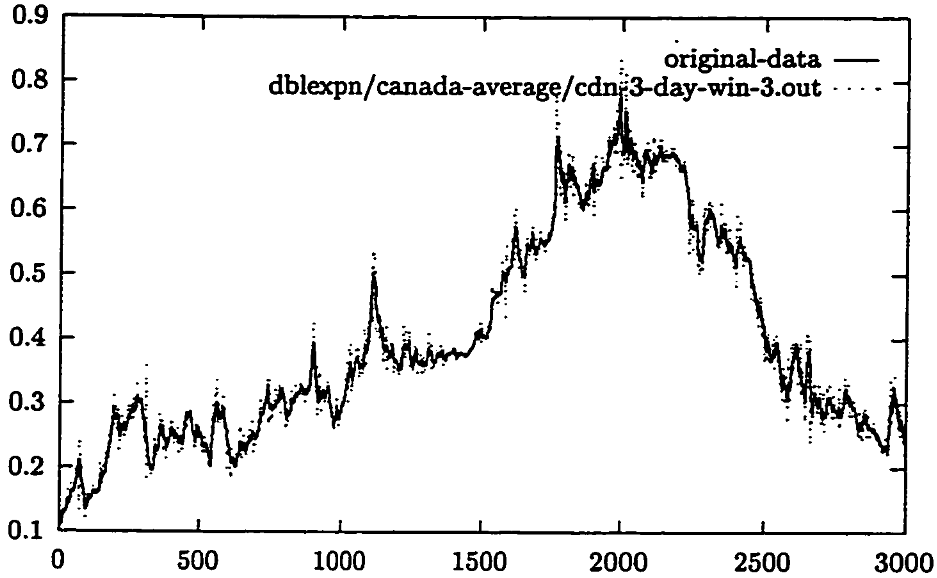


Figure 10: Predicting 3 days averages of Canadian-U.S. dollar exchange rates for window size 3 using double exponential smoothing

for the exchange rates and 2.15 for the averages of exchange rates. Finally, for the *DIR*, the value ranges from 0.4838 to 0.5075 for exchange rates and from 0.4715 to 0.5145. The overall mean for the exchange rates is about 0.49 whereas that for the averages of exchange rates is about 0.48. The results are summarized in tables.

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	1	0.00024	0.01535	0.00733	1.99216	0.4938
	3	2	0.00029	0.01697	0.00862	2.31598	0.4938
	3	3	0.00034	0.01841	0.00968	2.58769	0.4935
	4	1	0.00025	0.01578	0.00782	2.10415	0.4938
	4	2	0.00030	0.01741	0.00898	2.38901	0.4935
	4	3	0.00035	0.01878	0.00991	2.63934	0.4932
	5	1	0.00020	0.01406	0.00661	1.76573	0.4935
	5	2	0.00025	0.01578	0.00785	2.08388	0.4932
	5	3	0.00029	0.01711	0.00878	2.35021	0.4935
Britain	3	1	0.00015	0.01236	0.00743	2.31263	0.4838
	3	2	0.00019	0.01369	0.00859	2.67282	0.4838
	3	3	0.00022	0.01499	0.00966	3.00333	0.4842
	4	1	0.00016	0.01273	0.00773	2.40316	0.4838
	4	2	0.00020	0.01411	0.00892	2.76604	0.4842
	4	3	0.00024	0.01534	0.00989	3.06423	0.4845
	5	1	0.00013	0.01129	0.00654	2.03006	0.4842
	5	2	0.00016	0.01277	0.00782	2.42448	0.4845
	5	3	0.00020	0.01414	0.00896	2.76796	0.4845
Canada	3	1	0.00018	0.01349	0.00909	2.59633	0.4928
	3	2	0.00023	0.01506	0.01035	2.93700	0.4925
	3	3	0.00027	0.01651	0.01150	3.24070	0.4928
	4	1	0.00019	0.01374	0.00931	2.64004	0.4925
	4	2	0.00024	0.01534	0.01059	2.98560	0.4928
	4	3	0.00028	0.01667	0.01162	3.25367	0.4925
	5	1	0.00014	0.01171	0.00780	2.20632	0.4928
	5	2	0.00018	0.01347	0.00915	2.57265	0.4925
	5	3	0.00022	0.01499	0.01034	2.90155	0.4928
France	3	1	0.00013	0.01120	0.00671	2.13868	0.5005
	3	2	0.00015	0.01217	0.00753	2.38965	0.5008
	3	3	0.00017	0.01310	0.00830	2.63610	0.5012
	4	1	0.00013	0.01118	0.00681	2.16213	0.5008
	4	2	0.00015	0.01219	0.00762	2.41528	0.5012
	4	3	0.00017	0.01313	0.00840	2.66284	0.5008
	5	1	0.00010	0.00975	0.00564	1.78972	0.5012
	5	2	0.00012	0.01092	0.00663	2.10775	0.5008
	5	3	0.00014	0.01194	0.00762	2.41161	0.5005

Table 9: Predicting exchange rates using double exponential smoothing method

Country	Prediction Zones		Measures				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	1	0.00013	0.01123	0.00719	2.01502	0.4898
	3	2	0.00015	0.01240	0.00816	2.28606	0.4895
	3	3	0.00018	0.01351	0.00902	2.53169	0.4895
	4	1	0.00013	0.01140	0.00741	2.07949	0.4895
	4	2	0.00016	0.01251	0.00834	2.33967	0.4895
	4	3	0.00019	0.01369	0.00920	2.57564	0.4895
	5	1	0.00009	0.00969	0.00613	1.72730	0.4895
	5	2	0.00012	0.01109	0.00724	2.02966	0.4895
	5	3	0.00015	0.01238	0.00832	2.32319	0.4898
Japan	3	1	0.00013	0.01144	0.00784	1.95591	0.5042
	3	2	0.00016	0.01283	0.00894	2.22553	0.5045
	3	3	0.00020	0.01422	0.01004	2.51158	0.5048
	4	1	0.00014	0.01189	0.00815	2.03484	0.5045
	4	2	0.00018	0.01324	0.00919	2.29996	0.5048
	4	3	0.00021	0.01452	0.01017	2.55615	0.5048
	5	1	0.00010	0.01020	0.00677	1.70812	0.5048
	5	2	0.00014	0.01177	0.00797	2.00394	0.5048
	5	3	0.00017	0.01323	0.00916	2.31359	0.5048
Switzerland	3	1	0.00015	0.01220	0.00836	2.43323	0.5068
	3	2	0.00019	0.01364	0.00954	2.79113	0.5072
	3	3	0.00023	0.01507	0.01063	3.11297	0.5075
	4	1	0.00016	0.01262	0.00869	2.54258	0.5072
	4	2	0.00020	0.01400	0.00983	2.88496	0.5075
	4	3	0.00024	0.01549	0.01093	3.20486	0.5075
	5	1	0.00012	0.01077	0.00725	2.13305	0.5075
	5	2	0.00016	0.01252	0.00861	2.52892	0.5075
	5	3	0.00020	0.01422	0.00994	2.91257	0.5075

Table 10: Predicting exchange rates using double exponential smoothing method (contd.)

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	3	0.00041	0.02026	0.01065	2.92689	0.4718
	3	5	0.00026	0.01615	0.00788	2.14784	0.4882
	4	3	0.00041	0.02031	0.01089	3.00522	0.4715
	4	5	0.00031	0.01763	0.00909	2.48994	0.4882
	5	3	0.00039	0.01964	0.01053	2.93858	0.4715
	5	5	0.00027	0.01652	0.00842	2.33999	0.4882
Britain	3	3	0.00020	0.01430	0.00903	2.81095	0.5065
	3	5	0.00017	0.01314	0.00806	2.50639	0.4865
	4	3	0.00029	0.01717	0.01119	3.48272	0.5065
	4	5	0.00022	0.01470	0.00928	2.88428	0.4865
	5	3	0.00028	0.01685	0.01100	3.42627	0.5065
	5	5	0.00020	0.01399	0.00880	2.73807	0.4865
Canada	3	3	0.00025	0.01578	0.01078	3.05862	0.4832
	3	5	0.00021	0.01432	0.00971	2.76007	0.4822
	4	3	0.00036	0.01900	0.01330	3.76734	0.4832
	4	5	0.00026	0.01599	0.01105	3.12915	0.4822
	5	3	0.00034	0.01846	0.01301	3.68577	0.4832
	5	5	0.00022	0.01494	0.01038	2.93881	0.4822
France	3	3	0.00016	0.01261	0.00799	2.52130	0.4808
	3	5	0.00014	0.01163	0.00715	2.25964	0.4858
	4	3	0.00022	0.01484	0.00983	3.09663	0.4808
	4	5	0.00016	0.01268	0.00817	2.57507	0.4862
	5	3	0.00021	0.01446	0.00963	3.03998	0.4808
	5	5	0.00014	0.01194	0.00769	2.42885	0.4862

Table 11: Predicting averages of exchange rates using double exponential smoothing method

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	3	0.00017	0.01304	0.00877	2.44051	0.5075
	3	5	0.00014	0.01187	0.00778	2.16993	0.5035
	4	3	0.00025	0.01574	0.01092	3.03679	0.5075
	4	5	0.00018	0.01325	0.00906	2.52126	0.5035
	5	3	0.00024	0.01536	0.01072	2.98973	0.5075
	5	5	0.00016	0.01247	0.00857	2.39283	0.5035
Japan	3	3	0.00028	0.01678	0.01173	2.89722	0.5005
	3	5	0.00015	0.01238	0.00849	2.10563	0.4915
	4	3	0.00029	0.01696	0.01191	2.95245	0.5012
	4	5	0.00020	0.01421	0.00988	2.44901	0.4902
	5	3	0.00028	0.01674	0.01175	2.91964	0.4998
	5	5	0.00018	0.01354	0.00939	2.33738	0.4902
Switzerland	3	3	0.00022	0.01469	0.01024	2.96980	0.5145
	3	5	0.00020	0.01413	0.00977	2.83512	0.4962
	4	3	0.00032	0.01801	0.01288	3.73320	0.5145
	4	5	0.00023	0.01509	0.01065	3.09240	0.4962
	5	3	0.00031	0.01772	0.01273	3.69051	0.5145
	5	5	0.00021	0.01437	0.01017	2.95273	0.4962

Table 12: Predicting averages of exchange rates using double exponential smoothing method (contd.)

4.4 The Box-Jenkins Methodology

All the regression and exponential smoothing models used for forecasting made an assumption that the random error components and consequently the successive time series observations of the model

$$y_t = f(\beta_0, \beta_1, \cdots, \beta_p; t) + \varepsilon_t$$

are statistically independent of each other. In most real time series the successive observations are highly dependent. In these kind of observations, we cannot apply the smoothing techniques because they do not take the advantage of dependency in the most effective way. The Box-Jenkins Methodology [4] is used as an efficient tool for forecasting such data. This method involves mainly four steps:

1. Identification of a model: a tentative model is developed using the past historical data.
2. Estimation of unknown parameters: unknown parameters of the tentative model are estimated.
3. Diagnostic checking of the model: diagnostic checks are performed to test and improve (if need be) the adequacy of the model.
4. Forecasting: predictions of the future values of the time series.

Before proceeding to describe the Box-Jenkins model, some basic concepts are presented to get a clear understanding of the model.

Stationary Time Series and Nonstationary Time Series

A time series is called a *stationary* time series if its values fluctuate around a constant mean μ. And if there is no constant mean then it is called a *nonstationary* time series.

A time series may be represented by:

$$y_t = \mu + \psi_0 \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} \cdots$$

If a time series is nonstationary it is important to transform into a stationary time series for the purpose of building a model. This can be achieved sometimes by using the first differences of the original time series. i.e. using the transformation:

$$z_t = \nabla y_t = y_t - y_{t-1} \quad for \quad t = 2, \ldots n$$

However, if the resulting time series is still not stationary then the second difference is considered.

Autocorrelation and Partial Autocorrelation

Let us assume that the time series $z_a, z_{a+1}, \cdots z_n$ is a stationary time series, generated from differencing the original nonstationary time series. If the original time series was stationary itself, then $a = 1$. An important implication for a time series to be stationary is that the statistical properties of the time series are unaffected by a shift of the time origin. i.e. the relationships between n observations at origin t, say $z_t, z_{t+1}, \cdots, z_{t+n-1}$ are the same as the statistical relationships between n observations at origin $(t+j)$, say, $z_{t+j}, z_{t+j+1}, \cdots, z_{t+j+n-1}$. One of these important relationships is measured by ρ_k, which is the *autocorrelation* between any two time series observations separated by a "lag" of k time units. ρ_k is dimensionless and

$$-1 \leq \rho_k \leq 1 \quad and \quad \rho_k = \rho_{-k}$$

which means that we should consider only positive lags. Also, $\rho_k \to 1$ implies that observations separated by a lag of k units have a strong tendency to move together in a linear fashion with a positive slope and similarly for ρ_k having a value closer to -1 will indicate the same movement but with a negative slope. The estimate of ρ_k, namely, r_k is given by the sample autocorrelation at lag k, denoted by the symbol r_k, and is given by the formula:

$$r_k = \frac{\sum_{t=a}^{n-k}(z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=a}^{n}(z_t - \bar{z})^2} \tag{4.4.1}$$

where $\bar{z}$ is the average of the observations: $z_a, z_{a+1}, \cdots, z_n$ and is given by:

$$\bar{z} = \frac{\sum_{t=a}^{n} z_t}{n - a + 1} \tag{4.4.2}$$

The theoretical autocorrelation function is defined to be a listing, or graph, of ρ_k for lags of $k = 1, 2, \cdots$ and sample autocorrelation function is the listing of r_k for lags of $k = 1, 2, \cdots$. The theoretical autocorrelation function of a stationary time series tends to either *die down* with increasing lag k or *cut off* after a particular lag $k = q$, i.e., $\rho_k = 0$ for $k > q$. One of the ways to conclude if $\rho_k = 0$ or not is to calculate the "t-like statistic":

$$t_{rk} = \frac{r_k}{s_{rk}} \quad where$$

$$s_{rk} = \frac{1}{(n-a+1)^{\frac{1}{2}}}(1 + 2\sum_{j=1}^{q} r_j^2)^{\frac{1}{2}} \quad for \quad k > q$$

As a rule of thumb it can be concluded that:

$$\rho_k = 0 \quad if|t_{rk}| = |\frac{r_k}{s_{rk}}| \leq 2$$

It can be shown that if a time series is non-stationary then the sample autocorrelation function will neither cut off nor die down quickly, but rather will die down extremely slowly.

If the first differences are found to represent a non-stationary time series, then the second difference should be considered to produce a stationary time series.It is rarely necessary to consider more than the second differences to achieve stationarity for a time series not possessing seasonal variations.

The *partial autocorrelation* between any two time series observations separated by a lag of k time units defined by ρ_{kk} helps us in determining the particular time series model that can be assumed to have generated the observations $z_a, z_{a+1}, \cdots z_n$.

The estimate of ρ_{kk} is given by:

$$r_{kk} = \begin{cases} r_1 & \text{if k=1} \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & \text{if k= 2,3,...} \end{cases}$$

$$where \quad r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j}, \quad for \quad j = 1, 2, \cdots, k-1$$

The most common Box-Jenkins models for stationary time series are:

1. Moving-Average models of order (q):

$$z_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

2. Autoregressive models of order (p):

$$z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + \varepsilon_t$$

3. Mixed Moving-Average and Autoregressive models of order (p,q):

$$z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \varepsilon_t + \varepsilon_t$$

Of these, the third model has been chosen for our experiment. A detailed account of the first two models has been given in appendix A.3.

4.4.1 Mixed Auto-Regressive-Moving-Average Models

The model:

$$z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

is called a mixed *autoregressive-moving-average model of order (p,q).*

1. Stationarity condition: same as the autoregressive process of order p.

2. Invertibility condition: same as the moving-average model of order q.

Both the theoretical autocorrelation function and the theoretical partial autocorrelation function die down.

4.4.2 The Modelling

For the experiments on exchange rates, the following model has been considered:

$$z_t = \delta + \phi_1 z_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \tag{4.4.3}$$

The model described by 4.4.3 is a *mixed autoregressive and moving average model of order (1,1)* This model is stationary if

$$|\phi_1| < 1$$

and invertible if

$$|\theta_1| < 1$$

The mean of the model is given by

$$\mu = \frac{\delta}{1 - \phi_1}$$

This means that

$$\delta = \mu(1 - \phi_1) \qquad (4.4.4)$$

4.4.3 Estimation of Model Parameters

In order to develop the model described by equation 4.4.3 it is necessary to estimate the parameters ϕ , θ and δ.

Estimation of ϕ

In case of first order autoregressive models, the estimation of the parameter ϕ_1 is given by

$$\hat{\phi}_1 = r_1\left(\frac{1 - r_2}{1 - r_1^2}\right) \qquad (4.4.5)$$

The equation 4.4.5 follows from solving Yule-Walker equations described in appendix A.3 and assuming that r_k is an estimate of ρ_k.

Estimation of θ

Referring to the appendix A.3, we see that equation

$$\rho_1 = \frac{-\theta_1}{1+\theta_1^2} \tag{4.4.6}$$

can be used to calculate θ_1. Solution of which gives:

$$\theta_1 = -\frac{1}{2\rho_1} \pm [\frac{1}{(2\rho_1)^2} - 1]^{\frac{1}{2}} \tag{4.4.7}$$

Replacing the theoretical autocorrelation ρ_1 with sample autocorrelation r_1, a preliminary estimate of ρ_1, we get,

$$\hat{\theta}_1 = -\frac{1}{2r_1} \pm [\frac{1}{(2r_1)^2} - 1]^{\frac{1}{2}} \tag{4.4.8}$$

There are two values of $\hat{\theta}$ in the above equation 4.4.8. The value for which $\hat{\theta}$ satisfies the invertibility condition i.e., $|\hat{\theta}| < 1$ is used as the estimate of θ. (Please refer to appendix A.3 for further details) .

Estimation of δ

From equation 4.4.4 , a preliminary estimate of δ is given by:

$$\hat{\delta} = \bar{z}(1 - \hat{\phi}_1) \tag{4.4.9}$$

where $\bar{z}$ is the estimate of mean μ and is given by equation 4.4.2.

4.4.4 Forecasting the Future Values

After the estimation of the parameters ϕ , θ , and δ for the first-order mixed moving-average and autoregression model 4.4.3 considered for experiments, we are in a position to use the model for forecasting the future values of the time series. Now, we have,

$$z_t = \nabla y_t = y_t - y_{t-1}$$

Hence, we can write:

$$\begin{aligned} z_t &= \delta + \phi_1 z_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ \nabla y_t &= \delta + \phi_1 \nabla y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ y_t - y_{t-1} &= \delta + \phi_1 (y_{t-1} - y_{t-2}) + \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ y_t &= \delta + (\phi_1 + 1) y_{t-1} - \phi_1 y_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \end{aligned} \quad (4.4.10)$$

The model described by equation 4.4.10 has been used for carrying out experiments in forecasting future values of the exchange rates.

4.4.5 The Box-Jenkins Algorithm

As mentioned earlier, the first-order mixed autoregressive and moving-average model of order (1,1) has been considered for forecasting future values of exchange data. The algorithm is being stated below:

Step 1: Using a set of input time series data, compute the first differences to make the time series stationary.

Step 2: Calculate the correlation coefficients using equation 4.4.1.

Step 3: Calculate the estimates of ϕ , θ , and δ , using equations 4.4.5, 4.4.8 , and 4.4.9 respectively. Make sure that the stationarity and invertibility conditions are satisfied.

Step 4: Calculate the error terms ε by subtracting the forecasted value of the time series from the actual values.

Step 5: Substitute all the estimates in the modelling equation and perform forecasting for the remaining data.

Step 6: IF stopping condition is satisfied THEN stop ELSE go to Step 2.

Please note that a lag of 1 and first-order differences have been considered for the experiments.

4.4.6 Experimental Results

An example of the performance of Box-Jenkins method is illustrated in the following graph: The overall MSE range from 0.0023 to 0.003. The mean is around 0.002. The

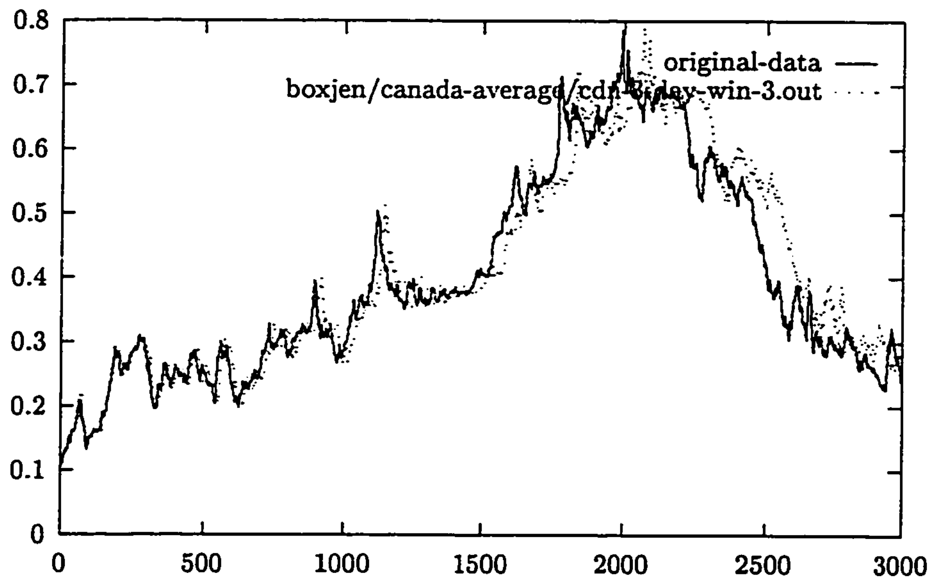


Figure 11: Predicting 3 days averages of Canadian-U.S. dollar exchange rates for window size 3 using Box-Jenkins method

range of MSE for the averages is from 0.0023 to 0.00299, the mean is about 0.0026. The range of $RMSE$ is from 0.0487 to 0.055 for exchange rates and from 0.048 to

0.054 for the averages of exchange rates. The mean of the former is around 0.05 and that for the latter is 0.048. MAE ranges from 0.0047 to 0.0058 for exchange rates and from 0.0029 to 0.0047 for the averages of exchange rates. The mean is about 0.005 for exchange rates and about 0.003 for the averages of exchange rates. Thus we see that the MAE values of the average of exchange rates is much lower than the exact exchange rates. $MAPE$ ranges from 1.14 to 2.00 for exchange rates and from 0.68 to 2.0 for the averages of exchange rates. The mean of the $MAPE$ values is around 1.55 for the exchange rates and 1.2 for the averages of exchange rates. But the overall $MAPE$ for averages is less than that of the exact exchange rates. Finally, for the DIR, the value ranges from 0.4753 to 0.50 for exchange rates and from 0.4766 to 0.50. The overall mean for the exchange rates is about 0.49 whereas that for the averages of exchange rates is about 0.48. The results are presented in the following tables.

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	1	0.00238	0.04877	0.00472	1.14679	0.4805
	3	2	0.00239	0.04890	0.00472	1.14802	0.4805
	3	3	0.00239	0.04887	0.00472	1.14843	0.4803
	4	1	0.00239	0.04890	0.00472	1.14802	0.4805
	4	2	0.00239	0.04887	0.00472	1.14843	0.4803
	4	3	0.00237	0.04869	0.00472	1.14600	0.4803
	5	1	0.00239	0.04887	0.00472	1.14843	0.4803
	5	2	0.00237	0.04869	0.00472	1.14600	0.4803
	5	3	0.00238	0.04878	0.00472	1.14650	0.4803
Britain	3	1	0.00263	0.05125	0.00499	1.50737	0.4753
	3	2	0.00263	0.05133	0.00499	1.51093	0.4750
	3	3	0.00262	0.05116	0.00499	1.50480	0.4753
	4	1	0.00263	0.05133	0.00499	1.51093	0.4750
	4	2	0.00262	0.05116	0.00499	1.50480	0.4753
	4	3	0.00262	0.05118	0.00499	1.50523	0.4753
	5	1	0.00262	0.05116	0.00499	1.50480	0.4753
	5	2	0.00262	0.05118	0.00499	1.50523	0.4753
	5	3	0.00262	0.05123	0.00499	1.50721	0.4753
Canada	3	1	0.00253	0.05032	0.00579	1.66220	0.4779
	3	2	0.00257	0.05074	0.00581	1.67176	0.4779
	3	3	0.00256	0.05058	0.00580	1.66729	0.4782
	4	1	0.00257	0.05074	0.00581	1.67176	0.4779
	4	2	0.00256	0.05058	0.00580	1.66729	0.4782
	4	3	0.00255	0.05052	0.00581	1.66828	0.4782
	5	1	0.00256	0.05058	0.00580	1.66729	0.4782
	5	2	0.00255	0.05052	0.00581	1.66828	0.4782
	5	3	0.00255	0.05047	0.00581	1.66593	0.4784
France	3	1	0.00287	0.05361	0.00449	1.55045	0.4895
	3	2	0.00288	0.05369	0.00449	1.55806	0.4895
	3	3	0.00288	0.05362	0.00449	1.55269	0.4897
	4	1	0.00288	0.05369	0.00449	1.55806	0.4895
	4	2	0.00288	0.05362	0.00449	1.55269	0.4897
	4	3	0.00288	0.05365	0.00449	1.55420	0.4895
	5	1	0.00288	0.05362	0.00449	1.55269	0.4897
	5	2	0.00288	0.05365	0.00449	1.55420	0.4895
	5	3	0.00288	0.05367	0.00449	1.55696	0.4892

Table 13: Predicting exchange rates using Box-Jenkins method

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	1	0.00301	0.05488	0.00476	1.80827	0.4782
	3	2	0.00302	0.05494	0.00476	1.82242	0.4782
	3	3	0.00301	0.05485	0.00476	1.80518	0.4782
	4	1	0.00302	0.05494	0.00476	1.82242	0.4782
	4	2	0.00301	0.05485	0.00476	1.80518	0.4782
	4	3	0.00301	0.05491	0.00476	1.81376	0.4782
	5	1	0.00301	0.05485	0.00476	1.80518	0.4782
	5	2	0.00301	0.05491	0.00476	1.81376	0.4782
	5	3	0.00301	0.05491	0.00474	1.81267	0.4784
Japan	3	1	0.00308	0.05546	0.00490	1.85489	0.5005
	3	2	0.00308	0.05550	0.00490	1.86863	0.5005
	3	3	0.00307	0.05544	0.00489	1.85052	0.5005
	4	1	0.00308	0.05550	0.00490	1.86863	0.5005
	4	2	0.00307	0.05544	0.00489	1.85052	0.5005
	4	3	0.00307	0.05542	0.00490	1.84250	0.5005
	5	1	0.00307	0.05544	0.00489	1.85052	0.5005
	5	2	0.00307	0.05542	0.00490	1.84250	0.5005
	5	3	0.00307	0.05544	0.00490	1.84835	0.5005
Switzerland	3	1	0.00298	0.05461	0.00538	2.00007	0.4950
	3	2	0.00299	0.05471	0.00538	2.01768	0.4953
	3	3	0.00298	0.05459	0.00538	1.99976	0.4955
	4	1	0.00299	0.05471	0.00538	2.01768	0.4953
	4	2	0.00298	0.05459	0.00538	1.99976	0.4955
	4	3	0.00298	0.05461	0.00538	2.00271	0.4955
	5	1	0.00298	0.05459	0.00538	1.99976	0.4955
	5	2	0.00298	0.05461	0.00538	2.00271	0.4955
	5	3	0.00299	0.05471	0.00538	2.01855	0.4955

Table 14: Predicting exchange rates using Box-Jenkins method (contd.)

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	3	0.00235	0.04843	0.00294	0.68019	0.4766
	3	5	0.00235	0.04851	0.00339	0.80813	0.4976
	4	3	0.00235	0.04843	0.00294	0.68019	0.4766
	4	5	0.00235	0.04851	0.00339	0.80813	0.4976
	5	3	0.00235	0.04843	0.00294	0.68019	0.4766
	5	5	0.00235	0.04851	0.00339	0.80813	0.4976
Britain	3	3	0.00261	0.05107	0.00320	0.94555	0.5068
	3	5	0.00262	0.05123	0.00458	1.38292	0.4829
	4	3	0.00261	0.05107	0.00320	0.94555	0.5068
	4	5	0.00262	0.05123	0.00458	1.38292	0.4829
	5	3	0.00261	0.05107	0.00320	0.94555	0.5068
	5	5	0.00262	0.05123	0.00458	1.38292	0.4829
Canada	3	3	0.00253	0.05033	0.00345	0.95906	0.4797
	3	5	0.00253	0.05035	0.00422	1.18611	0.4861
	4	3	0.00253	0.05033	0.00345	0.95906	0.4797
	4	5	0.00253	0.05035	0.00422	1.18611	0.4861
	5	3	0.00253	0.05033	0.00345	0.95906	0.4797
	5	5	0.00253	0.05035	0.00422	1.18611	0.4861
France	3	3	0.00286	0.05351	0.00274	0.98637	0.4900
	3	5	0.00286	0.05352	0.00323	1.13733	0.4929
	4	3	0.00286	0.05351	0.00274	0.98637	0.4900
	4	5	0.00286	0.05352	0.00323	1.13733	0.4929
	5	3	0.00286	0.05351	0.00274	0.98637	0.4900
	5	5	0.00286	0.05352	0.00323	1.13733	0.4929

Table 15: Predicting averages of exchange rates using Box-Jenkins method

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	3	0.00299	0.05472	0.00289	1.20538	0.5061
	3	5	0.00300	0.05476	0.00356	1.42232	0.4905
	4	3	0.00299	0.05472	0.00289	1.20538	0.5061
	4	5	0.00300	0.05476	0.00356	1.42232	0.4905
	5	3	0.00299	0.05472	0.00289	1.20538	0.5061
	5	5	0.00300	0.05476	0.00356	1.42232	0.4905
Japan	3	3	0.00306	0.05529	0.00318	1.36175	0.4955
	3	5	0.00308	0.05548	0.00472	1.80038	0.4926
	4	3	0.00306	0.05529	0.00318	1.36175	0.4955
	4	5	0.00308	0.05548	0.00472	1.80038	0.4926
	5	3	0.00306	0.05529	0.00318	1.36175	0.4955
	5	5	0.00308	0.05548	0.00472	1.80038	0.4926
Switzerland	3	3	0.00296	0.05440	0.00336	1.32163	0.5208
	3	5	0.00298	0.05455	0.00471	1.76606	0.4889
	4	3	0.00296	0.05440	0.00336	1.32163	0.5208
	4	5	0.00298	0.05455	0.00471	1.76606	0.4889
	5	3	0.00296	0.05440	0.00336	1.32163	0.5208
	5	5	0.00298	0.05455	0.00471	1.76606	0.4889

Table 16: Predicting averages of exchange rates using Box-Jenkins method (contd.)

4.5 Forecasting Seasonal Time Series

We consider time series that are seasonal in nature, i.e., time series which include seasonal variations. There are two types of seasonal variations:

1. Additive seasonal variations : the magnitude of the seasonal swing of the time series is independent of the average level as determined by the trend .
2. Multiplicative seasonal variations : the magnitude of the seasonal swing is proportional to the average level as determined by the trend. Thus if the average level is increased(decreased) so is the swing increased(decreased).

Very few real-world time series possesses seasonal variation that is precisely additive or multiplicative in nature.But it is necessary to classify a given time-series model for analysis. The time series models to be described in this context are special cases of the model:

$$y_t = f(TR_t, SN_t) + \varepsilon_t$$

where y_t : observed value of the time series in the time period t. TR_t : trend factor of the time series in time period t. SN_t : seasonal factor of the time series in time period t. f: a function relating the observed value of the time series to the trend and seasonal factors.

ε_t: irregular factor of the time series in time period t. Let tr_t and sn_t be the estimates of TR_t and SN_t respectively. Then the estimate of the value of the time series in period t is :

$$\hat{y}_t = f(tr_t, sn_t)$$

When analyzing a time series having additive seasonal variations, it is generally assumed that:

$$f(TR_t, SN_t) = TR_t + SN_t$$

This means that the magnitude of the seasonal swing of the time series is independent of the average level as determined by the trend.
The estimate of the value of y_t is :

$$\hat{y}_t = tr_t + sn_t$$

When analyzing a time series having multiplicative seasonal variations, it is generally assumed that :

$$f(TR_t, SN_t) = TR_t \times SN_t$$

This implies that SN_t is proportional to TR_t. The estimate is given by:

$$\hat{y}_t = tr_t \times sn_t$$

TR_t is generally assumed to be given by any one of the following equations:

1. $TR_t = \beta_0$ i.e. no trend.

2. $TR_t = \beta_0 + \beta_1 t$ i.e. linear trend.

3. $TR_t = \beta_0 + \beta_1 t + \beta_2 t^2$ i.e. quadratic trend.

Most of the time series exhibit linear trend. Therefore we shall assume: $TR_t = \beta_0 + \beta_1 t$

4.5.1 The Seasonal and Cyclical Factors

1. The Seasonal Factor : a correlation factor that accounts for or adjusts for the seasonality in the time series.

2. The Cyclical Factor : the seasonal variations in a time series, hence seasonal factor SN_t, reflects cyclical patterns in a time series that are completed within one calendar year. If a time series is also influenced by a cyclical pattern that has a duration of more than one year, then a cyclical pattern is said to exist and is denoted by CL_t.

For a real time series the factors TR_t, SN_t, CL_t are not known. They have to be estimated by the forecaster.

4.5.2 Forecasting using the Multiplicative Decomposition Metho

We discuss analyzing and forecasting a time series that has multiplicative seasonal variation.

$$y_t = TR_t \times SN_t + \varepsilon_t$$

The above model can be written as :

$$y_t = TR_t \times SN_t \times CL_t \times IR_t$$

where IR_t is the irregular factor of the time series in period t.

The forecast is done using:

$$\hat{y}_t = tn_t \times sn_t$$

The estimation of the seasonal factors can be made by using the "moving average technique" .This will remove the seasonal variation of a time series since each moving average is computed using exactly one observation from each season.Secondly, it has removed the effects of irregular factors hopefully.

Hence the centered moving averages represent a combination of trend and cyclical factors. We consider the centered moving average corresponding to period t as $tr_t \times cl_t$ the estimate of $TR_t \times CL_t$.

Since

$$y_t = TR_t \times SN_t \times CL_t \times IR_t \quad and$$
$$sn_t \times ir_t = \frac{tr_t \times sn_t \times cl_t \times ir_t}{tr_t \times cl_t}$$

We can obtain an estimate of $SN_t \times IR_t$ using the formula :

$$sn_t \times ir_t = \frac{y_t}{tr_t \times cl_t}$$

4.5.3 Winters' Method

This is an exponential smoothing procedure with a linear trend and multiplicative seasonal variations. It was developed by P. R. Winters [27] and is known as Winters' Method. The model can be described as below :

$$y_t = (\beta_0 + \beta_1 t) \times SN_t + \varepsilon_t$$

4.5.4 Updating the Estimates

Assume that at the end of period $(T-1)$ we have the estimates of the model parameters β_0 β_1 *and* SN_t . We have a new observation y_T in period T and want to estimate the parameters.

Let the updated estimates be $b_0(T)$, $b_1(T)$ and $sn_T(T)$ where $b_0(T)$ is the estimate of the intercept of the trend line, where we define the intercept to be the intercept at the original origin of time. The estimate of the intercept using the current time period as the origin shall be denoted by $a_0(T)$.

Assume that we have several seasonal factors. Let us denote the number of seasons that must occur before the first season repeats itself as L (for monthly time series L is 12 and for quarterly time series it is 4.) such that:

$$\sum_{t=1}^{L} SN_t = L$$

$$\sum_{t=1}^{L} sn_t(0) = L$$

Now, given the new observation y_T, we wish to obtain the estimates $a_0(T)$, $b_1(T)$, and $sn_T(T)$ by updating $a_0(T-1)$ and $b_1(T-1)$, the estimates obtained in the previous period, and by updating $sn_T(T-L)$, the last estimate of the current seasonal factor obtained L periods ago.

In order to update the estimate of β_0 , often called the *permanent component* we use the following equation:

$$a_0(T) = \alpha \frac{y_T}{sn_T(T-L)} + (1-\alpha)[a_0(T-1) + b_1(T-1)]$$

where $a_0(T)$ is the new estimate of β_0 and α is the smoothing constant $0 < \alpha < 1$, $a_0(T-1) + b_1(T-1)$ is simply the estimate of the average level of the time series at time T. y_T is divided by $sn_T(T-L)$ to get rid of the seasonal factor from influencing the estimate of the permanent component. This is called *deseasonalization* of y_T. We obtain the updated estimate of β_1, often called the *trend* by using the equation

$$b_1(T) = \beta[a_0(T) - a_0(T-1)] + (1-\beta)b_1(T-1)$$

where β is the smoothing constant $0 < \beta < 1$. and $a_0(T) - a_0(T-1)$ is the difference between the estimate of the current component and the one made in the last period.The updated estimate of the seasonal factor is obtained by using the following equation:

$$sn_T = \gamma\frac{y_T}{a_0(T)} + (1-\gamma)sn_T(T-L)$$

where $0 < \gamma < 1$

The estimate made L periods ago is used because that was the last time this particular season was observed.The current observed seasonal variation is obtained by dividing the observed y_T by the current estimate of the average level of the time series, $a_0(T)$

$$\hat{y}_{T+\tau}(T) = [a_0(T) + b_1(T)\tau]sn_{T+\tau}(T+\tau-L)$$

4.5.5 The Initial Estimates

To determine $a_0(0)$, $b_1(0)$ and $sn_T(0)$ for $t = 1, 2, \ldots L$, where L is the number of different seasons.

One way to find the initial estimates is to use the multiplicative decomposition method, then:

$a_0(0) = b_0$, $b_1(0) = b_1$, $sn_T(0) = sn_T$ for $t = 1, 2, \ldots L$. But the more frequently used method is the one given by Johnson and Montgomery [15] and is similar to the

Winters' method.

Suppose that the historical data for the last m years is available. Let $\bar{y}_i$ be the average of the observation in the i^{th} year, $i = 1, 2, \ldots m$.

Then the initial estimate of β_1, the *trend component* can be determined by:

$$b_1(0) = \frac{\bar{y}_m - \bar{y}_1}{(m-1)L}$$

where $\bar{y}_m$ is the average level of the time series in the middle of the year m

and $\bar{y}_1$ is the average level of the time series in the middle of year 1

The number of seasons between the middle of the year 1 and middle of year m is $(m-1)L$.

Thus $b_1(0)$is the change in average level per season from the middle of year 1 to the middle of year m. The initial estimate of the permanent component β_0 is given by:

$$a_0(0) = \bar{y}_1 - \frac{L}{2} b_1(0)$$

Obtaining the initial estimates of the L seasonal factors is done as follows: The expression

$$S_t = \frac{y_t}{\bar{y}_i - [(L+1)/2 - j] b_1(0)}$$

must be computed for each season t occurring in years 1 through m. Here $\bar{y}_i$ is the average of the observations for the year in which the season t occurs.

so, if $1 \leq t \leq L$, then $i = 1$; if $L + 1 \leq t \leq 2L$, then $i = 2$ etc.

Thus $\bar{y}_i$ measures the average level of the time series in the middle of the year in which t occurs. The letter j denotes the position of the season t within the year.

Thus the denominator of the above equation determines the average level of the time series in season t. If the season t occurs before the middle of the year, we subtract the appropriate trend from the average level at midyear in order to obtain the average level in season t. S_t is the ratio of the observation in season t to the average level of the time series in season t.It is the factor by which we must multiply the average level in order to obtain the observation, and hence S_t represents factors not accounted for in the average level of the time series.

Since the average level is determined by the permanent component and the trend, S_t represents the seasonal factors and the error term ε_t.

$$\bar{sn}_t = \frac{1}{m} \sum_{k=0}^{m-1} S_{t+kL}$$

$t = 1, 2, \ldots L$. This is the average seasonal index for each different season. i.e. for monthly data there are 12 seasonal factors , 1 for each month.

Finally the seasonal factors are normalized so that they add up to L. Thus:

$$sn_t(0) = \bar{sn}_t[\frac{L}{\sum_{t=1}^{L} \bar{sn}_t}]$$

$t = 1, 2, \ldots L$. Thus we have obtained the initial estimates of $a_0(0)$, $b_1(0)$ and $sn_t(0)$ $t = 1, 2, \ldots L$. Winters's method thus is an intuitive modification of either single or double exponential smoothing.

Chapter 5

Predicting Time Series using Neural Networks

Artificial neural net methods or "neural nets" are dense interconnections of computational elements called neurons or nodes. In this respect, artificial neural net structure is based on our present understanding of biological nervous systems. Neural nets are widely used in areas of speech and image recognition where many hypotheses are pursued in parallel, high computation rates. Neural net models explore many competing hypotheses simultaneously using massively parallel nets composed of many computational elements connected by links with variable weights.

The nodes used in a neural net are nonlinear and typically analog, and may be slower compared to digital circuitry. The simplest node sums N weighted inputs and passes the result through a non-linearity (we use the sigmoid function). The node is characterized by an internal threshold or offset θ and by the type of non-linearity. Note: the non-linearities may be hard limiters, threshold logic elements and sigmoidal.

Neural nodes are specified by the net topology, node characteristics, and training or learning rules. These rules specify an initial set of weights and indicate how weights should be adapted during use to improve performance. Most neural net algorithms also adapt connection weights in time to improve performance based on current results. Adaptation or learning is a major focus of neural net research. The ability to adapt and continue learning is essential in forecasting time series because the the net is trained with a limited number of data whereas in reality the time series behaves quite erratically. So, it is a challenge to choose a representative training set and adapt the weights of the interconnections to get an accurate prediction of the future values.

5.1 Training of artificial neural net

The training of a neural net depends on the structure of the network and the training algorithm. But each of these algorithms provide very similar training procedures. The differences between them lies on how the input patterns are presented to the net and how the weights adjust with each step of the presentation to give the closest output.

The training of a neural net takes place in different stages. First, a set of N inputs are fed in parallel to the first stage via N input connections. Each connection carries an analog value which may take on two levels or vary over a large range for a continuous valued inputs. The first stage computes matching scores and outputs these scores in parallel to the next stage over M analog output lines. Here the maximum of these values is selected and enhanced. The second stage has one output for each of the M classes. After classification is complete, only that output corresponding to the most likely class will be on strongly or "high"; other outputs will be "low". Note that in this design, outputs exist for every class and that this multiplicity of outputs must be preserved in further processing stages as long as the classes are considered distinct. If the correct class is provided, then this information and the classifier outputs can be fed back to the first stage of the classifier to adapt weights using a learning algorithm. Adaptation will make a correct response more likely for succeeding input patterns that are similar to the current pattern. Classifiers can perform three different tasks. First, they can identify which class best represents an input pattern, where it is assumed that inputs have been corrupted by noise or some other process. This is a classical decision theory problem. Second, the classifiers can be used as a content-addressable or associative memory, where the class exemplar is desired and the input pattern is used to determine which exemplar to produce. Finally, the third task is to vector quantize or cluster the N inputs into M clusters. These number of clusters can be pre-specified or may be allowed to grow up to a limit determined by the number of nodes available at the first stage.

As mentioned before, the input pattern for a neural net may be either binary or continuous-valued. The training for each set of inputs could either be *Supervised* or *Unsupervised*. Nets trained with supervision (e.g. Hopfield net and perceptrons) are used as associative memories or as classifiers. These nets are provided with side information or labels that specify the correct class for new input patterns during training. Nets trained without supervision such as the Kohonen's feature-map forming nets are not provided with any information about the correct class during training.

5.1.1 Single and Multi-layer Perceptrons

A single layer perceptron comprises of just one node. This net can be used for both binary and continuous valued inputs. The single node computes a weighted sum of the input elements, subtracts a threshold (θ) and passes the result through a hard limiting nonlinearity such that the output y is either $+1$ or -1.

Multi-layer perceptrons are feed-forward nets with one or more layers of nodes between the input and output nodes. These additional layers contain hidden units or nodes that are not directly connected to both the input and output nodes. Multiple-layer perceptrons overcome many of the limitations of single layer perceptrons, but were not popular in the past due to lack of effective training algorithms. This has changed after the development of new training algorithms. As noted above, a single-layer perceptron forms half-plane decision regions. A two-layer perceptron forms any, possibly unbounded, convex region in the space spanned by the inputs. Such regions include convex polygons called convex hulls, and unbounded convex regions. The convex regions are formed from intersections of the half-plane regions formed by each node in the first layer of the multiple-layer perceptron. Each node in the first layer behaves like a single-layer perceptron and has a "high" output only for points on one side of the hyper-plane formed by its weights and offset. If weights to an output node from N_1 first-layer nodes are all 1.0 and the threshold in the output node is $(N_1 - \varepsilon)$ where $0 < \varepsilon < 1$ then the output node will be "high" only if the outputs of all first-layer nodes are "high". This corresponds to performing a logical AND operation in the output node and results in a final decision region that is the intersection of all the half-plane regions formed in the first layer. Intersections of such half planes form convex regions as described above. These convex regions have at most as many sides as there are nodes in the first layer. This analysis provides some insight into the problem of selecting the number of nodes to use in a two-layer perceptron. The number of nodes must be large enough to form a decision region that is as complex as required by the given problem. It must not, however, be so large that the many weights required can not be reliably estimated from the available training data.

All the above discussion is based on multi-layer perceptrons with one output when hard limiting nonlinearities are used.

5.2 Design of Experiments

All the learning methods discussed in the subsequent sections have been implemented on the computer using the C programming language. The exchange rates of Australia, Britain, Canada, France, Germany, Japan and Switzerland with the U.S. dollar were used as input time series data. Experiments were conducted using 3000 data from each country and predictions were made in the following sequence:

1. Predicting the exchange rates of each country using window sizes of 3, 4, 5 and prediction horizon of 1 day, 2 days and 3 days.

2. Predicting the averages of the exchange rates of each country using window size of 3, 4 and 5 and prediction horizon of 3 and 5 days.

The results of predictions of each of the learning methods are presented in tables.

5.3 The Method of Backpropagation

The backpropagation training algorithm is an iterative gradient algorithm designed to minimize the mean square error between the actual output of a multi-layer feed-forward perceptron and the desired output. It requires continuous differentiable non-linearities. The following assumes a sigmoid logistic non-linearity function $f(\alpha)$ given by:

$$f(\alpha) = \frac{1.0}{1.0 + e^{-(\alpha-\theta)}}$$

The net is trained by initially selecting small random weights and internal thresholds and then presenting all the training data repeatedly. Weights are adjusted after every trial using side information specifying the correct class until weights converge and the cost function is reduced to an accepted value. The iterative steps of the algorithm as presented below propagates error terms required to adapt weights back from nodes in the output layer to nodes in the lower layers.

5.3.1 The Backpropagation Algorithm

The backpropagation algorithm can be stated as follows:

Step 1: (Initialize Weights and Offsets) Set all the weights and node offsets to small random values.

Step 2: (Present Inputs and Desired Outputs) Present a continuous valued input vector $\mathbf{x_0},\ \mathbf{x_1},\ \cdots\ \mathbf{x_{N-1}}$ and specify the desired outputs $\mathbf{d_0},\ \mathbf{d_1},\ \cdots\ \mathbf{d_{M-1}}$. Samples from the training set are presented cyclically until weights stabilize.

Step 3: (Calculate Actual Outputs)

The actual outputs are calculated based on the sigmoid function and relations described by the following equations :

$$y_l = f(\sum_{k=0}^{n-1} w_{kl} x_k - \theta_l)$$

Step 4: (Adapt Weights)

A recursive algorithm is used starting at the output node working back to the first hidden layer. The weights are adjusted by using the formula:

$w_{ij}(t+1) = w_{ij}(t) + \eta\delta x_j$

where w_{ij} is the weight from the hidden node i or from an input to a node j at time t, x_j' is either the output of node i or is an input. η is the gain term and δ is an error term for node j. If j is an internal hidden node, then

$\delta_j = y_j(1.0 - y_j)(d_j - y_j)$ where d_j is the desired output of the node j and y_j is the actual output. If node j is an internal hidden node, then $\delta_j = x_j'(1.0 - x_j') \sum_k \delta_k w_{jk}$ where k is over all nodes in the layers above node j. Internal node thresholds are adapted in a similar manner by assuming they are connection weights on links from auxiliary constant-valued inputs. Convergence is sometimes faster if a momentum term is added and weight changes are smoothed by:

$w_{ij}(t+1) = w_{ij}(t) + \eta\delta_j x_j' + \alpha(w_{ij}(t) - w_{ij}(t-1))$

where $0 < \alpha < 1$.

Step 5: (Repeat by going to Step 2)

5.3.2 Experimental Results

Out of the 3000 data, about 1000 were used for training the net and the rest 2000 were used for testing. A learning rate α of 0.8 was used.

1. For window of size 3, the number of input neurons chosen was 3, the number of hidden neurons was 6 and 1 output neuron.

2. For window of size 4, the number of input neurons chosen was 4, the number of hidden neurons was 7 and 1 output neuron.

3. For window of size 5, the number of input neurons chosen was 5, the number of hidden neurons was 8 and 1 output neuron.

All the weights were initialized randomly. The following graph depicts the performance of the backpropagation in predicting the 3-day-average of the Canadian-U.S. exchange rates. The overall MSE range from 0.0003 to 0.001. The mean is around

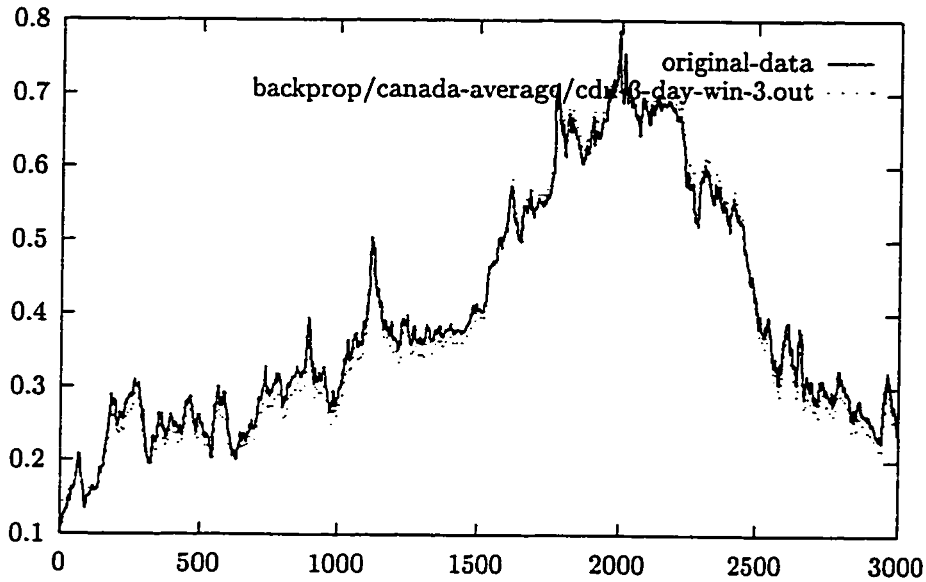


Figure 12: Predicting 3 days average of Can-U.S. exchange rates using the backpropagation method

0.0006. The range of MSE for the averages is from 0.00017 to 0.00055, the mean is about 0.00017. The range of $RMSE$ is from 0.01 to 0.04 for exchange rates and from 0.009 to 0.02 for the averages of exchange rates. The mean of the former is around 0.02 and that for the latter is 0.01. MAE ranges from 0.01 to 0.03 for exchange rates and from 0.006 to 0.01 for the averages of exchange rates. The mean is about 0.02

for exchange rates and about 0.01 for the averages of exchange rates. $MAPE$ ranges from 2.6 to 10.5 for exchange rates and from 1.75 to 6.51 for the averages of exchange rates. The mean of the $MAPE$ values is around 4.1 for the exchange rates and 3.5 for the averages of exchange rates. Finally, for the DIR, the value ranges from 0.49 to 0.518 for exchange rates and from 0.48 to 0.52. The overall mean for the exchange rates is about 0.49 whereas that for the averages of exchange rates is about 0.50. The results are tabulated below:

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	1	0.00054	0.02319	0.01448	4.85627	0.4947
	3	2	0.00095	0.03083	0.02107	6.23941	0.4950
	3	3	0.00111	0.03328	0.02415	7.57341	0.4953
	4	1	0.00050	0.02225	0.01410	5.32932	0.4950
	4	2	0.00104	0.03223	0.02207	6.36570	0.4953
	4	3	0.00145	0.03804	0.02692	7.67235	0.4957
	5	1	0.00065	0.02551	0.01647	5.01517	0.4953
	5	2	0.00110	0.03316	0.02228	6.15904	0.4957
	5	3	0.00135	0.03672	0.02593	7.40298	0.4953
Britain	3	1	0.00032	0.01798	0.01020	4.04879	0.5027
	3	2	0.00069	0.02631	0.01891	7.91936	0.5027
	3	3	0.00135	0.03679	0.02931	11.67030	0.5027
	4	1	0.00037	0.01914	0.01139	4.60607	0.5027
	4	2	0.00108	0.03280	0.02487	9.74489	0.5027
	4	3	0.00134	0.03662	0.02671	9.58364	0.5027
	5	1	0.00037	0.01921	0.01100	4.51840	0.5027
	5	2	0.00073	0.02708	0.01907	7.28880	0.5027
	5	3	0.00150	0.03877	0.02876	10.53460	0.5023
Canada	3	1	0.00041	0.02033	0.01324	3.45460	0.4947
	3	2	0.00084	0.02902	0.02434	7.46130	0.4947
	3	3	0.00122	0.03497	0.02926	8.59726	0.4947
	4	1	0.00045	0.02121	0.01739	5.39288	0.4947
	4	2	0.00073	0.02710	0.01751	4.03373	0.4947
	4	3	0.00175	0.04182	0.03535	10.52540	0.4947
	5	1	0.00057	0.02382	0.01712	4.85236	0.4947
	5	2	0.00092	0.03025	0.01891	4.04763	0.4947
	5	3	0.00134	0.03655	0.02612	6.15847	0.4947
France	3	1	0.00039	0.01970	0.01190	4.59060	0.5153
	3	2	0.00079	0.02809	0.01798	6.19398	0.5157
	3	3	0.00117	0.03425	0.02382	8.78002	0.5160
	4	1	0.00054	0.02333	0.01406	4.74266	0.5157
	4	2	0.00115	0.03385	0.02245	7.14418	0.5160
	4	3	0.00155	0.03932	0.02658	8.68075	0.5157
	5	1	0.00079	0.02803	0.01600	4.36051	0.5160
	5	2	0.00144	0.03795	0.02453	7.12387	0.5157
	5	3	0.00240	0.04898	0.03120	8.75925	0.5157

Table 17: Predicting exchange rates using backpropagation learning

Country	Prediction Zones		Measures				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	1	0.00016	0.01248	0.00873	2.62739	0.5093
	3	2	0.00025	0.01593	0.01195	3.47193	0.5093
	3	3	0.00069	0.02633	0.02014	5.17859	0.5093
	4	1	0.00014	0.01195	0.00877	2.71973	0.5093
	4	2	0.00026	0.01619	0.01221	3.52174	0.5093
	4	3	0.00052	0.02288	0.01745	4.58203	0.5090
	5	1	0.00015	0.01206	0.00880	2.72424	0.5093
	5	2	0.00033	0.01825	0.01374	3.79916	0.5090
	5	3	0.00102	0.03199	0.02565	6.54115	0.5087
Japan	3	1	0.00028	0.01683	0.01138	2.67773	0.5117
	3	2	0.00075	0.02741	0.02057	4.28728	0.5117
	3	3	0.00064	0.02524	0.01994	4.43512	0.5117
	4	1	0.00062	0.02483	0.01711	3.48249	0.5117
	4	2	0.00063	0.02514	0.01891	4.07529	0.5117
	4	3	0.00105	0.03240	0.02543	5.30260	0.5117
	5	1	0.00043	0.02080	0.01482	3.27662	0.5117
	5	2	0.00073	0.02696	0.02060	4.31716	0.5117
	5	3	0.00089	0.02980	0.02363	4.99340	0.5113
Switzerland	3	1	0.00028	0.01679	0.01360	4.33588	0.5173
	3	2	0.00042	0.02057	0.01431	3.96080	0.5173
	3	3	0.00049	0.02216	0.01834	5.60996	0.5177
	4	1	0.00022	0.01475	0.01137	3.66554	0.5173
	4	2	0.00043	0.02071	0.01700	5.26242	0.5177
	4	3	0.00052	0.02280	0.01895	5.83149	0.5177
	5	1	0.00021	0.01435	0.01086	3.48037	0.5177
	5	2	0.00049	0.02212	0.01832	5.60067	0.5177
	5	3	0.00055	0.02341	0.01945	5.90427	0.5180

Table 18: Predicting exchange rates using backpropagation learning (contd.)

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	3	0.00030	0.01730	0.00945	4.56127	0.4800
	3	5	0.00033	0.01828	0.00952	3.72249	0.4883
	4	3	0.00032	0.01785	0.01065	5.19780	0.4800
	4	5	0.00032	0.01788	0.00937	4.19723	0.4883
	5	3	0.00033	0.01813	0.01111	5.50543	0.4800
	5	5	0.00029	0.01713	0.00927	4.42731	0.4883
Britain	3	3	0.00030	0.01731	0.01301	4.72203	0.5063
	3	5	0.00020	0.01400	0.00797	3.51844	0.5103
	4	3	0.00037	0.01920	0.01510	5.41047	0.5063
	4	5	0.00023	0.01507	0.01034	3.97212	0.5103
	5	3	0.00032	0.01801	0.01388	4.97657	0.5063
	5	5	0.00028	0.01678	0.01252	4.57775	0.5103
Canada	3	3	0.00026	0.01597	0.01397	4.28454	0.5037
	3	5	0.00020	0.01419	0.00962	2.44134	0.5003
	4	3	0.00029	0.01717	0.01473	4.62887	0.5037
	4	5	0.00017	0.01289	0.01062	3.10447	0.5003
	5	3	0.00026	0.01619	0.01402	4.34835	0.5037
	5	5	0.00023	0.01504	0.01299	4.01296	0.5003
France	3	3	0.00021	0.01438	0.01081	4.29085	0.5117
	3	5	0.00022	0.01482	0.01002	4.89227	0.5120
	4	3	0.00023	0.01510	0.01230	5.00503	0.5117
	4	5	0.00020	0.01411	0.01032	4.77671	0.5120
	5	3	0.00029	0.01693	0.01426	5.70611	0.5117
	5	5	0.00021	0.01445	0.01150	4.87873	0.5120

Table 19: Predicting averages of exchange rates using backpropagation learning

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	3	0.00008	0.00920	0.00575	1.75202	0.5060
	3	5	0.00034	0.01854	0.01430	5.14011	0.5010
	4	3	0.00008	0.00878	0.00603	1.83180	0.5060
	4	5	0.00009	0.00974	0.00460	1.31736	0.5010
	5	3	0.00007	0.00846	0.00568	1.75125	0.5060
	5	5	0.00008	0.00913	0.00647	2.02970	0.5010
Japan	3	3	0.00051	0.02259	0.01457	2.82833	0.5140
	3	5	0.00020	0.01429	0.01100	2.68518	0.5197
	4	3	0.00042	0.02042	0.01684	3.60866	0.5140
	4	5	0.00069	0.02620	0.02242	4.63945	0.5197
	5	3	0.00156	0.03949	0.03374	6.51951	0.5140
	5	5	0.00050	0.02236	0.01882	3.96439	0.5197
Switzerland	3	3	0.00027	0.01643	0.01395	4.11947	0.5227
	3	5	0.00044	0.02094	0.01884	5.46716	0.5030
	4	3	0.00020	0.01419	0.01182	3.59338	0.5227
	4	5	0.00020	0.01420	0.01108	3.72766	0.5030
	5	3	0.00029	0.01716	0.01410	3.94991	0.5227
	5	5	0.00024	0.01556	0.01317	3.92474	0.5030

Table 20: Predicting averages of exchange rates using backpropagation learning (contd.)

5.4 The Method of Recurrent Nets

Time underlines many interesting human behaviour. Thus, the question of how to represent time in connectionist models is very important. One approach is to represent time implicitly by its effects rather than explicitly(spatial representation of time). This chapter deals with the use of recurrent links (first described by Jordan [16]) in order to provide networks with a dynamic memory. This means giving the processing system some dynamic properties that are responsive to temporal sequence.

There are many ways in which this can be accomplished. A number of interesting proposals have appeared in the literature, of which the connectionist model suggested by William and Zisper [26] has been chosen for this thesis work.

The design of networks having a memory was first suggested by Jordan [16]. It consisted of recurrent connections that were used to associate a static pattern with a serially ordered output pattern. The recurrent connections allow the network's hidden units to see its own previous output , so that the subsequent behaviour can be shaped by previous responses. These recurrent connections are responsible for the network memory.

This approach can be modified in the following way: Suppose a network is augmented at the input level by additional units (the state variables) . These units are hidden in the sense that they interact exclusively with other nodes internal to the network , and not the outside world. In this thesis, the units are considered to be scalars. Let the network be trained with the first set of input values at time t. Both the input units (which we call external input lines here after) and the units activate the hidden units; the hidden units then feed forward to activate the output units. The hidden units also feedback to activate the units. This constitutes the forward activation. Depending upon the task, there may or may not be a teacher forced learning, which means the output is compared with a desired output and the error is back-propagated to the network in order to adjust the connection weights incrementally. At the next time step i.e. at $t+1$ the units contain values which are exactly the hidden unit values at time t. Thus these units provide memory to the net in this manner. This is called back-propagation-through-time computation. One of the main advantages of these models is its generality, but on the other hand, its disadvantage is its growing memory requirement when given an arbitrarily long input sequence. And this was a matter of concern in the case of predicting future values of a time series which have very long training patterns which is the main focus of the current work.

Thus an algorithm was needed which had all the generalities of the backpropagation-through-time approach, yet did not suffer from growing memory requirement in arbitrarily long training sequence. The algorithm suggested by Williams and Zisper met the above requirements.

5.4.1 The Basic Algorithm

Let the network have n units, with m external input lines. Let y(t) denote the n-tuple of outputs of the units in time t , and let x(t) denote the m-tuple of external input signals to the network at time t. Let y(t) and x(t) be concatenated to form the $(m+n)$-tuple z(t), with U denoting the indices k such that z_{kt} is the output of a unit in the network and I the set of indices k for which z_{kt} is an external input. The indices on y and x are chosen to those of, so that

$$z_k(t) = \begin{cases} x_k(t) & \text{if } k \in I \\ y_k(t) & \text{if } k \in U \end{cases} \tag{5.4.1}$$

Let W be the weight matrix between every pair of the units and also from each input line to each unit. Hence the weight matrix has dimension $n \times (m+n)$. A bias weight of value 1 is added to each of the m input lines. All the computations of this network are based on the assumption that all the units are semi-linear in nature.

Let

$$s_k(t) = \sum_{U \cup I} w_{kt} z_t(t) \tag{5.4.2}$$

be the net input to the kth unit at time t, for $k \in U$. Then the output at the next time step is given by:

$$y_k(t+1) = f_k(s_k(t)) \tag{5.4.3}$$

where f_k is a squashing function.

The equations defined by 5.4.2 and 5.4.3 defines the entire dynamics of the network. The external input at time t does not influence the output of any unit until time $t+1$. The learning method adopted to train the recurrent net is called "temporal supervised learning". In this type of learning method, the output values of certain units should match a given desired(target) value. Let $T(t)$ be set of indices $k \in U$ for which there exists a specified target value $d_k(t)$. Then the error (a time dependent n-tuple) is given by:

$$e_k(t) = \begin{cases} d_k(t) - y_k(t) & \text{if } k \in T(t) \\ 0 & \text{otherwise} \end{cases} \tag{5.4.4}$$

It is to be noted that this formulation allows the possibility that target values are specified for different units at different times.

Let the overall network error at time t be given by:

$$J(t) = \frac{1}{2} \sum_{k \in U} |e_k(t)|^2 \tag{5.4.5}$$

The goal is to minimize the error function at any particular time. Let the total error be given by:

$$J_{total}(t_0, t_1) = \sum_{t=t_0+1}^{t_1} J(t) \tag{5.4.6}$$

where t_0 and t_1 are the initial and final times respectively that the network was operating.

Hence, our problem reduces to an optimization problem of minimization of the total error given by equation 5.4.6. This is done by the *gradient descent method* and the weights W are adjusted along the negative of $\nabla_W J_{total}(t_0, t+1)$.

Since the error function is a sum of squares of errors at each time step, it makes sense to calculate the gradient by accumulating the values of $\nabla_W J(t)$ for each time step along the trajectory. The overall weight change for any particular weight w_{ij} in the network can be written as:

$$\Delta w_{ij} = \sum_{t=t_0+1}^{t_1} \Delta w_{ij}(t) \tag{5.4.7}$$

where

$$\Delta w_{ij}(t) = -\alpha \frac{\partial J(t)}{\partial w_{ij}} \tag{5.4.8}$$

Here α is some fixed positive learning rate.

Now,

$$\begin{aligned} -\frac{\partial J(t)}{\partial w_{ij}} &= -\frac{\partial[\frac{1}{2}\sum_{k\in U}|e_k(t)|^2]}{\partial w_{ij}} \\ &= -\sum_{k\in U} \frac{\partial e_k(t)}{\partial w_{ij}}[d_k(t) - y_k(t)] \\ &= \sum_{k\in U} e_k(t) \frac{\partial y_k(t)}{\partial w_{ij}} \end{aligned} \tag{5.4.9}$$

Also,

$$\begin{aligned}
\frac{\partial y_k(t+1)}{\partial w_{ij}} &= \frac{\partial y_k(t)}{\partial s_k(t)} \frac{\partial s_k(t)}{\partial w_{ij}} \\
&= f_k'(s_k(t)) \; [\frac{\partial[\sum_{l\in U} w_{kl} z_l]}{\partial w_{ij}}] \\
&= f_k'(s_k(t)) \; [\sum_{l\in U} \frac{\partial w_{kl}}{\partial w_{ij}} z_l \; + \; \sum_{l\in U} w_{kl} \frac{\partial z_l}{\partial w_{ij}}] \\
&= f_k'(s_k(t)) \; [\delta_{ik} z_j(t) + \; \sum_{l\in U} w_{kl} \frac{\partial y_l(t)}{\partial w_{ij}}] \qquad (5.4.10)
\end{aligned}$$

where δ_{ik} is the *kronecker delta*.

Since, by our assumption, at $t = t_0$, the network has no functional dependence on the weights, we have:

$$\frac{\partial y_k(t_0)}{\partial w_{ij}} = 0 \qquad (5.4.11)$$

All the above mentioned equations hold for $i \in U$, $j \in U \cup I$, $k \in U$.

Let $\{p_{ij}^k\}$ define the dynamics of the network system for all $i \in U$, $j \in U \cup I$, $k \in U$. We can write:

$$p_{ij}^k(t+1) = f_k'(s_k(t)) \; [\delta_{ik} z_j(t) + \; \sum_{l\in U} w_{kl} p_{ij}^k(t)] \qquad (5.4.12)$$

with initial conditions:

$$p_{ij}^k(t_0) = 0 \qquad (5.4.13)$$

It follows that:

$$p_{ij}^k(t) = \frac{\partial y_k(t)}{\partial w_{ij}} \qquad (5.4.14)$$

Hence, we can finally say that:

$$\Delta w_{ij}(t) = \alpha \sum_{k \in U} e_k(t) p_{ij}^k(t) \tag{5.4.15}$$

In the case where each unit uses the logistic squashing function we can use:

$$f_k'(s_k(t)) = y_k(t+1)[1 - y_k(t+1)] \tag{5.4.16}$$

5.4.2 The Recurrent Nets Algorithm

The basic algorithm discussed in the preceding section can now be summarized as follows:

Step 1: Initialize the number of units and inputs and weights.

Step 2: Calculate $p_{ij}^k(t)$ using equations 5.4.12 and 5.4.13.

Step 3: Calculate the errors $e_k(t)$, i.e., the discrepancies between the desired and actual outputs.

Step 4: Calculate the weight changes $\Delta w_{ij}(t)$ using equation 5.4.15. Then calculate the overall weight change which is the sum of $\Delta w_{ij}(t)$.

Step 5: IF all data is used THEN stop., ELSE go to Step 2.

The above algorithm is implemented from time step t_0 to t_1.

5.4.3 Experimental Results

While conducting the experiments, the learning rate α was chosen to be 0.5. A fully connected recurrent net was taken into consideration for each experiment.

1. For window of size 3, the number of units chosen was 7, the number of inputs was 4, number of outputs was 1.

2. For window of size 4, the number of units chosen was 7, the number of inputs was 5 and the number of outputs was 1.

3. For window of size 5, the number of units chosen was 7, the number of inputs chosen was 6 and the number of outputs was 1.

In most of the experiments 800 data were used for training and the remaining 2200 data were used for testing. The performance of recurrent nets in predicting 3 days averages of the Canadian-U.S. exchange rates is shown in the following figure: The

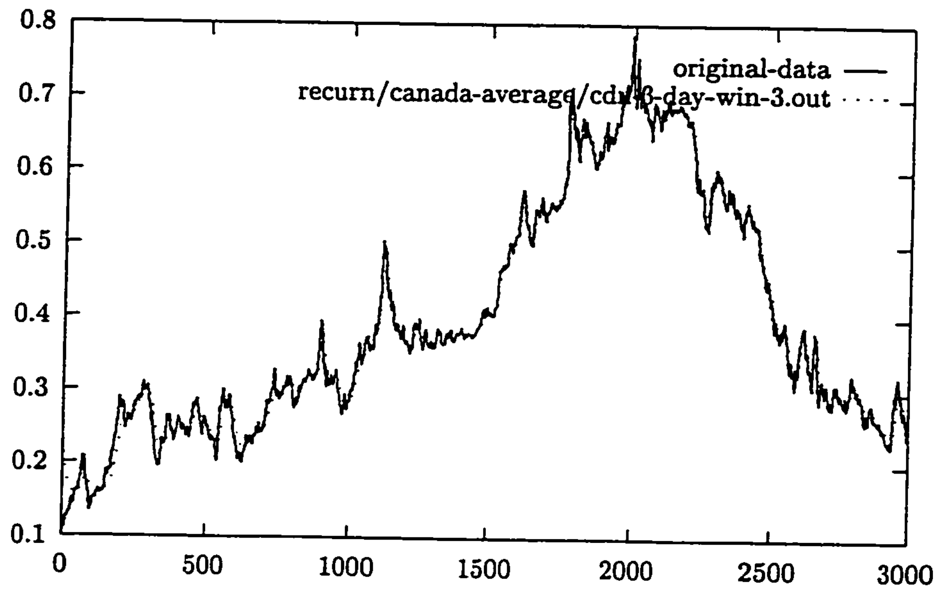


Figure 13: Predicting 3 days averages using recurrent nets with window size 3 on Canadian-U.S.exchange rates

overall MSE range from 0.00015 to 0.003. The mean is around 0.0002. The range of MSE for the averages is from 0.00012 to 0.003, the mean is about 0.002. The range of $RMSE$ is from 0.013 to 0.018 for exchange rates and from 0.012 to 0.07 for the averages of exchange rates. The mean of the former is around 0.01 and that for the latter is 0.01. MAE ranges from 0.011 to 0.08 for exchange rates and from 0.007 to 0.01 for the averages of exchange rates. The mean is about 0.01 for exchange rates and about 0.009 for the averages of exchange rates. $MAPE$ ranges from 2.7 to 4.2 for exchange rates and from 2.45 to 3.88 for the averages of exchange rates. The mean of the $MAPE$ values is around 3.4 for the exchange rates and 3.1 for the averages of exchange rates. Finally, for the DIR, the value ranges from 0.4947 to 0.5177 for exchange rates and from 0.4800 to 0.5120. The overall mean for the exchange rates is about 0.50 whereas that for the averages of exchange rates is about 0.50. The results are tabulated below:

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	1	0.00033	0.01830	0.00982	3.64269	0.4947
	3	2	0.00034	0.01840	0.00993	3.66497	0.4950
	3	3	0.00034	0.01857	0.01004	3.68592	0.4953
	4	1	0.00034	0.01840	0.00988	3.71291	0.4950
	4	2	0.00034	0.01832	0.00980	3.69041	0.4953
	4	3	0.00033	0.01828	0.00973	3.66846	0.4957
	5	1	0.00030	0.01745	0.00901	3.45559	0.4953
	5	2	0.00032	0.01777	0.00920	3.49215	0.4957
	5	3	0.00032	0.01797	0.00941	3.53040	0.4953
Britain	3	1	0.00031	0.01761	0.01074	4.06058	0.5027
	3	2	0.00033	0.01810	0.01118	4.20570	0.5027
	3	3	0.00033	0.01818	0.01129	4.23611	0.5027
	4	1	0.00034	0.01846	0.01118	4.25047	0.5027
	4	2	0.00034	0.01836	0.01112	4.23468	0.5027
	4	3	0.00033	0.01827	0.01105	4.21559	0.5027
	5	1	0.00029	0.01699	0.01010	3.89456	0.5027
	5	2	0.00030	0.01720	0.01032	3.94784	0.5027
	5	3	0.00030	0.01736	0.01049	3.99381	0.5023
Canada	3	1	0.00025	0.01591	0.01050	3.51022	0.4947
	3	2	0.00026	0.01622	0.01081	3.58105	0.4947
	3	3	0.00027	0.01648	0.01107	3.64201	0.4947
	4	1	0.00027	0.01639	0.01083	3.62309	0.4947
	4	2	0.00027	0.01648	0.01093	3.64444	0.4947
	4	3	0.00027	0.01649	0.01099	3.65142	0.4947
	5	1	0.00021	0.01464	0.00950	3.18991	0.4947
	5	2	0.00023	0.01511	0.00996	3.29638	0.4947
	5	3	0.00024	0.01546	0.01029	3.37482	0.4947
France	3	1	0.00019	0.01393	0.00922	3.58210	0.5153
	3	2	0.00020	0.01410	0.00933	3.60805	0.5157
	3	3	0.00020	0.01425	0.00941	3.62851	0.5160
	4	1	0.00020	0.01413	0.00929	3.65532	0.5157
	4	2	0.00020	0.01411	0.00923	3.63751	0.5160
	4	3	0.00020	0.01402	0.00918	3.62201	0.5157
	5	1	0.00018	0.01326	0.00846	3.37482	0.5160
	5	2	0.00018	0.01343	0.00863	3.41565	0.5157
	5	3	0.00018	0.01351	0.00876	3.44685	0.5157

Table 21: Predicting exchange rates using recurrent nets

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	1	0.00018	0.01328	0.00937	2.89034	0.5093
	3	2	0.00018	0.01345	0.00950	2.92458	0.5093
	3	3	0.00018	0.01352	0.00960	2.95381	0.5093
	4	1	0.00018	0.01354	0.00957	2.97581	0.5093
	4	2	0.00018	0.01339	0.00948	2.95640	0.5093
	4	3	0.00018	0.01324	0.00938	2.93185	0.5090
	5	1	0.00015	0.01233	0.00857	2.68529	0.5093
	5	2	0.00016	0.01253	0.00876	2.73485	0.5090
	5	3	0.00016	0.01263	0.00890	2.77308	0.5087
Japan	3	1	0.00020	0.01414	0.00976	2.98604	0.5117
	3	2	0.00021	0.01444	0.01001	3.04784	0.5117
	3	3	0.00021	0.01466	0.01020	3.09604	0.5117
	4	1	0.00020	0.01422	0.00970	3.04101	0.5117
	4	2	0.00020	0.01421	0.00970	3.04717	0.5117
	4	3	0.00020	0.01414	0.00966	3.04532	0.5117
	5	1	0.00017	0.01302	0.00876	2.78407	0.5117
	5	2	0.00018	0.01337	0.00908	2.86305	0.5117
	5	3	0.00018	0.01360	0.00932	2.92334	0.5113
Switzerland	3	1	0.00023	0.01532	0.01111	3.55045	0.5173
	3	2	0.00024	0.01553	0.01132	3.61106	0.5173
	3	3	0.00025	0.01570	0.01148	3.66041	0.5177
	4	1	0.00025	0.01567	0.01136	3.66276	0.5173
	4	2	0.00024	0.01557	0.01132	3.65967	0.5177
	4	3	0.00024	0.01553	0.01127	3.64629	0.5177
	5	1	0.00020	0.01419	0.01015	3.29750	0.5177
	5	2	0.00021	0.01459	0.01047	3.38643	0.5177
	5	3	0.00022	0.01491	0.01068	3.44908	0.5180

Table 22: Predicting exchange rates using recurrent nets (contd.)

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	3	0.00031	0.01751	0.00919	3.51430	0.4800
	3	5	0.00030	0.01720	0.00898	3.46326	0.4883
	4	3	0.00031	0.01751	0.00911	3.56930	0.4800
	4	5	0.00030	0.01722	0.00891	3.51834	0.4883
	5	3	0.00029	0.01701	0.00863	3.37488	0.4800
	5	5	0.00027	0.01644	0.00812	3.26851	0.4883
Britain	3	3	0.00028	0.01685	0.01008	3.88978	0.5063
	3	5	0.00028	0.01668	0.00991	3.84059	0.5103
	4	3	0.00029	0.01714	0.01003	3.92180	0.5063
	4	5	0.00029	0.01699	0.00986	3.87270	0.5103
	5	3	0.00025	0.01591	0.00909	3.57094	0.5063
	5	5	0.00022	0.01499	0.00883	3.50595	0.5103
Canada	3	3	0.00022	0.01477	0.00937	3.24747	0.5037
	3	5	0.00021	0.01451	0.00920	3.19451	0.5003
	4	3	0.00022	0.01486	0.00926	3.26460	0.5037
	4	5	0.00021	0.01466	0.00913	3.22202	0.5003
	5	3	0.00018	0.01345	0.00817	2.86724	0.5037
	5	5	0.00016	0.01282	0.00761	2.72583	0.5003
France	3	3	0.00017	0.01320	0.00866	3.43716	0.5117
	3	5	0.00017	0.01302	0.00851	3.39473	0.5120
	4	3	0.00018	0.01329	0.00864	3.50031	0.5117
	4	5	0.00017	0.01312	0.00846	3.44739	0.5120
	5	3	0.00016	0.01259	0.00799	3.25012	0.5117
	5	5	0.00015	0.01223	0.00763	3.16142	0.5120

Table 23: Predicting averages of exchange rates using recurrent nets

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Japan	3	3	0.00015	0.01236	0.00874	2.72614	0.5060
	3	5	0.00015	0.01213	0.00860	2.68972	0.5010
	4	3	0.00016	0.01253	0.00886	2.79731	0.5060
	4	5	0.00015	0.01232	0.00867	2.75039	0.5010
	5	3	0.00014	0.01163	0.00799	2.52047	0.5060
	5	5	0.00012	0.01116	0.00764	2.44117	0.5010
Switzerland	3	3	0.00017	0.01315	0.00896	2.80381	0.5140
	3	5	0.00017	0.01298	0.00883	2.78017	0.5197
	4	3	0.00017	0.01296	0.00866	2.79701	0.5140
	4	5	0.00016	0.01281	0.00852	2.77257	0.5197
	5	3	0.00015	0.01205	0.00793	2.56100	0.5140
	5	5	0.00013	0.01159	0.00752	2.48576	0.5197
	3	3	0.00021	0.01432	0.01028	3.31803	0.5227
	3	5	0.00020	0.01406	0.01010	3.26866	0.5030
	4	3	0.00021	0.01455	0.01039	3.38883	0.5227
	4	5	0.00020	0.01430	0.01023	3.34723	0.5030
	5	3	0.00018	0.01333	0.00930	3.03518	0.5227
	5	5	0.00016	0.01279	0.00888	2.93006	0.5030

Table 24: Predicting averages of exchange rates using recurrent nets (contd.)

5.5 An Adaptively Trained Neural Network

So far, in the previous chapter, stationary time series have been taken into consideration. In this section, attention has been drawn to predict slowly varying non-stationary time series data. An adaptively trained neural net is being used to learn and hence predict these kind of time series data. The theory behind this learning method has been adopted from the works of Dong C. Park,Mohammed A., El-Sharkawi,Robert J. Marks II [24]. The following two conditions should be satisfied for updating the weights of a layered perceptron.

1. the procedure should still respond appropriately to the previous training data if those data are not in conflict with the new training data.

2. the procedure should adapt to the new training data even when they are in conflict with portions of the old data.

5.5.1 Formulation of the Problem

Assume a layered perceptron artificial neural network trained with N sets of data. $(x(1), d(1)), (x(2), d(2)), ...(x(N), d(N))$ where x(i) and d(i) represent the input and desired output for the i^{th} data set and $1 \leq i \leq N$. We assume that $x(i)$ is an I-dimensional vector and $d(i)$ is a scalar. The layered perceptron is assumed to have one hidden layer with h hidden neurons. The matrix $\mathbf{W}$ represents the weight matrix between the input and hidden neurons and $\mathbf{v}$ denotes the weight vector which links the hidden and output neurons. The dimensions of $\mathbf{W}$ and $\mathbf{v}$ are $I \times h$ and $h \times 1$ respectively.

For a given input data vector, $x(i)$, the output of the layered perceptron of $y(i)$ is given by:

$$y(i) = f[v^T u] \tag{5.5.1}$$

$$u = f[W^T x(i)] \qquad (5.5.2)$$

where $u = [u_1, u_2, ... u_h]^T$, u_j, $1 \leq j \leq h$ represents the activation of the j^{th} hidden neuron; f[.] is the sigmoid function:

$$f[x] = 1/(1 + exp(-x)), \quad x \in R$$

and f[b] is the $h \times 1$ vector function

$$f[b] = [f_1[b].f_2[b], ..., f_h[b]]^T.$$

The sigmoid function for each hidden neurons is assumed to be identical.

$$f_1[.] = f_2[.] = \cdots = f_h[.] = f[.]$$

We assume that $W(N)$ and $v(N)$ are the weights that minimize the error function:

$$E(N) = \frac{1}{2}\sum_{i=1}^{n}(d(i) - y(i))^2. \qquad (5.5.3)$$

5.5.2 Problem Statement

The forecasting problem is effectively an optimization problem. It involves minimization of the error function described by equation 5.5.3. Mathematically, the problem can be stated as follows: Given $W(N)$, $v(N)$, the N sets of data, and $(x(N+1), d(N+1))$,determine $W(N+1)$ and $v(N+1)$ such that

$$E(N+1) = \frac{1}{2}\sum_{i=1}^{N+1}(d(i) - y(i))^2 = E(N) + \frac{1}{2}(d(N+1) - y(N+1))^2 \qquad (5.5.4)$$

is minimized in such a manner that $y(N+1) \approx d(N+1)$.

5.5.3 Linearization Process

In order to promote tractibility of analysis and implementation , it is recommended to linearize a given problem. Therefore, 5.5.1 and 5.5.2 gives:

$$d(N+1) = f[\mathbf{v}_T(N+1)\mathbf{u}] = f[(\mathbf{v}(N) + \Delta\mathbf{v})^T\mathbf{u}] \tag{5.5.5}$$

where

$$\mathbf{u} = \mathbf{f}[\mathbf{W^T}(N+1)\mathbf{x}(N+1)] = \mathbf{f}[(\mathbf{W}(N) + \Delta\mathbf{W})^T\mathbf{x}(N+1)]. \tag{5.5.6}$$

We expand the terms defined by equations 5.5.5 and 5.5.6 in a truncated Taylor's Series about $\{(\mathbf{x}(\text{N+1}),\mathbf{W}(\text{N+1})\}$, in the neighborhood of $\{(\mathbf{x}(N), \mathbf{W}(N)\}$. Such a linearization is used in Kalman filtering [14] and in quasi-linearization in [2], [17] and [25].

After a series of calculations (these are described in details in appendix B.1.1) we can derive the activations of the hidden neurons as:

$$\mathbf{u} \simeq \mathbf{u}^* + (\nabla_\mathbf{b}\mathbf{u}^*)\Delta\mathbf{W}^T\mathbf{x}(N+1) \tag{5.5.7}$$

Inverting 5.5.5 gives

$$f^{-1}[d(N+1)] = \mathbf{v}^T(N)\mathbf{u} + \Delta\mathbf{v}^T\mathbf{u} \tag{5.5.8}$$

where $f^{-1}[x] = ln(x/(1-x))$. From 5.5.7 and 5.5.8 we see that:

$$\begin{aligned} f^{-1}[d(N+1)] - \mathbf{v}^T(N)\mathbf{u}^* \;\; \simeq \;\; & \mathbf{v}^T(N)(\nabla_\mathbf{b}\mathbf{u}^*)\Delta\mathbf{W}^T\mathbf{x}(N+1) + \Delta\mathbf{v}^T\mathbf{u}^* + \\ & \Delta\mathbf{v}^T\nabla_\mathbf{b}\mathbf{u}^*\Delta\mathbf{W}^T\mathbf{x}(N+1) \end{aligned} \tag{5.5.9}$$

It is to be noted that the perturbation in $\Delta\mathbf{W}$ has to be small in order to use 5.5.7 .

If $\Delta\mathbf{v} \ll \Delta\mathbf{u}$ then the third term of the right hand side of 5.5.9 can be ignored and we can write:

$$f^{-1}[d(N+1)] - \mathbf{v}^T(N)\mathbf{u}^* \simeq \mathbf{v}^T(N)(\nabla_b\mathbf{u}^*)\Delta\mathbf{W}^T\mathbf{x}(N+1) + \Delta\mathbf{v}^T\mathbf{u}^* \quad (5.5.10)$$

$\Delta\mathbf{W}$ can be represented into a vector form as follows:

$$\Delta\mathbf{W}_{vec} = [\Delta\mathbf{w_1}^T, \Delta\mathbf{w_2}^T, \cdots \Delta\mathbf{w_I}^T]^T = [\Delta W_{vec,1} \cdots \Delta W_{vec,p}]$$

where $\Delta\mathbf{w}_i$ is the i^{th} row and $p = h \times I$ where p is the number of interconnections between the neurons in the input and the hidden layers.Then equation 5.5.10 can be rewritten as:

$$c_1 = [\Delta\mathbf{W}_{vec}^T : \Delta\mathbf{v}^T] \begin{bmatrix} \mathbf{u}^+ \\ .. \\ \mathbf{u}^* \end{bmatrix} = \mathbf{z}^T\mathbf{a} \quad (5.5.11)$$

where c_1, $\mathbf{a}$, $\mathbf{z}$ are vectors defined by:

$$c_1 \equiv f^{-1}[d(N+1)] - \mathbf{v}^T(N)\mathbf{u}^* \quad (5.5.12)$$

$$\mathbf{a} \equiv [\mathbf{u}^+ : \mathbf{u}^*]^T \quad (5.5.13)$$

$$\mathbf{z} \equiv [\Delta\mathbf{W}_{vec}^T : \Delta\mathbf{v}^T]^T \quad (5.5.14)$$

and $\mathbf{u}^+$ is a solution of

$$\Delta\mathbf{W}_{vec}^T\mathbf{u}^+ = \mathbf{v}^T(N)\mathbf{Q}\Delta\mathbf{W}^T\mathbf{x}(N+1) \quad (5.5.15)$$

Since there are $h \times (I+1)$ unknowns with one equation 5.5.11, there exist many solutions. In order to preserve uniqueness of solution we impose the additional constraint of our problem statement. Specifically, it is essential to change the weights of the $(N+1)st$ datum with minimum effect on the previous N data. Hence it is important to find the sensitivity of $y(i)$ over a weight change. The error term described by equation 5.5.4 can be rewritten as:

$$E(N) = \frac{1}{2}\sum_{i=1}^{N}(d(i) - y(i))^2 = \sum_{i=1}^{N} E_i$$

The sensitivity for the input weights and output weights are respectively as follows:

$$\frac{\partial E_i}{\partial w_{jk}} = -[d(i) - y(i)]\left(\frac{\partial y(i)}{\partial w_{jk}}\right) \quad (5.5.16)$$

$$\frac{\partial E_i}{\partial v_k} = -[d(i) - y(i)]\left(\frac{\partial y(i)}{\partial v_k}\right) \quad (5.5.17)$$

where w_{jk} , the weight of interconnection between input neuron j and the hidden neuron k is the jkth element of $\mathbf{W}$. Now, we can write:

$$\Delta E_i = \sum_{j,k}\left(\frac{\partial E_i}{\partial w_{jk}}\right)\Delta \mathbf{w}_{jk} + \sum_{k}\left(\frac{\partial E_i}{\partial w_{jk}}\right)\Delta \mathbf{v}_k \quad (5.5.18)$$

Equation 5.5.18 can be represented in a matrix form as below :

$$\mathbf{\Delta E} = \mathbf{ASz} \quad (5.5.19)$$

where the ith element ΔE is ΔE_i: The matrix $\mathbf{A}$ is given by:

$$\mathbf{A} = diag[\epsilon_1, \cdots, \epsilon_N] \quad (5.5.20)$$

here

$$\epsilon_i = -(d(i) - y(i)), \ 1 \leq i \leq N, \quad and \ p = I \times h.$$

and

$$\mathbf{S} = \begin{bmatrix} SW_{1,1} & \cdots & SW_{1,p} & SV_{1,1} & \cdots & SV_{1,h} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ SW_{N,1} & \cdots & SW_{N,p} & SV_{N,1} & \cdots & SV_{N,h} \end{bmatrix} \tag{5.5.21}$$

where $SV_{i,k}$ is the sensitivity caused by small changes in v_k and $SW_{i,jk}$ is the sensitivity of y(i) to the $w_{j,k}$'s. S is called the sensitivity matrix. and $\mathbf{z}$ is a $q \times 1$ vector defined in 5.5.11 and $q = p + h = (I + 1) \times h$.

The derivations of the sensitivity matrix and its associated terms are shown in details in the appendix B.1.2.

The weights $\mathbf{W}(N)$ and $\mathbf{v}(N)$ are optimal for minimizing $E(N)$. With the addition of the $(N+1)$st datum, $d(N+1) = y(N+1)$, the objective function of 5.5.4 can be changed to:

$$\begin{aligned} \mathbf{J} &= \frac{1}{2}\sum_{i=1}^{N}(E_{i,W(N)} - E_{i,W(N+1)})^2 \\ &= \frac{1}{2}\sum_{i=1}^{N}\Delta E_i^2 \end{aligned}$$

where $E_{i,W(N)}$ and $E_{i,W(N+1)}$ are the errors of the ith datum with $\{\mathbf{W}(N), \mathbf{v}(N)\}$ and $\{\mathbf{W}(N+1), \mathbf{v}(N+1)\}$, respectively. Equivalently, from 5.5.19

$$\begin{aligned} \mathbf{J} &= \frac{1}{2}(\Delta E)^T(\Delta E) \\ &= \frac{1}{2}(\mathbf{ASz})^{\mathbf{T}}(\mathbf{ASz}) \\ &= \frac{1}{2}\mathbf{z}^T\mathbf{K}\mathbf{z} \end{aligned} \tag{5.5.22}$$

where $\mathbf{K} = \mathbf{S}^T(\mathbf{A}^T\mathbf{A})\mathbf{S}$. Note that $\mathbf{K} = \mathbf{K}^T$. Since there exists only one equation with q unknowns in 5.5.11 , the solutions of the equations in 5.5.11 are on a $(q-1)$

dimensional hyper-plane. Since small perturbations for $\mathbf{W}$ and $\mathbf{v}$ are assumed in 5.5.7, among the solutions in 5.5.11, only those with small variation in weight space is allowable.

The problem now boils down to a standard nonlinear programming problem:

$$\begin{aligned} \textit{minimize} \ \ \mathbf{J}(\mathbf{z}) &= \frac{1}{2}\mathbf{z}^T\mathbf{K}\mathbf{z} \\ \textit{subject to} \ \ \mathbf{z}^T\mathbf{a} &= c_1 \end{aligned}$$

A equicost line of the cost function $\mathbf{J}(\mathbf{z})$ represents a q dimensional hyper-ellipse centered at the origin of the $\mathbf{z}$ plane. The constraint is a $(q-1)$ dimensional hyper-plane on the $\mathbf{z}$ plane. The shape of the cost function is given by the eigen values of $\mathbf{K}$. We need to introduce a boundary condition on $\mathbf{z}$ so that the perturbations are small enough to meet the linearization assumption and allow $\mathbf{z}$ to be large enough so that there is at least one solution on the $\mathbf{z}^T\mathbf{a} = c_1$ plane.

5.5.4 Boundary Constraint

We introduce the boundary as a set of points in the $\mathbf{z}$ plane:

$$\mathcal{B} = \{ \ \mathbf{z} : -\vec{\ell} \leq \mathbf{z} \leq \vec{\ell} \ \}$$

where

$$\vec{\ell} = [\ell_1, \ell_2, \cdots, \ell_q]^T$$

and

$$\ell_i \geq 0, \ \ i = 1, 2, \cdots, q$$

Let $\hat{\mathbf{z}}$ be included in the boundary constraint which is given by:

$$\hat{\mathbf{z}} = \frac{c_1\mathbf{a}}{\mathbf{a}^T\mathbf{a}} \tag{5.5.23}$$

Using the boundary condition our problem now becomes:

$$\begin{aligned} minimize \ \ \mathbf{J(z)} &= \frac{1}{2}\mathbf{z}^T\mathbf{K}\mathbf{z} \\ subject \ \ to \ \ \mathbf{z}^T\mathbf{a} &= c_1 \\ and \ \ \mathbf{z} \in \mathcal{B} \end{aligned} \tag{5.5.24}$$

We use the *Reduced Gradient Algorithm* to solve the above optimization problem.

5.5.5 The Reduced Gradient Algorithm

In order to understand the theory behind the performance of an ATNN network, it is very important to know the underlying mathematical concepts. In particular, one of the most important stages of learning in an ATNN network is finding the point where the error function is minimum. There exists a variety of nonlinear programming methods to solve the problem described by 5.5.24. Of these, the Reduced Gradient Method (RGM) using a linear constraint has been found to be most effective in this case. The method has been adopted from [2]. The basic idea behind this method involves partitioning the variables into 2 classes:

1. Basic Variables: dependent variables and the cardinality of the basic group is the number of constraints (which is 1 in our case).

2. Non-basic Variables: independent variables . In our case there are $(q-1)$ non-basic variables.

Assume we have a feasible solution $\mathbf{z}$ satisfying the constraint equation given by 5.5.24. Partitioning the vector $\mathbf{z}$ into basic and non-basic components yield:

$$(z_\alpha, z_\beta) = [z_\alpha, z_\beta^T]^T$$

where z_α is a one-dimensional basic variable and z_β is a $(q-1)$ dimensional non-basic variable. Hence our optimization problem 5.5.24 can be written as:

$$\begin{aligned} minimize \quad \mathbf{J}(z_\alpha, z_\beta) &= \frac{1}{2}(z_\alpha, z_\beta^T)\mathbf{K}(z_\alpha, z_\beta) && (5.5.25) \\ subject \ \ to \quad z_\alpha a_\alpha + \mathbf{a}_\beta^{\mathbf{T}}\mathbf{z}_\beta &= c_1 && (5.5.26) \\ and \quad (z_\alpha, z_\beta) &\in \mathcal{B} && (5.5.27) \end{aligned}$$

where

$$\mathbf{a} = [z_\alpha, z_\beta^T]^T$$

The RGM minimizes the objective function iteratively only in terms of the independent variables. The movement of the dependent variable is controlled by equation 5.5.26. Let

$$\begin{aligned} \mathbf{z}_\beta(k+1) &= \mathbf{z}_\beta(k) + \Delta\mathbf{z}_\beta \\ \mathbf{z}_\alpha(k+1) &= \mathbf{z}_\alpha(k) + \Delta\mathbf{z}_\alpha \end{aligned}$$

where k is an iteration index. Substituting the above in 5.5.26 we get:

$$\Delta z_\alpha a_\alpha + \mathbf{a}_\beta^{\mathbf{T}}\mathbf{z}_\beta = 0$$

this gives:

$$\Delta z_\alpha = -\frac{1}{a_\alpha}\mathbf{a}_\beta^T \Delta\mathbf{z}_\beta \qquad (5.5.28)$$

The new point about $(k+1)$, i.e., $(z_\alpha(k+1), \mathbf{z}_\beta)$ is always on the $(q-1)$ dimensional hyper-plane given by 5.5.26, this is ensured by equation 5.5.28.

The steepest descent method is now used to find the direction of the the independent variables in RGM. The gradient of each variable decides the direction of movement at each step. The gradient of $J(z)$ is given by:

$$\nabla_z J(z) \;=\; (\nabla_{z_\alpha} J(z), \nabla_{z_\beta} J(z)) = \mathbf{Kz} = \mathbf{K}(\mathbf{z}_\alpha, \mathbf{z}_\beta)$$

Since the relationships of movements of z_α should satisfy 5.5.28 , the amount of movement ∇_{z_α} , for z_α can be decomposed into the amount of movement for z_β . The *reduced gradient*, i.e., the gradient with respect to z_β is found to be:

$$r^T \;=\; \nabla_{z_\alpha} J(z_\alpha, z_\beta) - \frac{1}{a_\alpha} \nabla_{z_\alpha} J(z_\alpha, z_\beta) \mathbf{a}_\beta \tag{5.5.29}$$

where $r = [r_1, r_2, \cdots, r_{q-1}]^T$. Since the boundary is a box, each variable has a corresponding boundary defined by two sides of the box. In order to allow the possibility to add and delete elements from the boundary, the concept of *working set* is introduced. A constraint is defined as *active* if the corresponding variable lies on the boundary and is on the verge of violating the boundary conditions. At each step of the iteration, these active constraints are chosen and these constitute the working set and is denoted by $\mathcal{W}(\mathbf{z}_\beta)$. The vector $\mathbf{z}_\beta$ moves in the direction of its gradient unless it violates the boundary conditon, in which case it becomes a part of the working set. Hence, we can write:

$$\Delta z_{\beta,i} = \begin{cases} -r_i, & \text{if } i \notin \mathcal{W}(\mathbf{z}_\beta) \\ 0, & \text{otherwise} \end{cases}$$

If it happens that at a point $r_i = 0 \forall i \notin \mathcal{W}(\mathbf{z}_\beta)$ but there exists $j \in \mathcal{W}(\mathbf{z}_\beta)$ such that either $r_j < 0$ and j is in the working set because its corresponding variable violated its lower boundary constraint, or $r_j > 0$ and j has been put into the working set because its corresponding variable violated its upper boundary constraint. Then

j is deleted from the working set. Once the amount of movement of z_β is found, Δz_α is calculated using 5.5.28. The RGM usually converges to a local minimum. For the solution to reach the global minimum, it is necessary for the boundary constraint $\mathcal{B}$ to be a convex set and the objective function $J(\mathbf{z})$ to be a convex function. We are now in a position to state the *Reduced Gradient Algorithm*

Step 1: Set the initial feasible solution $\mathbf{z} = \hat{\mathbf{z}}$.

Step 2: Set $z_\alpha = z_1$ and $\mathbf{z}_\beta = [z_2, z_3, \cdots, z_q]^T$. Initialize $\mathcal{W}(\mathbf{z}_\beta) = \phi$., where ϕ denotes an empty set.

Step 3: Calculate r^T defined by 5.5.29.

Step 4: Find the direction of movement of $\mathbf{z}_\beta$ using the concept of working sets as mentioned above.

Step 5: Reconstruct the working set $\mathcal{W}(\mathbf{z}_\beta)$ if necessary.

Step 6: IF acceptable solution is reached , i.e., if $|\Delta z| \leq \epsilon$, where ϵ is the convergence measure, or $\forall \in \mathcal{W}(\mathbf{z}_\beta)$ THEN stop ELSE find Δz_α as described by equation 5.5.28.

Step 7: Find $\gamma_1, \gamma_2, \gamma_3$ such that

$$
\begin{aligned}
max\{\gamma_1 &: -l_\alpha \leq z_\alpha + \gamma_1 \Delta z_\alpha \leq l_\alpha : \gamma_1 \geq 0\} \\
max\{\gamma_2 &: -\vec{l}_\beta \leq \mathbf{z}_\beta + \gamma_2 \Delta \mathbf{z}_\beta \leq \vec{l}_\beta : \gamma_2 \geq 0\} \\
min\{\gamma_3 &: J(\mathbf{z} + \gamma_3 \Delta z) : 0 \leq \gamma_3 \leq \gamma_1 ; 0 \leq \gamma_3 \leq \gamma_2\}
\end{aligned}
$$

Step 8: Calculate $\mathbf{z} = \mathbf{z} + \gamma_3 \Delta \mathbf{z}$

Step 9: IF $\gamma_3 < \gamma_1$, THEN go to Step 3., ELSE declare the dependent variable to be independent and declare one of the independent variables which is positively inside of the non-linearity constraint to be dependent. Update a_α and $\mathbf{a}_\beta$. Go to Step 3.

5.5.6 The ATNN Algorithm

The following ATNN algorithm was used for conducting the experiments on exchange data.

Step 1: Begin

Step 2: Calculate $\mathbf{A}$ and $\mathbf{S}$ by using N data as described 5.5.20 and 5.5.21.

Step 3: Find $\mathbf{K}$ such that $\mathbf{K} = \mathbf{S}^{\mathbf{T}}(\mathbf{A}^{\mathbf{T}}\mathbf{A})\mathbf{S}$.

Step 4: By using 5.5.13 and 5.5.14 find the linear constraint equation

$$\mathbf{z}^{\mathbf{T}}\mathbf{a} = c_1$$

.

Step 5: Find the boundary constraint ,

$$\vec{\ell} = 2\hat{\mathbf{z}}$$

by using

$$\hat{\mathbf{z}} = \frac{c_1\mathbf{a}}{\mathbf{a}^{\mathbf{T}}\mathbf{a}}$$

Step 6: Perform the reduced gradient method as stated in the previous section and find $\mathbf{z}$ which minimizes the cost function

$$J(\mathbf{z}) = \frac{1}{2}\mathbf{z}^{\mathbf{T}}\mathbf{K}\mathbf{z}$$

with the constraints

$$\mathbf{z}^{\mathbf{T}}\mathbf{a} = c_1$$

and

$$-\vec{\ell} \leq \mathbf{z} \leq \vec{\ell}$$

.

Step 7: Find $\Delta\mathbf{W}$ and $\Delta\mathbf{v}$ by rearranging $\mathbf{z}$ and resulting ΔW_{vec}.

$$[\Delta\mathbf{W}_{\mathbf{vec}}^{\mathbf{T}} : \Delta\mathbf{v}^{\mathbf{T}}] = \mathbf{z}^T$$

.

Step 8: Update $\mathbf{W}(N)$ and $\mathbf{v}(N)$:

$$\begin{aligned} \mathbf{W}(N+1) &= \mathbf{W}(N) + \Delta\mathbf{W} \\ \mathbf{v}(N+1) &= \mathbf{v}(N) + \Delta\mathbf{v} \end{aligned}$$

.

Step 9: Stop.

5.5.7 Experimental Results

For the experiments conducted using the ATNN algorithm mentioned in the last section, the number of input neurons I was the same as the size of the window (i.e., 3, 4, or 5), number of hidden neurons h was 5 and 1 output neuron. The total number of input data N is 3000. The weights $\mathbf{W}$ and $\mathbf{v}$ were initialized randomly.

An example of performance of ATNN in predicting 3 days average of Canadian-U.S. exchange rates is shown by the following graph: The overall MSE range from 0.00007 to 0.0004. The mean is around 0.0001. The range of MSE for the averages is from 0.00005 to 0.00025, the mean is about 0.0001. The range of $RMSE$ is from 0.01 to 0.03 for exchange rates and from 0.006 to 0.02 for the averages of exchange rates. The mean of the former is around 0.001 and that for the latter is 0.008. MAE ranges from 0.004 to 0.014 for exchange rates and from 0.002 to 0.008 for the averages

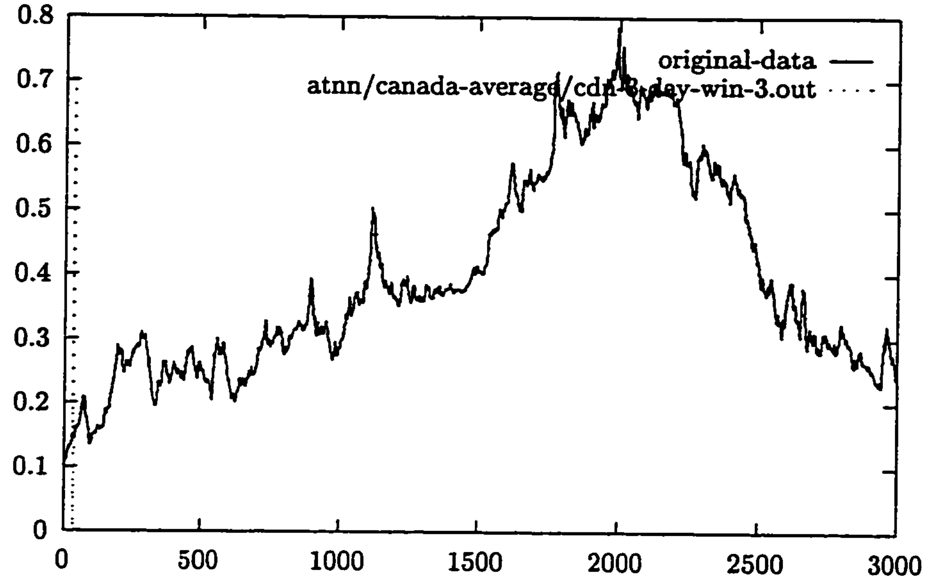


Figure 14: Predicting 3 days average of Canadian-U.S. exchange rates with window size 3

of exchange rates. The mean is about 0.005 for exchange rates and about 0.003 for the averages of exchange rates. $MAPE$ ranges from 1.2 to 3.08 for exchange rates and from 0.65 to 1.63 for the averages of exchange rates. The mean of the $MAPE$ values is around 1.4 for the exchange rates and 0.7 for the averages of exchange rates. Finally, for the DIR, the value ranges from 0.4712 to 0.5077 for exchange rates and from 0.4718 to 0.5145. The overall mean for the exchange rates is about 0.47 whereas that for the averages of exchange rates is about 0.48. The prediction results are summarized in the following tables:

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	1	0.00120	0.03463	0.00528	2.01662	0.4732
	3	2	0.00020	0.01425	0.00436	1.16808	0.4732
	3	3	0.00022	0.01493	0.00507	1.41514	0.4728
	4	1	0.00115	0.03391	0.00875	2.47993	0.4732
	4	2	0.00020	0.01418	0.00438	1.16963	0.4728
	4	3	0.00020	0.01421	0.00439	1.17144	0.4725
	5	1	0.00020	0.01422	0.00428	1.17986	0.4728
	5	2	0.00022	0.01475	0.00483	1.35499	0.4725
	5	3	0.00025	0.01590	0.00427	1.27183	0.4725
Britain	3	1	0.00083	0.02875	0.00508	1.91489	0.4715
	3	2	0.00026	0.01624	0.00469	1.50370	0.4712
	3	3	0.00055	0.02343	0.00487	1.57982	0.4715
	4	1	0.00013	0.01162	0.00448	1.36237	0.4712
	4	2	0.00014	0.01177	0.00449	1.36423	0.4715
	4	3	0.00014	0.01195	0.00450	1.36417	0.4715
	5	1	0.00018	0.01340	0.00452	1.45778	0.4715
	5	2	0.00026	0.01610	0.00450	1.41999	0.4715
	5	3	0.00016	0.01260	0.00435	1.35445	0.4715
Canada	3	1	0.00045	0.02110	0.00758	2.04880	0.4738
	3	2	0.00010	0.00985	0.00506	1.44254	0.4732
	3	3	0.00577	0.07597	0.04848	13.17730	0.4735
	4	1	0.00010	0.01014	0.00534	1.48122	0.4735
	4	2	0.00010	0.01020	0.00537	1.48714	0.4735
	4	3	0.00023	0.01531	0.01013	2.71704	0.4732
	5	1	0.00011	0.01053	0.00518	1.48292	0.4735
	5	2	0.00010	0.00983	0.00508	1.43503	0.4735
	5	3	0.00011	0.01055	0.00518	1.47537	0.4735
France	3	1	0.00396	0.06292	0.01481	8.87341	0.4855
	3	2	0.00015	0.01228	0.00442	1.41839	0.5008
	3	3	0.00125	0.03538	0.01474	4.07721	0.5012
	4	1	0.00012	0.01084	0.00402	1.23625	0.4858
	4	2	0.00009	0.00938	0.00409	1.27157	0.5012
	4	3	0.00009	0.00946	0.00410	1.27295	0.5012
	5	1	0.00012	0.01102	0.00388	1.24106	0.4858
	5	2	0.00014	0.01172	0.00501	1.55434	0.5012
	5	3	0.00016	0.01257	0.00422	1.36830	0.5008

Table 25: Predicting exchange rates using ATNN learning

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	1	0.00009	0.00963	0.00424	1.20394	0.4885
	3	2	0.00024	0.01549	0.00448	1.27676	0.4885
	3	3	0.00010	0.00975	0.00429	1.22306	0.4885
	4	1	0.00007	0.00844	0.00442	1.23239	0.4885
	4	2	0.00007	0.00842	0.00439	1.22729	0.4885
	4	3	0.00007	0.00831	0.00418	1.18827	0.4888
	5	1	0.00027	0.01649	0.00505	1.44618	0.4885
	5	2	0.00008	0.00917	0.00422	1.20161	0.4888
	5	3	0.00029	0.01705	0.00463	1.33308	0.4892
Japan	3	1	0.00041	0.02015	0.00541	1.36130	0.5058
	3	2	0.00016	0.01269	0.00497	1.23171	0.5055
	3	3	0.00833	0.09126	0.00691	0.00691	0.5058
	4	1	0.00010	0.00999	0.00529	1.30368	0.5055
	4	2	0.00008	0.00888	0.00491	1.21595	0.5055
	4	3	0.00008	0.00883	0.00489	1.21237	0.5055
	5	1	0.00025	0.01578	0.00490	1.21036	0.5055
	5	2	0.00417	0.06457	0.00564	1.16874	0.5055
	5	3	0.00048	0.02200	0.01324	3.08024	0.5055
Switzerland	3	1	0.00021	0.01441	0.00531	1.55766	0.5082
	3	2	0.00009	0.00939	0.00492	1.44208	0.5075
	3	3	0.00011	0.01036	0.00513	1.49152	0.5092
	4	1	0.00009	0.00935	0.00512	1.48290	0.5075
	4	2	0.00009	0.00930	0.00513	1.48467	0.5078
	4	3	0.00025	0.01582	0.00519	1.50975	0.5082
	5	1	0.00010	0.01010	0.00495	1.45123	0.5078
	5	2	0.00010	0.01007	0.00494	1.44963	0.5082
	5	3	0.00012	0.01114	0.00505	1.48060	0.5078

Table 26: Predicting exchange rates using ATNN learning (contd.)

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	3	0.00018	0.01353	0.00275	0.78277	0.4718
	3	5	0.00051	0.02269	0.00882	4.95732	0.4882
	4	3	0.00019	0.01376	0.00339	0.85054	0.4715
	4	5	0.00018	0.01335	0.00285	0.71251	0.4882
	5	3	0.00044	0.02097	0.00357	1.14064	0.4715
	5	5	0.00075	0.02736	0.00519	1.47522	0.4882
Britain	3	3	0.00010	0.00980	0.00268	0.83529	0.5065
	3	5	0.00078	0.02798	0.00313	1.01436	0.4865
	4	3	0.00079	0.02799	0.00415	1.00986	0.4865
	4	5	0.00010	0.00994	0.00279	0.81255	0.4865
	5	3	0.00009	0.00999	0.00234	0.98109	0.4865
	5	5	0.00009	0.00970	0.00228	0.71099	0.4865
Canada	3	3	0.00019	0.01396	0.00550	1.61218	0.4832
	3	5	0.00007	0.00806	0.00343	0.91019	0.4822
	4	3	0.00007	0.00814	0.00371	1.01980	0.4832
	4	5	0.00006	0.00760	0.00313	0.85459	0.4822
	5	3	0.00027	0.01634	0.00335	1.03086	0.4832
	5	5	0.00006	0.00744	0.00248	0.74230	0.4822
France	3	3	0.00035	0.01865	0.00387	1.12995	0.4808
	3	5	0.00007	0.00823	0.00195	0.64035	0.4858
	4	3	0.00007	0.00821	0.00287	0.85211	0.4808
	4	5	0.00006	0.00792	0.00242	0.71192	0.4862
	5	3	0.00006	0.00783	0.00232	0.74423	0.4808
	5	5	0.00008	0.00907	0.00287	0.88008	0.4862

Table 27: Predicting averages of exchange rates using ATNN learning

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	3	0.00005	0.00688	0.00288	0.80511	0.5075
	3	5	0.00005	0.00725	0.00230	0.65639	0.5035
	4	3	0.00005	0.00735	0.00321	0.87681	0.5075
	4	5	0.00005	0.00692	0.00278	0.75115	0.5035
	5	3	0.00007	0.00845	0.00265	0.76005	0.5075
	5	5	0.00010	0.00991	0.00288	0.78218	0.5035
Japan	3	3	0.00007	0.00807	0.00296	0.74551	0.5008
	3	5	0.00007	0.00829	0.00238	0.60069	0.4915
	4	3	0.00010	0.01015	0.00579	1.63235	0.5012
	4	5	0.00005	0.00719	0.00322	0.76505	0.4902
	5	3	0.00008	0.00910	0.00339	0.82240	0.4998
	5	5	0.00018	0.01332	0.00251	0.62206	0.4902
Switzerland	3	3	0.00010	0.01024	0.00320	0.93327	0.5145
	3	5	0.00010	0.00981	0.00253	0.74976	0.4962
	4	3	0.00008	0.00875	0.00453	1.31255	0.5145
	4	5	0.00011	0.01034	0.00261	0.76391	0.4962
	5	3	0.00007	0.00849	0.00302	0.89105	0.5145
	5	5	0.00025	0.01572	0.00285	0.82335	0.4962

Table 28: Predicting averages of exchange rates using ATNN learning (contd.)

5.6 The Fuzzy Learning Method

Fuzzy set theory has been used as a basis for pattern recognition by many scientists. It is highly mathematical in nature, hence it is not too popular with the non-experts who find the mathematical theories too mystifying. Nevertheless, fuzzy set theory can provide a robust and consistent foundation for information processing, including pattern recognition. In this chapter a fuzzy learning algorithm given by Tao et.al.,[19] has been presented. Before proceeding to present this algorithm, some relevant concepts in connection to fuzzy theory is being presented.

5.6.1 Basic Concepts in Fuzzy Set Theory

Most of the concepts used in this section are adopted from [23]

1. *Fuzzy set:* If X is a collection of objects x then a fuzzy set A in X is a set of ordered pairs:

   $A = \{(x,\ \mu_A(x))|\ x \,\varepsilon\, X\}$

The entity μ_A is called the "Membership function". The value of which is the grade of membership of x in A. It is also the degree to which the deterministic measurement x is compatible with (the vague concept) of A. $\mu_A(X)$ is a mapping from X to a membership space M. If M contains only 2 values 0 and 1, then A is not fuzzy. The range of the values of the membership function is a subset of non-negative real numbers with a finite least upper bound. If this bound is unity,i.e.,

$sup\{\mu_A(x) = 1\}$ then the fuzzy set A is called normal. Note that fuzzy sets can be defined in terms of continuous as well as discrete membership functions.

5.6.2 A Fuzzy Reasoning Expert System

Fuzzy systems contain knowledge bases that contain many rules. These rules are typically knowledge-based which incorporate imprecise knowledge. Most fuzzy reasoning systems use the inference schema of the following form, as in the seminal work of Zadeh [28].

The inference schema adopted for the fuzzy learning algorithm of Tao et.al.,[19] is the one given by Cao and Kandel [6]. The inference schema can be described as follows:

$$
\begin{aligned}
&IF\ \mathbf{X}\ \text{is}\ \mathbf{U_1}\quad then\ \mathbf{Y}\ \text{is}\quad w_{11}/V_1 + w_{12}/V_2 + \cdots + w_{1k}/V_k\\
&IF\ \mathbf{X}\ \text{is}\ \mathbf{U_2}\quad then\ \mathbf{Y}\ \text{is}\quad w_{11}/V_1 + w_{12}/V_2 + \cdots + w_{1k}/V_k\\
&IF\ \mathbf{X}\ \text{is}\ \mathbf{U_n}\quad then\ \mathbf{Y}\ \text{is}\quad w_{11}/V_1 + w_{12}/V_2 + \cdots + w_{1k}/V_k
\end{aligned}
\tag{5.6.1}
$$

where $w_{11}/V_1 + w_{12}/V_2 + \cdots + w_{1k}/V_k$ can be interpreted as V_1 with degree of confidence w_{11} , V_2 with degree of confidence w_{12}, $\cdots$ and V_k with degree of confidence w_{1k}.

Let an input value a be first mapped into a vector $< \mu_1(a), \mu_2(a), \quad \cdots \quad \mu_n(a) >$, where μ_i is the fuzzy membership function of U_i. This vector is then multiplied to the weight matrix $\mathbf{W}$, where

$$\mathbf{W} = w_{ij}, \quad i = 1, 2, \cdots n, \quad j = 1, 2, \cdots, k \tag{5.6.2}$$

The resulting vector $< y_1(a), y_2(a), \quad \cdots \quad y_k(a) >$, is defuzzified by the moment method to produce:

$$y = \frac{h_1 + h_2 + \cdots + h_k}{y_1 + y_2 + \cdots + y_k} \tag{5.6.3}$$

where $0 \leq h_i \leq 1$ are the central values of the fuzzy subsets $V_1, V_2, \cdots, V_k$. The above method makes sense for symmetric output fuzzy label subsets. The relational mapping from the input fuzzy concepts to the output fuzzy concepts is accomplished through the matrix vector product $\mu(a) \cdot \mathbf{W}$

This system is applicable to both *discrete* as well as *continuous* universe of discourse.

5.6.3 The Fuzzy Learning Algorithm

In the present context of predicting future values for exchange rates experiments have been carried out using the membership function stated below for the training of the neural net using fuzzy logic.

The membership function is the positive half of the sine function given by

$$f_i(x) = \begin{cases} sin(x - \theta_i), & \text{if } 0 \leq x - \theta_i \leq \pi \\ 0, & \text{otherwise} \end{cases}$$

The shape of the above function looks like

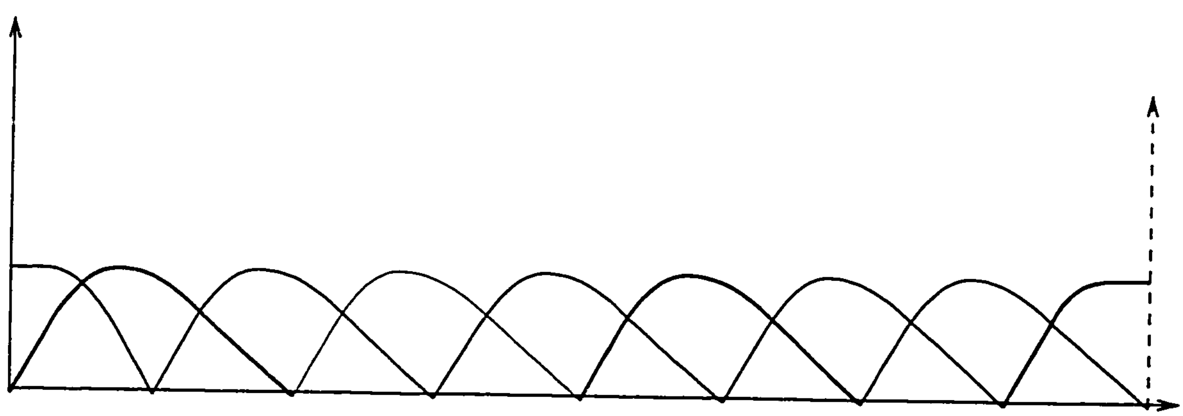


Figure 15: A fuzzy membership function

The widths of the membership functions are fixed and same throughout the experiment. It is to be noted that the columns of the weight matrix $\mathbf{W}$ must satisfy the condition:

$$\sum_{i=1}^{n} |w_{ij}| = 1, \quad for \ \ 1 \le j \le k; \tag{5.6.4}$$

Let $< h_1, h_2, \cdots, h_k >$ be the central values of the output fuzzy label subsets. The initial weight matrix **W** is generated randomly.

The fuzzy learning algorithm can be described in the following steps:

Step 1: (Forward propagation) Select an input x; $\xi =< \mu_1(a), \mu_2(a), \quad \cdots \quad \mu_n(a) >$ where the membership function μ_i is the same as defined at the beginning of the section. The function looks as shown in (15).

$< \gamma_1(a), \gamma_2(a), \quad \cdots \quad \gamma_k(a) > \leftarrow \xi \times W$

Here the h_i s are typically the central values of the k fuzzy sets generated by $< \gamma_1(a), \gamma_2(a), \quad \cdots \quad \gamma_k(a)$.

Step 2: (Backward error propagation) $e \leftarrow d - y$;

$< \varepsilon_1, \varepsilon_2, \quad \cdots, \quad \varepsilon_k > \leftarrow < F(h_1, e), F(h_2, e), \quad \cdots \quad F(h_k, e >$;

The functions $\varepsilon_i = F(h_i, e)$ chosen for the experiment is $h_i e$. This treats all the output fuzzy sets equally.

Step 3: Adjust weight matrix $w_{ij}(t+1) \leftarrow w_{ij}(t) + \alpha\mu_i(x)\varepsilon_i, \quad for \quad 1 \le j \le k, \quad 1 \le i \le n$;

$m_i \leftarrow \sum_{j=1}^{k} |w_{ij}(t+1)|, \quad for \ \ 1 \le i \le n; \quad w_{ij}(t+1) \leftarrow w_{ij}(t+1)/m_i; \quad for \ \ 1 \le i \le n; \ \ 1 \le j \le k$;

Here m_i is computed so that the constraint condition (5.6.4) is satisfied.

Step 4: IF stopping condition is satisfied THEN halt ELSE go to Step 1.

Note that the above mentioned algorithm holds true for multiple inputs too. The vector ξ is then the product of input and their corresponding membership functions.

5.6.4 Experimental Results

The above algorithm was implemented on the computer using the C programming language. The number of input neurons per run is the same as the window size. In constructing the fuzzy learning system for the experiments on exchange data, 5 input fuzzy labels (sets), i.e., the U_is and 6 output fuzzy labels(sets) are chosen i.e., the V_is. The performance of fuzzy learning in predicting the 3 days averages for Canadian-U.S. exchange rates is shown in the figure 16 The overall *MSE* range from 0.0004

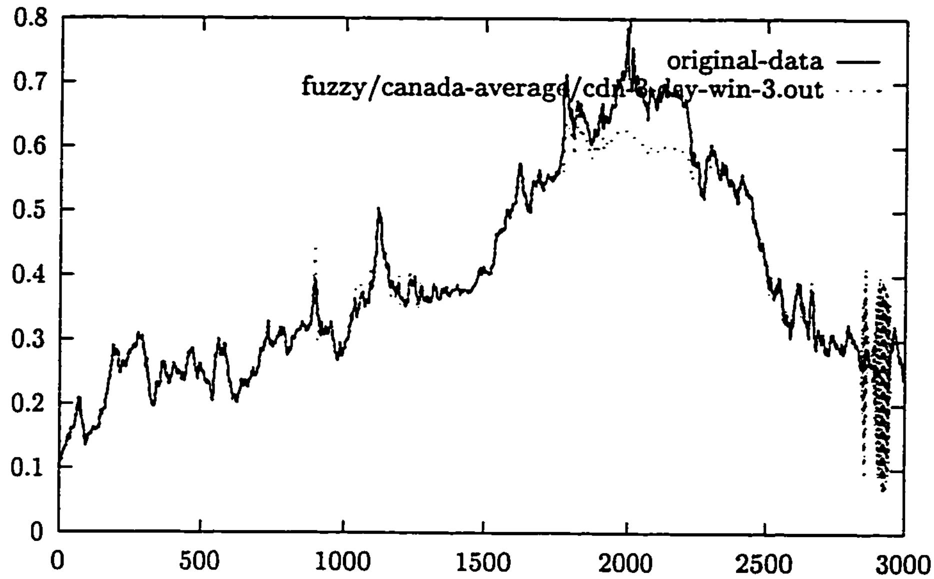


Figure 16: Predicting 3 days average of Can-U.S. exchange rates over window size 3 using fuzzy learning

to 0.002. The mean is around 0.0005. The range of *MSE* for the averages is from 0.0004 to 0.002, the mean is about 0.0004. The range of *RMSE* is from 0.02 to 0.04 for exchange rates and from 0.02 to 0.04 for the averages of exchange rates. The mean of the former is around 0.025 and that for the latter is 0.25. *MAE* ranges from 0.01 to 0.03 for exchange rates and from 0.009 to 0.01 for the averages of exchange rates. The mean is about 0.01 for exchange rates and about 0.01 for the averages of exchange rates. *MAPE* ranges from 2.87 to 7.2 for exchange rates and from 1.73 to 6.95 for the averages of exchange rates. The mean of the *MAPE* values is around

3.5 for the exchange rates and 3.5 for the averages of exchange rates. Finally, for the DIR, the value ranges from 0.48 to 0.50 for exchange rates and from 0.4720 to 0.5147. The overall mean for the exchange rates is about 0.49 whereas that for the averages of exchange rates is about 0.49. The results are described in the following tables:

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	1	0.00084	0.02897	0.01342	3.67570	0.4940
	3	2	0.00084	0.02898	0.01355	3.68765	0.4937
	3	3	0.00084	0.02900	0.01364	3.70435	0.4933
	4	1	0.00085	0.02917	0.01443	4.51882	0.4937
	4	2	0.00085	0.02920	0.01451	4.51498	0.4933
	4	3	0.00085	0.02923	0.01459	4.52035	0.4933
	5	1	0.00092	0.03038	0.01591	5.58995	0.4933
	5	2	0.00092	0.03041	0.01598	5.58640	0.4933
	5	3	0.00092	0.03041	0.01605	5.58677	0.4937
Britain	3	1	0.00044	0.02088	0.00983	3.48328	0.4840
	3	2	0.00044	0.02098	0.00999	3.52415	0.4840
	3	3	0.00044	0.02107	0.01006	3.52559	0.4843
	4	1	0.00049	0.02223	0.01073	4.07156	0.4840
	4	2	0.00050	0.02228	0.01086	4.08542	0.4843
	4	3	0.00050	0.02237	0.01094	4.09194	0.4843
	5	1	0.00059	0.02424	0.01186	4.77154	0.4843
	5	2	0.00059	0.02429	0.01199	4.79054	0.4843
	5	3	0.00059	0.02438	0.01207	4.80033	0.4843
Canada	3	1	0.00133	0.03646	0.01917	4.58846	0.4927
	3	2	0.00134	0.03655	0.01931	4.63336	0.4927
	3	3	0.00134	0.03660	0.01938	4.65808	0.4927
	4	1	0.00123	0.03511	0.01851	4.54201	0.4927
	4	2	0.00124	0.03519	0.01862	4.57732	0.4927
	4	3	0.00124	0.03523	0.01869	4.59349	0.4927
	5	1	0.00119	0.03449	0.01819	4.59261	0.4927
	5	2	0.00119	0.03455	0.01830	4.61916	0.4927
	5	3	0.00120	0.03461	0.01836	4.63233	0.4927
France	3	1	0.00058	0.02407	0.01235	4.24068	0.5007
	3	2	0.00058	0.02410	0.01251	4.29335	0.5010
	3	3	0.00059	0.02419	0.01258	4.30218	0.5010
	4	1	0.00060	0.02445	0.01296	4.96287	0.5010
	4	2	0.00060	0.02451	0.01308	4.98641	0.5010
	4	3	0.00060	0.02453	0.01314	4.99064	0.5007
	5	1	0.00071	0.02665	0.01414	5.88701	0.5010
	5	2	0.00071	0.02665	0.01425	5.90722	0.5007
	5	3	0.00071	0.02665	0.01432	5.91800	0.5007

Table 29: Predicting exchange rates using fuzzy learning

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	1	0.00070	0.02639	0.01306	3.65136	0.4897
	3	2	0.00071	0.02659	0.01327	3.72639	0.4897
	3	3	0.00072	0.02685	0.01346	3.79853	0.4893
	4	1	0.00056	0.02366	0.01196	3.19714	0.4897
	4	2	0.00056	0.02377	0.01208	3.23899	0.4893
	4	3	0.00057	0.02397	0.01219	3.27653	0.4897
	5	1	0.00048	0.02199	0.01120	2.87121	0.4893
	5	2	0.00049	0.02209	0.01135	2.91661	0.4897
	5	3	0.00050	0.02228	0.01147	2.95912	0.4897
Japan	3	1	0.00242	0.04920	0.03245	7.62744	0.5043
	3	2	0.00244	0.04937	0.03265	7.71978	0.5047
	3	3	0.00247	0.04965	0.03288	7.84541	0.5047
	4	1	0.00220	0.04688	0.03095	7.07335	0.5047
	4	2	0.00221	0.04697	0.03108	7.13383	0.5047
	4	3	0.00222	0.04713	0.03120	7.20701	0.5047
	5	1	0.00209	0.04567	0.03019	6.80225	0.5047
	5	2	0.00213	0.04614	0.03005	6.75668	0.5047
	5	3	0.00216	0.04652	0.03015	6.80946	0.5047
Switzerland	3	1	0.00069	0.02620	0.01285	3.76506	0.5070
	3	2	0.00069	0.02632	0.01302	3.81488	0.5073
	3	3	0.00070	0.02652	0.01317	3.87023	0.5073
	4	1	0.00053	0.02293	0.01172	3.24518	0.5073
	4	2	0.00053	0.02305	0.01192	3.30817	0.5073
	4	3	0.00054	0.02326	0.01210	3.36954	0.5077
	5	1	0.00043	0.02078	0.01095	2.86387	0.5073
	5	2	0.00044	0.02097	0.01112	2.90930	0.5077
	5	3	0.00045	0.02120	0.01128	2.96116	0.5073

Table 30: Predicting exchange rates using fuzzy learning (contd.)

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Australia	3	3	0.00082	0.02872	0.01291	3.54097	0.4720
	3	5	0.00080	0.02835	0.01255	3.47241	0.4883
	4	3	0.00084	0.02890	0.01387	4.38877	0.4717
	4	5	0.00082	0.02861	0.01356	4.31511	0.4883
	5	3	0.00091	0.03016	0.01538	5.43881	0.4717
	5	5	0.00089	0.02985	0.01504	5.37472	0.4883
Britain	3	3	0.00041	0.02022	0.00826	3.08358	0.5067
	3	5	0.00041	0.02017	0.00822	3.04082	0.4867
	4	3	0.00047	0.02174	0.00946	3.70942	0.5067
	4	5	0.00046	0.02155	0.00905	3.58511	0.4867
	5	3	0.00057	0.02378	0.01077	4.43784	0.5067
	5	5	0.00056	0.02359	0.01023	4.29299	0.4867
Canada	3	3	0.00131	0.03618	0.01732	3.95936	0.4833
	3	5	0.00130	0.03601	0.01728	3.95269	0.4823
	4	3	0.00121	0.03479	0.01693	3.98323	0.4833
	4	5	0.00119	0.03456	0.01650	3.85659	0.4823
	5	3	0.00116	0.03406	0.01665	4.03053	0.4833
	5	5	0.00114	0.03382	0.01601	3.83645	0.4823
France	3	3	0.00055	0.02349	0.01092	3.85204	0.4807
	3	5	0.00054	0.02322	0.01084	3.79356	0.4860
	4	3	0.00058	0.02411	0.01179	4.61674	0.4807
	4	5	0.00057	0.02393	0.01149	4.52965	0.4860
	5	3	0.00069	0.02627	0.01309	5.52632	0.4807
	5	5	0.00068	0.02613	0.01265	5.42259	0.4860

Table 31: Predicting averages of exchange rates using fuzzy learning

Country	*Prediction Zones*		*Measures*				
	window	horizon	MSE	RMSE	MAE	MAPE	Ave DIR
Germany	3	3	0.00065	0.02554	0.01066	2.75275	0.5073
	3	5	0.00064	0.02536	0.01038	2.63431	0.5033
	4	3	0.00054	0.02315	0.01007	2.49979	0.5077
	4	5	0.00052	0.02287	0.00947	2.29753	0.5033
	5	3	0.00044	0.02109	0.00935	2.17171	0.5077
	5	5	0.00043	0.02065	0.00877	1.98431	0.5033
Japan	3	3	0.00240	0.04895	0.03092	6.95412	0.5003
	3	5	0.00238	0.04883	0.03057	6.84215	0.4913
	4	3	0.00218	0.04667	0.02946	6.39976	0.5010
	4	5	0.00216	0.04647	0.02904	6.25286	0.4900
	5	3	0.00204	0.04517	0.02863	6.06158	0.4997
	5	5	0.00203	0.04503	0.02832	5.94812	0.4900
Switzerland	3	3	0.00064	0.02532	0.01007	2.82046	0.5147
	3	5	0.00063	0.02502	0.01003	2.77308	0.4960
	4	3	0.00048	0.02187	0.00925	2.38776	0.5147
	4	5	0.00047	0.02158	0.00866	2.19000	0.4963
	5	3	0.00038	0.01946	0.00878	2.10187	0.5147
	5	5	0.00035	0.01882	0.00770	1.73010	0.4963

Table 32: Predicting averages of exchange rates using fuzzy learning (contd.)

Chapter 6

Comparison of Statistical and Neural Network Methods in Forecasting

In chapters 4 and 5, some statistical and neural network methods of predicting the future values of a time series have been presented. (Please see Chapters 4 and 5). These include (in alphabetical order):

1. The Adaptively Trained Neural Network (*atnn*) method.
2. The method of Backpropagation (*backprop*).
3. The Box-Jenkins Methodology (*boxjen*).
4. The method of Double-exponential smoothing (*dblexpn*).
5. The method of Exponential smoothing (*exponen*).
6. The Fuzzy learning method (*fuzzy*).
7. The method of Recurrent nets (*recurn*).
8. The method of Regression (*regrsn*).

The acronyms for the different methods are mentioned in the corresponding brackets.

In this chapter a comparison of prediction accuracy of the above methods is being discussed. Most of the analysis has been carried out based on the performance metrics of each method over a given prediction horizon.

As mentioned in Chapter 3, the interest of the forecasters in the field of finance is to predict the directions of the data, i.e., the ups and downs of the curve instead of predicting the actual values. Since the time series data we are dealing in this thesis is related to finance, it is expected to focus our attention to the performance metric DS(directional symmetry) mainly, in order to compare the efficiency of the models in predicting future values. Nevertheless, the other measures also are important in determining the correctness of prediction, so they are also taken into consideration when comparing the performances of forecasting methods.

6.1 Prediction of Exchange Rates

In Chapter 2, we have seen that the patterns of the exchange rates of Britain (Fig.2), France (Fig.4), Germany (Fig.5) and Switzerland (Fig.7) are quite similar. In order to avoid repetition of analysis, the exchange rate of Britain versus United States has been chosen for comparative study. The behaviour pattern of Australian (Fig.1), Canadian (Fig.3), Japanese (Fig.6) exchange rates graphs are quite distinct from each other. Hence,the comparative study of the performances of different methods of prediction are made based on the time series data for exchange rates of Australia, Britain, Canada and Japan.

In this section we seek and compare the efficiencies of each of the forecasting techniques discussed so far in predicting the future values of exchange rates of Australia, Britain, Canada, France, Germany, Japan and Switzerland with respect to the U.S. \$. As mentioned before these exchange rates are vastly varying and exhibit complex patterns, almost chaotic in nature. Hence, it is really difficult to forecast the exchange rates over a given length of time.

Let us consider the following tables which describe the performance of the above mentioned methods in forecasting the future values of the exchange rates of Australia, Britain, Canada and Japan with respect to United States. All these tables represent a summary of the experimental findings. The remaining tables for exchange rates prediction of France, Germany and Switzerland is given in the appendix C.

For the purpose of organization of results, the performance of each of the methods are measured against each of the metrics MSE, RMSE, MAE, MAPE, and Ave.DIR. The last one denotes directional errors. All the analysis of the results have been made by studying tables 33- 41.

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00120	0.03463	0.00528	2.01662	0.4732
	backprop	0.00054	0.02319	0.01448	4.85627	0.4947
	boxjen	0.00238	0.04877	0.00472	1.14679	0.4805
	dblexpn	0.00024	0.01535	0.00733	1.99216	0.4938
	exponen	0.00508	0.07127	0.02856	9.57696	0.4940
	fuzzy	0.00084	0.02897	0.01342	3.67570	0.4940
	recurn	0.00033	0.01830	0.00982	3.64269	0.4947
	regrsn	0.00350	0.05919	0.00629	1.63987	0.4940
Britain	atnn	0.00083	0.02875	0.00508	1.91489	0.4715
	backprop	0.00032	0.01798	0.01020	4.04879	0.5027
	boxjen	0.00263	0.05125	0.00499	1.50737	0.4753
	dblexpn	0.00015	0.01236	0.00743	2.31263	0.4838
	exponen	0.00470	0.06856	0.02656	8.32390	0.4840
	fuzzy	0.00044	0.02088	0.00983	3.48328	0.4840
	recurn	0.00031	0.01761	0.01074	4.06058	0.5027
	regrsn	0.00371	0.06088	0.00639	1.95005	0.4840
Canada	atnn	0.00045	0.02110	0.00758	2.04880	0.4738
	backprop	0.00041	0.02033	0.01324	3.45460	0.4947
	boxjen	0.00253	0.05032	0.00579	1.66220	0.4779
	dblexpn	0.00018	0.01349	0.00909	2.59633	0.4928
	exponen	0.00539	0.07344	0.02832	9.42555	0.4927
	fuzzy	0.00133	0.03646	0.01917	4.58846	0.4927
	recurn	0.00025	0.01591	0.01050	3.51022	0.4947
	regrsn	0.00398	0.06311	0.00757	2.32648	0.4927
Japan	atnn	0.00041	0.02015	0.00541	1.36130	0.5058
	backprop	0.00028	0.01683	0.01138	2.67773	0.5117
	boxjen	0.00308	0.05546	0.00490	1.85489	0.5005
	dblexpn	0.00013	0.01144	0.00784	1.95591	0.5042
	exponen	0.00552	0.07427	0.03238	8.33185	0.5043
	fuzzy	0.00242	0.04920	0.03245	7.62744	0.5043
	recurn	0.00020	0.01414	0.00976	2.98604	0.5117
	regrsn	0.00397	0.06305	0.00668	1.89479	0.5043

Table 33: 1 day ahead forecasts of exchange rates over window size 3

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00020	0.01425	0.00436	1.16808	0.4732
	backprop	0.00095	0.03083	0.02107	6.23941	0.4950
	boxjen	0.00239	0.04890	0.00472	1.14802	0.4805
	dblexpn	0.00029	0.01697	0.00862	2.31598	0.4938
	exponen	0.00508	0.07125	0.02856	9.57738	0.4937
	fuzzy	0.00084	0.02898	0.01355	3.68765	0.4937
	recurn	0.00034	0.01840	0.00993	3.66497	0.4950
	regrsn	0.00355	0.05959	0.00787	2.04784	0.4937
Britain	atnn	0.00026	0.01624	0.00469	1.50370	0.4712
	backprop	0.00069	0.02631	0.01891	7.91936	0.5027
	boxjen	0.00263	0.05133	0.00499	1.51093	0.4750
	dblexpn	0.00019	0.01369	0.00859	2.67282	0.4838
	exponen	0.00470	0.06857	0.02655	8.32361	0.4840
	fuzzy	0.00044	0.02098	0.00999	3.52415	0.4840
	recurn	0.00033	0.01810	0.01118	4.20570	0.5027
	regrsn	0.00374	0.06119	0.00784	2.40467	0.4840
Canada	atnn	0.00010	0.00985	0.00506	1.44254	0.4732
	backprop	0.00084	0.02902	0.02434	7.46130	0.4947
	boxjen	0.00257	0.05074	0.00581	1.67176	0.4779
	dblexpn	0.00023	0.01506	0.01035	2.93700	0.4925
	exponen	0.00539	0.07340	0.02832	9.40732	0.4927
	fuzzy	0.00134	0.03655	0.01931	4.63336	0.4927
	recurn	0.00026	0.01622	0.01081	3.58105	0.4947
	regrsn	0.00403	0.06345	0.00924	2.78404	0.4927
Japan	atnn	0.00016	0.01269	0.00497	1.23171	0.5055
	backprop	0.00075	0.02741	0.02057	4.28728	0.5117
	boxjen	0.00308	0.05550	0.00490	1.86863	0.5005
	dblexpn	0.00016	0.01283	0.00894	2.22553	0.5045
	exponen	0.00551	0.07424	0.03235	8.32723	0.5047
	fuzzy	0.00244	0.04937	0.03265	7.71978	0.5047
	recurn	0.00021	0.01444	0.01001	3.04784	0.5117
	regrsn	0.00401	0.06332	0.00819	2.25459	0.5047

Table 34: 2 days ahead forecasts of exchange rates over window size 3

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00022	0.01493	0.00507	1.41514	0.4728
	backprop	0.00111	0.03328	0.02415	7.57341	0.4953
	boxjen	0.00239	0.04887	0.00472	1.14843	0.4803
	dblexpn	0.00034	0.01841	0.00968	2.58769	0.4935
	exponen	0.00507	0.07119	0.02856	9.57462	0.4933
	fuzzy	0.00084	0.02900	0.01364	3.70435	0.4933
	recurn	0.00034	0.01857	0.01004	3.68592	0.4953
	regrsn	0.00359	0.05991	0.00910	2.37831	0.4933
Britain	atnn	0.00055	0.02343	0.00487	1.57982	0.4715
	backprop	0.00135	0.03679	0.02931	11.67030	0.5027
	boxjen	0.00262	0.05116	0.00499	1.50480	0.4753
	dblexpn	0.00022	0.01499	0.00966	3.00333	0.4842
	exponen	0.00472	0.06869	0.02655	8.32836	0.4843
	fuzzy	0.00044	0.02107	0.01006	3.52559	0.4843
	recurn	0.00033	0.01818	0.01129	4.23611	0.5027
	regrsn	0.00380	0.06162	0.00913	2.80789	0.4843
Canada	atnn	0.00577	0.07597	0.04848	13.17730	0.4735
	backprop	0.00122	0.03497	0.02926	8.59726	0.4947
	boxjen	0.00256	0.05058	0.00580	1.66729	0.4782
	dblexpn	0.00027	0.01651	0.01150	3.24070	0.4928
	exponen	0.00539	0.07339	0.02832	9.39131	0.4927
	fuzzy	0.00134	0.03660	0.01938	4.65808	0.4927
	recurn	0.00027	0.01648	0.01107	3.64201	0.4947
	regrsn	0.00407	0.06383	0.01060	3.15433	0.4927
Japan	atnn	0.00833	0.09126	0.00691	0.00691	0.5058
	backprop	0.00064	0.02524	0.01994	4.43512	0.5117
	boxjen	0.00307	0.05544	0.00489	1.85052	0.5005
	dblexpn	0.00020	0.01422	0.01004	2.51158	0.5048
	exponen	0.00551	0.07422	0.03233	8.32378	0.5047
	fuzzy	0.00247	0.04965	0.03288	7.84541	0.5047
	recurn	0.00021	0.01466	0.01020	3.09604	0.5117
	regrsn	0.00405	0.06362	0.00955	2.60456	0.5047

Table 35: 3 days ahead forecasts of exchange rates over window size 3

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00115	0.03391	0.00875	2.47993	0.4732
	backprop	0.00050	0.02225	0.01410	5.32932	0.4950
	boxjen	0.00239	0.04890	0.00472	1.14802	0.4805
	dblexpn	0.00025	0.01578	0.00782	2.10415	0.4938
	exponen	0.00508	0.07125	0.02856	9.57738	0.4937
	fuzzy	0.00085	0.02917	0.01443	4.51882	0.4937
	recurn	0.00034	0.01840	0.00988	3.71291	0.4950
	regrsn	0.00352	0.05934	0.00703	1.82297	0.4937
Britain	atnn	0.00013	0.01162	0.00448	1.36237	0.4712
	backprop	0.00037	0.01914	0.01139	4.60607	0.5027
	boxjen	0.00263	0.05133	0.00499	1.51093	0.4750
	dblexpn	0.00016	0.01273	0.00773	2.40316	0.4838
	exponen	0.00470	0.06857	0.02655	8.32361	0.4840
	fuzzy	0.00049	0.02223	0.01073	4.07156	0.4840
	recurn	0.00034	0.01846	0.01118	4.25047	0.5027
	regrsn	0.00372	0.06100	0.00695	2.12547	0.4840
Canada	atnn	0.00010	0.01014	0.00534	1.48122	0.4735
	backprop	0.00045	0.02121	0.01739	5.39288	0.4947
	boxjen	0.00257	0.05074	0.00581	1.67176	0.4779
	dblexpn	0.00019	0.01374	0.00931	2.64004	0.4925
	exponen	0.00539	0.07340	0.02832	9.40732	0.4927
	fuzzy	0.00123	0.03511	0.01851	4.54201	0.4927
	recurn	0.00027	0.01639	0.01083	3.62309	0.4947
	regrsn	0.00400	0.06321	0.00822	2.49081	0.4927
Japan	atnn	0.00010	0.00999	0.00529	1.30368	0.5055
	backprop	0.00062	0.02483	0.01711	3.48249	0.5117
	boxjen	0.00308	0.05550	0.00490	1.86863	0.5005
	dblexpn	0.00014	0.01189	0.00815	2.03484	0.5045
	exponen	0.00551	0.07424	0.03235	8.32723	0.5047
	fuzzy	0.00220	0.04688	0.03095	7.07335	0.5047
	recurn	0.00020	0.01422	0.00970	3.04101	0.5117
	regrsn	0.00399	0.06314	0.00726	2.03426	0.5047

Table 36: 1 day ahead forecasts of exchange rates over window size 4

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00020	0.01418	0.00438	1.16963	0.4728
	backprop	0.00104	0.03223	0.02207	6.36570	0.4953
	boxjen	0.00239	0.04887	0.00472	1.14843	0.4803
	dblexpn	0.00030	0.01741	0.00898	2.38901	0.4935
	exponen	0.00507	0.07119	0.02856	9.57462	0.4933
	fuzzy	0.00085	0.02920	0.01451	4.51498	0.4933
	recurn	0.00034	0.01832	0.00980	3.69041	0.4953
	regrsn	0.00356	0.05969	0.00843	2.17705	0.4933
Briatin	atnn	0.00014	0.01177	0.00449	1.36423	0.4715
	backprop	0.00108	0.03280	0.02487	9.74489	0.5027
	boxjen	0.00262	0.05116	0.00499	1.50480	0.4753
	dblexpn	0.00020	0.01411	0.00892	2.76604	0.4842
	exponen	0.00472	0.06869	0.02655	8.32836	0.4843
	fuzzy	0.00050	0.02228	0.01086	4.08542	0.4843
	recurn	0.00034	0.01836	0.01112	4.23468	0.5027
	regrsn	0.00377	0.06144	0.00836	2.56919	0.4843
Canada	atnn	0.00010	0.01020	0.00537	1.48714	0.4735
	backprop	0.00073	0.02710	0.01751	4.03373	0.4947
	boxjen	0.00256	0.05058	0.00580	1.66729	0.4782
	dblexpn	0.00024	0.01534	0.01059	2.98560	0.4928
	exponen	0.00539	0.07339	0.02832	9.39131	0.4927
	fuzzy	0.00124	0.03519	0.01862	4.57732	0.4927
	recurn	0.00027	0.01648	0.01093	3.64444	0.4947
	regrsn	0.00404	0.06360	0.00980	2.92544	0.4927
Japan	atnn	0.00008	0.00888	0.00491	1.21595	0.5055
	backprop	0.00063	0.02514	0.01891	4.07529	0.5117
	boxjen	0.00307	0.05544	0.00489	1.85052	0.5005
	dblexpn	0.00018	0.01324	0.00919	2.29996	0.5048
	exponen	0.00551	0.07422	0.03233	8.32378	0.5047
	fuzzy	0.00221	0.04697	0.03108	7.13383	0.5047
	recurn	0.00020	0.01421	0.00970	3.04717	0.5117
	regrsn	0.00402	0.06342	0.00865	2.37822	0.5047

Table 37: 2 days ahead forecasts of exchange rates over window size 4

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00020	0.01421	0.00439	1.17144	0.4725
	backprop	0.00145	0.03804	0.02692	7.67235	0.4957
	boxjen	0.00237	0.04869	0.00472	1.14600	0.4803
	dblexpn	0.00035	0.01878	0.00991	2.63934	0.4932
	exponen	0.00506	0.07116	0.02856	9.57318	0.4933
	fuzzy	0.00085	0.02923	0.01459	4.52035	0.4933
	recurn	0.00033	0.01828	0.00973	3.66846	0.4957
	regrsn	0.00360	0.06004	0.00949	2.47147	0.4933
Britain	atnn	0.00014	0.01195	0.00450	1.36417	0.4715
	backprop	0.00134	0.03662	0.02671	9.58364	0.5027
	boxjen	0.00262	0.05118	0.00499	1.50523	0.4753
	dblexpn	0.00024	0.01534	0.00989	3.06423	0.4845
	exponen	0.00471	0.06865	0.02655	8.32513	0.4843
	fuzzy	0.00050	0.02237	0.01094	4.09194	0.4843
	recurn	0.00033	0.01827	0.01105	4.21559	0.5027
	regrsn	0.00381	0.06170	0.00956	2.93604	0.4843
Canada	atnn	0.00023	0.01531	0.01013	2.71704	0.4732
	backprop	0.00175	0.04182	0.03535	10.52540	0.4947
	boxjen	0.00255	0.05052	0.00581	1.66828	0.4782
	dblexpn	0.00028	0.01667	0.01162	3.25367	0.4925
	exponen	0.00539	0.07340	0.02832	9.37172	0.4927
	fuzzy	0.00124	0.03523	0.01869	4.59349	0.4927
	recurn	0.00027	0.01649	0.01099	3.65142	0.4947
	regrsn	0.00409	0.06398	0.01103	3.26093	0.4927
Japan	atnn	0.00008	0.00883	0.00489	1.21237	0.5055
	backprop	0.00105	0.03240	0.02543	5.30260	0.5117
	boxjen	0.00307	0.05542	0.00490	1.84250	0.5005
	dblexpn	0.00021	0.01452	0.01017	2.55615	0.5048
	exponen	0.00552	0.07427	0.03230	8.33161	0.5047
	fuzzy	0.00222	0.04713	0.03120	7.20701	0.5047
	recurn	0.00020	0.01414	0.00966	3.04532	0.5117
	regrsn	0.00407	0.06380	0.00988	2.70636	0.5047

Table 38: 3 days ahead forecasts of exchange rates over window size 4

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00020	0.01422	0.00428	1.17986	0.4728
	backprop	0.00065	0.02551	0.01647	5.01517	0.4953
	boxjen	0.00239	0.04887	0.00472	1.14843	0.4803
	dblexpn	0.00020	0.01406	0.00661	1.76573	0.4935
	exponen	0.00507	0.07119	0.02856	9.57462	0.4933
	fuzzy	0.00092	0.03038	0.01591	5.58995	0.4933
	recurn	0.00030	0.01745	0.00901	3.45559	0.4953
	regrsn	0.00349	0.05907	0.00630	1.61768	0.4933
Britain	atnn	0.00018	0.01340	0.00452	1.45778	0.4715
	backprop	0.00037	0.01921	0.01100	4.51840	0.5027
	boxjen	0.00262	0.05116	0.00499	1.50480	0.4753
	dblexpn	0.00013	0.01129	0.00654	2.03006	0.4842
	exponen	0.00472	0.06869	0.02655	8.32836	0.4843
	fuzzy	0.00059	0.02424	0.01186	4.77154	0.4843
	recurn	0.00029	0.01699	0.01010	3.89456	0.5027
	regrsn	0.00372	0.06103	0.00630	1.93083	0.4843
Canada	atnn	0.00011	0.01053	0.00518	1.48292	0.4735
	backprop	0.00057	0.02382	0.01712	4.85236	0.4947
	boxjen	0.00256	0.05058	0.00580	1.66729	0.4782
	dblexpn	0.00014	0.01171	0.00780	2.20632	0.4928
	exponen	0.00539	0.07339	0.02832	9.39131	0.4927
	fuzzy	0.00119	0.03449	0.01819	4.59261	0.4927
	recurn	0.00021	0.01464	0.00950	3.18991	0.4947
	regrsn	0.00398	0.06305	0.00738	2.25170	0.4927
Japan	atnn	0.00025	0.01578	0.00490	1.21036	0.5055
	backprop	0.00043	0.02080	0.01482	3.27662	0.5117
	boxjen	0.00307	0.05544	0.00489	1.85052	0.5005
	dblexpn	0.00010	0.01020	0.00677	1.70812	0.5048
	exponen	0.00551	0.07422	0.03233	8.32378	0.5047
	fuzzy	0.00209	0.04567	0.03019	6.80225	0.5047
	recurn	0.00017	0.01302	0.00876	2.78407	0.5117
	regrsn	0.00397	0.06302	0.00650	1.85115	0.5047

Table 39: 1 day ahead forecasts of exchange rates over window size 5

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00022	0.01475	0.00483	1.35499	0.4725
	backprop	0.00110	0.03316	0.02228	6.15904	0.4957
	boxjen	0.00237	0.04869	0.00472	1.14600	0.4803
	dblexpn	0.00025	0.01578	0.00785	2.08388	0.4932
	exponen	0.00506	0.07116	0.02856	9.57318	0.4933
	fuzzy	0.00092	0.03041	0.01598	5.58640	0.4933
	recurn	0.00032	0.01777	0.00920	3.49215	0.4957
	regrsn	0.00353	0.05944	0.00775	2.00245	0.4933
Britain	atnn	0.00026	0.01610	0.00450	1.41999	0.4715
	backprop	0.00073	0.02708	0.01907	7.28880	0.5027
	boxjen	0.00262	0.05118	0.00499	1.50523	0.4753
	dblexpn	0.00016	0.01277	0.00782	2.42448	0.4845
	exponen	0.00471	0.06865	0.02655	8.32513	0.4843
	fuzzy	0.00059	0.02429	0.01199	4.79054	0.4843
	recurn	0.00030	0.01720	0.01032	3.94784	0.5027
	regrsn	0.00376	0.06129	0.00780	2.39091	0.4843
Canada	atnn	0.00010	0.00983	0.00508	1.43503	0.4735
	backprop	0.00092	0.03025	0.01891	4.04763	0.4947
	boxjen	0.00255	0.05052	0.00581	1.66828	0.4782
	dblexpn	0.00018	0.01347	0.00915	2.57265	0.4925
	exponen	0.00539	0.07340	0.02832	9.37172	0.4927
	fuzzy	0.00119	0.03455	0.01830	4.61916	0.4927
	recurn	0.00023	0.01511	0.00996	3.29638	0.4947
	regrsn	0.00403	0.06345	0.00902	2.70462	0.4927
Japan	atnn	0.00417	0.06457	0.00564	1.16874	0.5055
	backprop	0.00073	0.02696	0.02060	4.31716	0.5117
	boxjen	0.00307	0.05542	0.00490	1.84250	0.5005
	dblexpn	0.00014	0.01177	0.00797	2.00394	0.5048
	exponen	0.00552	0.07427	0.03230	8.33161	0.5047
	fuzzy	0.00213	0.04614	0.03005	6.75668	0.5047
	recurn	0.00018	0.01337	0.00908	2.86305	0.5117
	regrsn	0.00402	0.06339	0.00806	2.24530	0.5047

Table 40: 2 days ahead forecasts of exchange rates over window size 5

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00025	0.01590	0.00427	1.27183	0.4725
	backprop	0.00135	0.03672	0.02593	7.40298	0.4953
	boxjen	0.00238	0.04878	0.00472	1.14650	0.4803
	dblexpn	0.00029	0.01711	0.00878	2.35021	0.4935
	exponen	0.00507	0.07121	0.02856	9.57505	0.4937
	fuzzy	0.00092	0.03041	0.01605	5.58677	0.4937
	recurn	0.00032	0.01797	0.00941	3.53040	0.4953
	regrsn	0.00358	0.05984	0.00887	2.32526	0.4937
Britain	atnn	0.00016	0.01260	0.00435	1.35445	0.4715
	backprop	0.00150	0.03877	0.02876	10.53460	0.5023
	boxjen	0.00262	0.05123	0.00499	1.50721	0.4753
	dblexpn	0.00020	0.01414	0.00896	2.76796	0.4845
	exponen	0.00472	0.06871	0.02655	8.32781	0.4843
	fuzzy	0.00059	0.02438	0.01207	4.80033	0.4843
	recurn	0.00030	0.01736	0.01049	3.99381	0.5023
	regrsn	0.00380	0.06165	0.00910	2.79540	0.4843
Canada	atnn	0.00011	0.01055	0.00518	1.47537	0.4735
	backprop	0.00134	0.03655	0.02612	6.15847	0.4947
	boxjen	0.00255	0.05047	0.00581	1.66593	0.4784
	dblexpn	0.00022	0.01499	0.01034	2.90155	0.4928
	exponen	0.00538	0.07336	0.02832	9.35014	0.4927
	fuzzy	0.00120	0.03461	0.01836	4.63233	0.4927
	recurn	0.00024	0.01546	0.01029	3.37482	0.4947
	regrsn	0.00407	0.06378	0.01043	3.09846	0.4927
Japan	atnn	0.00048	0.02200	0.01324	3.08024	0.5055
	backprop	0.00089	0.02980	0.02363	4.99340	0.5113
	boxjen	0.00307	0.05544	0.00490	1.84835	0.5005
	dblexpn	0.00017	0.01323	0.00916	2.31359	0.5048
	exponen	0.00551	0.07425	0.03228	8.32975	0.5047
	fuzzy	0.00216	0.04652	0.03015	6.80946	0.5047
	recurn	0.00018	0.01360	0.00932	2.92334	0.5113
	regrsn	0.00406	0.06370	0.00940	2.59280	0.5047

Table 41: 3 days ahead forecasts of exchange rates over window size 5

6.1.1 Analysis of the Results

Before drawing any conclusions about the efficiencies of the models, it is better to understand what the numbers in each of the tables represent. Closely studying the magnitude of the errors, the following deductions have been made:

1. In terms of the metric MSE, the *atnn* method performs the best, in the sense that it gives minimum MSE error. It is closely followed by *backprop*, *dblexpn*, and *recurn*. The methods *regrsn*, *boxjen*, *fuzzy* produced relatively high MSE values. And *exponen* by far exceeds the magnitude of MSE values from all others. Even then, the overall MSE values rarely exceed 0.5%. This means that most of the times the differences between the forecasted values and the actual values is very less. It is to be noted that RMSE varies according as MSE. i.e., RMSE increases/decreases when MSE increases/decreases and vice versa.

2. For the metric MAE, we notice that both *atnn* and *boxjen* perform well, the former having lower values in most of the cases. Apart from these, *regrsn*, *dblexpn*, *fuzzy* have relatively lower values. *backprop* and *recurn* have close values to each other, but the magnitudes are on the higher side. The highest MAE is exhibited by *exponen*. The average MAE is around 2%.

3. MAPE has similar outcome as MAE. *atnn* and *boxjen* gives smaller values than the rest. But unlike the former metrics, here the range of values is quite high, from 1.14 to 9.57 in the same prediction zone. This metric also determines how well the predicted curve fits the actual one to a certain extent.

4. Finally, let us consider the DS metric. The tables 42 - 47 shows very different results from the previous set of tables (33 - 41). Unlike previous case, here the DIR errors have very close values to each other. In fact, it may be concluded that both *atnn*, *fuzzy*, *dblexpn* and *boxjen* give quite similar values 47% to roundabout 50% errors. i.e., 53% to 50% correctness in predicting direction of movement.

The methods *backprop* and *recurn* give higher errors than the rest and these too have very close values. However, the overall error is not more than 51%.

Also, the performance is slightly better for short range forecasts than long range for the forecasting methods discussed in this thesis. The prediction errors for Canada and Japan are higher than Australia and Britain in a general sense.

6.2 Prediction of Averages of Exchange Rates

In this section, the performances of the methods in forecasting future values of the averages of 3 and 5 days (window sizes 3,4,5) have been presented. The exchange rates of Australia, Britain, Canada and Japan with respect to the United States dollar have been used as time series data inputs to the experiments. The results are summarized in tables 42 - 47.

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00018	0.01353	0.00275	0.78277	0.4718
	backprop	0.00030	0.01730	0.00945	4.56127	0.4800
	boxjen	0.00235	0.04843	0.00294	0.68019	0.4766
	dblexpn	0.00041	0.02026	0.01065	2.92689	0.4718
	exponen	0.00505	0.07105	0.02839	9.53248	0.4720
	fuzzy	0.00082	0.02872	0.01291	3.54097	0.4720
	recurn	0.00031	0.01751	0.00919	3.51430	0.4800
	regrsn	0.00359	0.05990	0.00882	2.34411	0.4720
Britain	atnn	0.00010	0.00980	0.00268	0.83529	0.5065
	backprop	0.00030	0.01731	0.01301	4.72203	0.5063
	boxjen	0.00261	0.05107	0.00320	0.94555	0.5068
	dblexpn	0.00020	0.01430	0.00903	2.81095	0.5065
	exponen	0.00472	0.06870	0.02649	8.31447	0.5067
	fuzzy	0.00041	0.02022	0.00826	3.08358	0.5067
	recurn	0.00028	0.01685	0.01008	3.88978	0.5063
	regrsn	0.00375	0.06128	0.00708	2.18199	0.5067
Canada	atnn	0.00019	0.01396	0.00550	1.61218	0.4832
	backprop	0.00026	0.01597	0.01397	4.28454	0.5037
	boxjen	0.00253	0.05033	0.00345	0.95906	0.4797
	dblexpn	0.00025	0.01578	0.01078	3.05862	0.4832
	exponen	0.00536	0.07320	0.02809	9.38083	0.4833
	fuzzy	0.00131	0.03618	0.01732	3.95936	0.4833
	recurn	0.00022	0.01477	0.00937	3.24747	0.5037
	regrsn	0.00398	0.06309	0.00825	2.48536	0.4833
Japan	atnn	0.00007	0.00807	0.00296	0.74551	0.5008
	backprop	0.00051	0.02259	0.01457	2.82833	0.5140
	boxjen	0.00306	0.05529	0.00318	1.36175	0.4955
	dblexpn	0.00028	0.01678	0.01173	2.89722	0.5005
	exponen	0.00552	0.07430	0.03239	8.32797	0.5003
	fuzzy	0.00240	0.04895	0.03092	6.95412	0.5003
	recurn	0.00017	0.01315	0.00896	2.80381	0.5140
	regrsn	0.00407	0.06379	0.00967	2.62656	0.5003

Table 42: 3 days ahead forecasts of averages of exchange rates over window size 3

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00051	0.02269	0.00882	4.95732	0.4882
	backprop	0.00033	0.01828	0.00952	3.72249	0.4883
	boxjen	0.00235	0.04851	0.00339	0.80813	0.4976
	dblexpn	0.00026	0.01615	0.00788	2.14784	0.4882
	exponen	0.00503	0.07094	0.02827	9.50143	0.4883
	fuzzy	0.00080	0.02835	0.01255	3.47241	0.4883
	recurn	0.00030	0.01720	0.00898	3.46326	0.4883
	regrsn	0.00349	0.05907	0.00616	1.60257	0.4883
Britain	atnn	0.00078	0.02798	0.00313	1.01436	0.4865
	backprop	0.00020	0.01400	0.00797	3.51844	0.5103
	boxjen	0.00262	0.05123	0.00458	1.38292	0.4829
	dblexpn	0.00017	0.01314	0.00806	2.50639	0.4865
	exponen	0.00467	0.06837	0.02636	8.25786	0.4867
	fuzzy	0.00041	0.02017	0.00822	3.04082	0.4867
	recurn	0.00028	0.01668	0.00991	3.84059	0.5103
	regrsn	0.00371	0.06088	0.00626	1.91456	0.4867
Canada	atnn	0.00007	0.00806	0.00343	0.91019	0.4822
	backprop	0.00020	0.01419	0.00962	2.44134	0.5003
	boxjen	0.00253	0.05035	0.00422	1.18611	0.4861
	dblexpn	0.00021	0.01432	0.00971	2.76007	0.4822
	exponen	0.00535	0.07312	0.02792	9.31501	0.4823
	fuzzy	0.00130	0.03601	0.01728	3.95269	0.4823
	recurn	0.00021	0.01451	0.00920	3.19451	0.5003
	regrsn	0.00396	0.06296	0.00735	2.24189	0.4823
Japan	atnn	0.00007	0.00829	0.00238	0.60069	0.4915
	backprop	0.00020	0.01429	0.01100	2.68518	0.5197
	boxjen	0.00308	0.05548	0.00472	1.80038	0.4926
	dblexpn	0.00015	0.01238	0.00849	2.10563	0.4915
	exponen	0.00549	0.07410	0.03222	8.28219	0.4913
	fuzzy	0.00238	0.04883	0.03057	6.84215	0.4913
	recurn	0.00017	0.01298	0.00883	2.78017	0.5197
	regrsn	0.00397	0.06304	0.00655	1.85245	0.4913

Table 43: 5 days ahead forecasts of averages of exchange rates over window size 3

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00019	0.01376	0.00339	0.85054	0.4715
	backprop	0.00032	0.01785	0.01065	5.19780	0.4800
	boxjen	0.00235	0.04843	0.00294	0.68019	0.4766
	dblexpn	0.00041	0.02031	0.01089	3.00522	0.4715
	exponen	0.00504	0.07101	0.02839	9.53125	0.4717
	fuzzy	0.00084	0.02890	0.01387	4.38877	0.4717
	recurn	0.00031	0.01751	0.00911	3.56930	0.4800
	regrsn	0.00360	0.06004	0.00944	2.52852	0.4717
Britain	atnn	0.00007	0.02799	0.00415	1.00986	0.4865
	backprop	0.00037	0.01920	0.01510	5.41047	0.5063
	boxjen	0.00261	0.05107	0.00320	0.94555	0.5068
	dblexpn	0.00029	0.01717	0.01119	3.48272	0.5065
	exponen	0.00469	0.06847	0.02645	8.29043	0.5067
	fuzzy	0.00047	0.02174	0.00946	3.70942	0.5067
	recurn	0.00029	0.01714	0.01003	3.92180	0.5063
	regrsn	0.00380	0.06168	0.00971	2.98732	0.5067
Canada	atnn	0.00007	0.00814	0.00371	1.01980	0.4832
	backprop	0.00029	0.01717	0.01473	4.62887	0.5037
	boxjen	0.00253	0.05033	0.00345	0.95906	0.4797
	dblexpn	0.00036	0.01900	0.01330	3.76734	0.4832
	exponen	0.00536	0.07320	0.02809	9.38083	0.4833
	fuzzy	0.00121	0.03479	0.01693	3.98323	0.4833
	recurn	0.00022	0.01486	0.00926	3.26460	0.5037
	regrsn	0.00408	0.06389	0.01136	3.36056	0.4833
Japan	atnn	0.00010	0.01015	0.00579	1.63235	0.5012
	backprop	0.00042	0.02042	0.01684	3.60866	0.5140
	boxjen	0.00306	0.05529	0.00318	1.36175	0.4955
	dblexpn	0.00029	0.01696	0.01191	2.95245	0.5012
	exponen	0.00553	0.07439	0.03249	8.34474	0.5010
	fuzzy	0.00218	0.04667	0.02946	6.39976	0.5010
	recurn	0.00017	0.01296	0.00866	2.79701	0.5140
	regrsn	0.00409	0.06396	0.01030	2.78903	0.5010

Table 44: 3 days ahead forecasts of averages of exchange rates over window size 4

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00018	0.01335	0.00285	0.71251	0.4882
	backprop	0.00032	0.01788	0.00937	4.19723	0.4883
	boxjen	0.00235	0.04851	0.00339	0.80813	0.4976
	dblexpn	0.00031	0.01763	0.00909	2.48994	0.4882
	exponen	0.00503	0.07094	0.02827	9.50143	0.4883
	fuzzy	0.00082	0.02861	0.01356	4.31511	0.4883
	recurn	0.00030	0.01722	0.00891	3.51834	0.4883
	regrsn	0.00354	0.05946	0.00767	2.02263	0.4883
Britain	atnn	0.00010	0.00994	0.00279	0.81255	0.4865
	backprop	0.00023	0.01507	0.01034	3.97212	0.5103
	boxjen	0.00262	0.05123	0.00458	1.38292	0.4829
	dblexpn	0.00022	0.01470	0.00928	2.88428	0.4865
	exponen	0.00467	0.06837	0.02636	8.25786	0.4867
	fuzzy	0.00046	0.02155	0.00905	3.58511	0.4867
	recurn	0.00029	0.01699	0.00986	3.87270	0.5103
	regrsn	0.00374	0.06119	0.00782	2.39684	0.4867
Canada	atnn	0.00006	0.00760	0.00313	0.85459	0.4822
	backprop	0.00017	0.01289	0.01062	3.10447	0.5003
	boxjen	0.00253	0.05035	0.00422	1.18611	0.4861
	dblexpn	0.00026	0.01599	0.01105	3.12915	0.4822
	exponen	0.00535	0.07312	0.02792	9.31501	0.4823
	fuzzy	0.00119	0.03456	0.01650	3.85659	0.4823
	recurn	0.00021	0.01466	0.00913	3.22202	0.5003
	regrsn	0.00401	0.06334	0.00913	2.73634	0.4823
Japan	atnn	0.00005	0.00719	0.00322	0.76505	0.4902
	backprop	0.00069	0.02620	0.02242	4.63945	0.5197
	boxjen	0.00308	0.05548	0.00472	1.80038	0.4926
	dblexpn	0.00020	0.01421	0.00988	2.44901	0.4902
	exponen	0.00552	0.07429	0.03238	8.31682	0.4900
	fuzzy	0.00216	0.04647	0.02904	6.25286	0.4900
	recurn	0.00016	0.01281	0.00852	2.77257	0.5197
	regrsn	0.00403	0.06347	0.00828	2.29017	0.4900

Table 45: 5 days ahead forecasts of averages of exchange rates over window size 4

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00044	0.02097	0.00357	1.14064	0.4715
	backprop	0.00033	0.01813	0.01111	5.50543	0.4800
	boxjen	0.00235	0.04843	0.00294	0.68019	0.4766
	dblexpn	0.00039	0.01964	0.01053	2.93858	0.4715
	exponen	0.00504	0.07101	0.02839	9.53125	0.4717
	fuzzy	0.00091	0.03016	0.01538	5.43881	0.4717
	recurn	0.00029	0.01701	0.00863	3.37488	0.4800
	regrsn	0.00360	0.05997	0.00933	2.52314	0.4717
Britain	atnn	0.00027	0.01634	0.00335	1.03086	0.4832
	backprop	0.00032	0.01801	0.01388	4.97657	0.5063
	boxjen	0.00261	0.05107	0.00320	0.94555	0.5068
	dblexpn	0.00028	0.01685	0.01100	3.42627	0.5065
	exponen	0.00469	0.06847	0.02645	8.29043	0.5067
	fuzzy	0.00057	0.02378	0.01077	4.43784	0.5067
	recurn	0.00025	0.01591	0.00909	3.57094	0.5063
	regrsn	0.00380	0.06167	0.00973	2.99691	0.5067
Canada	atnn	0.00027	0.01634	0.00335	1.03086	0.4832
	backprop	0.00026	0.01619	0.01402	4.34835	0.5037
	boxjen	0.00253	0.05033	0.00345	0.95906	0.4797
	dblexpn	0.00034	0.01846	0.01301	3.68577	0.4832
	exponen	0.00536	0.07320	0.02809	9.38083	0.4833
	fuzzy	0.00116	0.03406	0.01665	4.03053	0.4833
	recurn	0.00018	0.01345	0.00817	2.86724	0.5037
	regrsn	0.00408	0.06386	0.01133	3.35549	0.4833
Japan	atnn	0.00008	0.00910	0.00339	0.82240	0.4998
	backprop	0.00156	0.03949	0.03374	6.51951	0.5140
	boxjen	0.00306	0.05529	0.00318	1.36175	0.4955
	dblexpn	0.00028	0.01674	0.01175	2.91964	0.4998
	exponen	0.00552	0.07428	0.03239	8.32320	0.4997
	fuzzy	0.00204	0.04517	0.02863	6.06158	0.4997
	recurn	0.00015	0.01205	0.00793	2.56100	0.5140
	regrsn	0.00409	0.06392	0.01035	2.79959	0.4997

Table 46: 3 days ahead forecasts of averages of exchange rates over window size 5

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
Australia	atnn	0.00075	0.02736	0.00519	1.47522	0.4882
	backprop	0.00029	0.01713	0.00927	4.42731	0.4883
	boxjen	0.00235	0.04851	0.00339	0.80813	0.4976
	dblexpn	0.00027	0.01652	0.00842	2.33999	0.4882
	exponen	0.00503	0.07094	0.02827	9.50143	0.4883
	fuzzy	0.00089	0.02985	0.01504	5.37472	0.4883
	recurn	0.00027	0.01644	0.00812	3.26851	0.4883
	regrsn	0.00352	0.05931	0.00723	1.92855	0.4883
Britain	atnn	0.00009	0.00970	0.00228	0.71099	0.4865
	backprop	0.00028	0.01678	0.01252	4.57775	0.5103
	boxjen	0.00262	0.05123	0.00458	1.38292	0.4829
	dblexpn	0.00020	0.01399	0.00880	2.73807	0.4865
	exponen	0.00467	0.06837	0.02636	8.25786	0.4867
	fuzzy	0.00056	0.02359	0.01023	4.29299	0.4867
	recurn	0.00022	0.01499	0.00883	3.50595	0.5103
	regrsn	0.00374	0.06112	0.00753	2.30870	0.4867
Canada	atnn	0.00006	0.00744	0.00248	0.74230	0.4822
	backprop	0.00023	0.01504	0.01299	4.01296	0.5003
	boxjen	0.00253	0.05035	0.00422	1.18611	0.4861
	dblexpn	0.00022	0.01494	0.01038	2.93881	0.4822
	exponen	0.00535	0.07312	0.02792	9.31501	0.4823
	fuzzy	0.00114	0.03382	0.01601	3.83645	0.4823
	recurn	0.00016	0.01282	0.00761	2.72583	0.5003
	regrsn	0.00400	0.06322	0.00870	2.61636	0.4823
Japan	atnn	0.00018	0.01332	0.00251	0.62206	0.4902
	backprop	0.00050	0.02236	0.01882	3.96439	0.5197
	boxjen	0.00308	0.05548	0.00472	1.80038	0.4926
	dblexpn	0.00018	0.01354	0.00939	2.33738	0.4902
	exponen	0.00552	0.07429	0.03238	8.31682	0.4900
	fuzzy	0.00203	0.04503	0.02832	5.94812	0.4900
	recurn	0.00013	0.01159	0.00752	2.48576	0.5197
	regrsn	0.00402	0.06341	0.00799	2.22659	0.4900

Table 47: 5 days ahead forecasts of averages of exchange rates over window size 5

6.2.1 Analysis of the Results

In this section analysis of the tables for predicting averages of exchange rates (tables 42 - 47) have been made followed by conclusions from the analysis.

As done in previous section, let us consider the performances of the forecasting methods in terms of the performance metrics.

1. In terms of the metric MSE, *atnn* performs marginally better than *recurn*, *backprop* and *regrsn*. This outcome is different than the one obtained for regular exchange rates. And as before, *exponen* gives the highest MSE errors. The overall magnitude of MSE errors is less than regular exchange data. The average is roundabout 0.02%. Like before, it is to be noted that RMSE varies according as MSE. i.e., RMSE increases/decreases when MSE increases/decreases and vice versa.

2. For the metric MAE, we notice that both *atnn* and *boxjen* perform well, the former having lower values in most of the cases. Apart from these, *regrsn* and *dblexpn* have relatively lower values. *backprop* and *recurn* have close values to each other, but the magnitudes are on the higher side. The highest MAE is exhibited by *exponen*. The MAE values for *fuzzy* is very close to those of *exponen*. The average MAE is around 0.2%.

3. MAPE has similar outcome as MAE. *atnn* and *boxjen* gives smaller values than the rest. But unlike the former metrics, here the range of values is quite high, from 0.70 to 9.57 in the same prediction zone. This metric also determines how well the predicted curve fits the actual one to a certain extent.

4. Finally, let us consider the DS metric. The tables 42-47 show that *atnn* has the minimum directional errors, i.e., the prediction of movements of the data is most accurate compared to other methods. The values lie between 0.4725 (for Australia - U.S. conversion rates) to 0.5035 (for Japanese - U.S. conversion

rates) for *atnn*. This means that the prediction of direction is correct 53% to atleast 50% times. *fuzzy*, *boxjen*, *exponen*, *regrsn*, *dblexpn* gives about 51% accuracy on an average. Whereas, *backprop* and *recurn* gives about 49% accuracy in predicting directions. One point interesting to note is that the performance of *backprop* and *recurn* are quite close to each other where this metric is concerned.

Investigating the above 4 points, we can say that *atnn* performs marginally better than *boxjen*, *dblexpn* or even *regrsn*. Also, *fuzzy* gives good prediction for averages compared to regular exchange data. The methods *backprop* and *recurn* have similar performance results. And *exponen*, as before, is not very suitable for predicting averages of exchange rates. But, the crucial point to be noted is that, there is not a vast difference between the magnitudes of the performance metrics in a general sense for each of the methods.

Finally, as a follow-up of the above analysis, it is found that the performance of neural nets is better than statistical models.

Chapter 7

Prediction of White Noise and Chaotic Time Series

The first part of this chapter contains a discussion on prediction of continuous-time-white noise which is a derivative of a Brownian motion. The predictions can be made by using the method of *fractional differencing*, which is an extension of the Box-Jenkins methodology described in Chapter 4. Here, the first level of differences takes any real values rather than integral differences as considered by the Box-Jenkins method. This is followed by a discussion on predicting chaotic time series.

7.1 The Fractional Differencing Technique

It is a standard technique among statisticians to use the method of differencing in order to achieve *stationarity* in a given time series. Once a time series is adapted to become stationary, it becomes easier to apply standard methods of prediction, like the Box-Jenkins method. (Please refer Chapter 4). Box-Jenkins method is typically known as the ARIMA modelling and the statisticians have interpreted the technique in many ways to suit the time series data. The method of fractional differencing proposed by Hosking [13], Granger et. al.,[10] and Mandelbrot et.al., [22] has been presented in this section.

Most of the recent work in time series assumed that observations separated by a long time span are nearly independent. Yet in many empirical time series,this is not the case,there does exist a dependency between the distant observations. Such series appear to exhibit cycles and changes of levels of all orders of magnitude,and their spectral densities increases indefinitely as the frequency tends to zero. The present section focuses on the following aspects and introduce a family of models which meet the same requirement:

1. Explicitly modelling long term persistence;
2. Being flexible enough to explain both the short term and the long term correlation structure of the series.
3. Enabling synthetic series to be easily generated from the model.

Generalizing the Box-Jenkins ARIMA(p, d, q) by permitting the difference operator d to take any real value, rather than integers, it turns out that for $0 < d <= \frac{1}{2}$ these fractionally differenced processes are capable of modelling long-term persistence.

7.1.1 Derivation of Fractionally Differenced White Noise

Brownian motion is a continuous time stochastic process B(t) with independent Gaussian increments and spectral density ω^{-2}.Its derivative is the continuous time white noise process, which has constant spectral density. Fractional Brownian motion, $B_H(t)$,defined by Mandelbrot and Van Ness [22] is a generalization of these processes. The basic properties of Fractional Brownian motion (f.b.m) are:

1. Fractional Brownian motion with parameter H, usually $0 < H < 1$ is the ($\frac{1}{2}$ -H)th fractional derivative of Brownian motion,the derivative is in Weyll or Reimann-Liouville senses.
2. The spectral density of f.b.m. is proportional to ω^{-2H-1}.

3. The covariance function is proportional to $|K|^{2H-2}$

The continuous time fractional noise is then defined by $B'_H(t)$,the derivative of $B_H(t)$. It is also thought of as the $(\frac{1}{2}-H)^{th}$ fractional derivative of the continuous time white noise, to which it reduces when H $=\frac{1}{2}$. Derivative exists only in the sense of a random Schwarz distribution. Def: $\Delta B_H(t) = B_H(t) - B_H(t-1)$

The discrete time analogue of Brownian Motion is the random walk or ARIMA$(0,1,0)$ process $\{x_t\}$, defined by: $\Delta x_t = (1-B)x_t = a_t$ where $Bx_t = x_{t-1}$ and a_t s are i.i.d. random variables. The first difference of $\{x_t\}$ is the white noise process $\{a_t\}$. We define fractional differenced white noise with parameter H to be the ($\frac{1}{2}$ -H)th fractional difference of discrete time white noise.

$$\Delta^d = (1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k = 1 - dB - \frac{1}{2!}d(1-d)B^{2-\frac{1}{3!}d(1-d)(2-d)B^3-\cdots}$$

$d = (H-\frac{1}{2})$, so continuous-time fractional noise with parameter H has as its discrete time analogue the process: $x_t = \Delta^{-d}a_t \Delta^d x_t = a_t$ where $\{a_t\}$ is a white noise process. $\{x_t\}$:ARIMA$(0,d,0)$process with non integral d.

7.1.2 TheARIMA$(0,d,0)$ Process

Defn:

$\Delta^d x_t = a_t$. where a_t are i.i.d with mean 0 and variance σ_a^2

Basic properties of the process with $\sigma_a^2 = 1$

Theorem 1 *Let $\{x_t\}$ be an ARIMA(p, d, q) process. (a) When $d < \frac{1}{2}$,$\{x_t\}$ is a stationary process and has infinite moving average representation:*

$$x_t = \Psi(B)a_t = \sum_{k=0}^{\infty} \Psi_k a_{t-k}$$

$$where \quad \Psi_k = \frac{d(1+d)\cdots(k-1+d)}{k!} \tag{7.1.1}$$

$$as \quad k \to \infty \quad \Psi_k \rightsquigarrow \frac{k^{d-1}}{(d-1)!}$$

(b) When $d > \frac{1}{2}$ *,* $\{x_t\}$ *is invertible and has the infinite autoregressive representation*

$$\Pi(B)x_t = \sum_{k=0}^{\infty} \Pi_k x_{t-k} = a_t$$

$$where \quad \Pi_k = \frac{-d(-1-d)\cdots(k-1-d)}{k!} = \frac{(k-d-1)!}{k!(d-1)!}$$

$$= \begin{pmatrix} k-d-1 \\ -d-1 \end{pmatrix}$$

$$as \quad k \to \infty, \quad \Pi_k \rightsquigarrow \frac{k^{-d-1}}{(-d-1)!}$$

When $\frac{-1}{2} < d < \frac{1}{2}$

(c) The spectral density of $\{x_t\}$ *is*

$$s(\omega) = (2\sin\frac{1}{2}\omega)^{-2d} \quad for \quad 0 < \omega \le \pi \quad and \quad s(\omega) \rightsquigarrow \omega^{-2d} \quad as \quad \omega \to 0$$

(d) The covariance function of $\{x_t\}$ *is*

$$\gamma_k = E(x_t \quad x_{t-k}) = \frac{(-1)^k(-2d)!}{(k-d)!(-k-d)!} \tag{7.1.2}$$

the correlation function of $\{x_t\}$ *is:*

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{(-d)!(k+d-1)!}{(d-1)!(k-d)!} \quad (k = 0, \pm 1 \cdots)$$

$$\rho_k = \frac{d(1+d)\cdots(k-1+d)}{(1-d)(2-d)\cdots(k-d)} \quad (k = 1, 2, \cdots)$$

$$In \quad particular \quad \gamma_0 = \frac{(-2d)!}{(-d!)^2} \quad and$$

$$\rho_1 = \frac{d}{1-d} \quad as \quad k \to \infty \quad \rho_k \rightsquigarrow \frac{(-d)!}{(d-1)!} k^{2d-1}$$

$$(7.1.3)$$

(e) The inverse autocorrelations of $\{x_t\}$ are:

$$\rho_{inv,k} = \frac{d!(k-d-1)!}{(-d-1)!(k+d)!} \rightsquigarrow \frac{d!}{(-d-1)!} k^{-1-2d} \quad as \quad k \to \infty$$

(f) The partial correlations of $\{x_t\}$ are:

$$\phi_{kk} = \frac{d}{k-d} \quad (k = 1, 2, \cdots) \qquad (7.1.4)$$

7.1.3 Conclusions from the Theorem

For $\frac{-1}{2} < d < \frac{1}{2}$ $\{x_t\}$ is both stationary and invertible. ψ_k *and* π_k decay hyperbolically. since

$$\lim_{k\to\infty} \psi_k = \lim_{k\to\infty} \frac{k^{d-1}}{(d-1)!} = \frac{1}{(d-1)!} \lim_{k\to\infty} k^{d-1}$$

$$(d-1) < 0 \quad for \quad d < \frac{1}{2}$$

$$similarly \ for \ \pi_k$$

and not exponential decay as in ARIMA$(p, 0, q)$

The behaviour of the spectrum at low frequencies indicate that for $d > 0, \{x_t\}$ is a long term persistant. process. This is also characterised by the hyperbolic decay of ρ_k in (7.1.3).

(7.1.3) also implies that $\{x_t\}$ is asymptotically self similar. The partial and inverse correlations of $\{x_t\}$ decay hyperbolically and at different rates. The partial linear coefficients ϕ_{kj} *for* $1 \ll j \ll k$ we have, as $j, k \to \infty$ *with* $j/k \to 0$.

$$\phi_{kj} \sim \frac{-j^{-d-1}}{(-d-1)!}$$

McLeod & Hipel(1978) defined a stationary process having a long or short memory depending whether the correlations have finite or infinite sums. Theorem 1 implies that for $0 < d < \frac{1}{2}$ the ARIMA$(0, d, 0)$ process is a long-memory stationary process.

An ARIMA$(0, d, 0)$ process where d is any real number, may be summed or differnced so that $d \in [\frac{-1}{2}, \frac{1}{2}]$ and will then be both invertible and stationary. And when $d = \pm\frac{1}{2}$ then the process is either stationary or invertible.

When $d = \frac{1}{2}$ the spectral density of the process is:

$$s(\omega) = 1/\{2\sin(\frac{1}{2}\omega)\} \sim \omega^{-1} \quad as \quad \omega \to 0$$

Thus the ARIMA$(0, \frac{1}{2}, 0)$ is the discrete-time 1/f noise.

ARIMA$(0, 0, 0)$ is white noise with zero correlations and constant spectral density.

When $\frac{-1}{2} < d < 0$ the ARIMA$(0, d, 0)$ has a short memory and is antipersistent in the terminology of Mandelbrot(1977). The correlations and the partial correlations are all negative, except $\rho_0 = 1$ and decay monotonically and hyperbolically to zero. $s(\omega)$ is and increasing function of ω, and vanishes at $\omega = 0$ but has ∞ gradient there.

The ARIMA$(0, \frac{-1}{2}, 0)$ is stationary but not invertible.Hence the forecasts cannot be expressed as a convergent sum of past values of the process. ψ_k for infinite moving average is the same as π_k for an ARIMA$(0, \frac{1}{2}, 0)$ and decays as $k^{-3/2}$ for large k. $s(\omega) = 2\sin\{\frac{1}{2}\omega\} \rightsquigarrow 0 \quad as \omega \to 0$. The gradient is 1 at $\omega = 0$. $\rho_k = -1/(4k^2 - 1)$; $\gamma_0 = 4/\pi$; $\psi_{kk} = -1/(2k+1)$

7.2 Predicting Chaotic Time Series using Fractal Theory

This section deals with an approach of modelling chaotic time series. One of the central problems in science is forecasting: to predict the future from the past knowledge. One of the ways is to build a model from the *first principles* and *initial data*. This is a classical approach for forecasting and the predictions are not very accurate because of the following reasons:

1. The first principles are not always available. This is particularly true for economic time series.

2. Sometimes the initial data are difficult to obtain. e.g. in fluid flow problems.

Hence the classical approach of modelling is not applicable to many physical systems. Therefore it is necessary to find alternative means of prediction. One of these alternative approaches is building an *ad hoc* linear model directly from the data. The modern theory views a time series $x(t_i)$ as a realization of a random process. (A process is a procedure which generates time series. Each application of the procedure yields a time series, called a realization of the process.) This is applicable when randomness is generated due to complicated motions involving many independent, irreducible degrees of freedom. Another cause is *chaos* which may occur in very simple deterministic system.

7.2.1 Properties of a Chaotic Process

The properties of a chaotic process are characterized by the following:

1. *Disorder or apparent disorder*: A chaotic process is ordered in the sense that it follows deterministic dynamics in the state phase. But when viewed in terms of time series, the process masquerades as disordered.

2. *Determinism*: If the initial conditions are precisely repeated, the system evolution over time is identical.

3. *Sensitivity to initial conditions*: Little changes in the initial conditions will render radical differences in the solution of the system.

4. *Random initial conditions*: The initial conditions are randomly chosen.

5. *Vanishing Correlation Function*: The solutions are truly disordered and diverge from each other...never to return. If the correlation function is zero for all nonzero lags, the process is uncorrelated or "white" chaos in analogy to "white noise".

6. *Aperiodicity*: Even though a system when placed in its chaotic regime by a parametric value might have a periodic solution for the dynamical equation. Such a system is not regarded as a chaotic system.

7. *Stationarity*: The statistical properties remain unchanged with time. Hence it is not possible to generate randomness due to initial conditions because the probability distribution remains invariant. Past determines future: they are manifestations of the same phenomena linked together by stationarity.

7.2.2 Forecasting a Chaotic Time Series

If the data are a single time series, the first step is to embed it in a state space. We introduce a state vector $x(t)$ by assigning co-ordinates:

$$x_1(t) = x(t) \quad x_2(t) = x(t-\tau), \quad \cdots \quad x_d(t) = x(t-(d-1)\tau) \tag{7.2.1}$$

where τ is the delay time. If the attractor is of dimension D, a minimal requirement is that $d \geq D$ Let the functional relationship between $x(t)$ and $x(t+\tau)$ be:

$$x(t+\tau) = f_T(x(t)) \tag{7.2.2}$$

We have to find a predictor F_T which approximates f_T. For a chaotic data f_T is nonlinear. f_T may be expressed as a m^{th} order polynomial in d dimension. The coefficients are fitted to the data set using the least squares technique. Forecasts for longer times $2T$, $3T$, etc. can be done by either composing F_T or by introducing new F_T s for every T. The first has its disadvantage that the error of approximation grows exponentially with each composition. The second technique works well for smooth functions. But the higher iterates of chaotic mappings are not smooth. So, both approaches have their disadvantages. Another approach is to recast (7.2.2) as a differential equation and write $x(t+\tau)$ as its integral. However, the one drawback common to all these approaches is the fact that the number of free parameters for a general polynomial is

$$\frac{(m+d)!}{m!d!} \approx d^m$$

which is intractable for large d. One of the ways to overcome this problem is to use "local approximation": using only nearby states to make predictions. To predict $x(t+\tau)$, we introduce a metric $||.||$ in the state space, and find the k nearest neighbors of $x(t)$ i.e., k states $x(t')$ with $t' < t$ which minimizes $||x(t)-x(t')||$. We can either use the zeroth order approximation(i.e. $k = 1$ and $x_{pred}(t,T) = x(t'+T)$) or use linear approximation(i.e. $k > d$ and fitting the linear polynomials to the pairs $(x(t'), x(t'+T))$). The range is taken to be scalar for computational convenience. So the mapping is from a $d-dimensional$ state to $1-dimensional$ value. When $k = d+1$ linear interpolation may be used. But to ensure stability of the solution it is advantageous to take $k > d+1$. The computation involves N steps for N data. This can be reduced to $logN$ by partitioning the data in a decision tree. Furthermore, once the neighbors are found, predictors for the multiples of T can be computed in parallel.

There are many other models used for predicting chaotic time series in addition to the ones presented in this chapter. The discussion of those models are beyond the scope of the current thesis work.

Chapter 8

Conclusions

As noted in Chapter 6, the method of Adaptively Trained Neural Nets (ATNN) performs the best with regards to prediction of exchange rates conversion of Australia, Britain, Canada and Japan versus the United States. This may be due to the reason that the ATNN algorithm has the flexibility of adapting itself with changing data and the training takes place on the fly. This method of training differs from backpropagation and recurrent nets methods. Here, the network is first trained with some training data possessing the exemplar pattern as the original data and then the trained net is used for prediction. The performance of the backpropagation and recurrent nets methods are dependent on how well the weights were trained and also on the learning rates. It has been found that backpropagation performs better with higher learning rates (close to 1.0) whereas recurrent nets require comparatively lower learning rates (close to 0.5). Fuzzy method also needs higher learning rates. For the experiments in this thesis, a learning rate of 0.8 has been chosen for training using fuzzy learning. Fuzzy method gave good prediction results when the data values did not vary to a great extent, in other words, it did not give a good prediction of higher and lower values. The training of fuzzy learning is also on the fly as the ATNN, but it is not as fast as the latter.

The performance of the statistical methods of prediction of exchange rates have been found to be surprisingly good. In fact, the error values of Box-Jenkins is as close at the ATNN for certain cases. And this method sometimes performed better than either backpropagation or recurrent or fuzzy training methods. The performance of double exponential smoothing is comparatively worse than the neural nets, yet on the whole the predicted values using this method were not too far off the actual ones. The regression method, as used in this thesis, gave good results also. The figures of performance metrics closely followed those of fuzzy learning. The exponential smoothing technique had the worst performance. This may be due to the reason that the time series data used for the experiments did not exhibit smoothness, in fact, there were a lot of fluctuations in the pattern. Exponential smoothing performs better if the current pattern follows the historical one. This was not the case in our experimental data. Even changing the smoothing constant did not affect the performance of prediction significantly. A smoothing constant of 0.02 gave the least sum of squares errors (a measure to select an appropriate smoothing constant) and hence was used for the experiments.

The forecasting methods discussed in this thesis gave reasonably good predictions for the exchange rates. But, as it has been observed that these methods are not very dependable in terms of predicting the direction of the time series data. This is due to the fact that the time series data chosen for experimentation was chaotic in nature within close intervals. But the range of values of the data points is not high, hence, it was possible to get lower values for the other performance metrics. As a general observation: the errors in predicting the averages of exchange rates have lower values than those in predicting the exact exchange rates. Also, prediction results for Australia, Britain, France are better than Canada, Germany, Japan and Switzerland.

The methods of predicting white noise using fractional differencing and chaotic time series has been presented in some details, though the implementation of the method on the computer is beyond the scope of the current work.

A total of 105 experiments were conducted for each method of prediction which made the total number of experiments to be 840. All these experiments were conducted using the C Programming language.

Appendix A

Some Notes on Chapter 4

A.1 Contribution of an Independent Variable in Linear Regression

There are a number of statistical tools available to determine the contribution of an independent variable in a regression model. One of the popular ones is the t_{bj}*-statistic.* It measures the importance of a particular independent variable x_{tj} in describing the dependent variable y_t in the multiple regression model

$$y_t = \mu_t + \varepsilon_t = \beta_0 + \beta_1 x_{t1} + \cdots + \beta_{j-1} x_{t,j-1} + \beta_j x_{tj} + \cdots + \beta_p x_{tp} + \varepsilon_t$$

The t_{bj}-statistic is defined by the equation

$$t_{bj} = \frac{b_j}{s_{bj}}$$

where b_j is the least squares estimate of β_j and s_{bj} is the *standard error of the estimate* b_j. More precisely, the t_{bj}-statistic measures the *additional importance* of the independent variable x_{tj} over and above the combined importance of the other independent variables. As a rule of thumb, the independent variable x_{tj} is said have a significant importance if

$$|t_{bj}| > t_{2.5}(n-(p-1))$$

where $t_{2.5}(n-(p-1)$ is the point on the scale of t-distribution having (n-(p-1)) degrees of freedom such that an area of 0.025 exists under the curve of this t-distribution between $t_{2.5}$ and ∞. Here, n is the number of observations made and p is the number of independent variables in the regression model. But one thing is to be noted at this point. The t_{bj}-statistic is not always a very reliable tool always, hence it is to be used with caution. It might be possible that an independent variable is found to be significantly important according to the t_{bj}-statistic where in reality it is not. This might lead to misleading results. This is particularly valid when there exists *multicollinearity* among the independent variables, i.e., they are related to dependent upon each other. As a result, even though each independent variable is contributing some information for the prediction of the dependent variable, some of the information is overlapping. Hence, in addition to the t_{bj}-statistic we need to measure the combined importance of all the independent variables taken together in describing the dependent variable. One of the simple measures of correlation is the *simple correlation coefficient* which is defined by:

$$R_{x_{ti},x_{tj}} = \frac{\sum_{t=1}^{n}(x_{ti}-\bar{x}_i)(x_{tj}-\bar{x}_j)}{[\sum_{t=1}^{n}(x_{ti}-\bar{x}_i)^2\sum_{t=1}^{n}(x_{tj}-\bar{x}_j)^2]^{\frac{1}{2}}}$$

where

$$\bar{x}_i = \frac{\sum_{t=1}^{n} x_{ti}}{n} \quad and \quad \bar{x}_j = \frac{\sum_{t=1}^{n} x_{tj}}{n}$$

The value of $R_{x_{ti},x_{tj}}$ lies between -1 and 1. The value of $R_{x_{ti},x_{tj}}$ nearer to 1(-1) implies that the independent variables x_{ti} and x_{tj} move together in a straight line fashion with a positive(negative) slope and are strongly correlated. Whereas when

$R_{x_{ti},x_{tj}}$ has a value near 0, it imples that the independent variables involved do not have a linear correlation.

A.2 Impacts of the Smoothing Constant in Prediction

We can see that:

$$S_{T-2} = \alpha y_{T-2} - (1-\alpha)S_{T-3}$$

which when used in 4.2.5 gives :

$$S_T \;=\; \alpha y_T - \alpha(1-\alpha)y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + (1-\alpha)^3 S_{T-3}$$

Hence by putting $t = T-3,\ \ T-4\ \ \cdots\ \ 0$ and substituting recursively for each S_t in 4.2.4 for each t we obtain :

$$\begin{aligned} S_T \;&=\; \alpha y_T - \alpha(1-\alpha)y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \\ \cdots + &\;\; \alpha(1-\alpha)^{T-1} y_{T-1} + (1-\alpha)^T S_0 \end{aligned} \tag{A.2.1}$$

Thus we see that S_T, the estimate of β_0 in time period T, can be expressed in terms of the observations $y_1,\ \ y_2,\ \ y_3,\ \ \cdots y_T$ and the initial estimate S_0. The coefficients of the observations namely $\alpha,\ \ \alpha(1-\alpha),\ \ \alpha(1-\alpha)^2,\ \ \cdots\ \ \alpha(1-\alpha)^T$ measure the contribution of the observations for the respective time periods to the most recent estimate S_T. We see that these coefficients decrease geometrically with the age of the observations.The updating procedure described above is called *simple exponential smoothing* because these coefficients decrease exponentially.

Since the coefficients are decreasing, the most recent observation y_T makes the most significant contribution to S_T, the current estimate of β_0. The older observations make smaller and smaller contribution to the current estimate of β_0 at each successive time point. The rate at which the remote observations are *dampened out* depends on the smoothing constant α. For values of α near zero, remote observations are dampened out more slowly, and the rate becomes faster as α approaches 1.

So the choice of α has a great influence on the estimate S_T. In general when the time series is quite volatile, i.e., when the random component ϵ_t has a large variance, then we choose α to be small so that the smoothed estimate S_T in 4.2.2 will weight the estimate of the previous time period $T-1$, namely, S_{T-1}, to a greater extent than the current observation y_T. For a more stable time series, we can choose larger α.

A.2.1 Determination of an Appropriate Smoothing Constant

The choice of an appropriate smoothing constant is very important for forecasting future values. A smoothing constant determines the extent to which a past observation influence the forecast. A smaller value of α dampens out remote observations in the time series slowly. Hence the response to the changes in the parameters describing the average level of the time series is slow. On the other hand, if the value of α is large, it will dampen out the remote observations quickly. Since a large α gives larger weight to the more recent observations in the time series, it results in a more rapid response to changes in the time series. Unfortunately, this rapid response can cause the forecasting procedure to respond to the irregular movements in the time series that do not reflect changes in the parameters that do not describe the time series. This is not a favourable situation and might lead to misleading results. In practice, it has been found that the values of α ranging from 0.01 to 0.03 works quite well. Another approach is *simulation*. This procedure involves simulating a set of historical data using different values of α. That is, for each value of α a set of forecast is generated using the appropriate exponential smoothing procedure. These forecasts are then compared with the actual observations in the time series. The value of α

which gives the best forecast is chosen as the smoothing constant for forecasting the future values of the time series.

It has to be noted that in an effort to determine the appropriate smoothing constant using simulation of historical data, there might be a possibility of using an incorrect exponential smoothing model for the time series being analyzed. This is probably the case if the smoothing constant is greater than 0.3. In that case, either the observations are dependent on one another, i.e., autocorrelated, or, there is a cyclical or seasonal behaviour in the time series. In both cases we have to resort to alternative forecasting techniques which handles such cases appropriately.

A.3 Some Models in Box-Jenkins Methodology

The Box-Jenkins methodology focusses in choosing a particular time series model from a clas of stationary time series models which is then used for forecasting the future values. In this section *moving-average models* and *autoregressive models* will be discussed.

A.3.1 Moving-Average Models

The model:

$$z_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

is called a *moving-average model of order q*. The two most widely used moving-average model is the *first-order moving average model*

$$z_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} \tag{A.3.1}$$

and the *second-order moving-average model*

$$z_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \tag{A.3.2}$$

In the case of first-order moving-average models described by equation A.3.1, no conditions are necessary to be imposed on θ_1 to make the model stationary. The invertibility condition is $|\theta_1| < 1$. The conditions of stationarity and invertibility are important in estimation of parameters of a model. Because these conditions ensure uniqueness of the estimates. For the first-order moving-average model, the relationship between ρ_k and θ_1 is given by :

$$\rho_k = \begin{cases} \frac{-\theta_1}{1+\theta_1^2} & \text{for } k = 1 \\ 0 & \text{for } k > 1 \end{cases} \tag{A.3.3}$$

This gives :

$$\rho_1 = \frac{-\theta_1}{1+\theta_1^2} \tag{A.3.4}$$

The equation A.3.4 can be used to calculate θ_1. The estimate of the mean μ is given by the average $\bar{z}$ as follows :

$$\bar{z} = \frac{\sum_{t=a}^{n} z_t}{n-a+1} \tag{A.3.5}$$

If $\bar{z}$ is small relative to the time series values $z_a, \cdots, z_n$ then it is reasonable to assume $\mu = 0$. If the original time series $y_1, y_2, \cdots, y_n$ has a mean μ equal to 0 it means that the time series is fluctuating around a mean 0. However, if the first differences have a mean 0, it implies that there is no deterministic trend in the *original* time series values.

Likewise, there is a deterministic trend in the original time series for a non-zero value of μ.

A second-order moving-average model characterized by equation A.3.2, and other higher order moving-average models, the invertibility conditions are even more complicated.

1. Stationarity Conditions : No conditions are necessary to be imposed on the parameters $\theta_1, \theta_2, \cdots \theta_q$ to make the model stationary.

2. Invertibility Conditions (for a second-order moving-average model) :

$$\theta_1 + \theta_2 < 1 \quad \theta_2 - \theta_1 < 1 \quad |\theta_2| < 1 \tag{A.3.6}$$

It can be shown that the mean of the above time series models is μ and the theoretical partial autocorrelation function dies down according to a mixture of damped exponentials and/or damped sine waves, and the theoretical autocorrelation function cuts off after lag 2. In particular it can be shown that

$$\rho_1 = \frac{-\theta_1(1-\theta_1)}{1+\theta_1^2+\theta_2^2}$$
$$\rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}$$
$$\rho_k = 0 \quad for \quad k > 2$$

Let r_1 and r_2 be the estimates of ρ_1 and ρ_2 and $\hat{\theta}_1$ and $\hat{\theta}_2$ be the estimates of θ_1 and θ_2 respectively, then the invertibility conditions are given by :

$$\hat{\theta}_1 + \hat{\theta}_2 < 1 \quad \hat{\theta}_2 - \hat{\theta}_1 < 1 \quad |\hat{\theta}_2| < 1$$

A reasonable preliminary estimate of μ will be given by the sample mean defined in equation A.3.5.

A.3.2 Autoregressive Models

The model:

$$z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + \varepsilon_t \tag{A.3.7}$$

is called an *autoregressive process of order p*. The term "autoregressive" is used because z_t, the current value of the time series is "regressed" or expressed as a function of $z_{t-1}, z_{t-2}, \cdots, z_{t-p}$ which are the previous values of the same time series.Certain conditions are necessary to be imposed on $\phi_1, \phi_2, \cdots \phi_p$ to make this model stationary and no conditions are needed for invertibility.

It can be shown that the theoretical partial autocorrelation function of this model dies down after lag p, and that the theoretical autocorrelation function dies down, which is just the opposite case for a moving-average model.

The first order autoregressive model is given by:

$$z_t = \delta + \phi_1 z_{t-1} + \varepsilon_t$$

1. Stationarity Conditions :

   $$|\phi_1| < 1$$

   The mean is given by:

   $$\mu = \frac{\delta}{1 - \phi_1}$$

   The theoretical partial autocorrelation function of this model cuts off after lag 1.

A reasonable estimate of δ is given by:

$$\hat{\delta} = \bar{z}(1 - \hat{\phi}_1)$$

where $\bar{z}$ is given by equation A.3.5.

The *second-order autoregressive model* can be described by:

$$z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \varepsilon_t$$

1. Stationarity conditions:

$$\phi_1 + \phi_2 < 1 \quad \phi_2 - \phi_1 < 1 \quad |\phi_2| < 1$$

The mean of the model is given by:

$$\mu = \frac{\delta}{1 - \phi_1 - \phi_2}$$

The theoretical partial autocorrelation function cuts off after lag 2 and the theoretical autocorrelation function dies down according to a mixture of damped exponentials and/or damped sine waves. The Yule-Walker equations namely:

$$\rho_1 = \phi_1 + \phi_2 \rho_1 \quad \textit{and} \quad \rho_2 = \phi_1 \rho_1 + \phi_2$$

are used to solve ϕ_1 and ϕ_2 in terms of ρ_1 and ρ_2, which are estimated by r_1 and r_2.Thus we get:

$$\hat{\phi}_1 = r_1 \left(\frac{1 - r_2}{1 - r_1^2} \right) \quad \hat{\phi}_2 = \left(\frac{r_2 - r_1^2}{1 - r_1^2} \right)$$

where $\hat{\phi_1}$ and $\hat{\phi_2}$ are the estimates of ϕ_1 and ϕ_2 respectively.

2. Stationarity conditions :

$$\hat{\phi}_1 + \hat{\phi}_2 < 1 \quad \hat{\phi}_2 - \hat{\phi}_1 < 1 \quad |\hat{\phi}_2| < 1$$

The mean is given by:

$$\mu = \frac{\delta}{1 - \phi_1}$$

which gives,

$$\delta = \mu(1 - \phi_1 - \phi_2)$$

and thus a reasonable estimate of δ is given by:

$$\hat{\delta} = \bar{z}$$

where,

$$\bar{z} = \frac{\sum_{t=a}^{n} z_t}{n - a + 1}$$

Appendix B

Some Notes on Chapter 5

B.1 The ATNN network

B.1.1 Linearization Process

Let

$$\mathbf{b} = \mathbf{W}^T\mathbf{x}(N+1)$$

and

$$\Delta\mathbf{b} = \Delta\mathbf{W}^T\mathbf{x}(N+1)$$

Applying first order Taylor's expansion to 5.5.6 gives:

$$\mathbf{u} = \mathbf{f}[\mathbf{b}+\Delta\mathbf{b}] = \mathbf{f}[\mathbf{b}] + (\nabla_{\mathbf{b}}\mathbf{f}[\mathbf{b}])\Delta\mathbf{b}$$

where $\nabla_{\mathbf{b}}\mathbf{f}[\mathbf{b}]$ is the gradient of $\mathbf{f}[\mathbf{b}]$ with respect to $\mathbf{b}$ with elements $\partial f_i[\mathbf{b}]/\partial b_j$. This approximation is valid when the *Hessian* $H \ll 1$ and $\Delta\mathbf{b} \ll f[\mathbf{b}]/H$.

where the inequality applies to each component of the vector. Now,

$$\partial f_i[\mathbf{b}]/\partial \mathbf{b_j} = \partial \mathbf{f}[\mathbf{b_i}]/\partial \mathbf{b_j} = \mathbf{f}[\mathbf{b_i}](1-\mathbf{f}[\mathbf{b_i}])\delta_{\mathbf{i-j}} = \mathbf{u}_\mathbf{i}^*(1-\mathbf{u}_\mathbf{i}^*)\delta_{\mathbf{i-j}}$$

then

$$\nabla_b u^* = diag[u_1^*(1-u_1^*), u_2^*(1-u_2^*), u_h^*(1-u_h^*)]$$

where δ_k, the *kronecker delta* is 1 for $k = 0$ and is zero otherwise and

$$u_i^* = f[b_i]$$

is the activation of the i^{th} hidden neuron for the new input data with the old weight such as

$$\mathbf{u}^* = [u_1^*, u_2^*, \cdots, u_h^*]^T = \mathbf{f}[\mathbf{W}^T(N)\mathbf{x}(N+1)].$$

B.1.2 Derivation of the Sensitivity Matrix

$$y(i) = f[sum_y] \quad u_k = f[sum_u]$$

where

$$sum_y = \sum_k u_k v_k \quad sum_u = \sum_j w_{jk} x_j$$

Thus

$$\begin{aligned}\frac{\partial y_i}{\partial v_k} &= \left(\frac{\partial y_i}{\partial sum_y}\right)\left(\frac{\partial sum_y}{\partial v_k}\right) = \frac{\partial f[x]}{\partial x}|_{x=y(i)} u_k \\ &= y(i)(1-y(i))u_k \\ &\equiv SV_{i,k} \end{aligned} \tag{B.1.1}$$

and

$$\frac{\partial y_i}{\partial w_{jk}} = \left(\frac{\partial y_i}{\partial u_k}\right)\left(\frac{\partial u_k}{\partial w_{jk}}\right)$$

$$
\begin{aligned}
&= \left(\frac{\partial y_i}{\partial sum_y} \right) \left(\frac{\partial sum_y}{\partial u_k} \right) \left(\frac{\partial u_k}{\partial sum_u} \right) \left(\frac{\partial sum_u}{\partial w_{jk}} \right) \\
&= \frac{\partial f[x]}{\partial x}|_{x=y(i)} \quad v_k \quad \frac{\partial f[x]}{\partial x}|_{x=u_k} \quad x_j \\
&= y(i)(1 - y(i)) \;\; v_k u_k \;\; (1 - u_k) x_j \\
&\equiv SW_{i,jk} \qquad \text{(B.1.2)}
\end{aligned}
$$

where $SV_{i,k}$ is the sensitivity caused by small changes in v_k and $SW_{i,jk}$ is the sensitivity of y(i) to the $w_{j,k}$'s.

$$
\Delta E_i = \sum_{j,k} \left(\frac{\partial E_i}{\partial w_{jk}} \right) \; \Delta \mathbf{w}_{jk} \; + \; \sum_k \left(\frac{\partial E_i}{\partial w_{jk}} \right) \Delta \mathbf{v}_k \qquad \text{(B.1.3)}
$$

By combining B.1.1, B.1.2, and B.1.3 with 5.5.16 and 5.5.17, we obtain

$$
\begin{aligned}
\Delta E_i = (\epsilon_i)[\Delta W_{vec,1} \cdots \Delta W_{vec,p} : \Delta v_1 \cdots \Delta v_h]. \qquad \text{(B.1.4)} \\
[SW_{i,1} \cdots SW_{i,p} \cdots : SV_{i,1} \cdots SV_{i,h}]^T
\end{aligned}
$$

where

$$
\epsilon_i = -(d(i) - y(i)), \quad 1 \le i \le N, \quad and \;\; p = I \times h.
$$

Appendix C

Some notes on Chapter 6

A list of tables stating (in details) the performance of all the methods described in this thesis in predicting exchange rates of France, Germany, and Switzerland with respect to the United States has been given below.

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00396	0.06292	0.01481	8.87341	0.4855
	backprop	0.00039	0.01970	0.01190	4.59060	0.5153
	boxjen	0.00287	0.05361	0.00449	1.55045	0.4895
	dblexpn	0.00013	0.01120	0.00671	2.13868	0.5005
	exponen	0.00491	0.07006	0.02711	9.00454	0.5007
	fuzzy	0.00058	0.02407	0.01235	4.24068	0.5007
	recurn	0.00019	0.01393	0.00922	3.58210	0.5153
	regrsn	0.00385	0.06208	0.00592	1.93015	0.5007
Germany	atnn	0.00009	0.00963	0.00424	1.20394	0.4885
	backprop	0.00016	0.01248	0.00873	2.62739	0.5093
	boxjen	0.00301	0.05488	0.00476	1.80827	0.4782
	dblexpn	0.00013	0.01123	0.00719	2.01502	0.4898
	exponen	0.00493	0.07019	0.02646	7.56779	0.4897
	fuzzy	0.00070	0.02639	0.01306	3.65136	0.4897
	recurn	0.00018	0.01328	0.00937	2.89034	0.5093
	regrsn	0.00399	0.06316	0.00624	1.95369	0.4897
Switzerland	atnn	0.00021	0.01441	0.00531	1.55766	0.5082
	backprop	0.00028	0.01679	0.01360	4.33588	0.5173
	boxjen	0.00298	0.05461	0.00538	2.00007	0.4950
	dblexpn	0.00015	0.01220	0.00836	2.43323	0.5068
	exponen	0.00518	0.07197	0.02964	8.70733	0.5070
	fuzzy	0.00069	0.02620	0.01285	3.76506	0.5070
	recurn	0.00023	0.01532	0.01111	3.55045	0.5173
	regrsn	0.00396	0.06296	0.00707	2.21385	0.5070

Table 48: 1 day ahead forecasts of exchange rates over window size 3

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00015	0.01228	0.00442	1.41839	0.5008
	backprop	0.00079	0.02809	0.01798	6.19398	0.5157
	boxjen	0.00288	0.05369	0.00449	1.55806	0.4895
	dblexpn	0.00015	0.01217	0.00753	2.38965	0.5008
	exponen	0.00490	0.06997	0.02712	8.99938	0.5010
	fuzzy	0.00058	0.02410	0.01251	4.29335	0.5010
	recurn	0.00020	0.01410	0.00933	3.60805	0.5157
	regrsn	0.00387	0.06217	0.00706	2.27613	0.5010
Germany	atnn	0.00024	0.01549	0.00448	1.27676	0.4885
	backprop	0.00025	0.01593	0.01195	3.47193	0.5093
	boxjen	0.00302	0.05494	0.00476	1.82242	0.4782
	dblexpn	0.00015	0.01240	0.00816	2.28606	0.4895
	exponen	0.00494	0.07029	0.02647	7.58676	0.4897
	fuzzy	0.00071	0.02659	0.01327	3.72639	0.4897
	recurn	0.00018	0.01345	0.00950	2.92458	0.5093
	regrsn	0.00403	0.06349	0.00759	2.33723	0.4897
Switzerland	atnn	0.00009	0.00939	0.00492	1.44208	0.5075
	backprop	0.00042	0.02057	0.01431	3.96080	0.5173
	boxjen	0.00299	0.05471	0.00538	2.01768	0.4953
	dblexpn	0.00019	0.01364	0.00954	2.79113	0.5072
	exponen	0.00520	0.07210	0.02964	8.72693	0.5073
	fuzzy	0.00069	0.02632	0.01302	3.81488	0.5073
	recurn	0.00024	0.01553	0.01132	3.61106	0.5173
	regrsn	0.00402	0.06343	0.00867	2.70071	0.5073

Table 49: 2 days ahead forecasts of exchange rates over window size 3

Country	Method	Performance Measures				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00125	0.03538	0.01474	4.07721	0.5012
	backprop	0.00117	0.03425	0.02382	8.78002	0.5160
	boxjen	0.00288	0.05362	0.00449	1.55269	0.4897
	dblexpn	0.00017	0.01310	0.00830	2.63610	0.5012
	exponen	0.00489	0.06996	0.02712	8.99927	0.5010
	fuzzy	0.00059	0.02419	0.01258	4.30218	0.5010
	recurn	0.00020	0.01425	0.00941	3.62851	0.5160
	regrsn	0.00389	0.06235	0.00804	2.58443	0.5010
Germany	atnn	0.00010	0.00975	0.00429	1.22306	0.4885
	backprop	0.00069	0.02633	0.02014	5.17859	0.5093
	boxjen	0.00301	0.05485	0.00476	1.80518	0.4782
	dblexpn	0.00018	0.01351	0.00902	2.53169	0.4895
	exponen	0.00495	0.07038	0.02648	7.60720	0.4893
	fuzzy	0.00072	0.02685	0.01346	3.79853	0.4893
	recurn	0.00018	0.01352	0.00960	2.95381	0.5093
	regrsn	0.00408	0.06384	0.00874	2.67718	0.4893
Switzerland	atnn	0.00011	0.01036	0.00513	1.49152	0.5092
	backprop	0.00049	0.02216	0.01834	5.60996	0.5177
	boxjen	0.00298	0.05459	0.00538	1.99976	0.4955
	dblexpn	0.00023	0.01507	0.01063	3.11297	0.5075
	exponen	0.00520	0.07208	0.02964	8.72718	0.5073
	fuzzy	0.00070	0.02652	0.01317	3.87023	0.5073
	recurn	0.00025	0.01570	0.01148	3.66041	0.5177
	regrsn	0.00407	0.06376	0.01008	3.11202	0.5073

Table 50: 3 days ahead forecasts of exchange rates over window size 3

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00012	0.01084	0.00402	1.23625	0.4858
	backprop	0.00054	0.02333	0.01406	4.74266	0.5157
	boxjen	0.00288	0.05369	0.00449	1.55806	0.4895
	dblexpn	0.00013	0.01118	0.00681	2.16213	0.5008
	exponen	0.00490	0.06997	0.02712	8.99938	0.5010
	fuzzy	0.00060	0.02445	0.01296	4.96287	0.5010
	recurn	0.00020	0.01413	0.00929	3.65532	0.5157
	regrsn	0.00385	0.06203	0.00626	2.02691	0.5010
Germany	atnn	0.00007	0.00844	0.00442	1.23239	0.4885
	backprop	0.00014	0.01195	0.00877	2.71973	0.5093
	boxjen	0.00302	0.05494	0.00476	1.82242	0.4782
	dblexpn	0.00013	0.01140	0.00741	2.07949	0.4895
	exponen	0.00494	0.07029	0.02647	7.58676	0.4897
	fuzzy	0.00056	0.02366	0.01196	3.19714	0.4897
	recurn	0.00018	0.01354	0.00957	2.97581	0.5093
	regrsn	0.00401	0.06334	0.00667	2.09028	0.4897
Switzerland	atnn	0.00009	0.00935	0.00512	1.48290	0.5075
	backprop	0.00022	0.01475	0.01137	3.66554	0.5173
	boxjen	0.00299	0.05471	0.00538	2.01768	0.4953
	dblexpn	0.00016	0.01262	0.00869	2.54258	0.5072
	exponen	0.00520	0.07210	0.02964	8.72693	0.5073
	fuzzy	0.00053	0.02293	0.01172	3.24518	0.5073
	recurn	0.00025	0.01567	0.01136	3.66276	0.5173
	regrsn	0.00400	0.06323	0.00765	2.40993	0.5073

Table 51: 1 day ahead forecasts of exchange rates over window size 4

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00009	0.00938	0.00409	1.27157	0.5012
	backprop	0.00115	0.03385	0.02245	7.14418	0.5160
	boxjen	0.00288	0.05362	0.00449	1.55269	0.4897
	dblexpn	0.00015	0.01219	0.00762	2.41528	0.5012
	exponen	0.00489	0.06996	0.02712	8.99927	0.5010
	fuzzy	0.00060	0.02451	0.01308	4.98641	0.5010
	recurn	0.00020	0.01411	0.00923	3.63751	0.5160
	regrsn	0.00387	0.06220	0.00736	2.36786	0.5010
Germany	atnn	0.00007	0.00842	0.00439	1.22729	0.4885
	backprop	0.00026	0.01619	0.01221	3.52174	0.5093
	boxjen	0.00301	0.05485	0.00476	1.80518	0.4782
	dblexpn	0.00016	0.01251	0.00834	2.33967	0.4895
	exponen	0.00495	0.07038	0.02648	7.60720	0.4893
	fuzzy	0.00056	0.02377	0.01208	3.23899	0.4893
	recurn	0.00018	0.01339	0.00948	2.95640	0.5093
	regrsn	0.00405	0.06367	0.00791	2.44961	0.4893
Switzerland	atnn	0.00009	0.00930	0.00513	1.48467	0.5078
	backprop	0.00043	0.02071	0.01700	5.26242	0.5177
	boxjen	0.00298	0.05459	0.00538	1.99976	0.4955
	dblexpn	0.00020	0.01400	0.00983	2.88496	0.5075
	exponen	0.00520	0.07208	0.02964	8.72718	0.5073
	fuzzy	0.00053	0.02305	0.01192	3.30817	0.5073
	recurn	0.00024	0.01557	0.01132	3.65967	0.5177
	regrsn	0.00404	0.06353	0.00912	2.84087	0.5073

Table 52: 2 days ahead forecasts of exchange rates over window size 4

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00009	0.00946	0.00410	1.27295	0.5012
	backprop	0.00155	0.03932	0.02658	8.68075	0.5157
	boxjen	0.00288	0.05365	0.00449	1.55420	0.4895
	dblexpn	0.00017	0.01313	0.00840	2.66284	0.5008
	exponen	0.00491	0.07006	0.02713	9.01130	0.5007
	fuzzy	0.00060	0.02453	0.01314	4.99064	0.5007
	recurn	0.00020	0.01402	0.00918	3.62201	0.5157
	regrsn	0.00391	0.06251	0.00833	2.68148	0.5007
Germany	atnn	0.00007	0.00831	0.00418	1.18827	0.4888
	backprop	0.00052	0.02288	0.01745	4.58203	0.5090
	boxjen	0.00301	0.05491	0.00476	1.81376	0.4782
	dblexpn	0.00019	0.01369	0.00920	2.57564	0.4895
	exponen	0.00496	0.07040	0.02649	7.61726	0.4897
	fuzzy	0.00057	0.02397	0.01219	3.27653	0.4897
	recurn	0.00018	0.01324	0.00938	2.93185	0.5090
	regrsn	0.00409	0.06394	0.00905	2.76658	0.4897
Switzerland	atnn	0.00025	0.01582	0.00519	1.50975	0.5082
	backprop	0.00052	0.02280	0.01895	5.83149	0.5177
	boxjen	0.00298	0.05461	0.00538	2.00271	0.4955
	dblexpn	0.00024	0.01549	0.01093	3.20486	0.5075
	exponen	0.00518	0.07196	0.02963	8.71268	0.5077
	fuzzy	0.00054	0.02326	0.01210	3.36954	0.5077
	recurn	0.00024	0.01553	0.01127	3.64629	0.5177
	regrsn	0.00406	0.06375	0.01051	3.22563	0.5077

Table 53: 3 days ahead forecasts of exchange rates over window size 4

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00012	0.01102	0.00388	1.24106	0.4858
	backprop	0.00079	0.02803	0.01600	4.36051	0.5160
	boxjen	0.00288	0.05362	0.00449	1.55269	0.4897
	dblexpn	0.00010	0.00975	0.00564	1.78972	0.5012
	exponen	0.00489	0.06996	0.02712	8.99927	0.5010
	fuzzy	0.00071	0.02665	0.01414	5.88701	0.5010
	recurn	0.00018	0.01326	0.00846	3.37482	0.5160
	regrsn	0.00383	0.06190	0.00562	1.81886	0.5010
Germany	atnn	0.00027	0.01649	0.00505	1.44618	0.4885
	backprop	0.00015	0.01206	0.00880	2.72424	0.5093
	boxjen	0.00301	0.05485	0.00476	1.80518	0.4782
	dblexpn	0.00009	0.00969	0.00613	1.72730	0.4895
	exponen	0.00495	0.07038	0.02648	7.60720	0.4893
	fuzzy	0.00048	0.02199	0.01120	2.87121	0.4893
	recurn	0.00015	0.01233	0.00857	2.68529	0.5093
	regrsn	0.00401	0.06332	0.00600	1.91879	0.4893
Switzerland	atnn	0.00010	0.01010	0.00495	1.45123	0.5078
	backprop	0.00021	0.01435	0.01086	3.48037	0.5177
	boxjen	0.00298	0.05459	0.00538	1.99976	0.4955
	dblexpn	0.00012	0.01077	0.00725	2.13305	0.5075
	exponen	0.00520	0.07208	0.02964	8.72718	0.5073
	fuzzy	0.00043	0.02078	0.01095	2.86387	0.5073
	recurn	0.00020	0.01419	0.01015	3.29750	0.5177
	regrsn	0.00398	0.06308	0.00687	2.18530	0.5073

Table 54: 1 day ahead forecasts of exchange rates over window size 5

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00014	0.01172	0.00501	1.55434	0.5012
	backprop	0.00144	0.03795	0.02453	7.12387	0.5157
	boxjen	0.00288	0.05365	0.00449	1.55420	0.4895
	dblexpn	0.00012	0.01092	0.00663	2.10775	0.5008
	exponen	0.00491	0.07006	0.02713	9.01130	0.5007
	fuzzy	0.00071	0.02665	0.01425	5.90722	0.5007
	recurn	0.00018	0.01343	0.00863	3.41565	0.5157
	regrsn	0.00387	0.06221	0.00689	2.22470	0.5007
Germany	atnn	0.00008	0.00917	0.00422	1.20161	0.4888
	backprop	0.00033	0.01825	0.01374	3.79916	0.5090
	boxjen	0.00301	0.05491	0.00476	1.81376	0.4782
	dblexpn	0.00012	0.01109	0.00724	2.02966	0.4895
	exponen	0.00496	0.07040	0.02649	7.61726	0.4897
	fuzzy	0.00049	0.02209	0.01135	2.91661	0.4897
	recurn	0.00016	0.01253	0.00876	2.73485	0.5090
	regrsn	0.00404	0.06358	0.00742	2.31013	0.4897
Switzerland	atnn	0.00010	0.01007	0.00494	1.44963	0.5082
	backprop	0.00049	0.02212	0.01832	5.60067	0.5177
	boxjen	0.00298	0.05461	0.00538	2.00271	0.4955
	dblexpn	0.00016	0.01252	0.00861	2.52892	0.5075
	exponen	0.00518	0.07196	0.02963	8.71268	0.5077
	fuzzy	0.00044	0.02097	0.01112	2.90930	0.5077
	recurn	0.00021	0.01459	0.01047	3.38643	0.5177
	regrsn	0.00400	0.06328	0.00855	2.65413	0.5077

Table 55: 2 days ahead forecasts of exchange rates over window size 5

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00016	0.01257	0.00422	1.36830	0.5008
	backprop	0.00240	0.04898	0.03120	8.75925	0.5157
	boxjen	0.00288	0.05367	0.00449	1.55696	0.4892
	dblexpn	0.00014	0.01194	0.00762	2.41161	0.5005
	exponen	0.00491	0.07005	0.02713	9.01336	0.5007
	fuzzy	0.00071	0.02665	0.01432	5.91800	0.5007
	recurn	0.00018	0.01351	0.00876	3.44685	0.5157
	regrsn	0.00390	0.06241	0.00801	2.57662	0.5007
Germany	atnn	0.00029	0.01705	0.00463	1.33308	0.4892
	backprop	0.00102	0.03199	0.02565	6.54115	0.5087
	boxjen	0.00301	0.05491	0.00474	1.81267	0.4784
	dblexpn	0.00015	0.01238	0.00832	2.32319	0.4898
	exponen	0.00497	0.07050	0.02651	7.64026	0.4897
	fuzzy	0.00050	0.02228	0.01147	2.95912	0.4897
	recurn	0.00016	0.01263	0.00890	2.77308	0.5087
	regrsn	0.00409	0.06393	0.00871	2.68442	0.4897
Switzerland	atnn	0.00012	0.01114	0.00505	1.48060	0.5078
	backprop	0.00055	0.02341	0.01945	5.90427	0.5180
	boxjen	0.00299	0.05471	0.00538	2.01855	0.4955
	dblexpn	0.00020	0.01422	0.00994	2.91257	0.5075
	exponen	0.00518	0.07195	0.02964	8.71158	0.5073
	fuzzy	0.00045	0.02120	0.01128	2.96116	0.5073
	recurn	0.00022	0.01491	0.01068	3.44908	0.5180
	regrsn	0.00405	0.06361	0.01011	3.10161	0.5073

Table 56: 3 days ahead forecasts of exchange rates over window size 5

Country	Method	Performance Measures				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00035	0.01865	0.00387	1.12995	0.4808
	backprop	0.00021	0.01438	0.01081	4.29085	0.5117
	boxjen	0.00286	0.05351	0.00274	0.98637	0.4900
	dblexpn	0.00016	0.01261	0.00799	2.52130	0.4808
	exponen	0.00490	0.07002	0.02696	8.96108	0.4807
	fuzzy	0.00055	0.02349	0.01092	3.85204	0.4807
	recurn	0.00017	0.01320	0.00866	3.43716	0.5117
	regrsn	0.00387	0.06221	0.00641	2.07348	0.4807
Germany	atnn	0.00005	0.00688	0.00288	0.80511	0.5075
	backprop	0.00008	0.00920	0.00575	1.75202	0.5060
	boxjen	0.00299	0.05472	0.00289	1.20538	0.5061
	dblexpn	0.00017	0.01304	0.00877	2.44051	0.5075
	exponen	0.00492	0.07017	0.02631	7.53349	0.5073
	fuzzy	0.00065	0.02554	0.01066	2.75275	0.5073
	recurn	0.00015	0.01236	0.00874	2.72614	0.5060
	regrsn	0.00401	0.06336	0.00694	2.15152	0.5073
Switzerland	atnn	0.00010	0.01024	0.00320	0.93327	0.5145
	backprop	0.00027	0.01643	0.01395	4.11947	0.5227
	boxjen	0.00296	0.05440	0.00336	1.32163	0.5208
	dblexpn	0.00022	0.01469	0.01024	2.96980	0.5145
	exponen	0.00515	0.07179	0.02950	8.64364	0.5147
	fuzzy	0.00064	0.02532	0.01007	2.82046	0.5147
	recurn	0.00021	0.01432	0.01028	3.31803	0.5227
	regrsn	0.00398	0.06306	0.00790	2.43777	0.5147

Table 57: 3 days average forecasts of exchange rates over window size 3

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00007	0.00823	0.00195	0.64035	0.4858
	backprop	0.00022	0.01482	0.01002	4.89227	0.5120
	boxjen	0.00286	0.05352	0.00323	1.13733	0.4929
	dblexpn	0.00014	0.01163	0.00715	2.25964	0.4858
	exponen	0.00489	0.06990	0.02690	8.93564	0.4860
	fuzzy	0.00054	0.02322	0.01084	3.79356	0.4860
	recurn	0.00017	0.01302	0.00851	3.39473	0.5120
	regrsn	0.00385	0.06204	0.00570	1.84917	0.4860
Germany	atnn	0.00005	0.00725	0.00230	0.65639	0.5035
	backprop	0.00034	0.01854	0.01430	5.14011	0.5010
	boxjen	0.00300	0.05476	0.00356	1.42232	0.4905
	dblexpn	0.00014	0.01187	0.00778	2.16993	0.5035
	exponen	0.00490	0.07002	0.02620	7.48069	0.5033
	fuzzy	0.00064	0.02536	0.01038	2.63431	0.5033
	recurn	0.00015	0.01213	0.00860	2.68972	0.5010
	regrsn	0.00399	0.06315	0.00609	1.90925	0.5033
Switzerland	atnn	0.00010	0.00981	0.00253	0.74976	0.4962
	backprop	0.00044	0.02094	0.01884	5.46716	0.5030
	boxjen	0.00298	0.05455	0.00471	1.76606	0.4889
	dblexpn	0.00020	0.01413	0.00977	2.83512	0.4962
	exponen	0.00515	0.07178	0.02935	8.61415	0.4960
	fuzzy	0.00063	0.02502	0.01003	2.77308	0.4960
	recurn	0.00020	0.01406	0.01010	3.26866	0.5030
	regrsn	0.00398	0.06311	0.00755	2.35098	0.4960

Table 58: 5 days average forecasts of exchange rates over window size 3

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00007	0.00821	0.00287	0.85211	0.4808
	backprop	0.00023	0.01510	0.01230	5.00503	0.5117
	boxjen	0.00286	0.05351	0.00274	0.98637	0.4900
	dblexpn	0.00022	0.01484	0.00983	3.09663	0.4808
	exponen	0.00490	0.07002	0.02696	8.96108	0.4807
	fuzzy	0.00058	0.02411	0.01179	4.61674	0.4807
	recurn	0.00018	0.01329	0.00864	3.50031	0.5117
	regrsn	0.00393	0.06267	0.00868	2.78745	0.4807
Germany	atnn	0.00005	0.00735	0.00321	0.87681	0.5075
	backprop	0.00008	0.00878	0.00603	1.83180	0.5060
	boxjen	0.00299	0.05472	0.00289	1.20538	0.5061
	dblexpn	0.00025	0.01574	0.01092	3.03679	0.5075
	exponen	0.00491	0.07006	0.02627	7.49485	0.5077
	fuzzy	0.00054	0.02315	0.01007	2.49979	0.5077
	recurn	0.00016	0.01253	0.00886	2.79731	0.5060
	regrsn	0.00407	0.06380	0.00954	2.85554	0.5077
Switzerland	atnn	0.00008	0.00875	0.00453	1.31255	0.5145
	backprop	0.00020	0.01419	0.01182	3.59338	0.5227
	boxjen	0.00296	0.05440	0.00336	1.32163	0.5208
	dblexpn	0.00032	0.01801	0.01288	3.73320	0.5145
	exponen	0.00515	0.07179	0.02950	8.64364	0.5147
	fuzzy	0.00048	0.02187	0.00925	2.38776	0.5147
	recurn	0.00021	0.01455	0.01039	3.38883	0.5227
	regrsn	0.00407	0.06381	0.01103	3.34480	0.5147

Table 59: 3 days average forecasts of exchange rates over window size 4

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00006	0.00792	0.00242	0.71192	0.4862
	backprop	0.00020	0.01411	0.01032	4.77671	0.5120
	boxjen	0.00286	0.05352	0.00323	1.13733	0.4929
	dblexpn	0.00016	0.01268	0.00817	2.57507	0.4862
	exponen	0.00489	0.06994	0.02689	8.93676	0.4860
	fuzzy	0.00057	0.02393	0.01149	4.52965	0.4860
	recurn	0.00017	0.01312	0.00846	3.44739	0.5120
	regrsn	0.00388	0.06229	0.00705	2.27006	0.4860
Germany	atnn	0.00005	0.00692	0.00278	0.75115	0.5035
	backprop	0.00009	0.00974	0.00460	1.31736	0.5010
	boxjen	0.00300	0.05476	0.00356	1.42232	0.4905
	dblexpn	0.00018	0.01325	0.00906	2.52126	0.5035
	exponen	0.00490	0.07002	0.02620	7.48069	0.5033
	fuzzy	0.00052	0.02287	0.00947	2.29753	0.5033
	recurn	0.00015	0.01232	0.00867	2.75039	0.5010
	regrsn	0.00402	0.06341	0.00769	2.34901	0.5033
Switzerland	atnn	0.00011	0.01034	0.00261	0.76391	0.4962
	backprop	0.00020	0.01420	0.01108	3.72766	0.5030
	boxjen	0.00298	0.05455	0.00471	1.76606	0.4889
	dblexpn	0.00023	0.01509	0.01065	3.09240	0.4962
	exponen	0.00514	0.07171	0.02935	8.59970	0.4963
	fuzzy	0.00047	0.02158	0.00866	2.19000	0.4963
	recurn	0.00020	0.01430	0.01023	3.34723	0.5030
	regrsn	0.00400	0.06327	0.00882	2.71241	0.4963

Table 60: 5 days average forecasts of exchange rates over window size 4

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00006	0.00783	0.00232	0.74423	0.4808
	backprop	0.00029	0.01693	0.01426	5.70611	0.5117
	boxjen	0.00286	0.05351	0.00274	0.98637	0.4900
	dblexpn	0.00021	0.01446	0.00963	3.03998	0.4808
	exponen	0.00490	0.07002	0.02696	8.96108	0.4807
	fuzzy	0.00069	0.02627	0.01309	5.52632	0.4807
	recurn	0.00016	0.01259	0.00799	3.25012	0.5117
	regrsn	0.00393	0.06265	0.00868	2.79108	0.4807
Germany	atnn	0.00007	0.00845	0.00265	0.76005	0.5075
	backprop	0.00007	0.00846	0.00568	1.75125	0.5060
	boxjen	0.00299	0.05472	0.00289	1.20538	0.5061
	dblexpn	0.00024	0.01536	0.01072	2.98973	0.5075
	exponen	0.00491	0.07006	0.02627	7.49485	0.5077
	fuzzy	0.00044	0.02109	0.00935	2.17171	0.5077
	recurn	0.00014	0.01163	0.00799	2.52047	0.5060
	regrsn	0.00407	0.06379	0.00955	2.86534	0.5077
Switzerland	atnn	0.00007	0.00849	0.00302	0.89105	0.5145
	backprop	0.00029	0.01716	0.01410	3.94991	0.5227
	boxjen	0.00296	0.05440	0.00336	1.32163	0.5208
	dblexpn	0.00031	0.01772	0.01273	3.69051	0.5145
	exponen	0.00515	0.07179	0.02950	8.64364	0.5147
	fuzzy	0.00038	0.01946	0.00878	2.10187	0.5147
	recurn	0.00018	0.01333	0.00930	3.03518	0.5227
	regrsn	0.00407	0.06382	0.01111	3.36901	0.5147

Table 61: 3 days average forecasts of exchange rates over window size 5

Country	*Method*	*Performance Measures*				
		MSE	RMSE	MAE	MAPE	Ave DIR
France	atnn	0.00008	0.00907	0.00287	0.88008	0.4862
	backprop	0.00021	0.01445	0.01150	4.87873	0.5120
	boxjen	0.00286	0.05352	0.00323	1.13733	0.4929
	dblexpn	0.00014	0.01194	0.00769	2.42885	0.4862
	exponen	0.00489	0.06994	0.02689	8.93676	0.4860
	fuzzy	0.00068	0.02613	0.01265	5.42259	0.4860
	recurn	0.00015	0.01223	0.00763	3.16142	0.5120
	regrsn	0.00387	0.06222	0.00674	2.17746	0.4860
Germany	atnn	0.00010	0.00991	0.00288	0.78218	0.5035
	backprop	0.00008	0.00913	0.00647	2.02970	0.5010
	boxjen	0.00300	0.05476	0.00356	1.42232	0.4905
	dblexpn	0.00016	0.01247	0.00857	2.39283	0.5035
	exponen	0.00490	0.07002	0.02620	7.48069	0.5033
	fuzzy	0.00043	0.02065	0.00877	1.98431	0.5033
	recurn	0.00012	0.01116	0.00764	2.44117	0.5010
	regrsn	0.00401	0.06334	0.00739	2.27253	0.5033
Switzerland	atnn	0.00025	0.01572	0.00285	0.82335	0.4962
	backprop	0.00024	0.01556	0.01317	3.92474	0.5030
	boxjen	0.00298	0.05455	0.00471	1.76606	0.4889
	dblexpn	0.00021	0.01437	0.01017	2.95273	0.4962
	exponen	0.00514	0.07171	0.02935	8.59970	0.4963
	fuzzy	0.00035	0.01882	0.00770	1.73010	0.4963
	recurn	0.00016	0.01279	0.00888	2.93006	0.5030
	regrsn	0.00399	0.06320	0.00855	2.63371	0.4963

Table 62: 5 days average forecasts of exchange rates over window size 5

Bibliography

[1] Armstrong, J. Scott and Fred Collopy, "Error messages for generalizing about forecasting methods: Emperical comparisons", *International Journal of Forecasting*, **vol 8**, pp 69-80, 1992.

[2] Bellman, R., *Dynamic Programming*, Princeton University Press, NJ, 1957.

[3] Bowerman, Bruce, L. and O'Connell, Richard, T., *Time Series and forecasting, an applied approach*, Duxbury Press., Belmont, California, 1979.

[4] Box, G.E.P. and Jenkins G.M., *Time Series Analysis, Forecasting and Control*, Holden-Day, Inc., San Francisco, 1970.

[5] Caldwell, Randal, B., "Performance metrics for neural network-based trading system development", *NEUROVE$T JOURNAL*, **March/April**, pp 13-23, 1995.

[6] Cao, Z. Q., Kandel, A., and Li, L.H., "A new model of fuzzy reasoning", *Fuzzy Sets and Systems*, **vol 36**, pp 311-325, 1990.

[7] Elman, Jeffrey, L., "Finding structure in time", *Cognitive Science*, **vol 14**, pp 179-211, 1990.

[8] Fuller, Wayne A., *Introduction to statistical time series*, John Wiley & Sons, Inc., NY, 1976.

[9] Glass, Leon and Kaplan, Daniel, "Time series analysis of complex dynamics in physiology and medicine", *Medical Progress through Technology*, **vol 19**, pp 115-128, 1993.

[10] Granger, C.W.J. and Joyeux, Roselyne, "An introduction to long-memory time series models and fractional differencing", *Journal of Time Series Analysis*, **vol 1, no.1**, pp 15-29, 1980.

[11] Hinton, G.E., McClelland, J.L., and Rumelhart, D.E., "Learning internal representations by error propagation", *Parallel distributed processing: Explorations in the microstructure of cognition*, **vol 1**, pp 318-362, 1986.

[12] Hopfield, J.J., "Neural networks as physical systems with emergent collective computational abilities", *Proceedings of the National Academy of Sciences*, **vol 79**, pp 531-546, 1986.

[13] Hosking, J.R.M., "Fractional differencing", *Biometrika*, **vol 68**, pp 165-176, 1981.

[14] Jazwinski, A.H., *Stochastic Process and Filtering Theory*, New York Academy Press, 1957.

[15] Johnson, L.A. and Montgomery, D.C., *Forecasting and time series analysis*, McGraw-Hill Book Company, NY, 1976.

[16] Jordan, M.I., "Attractor dynamics and parallelism in a connectionist sequential machine", *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, **vol 8**, pp 531-546, 1986.

[17] Kumar, K.S.P. and Shridar, R., "On the identification of control systems by quasi-linearization method.", *IEEE Trans. Automat. Contr.*, **vol AC-9**, pp 151-154, 1964.

[18] Lewine, Donald, *POSIX Programmer's Guide*, O'Reilly & Associates, Inc. CA, 1992.

[19] Li, T., Guo, D., Jayakumar, R., Tang, Y.Y. and Klasa, S., "A fuzzy learning expert system and its applications", *Proceedings of the 6th Australian Joint Conference on Artificial Intelligence*, **Melbourne, November 1993**, pp 119-125, 1993.

[20] Lippmann, Richard P., "An introduction to computing with neural nets", *IEEE ASSP Magazine*, **vol 4 no. 2**, pp 4-20, 1987.

[21] Luenberger, D., *Linear and nonlinear programming*, Reading, Addison-Wesley, MA, 1984.

[22] Mandelbrot, Benoit B and Ness, John W., "Fractional brownian motions, fractional noises and applicaations", *SIAM REVIEW*, **vol 10, No. 4**, pp 422-437, 1968.

[23] Pao, Yoh-Han, *Adaptive pattern recognition and neural networks*, Addison-Wesley Publishing Company Inc., 1989.

[24] Park, Dong C., El-Sharkawi, Mohammed A., Marks II, Robert J., "An adaptively trained neural network", *IEEE Transcations on Neural Networks*, **vol 3**, pp 334-345, 1991.

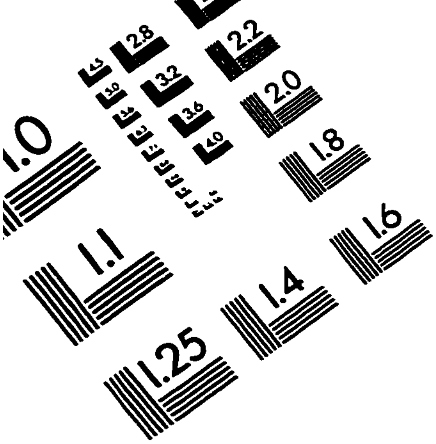
[25] Sage, S.P. and Eisenberg, B.R., "Experiments in nonlinear and non-stationary system identification via quasilinearization and differential approximation ", *Proc. Joint Automat. Contr. Conf.*, , pp 522-530, 1965.

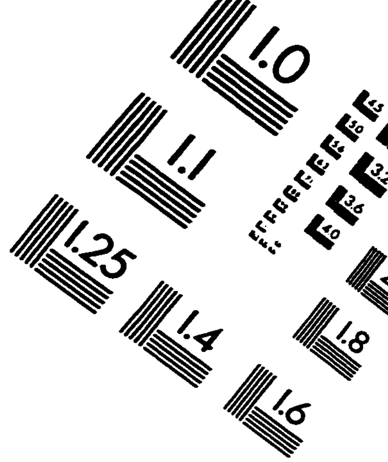
[26] Williams, Ronald J. and Zisper, David, "A learning algorithm for continually running fully recurrent neural networks", *Neural Computation*, **vol 1**, pp 270-280, 1989.

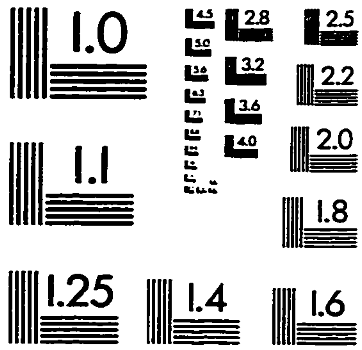
[27] Winters, P.R., "Forecasting sales by exponentially weighted moving averages", *Management Science*, vol **6** **no.3**, pp 324-342, 1960.

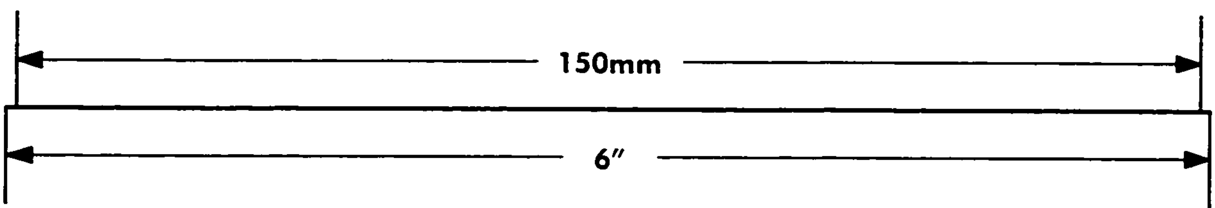
[28] Zadeh, L. A., "Outline of a new approach to the analysis of complex systems and decision processes", *IEEE Trans. Systems, Man, and Cybernetics*, vol **3** **no.1**, pp 28-44, 1973.

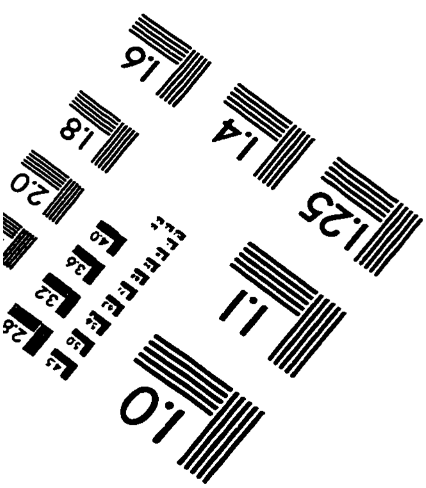
TEST TARGET (QA-3)

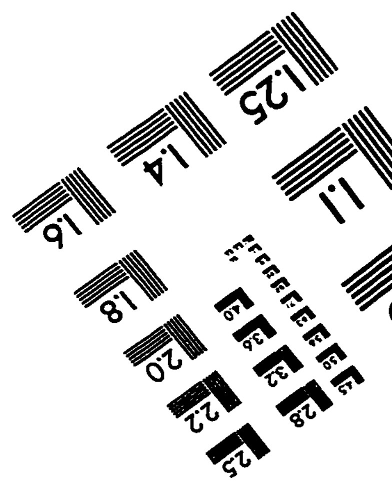












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved