College of Science and Technology

School of Science

Department of Mathematics

Forecasting Exchange Rates: *US dollar* (USD) versus *Rwandan franc* (RWF)

A dissertation submitted in partial fulfillment of the requirement for the award of Bachelor's Degree of Science with honor in Mathematics.

Author:
Aimable Parfait Kamugisha
Reg. No: 215009209

Supervisor:
Mr. Jean-Paul Murara

Academic year: 2017-2018

June 11, 2018

Declaration

I, **Aimable Parfait KAMUGISHA**, student from University of Rwanda, College of Science and Technology, with **registration number 215009209**, I hereby declare and confirm that, this research project entitled **Forecasting Exchange rates: *US dollar* (USD) versus *Rwandan franc* (RWF)** submitted to the College of Science and Technology(CST) in partial fulfillment of requirement for the award of the degree of Bachelor's of Science with honor in Mathematics, is my own contribution under the supervision of **Mr. Jean-Paul MURARA**. I further declare that the work reported in this project has never been submitted in any University or High Learning Institution for any award.

CANDIDATE	*SUPERVISOR*
Aimable Parfait KAMUGISHA	Mr. Jean-Paul MURARA
Date:........../........../2018	Date:........../........../2018
Signature:......................	Signature:......................

Approval

This research project has been submitted with approval as university supervisor

Mr. Jean-Paul Murara
Department of Mathematics
School of Science
College of science and Technology (CST)
P.O. Box 3900 Kigali
Website: www.cst.ur.ac.rw

UNIVERSITY OF RWANDA

College of Science and Technology

School of Science

Department of Mathematics

CERTIFICATE

This is to certify that this research project **Forecasting Exchange rates: *US dollar* (USD) versus *Rwandan franc* (RWF)** is original work done by: **Aimable Parfait KAMUGISHA, with registration number: 215009209**. As partial fulfillment of the requirements for the award of Bachelor's of Science in Mathematics at University of Rwanda, College of Science and Technology during the academic year 2017 − 2018.

SUPERVISOR
Mr. Jean-Paul MURARA
Date:........../........../2018
Signature:......................

HEAD OF DEPARTMENT
Dr. Wellars BANZI
Date:........../........../2018
Signature:......................

Dedication

I dedicate this project to:
My beloved mother,
My brothers,
My friends.

Acknowledgement

First of all, I would like to acknowledge and give thanks in a special way to the Almighty God,who helped me to understand myself and my worth,and i would like to thank my mother for believing me and supporting me throughout my days at college and for the period of research.

I would like to express sincere appreciation to my supervisor Mr. Jean-Paul MURARA, for his invaluable support, commitment, effort, wise advice, review, corrections, encouragement and guidance that made this work fruitful. I would like to express gratitude to Dr. BANZI Wellars, Dr. Japhet NIYOBUHUNGIRO, Mrs. Denyse UWAMARIYA, Dr. Denis NDANGUZA, Dr. Innocent RUSAGARA and Dr. KURUJYIBWAMI Celestin for their support, guidance, documentation and advice, I am very gratefully to the government of Rwanda for the financial support that has made my studies successful.

Special thanks to all my lecturers in the School of Science in CST, department of Mathematics for the intellectual package they provided to me. Without their guidance, support and advice I would not have successfully finished this work. I would like to forward my deepest gratitude to everyone, who contributed to this study in terms of prayers, ideas, technical knowledge and support, I would like to extend my deepest and most wholehearted gratitude.

Abstract

The work of this thesis primarily revolves around the concept of forecasting the daily exchange rates of the United States Dollar versus Rwandan franc. Forecasting is a relatively important issue in business operations, however it is also one of the most problematic. With the uncertainty of the future, forecasts are difficult to rely on. The aim of this thesis is to analyze some of the forecasting models used on exchange rates, and determine which one of the forecasting models properly work when applied to exchange rates, why or why not, and their measure of accuracy. The methodology of this thesis revolves extensively around quantitative research. Exchange rates values from January 2015 to the end of April 2018 were collected and recorded to form a data set. The exchange rate data was then used in the application of a variety of mathematical forecasting models to forecast the daily exchange rates for a future, one-month period. Upon measuring the accuracy of the forecasts, the forecasted exchange rates contained very little error. Therefore, the forecasts are considered to be successful, and the hypothesis that exchange rates could be determined with the aid of a mathematical forecasting model is accepted. Though it is very difficult to consistently estimate exchange rates successfully, the work of this thesis shows there is always a greater probability of benefiting from a forecast.

Contents

Declaration i
Approval ii
Certificate iii
Dedication iv
Acknowledgement v
Abstract vi
List of Figures ix
List of Tables xi

1 Introduction **1**
1.1 Exchange rates 1
1.1.1 Appreciation 1
1.1.2 Depreciation 2
1.1.3 Factors affecting the fluctuation of exchange rates 2
1.2 Forecasting 3
1.2.1 Forecasting methods 3
1.2.2 Forecasting Exchange rates 6
1.3 Components of time series 7
1.4 Time series models 8
1.4.1 Generalized autoregressive conditional heteroskedasticity (GARCH(p,q)) model 8
1.4.2 ARIMA(p,d,q) models 11
1.4.3 Exponential smoothing 18
1.4.4 ARIMA vs ETS 20
1.5 Problem statement 20
1.6 Objective of the study 21
1.6.1 General objective 21
1.6.2 Specific objectives 21
1.7 The limitation of the study 21
1.8 Justification of the study 21

2 Literature review **22**

3 Methodology **23**
3.1 Data source and Data collection 23
3.2 Sample size 23
3.3 Methodology of data collection 23
3.4 Methodology of forecasting 24
3.4.1 Stationarity test 24

3.4.2 Evaluating forecast accuracy 26

4 RESULTS AND ANALYSIS 28
4.1 Data description 28
4.2 Stationarity testing 29
4.2.1 Augmented Dicker-Fuller test (Adf test) 29
4.2.2 Autocorrelation functions 29
4.3 Forecasting models 30
4.3.1 Exponential smoothing 30
4.3.2 ARIMA(p,d,q) models 31
4.3.3 GARCH(p,q) models 33
4.4 Choosing the best model 35
4.4.1 Forecasting Accuracy 35
4.5 Forecasting Exchange rates Between US dollar (USD) and Rwandan franc of the month May,2018 36
4.5.1 Forecasts from HoltWinters 36
4.5.2 Forecasts from ARIMA(p,d,q) model 37
4.5.3 Forecasting using GARCH(p,q) model 39

5 DISCUSION, CONCLUSION AND RECOMMENDATION 42
5.1 DISCUSION 42
5.2 CONCLUSION 42
5.3 RECOMMENDATION 43
5.4 Appendices 44
5.4.1 Appendix A: The change of raw data into time series data, using R. . . . 44
5.4.2 Appendix B: Fitting an exponential smoothing model to our time series data, using R. 44
5.4.3 Appendix C: Fitting an ARIMA model to our time series data, using R. 44
5.4.4 Appendix D: Fitting a GARCH model to our time series data, using R. . 45
5.4.5 Appendix E: Exchange rates of USD Vs RWF of the month of May 2018 table. 46
REFERENCES 47

List of Figures

1.1 Components of a time series 8
1.2 Simulated GARCH(1,1) data with $\alpha_1 = 0.85$ and $\beta_1 = 0.10$. 10
1.3 Estimated volatility of the simulated GARCH(1,1) data with $\alpha_1 = 0.85$ and $\beta_1 = 0.10$. 10
1.4 Simulated GARCH(1,1) data with $\alpha_1 = 0.10$ and $\beta_1 = 0.85$. 10
1.5 Estimated volatility of the simulated GARCH(1,1) data $\alpha_1 = 0.10$ and $\beta_1 = 0.85$. 10
1.6 Which of these series are stationary? (a) Dow Jones index on 292 consecutive days; (b) Daily change in Dow Jones index on 292 consecutive days; (c) Annual number of strikes in the US; (d) Monthly sales of new one-family houses sold in the US; (e) Price of a dozen eggs in the US (constant dollars); (f) Monthly total of pigs slaughtered in Victoria, Australia; (g) Annual total of lynx trapped in the McKenzie River district of north-west Canada; (h) Monthly Australian beer production; (i) Monthly Australian electricity production. 12
1.7 The ACF of the Dow-Jones index (left) and of the daily changes in the Dow-Jones index (right). 13
1.8 Two examples of data from autoregressive models with different parameters. Left: *AR(1)* with $y_t = 18 - 0.8y_{t-1} + e_t$ Right: *AR(2)* with $y_t = 8 + 1.3y_{t-1} - 0.7y_{t-2} + e_t$. In both cases, e_t is normally distributed white noise with mean zero and variance one. 15
1.9 Two examples of data from moving average models with different parameters. Left: *MA(1)* with $y_t = 20+0.8e_{t-1}+e_t$ Right: *MA(2)* with $y_t = e_t - e_{t-1} + 0.8e_{t-2}$. In both cases, e_t is normally distributed white noise with mean zero and variance one. 16

4.1 A plot of Exchange rates of **US dollar** against **Rwandan franc** from January 2^{nd},2015 to April 30^{th},2018. 28
4.2 Augmented Dicker-Fuller test on our exchange rates data, using R 29
4.3 Autocorrelation function of exchange rates time series data, using R 30
4.4 Holt-Winter's exponential smoothing with trend and without seasonal component, using R 31
4.5 Autoregressive Integrated Moving Average model, using R 31
4.6 Plot of residuals of ARIMA(1,2,3) model, using Stata 32
4.7 Histogram of residuals of ARIMA(1,2,3) model, using R 33
4.8 Generalized Autoregressive Conditional Heteroscedasticity, using R 34
4.9 Accuracy of the Holt-Winters exponential smoothing, using R 35
4.10 Accuracy of the ARIMA(1,2,3), using R 35
4.11 Forecasts from HoltWinters, using R 36
4.12 The forecast results using HoltWinters exponential smoothing, using R 37
4.13 Plot of the forecast results using ARIMA (1,2,3) model, using R 38

4.14 Values of the forecast results,for the month of May 2018, using ARIMA (1,2,3) model, using R 39
4.15 Graph describing the upper and lower limits of the forecasts of GARCH(1,1) model, using R 40
4.16 Garch forecasts of the month of May 2018, using R 41

List of Tables

1.1 Estimates and t-ratios of both models with the same $\alpha_1 + \beta_1$ value 10
1.2 Cases of the ARIMA model 17
1.3 ETS and ARIMA 20

3.1 Critical Values for the DF Test 25

4.1 Aic and Bic 35

Chapter 1

Introduction

In modern days, where global economy is possible, exchange rate is a very usual concept that plays a big role in international trades. When you are on your way to the market, going to shop, and you have an encounter with a sign saying ”foreign exchange rates” with numbers written on it, if you are a finance or economy enthusiast you probably know what that sign means, but if you are not you probably pass the sign without even noticing it. Simply, the sign is trying to tell you how much it would cost you to get a foreign currency, a currency from a different country from yours, by exchanging it with your local currency. For example let us say that you have a relative in Burundi and you are trying to send him/her money so that he/she can bring you some Burundian Saladin, you have two ways for sending the money: the first one is that you can send him/her the money in your currency and him/her can exchange your currency into the Burundian currency in Burundi, depending on the present rates that can be done. The second choice is you can exchange your local currency into Burundian currency before sending the money to your relative. either way, the exchange rates plays a big role in such type of transaction. With that simple definition, someone can get an intuition of what is ”exchange rates” and what does it do. Exchange rates values vary with time and are unpredictable, they are stochastic processes, which means that you can't know for sure how the exchange rates values will behave let's say in the next two weeks. The only thing you can do is to use your instinct or even mathematical methods to predict those values for the future. Here we are interested on how to predict those values using mathematical methods and understanding those methods used. We got the following definitions with help from [1], [2], [3], [4] and [5].

1.1 Exchange rates

Exchange rate is the rate at which one currency will be exchanged for another. It means that it is the value of one country's currency in relation to another currency. For example, Bank of Kigali exchange rate of 859 Rwandan francs to the United States dollar means that RWF 859 will be exchanged for each USD 1 or that USD 1 will be exchanged for each RWF 859. Exchange rates rise and fall through time, which is a phenomenon called ”fluctuations in exchange rate”, and the rise and fall of the exchange rates introduces two new concepts which are called *Appreciation* and *Depreciation*.

1.1.1 Appreciation

Appreciation is an increase in the value of a currency relative to another currency.an appreciated currency is more valuable and therefore can be exchanged for more amount of foreign currency.For example if RWF 1 / USD 1 $\rightarrow$ RWF 0.80 / USD 1 means that the USD has appreciated against the RWF.

1.1.2 Depreciation

Depreciation is a decrease in the value of a currency relative to another currency.A depreciated currency is less valuable and therefore can be exchanged for a small amount of foreign currency. We give another example, consider that RWF 1 / USD 1 $\rightarrow$ RWF 1 / USD 1.60 it means that the *USD* has depreciated against the *RWF*, it takes USD 1.60 to buy RWF 1, so that the US dollar, *USD*, is less valuable. At the same time the Rwandan franc,*RWF*, has appreciated against the US dollar, it is now more valuable.

1.1.3 Factors affecting the fluctuation of exchange rates

Numerous factors determine exchange rates, and all are related to the trading relationship between two countries. Remember, exchange rates are relative, and are expressed as a comparison of the currencies of two countries. The following are some of the principal determinants of the exchange rate between two countries.these factors are in no particular order.

- Interest Rates Level
  Interest rates are the cost and profit of borrowing capital. When a country raises its interest rate or its domestic interest rate is higher than the foreign interest rate, it will cause capital inflow, thereby increasing the demand for domestic currency, allowing the currency to appreciate and the foreign exchange depreciate.

- Balance of payments
  When a country has a large international balance of payments deficit or trade deficit, it means that its foreign exchange earnings are less than foreign exchange expenditures and its demand for foreign exchange exceeds its supply, so its foreign exchange rate rises, and its currency depreciates.

- Inflation factor
  As a general rule, a country with a consistently lower inflation rate exhibits a rising currency value, as its purchasing power increases relative to other currencies. Those countries with higher inflation typically see depreciation in their currency in relation to the currencies of their trading partners.

- Fiscal and monetary policy
  Although the influence of monetary policy on the exchange rate changes of a country's government is indirect, it is also very important. In general, the huge fiscal revenue and expenditure deficit caused by expansionary fiscal and monetary policies and inflation will devalue the domestic currency. The tightening fiscal and monetary policies will reduce fiscal expenditures, stabilize the currency, and increase the value of the domestic currency.

- Venture capital
  If speculators expect a certain currency to appreciate, they will buy a large amount of that currency, which will cause the exchange rate of that currency to rise. Conversely, if speculators expect a certain currency to depreciate, they will sell off a large amount of the currency, resulting in speculation. The currency exchange rate immediately fell. Speculation is an important factor in the short-term fluctuations in the exchange rate of the foreign exchange market.

- Government market intervention
  When exchange rate fluctuations in the foreign exchange market adversely affect a country's economy, trade, or the government needs to achieve certain policy goals through exchange rate adjustments, monetary authorities can participate in currency trading, buying or selling local or foreign currencies in large quantities in the market. The foreign exchange supply and demand has caused the exchange rate to change.

- Political Stability and Economic Performance
  Foreign investors inevitably seek out stable countries with strong economic performance in which to invest their capital. A country with such positive attributes will draw investment funds away from other countries perceived to have more political and economic risk. Political turmoil, for example, can cause a loss of confidence in a currency and a movement of capital to the currencies of more stable countries.

1.2 Forecasting

Forecasting is a way of predicting, projecting, or estimating of some future activity, event, or occurrence.It is simply a statement about the future.Forecasts are possible only when a history of data exists. forecasting is used in business, finance and economy, where they need to know how data may behave in the future to make robust plans.
More generally, firms use forecasts to decide what to produce (What product or mix of products should be produced?), when to produce (Should we build up inventories now in anticipation of high future demand? How many shifts should be run?), how much to produce and how much capacity to build (What are the trends in market size and market share? Are there cyclical or seasonal effects? How quickly and with what pattern will a newly-built plant or a newly-installed technology depreciate?), and where to produce (Should we have one plant or many? If many, where should we locate them?). Firms also use forecasts of future prices and availability of inputs to guide production decisions.

1.2.1 Forecasting methods

Qualitative methods

These types of forecasting methods are based on judgments, opinions, intuition, emotions, or personal experiences and are subjective in nature. They do not rely on any rigorous mathematical computations.

- Grass roots
  Grass roots forecasting builds the forecast by adding successively from the bottom. The assumption here is that the person closest to the customer or end use of the product knows its future needs best. Though this is not always true, in many instances it is a valid assumption, and it is the basis for this method. Forecasts at this bottom level are summed and given to the next higher level. This is usually a district warehouse, which then adds in safely stocks and any effects of ordering quantity sizes. This amount is then fed to the next level, which may be a regional warehouse. The procedure repeat until it becomes an input at the top level, which, in the case of a manufacturing firm, would be the input to the production system.

- Market research
  Firms often hire outside companies that specialize in market research to conduct this type

of forecasting. You may have been involved in market surveys through a marketing class. Certainly you have not escaped telephone calls asking you about product preferences, your income, habits, and so on. Market research is used mostly for product research in the sense of looking for new product ideas, likes and dislikes about existing products, which competitive products within a particular class are preferred, and so on. Again, the data collection methods are primarily surveys and interviews.

- Panel consensus
  In a panel consensus, the idea that two heads are better than one is extrapolated to the idea that a panel of people from a variety of positions can develop a more reliable forecast than a narrower group. Panel forecasts are developed through open meetings with free exchange of ideas form all levels of management and individuals. The difficulty with this open style is that lower employee levels are intimidated by higher levels of management. For example, a salesperson in a particular product line may have a good estimate of future product demand but may not speak up to refute a much different estimate given by the vice president of marketing. The Delphi technique (which we discuss shortly) was developed to try to correct this impairment to free exchange. When decisions in forecasting are at a broader, higher level (as when introducing a new product line or concerning strategic product decisions such as new marketing areas) the term executive judgment is generally used. The term is self-explanatory: a higher level of management is involved.

- Historical analogy
  The historical analogy method is used for forecasting the demand for a product or service under the circumstances that no past demand data are available. This may specially be true if the product happens to be new for the organization. However, the organization may have marketed product(s) earlier which may be similar in some features to the new product. In such circumstances, the marketing personnel use the historical analogy between the two products and derive the demand for the new product using the historical data of the earlier product. The limitations of this method are quite apparent. They include the questionable assumption of the similarity of demand behaviors, the changed marketing conditions, and the impact of the capability of being substituted factor on the demand.

- Delphi method
  a statement or opinion of a higher-level person will likely be weighted more than that of a lower-level person. The worst case in where lower level people feel threatened and do not contribute their true beliefs. To prevent this problem, the Delphi method conceals the identity of the individuals participating in the study. Everyone has the same weight. A moderator creates a questionnaire and distributes it to participants. Their responses are summed and given back to the entire group along with a new set of questions.

Quantitative methods

These types of forecasting methods are based on mathematical (quantitative) models, and are objective in nature. They rely heavily on mathematical computations.

- **Average approach**

For this approach, the predictions of future values are equal to the mean of all past data.

$$\hat{y}_{T+h|T} = \frac{1}{T}\sum_{t=0}^{T} y_t = \bar{y} \tag{1.1}$$

- **Naïve approach**
  This approach is the most cost-effective, it is used on time series data only.the forecasts are produced that are equal to the last observed value.

$$\hat{y}_{T+h} = y_T \tag{1.2}$$

- **Drift method**
  In this approach, the forecast equal to the last value observed plus average change.

$$\hat{y}_{T+h|T} = y_T + \frac{h}{T-1}\sum_{t=2}^{T}(y_t - y_{t-1}) = y_T + \frac{h}{T-1}(y_T - y_1) \tag{1.3}$$

- **Seasonal naïve approach**
  This approach accounts for seasonality by setting each prediction to be equal to the last observed value of the same season.it is very useful for data that has a very high level of seasonality.

$$\hat{y}_{T+h|T} = y_{T+h-km} \tag{1.4}$$

  where m = seasonal period and $k = \left[\frac{(h-1)}{m}\right] + 1$.

- **casual or econometric forecasting methods**
  Some forecasting methods try to identify the underlying factors that might influence the variable that is being forecast. For example, including information about climate patterns might improve the ability of a model to predict umbrella sales. Forecasting models often take account of regular seasonal variations. In addition to climate, such variations can also be due to holidays and customs: for example, one might predict that sales of college football apparel will be higher during the football season than during the off season.casual method include regression analysis and Auto-regressive moving average with exogenous inputs (ARMAX)

- **Linear Regression**
  This method use a linear equation that show relationship between the independent variables and the dependent variables, in our case, which are the number of observed exchange rates and the observed exchange rates respectively. the simple linear regression uses the equation

$$y = \hat{\beta}_0 + \hat{\beta}_1 x \tag{1.5}$$

  where x is the independent variable and y is the dependent variables.$\hat{\beta}_0$ and $\hat{\beta}_1$ are estimated parameters.

- **Time series methods**
  time series methods use historical data to estimate future outcomes.One of the more popular time series approaches is called the autoregressive moving average (ARMA) process. The rationale for using this method is based on the idea that your data are stationary,

which means that they have no trends, no seasonality and no cyclical components and past behavior and price patterns can be used to predict future price behavior and patterns. All you need to use this approach is a time series of stationary data that can then be entered into a computer program to estimate the parameters and create a model suitable for the data.other methods will be described in the next section.

1.2.2 Forecasting Exchange rates

Firms such as multinational corporations (MNCs) which have facilities and other assets in at least one country other than its home country and investors need Exchange rates forecasts for their:

1. Hedging decisions,
2. Short-term financing decisions,
3. Short-term investment decisions,
4. Capital budgeting decisions,
5. Earnings assessments, and
6. Long-term financing decisions.

There are numerous methods used to forecast exchange rates, and they can be categorized into four general groups:

1. Technical,
2. Fundamental,
3. Market-based, and
4. Mixed.

Technical forecasting involves the use of historical data to predict future values.e.g Time series models.
Speculators may find the models useful for predicting day-to-day movements.However, since the models typically focus on the near future and rarely provide point or range estimates, they are of limited use to MNCs.

Fundamental forecasting is based on the fundamental relationships between economic variables and exchange rates.for example: subjective assessments, quantitative measurements based on regression models and sensitivity analyses.
Note that the use of PPP(Purchase Power Parity) to forecast future exchange rates is inadequate since PPP may not hold and future inflation rates are also uncertain.
In general, fundamental forecasting is limited by:

- The uncertain timing of the impact of the factors,
- The need to forecast factors that have an immediate impact on exchange rates,
- The omission of factors that are not easily quantifiable, and

- Changes in the sensitivity of currency movements to each factor over time.

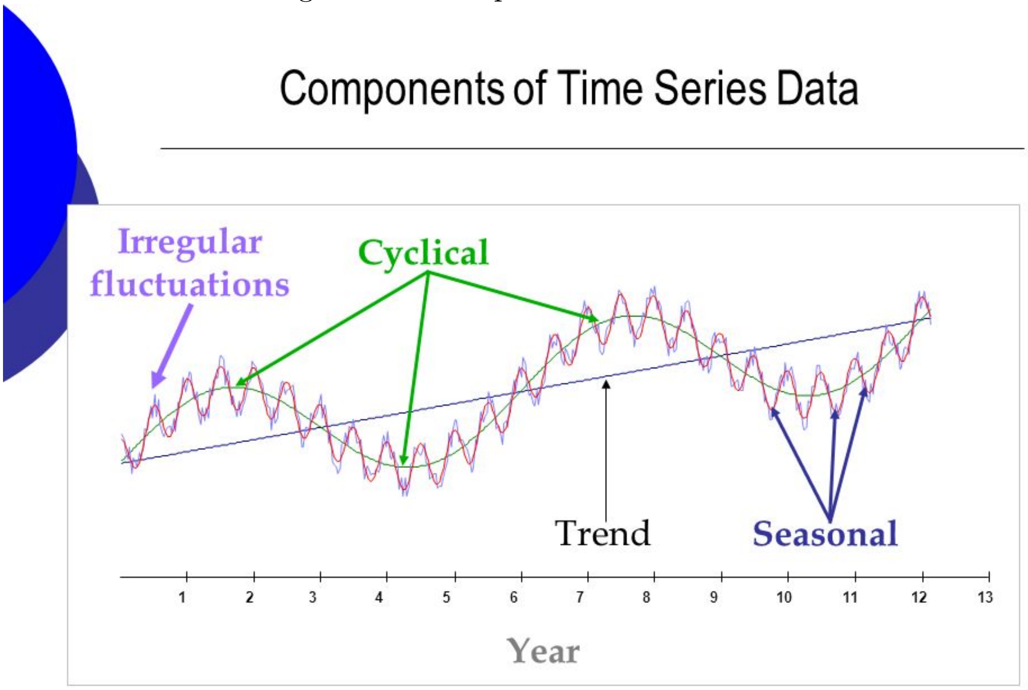
Market-Based Forecasting uses market indicators to develop forecasts.The current spot/forward rates are often used, since speculators will ensure that the current rates reflect the market expectation of the future exchange rate.
For long-term forecasting, the interest rates on risk-free instruments can be used under conditions of IRP.

Mixed Forecasting refers to the use of a combination of forecasting techniques.The actual forecast is a weighted average of the various forecasts developed.

1.3 Components of time series

- **Trend component**
  Data exhibit a steady growth or decline over time.
- **Seasonal component**
  Data exhibit upward and downward swings in a short to intermediate time frame (most notably during a year).
- **Cyclical component**
  Data exhibit upward and downward swings in over a very long time frame.
- **Irregular or random variations**
  These refer to the erratic fluctuations in the data which cannot be attributed to the trend, seasonal or cyclical factors. In many cases, the root cause of these variations can be isolated only after a detailed analysis of the data and the accompanying explanations, if any. Such variations can be due to a wide variety of factors like sudden weather changes, strike or a communal clash. Since these are truly random in nature, their future occurrence and the resulting impact on demand are difficult to predict. The effect of these events can be eliminated by smoothing the time series data.

Figure 1.1: Components of a time series



1.4 Time series models

An important part of working with time series data is the selection of a suitable probability model (or class of models) for the data.there exist many models, time series models,that can be used for time series analysis. For us, we will work with models that are suitable for our exchange rates data. There are two popular models for our type of data, non stationary data with a strong trend, which are GARCH (p,q) and ARIMA (p,d,q) models. In our case we will analyze three models, which are the GARCH model, ARIMA models and Exponential smoothing.

1.4.1 Generalized autoregressive conditional heteroskedasticity (GARCH(p,q)) model

generalized autoregressive conditional heteroskedasticity (GARCH(p,q)) model is a statistical model for time series data that describes the variance of the current error term or innovation as a function of the actual sizes of the previous time periods' error terms.an autoregressive moving average model (ARMA) model is assumed for the error variance.
The GARCH process is often preferred by financial modeling professionals because it provides a more real-world context than other forms when trying to predict the prices and rates of financial instruments.

ARCH(q) model

An ARCH(q) process is one for which the variance at time t is conditional on observations at the previous q times, and the relationship is

$$Var(y_t|y_{t-1},\dots,y_{t-q}) = \sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \dots + \alpha_q y_{t-q}^2. \tag{1.6}$$

With certain constraints imposed on the coefficients, the y_t series squared will theoretically be AR(q).

GARCH (p,q) model

Let (Z_n) be a sequence of i.i.d. random variables such that $Z_t \sim N(0,1)$. (X_t) is called **the generalized autoregressive conditionally heteroskedastic** or GARCH(q,p) process if

$$X_t = \sigma_t Z_t, \quad t \in \mathbb{Z} \tag{1.7}$$

where (σ_t) is a nonnegative process such that

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + ... + \alpha_q X_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + ... + \beta_p \sigma_{t-p}^2, \quad t \in \mathbb{Z} \tag{1.8}$$

and

$$\alpha_0 > 0, \alpha_i \geq 0 \;\; i = 1, ..., q \quad \beta_i \geq 0 \;\; i = 1, ..., p. \tag{1.9}$$

The conditions on parameters,(1.9), ensure strong positivity of the conditional variance (1.8).

GARCH (1,1) model

The most used heteroscedastic model in financial time series is a GARCH(1,1).
This particular model parameterises the conditional variance as

$$\begin{aligned} y_t &= \mu_t + X_t, \\ \mu_t &= \dots \text{(e.g. a constant or an ARMA equation without the term } u_t), \\ X_t &= \sigma_t Z_t, \\ \sigma_t^2 &= \omega + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \\ Z_t &\sim i.i.d(0,1). \end{aligned} \tag{1.10}$$

Assuming the process began infinitely far in the past with a finite initial variance, the sequence of the variances converge to a constant

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1},$$

if $\alpha_1 + \beta_1 < 1$
therefore, the GARCH process is unconditionally homoscedastic. The α_1 parameter indicates the contributions to conditional variance of the most recent news, and the β_1 parameter corresponds to the moving average part in the conditional variance, that is to say, the recent level of volatility.

Example 1. *In this example we generate several time series following a GARCH(1,1) model, with the same value of $\alpha_1 + \beta_1$, but different values of each parameter to reflect the impact of β_1 in the model. The example was from [5]*

Table 1.1: Estimates and t-ratios of both models with the same $\alpha_1 + \beta_1$ value

	α_0	α_1	β_1
Model A	0.05	0.85	0.10
estimates	0.0411	0.8572	0.0930
t-ratio	(1.708)	(18.341)	(10.534)
Model B	0.05	0.10	0.85
estimates	0.0653	0.1005	0.8480
t-ratio	1.8913	3.5159	18.439

Figure 1.2: Simulated GARCH(1,1) data with $\alpha_1 = 0.85$ and $\beta_1 = 0.10$.

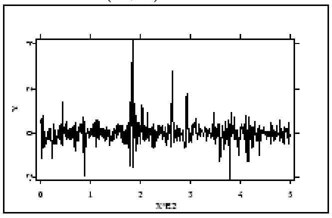


Figure 1.3: Estimated volatility of the simulated GARCH(1,1) data with $\alpha_1 = 0.85$ and $\beta_1 = 0.10$.

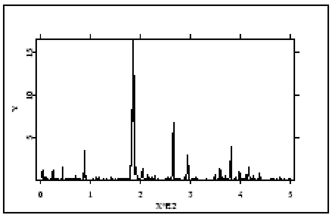


Figure 1.4: Simulated GARCH(1,1) data with $\alpha_1 = 0.10$ and $\beta_1 = 0.85$.

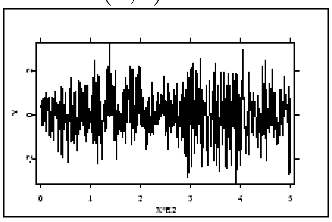
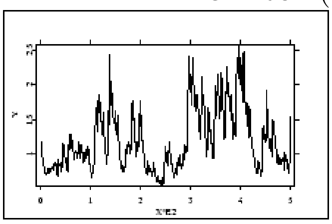


Figure 1.5: Estimated volatility of the simulated GARCH(1,1) data $\alpha_1 = 0.10$ and $\beta_1 = 0.85$.



All estimates are significant in both models. The estimated parameters $\hat{\alpha}_1$ and $\hat{\beta}_1$ indicate that the conditional variance is time-varying and strongly persistent ($\hat{\alpha}_1 + \hat{\beta}_1 \approx 0.95$). But the conditional variance is very different in both models as we can see in figures 6.10 and 6.12.

1.4.2 ARIMA(p,d,q) models

ARIMA models provide another approach to time series forecasting. Exponential smoothing and ARIMA models are the two most widely-used approaches to time series forecasting, and provide complementary approaches to the problem. While exponential smoothing models were based on a description of trend and seasonality in the data, ARIMA models aim to describe the autocorrelations in the data.
Before we introduce ARIMA models, we need to first discuss the concept of stationarity and the technique of differencing time series with help from [5].

Stationarity

A stationary time series is one whose properties do not depend on the time at which the series is observed.1 So time series with trends, or with seasonality, are not stationary — the trend and seasonality will affect the value of the time series at different times. On the other hand, a white noise series is stationary — it does not matter when you observe it, it should look much the same at any period of time.
Some cases can be confusing — a time series with cyclic behavior (but not trend or seasonality) is stationary. That is because the cycles are not of fixed length, so before we observe the series we cannot be sure where the peaks and troughs of the cycles will be.
In general, a stationary time series will have no predictable patterns in the long-term. Time plots will show the series to be roughly horizontal (although some cyclic behavior is possible) with constant variance.

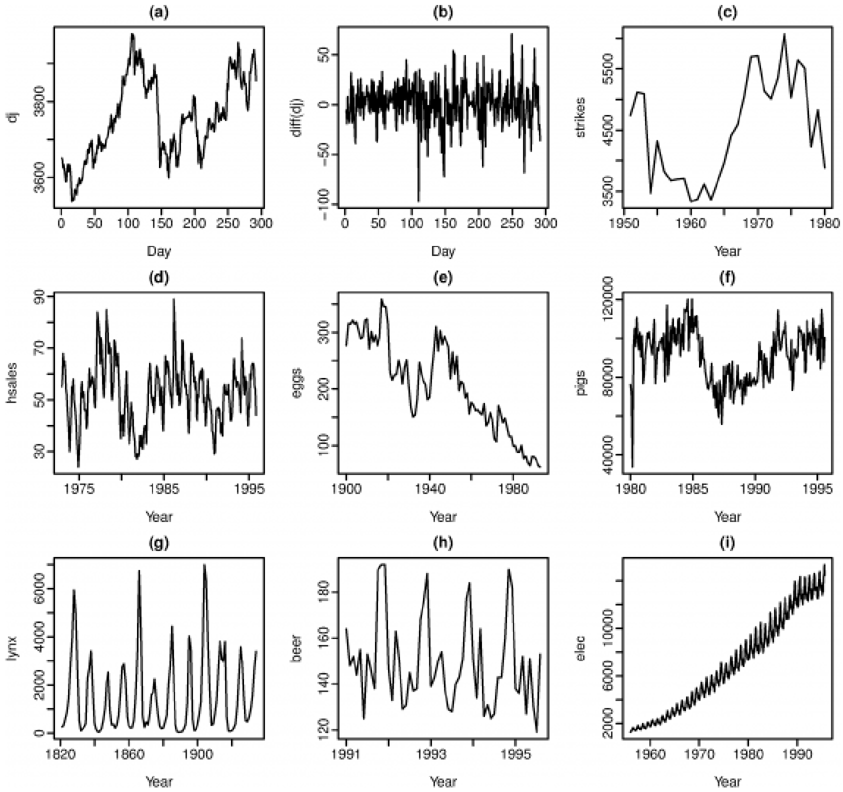


Figure 1.6: Which of these series are stationary? (a) Dow Jones index on 292 consecutive days; (b) Daily change in Dow Jones index on 292 consecutive days; (c) Annual number of strikes in the US; (d) Monthly sales of new one-family houses sold in the US; (e) Price of a dozen eggs in the US (constant dollars); (f) Monthly total of pigs slaughtered in Victoria, Australia; (g) Annual total of lynx trapped in the McKenzie River district of north-west Canada; (h) Monthly Australian beer production; (i) Monthly Australian electricity production.

Consider the nine series plotted in Figure 1.6. Which of these do you think are stationary? Obvious seasonality rules out series (d), (h) and (i). Trend rules out series (a), (c), (e), (f) and (i). Increasing variance also rules out (i). That leaves only (b) and (g) as stationary series. At first glance, the strong cycles in series (g) might appear to make it non-stationary. But these cycles are aperiodic — they are caused when the lynx population becomes too large for the available feed, so they stop breeding and the population falls to very low numbers, then the regeneration of their food sources allows the population to grow again, and so on. In the long-term, the timing of these cycles is not predictable. Hence the series is stationary.

Differencing

In Figure 1.6, notice how the Dow Jones index data was non-stationary in panel (a), but the daily changes were stationary in panel (b). This shows one way to make a time series stationary — compute the differences between consecutive observations. This is known as differencing.
Transformations such as logarithms can help to stabilize the variance of a time series. Differencing can help stabilize the mean of a time series by removing changes in the level of a time series, and so eliminating trend and seasonality.
As well as looking at the time plot of the data, the ACF plot is also useful for identifying non-stationary time series. For a stationary time series, the ACF will drop to zero relatively quickly, while the ACF of non-stationary data decreases slowly. Also, for non-stationary data,

the value of r_1 is often large and positive.

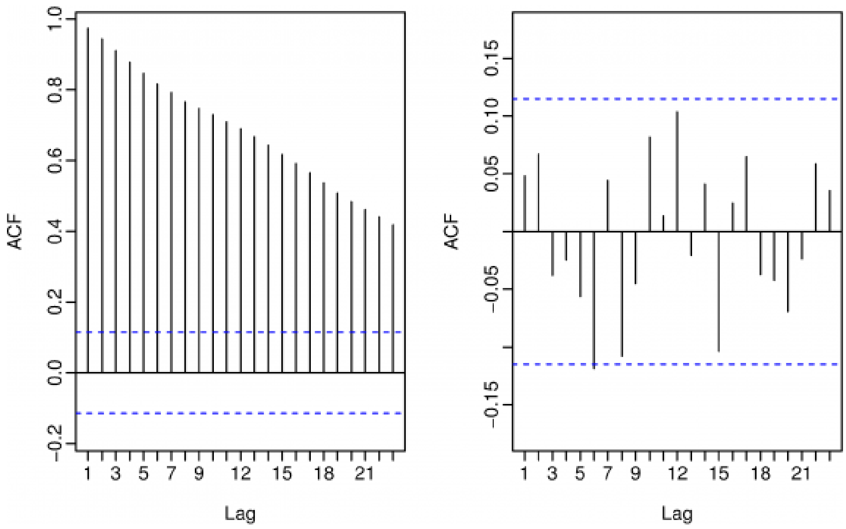


Figure 1.7: The ACF of the Dow-Jones index (left) and of the daily changes in the Dow-Jones index (right).

The ACF of the differenced Dow-Jones index looks just like that from a white noise series. There is only one autocorrelation lying just outside the 95% limits, and the Ljung-Box Q^* statistic has a p-value of 0.153 (for h=10). This suggests that the daily change in the Dow-Jones index is essentially a random amount uncorrelated with previous days.

Random walk model

The differenced series is the *change* between consecutive observations in the original series, and can be written as $y'_t = y_t - y_{t-1}$.
The differenced series will have only T-1 values since it is not possible to calculate a difference y'_1 for the first observation.
When the differenced series is white noise, the model for the original series can be written as $y_t - y_{t-1} = e_t$ or $y_t = y_{t-1} + e_t$ A random walk model is very widely used for non-stationary data, particularly finance and economic data. Random walks typically have:

- long periods of apparent trends up or down
- sudden and unpredictable changes in direction.

The forecasts from a random walk model are equal to the last observation, as future movements are unpredictable, and are equally likely to be up or down. Thus, the random walk model underpins naïve forecasts.
A closely related model allows the differences to have a non-zero mean. Then
$y_t - y_{t-1} = c + e_t$ or $y_t = c + y_{t-1} + e_t$ The value of c is the average of the changes between consecutive observations. If c is positive, then the average change is an increase in the value of y_t. Thus y_t will tend to drift upwards. But if c is negative, y_t will tend to drift downwards.

Second-order differencing

Occasionally the differenced data will not appear stationary and it may be necessary to difference the data a second time to obtain a stationary series:

$$\begin{aligned} y''_t &= y'_t - y'_{t-1} \\ &= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \\ &= y_t - 2y_{t-1} + y_{t-2}. \end{aligned}$$

In this case, y''_t will have $T-2$ values. Then we would model the "change in the changes" of the original data. In practice, it is almost never necessary to go beyond second-order differences.

Backshift notation

The backward shift operator **B** is a useful notational device when working with time series lags:

$$By_t = y_{t-1}$$

(Some references use **L** for "lag" instead of **B** for "backshift".) In other words, **B**, operating on y_t, has the effect of **shifting the data back one period**. Two applications of **B** to y_t **shifts the data back two periods**:

$$B(By_t) = B^2 y_t = y_{t-2}$$

For monthly data, if we wish to consider "the same month last year," the notation is

$$B^{12} y_t = y_{t-12}.$$

The backward shift operator is convenient for describing the process of differencing. A first difference can be written as

$$y'_t = y_t - y_{t-1} = y_t - By_t = (1-B)y_t.$$

Note that a first difference is represented by $(1-B)$. Similarly, if second-order differences have to be computed, then:

$$y''_t = y_t - 2y_{t-1} + y_{t-2} = (1 - 2B + B^2)y_t = (1-B)^2 y_t\ .$$

In general, a dth-order difference can be written as

$$(1-B)^d y_t.$$

Autoregressive models

In a multiple regression model, we forecast the variable of interest using a linear combination of predictors. In an autoregression model, we forecast the variable of interest using a linear combination of past values of the variable. The term autoregression indicates that it is a regression of the variable against itself.
Thus an autoregressive model of order pp can be written as

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t,$$

where c is a constant and e_t is white noise. This is like a multiple regression but with lagged values of y_t as predictors. We refer to this as an **AR($\boldsymbol{p}$) model**

Autoregressive models are remarkably flexible at handling a wide range of different time series patterns. The two series in Figure 1.3 show series from an AR(1) model and an AR(2) model. Changing the parameters $\phi_1, ..., \phi_p$ results in different time series patterns. The variance of the error term e_t will only change the scale of the series, not the patterns.

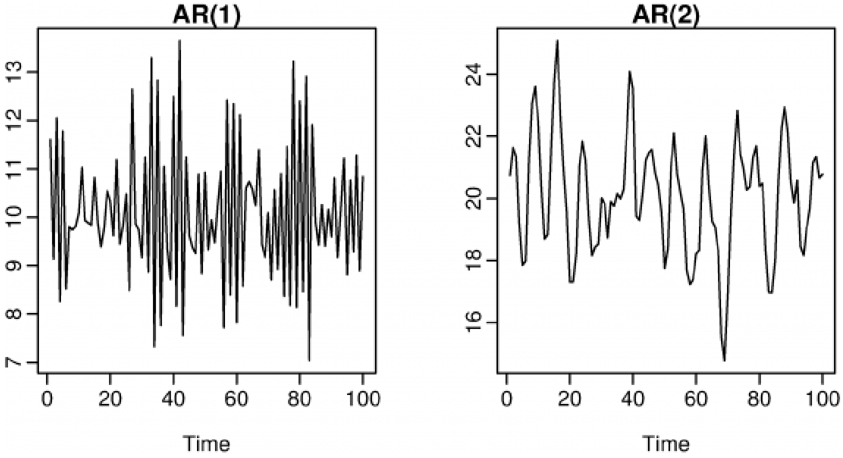


Figure 1.8: Two examples of data from autoregressive models with different parameters. Left: *AR(1)* with $y_t = 18 - 0.8y_{t-1} + e_t$ Right: *AR(2)* with $y_t = 8 + 1.3y_{t-1} - 0.7y_{t-2} + e_t$. In both cases, e_t is normally distributed white noise with mean zero and variance one.

For an AR(1) model:

- When $\phi_1 = 0$, y_t is equivalent to white noise
- When $\phi_1 = 1$ and $c = 0$, y_t is equivalent to a random walk
- When $\phi_1 = 1$ and $c \neq 0$, y_t is equivalent to a random walk with drift
- When $\phi_1 < 0$, y_t tends to oscillate between positive and negative values.

We normally restrict autoregressive models to stationary data, and then some constraints on the values of the parameters are required.

- For an *AR(1)* model:$-1 < \phi_1 < 1$
- For an *AR(2)* model:$-1 < \phi_2 < 1, \phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1$

When $p \geq 3$ the restrictions are much more complicated.

Moving average models

Rather than use past values of the forecast variable in a regression, a moving average model uses past forecast errors in a regression-like model.

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q},$$

where e_t is white noise. We refer to this as an $\mathbf{MA(q)}$ **model**. Of course, we do not observe the values of e_t, so it is not really regression in the usual sense.

Notice that each value of y_t can be thought of as a weighted moving average of the past few forecast errors. However, moving average models should not be confused with moving average smoothing.A moving average model is used for forecasting future values while moving average smoothing is used for estimating the trend-cycle of past values.

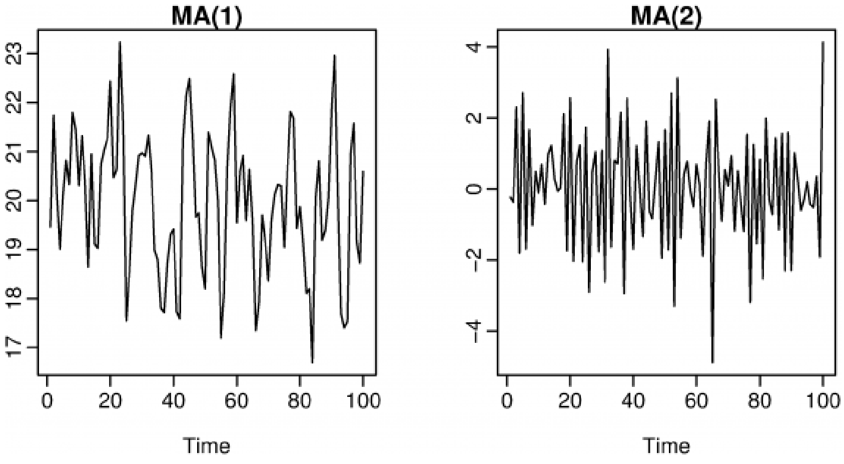


Figure 1.9: Two examples of data from moving average models with different parameters. Left: *MA(1)* with $y_t = 20 + 0.8e_{t-1} + e_t$ Right: *MA(2)* with $y_t = e_t - e_{t-1} + 0.8e_{t-2}$. In both cases, e_t is normally distributed white noise with mean zero and variance one.

Figure 1.9 shows some data from an MA(1) model and an MA(2) model. Changing the parameters $\theta_1, \ldots, \theta_q$ results in different time series patterns. As with autoregressive models, the variance of the error term e_t will only change the scale of the series, not the patterns.
It is possible to write any stationary AR(p) model as an MA(∞) model. For example, using repeated substitution, we can demonstrate this for an AR(1) model :

$$\begin{aligned} y_t &= \phi_1 y_{t-1} + e_t \\ &= \phi_1(\phi_1 y_{t-2} + e_{t-1}) + e_t \\ &= \phi_1^2 y_{t-2} + \phi_1 e_{t-1} + e_t \\ &= \phi_1^3 y_{t-3} + \phi_1^2 e_{t-2} + \phi_1 e_{t-1} + e_t \\ &\text{etc.} \end{aligned}$$

Provided $-1 < \phi_1 < 1$, the value of ϕ_1^k will get smaller as k gets larger. So eventually we obtain

$$y_t = e_t + \phi_1 e_{t-1} + \phi_1^2 e_{t-2} + \phi_1^3 e_{t-3} + \cdots ,$$

an MA(∞) process.
The reverse result holds if we impose some constraints on the MA parameters. Then the MA model is called "invertible". That is, that we can write any invertible MA(q) process as an AR(∞) process.
Invertible models are not simply to enable us to convert from MA models to AR models. They also have some mathematical properties that make them easier to use in practice.
The invertibility constraints are similar to the stationarity constraints.

- For an *MA(1)* model:$-1 < \theta_1 < 1$
- For an *MA(2)* model:$-1 < \theta_2 < 1, \theta_1 + \theta_2 > -1, \theta_1 - \theta_2 < 1$

When $q \geq 3$ the restrictions are much more complicated.

Non-seasonal ARIMA models

If we combine differencing with autoregression and a moving average model, we obtain a non-seasonal ARIMA model. ARIMA is an acronym for AutoRegressive Integrated Moving Average model ("integration" in this context is the reverse of differencing). The full model can be written as

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t, \tag{1.1}$$

where y'_t is the differenced series (it may have been differenced more than once). The "predictors" on the right hand side include both lagged values of y_t and lagged errors. We call this an **ARIMA(*p,d,q*) model**, where

p= order of the autoregressive part;
d= degree of first differencing involved;
q= order of the moving average part.

The same stationarity and invertibility conditions that are used for autoregressive and moving average models apply to this ARIMA model.

Once we start combining components in this way to form more complicated models, it is much easier to work with the backshift notation. Then equation (1.1) can be written as

$$\underset{\substack{\uparrow \\ \mathrm{AR}(p)}}{(1-\phi_1 B - \cdots - \phi_p B^p)} \quad \underset{\substack{\uparrow \\ d \text{ differences}}}{(1-B)^d y_t} \quad = \quad c + \underset{\substack{\uparrow \\ \mathrm{MA}(q)}}{(1+\theta_1 B + \cdots + \theta_q B^q) e_t}$$

Selecting appropriate values for ***p***, ***d*** and ***q*** can be difficult.

Many models are special cases of the ARIMA model as shown in the following table.

Table 1.2: Cases of the ARIMA model

White noise	ARIMA(0,0,0)
Random walk	ARIMA(0,1,0) with no constant
Random walk with drift	ARIMA(0,1,0) with a constant
Autoregression	ARIMA(p,0,0)
Moving average	ARIMA(0,0,q)

Estimation and order selection

1. **Maximum likelihood estimation**
   Once the model order has been identified (i.e., the values of ***p***, ***d*** and ***q***), we need to estimate the parameters c,$\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$. this technique finds the values of the parameters which maximize the probability of obtaining the data that we have observed. For ARIMA models,*maximum likelihood estimation* (MLE) is very similar to the *least squares* estimates that would be obtained by minimizing

   $$\sum_{t=1}^{T} e_t^2.$$

   Note that ARIMA models are much more complicated to estimate than regression models, and different softwares will give slightly different answers as they use different methods of estimation, or different estimation algorithms.

2. **Information Criteria**
   Akaike's Information Criterion (AIC) is useful for determining the order of an ARIMA

model. It can be written as

$$\text{AIC} = -2\log(L) + 2(p+q+k+1),$$

where $\boldsymbol{L}$ is the likelihood of the data,$k=1$ if $c \neq 0$ and $k=0$ if $c=0$.Note that the last term in parentheses is the number of parameters in the model (including σ^2, the variance of the residuals).
For ARIMA models, the corrected AIC can be written as

$$\text{AIC}_c = \text{AIC} + \frac{2(p+q+k+1)(p+q+k+2)}{T-p-q-k-2}.$$

and the Bayesian Information Criterion can be written as

$$\text{BIC} = \text{AIC} + (\log(T) - 2)(p+q+k+1).$$

Good models are obtained by minimizing either the AIC, AIC_c or BIC. Our preference is to use the AIC_c.

1.4.3 Exponential smoothing

Exponential smoothing was proposed in the late 1950s (Brown 1959, Holt 1957 and Winters 1960 are key pioneering works) and has motivated some of the most successful forecasting methods. Forecasts produced using exponential smoothing methods are weighted averages of past observations, with the weights decaying exponentially as the observations get older. In other words, the more recent the observation the higher the associated weight. This framework generates reliable forecasts quickly and for a wide spectrum of time series which is a great advantage and of major importance to applications in industry.

Simple exponential smoothing

For simple exponential smoothing,Forecasts are calculated using weighted averages where the weights decrease exponentially as observations come from further in the past,the smallest weights are associated with the oldest observations:

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha)y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + ...,$$

where $0 \leq \alpha \leq 1$ is the smoothing parameter. The one-step-ahead forecast for time T+1 is a weighted average of all the observations in the series y_1,...,y_T. The rate at which the weights decrease is controlled by the parameter α. For any α between 0 and 1, the weights attached to the observations decrease exponentially as we go back in time, hence the name "exponential smoothing". If α is small (i.e., close to 0), more weight is given to observations from the more distant past. If α is large (i.e., close to 1), more weight is given to the more recent observations. At the extreme case where $\alpha = 1$, $\hat{y}_{T+1|T} = y_T$ and forecasts are equal to the naïve forecasts.
Here are other forms of simple exponential smoothing:

- **Weighted average form**
  The forecast at time t+1 is equal to a weighted average between the most recent observation y_T and the most recent forecast $\hat{y}_{T+1|T}$,

$$\hat{y}_{T+1|T} = \alpha y_T + (1-\alpha)\hat{y}_{T|T-1}$$

for $t = 1, ..., T$, where $0 \leq \alpha \leq 1$ is the smoothing parameter. The process has to start somewhere, so we let the first forecast of y_1 be denoted by l_0. Then

$$\hat{y}_{2|1} = \alpha y_1 + (1-\alpha) l_0$$

$$\hat{y}_{3|2} = \alpha y_2 + (1-\alpha) \hat{y}_{2|1}$$

$$.$$

$$.$$

$$.$$

$$\hat{y}_{T+1|T} = \alpha y_T + (1-\alpha) \hat{y}_{T|T-1}$$

- **Component form**
  An alternative representation is the component form. For simple exponential smoothing the only component included is the level, l_t. (Other methods considered later in this chapter may also include a trend b_t and seasonal component s_t.) Component form representations of exponential smoothing methods comprise a forecast equation and a smoothing equation for each of the components included in the method. The component form of simple exponential smoothing is given by:
  Forecasting equation: $\hat{y}_{t+1|t} = l_t$
  Smoothing equation: $l_t = \alpha y_t + (1-\alpha) l_{t-1}$
  where l_t is the level (or the smoothed value) of the series at time t. The forecast equation shows that the forecasted value at time t+1 is the estimated level at time t. The smoothing equation for the level (usually referred to as the level equation) gives the estimated level of the series at each period t. Applying the forecast equation for time T gives, $\hat{y}_{T+1|T} = l_T$,the most recent estimated level. If we replace l_t by $\hat{y}_{t+1|t}$ and l_{t-1} by $\hat{y}_{t|t-1}$ in the smoothing equation, we will recover the weighted average form of simple exponential smoothing.

- **Error correction form**
  The third form of simple exponential smoothing is obtained by re-arranging the level equation in the component form to get what we refer to as the error correction form

$$\begin{aligned} l_t &= l_{t-1} + \alpha(y_t - l_{t-1}) \\ &= l_{t-1} + \alpha e_t \end{aligned}$$

  where $e_t = y_t - l_{t-1} = y_t - \hat{y}_{t|t-1}$ for $t = 1, ..., T$ That is, e_t is the one-step within-sample forecast error at time t. The within-sample forecast errors lead to the adjustment/correction of the estimated level throughout the smoothing process for $t = 1, ..., T$. For example, if the error at time t is negative, then $\hat{y}_{t|t-1} > y_t$ and so the level at time t−1 has been over-estimated. The new level l_t is then the previous level l_{t-1} adjusted downwards. The closer α is to one the "rougher" the estimate of the level (large adjustments take place). The smaller the α the "smoother" the level (small adjustments take place).

1.4.4 ARIMA vs ETS

It is a common myth that ARIMA models are more general than exponential smoothing. While linear exponential smoothing models are all special cases of ARIMA models, the non-linear exponential smoothing models have no equivalent ARIMA counterparts. There are also many ARIMA models that have no exponential smoothing counterparts. In particular, every ETS model is non-stationary, while ARIMA models can be stationary.

The ETS models with seasonality or non-damped trend or both have two unit roots (i.e., they need two levels of differencing to make them stationary). All other ETS models have one unit root (they need one level of differencing to make them stationary).

The following table gives some equivalence relationships for the two classes of models.

Table 1.3: classes of ETS and ARIMA models

ETS model	ARIMA model	Parameters
ETS(A,N,N)	ARIMA(0,1,1)	$\theta_1 = \alpha - 1$
ETS(A,A,N)	ARIMA(0,2,2)	$\theta_1 = \alpha + \beta - 2; \quad \theta_2 = 1 - \alpha$
ETS(A,A$_d$,N)	ARIMA(1,1,2)	$\phi_1 = \phi$; $\theta_1 = \alpha + \phi\beta - 1 - \phi$; $\theta_2 = (1-\alpha)\phi$
ETS(A,N,A)	ARIMA(0,0,m)$(0,1,0)_m$	
ETS(A,A,A)	ARIMA(0,0,m+1)$(0,1,0)_m$	
ETS(A,A_d,A)	ARIMA(1,0,$m+1)(0,1,0)_m$	

1.5 Problem statement

World currencies are being traded everyday against one another to the tune of trillions of dollars per day. Through this trading each currency is fixed at a particular level and measured against the other by an exchange rate. In most developing countries,like Rwanda ,exchange rates against the most developed countries, like the United States of America,have tendencies to depreciate. So it is wise for those interested to start their ventures in those countries, developing countries, to know how the currency might behave in the coming days or even years to start planning and implementing their ideas and visions.

The question that arises is: what causes exchange rates to change, and how does one predict future value?

Time series analysis involves both model identification and parameter estimation. Most analyses would agree the identification problem is more difficult. Once the functional form of a model is specified, estimating the model parameters is usually straightforward. To identify a model that best represents a time series, it is necessary to be clear about the purpose of the model. Is the model's chief objective to explain the nature of the system generating the series? Or, is the model to be judged on its ability to predict future values of the time series? Therefore, to arrive at a model that represents only the main features of the series, a selection criterion, which balances model, fit, and model complexity, must be used.

After all of this analysis we have to forecast the exchange rates.

1.6 Objective of the study

1.6.1 General objective

The general objective of this study is to choose appropriate time series model that best forecasts the exchange rates of ***US dollar*** to ***RWF*** so that we can make conclusion whether it is ideal for foreign investors, especially coming from the United States of America, to strongly consider a long term, or even a short term, investment in Rwanda and if the profit will be of a great significance in the future time.

1.6.2 Specific objectives

- Find the best model for our exchange rates data,
- Forecast the exchange rates data from January 1^{st} 2015 to April 30^{th} 2018,
- Analyze the forecasted exchange rates data and suggest whether it is appropriate or not to consider investing in Rwanda.

1.7 The limitation of the study

In this study we will analyze exchange rates data obtained from National Bank of Rwanda. The data that we are interested in are from January 1^{st} 2015 to April 30^{th} 2018. We will forecast 30 days of the month of May 2018.

1.8 Justification of the study

It is important for many businesses and Investors to analyze and to know how the currency of the country of their interest is performing with respect to the currency of their local country. This study will be interested in finding the best technical model to the exchange rates of USD vs RWF and forecast the 30 days of the month of May 2018.

Chapter 2

Literature review

For a country to be involved in international trade, finance, and investment, it is necessary to have access to foreign currencies of other countries. The sale and purchase of foreign currencies take place in the foreign exchange markets. This market allows for the movement of large volumes of funds for investment purposes around the world. Any changes in exchange rates are important because of the effect they have on the prices we pay for imports, the prices we receive for our exports, and the amount of money flowing into and out of the economy.
Exchange rate fluctuations play a key role in determining economic policy. These fluctuations have repercussions on economic performances. It is essentially the dependence with respect to imports and specialization in exports that account for exchange rate fluctuations on the economic performances of countries. In order to stabilize the economy during these fluctuations, government may increase or decrease money supplies, which, in turn, can weaken or strengthen the price of the exchange rate.
In fundamental models of exchange rate, macroeconomic variables such as interest rates, money supplies, gross domestic products, trade account balances, and commodity prices have long been perceived as the determinants of the equilibrium exchange rate. The foreign exchange rate in fundamental models is classified as a highly liquid market where all information is public, and traders in the market share the same expectations with no information advantage over the other.
After more than two decades of research since Meese and Rogoff's seminal work on exchange rates predictability (see Meese and Rogoff, 1983a, 1983b), the goal of exploiting foreign exchange rates forecasting model to beat naïve random walk forecasts remains as elusive as ever (Taylor, 1995) due to the fact that evidence supporting or refuting the exchange rate predictability seems plausible. For example, Bekaert and Hodrick (1992), Fong and Ouliaris (1995), LeBaron (1999), Levich and Thomas (1993), Liu and He (1991), McCurdy and Morgan (1988), Baillie and Bollerslev (1989), Sweeney (1986) and Soofi et al. (2006) found evidence contrary to the martingale (random walk or pure unit-root) hypothesis for nominal or real exchange rates, indicating that exchange rates are predictable. While Diebold and Nason (1990), Fong, Koh and Ouliaris (1997), Hsieh (1988, 1989, 1993), McCurdy and Morgan (1987),and Meese and Rogoff (1983a, 1983b) found little evidence against the martingale (random walk or pure unit-root) hypothesis for nominal or real exchange rates, implying that predictability of exchange rates is impossible. One simple and possible explanation is that traditional exchange rate forecasting models are inadequate. Due to the fact that exchange rate forecasting is of practical as well as theoretical importance, a large number of methods and techniques (including linear and nonlinear) were introduced to beat random walk model in foreign exchange markets.

Chapter 3

Research Methodology

Since the influential paper by Meese and Rogoff (1983), the random walk (RW) model remains the standard metric by which to judge the forecast accuracy of an exchange rate model.but for us we are going to use other time series models and analyze their accuracies so that we can identify which one makes good predictions.
This chapter presents the methods and model estimation techniques and types of data used in this study to forecast exchange rates, $\boldsymbol{USD}$ to $\boldsymbol{RWF}$ rates, for the given period of time.

3.1 Data source and Data collection

Data used in this research are secondary data provided by National Bank of Rwanda. The data constitute daily $\boldsymbol{USD}$ to $\boldsymbol{RWF}$ rates, buying rates, selling rates and their average rates.we are interested in the average rates. The data collected are then analyzed, modeled and interpreted, and forecasted for the given time in the future.

3.2 Sample size

During our research, we chose to use data from January 2^{nd},2015 to April 30^{th},2018. the sample size is 825 observations.

3.3 Methodology of data collection

There are many methods of collecting data such as interviewing, questionnaire, Internet, literature, and others with the purpose of getting primary data or secondary data. the primary data are to be originally collected by investigator, while secondary data are data which have been already collected by someone else and which have been passed through statistical process such as compilation. The later are used in this project.
The $\boldsymbol{USD} \backslash \boldsymbol{RWF}$ exchange rates data was released to the public by National Bank of Rwanda (BNR) through the Internet, by posting the data in the archive section to their website. So we got the data from their official website, website of (BNR).

3.4 Methodology of forecasting

As we have seen, forecasting can be done using different approaches, depending on your target and resources.
In this chapter we will explain the methods we chose and how they will be applied to our data. The methods chosen are all applied to time series data only.
The specific objective is to choose appropriate time series method\model that best describe daily exchange rates.
The first step is to plot daily exchange rates data against time and see whether the plot shows trend or not, and hence, with its behavior, one could determine which type of model approach would be appropriate. Then the series, if necessary,is transformed to induce stationarity.
For our exchange rates data, we chose to use between Exponential smoothing approach, ARIMA(p,d,q) and GARCH(p,q) model when forecasting.depending on the accuracy of each model's forecast.

3.4.1 Stationarity test

For us to find the best fit ARIMA(p,d,q) model we have to test whether our data are stationary or not.

Dicker-Fuller test

We know the statistical basis for our estimation and forecasting depends on series being covariance stationary. Remember that for covariance stationarity, we said all roots of the autoregressive lag polynomial must be less than 1.
for example,

$$y_t = y_{t-1} + \mathcal{E}_t \tag{3.1}$$
$$\mathcal{E}_t = \mathcal{N}(0, \sigma^2) \tag{3.2}$$

Because the autoregressive lag polynomial has one root equal to one, we say it has a **unit root**. The early and pioneering work on testing for a unit root in time series was done by Dickey and Fuller (Dickey and Fuller 1979, Fuller 1976). The basic objective of the test is to test the null hypothesis that $\phi = 1$ in:

$$y_t = \phi y_{t-1} + \mathcal{E}_t$$

against the one-sided alternative $\phi < 1$. Hence we have H_0: series contains a unit root vs. H_1: series is stationary.
We usually use the regression: $\Delta y_t = \psi y_{t-1} + \mathcal{E}_t$ so that a test of $\phi = 1$ is equivalent to a test of $\psi = 1$ (since $\phi = 1 - \psi$).

- **Different forms for the DF Test Regressions**
  Dickey Fuller tests are also known as τ tests: τ; τ_μ; τ_τ . The null (H_0) and alternative (H_1) models in each case are:

  1.

  $$H_0 : y_t = y_{t-1} + \mathcal{E}_t \tag{3.3}$$
  $$H_1 : y_t = \phi y_{t-1} + \mathcal{E}_t, \phi < 1 \tag{3.4}$$

  This is a test for a random walk against a stationary autoregressive process of order one (AR(1))

2.

$$H_0 : y_t = y_{t-1} + \mathcal{E}_t \tag{3.5}$$
$$H_1 : y_t = \phi y_{t-1} + \mu + \mathcal{E}_t, \phi < 1 \tag{3.6}$$

This is a test for a random walk against a stationary AR(1) with drift.

3.

$$H_0 : y_t = y_{t-1} + \mathcal{E}_t \tag{3.7}$$
$$H_1 : y_t = \phi y_{t-1} + \mu + \lambda t + \mathcal{E}_t, \phi < 1 \tag{3.8}$$

This is a test for a random walk against a stationary AR(1) with drift and a time trend.

- **Computing the DF Test Statistic**
  We can write $\Delta y_t = \mathcal{E}_t$ where $\Delta y_t = y_t - y_{t-1}$, and the alternatives may be expressed as $y_t = \psi y_{t-1} + \mu + \lambda t + \mathcal{E}_t$ with $\mu = \lambda = 0$ in case 1), and $\lambda = 0$ in case 2) and $\psi = \phi - 1$. In each case, the tests are based on the t-ratio on the y_{t-1} term in the estimated regression of Δy_t on y_{t-1}, plus a constant in case 2) and a constant and trend in case 3). The test statistics are defined as test statistic$=\frac{\hat{\psi}}{SE_{\psi}}$ The test statistic does not follow the usual t-distribution under the null, since the null is one of non-stationarity, but rather follows a non-standard distribution. Critical values are derived from Monte Carlo experiments in, for example, Fuller (1976). Relevant examples of the distribution are shown in table below. table below shows Critical values for DF and ADF tests (Fuller 1976, p373).

Table 3.1: Critical Values for the DF Test

Significance level	10%	5%	1%
C.V. for constant but no trend	-2.57	-2.86	-3.43
C.V. for constant and trend	-3.12	-3.41	-3.96

The null hypothesis of a unit root is rejected in favor of the stationary alternative in each case if the test statistic is more negative than the critical value.

Augmented Dicker-Fuller test

The DF test only valid if $\mathcal{E}_t$ is white noise. In particular, $\mathcal{E}_t$ will be autocorrelated if there was autocorrelation in the dependent variable of the regression Δy_t which we have not modeled. The solution is to augment the test using p lags of the dependent variable. The alternative model in case (i) is now written:

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + \mathcal{E}_t$$

The same critical values from the DF tables are used as before. A problem now arises in determining the optimal number of lags of the dependent variable. There are 2 ways:

- use the frequency of the data to decide
- use information criteria.

3.4.2 Evaluating forecast accuracy

Forecast accuracy measures

Let y_i denote the i^{th} observation and $\hat{y}_i$ denote a forecast of y_i.

1. **Scale-dependent errors**
   The forecast error is simply $e_i = y_i - \hat{y}_i$, which is on the same scale as the data. Accuracy measures that are based on e_i are therefore scale-dependent and cannot be used to make comparisons between series that are on different scales.
   The two most commonly used scale-dependent measures are based on the absolute errors or squared errors:

$$\text{Mean absolute error: MAE} = \text{mean}(|e_i|),$$

$$\text{Root mean squared error: RMSE} = \sqrt{\text{mean}(e_i^2)}.$$

   When comparing forecast methods on a single data set, the MAE is popular as it is easy to understand and compute.

2. **Percentage errors**
   The percentage error is given by $p_i = 100e_i/y_i$. Percentage errors have the advantage of being scale-independent, and so are frequently used to compare forecast performance between different data sets. The most commonly used measure is:

$$\text{Mean absolute percentage error: MAPE} = \text{mean}(|p_i|).$$

   Measures based on percentage errors have the disadvantage of being infinite or undefined if $y_i = 0$ for any i in the period of interest, and having extreme values when any y_i is close to zero. Another problem with percentage errors that is often overlooked is that they assume a meaningful zero. For example, a percentage error makes no sense when measuring the accuracy of temperature forecasts on the Fahrenheit or Celsius scales.

   They also have the disadvantage that they put a heavier penalty on negative errors than on positive errors. This observation led to the use of the so-called "symmetric" MAPE (sMAPE) proposed by Armstrong (1985, p.348), which was used in the M3 forecasting competition. It is defined by

$$\text{sMAPE} = \text{mean}\left(200|y_i - \hat{y}_i|/(y_i + \hat{y}_i)\right).$$

   However, if y_i is close to zero, $\hat{y}_i$ is also likely to be close to zero. Thus, the measure still involves division by a number close to zero, making the calculation unstable. Also, the value of sMAPE can be negative, so it is not really a measure of "absolute percentage errors" at all.

   Hyndman and Koehler (2006) recommend that the sMAPE not be used.

3. **Scaled errors**
   Scaled errors were proposed by Hyndman and Koehler (2006) as an alternative to using percentage errors when comparing forecast accuracy across series on different scales. They

proposed scaling the errors based on the training MAE from a simple forecast method. For a non-seasonal time series, a useful way to define a scaled error uses naïve forecasts:

$$q_j = \frac{e_j}{\frac{1}{T-1}\sum_{t=2}^{T}|y_t - y_{t-1}|}.$$

Because the numerator and denominator both involve values on the scale of the original data, q_J is independent of the scale of the data. A scaled error is less than one if it arises from a better forecast than the average naïve forecast computed on the training data. Conversely, it is greater than one if the forecast is worse than the average naïve forecast computed on the training data. For seasonal time series, a scaled error can be defined using seasonal naïve forecasts:

$$q_j = \frac{e_j}{\frac{1}{T-m}\sum_{t=m+1}^{T}|y_t - y_{t-m}|}.$$

For cross-sectional data, a scaled error can be defined as

$$q_j = \frac{e_j}{\frac{1}{N}\sum_{i=1}^{N}|y_i - \bar{y}|}.$$

In this case, the comparison is with the mean forecast. (This doesn't work so well for time series data as there may be trends and other patterns in the data, making the mean a poor comparison. Hence, the naïve forecast is recommended when using time series data.)

The *mean absolute scaled error* is simply

$$\text{MASE} = \text{mean}(|q_j|).$$

Similarly, the mean squared scaled error (MSSE) can be defined where the errors (on the training data and test data) are squared instead of using absolute values.

Chapter 4

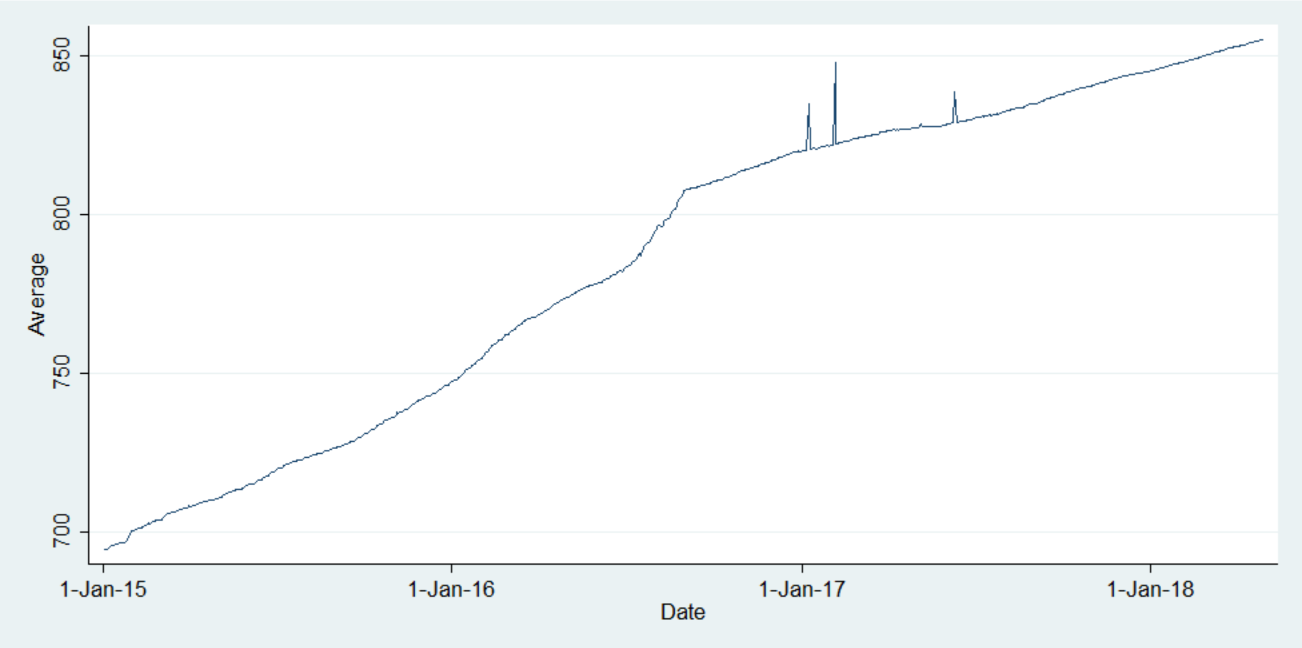
RESULTS AND ANALYSIS

In this chapter, the data collected from National Bank of Rwanda is analyzed and then interpreted. The results are illustrated using tables and graphs for a better understanding.

4.1 Data description

Exchange rates data are known to be volatile. They have many fluctuations, so it is hard to make a very sound prediction on how they will behave in the coming days.
The data we chose to work with, given how the simple size is large, have a strong volatility as we will see in figure 4.1

Figure 4.1: A plot of Exchange rates of **US dollar** against **Rwandan franc** from January 2^{nd},2015 to April 30^{th},2018.



From figure 4.1 we can automatically detect the trend showing how the US dollar has appreciated against the Rwandan franc over the past three years.
the first thing to do is to determine which model can best fit this type of data we have so that we can make good prediction.

As we have said, before making forecasts we will have to determine the best fit model to our type of data using AIC and\or BIC.

4.2 Stationarity testing

From figure 4.1 we can confirm with our naked eye that our time series data are not stationary. but, being technical, let us run some test to confirm with certainty that we have non-stationary time series data

4.2.1 Augmented Dicker-Fuller test (Adf test)

Using R, we run an Adf test to test whether our data were stationary or non-stationary, and the result we got are described in figure 4.2

Figure 4.2: Augmented Dicker-Fuller test on our exchange rates data, using R

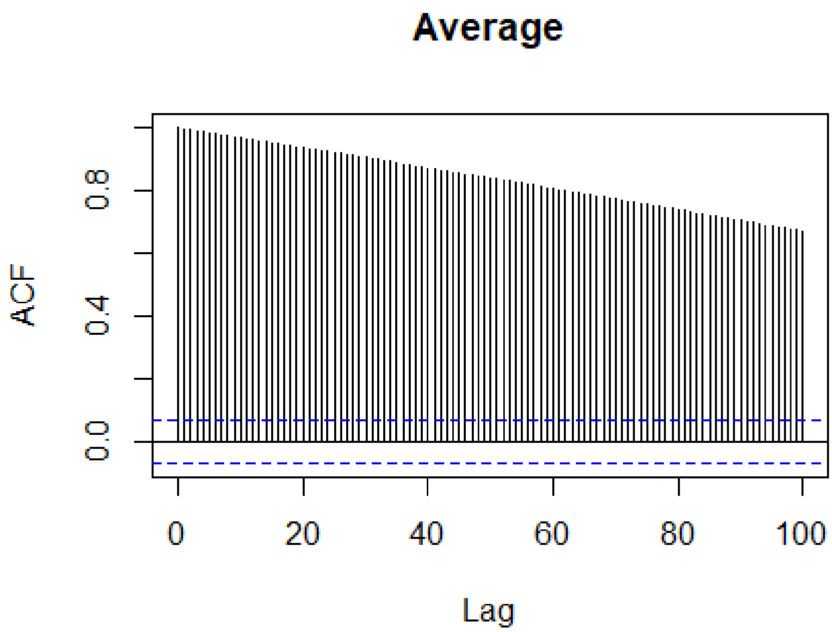
```
        Augmented Dickey-Fuller Test

data:  exchangerates.ts
Dickey-Fuller = -1.9128, Lag order = 0, p-value = 0.6152
alternative hypothesis: stationary
```

From the analysis described in figure 4.2, we see that there is a probability of 0.6152 that our time series data have a unit root, which implies that they are non stationary, which is greater than 0.05. so we accept the null hypothesis that states that there is a unit root, hence our time series are non-stationary.

4.2.2 Autocorrelation functions

These functions can also become handy in determining the stationarity of our data. let us check the correlograms of our data using R

Figure 4.3: Autocorrelation function of exchange rates time series data, using R



From figure 4.3 it is clear that from all lags not a single spike falls between the confident limits. so we conclude that our data are non-stationary

4.3 Forecasting models

In this part, we are going to analyze models that best fit our time series data.
We are interested in the best models to use while dealing with volatile financial data such as Exchange rates. Hence we chose to consider only three models, which are: Exponential smoothing;ARIMA(p,d,q) models and GARCH(p,q) models.
Let us analyze how each of this fit our Exchange rates time series data.

4.3.1 Exponential smoothing

Using R, we used Holt-Winter's exponential smoothing with trend and without seasonal component command to create a model for our data, as seen in figure 4.4

Figure 4.4: Holt-Winter's exponential smoothing with trend and without seasonal component, using R

```
Holt-Winters exponential smoothing with trend and without seasonal component

Call:
HoltWinters(x = exchangerates.ts, gamma = F)

Smoothing parameters:
 alpha: 0.2264808
 beta : 0.08633584
 gamma: FALSE

Coefficients:
         [,1]
a 855.0117952
b   0.1165015
```

Where *alpha* (α) is the smoothing parameter for the level,$0 \leq \alpha \leq 1$ and *beta* (β) is the smoothing parameter for the trend, $0 \leq \beta \leq 1$.
a and b are given by the following formulas:

$$\begin{aligned} \text{Forecast equation} \qquad & \hat{y}_{t+h|t} = a_t + hb_t \\ \text{Level equation} \qquad & a_t = \alpha y_t + (1-\alpha)(a_{t-1} + b_{t-1}) \\ \text{Trend equation} \qquad & b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1} \end{aligned}$$

4.3.2 ARIMA(p,d,q) models

After analyzing the time series data, we tried to figure out which ARIMA(p,d,q) model best fitted our data. Using R command, auto.arima, we got this model presented in figure 4.5

Figure 4.5: Autoregressive Integrated Moving Average model, using R

```
ARIMA(1,2,3)

Coefficients:
         ar1      ma1     ma2      ma3
      0.6695  -2.4289  1.9478  -0.5127
s.e.  0.1414   0.1398  0.2503   0.1134

sigma^2 estimated as 1.511:  log likelihood=-1338.23
AIC=2686.46   AICc=2686.53   BIC=2710.02
```

Figure 4.5 can be interpreted as follow:

$$(1-B)^2\hat{y}_t = 0.6695(1-B)^2 y_{t-1} + 2.4289e_{t-1} - 1.9478e_{t-2} + 0.5127e_{t-3} + \mathcal{E}_t$$

where, B is called Backshift Operator;

$$By_t = y_{t-1}$$

We can even check the residues to see if our model's residues are normally distributed. From the following figures we can see that they are all distributed around mean zero.

Figure 4.6: Plot of residuals of ARIMA(1,2,3) model, using Stata

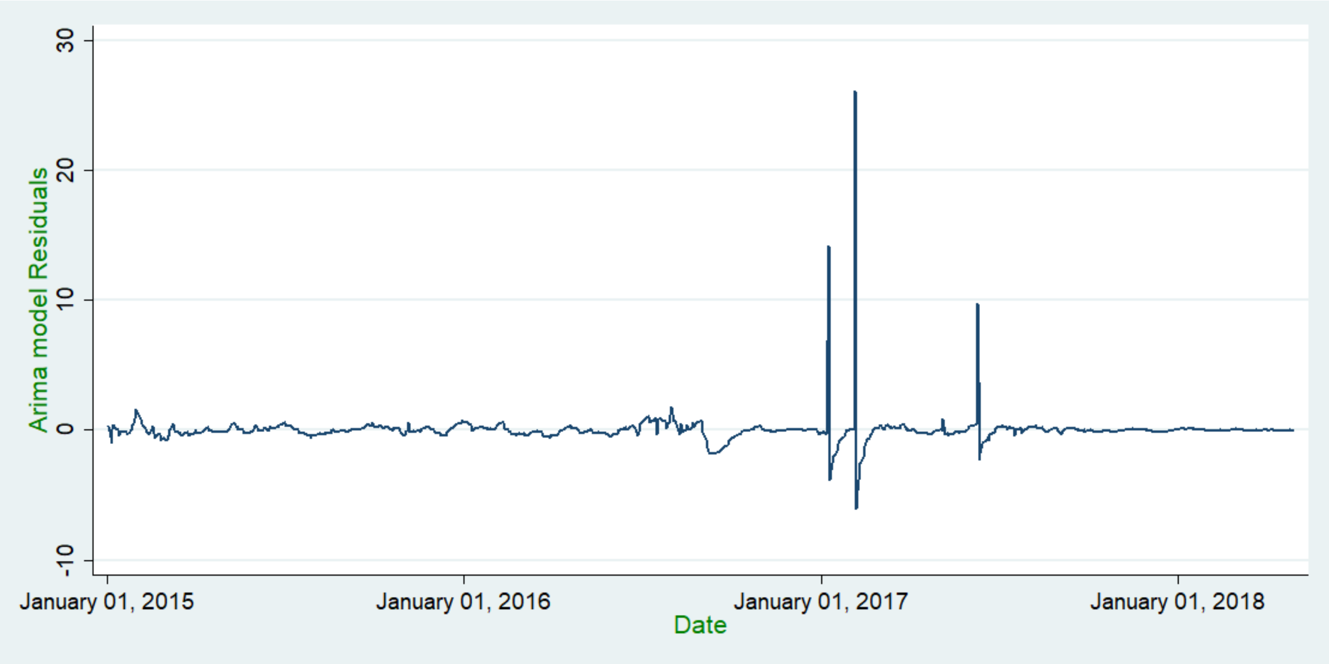
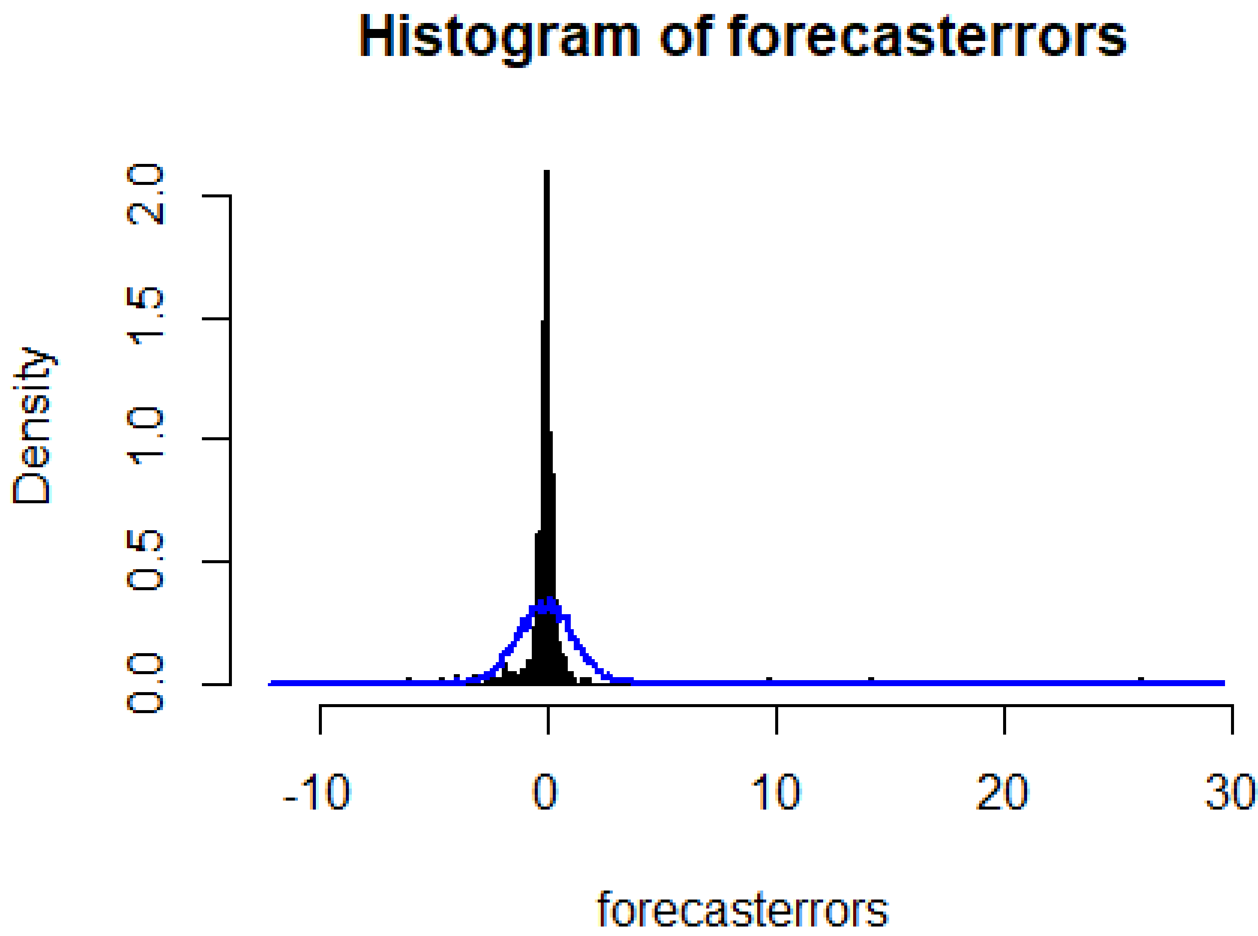


Figure 4.7: Histogram of residuals of ARIMA(1,2,3) model, using R



4.3.3 GARCH(p,q) models

This is the most useful model when dealing with financial time series data. We applied it to our data and it gave us a model of the form:

Figure 4.8: Generalized Autoregressive Conditional Heteroscedasticity, using R

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error    t value Pr(>|t|)
mu     826.740106    0.760145 1087.60793 0.000000
ma1      0.849980    0.014165   60.00527 0.000000
omega    0.006627    0.006831    0.97012 0.331988
alpha1   0.440491    0.064535    6.82562 0.000000
beta1    0.558509    0.064891    8.60683 0.000000
shape   26.569778    6.249741    4.25134 0.000021

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu     826.740106   14.668019 56.36344 0.000000
ma1      0.849980    0.035379 24.02490 0.000000
omega    0.006627    0.007840  0.84527 0.397962
alpha1   0.440491    0.104886  4.19971 0.000027
beta1    0.558509    0.100202  5.57384 0.000000
shape   26.569778   26.611763  0.99842 0.318075

LogLikelihood : -3399.456

Information Criteria
------------------------------------

Akaike       8.2557
Bayes        8.2899
Shibata      8.2555
Hannan-Quinn 8.2688
```

Which can be interpreted as follow

$$
\begin{aligned}
y_t &= \theta_1 e_{t-1} - \mu + e_t, \\
e_t &= \sigma_t \varepsilon_t, \\
\sigma_t^2 &= \omega + \alpha_1 e_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \\
\varepsilon_t &\sim i.i.d(0, 1, \gamma)
\end{aligned}
$$

Hence,our fitted model sGARCH(1,1)-ARFIMA(0,0,1) becomes

$$
\begin{aligned}
y_t &= 0.84998e_{t-1} - 826.740106 + e_t, \\
e_t &= \sigma_t \varepsilon_t, \\
\sigma_t^2 &= 0.006627 + 0.440491e_{t-1}^2 + 0.558509\sigma_{t-1}^2, \\
\varepsilon_t &\sim i.i.d(0, 1, \gamma)
\end{aligned}
$$

4.4 Choosing the best model

From the above three models, we can now analyze the fitness of each model, ARIMA(p,d,q) and GARCH(p,q) using the results of AIC and BIC.
For the Holt-Winter's model, we will check the accuracy of its forecasts and then compare it to the others.
We compared the Akaike Information Criteria,AIC, and the Bayesian Information Criteria of the ARIMA(1,2,3) model and GARCH(1,1) model;

Table 4.1: Akaike Information Criteria,AIC, and the Bayesian Information Criteria,BIC, of the ARIMA(1,2,3) model and GARCH(1,1) model

	AIC	BIC
ARIMA(1,2,3) model	2686.46	2710.02
GARCH(1,1) model	8.2557	8.2899

From the values described in the table we conclude that the best fit model is GARCH(1,1) model.

4.4.1 Forecasting Accuracy

After forecasting using different models, we tried to analyze their errors to check the ones with fewer errors. The models that were analyzed are the Holt-Winter exponential smoothing and the ARIMA(p,d,q) model in the following figures.

Figure 4.9: Accuracy of the Holt-Winters exponential smoothing, using R

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	0.0708545	0.111007	0.1078655	0.008284275	0.01261312	1.000021	-0.002954092

Figure 4.10: Accuracy of the ARIMA(1,2,3), using R

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-0.008205059	1.224705	0.3526752	-0.001106601	0.0444608	1.061147	-0.0116247

From the above 2 figures, we can see that the most accurate is the Holt-Winter's exponential smoothing.
With all this analysis now we understand how each model will execute the command of forecasting.

4.5 Forecasting Exchange rates Between US dollar (USD) and Rwandan franc of the month May,2018

With all the three models, we managed to make different forecasts for the month of May 2018. here are the three forecasts gotten from the three models.

4.5.1 Forecasts from HoltWinters

Here is the figures showing results we got when we forecasted using HoltWinters exponential smoothing method.

Figure 4.11: Forecasts from HoltWinters, using R

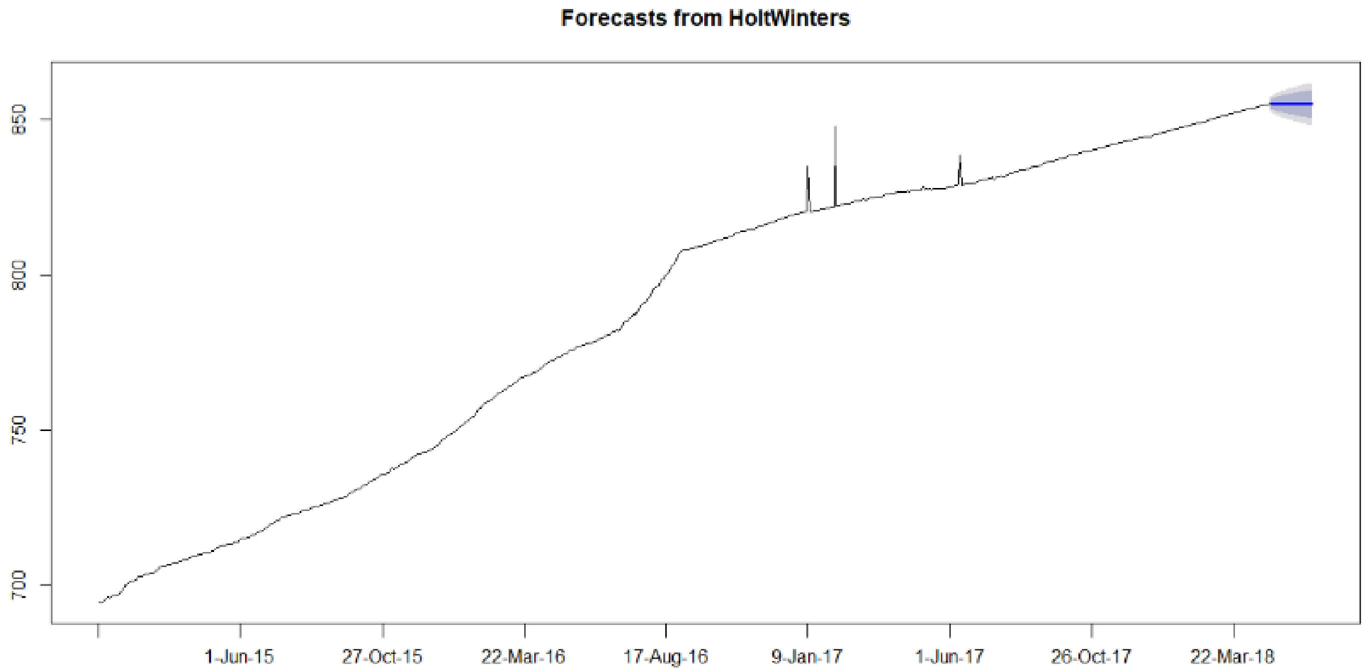


Figure 4.12: The forecast results using HoltWinters exponential smoothing, using R

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
826	855.1283	853.5446	856.7120	852.7063	857.5503
827	855.2448	853.6139	856.8757	852.7505	857.7391
828	855.3613	853.6770	857.0456	852.7854	857.9372
829	855.4778	853.7340	857.2216	852.8109	858.1447
830	855.5943	853.7850	857.4036	852.8272	858.3614
831	855.7108	853.8300	857.5916	852.8344	858.5872
832	855.8273	853.8693	857.7853	852.8328	858.8218
833	855.9438	853.9030	857.9846	852.8227	859.0649
834	856.0603	853.9313	858.1893	852.8043	859.3163
835	856.1768	853.9545	858.3992	852.7780	859.5756
836	856.2933	853.9726	858.6140	852.7442	859.8425
837	856.4098	853.9861	858.8336	852.7030	860.1166
838	856.5263	853.9949	859.0577	852.6549	860.3977
839	856.6428	853.9995	859.2862	852.6002	860.6855
840	856.7593	853.9998	859.5188	852.5390	860.9796
841	856.8758	853.9962	859.7555	852.4717	861.2799
842	856.9923	853.9886	859.9960	852.3986	861.5860
843	857.1088	853.9775	860.2402	852.3198	861.8978
844	857.2253	853.9627	860.4879	852.2356	862.2151
845	857.3418	853.9445	860.7391	852.1461	862.5375
846	857.4583	853.9231	860.9936	852.0516	862.8650
847	857.5748	853.8984	861.2512	851.9522	863.1974
848	857.6913	853.8707	861.5120	851.8481	863.5345
849	857.8078	853.8399	861.7757	851.7394	863.8762
850	857.9243	853.8063	862.0424	851.6263	864.2224
851	858.0408	853.7698	862.3119	851.5088	864.5728
852	858.1573	853.7306	862.5841	851.3872	864.9275
853	858.2738	853.6887	862.8590	851.2614	865.2862
854	858.3903	853.6442	863.1365	851.1317	865.6490
855	858.5068	853.5971	863.4166	850.9981	866.0156

4.5.2 Forecasts from ARIMA(p,d,q) model

From the forecasting models section, we saw that the best ARIMA (p,d,q) model to our exchange rates data is ARIMA(p,d,q).
After knowing the order of or ARIMA model, we went on and forecasted for the coming 30 days, which can be described as the month of May 2018. below are the figures representing the plot and the values of our forecasted values.

Figure 4.13: Plot of the forecast results using ARIMA (1,2,3) model, using R

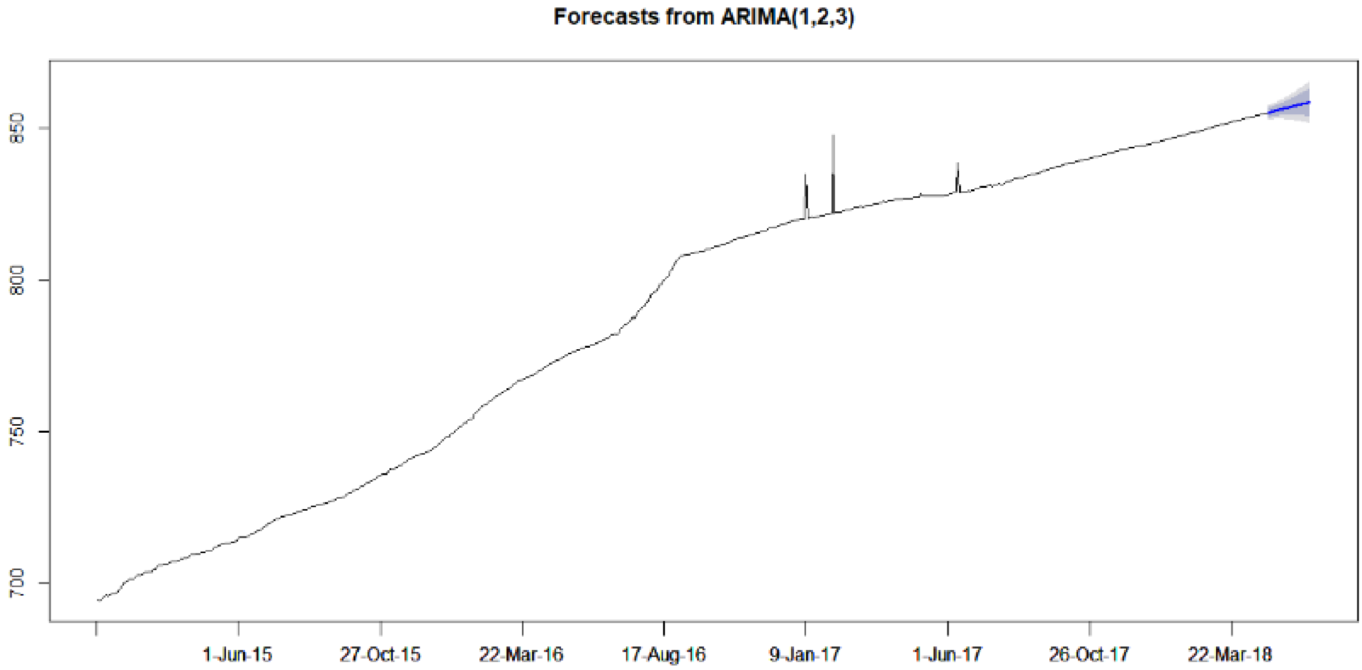


Figure 4.14: Values of the forecast results,for the month of May 2018, using ARIMA (1,2,3) model, using R

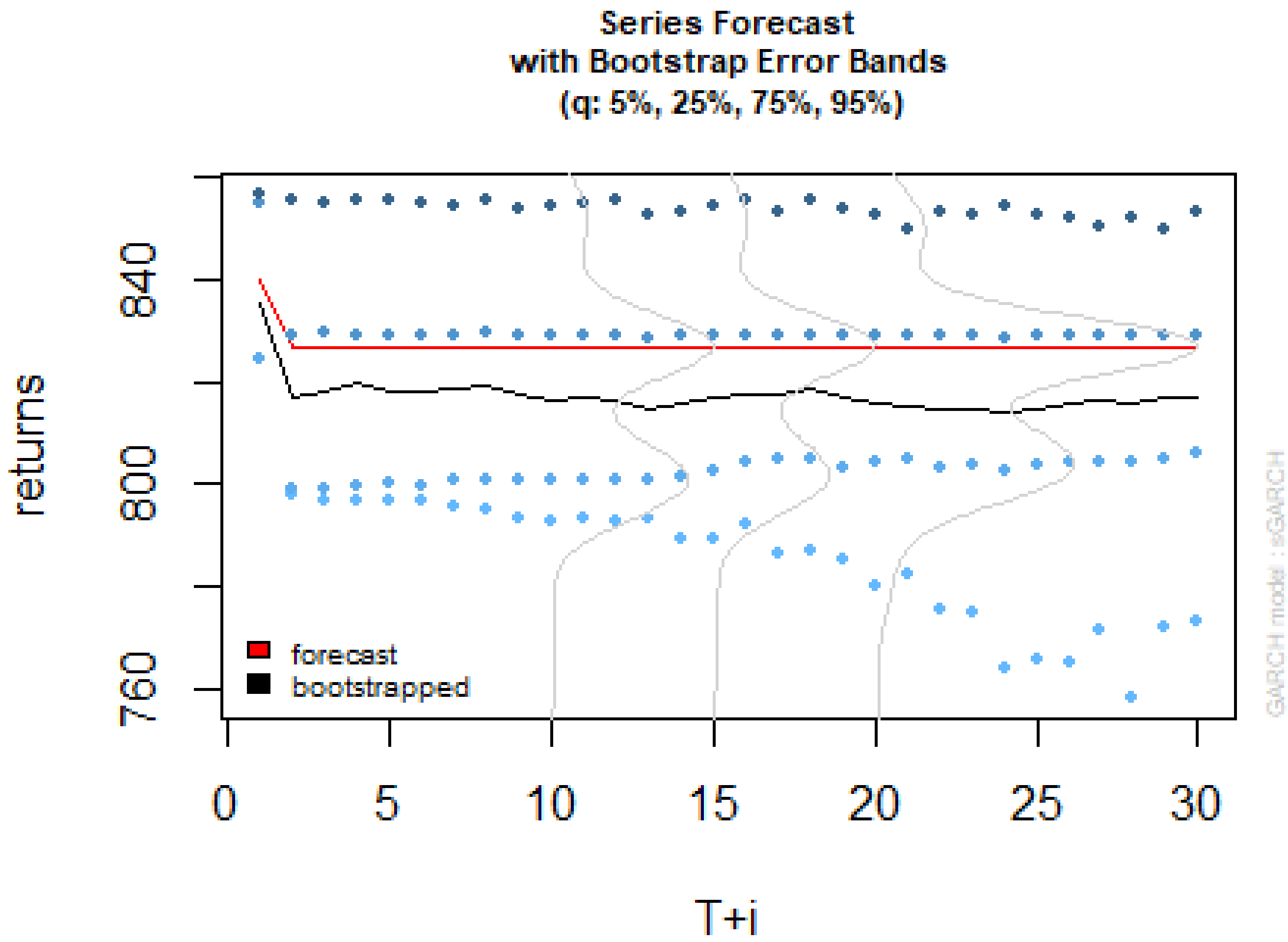
	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
826	855.1311	853.5559	856.7064	852.7220	857.5403
827	855.2478	853.6276	856.8681	852.7699	857.7258
828	855.3643	853.6965	857.0321	852.8136	857.9150
829	855.4806	853.7617	857.1995	852.8517	858.1095
830	855.5968	853.8225	857.3711	852.8832	858.3104
831	855.7129	853.8785	857.5472	852.9075	858.5183
832	855.8289	853.9297	857.7282	852.9243	858.7336
833	855.9450	853.9758	857.9141	852.9334	858.9565
834	856.0610	854.0169	858.1050	852.9349	859.1871
835	856.1770	854.0530	858.3009	852.9287	859.4252
836	856.2929	854.0843	858.5016	852.9151	859.6708
837	856.4089	854.1108	858.7071	852.8942	859.9237
838	856.5249	854.1326	858.9171	852.8663	860.1835
839	856.6409	854.1501	859.1316	852.8316	860.4502
840	856.7568	854.1632	859.3504	852.7903	860.7234
841	856.8728	854.1723	859.5733	852.7427	861.0029
842	856.9888	854.1774	859.8002	852.6891	861.2884
843	857.1047	854.1786	860.0308	852.6296	861.5798
844	857.2207	854.1762	860.2652	852.5645	861.8768
845	857.3366	854.1702	860.5031	852.4940	862.1793
846	857.4526	854.1609	860.7443	852.4183	862.4869
847	857.5686	854.1482	860.9889	852.3376	862.7995
848	857.6845	854.1324	861.2367	852.2520	863.1170
849	857.8005	854.1135	861.4875	852.1617	863.4393
850	857.9165	854.0916	861.7413	852.0669	863.7660
851	858.0324	854.0669	861.9979	851.9677	864.0972
852	858.1484	854.0393	862.2574	851.8641	864.4326
853	858.2643	854.0091	862.5196	851.7565	864.7722
854	858.3803	853.9762	862.7845	851.6447	865.1159
855	858.4963	853.9407	863.0519	851.5291	865.4635

From this result we can see that they look like the values gotten when using HoltWinters forecasts.

4.5.3 Forecasting using GARCH(p,q) model

After getting the GARCH(p,q) model that best fit our data, the next step is to forecast the month of May 2018, using the model we got, GARCH(1,1). From the following graph we will get the forecasted values and their limits;

Figure 4.15: Graph describing the upper and lower limits of the forecasts of GARCH(1,1) model, using R



Where the dark blue dots show the upper limit of our forecasts, and the clear blue dots show the lower limits of the forecasts.
The simple explanation is that in the upcoming 30 days the exchange rates values will neither exceed 850 *RWF* nor fall below 760 *RWF*.
The following figure shows the forecasted values of the month of May 2018;

Figure 4.16: Garch forecasts of the month of May 2018, using R

```
0-roll forecast:
      Series Sigma
T+1    839.7 15.20
T+2    826.7 15.20
T+3    826.7 15.19
T+4    826.7 15.18
T+5    826.7 15.17
T+6    826.7 15.17
T+7    826.7 15.16
T+8    826.7 15.15
T+9    826.7 15.14
T+10   826.7 15.14
T+11   826.7 15.13
T+12   826.7 15.12
T+13   826.7 15.12
T+14   826.7 15.11
T+15   826.7 15.10
T+16   826.7 15.09
T+17   826.7 15.09
T+18   826.7 15.08
T+19   826.7 15.07
T+20   826.7 15.06
T+21   826.7 15.06
T+22   826.7 15.05
T+23   826.7 15.04
T+24   826.7 15.03
T+25   826.7 15.03
T+26   826.7 15.02
T+27   826.7 15.01
T+28   826.7 15.01
T+29   826.7 15.00
T+30   826.7 14.99
```

Chapter 5

DISCUSION, CONCLUSION AND RECOMMENDATION

This chapter is about understanding the whole process that took place while analyzing the Exchange rates data by summarizing and comparing ,this research, to previous studies on the same subject.

5.1 DISCUSION

In chapter 4, we used three models to analyze our exchange rates data and the output were different. we compared those models and calculated the forecasts of 30 days of the month of may 2018, using each model, and still the results were different.
This process was done mainly to check if the GARCH model is the perfect model to use while dealing with exchange rates data, as suggested by previous studies. as shown in chapter 4, the analysis showed that GARCH(1,1) was the perfect model to fit our data compared to the other two.
However, this doesn't imply that GARCH models give perfect forecasts, since the data we are dealing with are randomly generated data. which means that the results we got are not to be relied on entirely.
Moreover, there were other models we ignored, due to the lack of thorough understanding of the models. The most important one is **Artificial Neural Network model**(ANN_m), which became popular in this century. this method is non-linear.
As seen in previous studies, GARCH models deal better with financial time series data than any other model, except ANN_m, since those data have volatility which means that their variances and standard deviations are not constant.

5.2 CONCLUSION

This study aimed generally on understanding the importance of exchange rates in our national economy and more specifically on forecasting exchange rates of USD versus RWF using technical methods.
In the study, we found that they were various methods that can be used to forecast exchange rates, whether for a group of investors who want to predict if their investment will be profitable in a given future, or even for financial enthusiasts who are planning for a long term venture in a foreign country.

The method used were three: Exponential smoothing, ARIMA models, GARCH models. From which we looked for the best one to fit our exchange rates data. the best fitted model was **GARCH(1,1)**.
After all that process we came to a conclusion that the previous studies were correct to say that GARCH models are the best fit to time series data with volatility. However, with little resources, we didn't explore all methods that are available to deal with volatile time series data, which I hope will no longer be a problem in the next studies about this subject.

5.3 RECOMMENDATION

After this study, we came up with advices and solutions to the next research who would like to tackle the same subject, to the foreign and local investors interested to invest in Rwanda and The United States of America respectively, to the Ministry of Finance and Economic Planning (MINECOFIN), to the Department of Mathematics in CST and lastly to young entrepreneurs.
I would like to recommend the next researcher who is interested in this study to try and analyze the Artificial Neural Network model (ANN_m) and to use a very comprehensible mathematical program when analyzing the data.
I advise and recommend foreign investors to try to analyze all factors that would affect their visions and budgets,including exchange rates predictions, rigorously. Since our economy still has a long way to go.
I advise local investors to try to make transactions using our currency, so that the value of our currency increases and hence our economic growth reaches greatness.
I recommend the Ministry of Finance and Economic Planning to try to join forces with the Ministry of Youth on a mission to include youth in our Economic Growth by investing in the youth and teaching them how local businesses can help in making our country one of the most developed countries.
I urge young entrepreneurs to valorize our currency by trying to attract foreing investors into investing in our country so that our currency appreciates. I recommend them to invest in local areas so that we can become a resourceful country that attracts foreign investors.
Lastly, I appreciate what the department has done for us, by helping us to see wider and to understand our surroundings. it was a great experience under your guidance.I would also like to suggest that you encourage students to be more interested in Financial Mathematics since the vision of our country is towards eradicating poverty through the youth.

5.4 Appendices

5.4.1 Appendix A: The change of raw data into time series data, using R.

```
exchangerates.ts <- ts(exchangerates,start=1, end = 825, frequency = 1)
plot.ts(exchangerates.ts)
```

5.4.2 Appendix B: Fitting an exponential smoothing model to our time series data, using R.

```
library(forecast)
exchangerates.ts_ttrd <- HoltWinters(exchangerates.ts, gamma = F)
exchangerates.ts_ttrd
exchangerts_trd <- forecast:::forecast.HoltWinters(exchangerates.ts_ttrd, h=30)
forecast:::plot.forecast(exchangerts_trd)
exchangerts_trd
accuracy(exchangerts_trd)
```

5.4.3 Appendix C: Fitting an ARIMA model to our time series data, using R.

```
exchange_Arima <- auto.arima(exchangerates.ts)
exchange_Arima
```

5.4.4 Appendix D: Fitting a GARCH model to our time series data, using R.

```
library(rugarch)

x2 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1,1)),mean.model =
list(armaOrder = c(1,1)),distribution.model = "std")

xgarch1 <- ugarchfit(spec = x2,data = x1)

x3 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1,1)),mean.model =
list(armaOrder = c(0,0)),distribution.model = "std")

xgarch3 <- ugarchfit(spec = x3,data = x1)

xgarch3

x4 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1,1)),mean.model =
list(armaOrder = c(0,1)),distribution.model = "std")

xgarch2 <- ugarchfit(spec = x4,data = x1)

xgarch2

x5 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1,1)),mean.model =
list(armaOrder = c(1,0)),distribution.model = "std")

xgarch4 <- ugarchfit(spec = x5,data = x1)

xpredict <- ugarchboot(xgarch2, n.ahead = 30, method = c("partial","full")[1])

plot(xpredict,which = 2)

xpredict
```

5.4.5 Appendix E: Exchange rates of USD Vs RWF of the month of May 2018 table.

Date	Average
2-May-18	855.060472
3-May-18	854.98645
4-May-18	855.086574
7-May-18	855.196787
8-May-18	855.343881
9-May-18	855.354145
10-May-18	855.492713
11-May-18	855.600926
14-May-18	855.721247
15-May-18	855.821461
16-May-18	855.951782
17-May-18	856.091996
18-May-18	856.202209
21-May-18	856.338345
22-May-18	856.468657
23-May-18	856.568979
24-May-18	856.6893
25-May-18	856.789622
28-May-18	856.899835
29-May-18	857.020049
30-May-18	857.128719
31-May-18	857.267574

Bibliography

[1] Williamson, John.(2007).*Exchange Rate Economics.*www.growthcommission.org.

[2] Pacelli,Vincenzo.(2012). *Forecasting Exchange Rates: a Comparative Analysis.* International Journal of Business and Social Science,3(10),145-148..

[3] Sullivan, W. G., & Claycombe, W. W.(1977). *Fundamentals of forecasting.*Prentice Hall.

[4] Cross, Deirdre.(2006).*Forecasting Exchange Rates.* retrieved from http://slideplayer.com/slide/8333442/#.

[5] Hyndman, R. J., Athanasopoulos, G. (2014). Forecasting: principles and practice. OTexts.

[6] Van Bergen, Jason.(March 1, 2018). *6 factors that influence exchange rates.* retrieved from https://www.investopedia.com/articles/basics/04/050704.asp

[7] National Bank of Rwanda(BNR). *Exchange rates archive.* retrieved from https://www.bnr.rw/index.php?id=89&tx_excratearch_excratearch%5Baction%5D=&tx_excratearch_excratearch%5Bcontroller%5D=ExchRateArchive&cHash=ee79272d4dce48107697c3ef24328f01.