Forecasting and Analysis of Major Forex Currency Pairs

REPORT

MFM 3151

INTRODUCTION TO CASE STUDY II

DEPARTMENT OF MATHEMATICS

UNIVERSITY OF RUHUNA

Forecasting and Analysis of Major Forex Currency Pairs

REPORT

MFM 3151
INTRODUCTION TO CASE STUDY II

GROUP No.17
A.K.S.S.Wimalasena - SC / 2021 / 12469

SUPREVISOR
Mrs. A.W.S.P. Karunarathne

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF RUHUNA

Abstract

Acknowledgment

We want to extend our sincere gratitude to Mrs. A.W.S.P. Karunarathne for her exceptional mentorship, unwavering support, and invaluable guidance throughout this case study. Her expertise, insight, and encouragement have been instrumental in shaping the direction and outcomes of this study. Her dedication to excellence and commitment to fostering academic growth have inspired and motivated me to achieve my best. We are profoundly grateful for her mentorship.

we would also like to express our gratitude to Mrs. I.C. Liyanage for her valuable contributions and support during this research. Her advice and guidelines have been appreciated and have contributed to the overall success of this case study.

Also, our sincere gratitude goes to lecturers of the Department of Mathematics, for their support and guidance in keeping us progressing in conducting this case study. We would like to extend our thanks to our colleagues, who provided valuable feedback and engaged in helpful discussions, enhancing the case study process.Our special thanks go to every individual who is involved in this case study for their contributions and support throughout the journey.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background of the Study

A time series is the measurement of a sequential set of variables over identical, consistent amounts of time, recorded as observations. Time series is frequently applied in areas like economics, history, sociology, and other disciplines. The reason is that the time series contains features such as trends, cycles, and seasonal variation that enables understanding of the past behavior of such data and assists in predicting future events. In finance, time series forecasting enables studies of how future prices may change based on past data trends for the two most extreme cases where price action is especially sensitive to change, the stock market and the Forex market.

As a result of the nature of time series data, it is inherently the case that separate observations depend upon the previous one, due to which the emergence of patterns is such that they are usually autocorrelated. That's where traditional methods of data analysis can hardly work, given that assumptions of independence among data points clearly cannot hold in this context. Time series analysis offers a formal modeling approach toward such temporal dependence, which could show up in forms such as trend components, seasonal effects, or even cyclic behavior. All these patterns can be captured by an appropriate model, which in turn offers a very good scope for robust forecasting, especially in applications where the value of being able to understand future values is high.

1.2 Forex Market and Major Currency Pairs

The forex market represents the most extensive and most liquid of any financial markets globally, with over $6 trillion worth of transactions taking place every day. This is an international, decentralized market where anyone can buy or sell currencies while maintaining many activities, including international trade and investment, as well as speculative trading.

Contrary to traditional stock exchanges, Forex is a 24/5 market that moves on the trading of New York, London, Tokyo, and Sydney. The round-the-clock operation blends with vast diversity to make Forex highly liquid, featuring its special price dynamics.

The major currency pairs are especially prominent within the Forex market,

- EUR/USD – Shows how many U.S. Dollars (USD) are required to buy one Euro (EUR). For example, if EUR/USD = 1.10, it means 1 Euro costs 1.10 U.S. Dollars.
- GBP/USD – Shows how many U.S. Dollars (USD) are needed to buy one British Pound (GBP). If GBP/USD = 1.30, it means 1 British Pound costs 1.30 U.S. Dollars.
- USD/JPY – Shows how many Japanese Yen (JPY) are needed to buy one U.S. Dollar (USD). For instance, if USD/JPY = 110.00, it means 1 U.S. Dollar costs 110 Japanese Yen.
- USD/CHF – Shows how many Swiss Francs (CHF) are required to buy one U.S. Dollar (USD). If USD/CHF = 0.92, it means 1 U.S. Dollar costs 0.92 Swiss Francs.
- AUD/USD – Shows how many U.S. Dollars (USD) are needed to buy one Australian Dollar (AUD). For example, if AUD/USD = 0.75, it means 1 Australian Dollar costs 0.75 U.S. Dollars.
- USD/CAD – Shows how many Canadian Dollars (CAD) are needed to buy one U.S. Dollar (USD). If USD/CAD = 1.25, it means 1 U.S. Dollar costs 1.25 Canadian Dollars.
- NZD/USD – Shows how many U.S. Dollars (USD) are needed to buy one New Zealand Dollar (NZD). For example, if NZD/USD = 0.65, it means 1 New Zealand Dollar costs 0.65 U.S. Dollars.

These are prominent because they have large volumes of trading and stability, pairing up the most-used global currencies: the economy of the United States, the European Union, Japan, the United Kingdom, and Switzerland. Specifically, these pairs of currencies are in greater demand among traders because their economic strength and geopolitical stability can provide traders with more liquidity and lower transaction costs, which go hand in glove with high trading activity.

Major currency pairs are the first choice due to huge liquidity, economic stability, and major influence in world finance. Their currencies represent the most stable and advanced economies in the world, which means huge trading volumes and tighter spreads-a scenario that can guarantee more reliable analysis. Their movements are based on well-documented economic indicators, among other macroeconomic factors, such as the actions of central banks. It is for these reasons that such pairs are more predictable and suitable for time series analysis.

Any efforts at understanding and predicting price trends in the Forex market, especially for major currency pairs, need strong analytical tools. Results from time series analysis pinpoint the underlying patterns and trends in the historical prices, thus providing insight that will be useful to traders and policymakers in making better decisions.

1.3 Research Problem

In spite of this broad application, however, a big gap in understanding remains concerning their effectiveness in forecasting the major Forex currency pairs. Classic models do indeed have certain limitations when being applied to Forex data in capturing volatility and external influences that hit the market. This paper, therefore, investigates the limitation and tests the predictive accuracy of various time-series models within the volatile and economically influenced contexts of the Forex market.

The primary problem to address is: "Can accurate models for traders be built using traditional time series methods?"

1.4 Significance of the Study

The findings of this study will provide an understanding of how effective traditional time series models perform in the Forex market, highlighting both strengths and weaknesses. It will be able to identify any weaknesses in the models discussed and mention any alternative methods that may be used in improving forecasting accuracy in financial forecasting. The outcome will, hence, provide good grounds for traders and financial analysts in making better decisions within a very volatile environment.

1.5 Objectives of the study

1. Evaluate the performance of traditional time series models to generate the forecast for major pairs using the Forex market.
2. Discuss any weaknesses that could be arising out of the traditional time series model fitted on the Forex data.
3. Suggest improvements in the forecasting methods of the Forex market, based on the results obtained from the traditional time series model analysis.
4. Provide significant input related to financial forecasting, which might be helpful for traders and analysts to make insightful decisions.

Chapter 2

Literature Review

This literature review provides a solid understanding of the existing knowledge around using time series approaches to forecast major currency pairs. By looking at previous studies, key theories, and important findings, this section highlights where the research stands today, points out gaps in the literature, and emphasizes the need for further exploration. It also explains how the research fits into the larger academic context and supports the choice of methodology.

Historically, time series models like the Autoregressive Integrated Moving Average (ARIMA) have been widely used for predicting exchange rates.As a example, in forecasting the EUR/USD exchange rate,Pacelli and Vincenzo [1].applied a hybrid ARIMA model. This method combined ARIMA's ability to handle linear trends with Long Short-Term Memory (LSTM) models that capture non-linear patterns. The study found that this combination significantly improved forecasting accuracy for both short- and long-term predictions.

When looking at the AUD/USD exchange rate, research by Kontopoulou et al [2]. showed that while ARIMA models can offer reliable short-term forecasts, they often struggle with the more complex and unpredictable nature of exchange rate movements.On the other hand, GARCH models, introduced by Bollerslev [3]. have been more successful in addressing the market's volatile behavior, making them useful for forecasting fluctuations in the AUD/USD exchange rate.

The GBP/USD pair, one of the most traded globally, was initially studied using traditional linear models like ARIMA. Meese and Rogoff's [4]. influential research applied ARIMA to various exchange rates, including GBP/USD, showing that ARIMA models were effective for short-term predictions.

Similarly, the USD/CHF exchange rate—particularly interesting due to the Swiss franc's safe-haven status—has been a key area of research. Box and Jenkins [5].introduced ARIMA, but later studies, including one by Meese and Rogoff [4]. revealed that while ARIMA could capture short-term trends, it often failed to outperform the simpler random walk model in longer-term forecasts.

Also forecasting the USD/JPY exchange rate has involved a range of time series models, from traditional econometric approaches to modern machine learning techniques.Meese and

Rogoff [4]. found that, in the short term, simpler time series models like ARIMA sometimes performed better than more complex structural models.Additionally, researchers have used GARCH models, developed by Bollerslev [3]. to address the volatility commonly seen in exchange rates. Engle [6]. applied the ARCH model,a precursor to GARCH, to capture the time-varying volatility of USD/JPY, demonstrating its effectiveness in managing unpredictable market behavior.

In conclusion, the literature on time series forecasting of major currency pairs, such as EUR/USD,GBP/USD,USD/JPY,USD/CHF etc.demonstrates the evolution of forecasting techniques from traditional linear models like ARIMA, while ARIMA and similar models have proven effective in short-term forecasting, their limitations in capturing the non-linear complexities and volatility of exchange rate movements have led to the adoption of more advanced models.

Chapter 3

Methodology

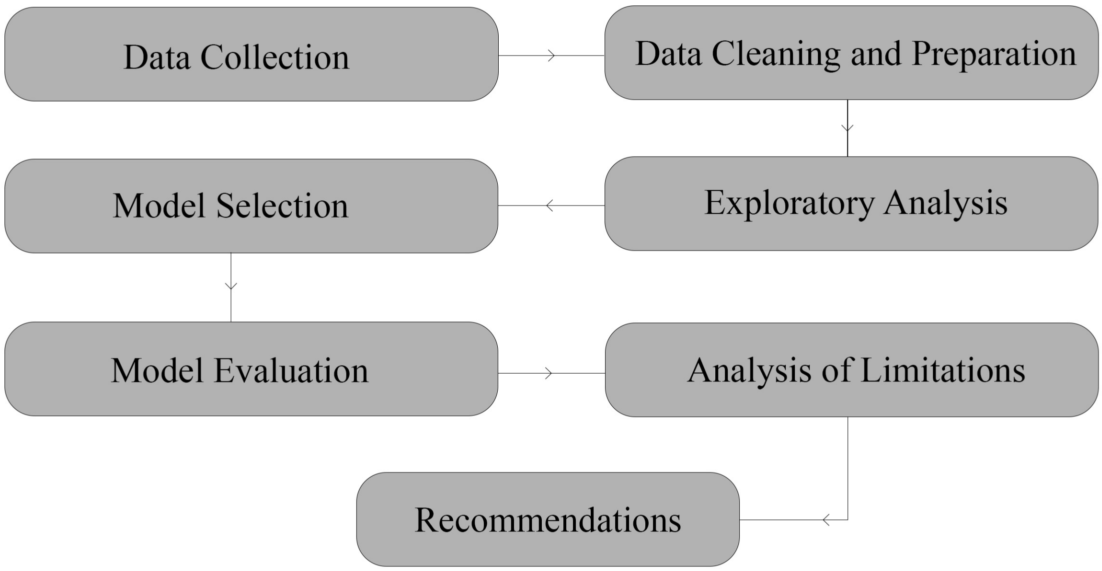


Figure 3.1: General methodology flowchart

3.1 Data Collection

For this study, historical forex data was gathered from Yahoo Finance, a widely recognized and reliable source for financial data. In this study, weekdays are used for the analysis.

The dataset utilized for this study is a time series dataset comprising historical forex rates for major currency pairs. The data spans from 2023-01-02, to 2024-08-30. Each data point represents the daily closing exchange rate, capturing fluctuations and trends over time.

This extensive period enables a thorough analysis of long-term trends, periodicities, and volatility within the forex market.

3.2 Data cleaning and preparation

The data is cleaned regarding missing values, outliers, and inconsistencies. From here, if there are any missing day values, these will be determined and replaced with the previous day's closing price value for continuity issues in this stage. The data is then transformed into an analysis-ready format.

3.3 Exploratory data analysis

To initiate the analysis, an Exploratory Data Analysis was conducted to understand the characteristics of the datasets. This critical phase involved cleaning and preparing the data, performing descriptive statistics, and generating visual representations to gain insights into the forex time series data. In this stage, The forex currency pairs was described one by one using time series plot and decomposition method.

3.3.1 EUR/USD – Euro / US Dollar

The Currency Pair EUR/USD is the shortened term for the euro against the U.S. dollar pair, or cross for the currencies of the European Union (EU) and the United States (USA). The currency pair indicates how many U.S. dollars (the quote currency) are needed to purchase one euro (the base currency).

Figure 3.2: EUR/USD - Daily closing prices

The time series plot clearly shows there is no specific trend or seasonality but in 2023 July to 2023 October shows strong downtrend and then the close price is fluctuating around 1.08.

Tabla 3.1: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-2.7496	0.2605	False
Kwiatkowski-Phillips-Schmidt-Shin	0.20946	0.1	False
Phillips-Perron (PP)	-16.851	0.1665	False

The above table shows test results for stationary of the dataset and majority of tests say the dataset isn't stationary so time series difference is used to make the dataset stationary. The following figure shows the first order difference close price data.

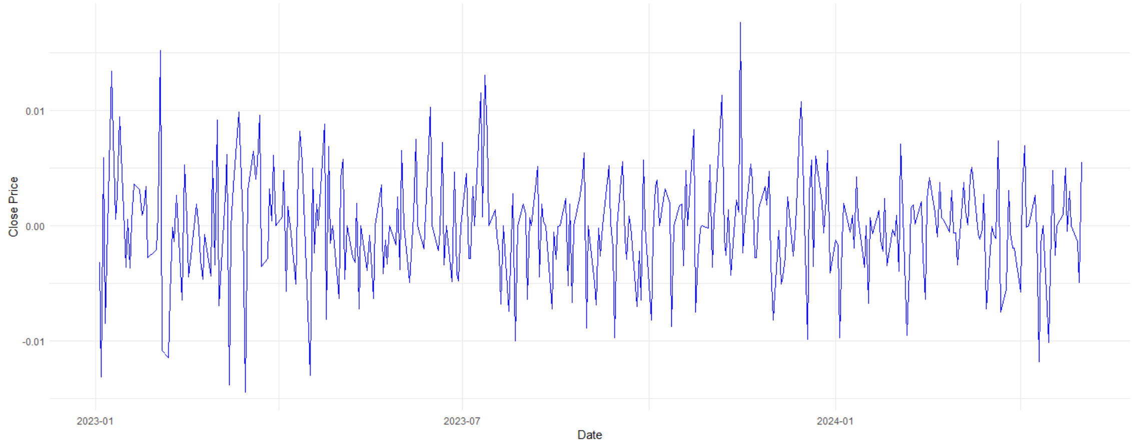


Figure 3.3: EUR/USD - Order one difference daily closing prices

Tabla 3.2: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-6.321	0.01	True
Kwiatkowski-Phillips-Schmidt-Shin	0.050041	0.1	True
Phillips-Perron (PP)	-352.87	0.01	True

The majority of the test for stationary are saying the time series is stationary so the first difference is sufficient for make the dataset stationary.

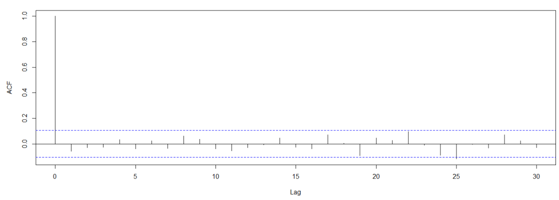


Figure 3.4: ACF plot - Order one difference daily closing prices

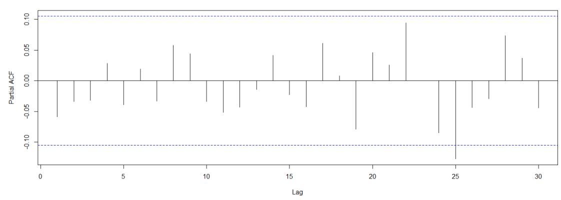


Figure 3.5: PACF plot - Order one difference daily closing prices

The above ACF and PACF plots confidence bounds plots for 0.95 significance level and the plots show there is no significance lags. And there is no any seasonal in the datset.

3.3.2 GBP/USD – British Pound / US Dollar

The GBP/USD (British pound/U.S. dollar) is an abbreviation for the British pound and U.S. dollar currency pair, or cross. The currency pair shows how many U.S. dollars (the quote currency) are needed to purchase one British pound (the base currency).

Figure 3.6: GBP/USD - Daily closing prices

The time series plot clearly shows there is no specific trend or seasonality but in 2023 March to 2023 August shows strong uptrend and then the close price is fluctuating around 1.26.

Tabla 3.3: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-2.1712	0.5046	False
Kwiatkowski-Phillips-Schmidt-Shin	1.5814	0.01	False
Phillips-Perron (PP)	-11.523	0.4649	False

The above table shows test results for stationary of the dataset and all of tests say the dataset isn't stationary so time series difference is used to make the dataset stationary. The following figure shows the first order difference close price data.

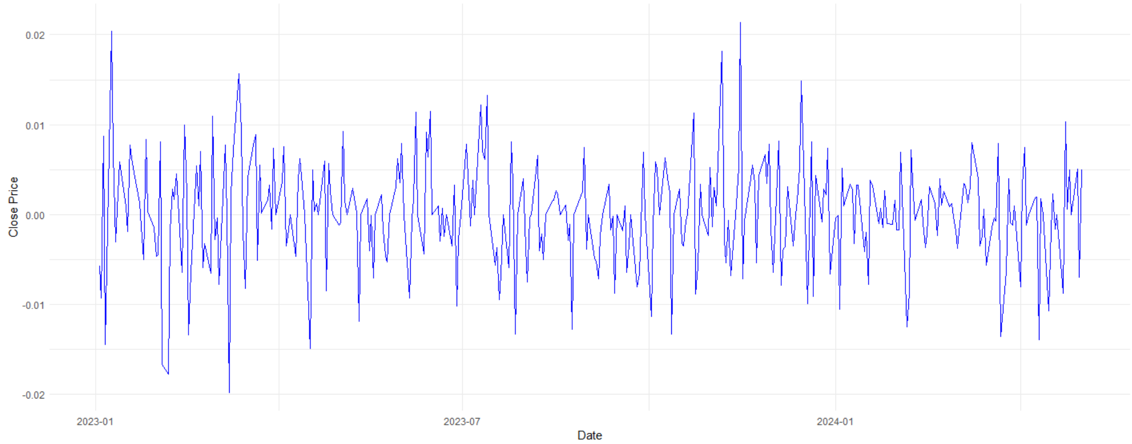


Figure 3.7: GBP/USD - Order one difference daily closing prices

Tabla 3.4: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-6.3887	0.01	True
Kwiatkowski-Phillips-Schmidt-Shin	0.073017	0.1	True
Phillips-Perron (PP)	-325.4	0.01	True

All of the test for stationary are saying the time series is stationary so the first difference is sufficient for make the dataset stationary.

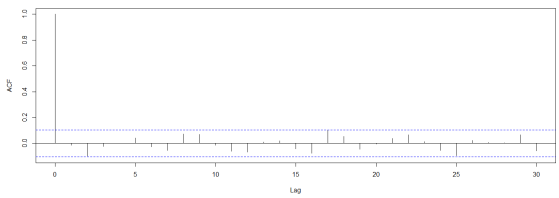


Figure 3.8: ACF plot - Order one difference daily closing prices

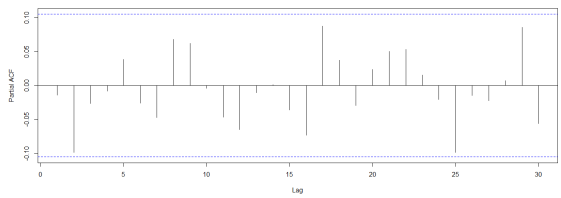


Figure 3.9: PACF plot - Order one difference daily closing prices

The above ACF and PACF plots confidence bounds plots for 0.95 significance level and the plots show there is no significance lags. And there is no any seasonal in the datset.

3.3.3 USD/JPY – US Dollar / Japanese Yen

USD/JPY is the abbreviation used to denote the currency exchange rate for the U.S. dollar and Japanese yen. The currency pair shows how many Japanese yen (the quote currency) are needed to purchase one U.S. dollar (the base currency).

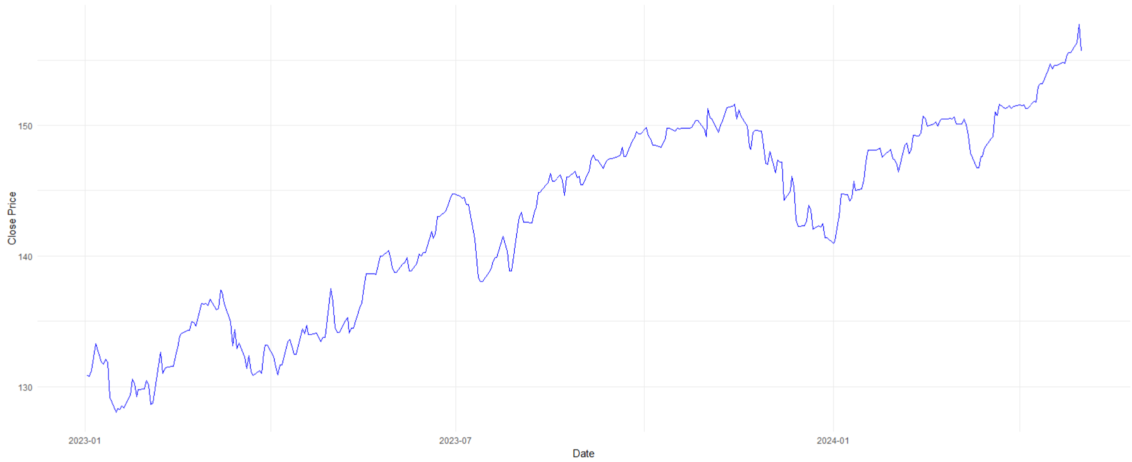


Figure 3.10: USD/JPY - Daily closing prices

The time series plot clearly shows a long-term uptrend in the entire dataset.

Tabla 3.5: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-2.5241	0.3556	False
Kwiatkowski-Phillips-Schmidt-Shin	4.9689	0.01	False
Phillips-Perron (PP)	-12.275	0.4228	False

The above table shows test results for stationary of the dataset and all of tests say the dataset isn't stationary so time series difference is used to make the dataset stationary. The following figure shows the first order difference close price data.

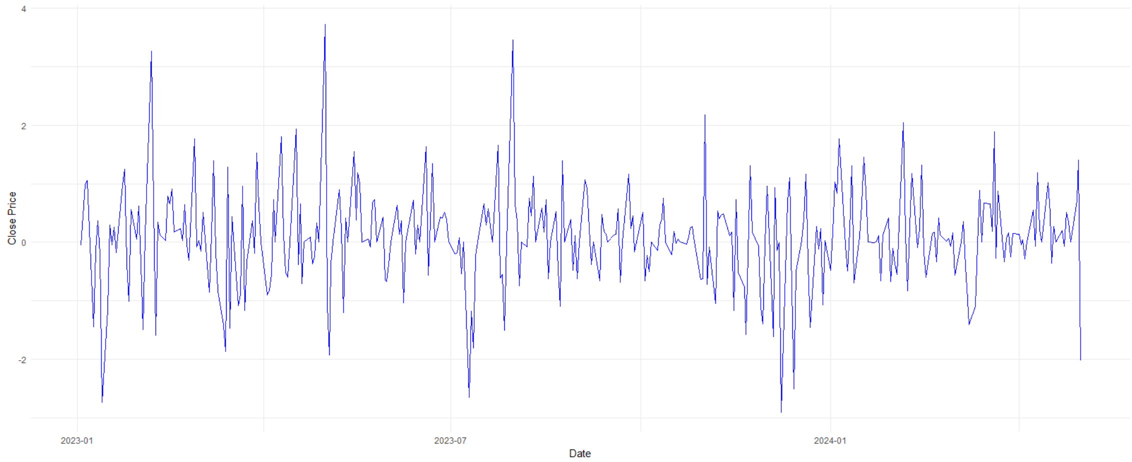


Figure 3.11: USD/JPY - Order one difference daily closing prices

Tabla 3.6: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-6.1475	0.01	True
Kwiatkowski-Phillips-Schmidt-Shin	0.04657	0.1	True
Phillips-Perron (PP)	-298.9	0.01	True

All of the test for stationary are saying the time series is stationary so the first difference is sufficient for make the dataset stationary.

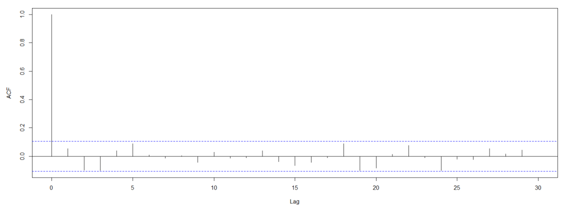


Figure 3.12: ACF plot - Order one difference daily closing prices

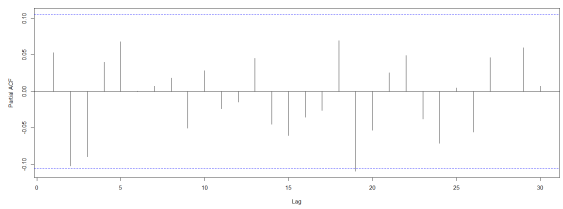


Figure 3.13: PACF plot - Order one difference daily closing prices

The above ACF and PACF plots confidence bounds plots for 0.95 significance level and the plots show there is no significance lags. And there is no any seasonal in the datset.

3.3.4 USD/CHF- US Dollar / Swiss Frank

The USD/CHF represents the exchange rate between the Swiss Franc (CHF) and the U.S. Dollar (USD). Specifically, it indicates how many U.S. dollars are needed to purchase one Swiss Franc.

Figure 3.14: USD/CHF - Daily closing prices

When considering the CHF and USD ,There was a approximately downward trend from January 2023 to May 2024, also significant downward decline was recorded on July 2024. But after that exchange rates gradually increased on next 4 months, also between the October 2024 to January 2024 there were strictly negative trend pattern. But afterward CHF/USD exchange rate gradually increased over the time.

Tabla 3.7: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-1.5167	0.7807	False
Kwiatkowski-Phillips-Schmidt-Shin	2.1935	0.01	False
Phillips-Perron (PP)	-6.1766	0.7643	False

The above table shows test results for stationary of the dataset and all of tests say the dataset isn't stationary so time series difference is used to make the dataset stationary. The following figure shows the first order difference close price data.

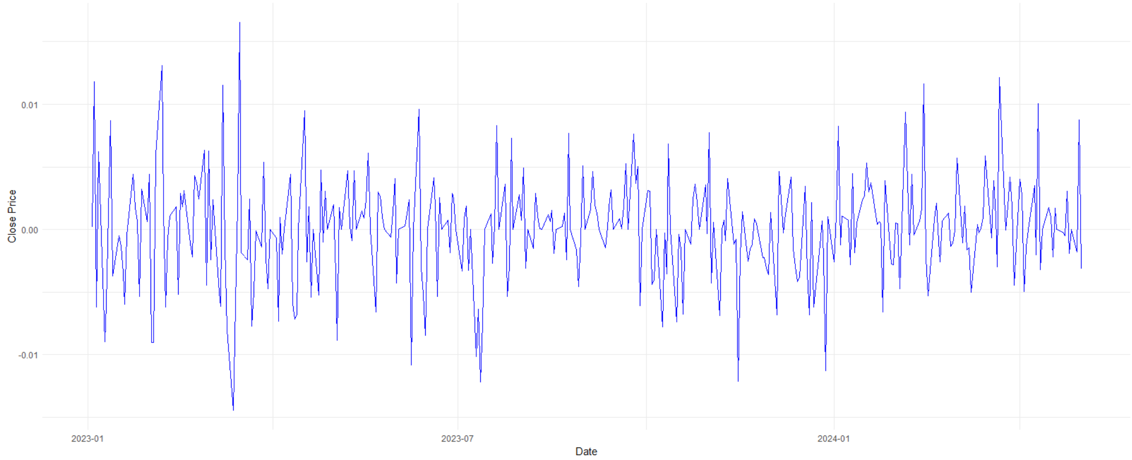


Figure 3.15: USD/CHF - Order one difference daily closing prices

Tabla 3.8: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-5.9974	0.01	True
Kwiatkowski-Phillips-Schmidt-Shin	0.19317	0.1	True
Phillips-Perron (PP)	-319.01	0.01	True

All of the test for stationary are saying the time series is stationary so the first difference is sufficient for make the dataset stationary.

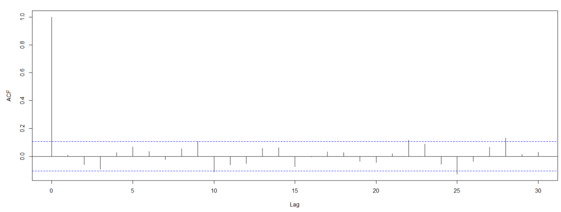


Figure 3.16: ACF plot - Order one difference daily closing prices

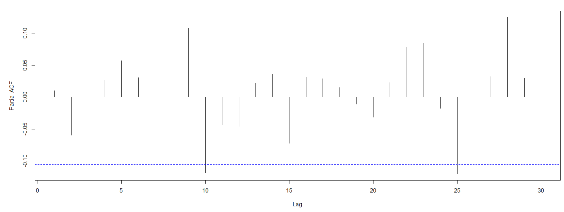


Figure 3.17: PACF plot - Order one difference daily closing prices

The above ACF and PACF plots confidence bounds plots for 0.95 significance level and the ACF plot, 9,10,22,25 and 28 lags are significance. According to the PACF plot 9,10,25 and 28 lags are significant. And there is no any seasonal in the datset.

3.3.5 USD/CAD – US Dollar / Canadian Dollar

The currency pair USD/CAD represents the value of the U.S. dollar (USD) in terms of the Canadian dollar (CAD) and is a significant pair in Forex trading. They are also referred to as the cross between the currencies of Canada and the United States.

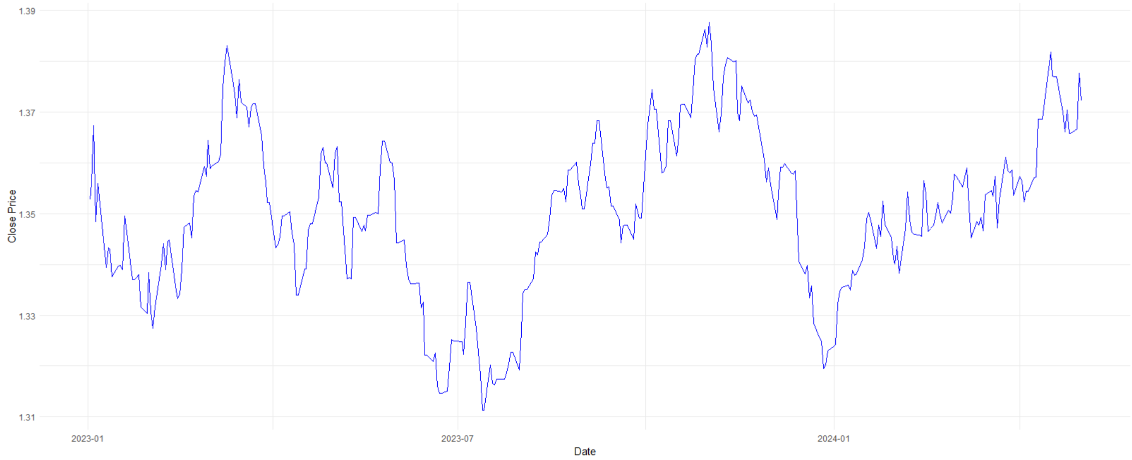


Figure 3.18: USD/CAD - Daily closing prices

According to the time series plot, the chart shows multiple cycles of rise and fall over the period. Each peak and trough indicates changes in market conditions or events that may have influenced the price of this currency pair, with significant volatility observed throughout the year.

Tabla 3.9: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-2.6707	0.2938	False
Kwiatkowski-Phillips-Schmidt-Shin	0.50702	0.04009	False
Phillips-Perron (PP)	-16.186	0.2038	False

The above table shows test results for stationary of the dataset and all of tests say the dataset isn't stationary so time series difference is used to make the dataset stationary. The following figure shows the first order difference close price data.

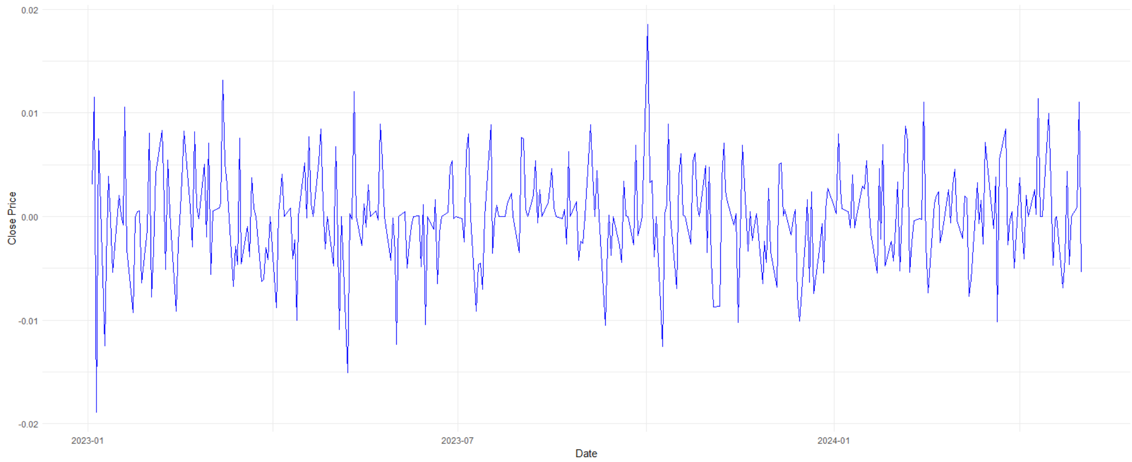


Figure 3.19: USD/CAD - Order one difference daily closing prices

Tabla 3.10: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-6.8269	0.01	True
Kwiatkowski-Phillips-Schmidt-Shin	0.054536	0.1	True
Phillips-Perron (PP)	-328.11	0.01	True

All of the test for stationary are saying the time series is stationary so the first difference is sufficient for make the dataset stationary.

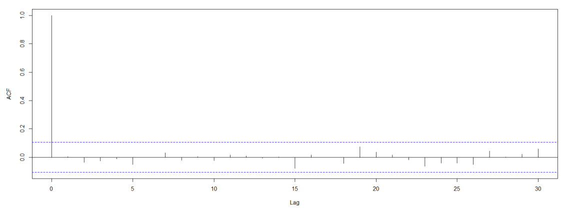


Figure 3.20: ACF plot - Order one difference daily closing prices

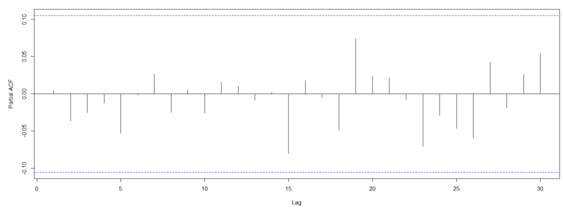


Figure 3.21: PACF plot - Order one difference daily closing prices

The above ACF and PACF plots confidence bounds plots for 0.95 significance level and there is no significance lags in ACF or PACF plots. And there is no any seasonal in the datset.

3.3.6 AUD/USD – Australian Dollar / US Dollar

The AUD/USD represents the exchange rate between the Australian Dollar(AUD) and the U.S. Dollar (USD). Specifically, it indicates how many U.S. dollars are needed to purchase one Australian Dollar.

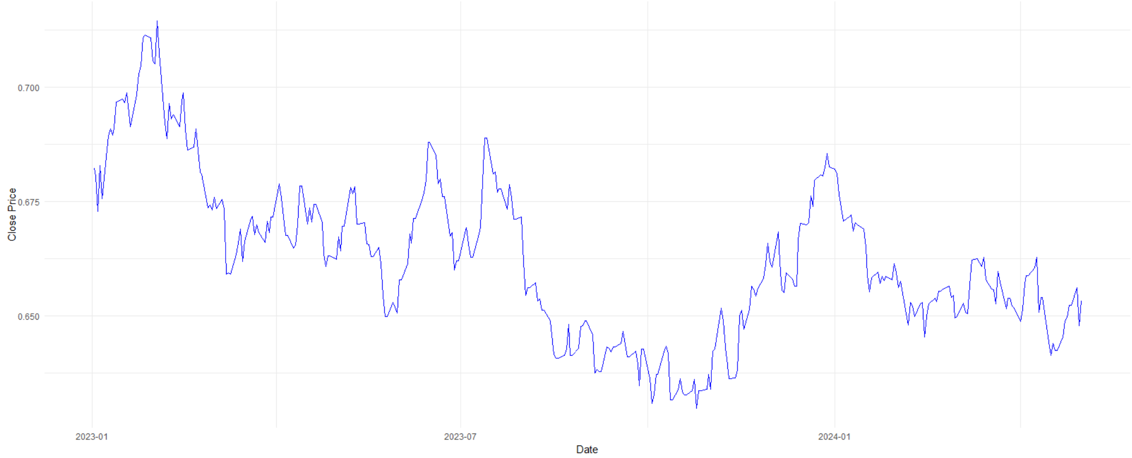


Figure 3.22: AUD/USD - Daily closing prices

Following time series plot described fluctuations of the exchange rate between the AUD and USD, at a glance view its difficult to find any seasonal or trend within the observed time frame. As well as there was no any clear long term trend. But during the July -November there were downward trend. In January 2023 there was a significant upward spike in the closing price of over 0.7 dollars also it hit around 0.625 dollars in November 2023 as the lowest closing price.

Tabla 3.11: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-2.5791	0.3325	False
Kwiatkowski-Phillips-Schmidt-Shin	2.382	0.01	False
Phillips-Perron (PP)	-15.135	0.2626	False

The above table shows test results for stationary of the dataset and all of tests say the dataset isn't stationary so time series difference is used to make the dataset stationary. The following figure shows the first order difference close price data.

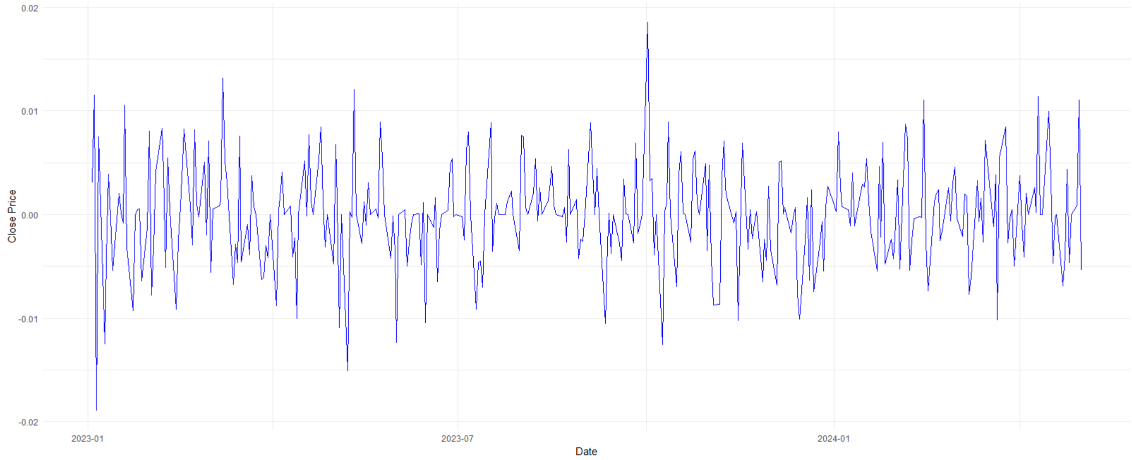


Figure 3.23: USD/CAD - Order one difference daily closing prices

Tabla 3.12: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-7.0093	0.01	True
Kwiatkowski-Phillips-Schmidt-Shin	0.042102	0.1	True
Phillips-Perron (PP)	-334.29	0.01	True

All of the test for stationary are saying the time series is stationary so the first difference is sufficient for make the dataset stationary.

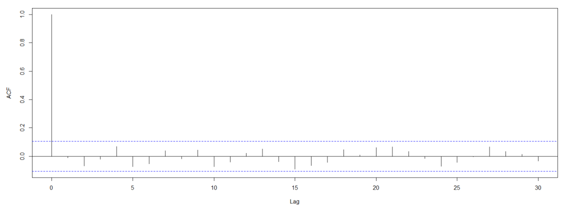


Figure 3.24: ACF plot - Order one difference daily closing prices

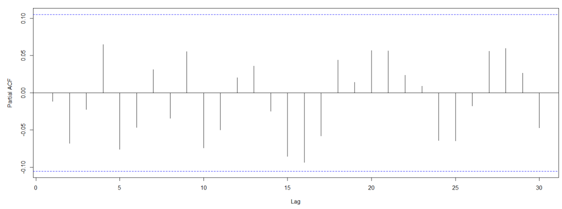


Figure 3.25: PACF plot - Order one difference daily closing prices

The above ACF and PACF plots confidence bounds plots for 0.95 significance level and there is no significance lags in ACF or PACF plots. And there is no any seasonal in the datset.

3.3.7 NZD/USD – New Zealand Dollar / US Dollar

The NZD/USD currency pair represents the exchange rate between the New Zealand Dollar (NZD) and the United States Dollar (USD).

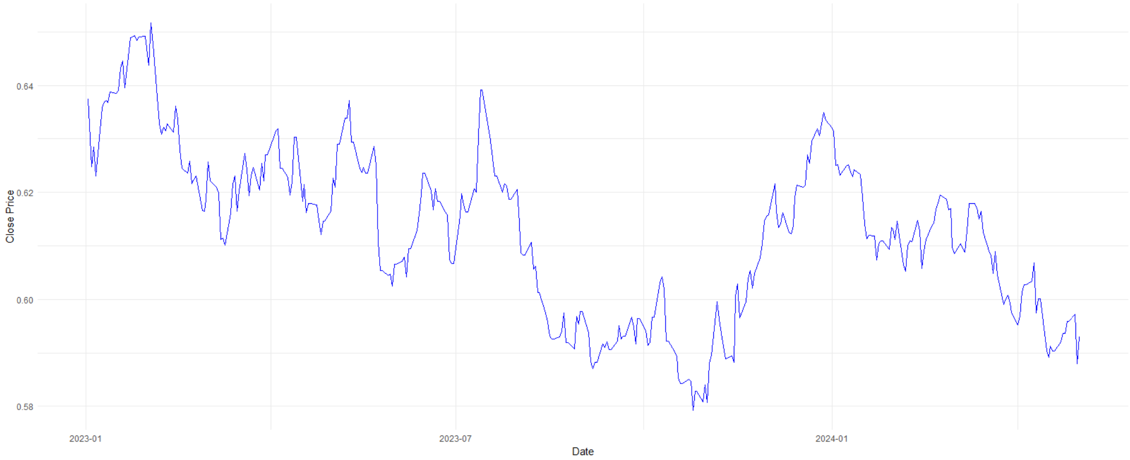


Figure 3.26: NZD/USD - Daily closing prices

The line chart presents a detailed overview of the closing price fluctuations of an asset over a 17-month period, spanning from January 2023 to May 2024. While the general trend suggests a decline in the asset's value, with the closing price dipping below the 0.60 mark in the latter half of the year, several notable upward spikes disrupt this trajectory. The most significant spike occurs during the early months of 2023, reaching a peak just shy of 0.64.

Tabla 3.13: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-2.5485	0.3454	False
Kwiatkowski-Phillips-Schmidt-Shin	2.0727	0.01	False
Phillips-Perron (PP)	-14.766	0.2833	False

The above table shows test results for stationary of the dataset and all of tests say the dataset isn't stationary so time series difference is used to make the dataset stationary. The following figure shows the first order difference close price data.

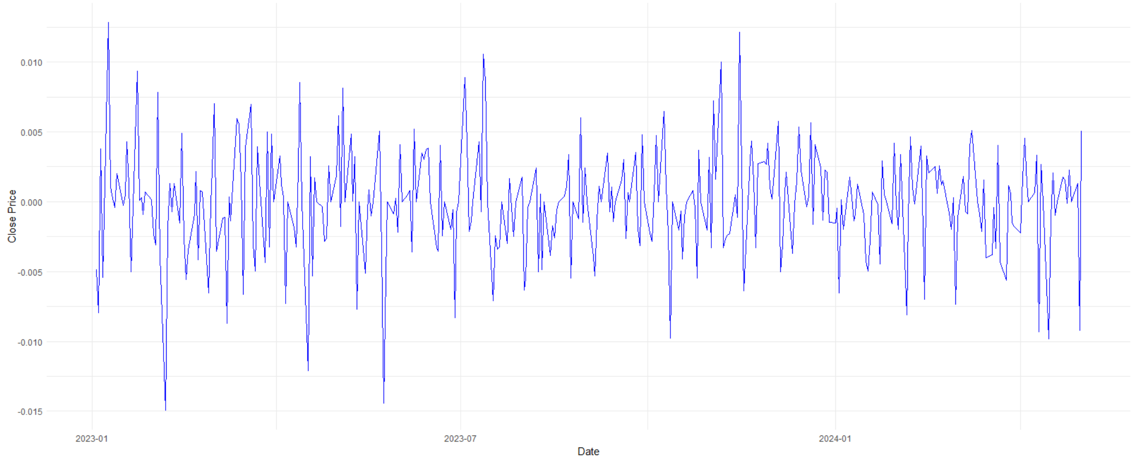


Figure 3.27: NZD/USD - Order one difference daily closing prices

Tabla 3.14: Stationarity Test Results

Test	Test Statistic	p-value	Stationary
Augmented Dickey-Fuller	-6.3987	0.01	True
Kwiatkowski-Phillips-Schmidt-Shin	0.04416	0.1	True
Phillips-Perron (PP)	-329.48	0.01	True

All of the test for stationary are saying the time series is stationary so the first difference is sufficient for make the dataset stationary.

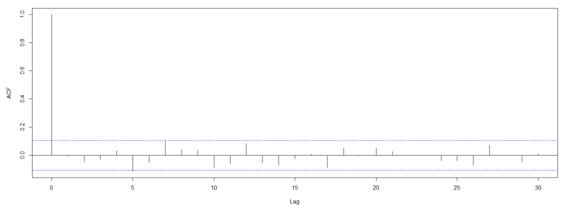


Figure 3.28: ACF plot - Order one difference daily closing prices

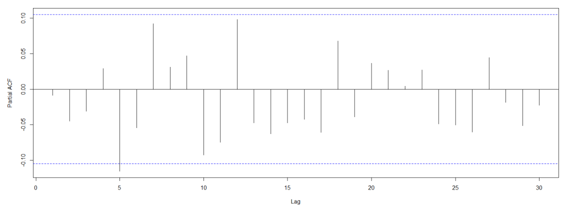


Figure 3.29: PACF plot - Order one difference daily closing prices

The above ACF and PACF plots confidence bounds plots for 0.95 significance level and the lag 5 is significance in both ACF and PACF. And there is no any seasonal in the datset.

3.4 Model Selection

Traditional time series models comprise an Auto Regressive model, Moving Average model (MA), Auto Regressive Integrated Moving Average (ARIMA), Seasonal Retrogressive Integrated Moving Average (SARIMA), and Generalized Auto regressive Conditional Heteroskedasticity (GARCH), which will be chosen to implement the forecast. Every model will be chosen based on appropriateness with regard to characteristics identified in the data through an exploratory analysis.

3.5 Model Evaluation

These chosen models will then be further evaluated statistically using the Akaike Information Criterion (AIC), for predictive accuracy. Forecasts from these models will be compared against actual closing prices to determine the strength of each model in handling Forex data.

3.6 Analysis of Limitations

In this study only use the traditional time series models and the model evaluate using only the AIC value.

Chapter 4

Quantitative analysis

4.1 Statistical Tests

4.1.1 Test for Stationary

Most of the statistical time series models assumes the dataset is stationary so the following test are used for test the stationarity of the dataset.

Augmented Dickey-Fuller (ADF) Test

The Augmented Dickey-Fuller (ADF) test is used to test the null hypothesis that a unit root is present in a time series, indicating that the series is non-stationary.

The hypotheses for the ADF test are as follows:

- Null Hypothesis (H_0): The time series has a unit root, i.e., it is non-stationary.
- Alternative Hypothesis (H_1): The time series does not have a unit root, i.e., it is stationary.

The ADF test involves fitting a regression model to the time series and testing whether the coefficient of the lagged value of the series is significantly different from zero. If the null hypothesis is rejected, the series is considered stationary.

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test

The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test is used to test for stationarity around a deterministic trend in a time series.

The hypotheses for the KPSS test are as follows:

- Null Hypothesis (H_0): The time series is stationary around a deterministic trend (trend stationary).
- Alternative Hypothesis (H_1): The time series is non-stationary (contains a unit root).

The KPSS test involves regressing the time series on a deterministic trend (if specified) and testing the null hypothesis. If the null hypothesis is rejected, it indicates that the series is non-stationary and may contain a unit root.

Phillips-Perron (PP) Test

The Phillips-Perron (PP) test is used to test for a unit root in a time series, which helps determine whether the series is stationary or non-stationary.

The hypotheses for the PP test are as follows:

- Null Hypothesis (H_0): The time series has a unit root, i.e., it is non-stationary.
- Alternative Hypothesis (H_1): The time series does not have a unit root, i.e., it is stationary.

The PP test is a modification of the Dickey-Fuller test and takes into account heteroscedasticity and autocorrelation of the residuals. If the null hypothesis is rejected, the time series is considered stationary.

4.1.2 Test for ARCH Effect

The Box-Ljung test can be used to test for the presence of autocorrelation in squared residuals, which helps in detecting ARCH effects in a time series.

The hypotheses for the Box-Ljung test for ARCH effect are as follows:

- Null Hypothesis (H_0): There is no autocorrelation in the squared residuals (i.e., no ARCH effect is present).
- Alternative Hypothesis (H_1): There is autocorrelation in the squared residuals (i.e., ARCH effect is present).

The test involves applying the Box-Ljung test to the squared residuals of the model. If the p-value is smaller than the significance level (e.g., 0.05), the null hypothesis is rejected, indicating the presence of ARCH effects in the data.

4.2 Auto-Regressive and Moving Average Models

Autocorrelation analysis is an important step in the Exploratory Data Analysis (EDA) of time series. The autocorrelation analysis helps in detecting hidden patterns and seasonality and in checking for randomness. It is especially important when you intend to use an ARIMA model for forecasting because the autocorrelation analysis helps to identify the AR and MA parameters for the ARIMA model.

4.2.1 Auto-Regressive (AR) Model

$$\hat{y}_t = \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} \tag{4.1}$$

The AR model assumes that the current value ($\hat{y}_t$) is dependent on previous values (y_{t-1}, y_{t-2}, y_{t-3}, $\ldots$). Because of this assumption, we can build a linear regression model. To figure out the order of an AR model, you would use the PACF.

4.2.2 Moving Average (MA) Model

$$\hat{y}_t = \epsilon_t + \beta_1 \epsilon_{t-1} + \cdots + \beta_q \epsilon_{t-q} \tag{4.2}$$

The MA model assumes that the current value ($\hat{y}_t$) is dependent on the error terms including the current error (ϵ_t, ϵ_{t-1}, ϵ_{t-2}, ϵ_{t-3}, $\ldots$). Because error terms are random, there is no linear relationship between the current value and the error terms. To figure out the order of an MA model, you would use the ACF.

4.2.3 AutoRegressive Moving Average (ARMA) Model

The ARMA (AutoRegressive Moving Average) model is a combination of two key models used in time series analysis: the Autoregressive (AR) model and the Moving Average (MA) model. The formula for the ARMA(p, q) model is:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t \tag{4.3}$$

where:

- x_t is the current value of the time series.
- ϕ_i are the parameters of the autoregressive part (AR).
- θ_j are the parameters of the moving average part (MA).
- ϵ_t is the white noise error term.

In summary, the ARMA model captures both the momentum and shock effects in time series data by combining the effects of past values and

4.2.4 Order of AR, MA, and ARMA Model

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off (Geometric decay)	Significant at lag q / Cuts off after lag q	Tails off (Geometric decay)
PACF	Significant at each lag p / Cuts off after lag p	Tails off (Geometric decay)	Tails off (Geometric decay)

Tabla 4.1: Comparison of ACF and PACF for AR(p), MA(q), and ARMA(p, q) models

4.3 AutoRegressive Integrated Moving Average (ARIMA) Model

Autoregressive Moving Average models:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t.$$

- Predictors include both lagged values of y_t and lagged errors.
- Conditions on coefficients ensure stationarity.
- Conditions on coefficients ensure invertibility.

Autoregressive Integrated Moving Average models

- Combine ARMA model with differencing.
- $(1 - B)^d y_t$ follows an ARMA model.

4.3.1 ARIMA(p, d, q) model

- AR: p = order of the autoregressive part
- I: d = degree of first differencing involved
- MA: q = order of the moving average part

- White noise model: ARIMA(0,0,0)
- Random walk: ARIMA(0,1,0) with no constant
- Random walk with drift: ARIMA(0,1,0) with const.
- AR(p): ARIMA(p,0,0)
- MA(q): ARIMA(0,0,q)

4.4 ARCH-GARCH models

Autoregressive integrated moving average (ARIMA) models that allow modeling of volatility are unable to deal with volatility over time.Monetary and financial series are characterized by volatility clustering means periods of high volatility alternate with periods of low volatility. This phenomenon is called conditional heterocedents that is particularly common in market data. From an empirical point of view, we can note the presence of nonlinear phenomena in the time series such as the presence of non-constant variability, the presence of the evolution cycle of volatility.

4.4.1 The ARCH (q) model

Let Y_t the return be explained by some variables X_t which can be an ARMA model (p, q).

$$Y_t = X_t \beta + \epsilon_t \tag{4.4}$$

Where $\varepsilon_t / I_{t-1} \sim N(0, \sigma_t)$ I_{t-1} refers to all the information available in the returns data up to the time $t - 1$.

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2$$

$$\varepsilon_t = \sigma_t z_t \quad z_t \sim N(0,1) \quad \text{where:}$$

σ_t is the conditional variance
z_t is the standard residue
ε_t is conditional errors of the asset at time t
$\alpha_0, \ldots, \alpha_i$ are real parameters
q is the order of the ARCH process

4.4.2 The GARCH (q,p) model

GARCH modeling has become an essential tool in finance for analyzing and predicting volatility. It was introduced to correct the weaknesses of the ARCH model at the estimation level with a high number of parameters.

In order to highlight the long memory of volatility Bollerslev (1986) generalized the ARCH (q) model by proposing the GARCH model (p, q) which consists in adding the delayed variance in its equation.

- Let $Y_t = X_t\beta + \varepsilon_t$ the return be explained by some variables X_t. Where, $\varepsilon_t / I_{t-1} \sim N(0, \sigma_t)$
- The GARCH process allows you to set the conditional variance σ_t^2

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$$

- α_0 : Parameter representing the long-term average variance
- α_i : Parameter that measures the sensitivity to conditional volatility
- β_j : Parameter that measures the persistence of conditional volatility
- To obtain a positive conditional variance, one must respect these conditions : $\alpha_0 > 0$ and $\alpha_i \geq 0$ for all $i = 1, 2, 3, ..., q$ $\beta_j \geq 0$ $for\ all$ j = 1, 2, 3, ..., p q > 0 and p ≥ 0

This study aims at the prediction of the Forex major currency pairs using the modeling techniques ARIMA and GARCH.Let P_t be the Close price at the end of the trading day of a currency pair.

4.5 EUR/USD – Euro / US Dollar

Modeling forex prices data usually do not follow and AR, MA or ARMA process but the ARIMA model shows some useful results compared with other time series models. The following table shows the all possible models that can design using ARIMA, AIC (Akaike Information Criterion) is used to measure the accuracy of the model.

Tabla 4.2: A sample long table.

p	d	q	AIC
0	1	0	-2720.598
0	1	1	-2719.888
0	1	2	-2718.266
0	1	3	-2716.495
0	1	4	-2714.761
0	1	5	-2713.363
1	1	0	-2719.801
1	1	1	-2718.348
1	1	2	-2716.351
1	1	3	-2714.346
1	1	4	-2712.579
1	1	5	-2712.093
2	1	0	-2718.206
2	1	1	-2716.793
2	1	2	-2715.523
2	1	3	-2714.101
2	1	4	-2712.104
2	1	5	-2710.088
3	1	0	-2716.550
3	1	1	-2716.113
3	1	2	-2714.105
3	1	3	-2716.891
3	1	4	-2714.891
3	1	5	-2713.297
4	1	0	-2714.841
4	1	1	-2714.113
4	1	2	-2714.819
4	1	3	-2719.967
4	1	4	-2709.527
4	1	5	-2709.154
5	1	0	-2713.404
5	1	1	-2712.046
5	1	2	-2715.161
5	1	3	-2712.982

Continued on next page

Tabla 4.2 – continued from previous page

p	d	q	AIC
5	1	4	-2721.734
5	1	5	-2710.959

So the best model among this all possible model is ARIMA(5,1,4) and that has the minimum AIC value.

```
ARIMA(5,1,4)
```

Coefficients

	ar1	ar2	ar3	ar4	ar5	ma1	ma2	ma3	ma4
Estimate	0.4316	-0.0921	-0.3296	0.8887	-0.0321	-0.5187	0.1218	0.3121	-0.9151
Std. Error	0.0984	0.1043	0.0986	0.0698	0.0569	0.0833	0.1185	0.1184	0.0833

Sigma Squared

$$\sigma^2 = 2.125 \times 10^{-5}$$

Log Likelihood

$$\text{log likelihood} = 1370.87$$

AIC

$$\text{AIC} = -2721.73$$

Training Set Error Measures

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
0.0001	0.0046	0.0034	0.0078	0.3151	0.9693	-0.0007

Model fitted values for the train dataset are shown below.

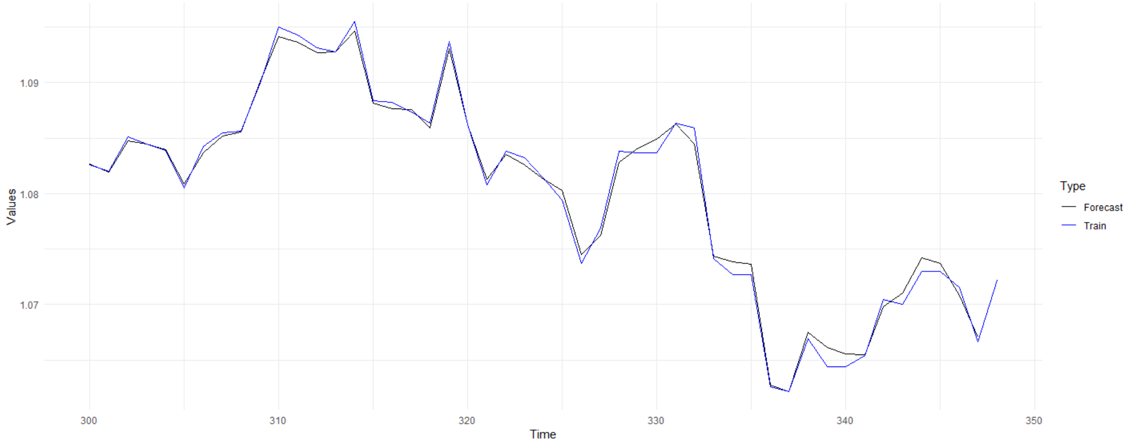


Figure 4.1: ARIMA(5,1,4) Model fit for 80% of dataset (train data)

Forecast the first ten days of the test data using the model and the previous fifteen days fitted values and the train data values are plotted below.

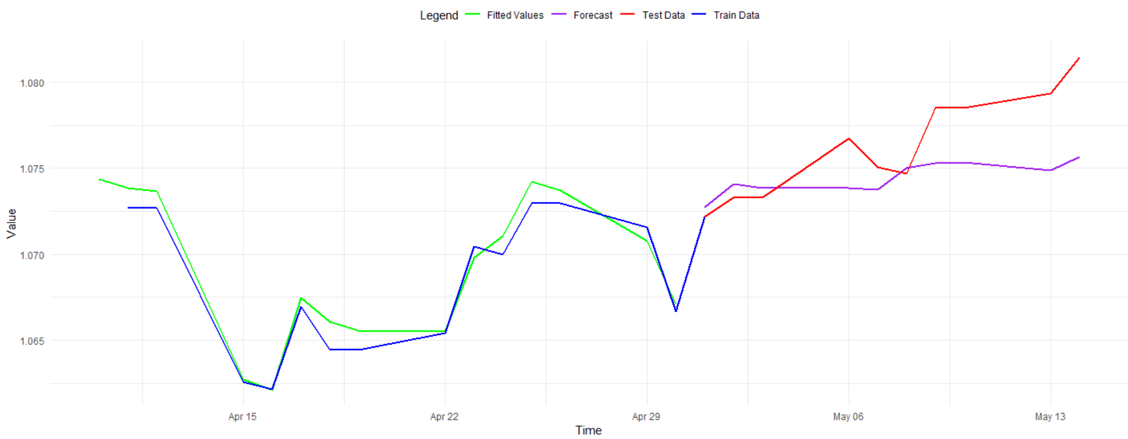


Figure 4.2: Test data and model forecast for ten days

The ARIMA model assume the residuals of the model is normally distributed with mean zero so the following figures shows the model residuals characteristics.

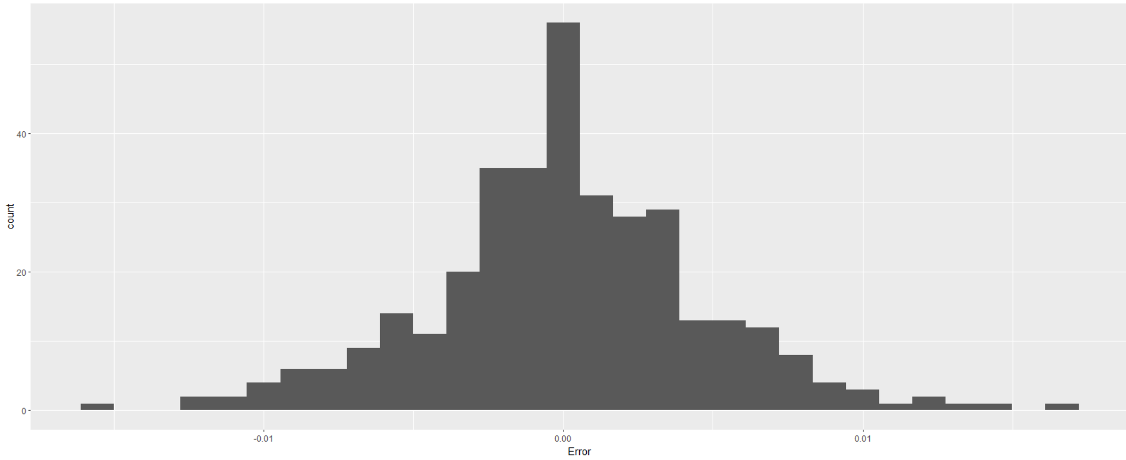


Figure 4.3: ARIMA(5,1,4) model residuals histogram

Statistic	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Value	-1.513e-02	-2.489e-03	1.053e-05	1.017e-04	2.784e-03	1.707e-02

Tabla 4.3: Summary statistics of residuals

So, the above histogram and the summary statistics of the residuals shows that the model is normally distributed with nearly mean zero, so the model fulfill the assumptions of the ARIMA model.

The results of the Box-Ljung test for squared residuals (y^2) are summarized in the table below:

Statistic	Value
Test Statistic (X^2)	0.56065
Degrees of Freedom (df)	1
p-value	0.454
Alternative Hypothesis	y is heteroscedastic

Tabla 4.4: Box-Ljung Test Results

Since the p-value (0.454) is greater than the typical significance level (e.g., 0.05), we fail to reject the null hypothesis. This indicates insufficient evidence to suggest that there is no any ARCH effect.

4.6 GBP/USD – British Pound / US Dollar

The same as before AR, MA or ARMA process is not followed by the dataset but the ARIMA model shows some useful results compared with other time series models. The following table shows the all possible models that can design using ARIMA, AIC (Akaike Information Criterion) is used to measure the accuracy of the model.

Tabla 4.5: All possible ARIMA models

p	d	q	AIC
0	1	0	-2574.346
0	1	1	-2572.427
0	1	2	-2573.851
0	1	3	-2571.924
0	1	4	-2569.924
0	1	5	-2568.537
1	1	0	-2572.411
1	1	1	-2571.826
1	1	2	-2571.916
1	1	3	-2569.882
1	1	4	-2567.968
1	1	5	-2566.544
2	1	0	-2573.775
2	1	1	-2571.963
2	1	2	-2567.583
2	1	3	-2567.933
2	1	4	-2569.464
2	1	5	-2564.338
3	1	0	-2571.993
3	1	1	-2569.777
3	1	2	-2568.020
3	1	3	-2572.863
3	1	4	-2571.479
3	1	5	-2569.343
4	1	0	-2570.016
4	1	1	-2568.181
4	1	2	-2569.764
4	1	3	-2571.581
4	1	4	-2567.786
4	1	5	-2572.801
5	1	0	-2568.611
5	1	1	-2566.677
5	1	2	-2571.460
5	1	3	-2568.621
Continued on next page			

Tabla 4.5 – continued from previous page

p	d	q	AIC
5	1	4	-2569.119
5	1	5	-2567.001
5	1	4	-2721.734
5	1	5	-2710.959

So the best model among this all possible model is ARIMA(0,1,0) and that has the minimum AIC value that means that the ARIMA model parameters according to random walk are only fitted for this currency pair.

```
ARIMA(0,1,0)
```

Sigma Squared

$$\sigma^2 = 3.492 \times 10^{-5}$$

Log Likelihood

$$\text{log likelihood} = 1288.17$$

AIC

$$\text{AIC} = -2574.35$$

Training Set Error Measures

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
0.0001	0.0059	0.0043	0.0089	0.3486	0.9979	-0.0146

Model fitted values for the train dataset are shown below.

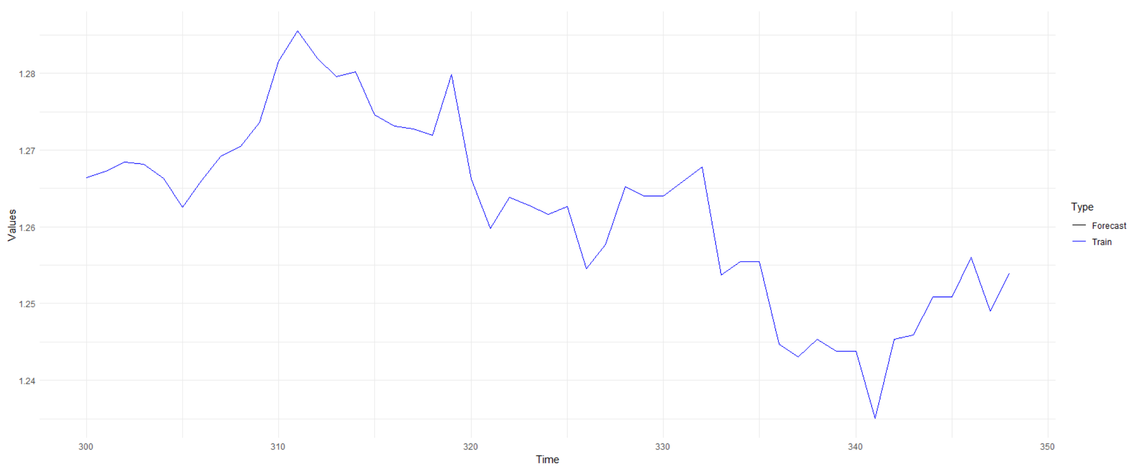


Figure 4.4: ARIMA(0,1,0) Model fit for 80% of dataset (train data)

Forecast the first ten days of the test data using the model and the previous fifteen days fitted values and the train data values are plotted below.

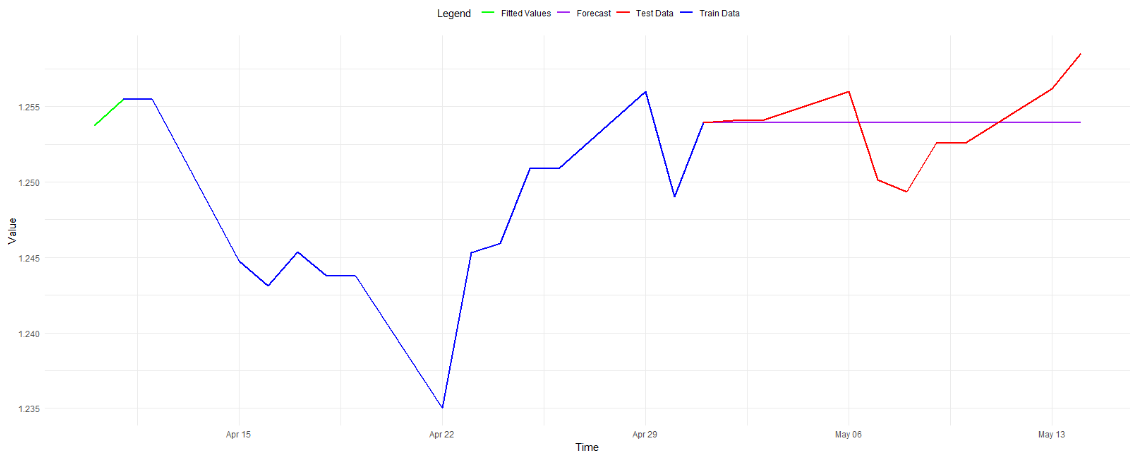


Figure 4.5: Test data and model forecast for ten days

The ARIMA model assume the residuals of the model is normally distributed with mean zero so the following figures shows the model residuals characteristics.

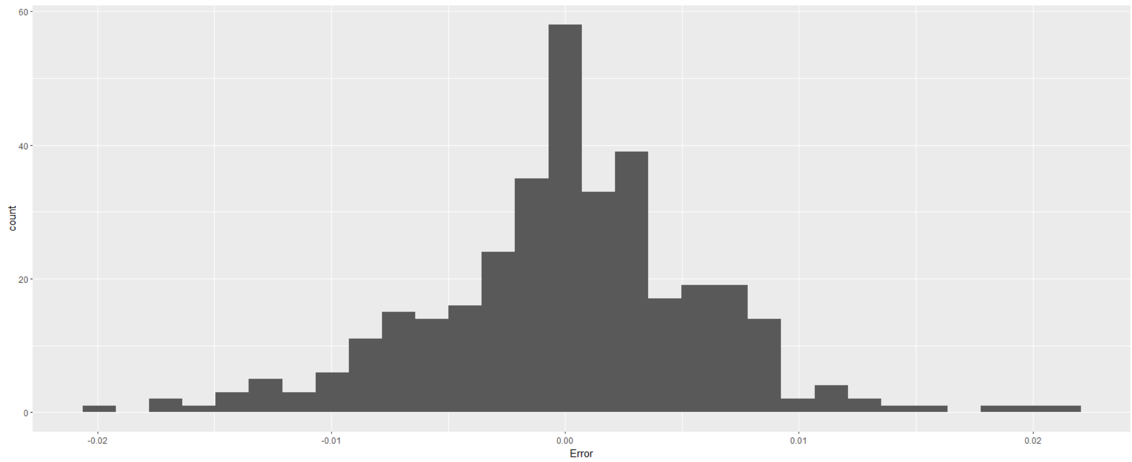


Figure 4.6: ARIMA(0,1,0) model residuals histogram

Statistic	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Value	-0.0198717	-0.002964	0.000000	0.0001239	0.0033707	0.0213896

Tabla 4.6: Summary statistics of residuals

So, the above histogram and the summary statistics of the residuals shows that the model is normally distributed with nearly mean zero, so the model fulfill the assumptions of the ARIMA model.

The results of the Box-Ljung test for squared residuals (y^2) are summarized in the table below:

Statistic	Value
Test Statistic (X^2)	3.674
Degrees of Freedom (df)	1
p-value	0.05527
Alternative Hypothesis	y is heteroscedastic

Tabla 4.7: Box-Ljung Test Results

Since the p-value (0.05527) is greater than the typical significance level (e.g., 0.05), we fail to reject the null hypothesis. This indicates insufficient evidence to suggest that there is no any ARCH effect.

4.7 USD/JPY – US Dollar / Japanese Yen

The ACF and PACF plots of USD/JPY currency pair isn't show any AR or MA process and there is no significance lags so explore a range of potential time series models. Select the model with the lowest AIC value as the most suitable representation of the data.

Tabla 4.8: All possible ARIMA models

p	d	q	AIC
0	1	0	862.4115
0	1	1	862.8473
0	1	2	862.7086
0	1	3	862.2603
0	1	4	863.8423
0	1	5	862.4385
1	1	0	863.1234
1	1	1	864.2763
1	1	2	863.8920
1	1	3	864.1358
1	1	4	864.0243
1	1	5	864.4196
2	1	0	861.8677
2	1	1	862.9610
2	1	2	861.9695
2	1	3	863.7056
2	1	4	864.8891
2	1	5	865.3390
3	1	0	861.5449
3	1	1	863.2430
3	1	2	862.2679
3	1	3	864.1481
3	1	4	865.1578
3	1	5	867.2365
4	1	0	862.6826
4	1	1	865.5224
4	1	2	864.4218
4	1	3	865.1857
4	1	4	865.6519
4	1	5	871.1308
5	1	0	862.6494
5	1	1	864.6292
5	1	2	866.5808
5	1	3	867.9231
5	1	4	869.5778

Continued on next page

Tabla 4.8 – continued from previous page

p	d	q	AIC
5	1	5	872.7277

So the best model among this all possible model is ARIMA(3,1,0) and that has the minimum AIC value.

```
ARIMA(3,1,0)
```

Coefficients

	ar1	ar2	ar3
Estimate	0.0609	-0.0915	-0.0827
Std. Error	0.0539	0.0540	0.0541

Sigma Squared

$$\sigma^2 = 0.685$$

Log Likelihood

$$\text{log likelihood} = -426.77$$

AIC

$$\text{AIC} = 861.54$$

Training Set Error Measures

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
0.0796	0.8265	0.5826	0.0538	0.4119	1.0123	-0.0057

Model fitted values for the train dataset are shown below.

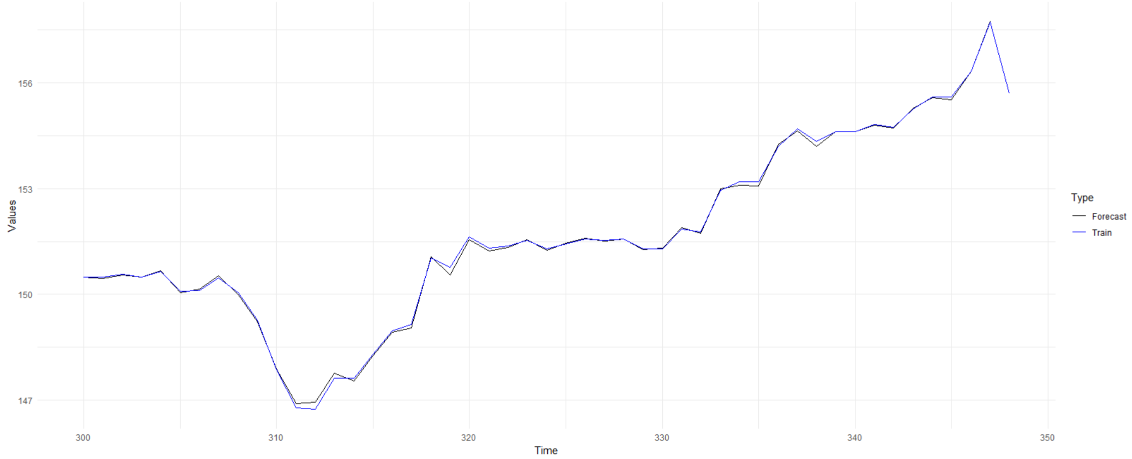


Figure 4.7: ARIMA(3,1,0) Model fit for 80% of dataset (train data)

Forecast the first ten days of the test data using the model and the previous fifteen days fitted values and the train data values are plotted below.

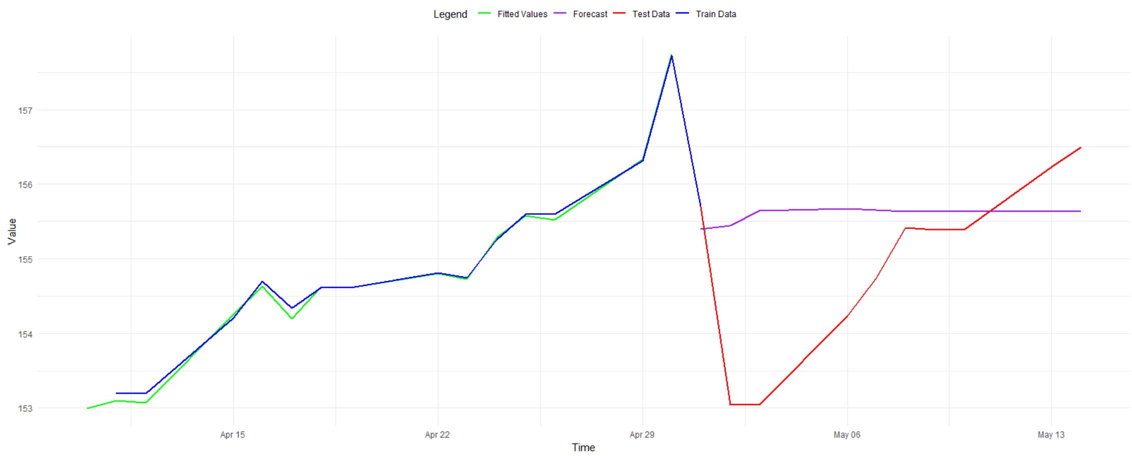


Figure 4.8: Test data and model forecast for ten days

The ARIMA model assume the residuals of the model is normally distributed with mean zero so the following figures shows the model residuals characteristics.

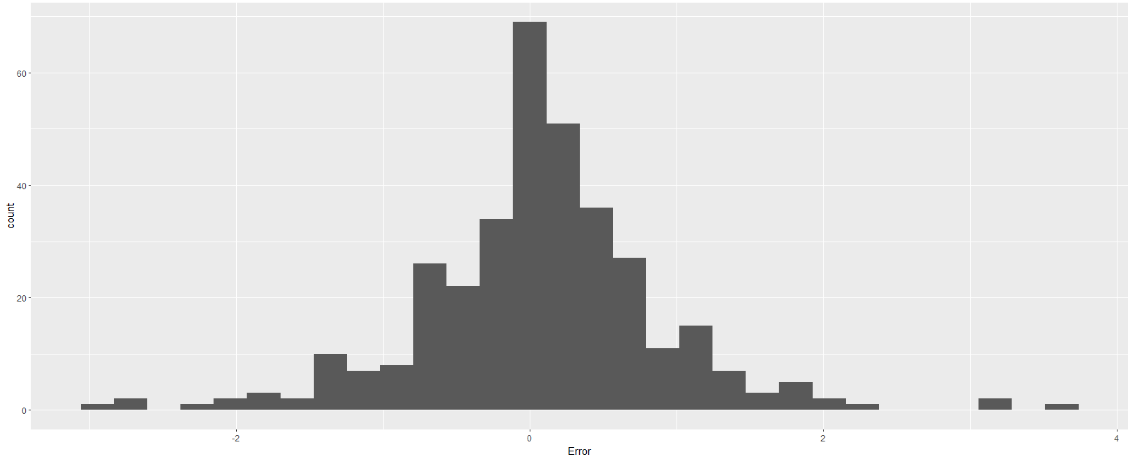


Figure 4.9: ARIMA(3,1,0) model residuals histogram

Statistic	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Value	-2.84174	-0.30975	0.08824	0.07966	0.51774	3.72950

Tabla 4.9: Summary statistics of residuals

So, the above histogram and the summary statistics of the residuals shows that the model is normally distributed with nearly mean zero, so the model fulfill the assumptions of the ARIMA model.

The results of the Box-Ljung test for squared residuals (y^2) are summarized in the table below:

Statistic	Value
Test Statistic (X^2)	0.083419
Degrees of Freedom (df)	1
p-value	0.7727
Alternative Hypothesis	y is heteroscedastic

Tabla 4.10: Box-Ljung Test Results

Since the p-value (0.7727) is greater than the typical significance level (e.g., 0.05), we fail to reject the null hypothesis. This indicates insufficient evidence to suggest that there is no any ARCH effect.

4.8 USD/CHF- US Dollar / Swiss Frank

The ACF and PACF plots of USD/CHF currency pair isn't show any AR or MA process but there are significance lags so explore a range of potential time series models. Select the model with the lowest AIC value as the most suitable representation of the data.

Tabla 4.11: All possible ARIMA models

p	d	q	AIC
0	1	0	-2796.694
0	1	1	-2794.735
0	1	2	-2793.893
0	1	3	-2794.223
0	1	4	-2792.518
0	1	5	-2792.627
1	1	0	-2794.730
1	1	1	-2792.806
1	1	2	-2792.606
1	1	3	-2792.340
1	1	4	-2792.317
1	1	5	-2791.805
2	1	0	-2794.006
2	1	1	-2790.730
2	1	2	-2789.181
2	1	3	-2790.333
2	1	4	-2788.361
2	1	5	-2788.660
3	1	0	-2794.881
3	1	1	-2793.016
3	1	2	-2795.973
3	1	3	-2801.771
3	1	4	-2799.988
3	1	5	-2788.054
4	1	0	-2793.171
4	1	1	-2790.893
4	1	2	-2793.980
4	1	3	-2791.975
4	1	4	-2804.851
4	1	5	-2804.324
5	1	0	-2792.276
5	1	1	-2790.893
5	1	2	-2792.016
5	1	3	-2790.161
5	1	4	-2788.081
Continued on next page			

Tabla 4.11 – continued from previous page

p	d	q	AIC
5	1	5	-2799.368

So the best model among this all possible model is ARIMA(4,1,4) and that has the minimum AIC value.

```
ARIMA(4,1,4)
```

Coefficients

	ar1	ar2	ar3	ar4	ma1	ma2	ma3	ma4
Estimate	-1.2044	-1.2913	-1.2228	-0.9191	1.2325	1.2894	1.2258	0.9958
Std. Error	0.0287	0.0287	0.0258	0.0252	0.0262	0.0256	0.0558	0.0458

Sigma Squared

$$\sigma^2 = 1.675 \times 10^{-5}$$

Log Likelihood

$$\text{log likelihood} = 1411.43$$

AIC

$$\text{AIC} = -2804.85$$

Training Set Error Measures

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-1.9099 $\times 10^{-5}$	0.004	0.003	-0.003	0.3392	0.9810	0.0002

Model fitted values for the train dataset are shown below.

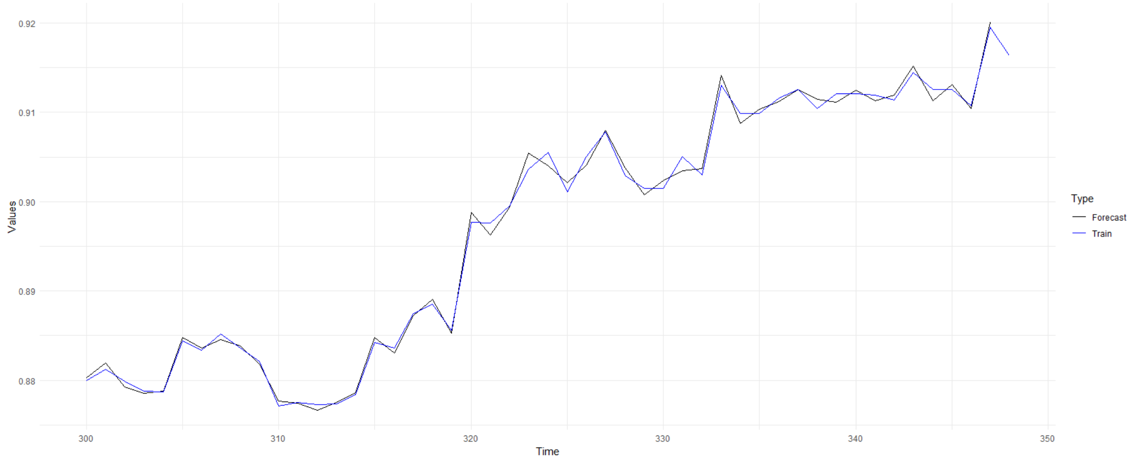


Figure 4.10: ARIMA(4,1,4) Model fit for 80% of dataset (train data)

Forecast the first ten days of the test data using the model and the previous fifteen days fitted values and the train data values are plotted below.

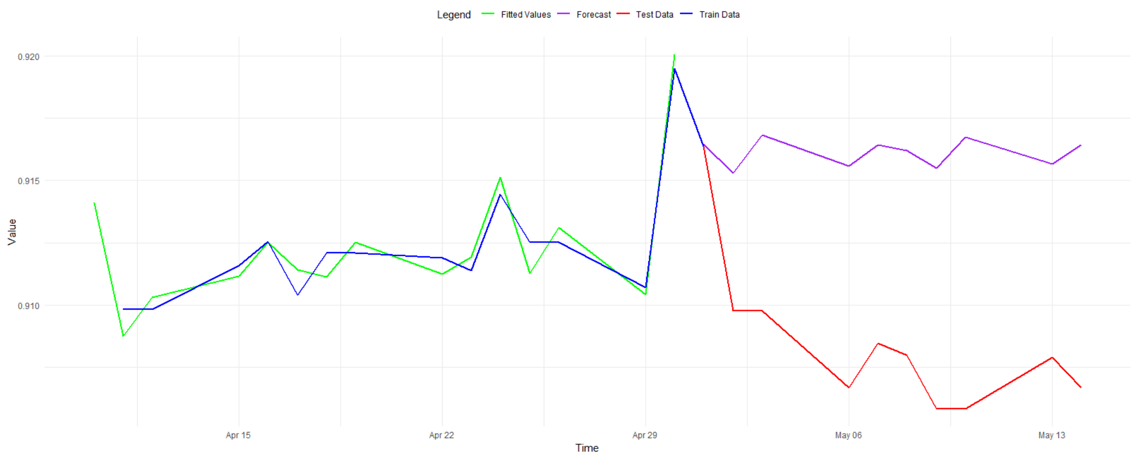


Figure 4.11: Test data and model forecast for ten days

The ARIMA model assume the residuals of the model is normally distributed with mean zero so the following figures shows the model residuals characteristics.

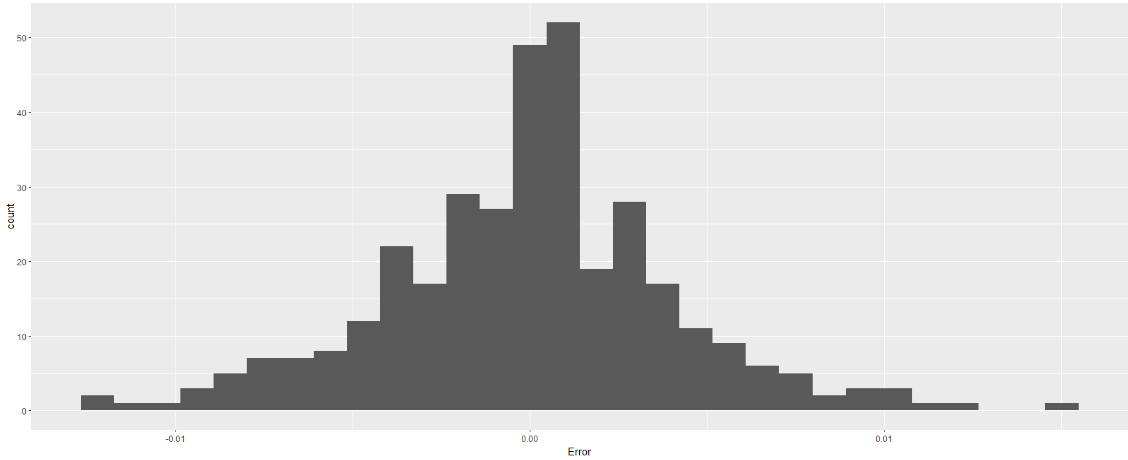


Figure 4.12: ARIMA(4,1,4) model residuals histogram

Statistic	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Value	-0.0125180	-0.0022678	0.0001106	-0.0000191	0.0023471	0.0146986

Tabla 4.12: Summary statistics of residuals

So, the above histogram and the summary statistics of the residuals shows that the model is normally distributed with nearly mean zero, so the model fulfill the assumptions of the ARIMA model.

The results of the Box-Ljung test for squared residuals (y^2) are summarized in the table below:

Statistic	Value
Test Statistic (X^2)	0.33659
Degrees of Freedom (df)	1
p-value	0.5618
Alternative Hypothesis	y is heteroscedastic

Tabla 4.13: Box-Ljung Test Results

Since the p-value (0.5618) is greater than the typical significance level (e.g., 0.05), we fail to reject the null hypothesis. This indicates insufficient evidence to suggest that there is no any ARCH effect.

4.9 USD/CAD – US Dollar / Canadian Dollar

The ACF and PACF plots of USD/CAD currency pair isn't show any AR or MA process and there is no any significance lags in PACF or ACF plots so explore a range of potential ARIMA models. Select the model with the lowest AIC value as the most suitable representation of the data.

Tabla 4.14: All possible ARIMA models

p	d	q	AIC
0	1	0	-2707.104
0	1	1	-2705.109
0	1	2	-2703.600
0	1	3	-2701.896
0	1	4	-2699.919
0	1	5	-2698.954
1	1	0	-2705.109
1	1	1	-2708.040
1	1	2	-2702.213
1	1	3	-2700.205
1	1	4	-2698.178
1	1	5	-2696.957
2	1	0	-2703.592
2	1	1	-2701.153
2	1	2	-2703.586
2	1	3	-2701.880
2	1	4	-2700.155
2	1	5	-2698.206
3	1	0	-2701.808
3	1	1	-2699.831
3	1	2	-2698.221
3	1	3	-2704.824
3	1	4	-2702.843
3	1	5	-2700.864
4	1	0	-2699.854
4	1	1	-2697.837
4	1	2	-2700.441
4	1	3	-2702.841
4	1	4	-2700.854
4	1	5	-2698.852
5	1	0	-2698.919
5	1	1	-2696.920
5	1	2	-2700.031
5	1	3	-2702.157

Continued on next page

Tabla 4.14 – continued from previous page

p	d	q	AIC
5	1	4	-2700.456
5	1	5	-2694.795

So the best model among this all possible model is ARIMA(1,1,1) and that has the minimum AIC value.

```
ARIMA(1,1,1)
```

Coefficients

	ar1	ma1
Estimate	0.9605	-1.0000
Std. Error	0.0165	0.0105

Sigma Squared

$$\sigma^2 = 2.334 \times 10^{-5}$$

Log Likelihood

$$\text{log likelihood} = 1357.02$$

AIC

$$\text{AIC} = -2708.04$$

Training Set Error Measures

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
0.0001	0.0048	0.0035	0.0073	0.2609	1.0048	0.02318907

Model fitted values for the train dataset are shown below.

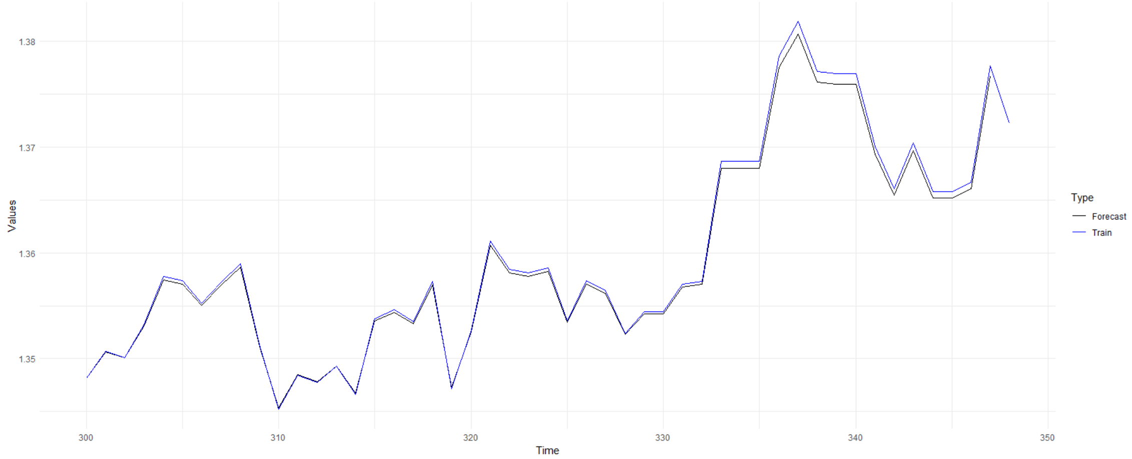


Figure 4.13: ARIMA(1,1,1) Model fit for 80% of dataset (train data)

Forecast the first ten days of the test data using the model and the previous fifteen days fitted values and the train data values are plotted below.

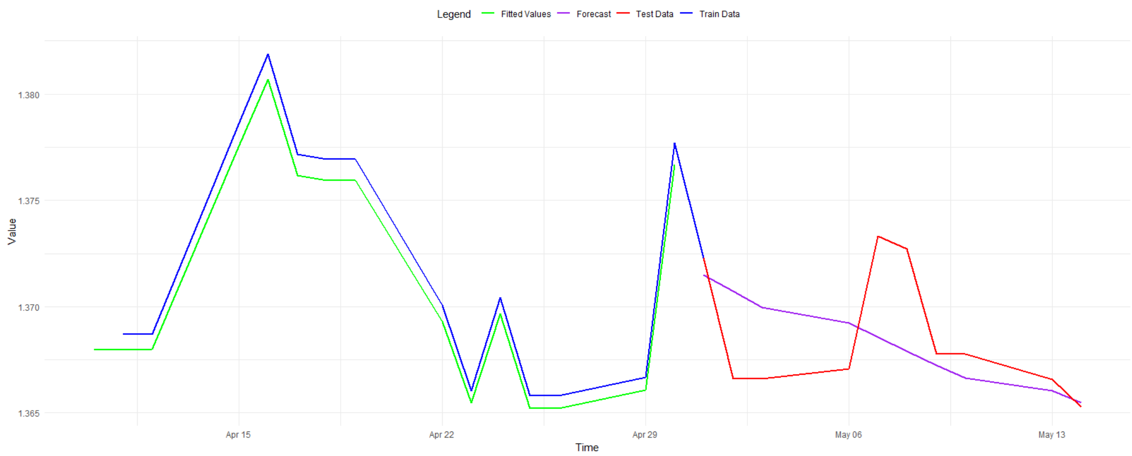


Figure 4.14: Test data and model forecast for ten days

The ARIMA model assume the residuals of the model is normally distributed with mean zero so the following figures shows the model residuals characteristics.

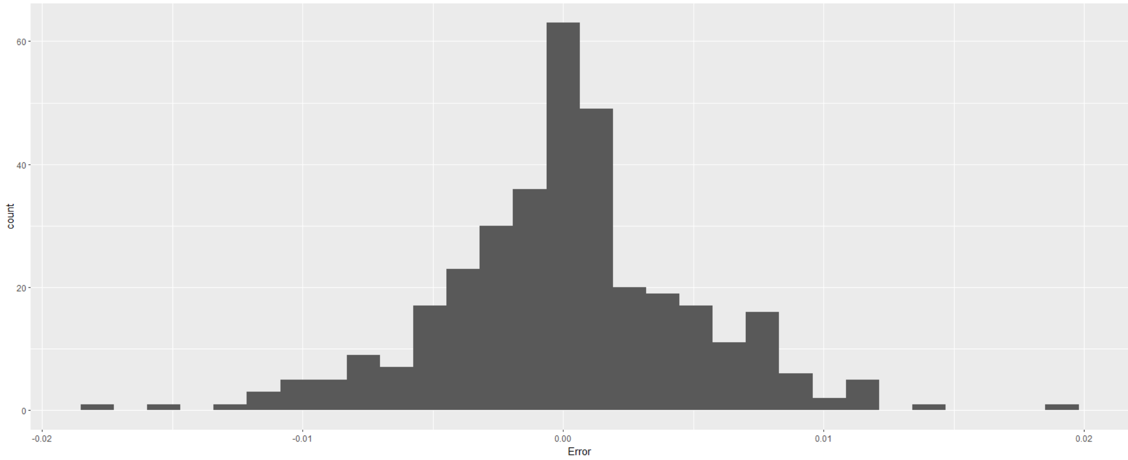


Figure 4.15: ARIMA(1,1,1) model residuals histogram

Statistic	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Value	-0.0184304	-0.0025425	0.0001136	0.0001149	0.0025380	0.0185987

Tabla 4.15: Summary statistics of residuals

So, the above histogram and the summary statistics of the residuals shows that the model is normally distributed with nearly mean zero, so the model fulfill the assumptions of the ARIMA model.

The results of the Box-Ljung test for squared residuals (y^2) are summarized in the table below:

Statistic	Value
Test Statistic (X^2)	0.032207
Degrees of Freedom (df)	1
p-value	0.8576
Alternative Hypothesis	y is heteroscedastic

Tabla 4.16: Box-Ljung Test Results

Since the p-value (0.8576) is greater than the typical significance level (e.g., 0.05), we fail to reject the null hypothesis. This indicates insufficient evidence to suggest that there is no any ARCH effect.

4.10 AUD/USD – Australian Dollar / US Dollar

The ACF and PACF plots of AUD/USD currency pair isn't show any AR or MA process and there is no any significance lags in PACF or ACF plots so explore a range of potential ARIMA models. Select the model with the lowest AIC value as the most suitable representation of the data.

Tabla 4.17: All possible ARIMA models

p	d	q	AIC
0	1	0	-2809.946
0	1	1	-2807.996
0	1	2	-2807.444
0	1	3	-2805.721
0	1	4	-2805.508
0	1	5	-2806.565
1	1	0	-2807.989
1	1	1	-2809.100
1	1	2	-2805.531
1	1	3	-2805.044
1	1	4	-2804.558
1	1	5	-2804.867
2	1	0	-2807.601
2	1	1	-2804.072
2	1	2	-2808.044
2	1	3	-2807.333
2	1	4	-2805.041
2	1	5	-2803.256
3	1	0	-2805.749
3	1	1	-2804.710
3	1	2	-2803.161
3	1	3	-2807.382
3	1	4	-2805.811
3	1	5	-2803.893
4	1	0	-2805.302
4	1	1	-2804.117
4	1	2	-2805.350
4	1	3	-2805.846
4	1	4	-2803.841
4	1	5	-2802.096
5	1	0	-2805.644
5	1	1	-2803.977
5	1	2	-2807.348
5	1	3	-2803.846

Continued on next page

Tabla 4.17 – continued from previous page

p	d	q	AIC
5	1	4	-2801.848
5	1	5	-2809.195

So the best model among this all possible model is ARIMA(0,1,0) and that has the minimum AIC value that means that the ARIMA model parameters according to random walk are only fitted for this currency pair.

```
ARIMA(0,1,0)
```

Sigma Squared

$$\sigma^2 = 1.771 \times 10^{-5}$$

Log Likelihood

$$\text{log likelihood} = 1405.97$$

AIC

$$\text{AIC} = -2809.95$$

Training Set Error Measures

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-8.1716 $\times 10^{-5}$	0.0042	0.0031	-0.0142	0.4679	0.9977	-0.0118

Model fitted values for the train dataset are shown below.

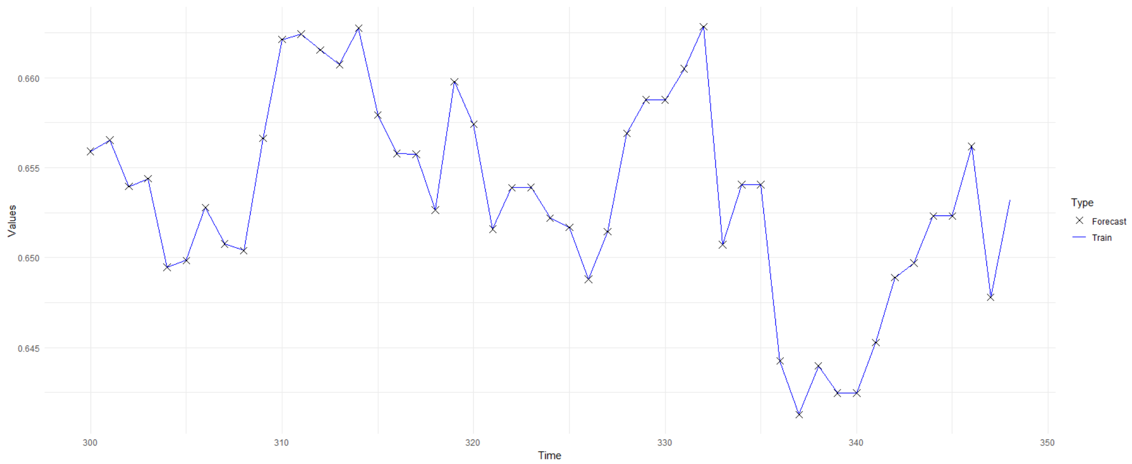


Figure 4.16: ARIMA(0,1,0) Model fit for 80% of dataset (train data)

Forecast the first ten days of the test data using the model and the previous fifteen days fitted values and the train data values are plotted below.

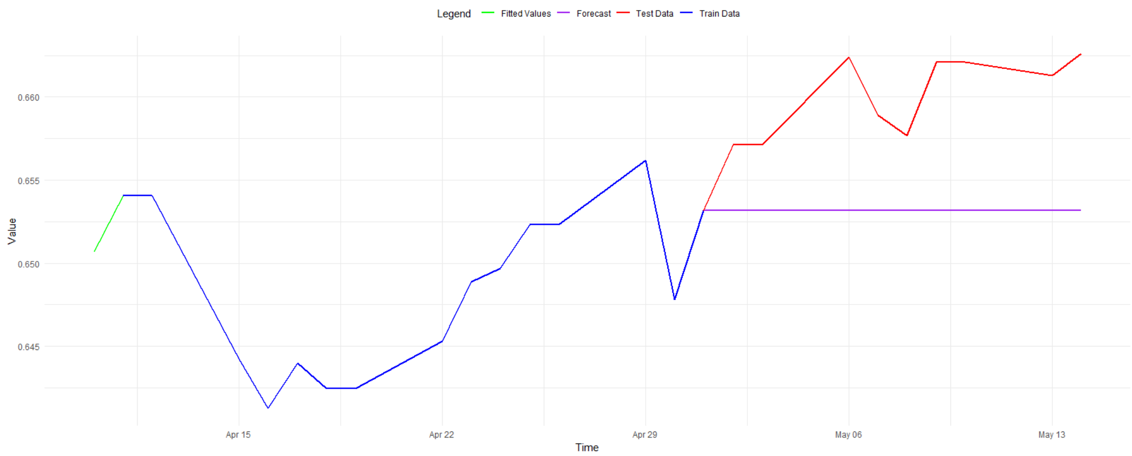


Figure 4.17: Test data and model forecast for ten days

The ARIMA model assume the residuals of the model is normally distributed with mean zero so the following figures shows the model residuals characteristics.

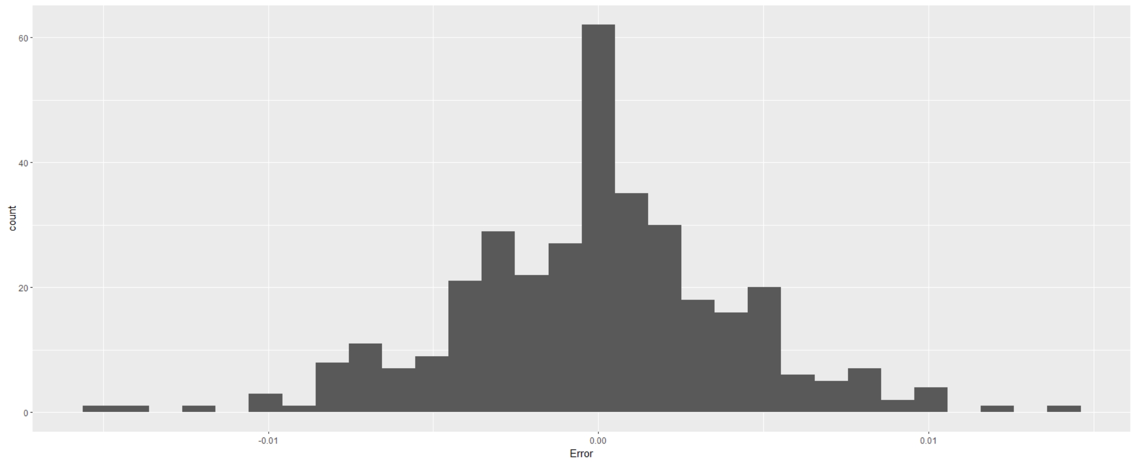


Figure 4.18: ARIMA(0,1,0) model residuals histogram

Statistic	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Value	-1.539 $\times 10^{-2}$	-2.748$\times 10^{-3}$	0.000	-8.172$\times 10^{-5}$	2.202$\times 10^{-3}$	1.383$\times 10^{-2}$

Tabla 4.18: Summary statistics of residuals

So, the above histogram and the summary statistics of the residuals shows that the model is normally distributed with nearly mean zero, so the model fulfill the assumptions of the ARIMA model.

The results of the Box-Ljung test for squared residuals (y^2) are summarized in the table below:

Statistic	Value
Test Statistic (X^2)	0.012356
Degrees of Freedom (df)	1
p-value	0.9115
Alternative Hypothesis	y is heteroscedastic

Tabla 4.19: Box-Ljung Test Results

Since the p-value (0.8576) is greater than the typical significance level (e.g., 0.05), we fail to reject the null hypothesis. This indicates insufficient evidence to suggest that there is no any ARCH effect.

4.11 NZD/USD – New Zealand Dollar / US Dollar

The ACF and PACF plots of AUD/USD currency pair isn't show any AR or MA process but lag 5 is a significance lag in PACF and ACF plots so explore a range of potential ARIMA models. Select the model with the lowest AIC value as the most suitable representation of the data.

Tabla 4.20: All possible ARIMA models

p	d	q	AIC
0	1	0	-2873.281
0	1	1	-2871.303
0	1	2	-2869.979
0	1	3	-2868.557
0	1	4	-2867.042
0	1	5	-2870.563
1	1	0	-2871.301
1	1	1	-2871.814
1	1	2	-2869.850
1	1	3	-2868.113
1	1	4	-2864.562
1	1	5	-2868.939
2	1	0	-2869.997
2	1	1	-2867.333
2	1	2	-2868.888
2	1	3	-2865.054
2	1	4	-2871.335
2	1	5	-2874.570
3	1	0	-2868.311
3	1	1	-2866.683
3	1	2	-2865.022
3	1	3	-2872.979
3	1	4	-2862.673
3	1	5	-2868.900
4	1	0	-2866.636
4	1	1	-2865.502
4	1	2	-2870.070
4	1	3	-2874.425
4	1	4	-2870.585
4	1	5	-2874.497
5	1	0	-2869.580
5	1	1	-2867.892
5	1	2	-2874.220
5	1	3	-2872.382
5	1	4	-2870.385
Continued on next page			

Tabla 4.20 – continued from previous page

p	d	q	AIC
5	1	5	-2879.736

So the best model among this all possible model is ARIMA(5,1,5) and that has the minimum AIC value.

```
ARIMA(5,1,5)
```

Coefficients

	ar1	ar2	ar3	ar4	ar5	ma1	ma2
Estimate	0.0757	-0.4052	0.3982	-0.0438	0.7722	-0.0825	0.4243
Std. Error	0.0643	0.0570	0.0602	0.0650	0.0512	0.0433	0.0311

	ma3	ma4	ma5
Estimate	-0.4274	0.1226	-0.9568
Std. Error	0.0383	0.0416	0.0346

Sigma Squared

$$\sigma^2 = 1.344 \times 10^{-5}$$

Log Likelihood

$$\text{log likelihood} = 1450.87$$

AIC

$$\text{AIC} = -2879.74$$

Training Set Error Measures

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-0.0002	0.0036	0.0027	-0.0435	0.4477	0.9900	0.0116

Model fitted values for the train dataset are shown below.

Figure 4.19: ARIMA(5,1,5) Model fit for 80% of dataset (train data)

Forecast the first ten days of the test data using the model and the previous fifteen days fitted values and the train data values are plotted below.

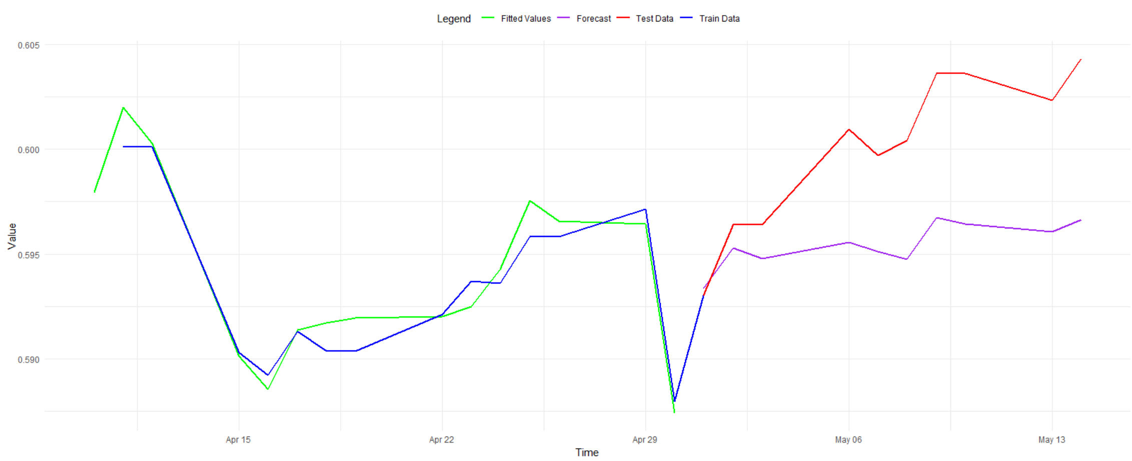


Figure 4.20: Test data and model forecast for ten days

The ARIMA model assume the residuals of the model is normally distributed with mean zero so the following figures shows the model residuals characteristics.

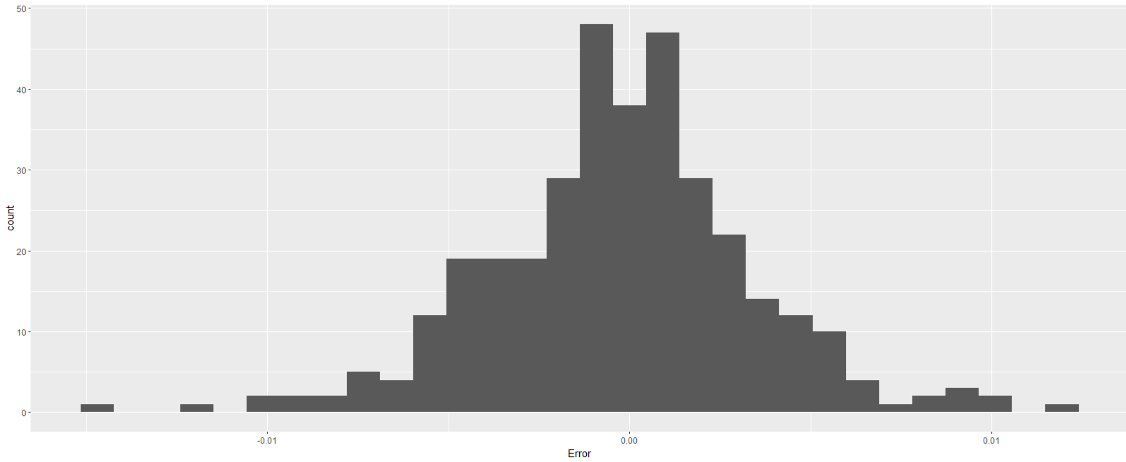


Figure 4.21: ARIMA(5,1,5) model residuals histogram

Statistic	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Value	-0.0142946	-0.0022426	-0.0001840	-0.0002516	0.0017563	0.0123491

Tabla 4.21: Summary statistics of residuals

So, the above histogram and the summary statistics of the residuals shows that the model is normally distributed with nearly mean zero, so the model fulfill the assumptions of the ARIMA model.

The results of the Box-Ljung test for squared residuals (y^2) are summarized in the table below:

Statistic	Value
Test Statistic (X^2)	0.13957
Degrees of Freedom (df)	1
p-value	0.7087
Alternative Hypothesis	y is heteroscedastic

Tabla 4.22: Box-Ljung Test Results

Since the p-value (0.7087) is greater than the typical significance level (e.g., 0.05), we fail to reject the null hypothesis. This indicates insufficient evidence to suggest that there is no any ARCH effect.

4.12 Results

The following table presents a summary of the fitted classical time series models along with their corresponding AIC values.

Currency pair	Fitted best model	AIC
EUR/USD – Euro / US Dollar	ARIMA(5,1,4)	2721.73
GBP/USD – British Pound / US Dollar	ARIMA(0,1,0)	2574.35
USD/JPY – US Dollar / Japanese Yen	ARIMA(3,1,0)	861.54
USD/CHF- US Dollar / Swiss Frank	ARIMA(4,1,4)	2804.85
USD/CAD – US Dollar / Canadian Dollar	ARIMA(1,1,1)	2708.04
AUD/USD – Australian Dollar / US Dollar	ARIMA(0,1,0)	2809.95
NZD/USD – New Zealand Dollar / US Dollar	ARIMA(5,1,5)	2879.74

Tabla 4.23: Fitted models for each currency pair

Chapter 5

Discussion and Conclusion

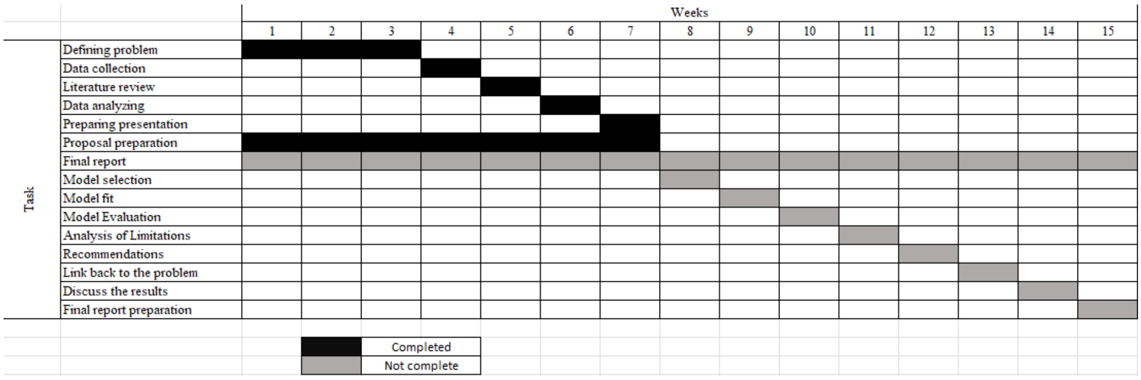
5.1 Discussion

AR, MA, ARMA, ARIMA, and ARCH were some of the models during the quantitative analysis. There was no presence of any AR or MA processes in the data set, due to which no appropriate model was identified with the use of ARMA. Then, the ARIMA model was applied, and all possible models were checked; the best model was selected using the lowest AIC value. The assumptions for the ARIMA model were checked, and any ARCH effect in the residuals was checked. But forecasts obtained from these models did not give significant results when compared to the test data. This was mainly due to the very high dependence of prices in a forex market which the analysis failed to capture.

5.2 Conclusion

Quantitative results therefore show that classical time series models cannot provide good results for the main forex currency pairs forecast. Thus, in this type of forecasting, more advanced models are called for to get meaningful and accurate results.

Time Schedule



Bibliography

[1] Vincenzo Pacelli. Forecasting exchange rates: A comparative analysis. International Journal of Business and Social Science, 3(10), 2012.

[2] Vaia I Kontopoulou, Athanasios D Panagopoulos, Ioannis Kakkos, and George K Matsopoulos. A review of arima vs. machine learning approaches for time series forecasting in data driven networks. Future Internet, 15(8):255, 2023.

[3] Robert F Engle and Tim Bollerslev. Modelling the persistence of conditional variances. Econometric reviews, 5(1):1–50, 1986.

[4] Martin D D Evans and Richard K Lyons. Meese-rogoff redux: Micro-based exchange-rate forecasting. American Economic Review, 95(2):405–414, 2005.

[5] George Box. Box and jenkins: time series analysis, forecasting and control. In A Very British Affair: Six Britons and the Development of Time Series Analysis During the 20th Century, pages 161–215. Springer, 2013.

[6] Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica: Journal of the econometric society, pages 987–1007, 1982.