MTH442 PROJECT: TIME SERIES ANALYSIS AND FORECASTING

Sanat Goel
Department of Mathematics & Statistics
Indian Institute of Technology Kanpur
sanatg21@iitk.ac.in

Shrish Shete
Department of Mathematics & Statistics
Indian Institute of Technology Kanpur
shrishs21@iitk.ac.in

Lakshika
Department of Mathematics & Statistics
Indian Institute of Technology Kanpur
lakshika21@iitk.ac.in

Om Shivam Verma
Department of Mathematics & Statistics
Indian Institute of Technology Kanpur
omsv21@iitk.ac.in

Pankajh Jhamtanil
Department of Mathematics & Statistics
Indian Institute of Technology Kanpur
pankajhj21@iitk.ac.in

November 13, 2023

ABSTRACT

This project focuses on the comprehensive analysis of historical foreign exchange (Forex) time series data to uncover valuable insights, trends, and patterns in currency markets. The objective is to provide a thorough understanding of currency exchange rate movements, develop predictive models, and explore risk management strategies.

1 Introduction

This project aims to analyze historical foreign exchange (Forex) time series data to understand currency market trends, develop forecasting models, and derive actionable insights. We have collected data from IMF This data includes daily values of different currencies. We collected the data for several currencies but mainly focused USD and INR for our analysis. We have done systematic analysis starting from data pre-processing, going through various types of statistical mainly time series analysis and end up having some important results and conclusions.

2 Data Preprocessing and Cleaning

We had data of almost 7600 days starting from 1994 to 2023. it consists of 11 variables, Canadian Dollars(CAD), Indian rupee(INR), Japanese yen (JPY), Kuwaiti dinar(KWD), Malaysian ringgit (MYR), New Zealand dollar (NZD), Singapore dollar (SGD), Thai baht (THB), U.A.E. dirham (AED), U.K. pound (GBP), U.S. dollar (USD). For the sake of simplicity we just opted for USD and INR, and did extensive analysis in these 2 variables.

Handling missing values in time series analysis requires careful consideration to ensure accurate and meaningful results. In the dataset at our disposal, nearly daily data was available, except for occasional missing values. To address this, we adopted a weekly aggregation approach, where we calculated the weekly average by aggregating the available data for each week. Subsequently, our analysis was conducted based on these weekly averages.

3 Time Series Analysis

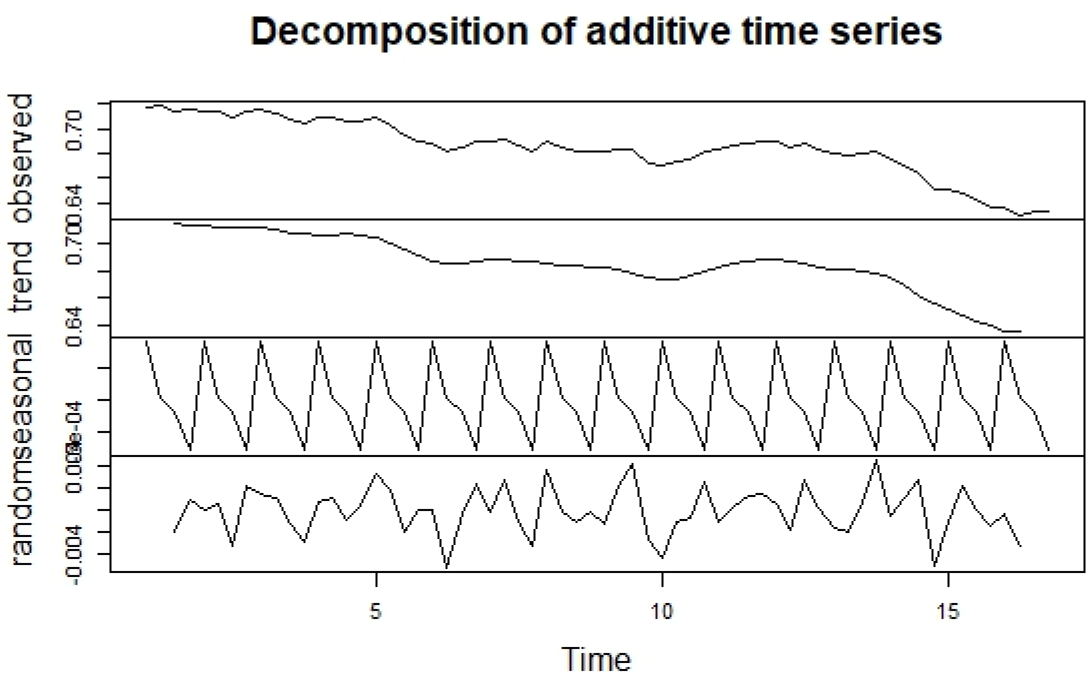
After having the preprocessed data the very first thing we have decided to do was the extraction of **trend**, **seasonality**, **cyclic component**, and **random component**.

Then we examined the time series, using various tests, differencing and then we removed seasonality and trend from the data> after having de-trended and de-seasonal data we looked for model fitting through in ARIMA through the inspection of AIC values. We found that the differencing of order 1 gave us a stationary time series, so the 2nd parameter in the ARIMA model becomes 1. and then at last we did the residual analysis to test the quality of the model.

Now we are going to describe each step explicitly.

3.1 Seasonal Decomposition

The decomposition technique is used to deconstruct the time series data into several components for visualization of time series characteristics. The period is chosen as 1 year, around 52 weeks. We will perform the decomposition using the additive model as the trend and seasonal variation vary over time.



Trend: The trend component is a general decreasing curve.

Seasonality: The seasonal component cycles on a yearly basis, which is a reasonable result.

Residuals: The residual component averages around the value of 0, which makes sense since we are using the additive model

The provided R code performs time series analysis on a dataset, presumably containing Indian Rupee exchange rate data.

3.2 Relative Ordering Test

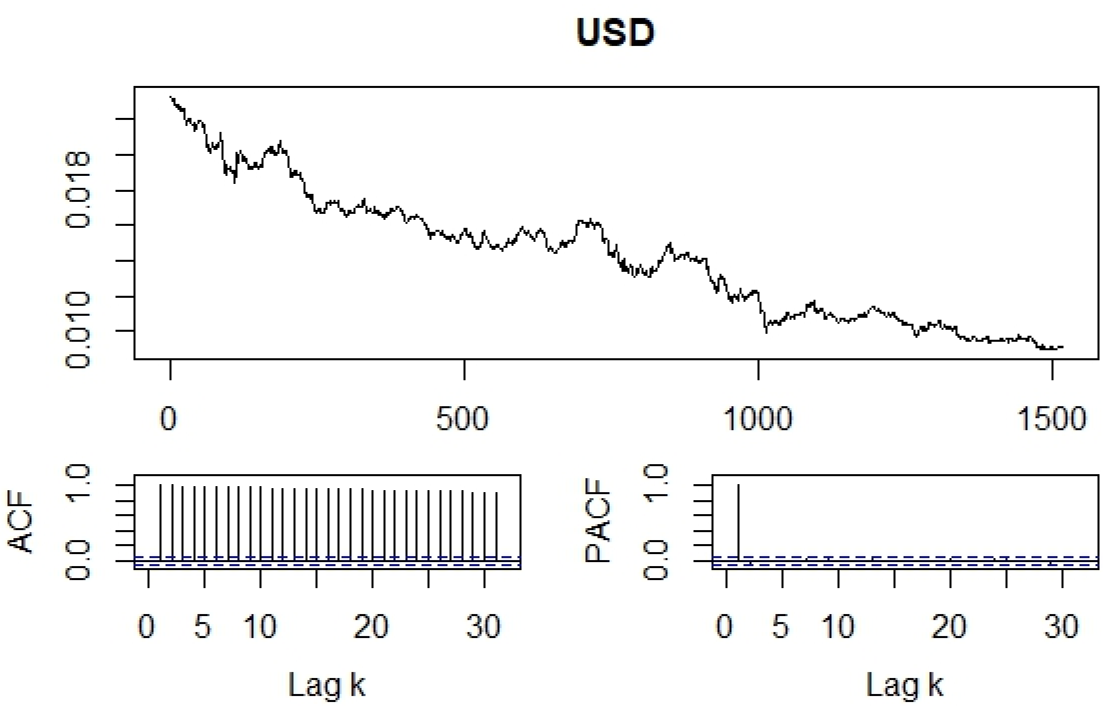
Relative Ordering Test is a non-parametric test procedure sed for testing existence of trend component

3.3 ADF Test

We used ADF test to check whether the time series is stationary or not.The output of the ADF test

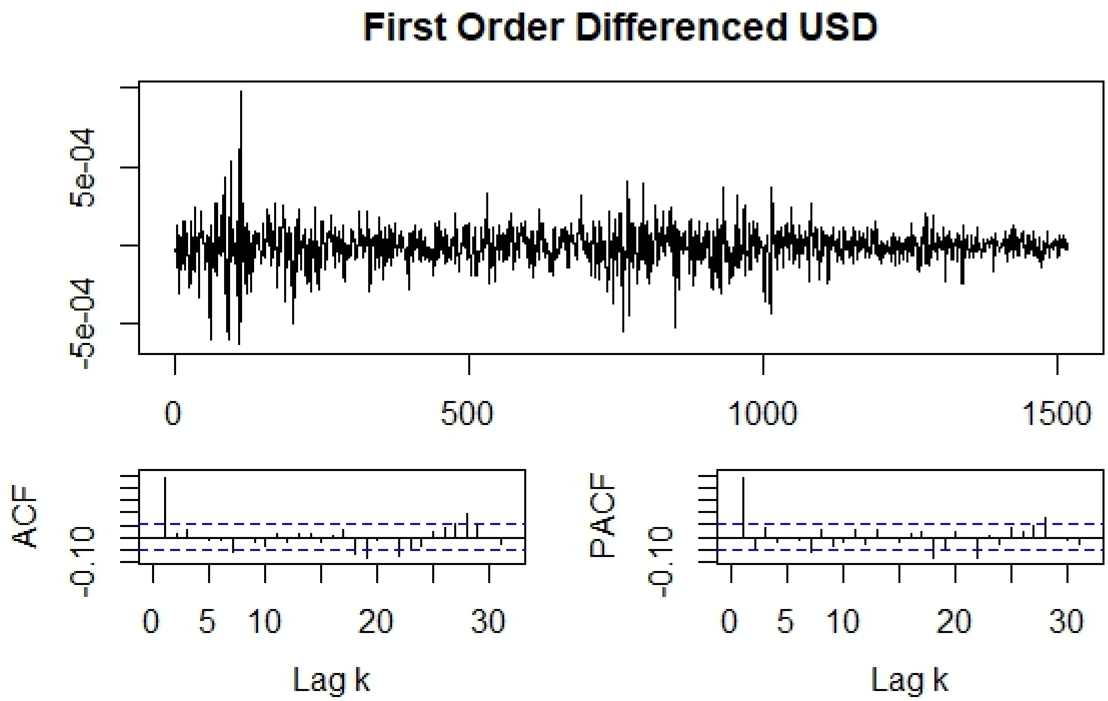
```
Augmented Dickey-Fuller Test

data:  timeseries
Dickey-Fuller = -1.8971, Lag order = 11, p-value = 0.6219
alternative hypothesis: stationary
```

By this, we found out that the time series is non-stationary as the p-value > 0.05.



3.4 Estimating d

When we did the test to check the presence of trend and seasonality then we found both of them are present in the data, so we applied first order differencing in the data and the time series we obtained was free from trend and seasonality. we tested the absence of seasonality and trend using the ADF test and ACF and PACF plots.



3.5 Constructing ARIMA Model using AIC

We have divided data into 2 parts i.e. test data and training data, we planned to find the parameters p and q from the training data and test the validation of model through the test data.

In the training data we were trying the optimum value for ARIMA model, by iterating through various combinations of p, q, and d. and choosing the one with the minimum AIC value.

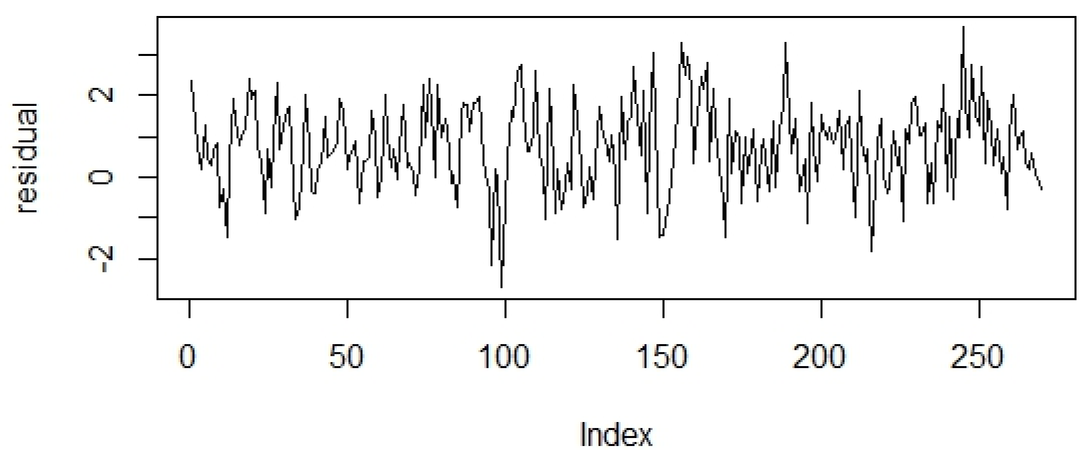
```
### AIC
p <- c(0, 1, 2, 3, 4)
d <- 1
q <- c(0, 1, 2, 3)

min_aic <- Inf
optimal_params <- NULL

for (i in p) {
  for (j in q) {
    tryCatch({
      fit <- arima(train$‘U.S. dollar (USD)                  ‘, order=c(i, d, j))
      current_aic <- AIC(fit)
      if (current_aic < min_aic) {
        min_aic <- current_aic
        optimal_params <- c(i, d, j)
      }
      print(paste("ARIMA(", i, ",", d, ",", j, ") - AIC:", current_aic, sep=""))
    }, error=function(e){})
  }
}
```

Output from this AIC test gives :

```
Best params: (1,1,0), min_aic: -10509.2178833479"
```

3.6 Validation

As mentioned in the previous section now we tried to validate the fitted model, we get p, d, q as 1, 1, 0respectively. Now we test the accuracy of the model by test data, we found the residuals and performed the residual analysis. The following graph for the Residual Analysis :



4 Conclusion

- Price of USD follows ARIMA(1, 1, 0)
- Seasonality and trend both were present but both vanished after 1st order differencing itself.