MODELING OF EXCHANGE RATES BY MULTIVARIATE ADAPTIVE REGRESSION SPLINES AND COMPARISON WITH CLASSICAL STATISTICAL METHODS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ECE KÖKSAL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
FINANCIAL MATHEMATICS

JUNE 2017

Approval of the thesis:

MODELING OF EXCHANGE RATES BY MULTIVARIATE ADAPTIVE REGRESSION SPLINES AND COMPARISON WITH CLASSICAL STATISTICAL METHODS

submitted by **ECE KÖKSAL** in partial fulfillment of the requirements for the degree of **Master of Science in Department of Financial Mathematics, Middle East Technical University** by,

Prof. Dr. Bülent Karasözen
Director, Graduate School of **Applied Mathematics** ______________

Assoc. Prof. Dr. Yeliz Yolcu Okur
Head of Department, **Financial Mathematics** ______________

Prof. Dr. Gerhard-Wilhelm Weber
Supervisor, **Financial Mathematics, IAM, METU** ______________

Examining Committee Members:

Prof. Dr. Gerhard-Wilhelm Weber
Financial Mathematics, IAM, METU ______________

Assoc. Prof. Dr. Sevtap Kestel
Actuarial Sciences, IAM, METU ______________

Assist. Prof. Dr. Özlem Defterli
Mathematics, Çankaya University ______________

Date: ______________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: ECE KÖKSAL

Signature :

ABSTRACT

MODELING OF EXCHANGE RATES BY
MULTIVARIATE ADAPTIVE REGRESSION SPLINES AND COMPARISON
WITH CLASSICAL STATISTICAL METHODS

Köksal, Ece
M.S., Department of Financial Mathematics
Supervisor : Prof. Dr. Gerhard-Wilhelm Weber

June 2017, 77 pages

Economic factors like inflation, interest rates and exchange rates are among the leading indicators of a country's relative level of economic health. With the help of technological improvements and global requirements, trading volume and a wide range of commerce network, exchange rates play a vital role in economics and finance since a higher exchange rate may result in a lower trade balance of a country, whereas a lower rate may cause an increase. Inflation, interest rates, domestic money supply growth, a country's balance of payments' size and trend, a country's economic growth, dependency on outside sources and central bank intervention, are the factors which affect an exchange rate. Since many dependent and independent factors affect exchange rates, it is difficult to predict them. In areas of application, data mining is frequently used for decision support, financial forecasting, marketing strategy, prediction, etc. The method of data mining and machine learning is applied to analyze and forecast the future behavior of such complex systems. Modeling and prediction of exchange rates are still a challenge, although mathematicians, economists and statisticians have worked to reach a model with a superior forecasting ability for many years. Therefore, in this study, we aim to generate mathematical models to forecast the monthly US Dollar (USD) / Turkish Lira (TRY) and Euro (EUR) / Turkish Lira (TRY) exchange rates via data mining tools. For this purpose, we apply a flexible model Multivariate Adaptive Regression Splines (MARS) and widely used models Linear Regression (LR) and Support Vector

Regression (SVR). In this study, MARS, LR and SVR models applied on USD / TRY and EUR / TRY exchange rate data sets in the period of 01/01/2007 and 30/04/2015; then the results of these models are compared and found out that MARS method has superior forecasting ability over LR and SVR methods for USD / TRY and EUR / TRY exchange rates. The thesis ends with a conclusion and an outlook to future investigations.

Keywords : Exchange-Rate Forecast, Linear Regression, Support Vector Regression, Multivariate Adaptive Regression Splines, Optimization, Economics, Finance.

ÖZ

ÇOK DEĞİŞKENLİ UYARLANABİLİR REGRESYON EĞRİLERİ İLE DÖVİZ KURU MODELLEMESİ VE KLASİK İSTATİSTİKSEL YÖNTEMLERLE KARŞILAŞTIRILMASI

Köksal, Ece
Yüksek Lisans, Finansal Matematik Bölümü
Tez Yöneticisi : Prof. Dr. Gerhard-Wilhelm Weber

Haziran 2017, 77 sayfa

Enflasyon, faiz oranları ve döviz kuru gibi ekonomik faktörler, bir ülkenin ekonomik düzeyinin önde gelen göstergelerindendir. Küresel gereksinimler ve teknolojik gelişim sayesinde ve genişleyen ticari hacim ve ağlar ile döviz kurları, ekonomi ve finans alanlarında önemli bir rol oynar; çünkü yüksek bir döviz kuru, bir ülkenin dış ticaret dengesinde düşüşe neden olabilirken, daha düşük bir oran artışa neden olabilir. Enflasyon, faiz oranları, yerel para arzındaki artış, bir ülkenin ödemeler dengesi ve eğilimi, ekonomik büyümesi, dış kaynaklara bağımlılığı ve merkez bankası müdahalesi gibi etkenler döviz kurunu etkileyen faktörlerdir. Bunlara benzer birçok bağımlı ve bağımsız faktör döviz kurunu etkiler ve bu nedenle döviz kurlarını tahmin etmek zordur. Döviz kurlarını tahminlemek amacıyla veri madenciliği, mali tahminler, pazarlama stratejileri benzeri metodlar sıklıkla kullanılır. Veri madenciliği ve makine öğrenimi yöntemi, bu karmaşık sistemlerin gelecekteki davranışlarını analiz etmek ve tahmin etmek için uygulanabilir. Matematikçiler, ekonomistler ve istatistikçiler yıllarca üstün tahmin yeteneği olan bir modele erişmek için uğraşmış olsalar da döviz kurlarının modellenmesi ve tahmini çalışmaları hala devam etmektedir. Bu nedenle, bu çalışmada, veri madenciliği araçları aracılığıyla aylık ABD Doları (USD) / Türk Lirası (TRY) ve Avro (EUR) / Türk Lirası (TRY) kurlarını tahmin etmek için matematiksel modeller üretmeyi amaçlıyoruz. Bu amaçla esnek bir model olan Çok Değişkenli Uyarlamalı Regresyon Şemaları (MARS) ve yaygın olarak kullanılan modeller olan Lineer Regression (LR) ve Support Vector Regression (SVR) methodları uygulanmaktadır.

Bu çalışmada, 01/01/2007 ve 30/04/2015 dönemine ait USD / TRY ve EUR /TRY kur verilerine uygulanan MARS, LR ve SVR modelleri ve bu modellerin sonuçları karşılaştırılmıştır. Karşılaştırma sonucunda MARS methodunun LR ve SVR methodlarına göre USD / TRY ve EUR /TRY kur modelleri için daha üstün modelleme yeteneği olduğu gözlemlenmiştir. Tez, sonuç ve gelecek araştırmalara yönelik bir görüş ile sona ermektedir.

Anahtar Kelimeler: Döviz Kuru Tahminleri, Lineer Regresyon, Destek Vektör Regresyonu, Çok Değişkenli Uyarlamalı Regresyonlu Şemaları, Optimizasyon, Ekonomi, Finans.

To

My parents: Hasan & Emel
This thesis is dedicated to my father, who courages me and advises me to do my best in the field of studies. It is also dedicated to my mother, who supports and uses every means available to help me during my thesis study. Thank you for your supports with my studies and I am honored to have you as my parents.

My brother: Can
You're not only my brother; you are also my colleague, my defender, my mentor, and my help in time of need. I am grateful to you for your precious ideas and helps.

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my thesis supervisor Prof. Dr. Gerhard-Wilhelm Weber for his patient guidance, continuous support of my graduate study and research, enthusiastic encouragement , motivation, immense knowledge and valuable advices during the development and preparation of this thesis. His willingness to give his time and to share his experiences has brightened my path and helped me in all the time of research and writing of this thesis. It is my honor to be a student of such an excellent and productive academician and great person.

Besides my advisor, I would like to thank to Dr. Ayşe Özmen for her share of experiences and expertise on modelling, and I express many great appreciation to her patience, endless guidance and valuable advices during this study.

I also would like to thank Assist. Prof. Dr. Seza Danışoğlu who kindly answered all my questions for helping me in data collection and sharing her precious finance knowledge.

Many thanks to Mertcan Tarhan, my friends Oğuzhan Özcan, Nurhan Tütüncü, Yahya Karakuş, Ali Eatamadi, Ali Afshar and all my further friends for their invaluable friendship and motivation. Their efforts helped me to overcome my concerns and helped me to move on. Also I would like to thank Mr. Udo Bassey for his valuable supports.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ε	Noise Term
σ^2	Variance (Error Variance)
AR	Auto Regression
ARFIMA	Auto Regressive Fractional Integrated Moving Average
ARIMA	Auto Regressive Integrated Moving Average
BF	Basis Function
CSI	Cubic Spline Interpolation
EEMD	Ensemble Empirical Mode Decomposition
EMS	European Monetary System
EUR	Euro
GARCH	Generalized Auto Regressive Conditional Heteroskedasticity
GCV	Generalized Cross Validation
GDP	Gross Domestic Product
INF	Inflation Rate
LOF	Lack of Fit
LR	Linear Regression
LS	Least Square
MO	Months
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MARS	Multivariate Adaptive Regression Splines
ME	Mean Error
MSE	Mean Square Error
NN	Neural Network
R^2	Coefficient of Determination
RMSE	Root Mean Square Error
RW	Random Walk
SVM	Support Vector Machine
SVR	Support Vector Regression
TRY	Turkish Lira

USD	US Dollar
VAR	Vector Auto Regression

CHAPTER 1

INTRODUCTION

For the most basic definition, exchange rate is the quotient of one national currency with respect to another foreign currency. In other words, exchange rates can be regarded as the values of one nation's currency in terms of other countries' currencies. There are various types of exchange rates which will be displayed subsequently.

A first categorization is set according to forex market which is accessible for a wide range of buyers and sellers where currency trading continues for 24 hours a day, except for holidays and weekends. In this category, it is usually hard for traders to distinguish nominal exchange rate and real exchange rate. Forex market set exchange rates are in continuous quotation. The rates which we follow with news or newspapers are this kind of rates. However, real exchange rates are the adjusted nominal exchange rates by inflation measure.

A second categorization of exchange rates is based on the number of currencies which are taken into account. Bilateral exchange rates are related to two countries' exchange rates. These rates are the results of matching demand and supply balance in banking transaction or in financial markets. Multilateral exchange rates, on the other hand, are used to judge the general dynamics of a country's currency toward other nations' currencies. If a country enforces the existence of more than one exchange rate, then multiple exchange rate s exist. But it requires the degree of capital control.

A third categorization is set according to exchange rates movements. Fluctuation of exchange rates have an important effect on companies' returns, especially, for multinational companies. For this reason, many companies can choose to use the forward exchange rate to protect themselves from financial risks. The forward exchange rate is fixed today to exchange the currency according to that rate on a specific future date while the spot rate is the current exchange rate in the market.

After economic factors like inflation and interest rates, exchange rates are one of the leading indicators of a country's relative level of economic health. With the help of technological improvements and global requirements, trading volume and a wide range of commerce network, exchange rates play a vital role in any country's level of trade. Moreover, they affect and are affected by international trade in free-market which helps to sustain a balance of trade and a balance of capital. Because of this reason, exchange rates can be analyzed and manipulated since they have an effect on an investor's real rate of return. For instance, a higher currency rate results in more expensive exports and cheaper imports in foreign markets, while a lower currency rate results in cheaper exports and more expensive imports with respect to a foreign market. As a consequecny, a higher exchange rate may result in a lower balance trade of a country, whereas a lower rate may cause an increase. For these reasons, exchange rates are quite important in economics and finance.

Buying and selling of a currency determine the actual price of an exchange rate. Moreover, the exchange rates are also fixed by demand and supply power like commodities. However, supply and demand can be affected by many factors. Inflation, interest rates, domestic money supply growth, a country's balance of payments' size and trend, a country's economic growth, dependency on outside sources and central bank intervention are the factors which affect an exchange rate. Since many dependent and independent factors affect exchange rates, it is difficult to predict them. For instance, higher interest rates create money saving among investors, and for this reason there exists an inflow of hot money and appreciation in exchange rates. Similarly, if a country's economy is going down, interest rates will also go down since the Central Bank cuts interest rates due to inflation decrease. As a result, the currency of a country begins to depreciate. For these reasons, it is very difficult to predict exchange rates.

In our study, we tried to introduce an innovative approach to exchange rate modeling by using MARS methodology. Since real-life problems and natural phenomena show a nonlinear behavior, nonparametric regression techniques - including MARS - are the most powerful tools to build flexible models for high-dimensional nonlinear data [18]. The main feature of MARS is the ability to automatically build nonlinear models and models with interactions, which is very important in classification and regression. MARS has this ability with the help of piecewise linear one-dimensional basis functions BFs which take into account additive and interactive effects of predictors to reach the response variable.

To do that, MARS has two stages, known as forward and backward stage. In the forward stage, MARS adds Basis Functions (BFs) until it reaches a highest complexity which causes an over-fitted model. At this point, MARS passes to the backward stage by eliminating multi-variate BFs to get the eventual model. This elimination is data-based and also specific for the analysis of MARS, hence MARS is a very useful and helpful adaptive regression algorithm for high-dimensional data modelling. Hitherto, MARS has been employed for applications in many areas, including energy market [49, 50], banking sector [1], engineering [48], marketing and finance sector [2, 32], medical science [9], but MARS has not been used to model exchange rates. However, MARS method was employed for most appropriate variable set selection to propose and test exchange rate forecasting models by Plakandaras, Papadimitriou and Gogas in their study "Forecasting Daily and Monthly Exchange Rates with Machine Learning Techniques" [37].

The aim of the study [37] was to develop, test and compare a forecasting exchange rate model for 5 selected currencies at both monthly and daily bases. To do that, the authors employed different Machine Learning (ML) methodologies and also employed MARS method for variable selection [40]. They employed Ensemble Empirical Mode Decomposition (EEMD) method on exchange rate data for decomposition, then MARS worked on smoothed data to select the most proper input data set to be fed into two diversified Support Vector Regression (SVR) model [11, 51]. These two SVR models provided one-period-ahead forecasts and a Neural Network (NN) was also used as an alternative to SVR for comparison. At the final stage of their study, the authors of [37] compared the sum of forecasted components to decide about the superior forecasting model for these selected 6 currencies [37]. Differently from these studies, we attempted to employ MARS as a modelling tool to forecast exchange rates. To do that we used the TL/USD and TL/EUR rates [18].

Chapter 2 prepared for a literature review of exchange rate modelling. In Chapter 3, we explain the methodologies which we applied in our study including MARS, Cubic Spline Interpolation (CSI), Linear Regression and Support Vector Regression (SVR), respectively [24]. Chapter 4 covers the data used in MARS model. We introduced our eventual exchange rate forecasting model by MARS, Linear Regression model and Support Vector Regression model in Chapter 5 and compared the performances of these models in Chapter 6. Eventually, Chapter 7 stands for our conclusions and the overall findings of our study, as well as for an outlook to future works.

CHAPTER 2

LITERATURE REVIEW

Even if it is hard to predict exchange rates, economists and mathematicians have been trying to establish an exchange rate prediction model. The first attempt made to predict exchange rates is the "Bretton Wood Fixed Exchange Rate System" [5]. Since the end of World War 2, the US and the US dollar have held powerful positions in international trade. In 1944, the Bretton Wood Agreement was signed at United Nations Monetary and Financial Conference in Bretton Wood, New Hampshire. According to this agreement, the International Monetary Fund, International Bank for Reconstruction and Development and the fixed exchange rate system were formed. Since the US was the superior power and holding most of the gold in the world, it was decided that all currencies tie to dollar, and are convertible to gold at 35 per ounce. Moreover, central banks of other countries could maintain a fixed exchange rate between dollar and their currencies.

In closer detail, if the currency of a country is less valuable against dollar, its central bank could sell its own currency in exchange for dollar. This fixed exchange rate system continued until 1971. Growing trading deficit and inflation rates in the US were determining the value of the US dollar in those years. Japan and Germany were two powerful competitors for the US and the value of their currencies started to increase. With these developments, the US left the fixed value of exchange rate. In 1971, the Smithsonian Agreement was signed instead of the Bretton Wood agreement but it did not work [20]. In 1973, all nations agreed to use floating exchange rates. However, this floating system was called managed float regime by economists since central banks had the right to interfere in the rates. Since exchange rate forecasting has a quite high importance for an economy, following the breakdown of Bretton Woods fixed exchange rate system, economists, mathematicians and statisticians tried to build large numbers of models to forecast exchange rates. All these model building attempts can be separated into two main categories - monetary exchange rate models and models - that are built on a micro-structural approach.

The first category, known as monetary exchange rate models, mainly focuses on the effect of macroeconomic variables on exchange rates and for this reason they are also called monetary exchange rate models. The leading studies on this category are Mundell's classic model published in 1968 [29], Fleming's work published in 1962 [16] and Dornbuch's sticky price monetary model built in 1976 [14]. However, the flexible price monetary model studied by Bilson in 1978 [4], Stockman in 1980 [45], Lucas in 1982 [22] and the extension of flexible price monetary model with interest rates by Frankel in 1979 [17] were used as a benchmark for years. Nevertheless, the article “Empirical Exchange Rate Models of the Seventies” by Meese and Rogoff changed the benchmark [25]. The aim of their work was to test different monetary exchange rate models against random walk (RW) with drift model in out-of-sample forecasting. To do that, they tested and compared time series and structural exchange rates with regard to their out-of-sample forecasting accuracy. Their study revealed that regardless of the version of the model used, none outperforms the random walk in terms of Mean Square Error (MSE). This outcome was applied as new benchmark in finance and economy for years to forecast exchange rates.

Firstly, they started with a structural model and selected three assets models to eventually dominate the current literature on exchange rate determination. They made this selection with the consideration of relative tractability of data requirements. These selected models were flexible price monetary model of Frankel and Bilson, the “sticky price” monetary model of Dornbuch and Frankel and the “sticky price” monetary model of Hooper and Morton. All these selected models suggested that the exchange rate exhibits first degree homogeneity in relative money supplies. However, the Frankel-Bilson model assumes that the purchasing power parity is zero while the Dornbuch-Frankel model allows for a slow domestic price adjustment and the consequent deviations from purchasing power parity is zero. Similarly, the Hooper-Morton model assumes that none of the variables in their model is equal to zero. Hence, it extends the Dornbuch-Frankel model to allow for changes in the long-run real exchange rate. By these models, Meese and Rogoff tried to forecast with the structural models using a grid of coefficient constraints drawn from the theoretical and empirical literature on money demand and purchasing power parity.

Secondly, Meese and Rogoff worked with univariate and multivariate time-series models. Many univariate time-series models are estimated for the logarithm of exchange rate. Some pre-filtering techniques like differencing, depersonalizing and removing time trends were applied on the data, then both actual and pre-filtered data were used in the analysis. In that part of their study, Meese and Rogoff went into details of mathematical techniques of analysis like Auto regression (AR), long AR, and Vector Auto Regression (VAR). Their data were chosen to conform theoretical assumptions underlying the specification of the structural models and all raw data were unadjusted. Meese and Rogoff used monthly Dollar/Mark, Dollar/Pound and Dollar/Yen spot exchange rates, average value of trade weight dollar, forward rates of one, three, six and twelve months’ maturities; short- and long-term interest rates started from March 1973. Then they made their analysis by using some statistical performance measures like Mean Error (ME), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). After getting empirical results, they concluded that their study with the Random Walk Models performs no worse than the estimated univariate time-series models, an unconstrained vector auto regression or candidate structural models in forecasting three bilateral rates and the trade weighted dollar. They concluded that regardless of the version of the monetary model used, none outperforms the Random Walk in terms of Mean Square Error.

However, in 2001, Mark and Sul rejected the univariate framework of Meese and Rogoff by using Panel regression model [23]. Mark and Sul argued that fundamentals possess the important forecasting ability on cross-sectional data examination. Discussion on superior exchange rates forecasting continued with Cheung in 2005 [8]. He argued against the ability of the widely used Mean Square Error metric for evaluating the forecasting power of structural models. Cheung suggested the that forecasting ability of each model depends on the time period considered during the evaluation. Moreover, in 2009, Molodtsova and Papell worked on out-of-sample predictability for monetary structural models compared to Random Walk for eleven different currencies for short-time period forecasting [28]. They found that the fundamentals have not established yet their value in exchange rate forecasting since their drawbacks have not been overturned.

The second category is the micro-structural approach. This category mainly focuses on short-term forecasts like daily forecasts. In 2002 and 2005, Evans and Lyons worked on this category and found that institutional aims of exchange rate forecasting like order flow have a great importance in forecasting. Before this study, in 2001, Evans also found that order flow is a more reliable measurement in the exchange rate market [37]. In addition to this, Sager and Taylor showed a forecasting ability of different countries' currencies and forecasting horizons. Then they rejected the superiority of order-flow models according to their findings, while Killeen stated that predictive information content in order-flow models decayed rapidly over time in 2006. He concluded his study by stating that the forecasting ability of exchange rates is time limited since they revert back to a Random Walk Model.

However, a main development was made by Karemera and Kim in 2006 [21]. They introduced Auto Regressive Integrated Moving Average (ARIMA) models which outperform random walk models for many currencies on a monthly forecasting time period; but, the forecasting ability strongly depended on the time period under the evaluation. Moreover in 1993, Cheung already realized that long-run memory in exchange rates and proposed the use of Auto Regressive Fractional Integrated Moving Average (ARFIMA) [7]. Moreover in 2001, Bollerslev and Wright mentioned that the Generalized Auto Regressive Conditional Heteroscedasticity models have a strong forecasting ability, but they are tending to forecast worse than Auto Regressive models with high frequency series [37].

Apart from these studies, many economists and statisticians are still working on exchange rate forecasting models. Obstfeld, for instance, focused on the European Monetary System (EMS) to understand the behavior of exchange rates. According to him, speculations against currency results objective economic conditions which make a liability devaluation. As a result, even fixed rates can be sustained indefinitely in the absence of speculative attacks. On the other hand, West and Cho aimed to compare out-of-sample forecasting ability of univariate Generalized Auto Regressive Conditional Heteroskedasticity (GARCH), Autoregressive and Nonparametric models for conditional variances by using weekly rates of dollar for ten years. Their study proved that GARCH models are superior to the other methods since GARCH has the smallest mean square error.

Clarida, Sarno, Taylor, Valento worked together to find whether allowing for non-linearly in underlying data to generate process for term structure yield rate forecasting is a better alternative, and to do this they used three regime Markov switching vectors for spot rates of currencies. Their outputs of analysis revealed that preferences of non-linearity in term structure appeared to be modelled well by a multivariate three-regime Markov switching Vector Equilibrium Correction Model which allows for shifts in both intercept and covariance structure. Hauner, Lee and Takizawa approached from a different point of view; they tried to compare the extents which several models of exchange rates determinations can take account of market exchange rate by using 50 currencies. They believed that their study is the first comprehensive investigation of this motivation; so they mainly focused on which factors appear to enter the formation of forecasts. Their finding is the view that the correct criterion is not mean square error alone, but the contribution of forecast store aliasing the objective agents. A Turkish scientist, Erataman [15], tried to focus on different aspects of exchange rate modeling; he investigated the demand of a currency by dynamic programing approach and to do that he used the Monte-Carlo simulation method [15]. In addition to all these studies, Plakandras, Papadimitriou and Gogas [37] used a hybrid model which contains machine learning and statistical learning to model exchange rates. Multivariate Adaptive Regression Splines (MARS), Support Vector Regression (SVR) and Ensemble Empirical Mode Decomposition (EEMD) are in their methodology [37]. Their hybrid system modelled exchange rates for 5 currencies, daily and monthly, by selecting variables from input variables which are both macroeconomic and financial variables.

To conclude, since exchange rates have a significant role in finance and economics, it is vital to model them sufficiently correctly and reliably. To reach a model with a superior forecasting ability, mathematicians, economists and statisticians have been working for years. However, since the exchange rates are depending on many independent, dependent and unpredictable variables, a model with a superior forecasting ability has not been constructed yet even though these previous studies are loadstar for future studies promisingly.

CHAPTER 3

METHODOLOGY

In this chapter, we explain the methods which we used in our study to model exchange rates with their brief descriptions and algorithms, thoroughly.

3.1 Linear Regression

Regression methods are basic and quite useful statistical tools in many areas, including engineering, science and economics for prediction. Besides, it is easy to establish a relationship between dependent variables and a model. With these features, we employed linear regression as a simplex tool to build a forecasting model for exchange rates and to compare with MARS method.

The main aim of the linear regression (LR) is to model a relationship between variables by fitting a linear equation. In other words, we can say that linear regression is a statistical tool to investigate and model the relationship between variables. Moreover, it is the most basic and common analysis for prediction based on observations. The variable, which we are trying to predict, is called dependent (response) variable referred to as Y, the variable our prediction is based on is called independent vector variable (predictor or regressor) and is referred to as $\boldsymbol{X}$. There are several linear regression analyses including simple linear regression, multiple linear regression, logistic regression, ordinal regression, multinominal regression and discriminant analysis. If there exists only one independent variable, then the prediction model is simple regression.

Multiple linear regression involves one independent variable and more than two dependent variables; logistic regression involves one binary independent variable and more than two dependent variables; ordinal regression contains one ordinal independent variable and one or more than one nominal independent variables, multinominal regression contains one nominal independent variable and one or more than one dependent variables; finally, discriminant analysis contains one nominal dependent variable and more than one independent variables.

The idea behind linear regression is to fit a single line through a scatter plot. We explain it in the following numerical example as expressed and illustrated by Table 3.1, Figure 3.1 and Figure 3.2 for simple linear regression.

Table 3.1: Example of data.

X	y
1	2,25
2	0,75
3	3,00
4	5,12
5	4,45

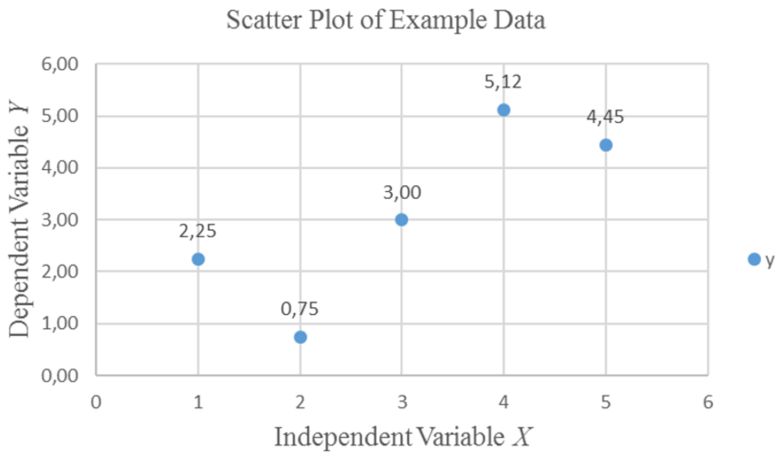

Figure 3.1: Example of scatter plot.

We can represent simple linear regression as

$$Y = mX + m_0 + \varepsilon, \tag{3.1}$$

where y is the estimated dependent random variable, m is the regression coefficient, m_0 is the intercept, ε is the noise term ϵ, N(0, σ^2)and X is the independent random variable. The above model involves only one independent variable, for this reason it is called "Simple Linear Regression Model".

Sometimes, we write this equation with a common mathematical expression shortly as

$$\hat{Y} = \beta_0 + \beta_1 X, \tag{3.2}$$

where β_0 is the intercept and β_1, the slope, is the change in the mean of the distribution of the response produced by a unit change in X, and the X term is an independent variable.

However, not all the observations are exactly on a straight line. In this case, the error term ε should be included in the regression equation:

$$Y = \beta_0 + \beta_1 X + \varepsilon. \tag{3.3}$$

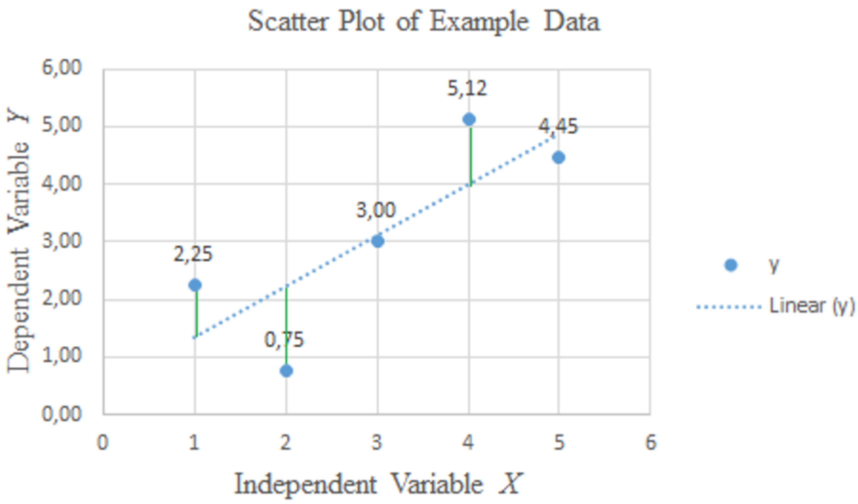

Figure 3.2: Example of scatter plot with regression line. The red line consists of the predictions, the blue points are the actual observations, the vertical lines between observation points and the red line presents the errors of prediction.

The error term ε is a random variable which accounts for the failure of the model to fit the data correctly and it is normally distributed ($\varepsilon \sim N(0, \sigma^2)$).

Since the response variable, Y, is a random variable, there is a probability distribution for Y at each value of X. Moreover, mean and the variance of independent variable can be stated as

$$E(Y \mid X = x) = \beta_0 + \beta_1 x \tag{3.4}$$

and

$$Var(Y \mid X = x) = Var(\beta_0 + \beta_1 x + \varepsilon) = \sigma^2. \tag{3.5}$$

In simple linear regression equation, β_0 and β_1 are unknown parameters and they need to be estimated. For this purpose, least-squares estimation seeks to minimize the sum of squares, $S(\beta_0, \beta_1)$, of the difference between the observed response variable y_i and the straight line.

Least-Squares (LS) criterion is to minimize $S(\beta_0, \beta_1) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$. The function S must be minimized with respect to the coefficients. Here, (x_i, y_i) $(i = 1, 2, ..., n)$ are the given data.

By least-squares approach, the unknown parameters β_0, and β_1, which minimize the objective function $S(\beta_0, \beta_1)$, can be estimated as given below.

$$\begin{aligned}
b_1 &= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\sum_{i=1}^{n}(x_i(x_i y_i - x_i\bar{y} - \bar{x}y_i + \bar{x}\bar{y})}{\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2)} \\
&= \frac{\sum_{i=1}^{n}(x_i y_i) - \bar{y}\sum_{i=1}^{n} x_i - \bar{x}\sum_{i=1}^{n} y_i + n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2} \\
&= \frac{\frac{1}{n}\sum_{i=1}^{n}\bar{xy} - \bar{x}\bar{y}}{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2} \\
&= \frac{\bar{xy} - \bar{x}\bar{y}}{\bar{x^2} - \bar{x}^2} = \frac{Cov(x,y)}{Var(x)} \\
&= r_{xy}\frac{S_y}{S_x}
\end{aligned} \tag{3.6}$$

and

$$b_0 = \bar{y} - b_1\bar{x} = \frac{\sum_{i=1}^{n} y_i}{n} - b_1\frac{\sum_{i=1}^{n} x_i}{n}, \tag{3.7}$$

where r_{xy} stands for sample correlation coefficient between *X* and *Y*; s_x and s_y are the notations of sample standard deviation of *X* and *Y*, respectively. The average value of the quantities is notated by the horizontal bar over the quantity.

The main difference between simple linear regression and multiple linear regression is the number of independent variables. Multiple linear regression contains more than one regressor and, in general, a multiple linear regression model with *p* regressors can be represented as

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon. \tag{3.8}$$

Similar to the simple linear regression, the error term ε is a random variable which accounts for the failure of the model to fit the data correctly and it is normally distributed ($\varepsilon \sim N(0, \sigma^2)$).

Multiple regression model can be formulized as

$$\begin{aligned}
Y_i &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + \varepsilon_i \\
&= \beta_0 + \sum_{j=1}^{p}\beta_j X_{ij} + \varepsilon_i \qquad (i{=}1, 2, \ldots, n).
\end{aligned} \tag{3.9}$$

In simple linear regression, a single response variable *Y* is represented by a single predictor variable X_i for each observation; however, in multiple linear regression, more than one predictor variable is available. For this reason, a multiple linear mean function can be written as

$$E(Y \mid \boldsymbol{X} = \boldsymbol{x}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p, \tag{3.10}$$

where β_0 is the intercept and the further parameters β_j are univariate slopes.

Similar to simple linear regression, all the coefficients β_j are unknown parameters and they are needed to be estimated. For this purpose, least-squares estimation seeks to minimize the sum of squares of the differences between the observed response data y_i and regression line.

LS procedure seeks to to minimize $S(\beta_0, \beta_1, \ldots, \beta_p) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij})^2$. The function *S* must be minimized with respect to the coefficients. Here, $(\boldsymbol{x}_i, y_i)$ $(i = 1, 2, \ldots, n)$ again stand for the given data.

However, for simplicity in representations and calculations, we can unfold the responses for all observations into an *n*-dimensional vector which is called the response data vector, written as

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \tag{3.11}$$

we can unfold all predictor data into $n \times (p+1)$-vector which is called the design matrix written as

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1p} \\ 1 & x_{21} & x_{22} & \ldots & x_{2p} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{np} \end{bmatrix}. \tag{3.12}$$

In this design matrix, the initial column of entries 1 stands for the constant factor of the intercept coefficient β_0.

We can unfold intercept and slopes into a $(p+1)$-dimensional vector which is called slope vector and written as

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}. \tag{3.13}$$

Lastly, we can unfold all error terms into *n*-dimensional vector which is called as error or residual vector and written as

$$\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}. \tag{3.14}$$

Finally, by using linear algebra notation, we can compactly summarize the linear models $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + \varepsilon_i$ as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{3.15}$$

where $\boldsymbol{X\beta}$ is the matrix-vector product.

The unknown parameters β_j need to be estimated, the least-squares approach is used to estimate them. For this purpose, least-squares estimation seeks to minimize the sum of squares of the difference between the observed response data y_i and the regression line, where $S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon^T \varepsilon = (\boldsymbol{y} - \boldsymbol{X\beta})^T (\boldsymbol{y} - \boldsymbol{X\beta})$.

After solving the least-square equation where we assume that the matrix $\boldsymbol{X}^T\boldsymbol{X}$ is invertible, the solution is

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}. \tag{3.16}$$

In addition to these explanations, we should recall the basic assumptions of linear regression which are

- Linearity and Additivity,
- Independence of errors,
- Normality of errors,
- Homogeneous dispersion of errors.

Until now, we have discussed the definition and main properties of simple linear regression and multiple linear regression which we employed in our study for comparison of model performances. Apart from SLR and MLR, also there exist other linear regression analysis which are Logistic Regression, Ordinal Regression, Multinominal Regression and Discriminant Analysis.

3.2 Support Vector Regression

Another technique we employed in our study to model exchange rates and to compare forecasting performances is Support Vector Regression. Before explaining Support Vector Regression (SVR), it is helpful to shortly explain Support Vector Machines (SVM) initially. SVM techniques are used to generate nonlinear boundaries by creating a linear boundary in a large feature space. In the next subsection, we explain SVM and SVR techniques briefly.

3.2.1 Support Vector Machines

Support Vector Machines approach, introduced by Boser, Guyon and Vapnik in 1992, suggested to create nonlinear classifiers to maximum-margin hyperplanes with the help of the kernel trick [6]. SVM technique, developed from Statistical Learning Theory of Vapnik and Chervonenkis, invented by 1963 [44]; however, the technique gained popularity several years later. Kernel functions' large margin hyperplanes, geometrical explanation of kernel functions for inner products in feature space, slack variable usage to deal with noise problem and similar optimization techniques have been widely employed to recognize patterns since the 1960s. However, all these features were not applied together to find maximal margin classifier called SVM until 1995, the date that the soft margin version was invented. SVMs provide useful machine learning algorithms for classifications and regression analysis. An SVM model is a representation of data points in space which are the examples of separate categories divided by an as wide as possible clear gap. The algorithm maps new examples into the space and predicts which side of the gap they belong to.

Basically, a hyperplane or a set of hyperplanes in the high-dimensional space or in an infinite-dimensional space are built by the SVM to use it for regression or classification aspect. The algorithm succeeds to separate data points almost perfectly with the help of hyperplanes. However, it is not often possible to separate points linearly. Because of this, the algorithm maps the finite-dimensional space into a relatively higher-dimensional space which makes the separation task easier. Dot products are defined by kernel functions $k(x, \boldsymbol{y})$ which is suitable for the problem. An SVM defines the hyperplanes as the set of points with their dot products with an element being a constant in a higher-dimensional space. This vector which defines the hyperplane can be a linear combination with parameters a_i of the image of feature vectors x_i occurred in the dataset. As a result, after choosing the hyperplane, data points in a feature space are mapped into the hyperplane as stated by

$$\sum_{i=1}^{n} \alpha_i k(\boldsymbol{x}_i, \boldsymbol{x}) = \text{constant.} \tag{3.17}$$

We should be aware that each term in the equation given above measures the closeness of the test point $\boldsymbol{x}$ to the corresponding data set point $\boldsymbol{x}_i$, if $k(x, \boldsymbol{y})$ is getting smaller as $\boldsymbol{y}$ is growing further from $\boldsymbol{x}$. Hence, the algorithm can measure the closeness of each test point to the data points. Classifying data is the core goal of machine learning. In order to separate data points into classes, an SVM regards a data points as a p-dimensional vector. Then, we seek to separate data points with a (p-1)-dimensional hyperplane.This application is called linear classifier. However, there may be many hyperplanes which can classify the data, so that it is important to find the most reasonable or "optimal" hyperplane. Choosing the hyperplane which represents the larger separation or the margin between data points is one reasonable choice for the best hyperplane. The hyperplane with the distance from itself to the nearest data point being maximal is the best choice and it is called *maximum-margin hyperplane*. Moreover, the linear classifier defined by this hyperplane is called *maximum margin classifier or perception of optimal stability*. Figure 3.3 shows three different hyperplanes and according to the explanation given above; hyperplane H_3 is the best choice for given data points.

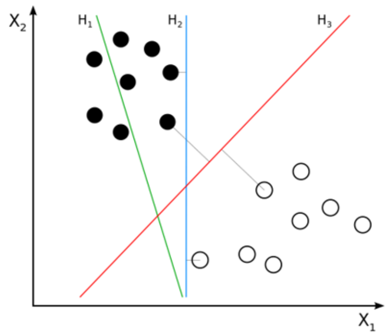

Figure 3.3: H_1 does not separate data points, H_2 separates data points with small margin, H_3 separates data points with maximum margin [26].

An SVM makes the optimal hyperplane choice by employing an iterative training algorithm which minimizes the error function. By the form of this error function, SVM models are categorized into four groups distinctly:

- *Classification SVM Type 1*, which is also known as *C-SVM classification*,
- *Classification SVM Type 2*, which is also known as *v-SVM classification*,
- *Regression SVM Type 1*, which is also known as *ε-SVM regression*,
- *Regression SVM Type 2*, which is also known as *v-SVM regression.*

For *C-SVM* type of SVM, the training algorithm leads to a minimization of the error function:

$$\underset{\boldsymbol{\omega},\boldsymbol{\xi}}{\text{minimize}} \quad \tfrac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C\sum_{i=1}^{n}\xi_i$$

subject to the constraints:

$$y_i(\boldsymbol{\omega}^T\Phi(\boldsymbol{x}_i)+b) \geq 1-\xi_i \qquad \text{and} \qquad \xi_i \geq 0 \qquad (i=1,2,\ldots,n). \tag{3.18}$$

Here C is the capacity constant, $\boldsymbol{\omega}$ the coefficient vector, b a constant, stands for the parameters of nonseparable data points, $\boldsymbol{\xi} = (\xi_1,\xi_2,\ldots,\xi_n)^T$ and n is the number of training cases. It is important that $y \in \{\pm 1\}$ shows the class labels and x_i shows the independent variables. The kernel Φ is employed to transform the input data to the feature space. Another important property about the error function of *C-SVM* type is that the larger the C value becomes, the larger the penalized error grows. Hence, C should be chosen carefully in order to avoid over-fitting.

In contrast to *C-SVM* type, for *v-SVM* type of SVM, the training algorithm aims at the minimization of the error function:

$$\underset{\boldsymbol{\omega}, \boldsymbol{\xi}, \rho}{\text{minimize}} \quad \tfrac{1}{2}\boldsymbol{\omega}^T\boldsymbol{w} - v\rho + \tfrac{1}{n}\sum_{i=1}^{n}\xi_i,$$

subject to the constraints:

$$y_i(\boldsymbol{\omega}^T\Phi(\boldsymbol{x}_i) + b) \geq \rho - \xi_i, \qquad \xi_i \geq 0 \qquad (i = 1, 2, \ldots, n) \qquad \text{and} \quad \rho \geq 0. \quad (3.19)$$

On the other hand, in Regression SVM, the functional dependence of response variable Y on the vector of regressor variables $\boldsymbol{X}$ should have to been estimated. As being the case with all regression concepts, it is considered as the relationship between response variable and the regressor variables is defined by a deterministic function f and some additive noise ε:

$$Y = f(\boldsymbol{X}) + \varepsilon. \quad (3.20)$$

Then, the SVR algorithm comes into use in order to find the functional form of f that is able to predict new cases which have not been used in the SVM before. For assessing f, the algorithm employs training a SVM model based on the sample data set. According to its definition, $\varepsilon - SVM$ regression or $v - SVM$ regression models are used to find an error function.

For $\varepsilon - SVM$ regression type of SVM the error function is:

$$\tfrac{1}{2}\boldsymbol{\omega}^T\boldsymbol{\omega} + C\sum_{i=1}^{n}\xi_i + C\sum_{i=1}^{n}\xi_i^*,$$

which we minimize subject to:

$$\begin{aligned} &\boldsymbol{\omega}^T\Phi(\boldsymbol{x}_i) + b - y_i \leq \varepsilon + \xi_i^*, \\ &y_i - \boldsymbol{\omega}^T\Phi(\boldsymbol{x}_i) + b_i \leq \varepsilon + \xi_i, \\ &\xi_i^*, \xi_i \geq 0 \quad (i=1, 2, \ldots, n). \end{aligned} \quad (3.21)$$

For the $v - SVM$ regression model, the error function is given by:

$$\tfrac{1}{2}\boldsymbol{\omega}^T\boldsymbol{\omega} - C[v\varepsilon + \tfrac{1}{n}\sum_{i=1}^{n}(\xi_i + \xi_i^*)],$$

which we minimize subject to:

$$\begin{aligned} &(\boldsymbol{\omega}^T\Phi(\boldsymbol{x}_i) + b) - y_i \leq \varepsilon + \xi_i, \\ &y_i - (\boldsymbol{\omega}^T\Phi(\boldsymbol{x}_i) + b) - y_i \leq \varepsilon + \xi_i^*, \\ &\xi_i^*, \xi_i \geq 0 \quad (i=1, 2, \ldots, n), \quad \varepsilon \geq 0. \end{aligned} \quad (3.22)$$

for some $\varepsilon \geq 0$, there are various types of kernel functions which can be employed in SVM models, including linear, polynomial, radial basis function (RBF) and sigmoid given below:

$$\begin{aligned} &k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i^T\boldsymbol{x}_j : \quad \text{Linear}, \\ &k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \mathrm{e}^{(-v\|x_i - x_j\|_2)} : \quad \text{RBF}, \\ &k(\boldsymbol{x}_i, \boldsymbol{x}_j) = (\gamma\boldsymbol{x}_i^T\boldsymbol{x}_j + C)^d : \quad \text{Polynomial}, \\ &k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \tanh(\gamma\boldsymbol{x}_i^T\boldsymbol{x}_j + C) : \quad \text{Sigmoidal}. \end{aligned} \quad (3.23)$$

where the terms *C*, *d* and γ stand for the adjustable kernel parameters in formulas. Among these types of kernel functions, RBF is the most popular choice of kernel function type employed in SVM due to its localization and the finite responses across the whole range of the real *x*-axis.

3.2.2 Support Vector Regression

SVM methodology can be applied both to classification and regression problems. Similar to the classification approach, the SVR algorithm runs to find and optimize generalization bounds for regression. SVR algorithm resists a loss function which ignores noise terms that stands for a certain distance from the real data points. This kind of loss functions are called *epsilon intensive* loss functions. Examples given in Figure 3.4, Figure 3.5 and Figure 3.6 show the one-dimensional linear regression case and the nonlinear regression case for an epsilon intensive loss function, respectively.

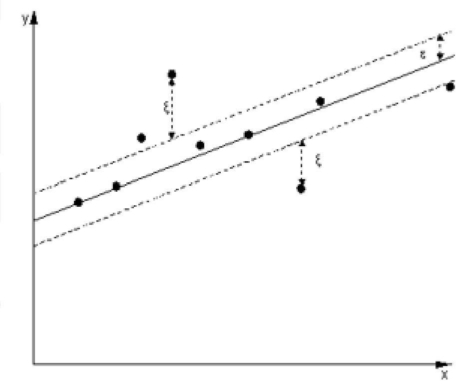

Figure 3.4: One-dimensional linear regression case for epsilon intensive loss function [44].

By using epsilon intensive loss function, it is certain to have both a global minimum and an optimum generalized bound at the same time as illustrated in Figure 6.

SVM Regression algorithm is started by mapping the input variable $\boldsymbol{X}$ onto an *m*-dimensional feature space with the help of some fixed nonlinear mapping. After this, the linear model, $f(\boldsymbol{x}, \boldsymbol{\omega})$, is built on this feature space which is stated by

$$f(\boldsymbol{x}, \boldsymbol{\omega}) = \sum_{j=1}^{m} \omega_j h_j(\boldsymbol{x}) + b, \tag{3.24}$$

where $h_j(\boldsymbol{x})$ (j = 1, 2, ..., m) stand for a nonlinear transformations set, and the term *b* is the "bias" term. In this equation, it is assumed that data have zero mean; as a result, the bias term is ignored.

The main aim of regression is to estimate a model which gives a rather true information about the data. For this purpose, the SVM Regression procedure employs some

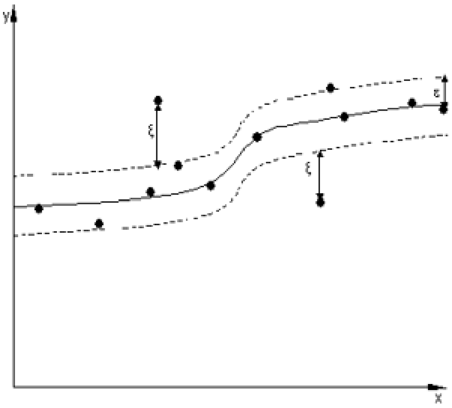

Figure 3.5: Nonlinear regression case for epsilon intensive loss function [44].

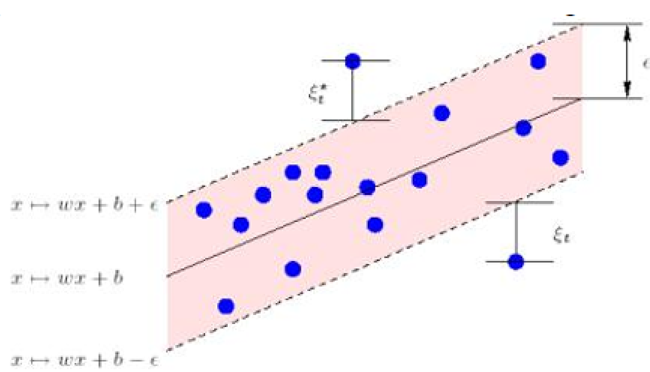

Figure 3.6: Epsilon band with slack variables with selected data points [27].

loss function $L(y, f(\boldsymbol{x}, \boldsymbol{\omega}))$ to measure the quality of the estimation. The type of the loss function used for SVM Regression algorithm is suggested by Vapnik and called an ε–intensive loss function:

$$L_\varepsilon(y, f(\boldsymbol{x}, \boldsymbol{\omega})) = \begin{cases} 0 & \text{, if } |y - f(\boldsymbol{x}, \boldsymbol{\omega})| \leq \varepsilon, \\ |y - f(\boldsymbol{x}, \boldsymbol{\omega})| - \varepsilon & \text{, otherwise.} \end{cases} \tag{3.25}$$

The empirical risk for the regression is given by

$$R_{emp}(\boldsymbol{\omega}) = \frac{1}{n} \sum_{i=1}^{n} L_\varepsilon(y_i, f(x_i, \boldsymbol{\omega})). \tag{3.26}$$

With the help of ε–intensive loss function, the SVM Regression algorithm builds a linear regression model for a high-dimensional feature space, and also lowers the complexity of the model by minimizing the Euclidean norm $||\boldsymbol{\omega}||_2$. This procedure can be explained by introducing the nonnegative slack variables ξ_i and ξ_i^* (i = 1,2,..., n) which are used to measure the deviation of the following training samples outside the ε–zone. As a result, SVM Regression can be described as the following minimization problem:

$$\underset{\boldsymbol{\omega}, \boldsymbol{\xi}, \boldsymbol{\xi}^*}{\text{minimize}} \quad \tfrac{1}{2}||\boldsymbol{\omega}||_2 + C\textstyle\sum_{i=1}^{n}(\xi_i + \xi_i^*),$$

$$\text{subject to} = \begin{cases} y_i - \boldsymbol{\omega}^T\boldsymbol{x}_i - b \leq \varepsilon + \xi_i^*, \\ \boldsymbol{\omega}^T\boldsymbol{x}_i + b - y_i \leq \varepsilon + \xi_i, \\ \xi_i, \xi_i^* \geq 0 \quad (i = 1, 2, \ldots, n), \end{cases} \tag{3.27}$$

where ξ_i and ξ_i^* are slack variables which are controlled by the penalty parameter *C* and ε stands for the tolerance cord around the regression line, as described in Figure 6.

This optimization problem can be converted to a dual problem and its solution is given below:

$$\begin{aligned} &\text{maximize} && f(\boldsymbol{x}) = \sum_{i=1}^{n_{sv}}(\alpha_i - \alpha_i^*)K(\boldsymbol{x}_i - \boldsymbol{x}), \\ &\text{subject to} && 0 \leq \alpha_i^* \leq C, \quad 0 \leq \alpha_i \leq C \quad (i = 1, 2, \ldots, n_{sv}), \end{aligned} \tag{3.28}$$

where n_{sv} denotes the number of Support Vectors (SV) and the Kernel Function $K(\boldsymbol{x_i}, \boldsymbol{x})$ can be defined as given in the subsequent form:

$$K(\boldsymbol{x}_i, \boldsymbol{x}) = \sum_{j=1}^{m} h_j(\boldsymbol{x})h_j(\boldsymbol{x}_i). \tag{3.29}$$

It is obvious that the performance of an SVM regression equation depends on the choice of the meta-parameters *C*, ε and the parameters in the kernel function. Hence, the selection of optimal parameters is more difficult and complicated. Software implementations to be used for SVM Regressions often view meta-parameters as user-defined input parameters. The parameter *C* describes the trade-off between the flatness of the model which is the complexity, and it measures the deviations bigger than the ε term which are extinguished in optimization. For instance, if the *C* parameter converges to infinity, then the objective only minimizes the empirical risk, regardless of model complexity in the optimization. Beside these parametric properties, the parameter ε stands in the formula to control the width of the insensitivity zone of ε which is used to fit the training data. Moreover, the number of support vectors in the regression equation depends on the value of ε. For example, a greater ε causes less support vectors in the equation. Also, greater ε values lead to more flat estimates. As a result, the complexity of the model is affected by both the parameters *C* and ε differently.

The advantage of using the Support Vector Machine and Support Vector Regression is that it prevents from the difficulties in the usage of linear functions for a high-dimensional feature space and dual convex quadratic transformed optimization problems [3]. In addition to this, the advantage of using Support Vector Regression is the penalization of the error terms which are greater than the threshold ε by the help of a loss function. These loss functions mean advantages to the rare description of some decision rule indeed, by permitting substantial algorithmic and representational benefits.

3.3 MARS Method

The third method we applied to model exchange rates and to compare forecasting performances in our study is Multivariate Adaptive Regression Splines. The proper selection of input variables is the main difficulty in modeling. We use *Multivariate Adaptive Regression Splines (MARS)* for automatic variable selection based on the statistical loss metrics. Jerome H. Friedman who is a renowned statistician and physicist, introduced Multivariate Adaptive Regression Splines (MARS) in 1991 [18] as a form of regression analysis method which is both linear and nonlinear nonparametric function estimation and shows a great performance for fitting general nonlinear multivariate functions. The original MARS algorithm was implemented in and produced by *Salford System Software called Salford Prediction Modeler (SPM)* for the user's benefit [39].

As distinct from other model-driven or supervised statistical learning methods, MARS is a regression model, fundamentally. In addition to this, the MARS method is powerful to capture significant nonlinearities and interactions. It was designed to predict continuous numeric outcomes; however, it is useful to perform variable selection, variable transformation, interaction detection and self-testing for higher dimensions at high speed, and all of this "automatically".

An easy, in fact, uni-variate representation of a MARS model compared to a parametric linear regression model is depicted in Figure 3.7.

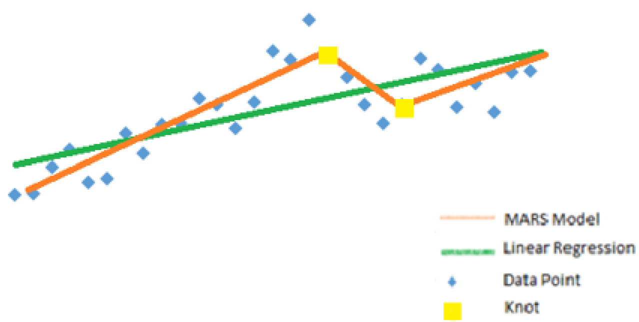

Figure 3.7: Visualization of the MARS model.

Basically, MARS is a non-parametric regression technique and it may be treated as an extension of linear models that can well approximate nonlinearities and interactions between variables automatically through an adaptive algorithm. Its main ability is to capture essential nonlinearities and interactions, and to predict continuous numeric outcomes. MARS selects best variables in terms of high forecasting accuracy ability according to empirical evidence of relevant literature.

MARS algorithm uses Basis Functions (BFs), which are also called splines, to have flexible regression models and these BFs are included into the algorithm as predictors based on the initial data set. MARS model contains all possible knots and all possible predictors with every possible interaction among themselves, and all these interactions are shown by combination of BFs. After filling the MARS model with the optimum quantities of BFs and knots, the algorithm uses the least-squares estimator method to shape the final model which reveals the best approximating model for the original data set with the remaining BFs. This algorithm is explained below applied by the usage of *Forward Stage* and *Backward Stage*.

In the forward stage, the model is generally an over-fit model which contains a large set of BFs. The algorithm adds basis functions to the model fast and continuously until reaching the maximum number of basis functions in the model. Hence, the obtained model includes all possible basis functions regardless of the contribution of these basis functions to the performance of the model. Forward stage causes the model to be an over-fit model; as a result, it needs to be cleaned from redundant BFs, and this necessitates the called backward stage.

The need for the backward stage is to reduce the complexity of the model by deleting the redundant BFs from the over-fit model. Backward stage algorithm keeps the BFs which promise the smallest increase in Residual Sum of Squares (RSS) and provides the model with the best approximation of the initial data set. These BFs taken-off procedure continues until achieving an optimal balance between variance and bias.

MARS sets up a model in two phases: the forward stage and the backward stage. The forward stage starts with an intercept term which is the mean value of response variables, then the basis functions in pairs are added to the model by MARS recurrently to get a better performance. While applying forward and backward stage algorithm, MARS splits the whole space of input data into several sub-regions and states different mathematical equations for each sub-region. In other words, by these mathematical equations, MARS creates links between sub-regions of input variables and outputs. To do this, MARS benefits from the two-sided truncated power functions given below [43]:

$$(x-t)_+ = \begin{cases} x-t, & \text{if} \quad x>t, \\ 0, & \text{otherwise}, \end{cases} \tag{3.30}$$

$$(x-t)_- = (t-x)_+ = \begin{cases} t-x, & \text{if} \quad t>x, \\ 0, & \text{otherwise}, \end{cases} \tag{3.31}$$

where x, $t \in \mathbb{R}$. These two functions given above are also called as *reflected pair* and the symbols "+" and "-" explain that positive and negative parts are employed, respectively.

Figure 3.8 shows a basis function pair as an example. This truncated power functions or, in other words, hinge functions, are the key elements of MARS models. They are used to partition the data into disjoint regions in order to create piecewise linearity. These both function are a reflected pair together.

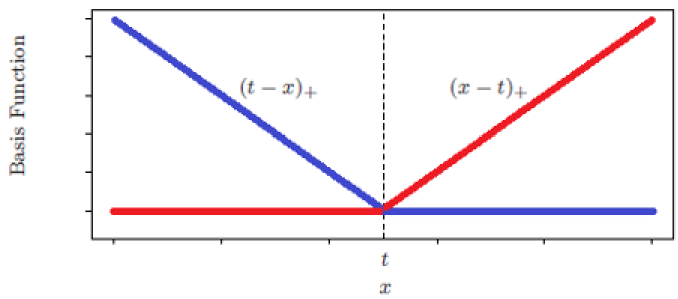

Figure 3.8: Sample Truncated Function $(x-t)_+$ (solid red) and $(x-t)_-$ (solid blue) used by MARS algorithm [43].

The general model given below shows the relation between dependent or response variable and the corresponding predictor variables:

$$Y = f(\boldsymbol{X}) + \varepsilon, \tag{3.32}$$

where the response variable showed by Y, $\boldsymbol{X} = (x_1, x_2, \ldots, x_p)^T$ is a vector standing for all the predictors and ε is an additive stochastic component with mean of zero and finite variance. Then the main aim of MARS algorithm is to build reflected pairs of each input x_j $(j = 1, 2, \ldots, p)$ with p-dimensional knots $\boldsymbol{t}_i = (t_{i,1}, t_{i,2}, \ldots, t_{i,p})^T$ at input data vectors $\boldsymbol{X}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^T$ (i=1,2,...,n) with the input data values x_{ij}. Hence, the collection of Basis Functions (BFs) can be defined as

$$S : \{(x_j - t)_+, (x_j - t)_- \quad | \quad t \in \{x_{1j}, x_{2j}, \ldots, x_{nj}\}, \quad j = 1, 2, \ldots, p\}, \tag{3.33}$$

where n is the number of observations (data). If all of the inputs are distinct, then there exists 2np many BFs.

The model building procedure of MARS algorithm is similar to forward stepwise linear regression. However, instead of merely using the 1-dimensional linear coordinate variables x_j, MARS allows us to employ functions in set S and their products. The MARS model at the end of the forward stage contains hinge functions, yielding the model form

$$Y = a_0 + \sum_{m=1}^{M} a_m B_m(\boldsymbol{X}) + \varepsilon, \tag{3.34}$$

with a vector of random variables $\boldsymbol{X}$ and a vector and noise term ε. In fact, Y is the output variable, $\boldsymbol{X}$ is the input variable of vector form $\boldsymbol{X} = (x_1, x_2, \ldots, x_p)^T$ which contributes to the function B_m, ε is the additive stochastic noise term which is supposed to have normal distribution with finite variance and zero mean, a_0 is the constant term also called as intercept, M is the number of basis functions, a_m is the m^{th} coefficient of spline function, B_m is the m^{th} basis function or product of two or more basis functions from the set S, which is taken from a set of M linearly independent BFs [18]. Existing BFs are multiplied by another reflected pair which joins further variable to create a new BF that reveals the interaction between different variables. Moreover, these existing

BFs stand in the model together with the newly created BFs, so that the so-called spline fitting becomes a higher-dimension spline fitting. For the observations, which are represented by (x_i, y_i) (i=1,2,...,n), the m^{th} multivariate BF is defined as:

$$B_m(\boldsymbol{x}) := \prod_{k=1}^{Km} (S_{km} \cdot (x_{v(k,m)} - t_{km}))_+, \tag{3.35}$$

where K is the number of truncated linear function multiplicated in the m^{th} BF, $x_{v(k,m)}$ is the input variable related to the k^{th} truncated linear function in the m^{th} BF, t_{km} is the knot value of variable $x_{v(km)}$, and $S_{k,m} = \pm 1$.

MARS algorithm is initiated with the constant function $B_0(x_0) = 1$ for a_0 estimation at the forward step. According to the definition, in the set S, all functions of $B_m(\boldsymbol{x})$ are candidate functions and, as a result, basis functions can be of one of the following forms:

1. Can be a constant, 1, which means that there only exists an intercept term,
2. x_k,
3. $(x_k - t_j)_+$,
4. $x_k x_l$,
5. $(x_k - t_i)_+ x_l$,
6. $(x_k - t_i)_+ (x_l - t_j)_+$.

Self-interaction between input variables are not allowed by MARS algorithm; consequently, any of the input variable cannot be the same for each BF. In other words, MARS never uses the same input variables for each BFs; hence, different input variables x_n and x_l, and their corresponding knots t_i and t_j, are used by the BFs. At each step, MARS model finds the knots and the corresponding pair of basis functions which gives a maximum reduction in sum of squares of residual errors. These two basis functions in the pair are superposable except that a different side of a mirror image of hinge function is used for each function. The forward stage regards all product of BFs $B_M(\boldsymbol{x})$ as new BFs and adds them into the model set. This term produces the largest decrease in training terms and in the form of

$$a_{M+1} B_n(\boldsymbol{x})(x_j - t)_+ + a_{M+2} B_n(\boldsymbol{x})(t - x_j)_+, \tag{3.36}$$

where a_{M+1} and a_{M+2} are coefficients which are determined by least-square estimation. Correspondingly, the BFs given below are potential candidates for the MARS algorithm:

1. 1,
2. x_k,

3. $(x_k - t_j)_+$, if x_k exists in the model,

4. $x_k x_l$, if x_k and x_l exist in the model,

5. $(x_k - t_i)_+ x_l$, if $x_k x_l$ and $(x_k - t_i)_+$ exist in the model,

6. $(x_k - t_i)_+(x_l - t_j)_+$, if $(x_k - t_i)_+ x_l$ and $(x_l - t_j)_+ x_k$ exist in the model.

Then, the forward stage adds the winning products to the model. MARS searches all combinations of existing terms, all variables and all values of each variable at the forward stage. This adding-of-basis functions procedure continues until the change of residual error becomes too small or the maximum number of terms added has reached an over-fitted and complicated model. As a consequence, the model contains incorrect terms and the model becomes highly complex. At the end of the forward stage, we can see the maximal model which over-fits the given data.

At this point, the backward stage takes in charge. Since the current MARS model is over-fit, the backward stage prunes the model in order to build a model with a more accurate forecasting ability. To do that, it removes redundant basis functions which contribute the smallest increase in residual squared error from the model at each step one by one. This removing continues until getting the best sub-model, $\hat{f}_a$, of each size (number of terms) α. The backward stage can select any term to delete while the forward stage can only deal with pairs at each step, so that the backward stage has an advantage over the forward stage. In other words, the backward stage eliminates one side of pairs generally and terms in pairs appear rarely in the model. However, the forward stage adds terms in pairs to the model. MARS uses the lack-of-fit (LOF) criterion defined by Generalized Cross Validation (GCV) at the forward stage in order to find the optimal number of terms, α, in the best-fit model as suggested by Friedman [18] and GCV reveals the lack of fit in the model. The GCV expression is written as follows [12]:

$$\text{LOF}(\hat{f}_a) = \text{GCV} := \sum_{i=1}^{n} (y_i - \hat{f}_a(x_i))^2 / (1 - C(\alpha)/n)^2, \tag{3.37}$$

where n is the number of data, $\hat{f}_a$ is the estimated best model and $C(\alpha)$ is the effective number of parameters, which stands for the number of terms in the model and the number of parameters used in the selection of the optimal positions of the knots. It is also a kind of complexity penalty that increases with the number of basis functions (BFs) in the model, which is defined as:

$$C(\alpha) = v + dK, \tag{3.38}$$

where K is the number of knots selected at the forward stage, v is the number of linearly independent functions in the model, d is a kind of penalty for each BF in the model which is generally set as d=3 (d=2 is used for additive models). The details about d are given by [18]. The numerator of the GCV term is the residual sum of squares; the denominator penalizes any increasing variance in case of increasing model complexity. With the help of the lack-of-fit criterion, the best-fit model is selected which minimizes

GCV at the backward stage.

We can evaluate the performance of a developed MARS model in terms of Coefficient of Determination (R^2). For an accurate model, the value of R^2 should be close to one. Figure 3.9 shows a the visualization of the MARS model.

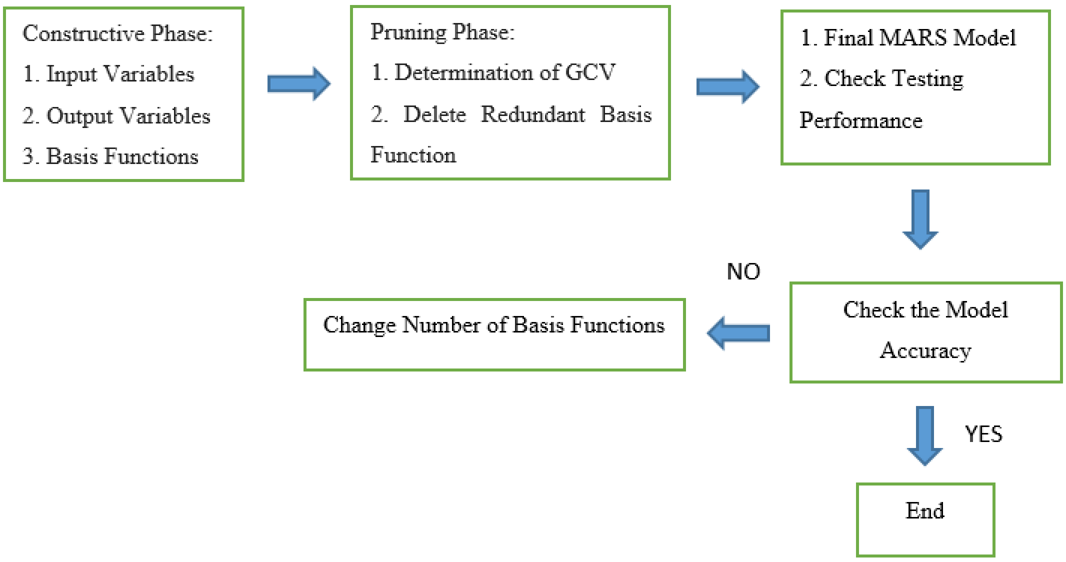

Figure 3.9: Visualization of the MARS model [41].

The main advantage of the MARS method is that it does not require any assumptions and it is more flexible than linear regression models. MARS can deal with both continuous and discrete data. Moreover, it is simple to understand and interpret. However, the MARS method is sensitive to outliers and requires a significant amount of data and a large number of variables. In addition to these items, there can be problems in choosing predictor variables when a multi-collinearity problem exists. MARS eliminates such a multi-collinearity problem at its backward stage. Friedman offers either fitting a series of increasing-interaction order between models and to compare GCV to select the model with lowest interaction order and an acceptable fit. The other solution proposed by Friedman is involving a penalty on inducing new variables into the model so that the change by entering two inputs which are highly collinear would decrease [48].

Another main advantage of MARS algorithm comes from the use of certain piecewise linear one-dimensional functions. This piecewise linearity allows MARS algorithm to operate locally. The particular piecewise linear basis functions are zero over a part of their range and this results in being nonzero over some small part of space only where both components are nonzero when they become multiplied. In this way, one can use nonzero components only where they are needed and build up the regression surface with sufficient components. Moreover, piecewise linear basis functions do not allow for self-interaction between input variables. At each step, the MARS model finds the knots and the corresponding pair of basis functions which give a maximum reduction in terms of sum-of-squares residual error. These two basis functions in a pair are identical except that the different side of a mirror image of hinge function is used for each function. This restriction permits prevention from the formation of high-order powers of any input variable.

3.4 Cubic Spline Interpolation

Numerical or non-numerical observations collected for statistical analysis which are broken down into components or units of data are called *"disaggregated data"*. Cubic Spline Interpolation is one of the most common methods for disaggregation problems [24]. The major challenge that we faced with our data was a disaggregation problem for GDP rate data. In order to abolish disaggregation, we employed Cubic Spline Interpolation before LR, MARS and SVR applications.

This method cannot create new observations; it can only give estimated data points from original data. As alternative methods to cope with the disaggregation problem, Denton (1970) approach, Chow-Lin (1971) framework and using quarterly GDP observations for that quarters' months can be applied. Abdul Rashid and Zanaib Jehan (2013) compared Denton (1970) approach, Chow-Lin (1971) framework and Cubic Spline Interpolation [10, 13, 38]. According to their study, these three methods give similar and highly correlated results. Therefore, we used Cubic Spline Interpolation as a remedy for the disaggregation problem.

Cubic Spline Interpolation (CSI) provides an almost perfect piecewise curve that passes through the observations which are underlying the sample time period. The main aim of this method is to receive a continuous interpolation formula in both first- and second-order derivatives and both within the intervals and at interpolating nodes. CSI does not need any high-frequency indicator variable observation related to low-frequency series. As a result, it outfaces low-frequency-variable related observed indicators at high-frequency choice. Moreover, it correlates each observed data point efficiently and effectively. Beside these advantages, implementing CSI is relatively easy and time saving.

Starting point of CSI is an engineering tool to draw a smooth curve which passes though different observation points. The spline interpolation process uses estimated coefficients of the cubic polynomial as weights of each interval. We can define the piecewise function, say, $S(y)$, to express n equally spaced internals as

$$\begin{aligned} S(y) = S_1(y), &\quad \text{if} \quad y_1 \leq y \leq y_2, \\ S_2(y), &\quad \text{if} \quad y_2 \leq y \leq y_3, \\ &: \\ &: \\ S_{n-1}(y), &\quad \text{if} \quad y_{n-1} \leq y \leq y_n. \end{aligned} \tag{3.39}$$

In order to state the splines which notated by *S*(*y*), we require *4n* parameters to estimate totally, as there are n evenly spaced intervals and 4 coefficients for each interval. These four coefficients cut the curve; therefore, it should pass each of the observations. As a result, the curve does not have any breaks in continuity.

The third-degree polynomial function $S_i(y)$ is defined as

$$S_i(y) = \beta_{3i}(y-y_i)^3+\beta_{2i}(y-y_i)^2+\beta_{1i}(y-y_i)^1+\beta_{0i}, \quad \text{for} \quad y \in (y_i, y_i+1). \tag{3.40}$$

In addition, there are two main conditions to receive cubic polynomial interpolation which should be satisfied by the values of low-frequency series at both ends of sub-interval. We can define these conditions as

$$\begin{aligned} S_i(y_i) &= x_i, \\ S_i(y_{i+1}) &= x_{i+1}, \end{aligned} \tag{3.41}$$

where x_i can be found by Equation (3.40).

These conditions provide a piecewise continuous function where each sub-functions should unite the data points at each end of sub-intervals. Furthermore, we need to enforce the first and second derivatives to make curves being smooth and almost perfect for our purposes:

$$S'_{i-1}(y_i) = S'_{i-1}(y_i), \quad S''_{i-1}(y_i) = S''_{i-1}(y_i) \quad (n = 1, 2, \ldots, n-1). \tag{3.42}$$

When we also satisfy this condition of first- and second-order derivatives to be continuous, it results in an unbroken smooth curve over all the sub-intervals which is passing though each data point.

CHAPTER 4

EMPRICAL ANALYSES AND PREDICTION MODELS

The aim of our study is to build a model to forecast exchange rates by using monthly data. We gathered together a variable dataset according to literature suggestions. All real-life observation data set are compiled from Thompson Reuters data base for EUR/TRY and USD/TRY can be found in Appendix B. However, we needed to eliminate the observation data set since some of the observations are highly correlated among themselves. The eliminated data span the period from 01/01/2007 to 30/04/2015, not including weekends and holidays. In order to model EUR/TRY exchange rate, we used variables with their data given by EUR/TRY exchange rates, historical data of EUR/TRY exchange rate, closing prices of major stock indices in Turkey and the Eurozone, spot prices of five precious and non-precious metals and seventeen commodities and macroeconomic variables including the unemployment rate for Eurozone and Turkey, the inflation rate for Eurozone and Turkey, and the monthly GDP rate for Eurozone and Turkey. We also included EURIBOR rates of 3-months and 6-months maturity, refinancing rate for Eurozone, and TRLIBOR rates of 3-months and 6-months maturity. All those real-life observation data sets are compiled from Thompson Reuters data base.

Similarly, to model USD/TRY exchange rate, we employed as variables, based on the data of USD/TRY exchange rates, the historical data of USD/TRY exchange rate, closing prices of major stock indices in Turkey and the United States, spot prices of five precious and non-precious metals and seventeen commodities and macroeconomic variables including unemployment rate for US and Turkey, inflation rate for US and Turkey, and monthly GDP rate for US and Turkey. We also include the USLIBOR rates of 6-months and 3-months maturity, US target rate and TRLIBOR rates of 6-months and 3-months maturity. We collected all variables used in the model listed as given in Table 4.1 and Table 4.2 below.

Moreover, we created dummy variables for the GDP rates and the inflation rate in Turkey, US and Eurozone to detect the effect of negative and positive trends in these variables on exchange rates.

In order to build a model with these variables, we required consistent time intervals. Since our whole data observations do not provide the same time frequency, we need to start with converting all data points to a monthly data basis. For instance, GDP rates are calculated from an annual and a quarterly basis. Before adding GDP rates as input vari-

Table 4.1: Input Variables for USD/TRY Model.

Commodities	**Metals**	**Stock Indices**	**Interest Rates**	**Macroeconomic Variables**
Ethanol	Palladium	S&P 500	TRLIBOR 6 months	GDP Rate TR
Soybean Oil	Gold	BIST All	USLIBOR 3 months	Unemployment Rate US
Canola	Iron	BIST 100	USLIBOR 6 months	Inflation Rate US
Lumber	ø	BIST30	USLIBOR 3 months	GDP Rate US
Wheat	ø	ø	ø	History of USD/TRY Rate
Feeder Cattle	ø	ø	ø	ø
Oat	ø	ø	ø	ø
Lean Hogs	ø	ø	ø	ø
Gasoil	ø	ø	ø	ø
Coffee	ø	ø	ø	ø
Cocoa	ø	ø	ø	ø
Sugar	ø	ø	ø	ø
Rough Rice	ø	ø	ø	ø
Corn	ø	ø	ø	ø

ø: not included in the model.

ables into our model, we need to convert quarterly GDP observations into monthly observations. This disaggregation problem can be treated by Cubic Spline Interpolation. As alternatives for coping with the disaggregation problem, Denton (1970) approach, Chow-Lin (1971) framework and using quarterly GDP observations for that quarters months can be used [10,13]. Abdul Rashid and Zanaib Jehan compared Denton (1970) approach, Chow-Lin (1971) framework and Cubic Spline Interpolation. According to their study, these three methods give similar and highly correlated results. In the light of Rashid's and Jehan's study, we chose to use Cubic Spline Interpolation [38]. This method cannot create new observations; it can only give estimated data points from original data. We applied Cubic Spline Interpolation method by the help of Matlab; then we compared results both statistically and visually.

Another problem which we faced during our study is correlation and multi-collinearity. Financial variables tend to move together because of the interaction between them. As a result, this movement creates a linear relationship among variables. Linear relationship between independent variables with a higher level causes the multi-collinearity problem while building an econometric model. In order to deal with these difficulties, we eliminated the input variables according to correlation among them. Moreover, the created five alternative models for both the USD/TRL and EUR/TRL analysis. The aim of these models is to overcome multi-collinearity among input variables. Lists of input variables used in alternative models are given in Table 4.3 and Table 4.4

Table 4.2: Input Variables for EUR/TRY Model.

Commodities	**Metals**	**Stock Indices**	**Interest Rates**	**Macroeconomic Variables**
Ethanol	Platinum	EUROSTOXX50	ECB Refinancing Rate	Unemployment Rate TR
Soybean Oil	Silver	BIST All	TRLIBOR 3 months	Inflation Rate TR
Canola	Palladium	BIST 100	TRLIBOR 6 months	GDP Rate TR
Lumber	Gold	BIST30	EURIBOR 3 months	Unemployment Rate EU
Wheat	Iron	ø	EURIBOR 6 months	Inflation Rate US
Feeder Cattle	ø	ø	ø	GDP Rate EU
Oat	ø	ø	ø	History of EUR/TRY Rate
Lean Hogs	ø	ø	ø	ø
Gasoil	ø	ø	ø	ø
Coffee	ø	ø	ø	ø
Cocoa	ø	ø	ø	ø
Sugar	ø	ø	ø	ø
Rough Rice	ø	ø	ø	ø
Corn	ø	ø	ø	ø

ø: not included in the model.

Alternative models created by the variables listed in Table 4.3 and Table 4.4 are given in Table 4.5 and Table 4.6.

Table 4.3: Input Variables for USD/TRY Alternative Models.

Commodities	**Metals**	**Stock Indices**	**Interest Rates**	**Macroeconomic Variables**
Wheat	Gold	BIST 100	TRLIBOR 6 months	Unemployment Rate TR
Sugar	Iron	S&P 500	USDIBOR 6 months	Inflation Rate TR
Rough Rice	ø	ø	ø	GDP Rate TR
Gasoil	ø	ø	ø	Unemployment Rate US
Feeder Cattle	ø	ø	ø	Inflation Rate US
Corn	ø	ø	ø	GDP Rate US
ø	ø	ø	ø	History of USD/TRY Rate

ø: not included in the model.

Table 4.4: Input Variables for EUR/TRY Alternative Models.

Commodities	**Metals**	**Stock Indices**	**Interest Rates**	**Macroeconomic Variables**
Wheat	Gold	BIST All	TRLIBOR 3 months	Unemployment Rate TR
Sugar	Iron	EUROSTOXX 50	EURIBOR 3 months	Inflation Rate TR
Rough Rice	ø	ø	ø	GDP Rate TR
Gasoil	ø	ø	ø	Unemployment Rate EU
Feeder Cattle	ø	ø	ø	Inflation Rate EU
Corn	ø	ø	ø	GDP Rate EU
ø	ø	ø	ø	History of EUR/TRY Rate

ø: not included in the model.

Table 4.5: Alternative Models for USD/TRY data.

Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
History of USD/TRY Rate	History of USD/TRY Rate	History of USD/TRY Rate	History of USD/TRY Rate	History of USD/TRY Rate
Rough Rice	Gold	Gasoil	Inf US^+ Dummy	Rough Rice
BIST 100	Inf TR^+ Dummy	6 mo Libor TR	Inf TR^+ Dummy	BIST 100
6 mo Libor US	Monthly GDP US^- Dummy	Monthly GDP US^- Dummy	Inf TR^- Dummy	Inf TR^+ Dummy
ø	ø	ø	Monthly GDP US^- Dummy	Inf TR^- Dummy
ø	ø	ø	Monthly GDP US^+ Dummy	Monthly GDP US^+ Dummy
ø	ø	ø	ø	Monthly GDP TR^+ Dummy

ø: not included in the model.

Table 4.6: Alternative Models for EUR/TRY data.

Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
History of EUR/TRY Rate	History of EUR/TRY Rate	History of EUR/TRY Rate	History of EUR/TRY Rate	History of EUR/TRY Rate
Rough Rice	Wheat	Rough Rice	BIST All	Rough Rice
Gasoil	3 mo Libor TR	Wheat	Inf EU^- Dummy	Wheat
Inf EU^- Dummy	Inf TR^+ Dummy	Inf EU^- Dummy	Inf EU^+ Dummy	Inf EU^- Dummy
Inf TR^+ Dummy	Inf TR^- Dummy	Inf EU^+ Dummy	Inf TR^+ Dummy	Inf EU^+ Dummy
Inf TR^- Dummy	Monthly GDP EU^- Dummy	Inf TR^+ Dummy	Inf TR^- Dummy	Inf TR^+ Dummy
Monthly GDP TR^+ Dummy	ø	Inf TR^- Dummy	Monthly GDP TR^+ Dummy	Inf TR^- Dummy
ø	ø	Monthly GDP EU^- Dummy	Monthly GDP TR^- Dummy	Iron

ø: not included in the model.

In order to distinguish the effect of increase and decrease on exchange rates, we generated dummy variables for percentage change in inflation and monthly GDP rates. Alternative variables were employed to solve collinearity problem contains these dummy variables. In Table 4.5 and 4.6, these variables notated with abbreviations. Increase in inflation for Eurozone is stated by Inf EU+ Dummy while decrease is stated by Inf EU- Dummy. Similarly, these abbreviation is employed for inflations of Turkey and US. Moreover, increase in monthly GDP rate for Eurozone is stated by Monthly GDP EU+ Dummy while decrease is stated by Monthly GDP- Dummy. Similarly, these abbreviations is employed for the monthly GDP rates of Turkey and US. Addition to these, the exchange rate data are a time-series data, exchange rates are affected by the histories of themselves according to the nature of time series-data. History of EUR/TRY Rate and History of USD/TRY notations are stands for the historical data of EUR/TRY and USD/TRY exchange rates, respectively.

We modeled USD/TRY and EUR/TRY during the period from 01/01/2007 to 30/4/2015 by Linear Regression (LR), Support Vector Regression (SVR) and Multivariate Adaptive Regression Splines (MARS), with alternative models for our study. In order to decide which forecasting method among those explained above is superior to the others, we compared them statistically. For literal results, we split our dataset into testing and training data. Then we fed test and train data into our different models.

4.1 Linear Regression Models

We started our modelling procedure by using Linear Regression as a primitive modelling tool. After checking Linear Regression assumptions, including linear relationship, normality, multi-collinearity, autocorrelation and homoscedasticity, we fed our input variables of alternative models into the regression algorithm.

All alternative models were compared according to statistical modelling performance measurements, including R, adjusted R^2, Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE). This comparison resulted in the insight that Model 4 for EUR/TRY data and Model 1 for USD/TRY data have a better modelling performance among other alternative models.

According to Linear Regression algorithm results, we can model the EUR/TRY exchange rate by the input variables' history of EUR/TRY exchange rate, BIST stock indices which is notated as "BIST All" in our study, negative trend in inflation rate for Turkey, positive trend in inflation rate of Turkey, negative trend in inflation rate of Eurozone, positive trend in inflation rate of Eurozone, negative trend in monthly GDP rate of Turkey and positive trend in monthly rate of Turkey within the time period from 01/01/2007 to 30/04/2015. However, the statistical importance of these input variables should be checked to get an accurate model. For these purposes, we checked which of those input variables are statistically important to model EUR/TRY data. The accurate model was built by history of the EUR/TRY exchange rate, BIST All (Turkish Stock Market) scores, positive trend in inflation rate of Eurozone and negative trend in GDP rate of Turkey. Linear regression method modelled the EUR/TRY exchange rate by these 4 input variables. However, these input variables are also meaningful financially. The link between inflation and exchange rates is known as it was mentioned in our literature review part. In addition to this, Taylor and Sarno in their study explained that nominal exchange rates are deflated by GDP deflators [46]. Stock markets, stock prices and exchange rates are further interrelated financial indicators. It is not surprising that LR model used BIST All variable to model the desired exchange rates [19]. Moreover, since the exchange rate data are a time-series data, exchange rates are affected by the histories of themselves according to the nature of time series-data.

Similarly, we can model the USD/TRY exchange rate by the input variables history of the USD/TRY exchange rate, price of commodity rough rice, BIST 100 scores of Turkey and 6 months LIBOR rate of US within the time period from 1/1/2007 to 30/4/2015. However, we checked the statistical importance of these variables in our linear regression model and found that the history of the USD/TRY exchange rate, price of commodity rough rice and 6 months LIBOR rate of US have a statistical importance in our model. This means that macroeconomic variables of 6 months LIBOR rate of US and price of rough rice as a commodity affect the USD/TRY exchange rate. Moreover, since exchange rates are stochastic processes, the USD/TRY rate is also affected by the history of itself. The link between interest rates and exchange rates are known [42]. Besides, rough rice is an important import and export product. Turkey also imports rough rice. According to ZMO (Agriculture Engineers Chamber of Turkey) [47], Turkey imports rough rice mostly from Italy, United States, India, Egypt, Thailand and Vietnam. This trade affects the USD/TRY exchange rate.

Table 4.7: Input variables used in Linear Regression Models for EUR/TRY and USD/TRY data.

Model	EUR/TRY Model	USD/TRY Model
Input Variable	History of EUR/TRY Data	History of USD/TRY Data
	BIST All	Rough Rice
	Inf EU$^+$ Dummy	6 mo Libor US
	GDP TR$^-$ Dummy	ø

ø: not included in the model.

Final model found for Linear Regression methods includes 4 independent variables for EUR/TRY data and 3 independent variables for USD/TRY data. The forecasting models for EUR/TRY and USD/TRY data are given below respectively.

$$\begin{aligned} \hat{Y}_{EUR/TRY} &= 0,019 + 0,986 \cdot \{\text{history of EUR/TRY rate}\} - 0,000 \cdot \{\text{BIST All}\} \\ &\quad + 0,016 \cdot \{\text{Inf EU}^+ \text{ Dummy}\} + 0,008 \cdot \{\text{GDP TR}^- \text{ Dummy}\}, \\ \hat{Y}_{USD/TRY} &= 0,121 + 0,884 \cdot \{\text{history of USD/TRY rate}\} - 0,038 \cdot \{\text{Rough Rice}\} \\ &\quad + 0,006 \cdot \{\text{6 mo Libor US}\}. \end{aligned} \tag{4.1}$$

4.2 Support Vector Regression Models

After LR models, we fed input variables of alternative models into SVR algorithm to get the final model. Similar to the LR procedure, we checked our alternative models' performance measures. According to these measures, Model 3 for the EUR/TRL data and Model 5 for the USD/TRL data resulted in better modeling performances than the others.

According to Support Vector Regression algorithm results, we can model the EUR/TRL exchange rate by the following input variables' history of the EUR/TRL exchange rate, price of rough rice, price of wheat, negative trend in inflation rate of Eurozone, positive trend in inflation rate of Eurozone, negative trend in inflation rate of Turkey, positive trend in inflation rate of Turkey, negative trend in the monthly GDP rate of Eurozone within the time period 01/01/2007 to 30/04/2015. We needed to check the statistical importance of these inputs for the model accuracy. The history of EUR/TRL exchange rate, price of rough rice, negative trend in monthly GDP rate of Eurozone, positive trend in inflation rate of Eurozone and Turkey are important variables to build an accurate model statistically. These input variables used in SVR model are also financially important as mentioned in our LR modeling part.

We did likewise with the EUR/TRL model, following the same procedure for USD/TRL. SVR algorithm used the history of USD/TRL exchange rate, price of rough rice, BIST100 score of Turkey, negative trend in inflation rate of Turkey, positive trend in inflation rates for Turkey, positive trend in monthly GDP rate of US and positive trend in monthly GDP rate of Turkey as input variables. Similarly, the linear regression procedure, we have checked the statistical importance of these input variables in the forecasting model. We observed that the input variables' history of USD/TRL exchange rate, price of rough rice, negative trend in inflation rate of Turkey, positive trend in inflation rates for Turkey, positive trend in monthly GDP rate of US, are statistically important in the SVR model. This means that, according to SVR algorithm, inflation rate of Turkey, positive trend in monthly GDP rate of US, price of rough rice and history of USD/TRL exchange rate, affect the USD/TRL rate. This model is also financially meaningful with these input variables.

Table 4.8: Input variables used in Support Vector Regression Models for EUR/TRL and USD/TRL data.

Model	EUR/TRY Model	USD/TRY Model
Input Variable	History of EUR/TRY Data	History of USD/TRY Data
	Rough Rice	Rough Rice
	Inf EU^+ Dummy	Inf TR^+ Dummy
	Inf TR^+ Dummy	Inf TR^- Dummy
	GDP EU^- Dummy	GDP US^+ Dummy

4.3 Multivariate Adaptive Regression Splines Models

The main aim of our study has been to build MARS models on exchange rates and to compare the performance of the MARS models with classical statistical methods. For this purpose, we eliminated input variables and created alternative models according to collinearity and correlation criteria. After having got rid of the variables with higher collinearity and correlation, we split the input variable data set into as test data set and train data set. These test and train data sets without collinearity problem are fed into MARS algorithm. At the end of this procedure, MARS gave us models for the USD/TRY and EUR/TRY rates. Besides of conducting this, we tested the modeling performances of alternative models according to R, Adjusted R^2, MAPE, MAE and RMSE scores. Model 4 for the EUR/TRY data and Model 5 for USD/TRY data performed a better modeling among the other alternative models.

Model 4 for the EUR/TRY data includes the input variables history of the EUR/TRY exchange rate, BIST All, negative trend in inflation rate of Turkey, positive trend in inflation rate of Turkey, negative trend in inflation rate of Eurozone, positive trend in inflation rate of Eurozone, positive trend in monthly GDP rate of Turkey and negative trend in monthly GDP rate of Turkey within the time period between 01/01/2007 to 30/04/2015. However, MARS algorithm made an elimination among these input variables according to the lack-of-fit criterion at its backward stage. After these eliminations, the final MARS model for the EUR/TRY data includes the input variables history of EUR/TRY exchange rate, positive trend in monthly GDP rate of Turkey and positive trend in inflation rate of Turkey.

Similarly, Model 5 for USD/TRY data includes the input variables history of USD/TRY exchange rate, price of rough rice, BIST 100, positive trend in GDP rate of US, positive trend in GDP rate of Turkey, positive and negative trend in inflation rate of Turkey, within the time period from 01/01/2007 to 30/04/2015. At the backward stage of MARS algorithm, these input variables are eliminated according to the lack-of-fit criterion and the accurate model of the USD/TRY exchange rate contains the history of USD/TRY exchange rate, positive trend in the GDP rate of Turkey and positive trend in the inflation rate of Turkey. MARS algorithm suggests the modeling of USD/TRY exchange rate by using these input variables.

Both the EUR/TRY exchange rate model and USD/TRY exchange rate model, which were built by MARS method, are financially meaningful. The exchange rate data are time-series data; as a result, exchange rates are affected by history of itself through the nature of time-series data. Moreover, the financial effects of inflation rates and GDP rates on exchange rates are mentioned in previous sections and sub-sections.

Table 4.9: Input variables used in Multivariate Adaptive Regression Splines Models for EUR/TRY and USD/TRY data.

Model	EUR/TRY Model	USD/TRY Model
Input Variable	History of EUR/TRY Data	History of USD/TRY Data
	Inf TR^{+} Dummy	Inf TR^{+} Dummy
	GDP TR^{+} Dummy	GDP TR^{+} Dummy

Final model found for Multivariate Adaptive Regression Spline methods by Salford Prediction Modeler (SPM) includes 6 basis functions for both EUR/TRY and USD/TRY data [39]. The forecasting models for EUR/TRY and USD/TRY data are given below respectively.

$$
\begin{aligned}
\hat{Y}_{\text{EUR/TRY}} = \; & 0.516 + 0.948 \max\{0, \text{history of EUR/TRY rate} - 0.534\} \\
- \; & 0.744 \max\{0, 0.534 - \text{history of EUR/TRY rate}\} \\
+ \; & 0.032 \max\{0, \text{GDP TR}_{+} + 0.654\} + \max\{0, - 0.654 - \text{GDP TR}_{+}\} \\
- \; & 0.381 \max\{0, 0.557 - \text{history of EUR/TRY rate}\} \cdot \max\{0, \text{GDP TR}_{+} + 0.654\} \\
- \; & 0.010 \max\{0, \text{INF TR}_{+} - 8.675\} \cdot \max\{0, - 0.654 - \text{GDP TR}_{+}\} \\
- \; & 0.366 \max\{0, \text{history of EUR/TRY rate} - 0.412\} \cdot \max\{0, - 0.654 - \text{GDP TR}_{+}\} ,
\end{aligned} \tag{4.2}
$$

where GDP TR_{+} stands for positive trend in GDP rate of Turkey and INF TR_{+} stands for positive trend in inflation rate of Turkey.

$$
\begin{aligned}
\hat{Y}_{\text{USD/TRY}} = \; & 0.785 - 2.581 \max\{0, 0.748 - \text{history of USD/TRY}\} \\
+ \; & 0.014 \max\{0, \text{GDP TR}_{+} + 0.841\} - 0.319 \max\{0, \text{INF TR}_{+} - 5.052\} \\
+ \; & 0.016 \max\{0, 5.052 - \text{INF TR}_{+}\} \\
+ \; & 0.325 \max\{0, 0.668 - \text{history of USD/TRY}\} \cdot \max\{0, 5.052 - \text{INF TR}_{+}\} \\
+ \; & 0.521 \max\{0, \text{INF TR}_{+} - 0.532\} \cdot \max\{0, - 0.748 - \text{history of USD/TRY}\} \\
- \; & 0.176 \max\{0, \text{history of USD/TRY} - 0.591\} \cdot \max\{0, - 0.841 - \text{GDP TR}_{+}\} ,
\end{aligned} \tag{4.3}
$$

where GDP TR_{+} stands for positive trend in GDP rate of Turkey and INF TR_{+} stands for positive trend in inflation rate of Turkey.

CHAPTER 5

STATISTICAL EVALUATION

In our study, we aimed to model USD/TRY and EUR/TRY exchange rates by using different regression models which are LR, SVR and MARS. First, we created 5 alternative models with various input variables according to correlation and multi-collinearity criteria; then we fed our alternative models' data into these model algorithms separately. As a result, we received five forecasting models according to all these regression models for the USD/TRY and EUR/TRY data.

However, we needed to select one model among the other alternative models which shows better statistical model accuracy performance measures, before making a comparison of forecasting ability of regression models. To do that, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), R, Adjusted R^2 and Root Mean Square Error (RMSE) were used as statistical model accuracy measurements.

Table 5.1 presents the alternative models selected by statistical model accuracy measurements for USD/TRY and EUR/TRY data, respectively.

Table 5.1: Alternative models used in forecasting regression model comparison.

	Forecasting Regression Models		
Exchange Rate	**LR**	**SVR**	**MARS**
USD/TRY	Model 1	Model 5	Model 5
EUR/TRY	Model 4	Model 3	Model 4

We found that the models given in Table 6.1 are most accurate models to compare LR, SVR and MARS methods. Before this comparison, we checked the accuracy of input variables in forecasting models, respectively. After the input variable elimination, the EUR/TRY exchange rate models for comparison include History of EUR/TRY Data, BIST All, Inf EU$^+$ Dummy and GDP TR$^-$ Dummy variables for LR model, History of EUR/TRY Data, Rough rice, Inf TR$^+$ Dummy, Inf EU$^+$ Dummy and GDP EU$^-$ Dummy for SVR model, and history of EUR/TRY Data, Inf TR$^+$ Dummy and GDP TR$^+$ Dummy for MARS model as our input variables.

Similarly, the USD/TRY exchange rate models for comparison include History of

USD/TRY Data, Rough rice, 6 Months LIBOR US for LR model, History of USD/TRY Data, Rough rice, Inf TR$^+$ Dummy, Inf TR$^-$ Dummy, GDP US$^+$ Dummy for SVR model, History of USD/TRY Data, Inf TR$^+$ Dummy and GDP TR$^+$ Dummy for MARS model as input variables. After these input variable eliminations, our models were prepared for method comparison.

Before comparison, in order to receive more accurate results, we split the data into train data set and test data set; then we fed our data into the model algorithms. The aim of using train data set is to find potential predictive relationships. In other words, we built forecasting models with train data. On the other hand, the test data were used to evaluate model performances. In machine learning studies and statistics, train and test data are widely employed to get accurate results. We divided data as 60% of the observations for train data and 40% of the observations for test data.

Forecasting models for the EUR/TRY and USD/TRY exchange rates were built by train data with LR, SVR and MARS algorithm. In order to see which model has a superior forecasting ability among the others, we need to check their R^2, Adjusted R^2, Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) results. We turn to account and compare the model performances by using these measurements.

MAPE is the most common measure to check the forecast error. It measures the accuracy of the fitted model values statistically. In order to apply MAPE, time-series data should be homogeneous or equally spaced or they should be of identical size. MAPE does not work if there exists a 0 value of the response the variable in data set. Moreover, if the predicted response values are too high, the MAPE value will exceed 100%, and if the predicted response values are too low, there will be no upper limit for the MAPE value. MAPE gives best results if there are no extreme observations in dataset. Formulation of the MAPE given in below the equation,

$$\text{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{(\hat{y}_i - y_i)}{y}, \tag{5.1}$$

where $\hat{y}_i$ is the forecast of the actual value of y_i, and n is the number of the observations.

Similarly, to MAPE, MAE also measures forecasting error. It calculates the difference between forecasts and eventual outcomes. Equation (6.2) shows the calculation of the MAE result:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) = \frac{1}{n}\sum_{i=1}^{n}|e_i|, \tag{5.2}$$

again $\hat{y}_i$ being the forecast of the actual value of y_i, n being the number of the observations and e_i being the error term.

MAPE and MAE are independent of the input variables' magnitude; as a result, they allow us to compare Linear Regression Models, SVR models and MARS models to forecast EUR/TL and USD/TL exchange rates.

Both R and Adjusted R^2, on the other hand, measure how well the data fit a curve or a line as follows:

$$R = \frac{1}{1-n}\sum_{i=1}^{n}\Big(\frac{x_i-\bar{x}}{s_x}\Big)\Big(\frac{y_i-\bar{y}}{s_y}\Big), \tag{5.3}$$

where $s_x = \sqrt{\frac{1}{1-n}(\sum_{i=1}^{n} x_i - \bar{x})^2}$, $\quad s_y = \sqrt{\frac{1}{1-n}(\sum_{i=1}^{n} y_i - \bar{y})^2}$ and

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}, \tag{5.4}$$

$$\text{Adjusted } R^2 = 1 - \bigg(\frac{(n-1)(1-R^2)}{n-k-1}\bigg), \tag{5.5}$$

where n is the number of observations, k is the number of independent variables in the regression equation. RMSE measures the squared root of residuals' variance as given below:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n}}, \tag{5.6}$$

where n is the number of observations, and $\hat{y}$ is the forecast of the actual response.

It should be keep in mind that Adjusted R and R values which are closer to 1 are better, while RMSE, MAPE and MAE values close to 0 provide better results.

The model accuracy performances of LR, SVR and MARS regression models for EUR/TRY and USD /TRY are reported in Table 5.2 and Table 5.3, respectively.

Table 5.2: Comparison of the EUR/TRY models performances.

	Forecasting Regression Models					
	LR		SVR		MARS	
	Train	Test	Train	Test	Train	Test
MAE	0.003	0.009	0.003	0.018	0.002	0.008
MAPE	0.006	0.025	0.005	0.048	0.004	0.020
R	0.996	0.964	0.997	0.952	0.998	0.971
Adjusted R^2	0.991	0.915	0.993	0.680	0.996	0.929
RMSE	0.004	0.011	0.004	0.021	0.003	0.009

Table 5.3: Comparison of the USD/TRY models performances.

	Forecasting Regression Models					
	LR		SVR		MARS	
	Train	Test	Train	Test	Train	Test
MAE	0.006	0.018	0.006	0.023	0.003	0.012
MAPE	0.009	0.039	0.008	0.051	0.005	0.025
R	0.993	0.962	0.993	0.914	0.998	0.970
Adjusted R^2	0.985	0.819	0.985	0.6477	0.996	0.917
RMSE	0.008	0.024	0.008	0.032	0.004	0.015

The training data have been gathered for the monthly USD/TRY and EUR/TRY exchange rates of 2007-2011, and the test data have been provided for the monthly USD/TRY and EUR/TRY exchange rates of 2012-2015. For both LR, SVR and MARS models, the predicted and observed values are given at Figures 5.1-5.4 on the same graphs.

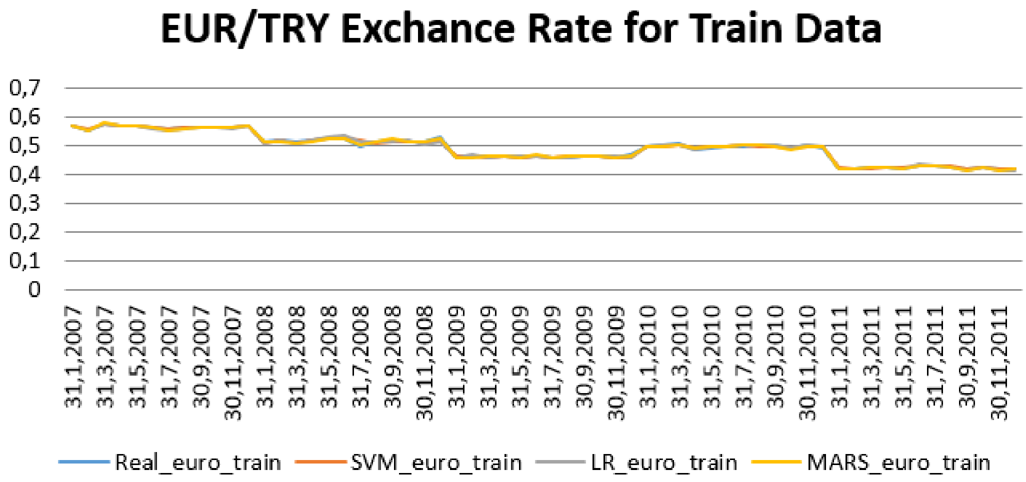

Figure 5.1: Real and Forecast Values of the EUR/TRY Exchange Rate for Training Data.

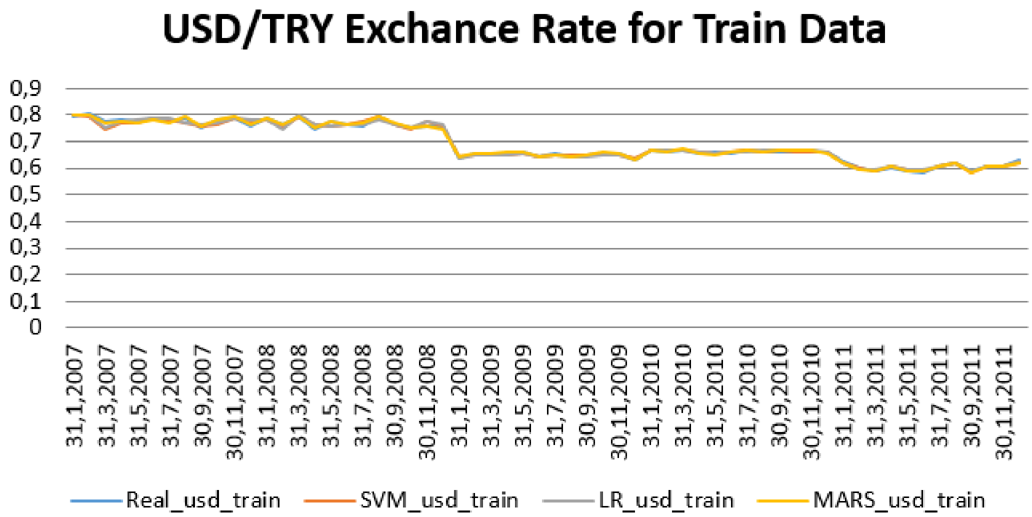

Figure 5.2: Real and Forecast Values of the USD/TRY Exchange Rate for Training Data.

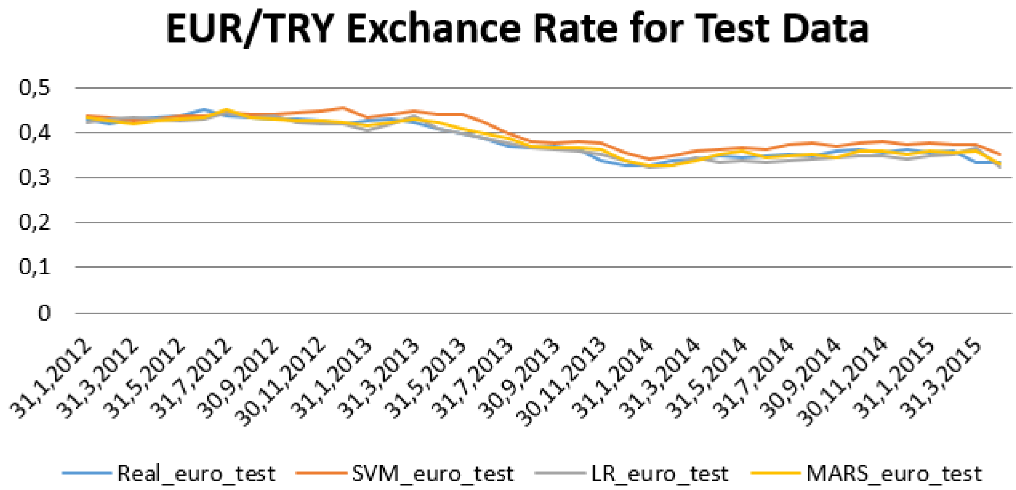

Figure 5.3: Real and Forecast Values of the EUR/TRY Exchange Rate for Test Data.

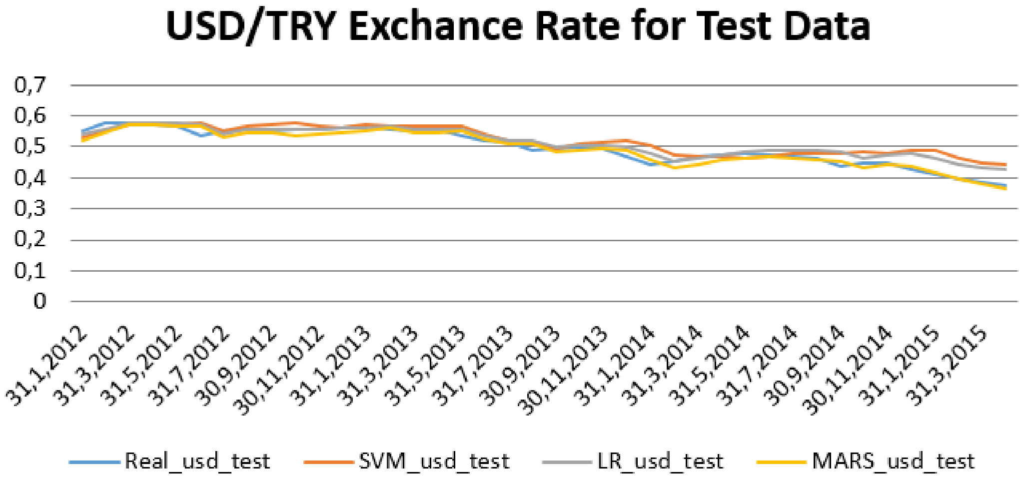

Figure 5.4: Real and Forecast Values of the USD/TRY Exchange Rate for Test Data.

As we may make a deduction from Figure 5.1 - Figure 5.4, when the real and forecast values of the Exchange Rates for LR, SVR and MARS algorithms are taken into account, these can provide adequate results. Considering the modeling phases, LR and MARS algorithms take less time and saving on time. When LR and MARS models are compared, for the test data, MARS model performed better than Linear Regression model for USD/TRY exhange rate model when all measures are considered. More specifically, MARS method provides higher R and Adjusted R2 results and lower MAE, MAPE and RMSE results for test data. However, for EUR/TRY exchange rate modeling, performance measures of LR and MARS models are close. In this circumstances, it depends on the modelers choice to select the method to model EUR/TRY exchange rate.

CHAPTER 6

CONCLUSION AND OUTLOOK

In this study, LR, SVR and MARS algorithms are applied to forecast EUR/TRY and USD/TRY exchange rate models. Forecasting exchange rates are built by a training dataset which has been collected from Thompson Reuters database, including the years 2007-2011. The testing dataset consist of monthly data, including the years 2012-2015. Before using training data to build models, we needed to overcome the data disaggregation problem and multi-collinearity between independent variables. The Cubic Spline Interpolation method enabled us to convert quarterly GDP rate data into monthly GDP rate. Hence, dataset variables were represented as monthly basis. Moreover, to detect the effect of change in inflation rates and GDP rates on the exchange rates, dummy variables were employed. Addition to these, 5 alternative models were taken into account to overcome the multi-collinearity problem. These alternative models contain input variables which do not have any correlation and multi-collinearity among themselves. Selection of the input variables for the alternative models were done according to a literature suggestion.

Following these preparations, we eliminated the number of alternative models according to performance measurements for LR, SVR and MARS, respectively. For EUR/TRY exchange rate modeling, Model 4 for LR modeling, Model 3 for SVR modeling and Model 4 for MARS modeling gave more accurate results when being compared with the others. Similarly, for USD/TRY exchange ling, we received more accurate results among the others in Model 1 for LR modeling, Model 5 for SVR modeling and Model 5 for MARS modeling. However, input variables were eliminated according to their statistical accuracy in the models. The final alternative models with input variable eliminations are also financially meaningful. Alternative models for all methods contain history of the exchange rates on the account of the fact that the nature of being time series data. Except the history of the rates, for EUR/TRY exchange rate modeling, LR models the rate by the input variables BIST All, Inflation and Monthly GDP variables, SVR models the rate by price of commodity rough rice, Inflation and Monthly GDP variables and MARS models the rate by Inflation and Monthly GDP. The effect of inflation on exchange rates are known by Fisher Equation. According to the equation, real interest rates are calculated by subtracting the inflation rate from nominal interest rate. Moreover, forward rates can be calculated by inflation and the spot rates. GDP rate also another key figure for exchange rates since the balance of trade between imports and exports are crucial for economy. Besides, rough rice is an important import and export product. According to ZMO (Agriculture Engineers Chamber of Turkey) [47], Turkey imports rough rice mostly from Italy, United States, India, Egypt, Thailand and Vietnam. This trade affects the USD/TRY exchange rate. Another factor which affects the exchange rates is stock prices. From asset pricing viewpoint, the correlation between exchange rates and equity returns are known. From foreign investors viewpoint, before investing on a foreign market they need the domestic currency. This requirement causes increase in demand of the domestic currency. Hence, the domestic currency gain power over other currencies. After this elimination, we compared the model performances according to R2, Adjusted R, MAE, MAPE and RMSE scores criteria. According to model performance comparison, it is realized that MARS method outperform among LR and SVR methods for the EUR/TRY and USD/TRY exchange rate for our data set.

In this study, we aimed to model EUR/TRY and USD/TRY exchange rates by the Linear Regression (LR) method, Support Vector Regression (SVR) method and Multivariate Adaptive Regression Splines (MARS) method and compared these methods' forecasting ability. According to our dataset and our forecasting models comparison, we found out that MARS method has a superior forecasting ability over LR and SVR methods for EUR/TRY and USD/TRY exchange rates.

This study may be extended for theoretical and practical purposes. Exchange rates are in constant fluctuation. For this reason, modeling the exchange rates with daily or hourly data can be one possible future study. Another issue is that, exchange rates are quite sensitive to current-account deficits, public debt and political stability. In order to see the effect of these variables on exchange rate modeling, we shall employ them as input variables can be another possible future work. Finally, modeling performances of MARS methods including MARS, Robust Multivariate Adaptive Regression Splines (RMARS), Robust Conic Multivariate Adaptive Regression Splines (RCMARS) and Conic Multivariate Adaptive Regression Splines (CMARS) can be compared [2, 30–36].

REFERENCES

[1] Z. Affes and R. H. Kaffel, *Forecast Bankruptcy Using a Blend of Clustering and MARS Model - Case of US Banks*, SSRN Electronic Journal, 2016.

[2] O. Alp, E. Buyukbebeci, A. I. Cekic, F. Ozkurt, P. Taylan, and G. W. Weber, *CMARS and GAM & CQP - Modern Optimization Methods Applied to International Credit Default Prediction*, Journal of Computational and Applied Mathematics, vol. 235, pp. 4639-4651, 2011.

[3] L. Auria and R. Moro, *Support Vector Machines (SVM) as a Technique for Solvency Analysis*, DIW Berlin Discussion Paper No. 811, 2008.

[4] J. Bilson, *The Monetary Approach to the Exchange Rate: Some Empirical Evidence*, IMF Staff Papers, vol. 25, pp. 48-75, 1978.

[5] M. Bordo, *Tales from the Bretton Woods*, Open Economies Review, vol. 25, no. 5, pp. 981-991, 2014.

[6] B. E. Boser, I. M. Guyon, and V. N. Vapnik, *A Training Algorithm for Optimal Margin Classifiers, in Proceedings of the Fifth Annual Workshop on Computational Learning Theory - COLT '92*, Pittsburgh, 1992.

[7] Y.-W. Cheung, *Long Memory in Foreign-Exchange Rates*, Journal of Business and Economic Statistics, vol. 11, pp. 93-101, 1993.

[8] Y.-W. Cheung, M. Chinn, and A. G. Pascual, *Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?*, Journal of International Money and Finance, vol. 24, no. 7, pp. 1150-1175, 2005.

[9] S. M. Chou, T. S. Lee, Y. E. Shao, and I. F. Chen, *Mining the Breast Cancer Pattern Using Artificial Neural Networks and Multivariate Adaptive Regression Splines*, Expert Systems with Applications, vol. 27, pp. 133-142, 2004.

[10] G. C. Chow and A. Lin, *Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series*, Review of Economics and Statistics, vol. 53, pp. 372-375, 1971.

[11] C. Cortes and V. Vapnik, *Support-Vector Networks*, Machine Learning, vol. 20, p. 273–297, 1995.

[12] P. Craven and G. Wahba, *Smoothing Noisy Data with Spline Functions: Estimating the Correct Degree of Smoothing by the Method of Generalized Cross-Validation*, Numer Math, vol. 31, pp. 317-403, 1979.

[13] F. T. Denton, *Quadratic Minimization Procedures for Adjusting Monthly or Quarterly Time Series to Independent Annual Totals*, Hamilton: McMaster University Department of Economics, 1970.

[14] R. Dornbusch, *Expectations and Exchange Rate Dynamics*, Journal of Political Economy, vol. 86, pp. 1161-1176, 1976.

[15] C. Erataman, *Two Essays in Modeling and Analysis of Exchange Rate*, Ankara: Middle East Technical University, 2009.

[16] J. M. Fleming, *Domestic Financial Policies Under Fixed and Floating Exchange Rates*, IMF Staff Papers, pp. 369-379, 1962.

[17] J. A. Frankel, *On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials*, American Economic Review, vol. 69, pp. 610-622, 1979.

[18] J. H. Friedman, *Multivariate Adaptive Regression Splines*, The Annals of Statistics, vol. 19, no. 1, pp. 1-67, 1991.

[19] M. Gavin, *The Stock Market and Exchange Rate Dynamics*, Journal of International Money and Finance, vol. 8, no. 2, pp. 181-200, 1989.

[20] L. S. Goldberg and C. D. Kolstad, *Foreign Direct Investment, Exchange Rate Variability and Demand Uncertainty*, International Economic Review, vol. 36, no. 4, pp. 855-873, 1995.

[21] D. Karemera and B. Kim, *Assessing the Forecasting Accuracy of Alternative Nominal Exchange Rate Models: The Case of Long Memory*, Journal of Forecasting, vol. 25, no. 5, pp. 369-380, 2006.

[22] R. E. Lucas, *Interest Rates and Currency Prices in a Two-Country World*, Journal of Monetary Economics, vol. 10, pp. 335-360, 1982.

[23] N. C. Mark and D. Sul, *Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post-Bretton Woods Panel*, Journal of International Economics, vol. 53, no. 1, pp. 29-52, 2001.

[24] S. McKinley and M. Levine, *Cubic Spline Interpolation*, Eureka: College of the Redwoods, 1998.

[25] R. Meese and K. Rogoff, *Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?*, Journal of International Economics, vol. 14, pp. 3-24, 1983.

[26] Mlcorner, *Machine Learning Corner*, Wordpress, 31 July 2013. [Online]. Available: https://mlcorner.wordpress.com/tag/support-vector-machine/. [Accessed 07 May 2016], 1983.

[27] A. Mohamed, *On the Prediction Accuracy of the Mid-term Electricity Prices Using Optimized Support Vector Machine*, Halifax: Dalhousie University, 2016.

[28] T. Molodtsova and D. H. Papell, *Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals*, Journal of International Economics, vol. 77, pp. 167-180, 2009.

[29] R. Mundell, *International Economics*, New York: Macmillan: New York, 1998.

[30] A. Ozmen, *Robust Conic Quadratic Programming Applied to Quality Improvement – A Robustification of CMARS*, M.Sc. Thesis, Middle East Technical University, 2010.

[31] A. Ozmen, *Advences in Robust Identification of Spline Models and Networks by Robust Conic Optimization, with Applications to Different Sectors*, Phd thesis, Scientific Computing, Institute of Applied Mathematics, Middle East Technical University, Scientific Computing, Institute of Applied Mathematics, Middle East Technical University, Ankara, Turkey, 2015.

[32] A. Ozmen, E. Kropat, and G. W. Weber, *Spline Regression Models for Complex Multimodal Regulatory Networks*, Optimization Methods and Software, vol. 29, pp. 515-534, 2014.

[33] A. Ozmen and G.-W. Weber, *A New Robust Optimization Tool Applied on Financial Data*, Pacific Journal of Optimization, 9(3), pp. 535–552, 2013.

[34] A. Ozmen, G.-W. Weber, and I. Batmaz, *The New Robust CMARS (RCMARS) Method*, preprint at Institute of Applied Mathematics, METU, ISI Proceedings of 24th MEC-EurOPT 2010 –Continuous Optimization and Information-Based Technologies In the Financial Sector, Izmir, Turkey,362-368; ISBN 978- 9955-28-598-4, June 23-26, 2010.

[35] A. Ozmen, G.-W. Weber, Z. Cavusoglu, and O. Defterli, *The New Robust Conic GPLM Method With An Application to Finance, Prediction of Credit Default*, Journal of Global Optimization, 56(2), pp. 233–249, June 2013.

[36] A. Ozmen, G.-W. W. Weber, I. Batmaz, and E. Kropat, *RCMARS: Robustification of CMARS With Different Scenarios Under Polyhedral Uncertainty Set.*, Communications in Nonlinear Science and Numerical Simulation 4780-787. Web., December 16, 2011.

[37] V. Plakandaras, T. Papadimitriou, and P. Gogas, *Forecasting Daily and Monthly Exchange Rates with Machine Learning Techniques*, Journal of Forecasting, vol. 34, no. 7, pp. 560-573, 16 August 2015.

[38] A. Rashid and Z. Jehan, *Derivation of Quarterly GDP, Investment Spending, and Government Expenditure Figures from Annual Data: The Case of Pakistan*, Munich Personal RePEc Archive, 13 May 2013. [Online]. Available: https://mpra.ub.uni-muenchen.de/46937/. [Accessed 27 June 2016], 2014.

[39] Salford-Systems, *SPM Software for MARS*, Salford-Systems, [Online], Available: `http://www.salfordsystems.com`. [Accessed 25 April 2016].

[40] A. L. Samuel, *Some Studies in Machine Learning Using the Game of Checkers*, IBM Journal of Research and Development, vol. 3, no. 3, pp. 210-229, 1959.

[41] P. Samui and D. P. Kothari, *A Multivariate Adaptive Regression Spline Approach for Prediction of Maximum Shear Modulus (Gmax) and Minimum Damping Ratio (εmin)*, Engineering Journal, vol. 16, no. 5, 2012.

[42] M. Sanchez, *The Link Between Interest Rates and Exchange Rates: Do Dontractionary Depreciations Make a Difference?*, International Economic Journal, vol. 22, no. 1, pp. 43-61, 2008.

[43] S. Sekulic and B. R. Kowalski, *MARS: A Tutorial*, J. Chemom., vol. 6, pp. 199-216, 1992.

[44] A. J. Smola and B. Schölkopf, *A Tutorial on Support Vector Regression*, Statistics and Computing, vol. 14, no. 3, p. 199–222, 2004.

[45] A. Stockman, *A Theory of Exchange Rate Determination*, Journal of Political Economy, vol. 88, pp. 673-698, 1980.

[46] M. P. Taylor and L. Sarno, *The Behavior of Real Exchange Rates During the Post-Bretton Woods Period*, Journal of international Economics, vol. 46, no. 2, p. 281–312, 1998.

[47] Z. M. O. TMMOB, *Turkiye'de Celtik Uretimi ve Ithalati*, Available: `http://www.zmo.org.tr/genel/bizden_detay.php?kod=23114`, 08 October 2016. Online.

[48] R. D. Veaux, D. Psichogios, and L. Ungar, *A Comparison of Two Nonparametric Estimation Schemes: MARS and Neural Networks, Computers & Chemical Engineering*, International Journal of Computer Applications in Chemical Engineering, vol. 17, no. 8, p. 819–837, 1993.

[49] Y. Yilmaz, *Daily Natural Gas Consumption Prediction by MARS and CMARS Models for Residential Users in Ankara*, Ankara: Middle East Technical University, 2015.

[50] H. Zareipour, K. Bhattacharya, and C. Canizares, *Forecasting the Hourly Ontario Energy Price by Multivariate Adaptive Regression Splines*, in Power Engineering Society General Meeting, 2006.

[51] W. Zhaohua and N. E. Huang, *Ensemble Empirical Mode Decomposition: A Noise-Assited Data Analysis Method*, Advances in Adaptive Data Analysis, vol. 1, no. 1, pp. 1-41, 2009.

APPENDIX A

Cubic Spline Interpolation Results

A.1 GDP for Turkey

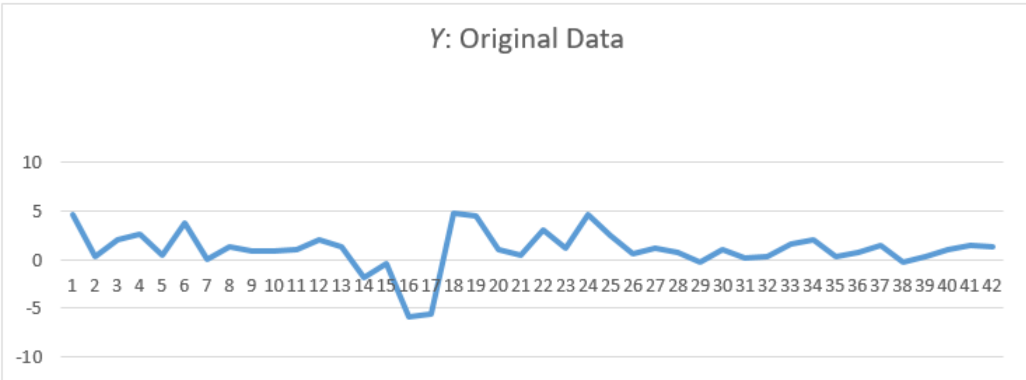

Figure A.1: Original data for GDP rate for Turkey.

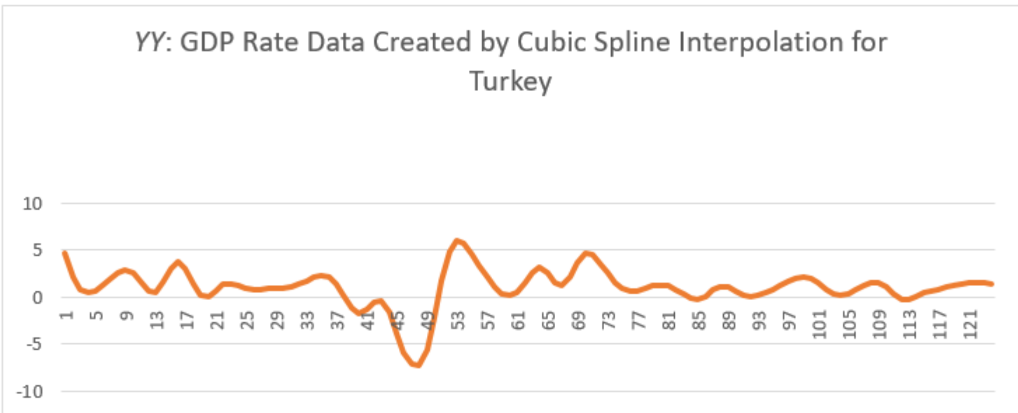

Figure A.2: GDP rate created by cubic spline interpolation for Turkey.

Table A.1: Descriptive statistics of original data for GDP rate for Turkey.

Y: Original Data	
Mean	1.054762
Standard Error	0.325087
Median	1.05
Mode	0.4
Standard Deviation	2.106807
Sample Variance	4.438635
Kurtosis	4.106684
Skewness	-1.2144
Range	10.7
Minimum	-5.9
Maximum	4.8
Sum	44.3
Count	42
Confidence Level (95%)	0.656527

Table A.2: Descriptive statistics of GDP rate created by cubic spline interpolation for Turkey.

YY: Data created by CSI	
Mean	1.008962
Standard Error	0.179735
Median	1.05292
Mode	0.4
Standard Deviation	2.001443
Sample Variance	4.005775
Kurtosis	5.623044
Skewness	-1.41644
Range	13.20093
Minimum	-7.22412
Maximum	5.97681
Sum	125.1113
Count	124
Confidence Level (95%)	0.355774

GDP rate for Turkey is distributed with mean of 1.054762and variance of 4.438635 within the range of -5.9 and 4.8. GDP rate created by cubic spline interpolation for Turkey is distributed with mean of 1.054762 and variance of 4.005775 within the range of -7.22412 and 5.97681.

t-Test: Two-Sample Assuming Unequal Variances

Table A.3: *t*-Test: Two-Sample Assuming Unequal Variances for GDP rate of Turkey.

	Y	*YY*
Mean	1.054762	1.008962
Variance	4.438635	4.005775
Observations	42	124
Hypothesized Mean Difference	0	
df	68	
t Stat	0.123295	
$P(T \leq t)$ one-tail	0.451119	
t Critical one-tail	1.667572	
$P(T \leq t)$ two-tail	0.902237	
t Critical two-tail	1.995469	

In order to test the difference between data *Y* and *YY* *t*-test is used. For this purpose, we state our hypothesis as,

H_0: $\mu_y - \mu_{yy} = 0$,

$p_{val} > \alpha = 0.05$,

$t_{critical} = -1.995469 < t_{stat} = 0.123295 < t_{critical} = 1.995469$.

Therefore, we do not reject the null hypothesis.
The observed difference between the sample means is not convincing enough to say that the original data and the data created by cubic spline differ significantly. Data and the data created by cubic spline differ significantly.

A.2 GDP for US

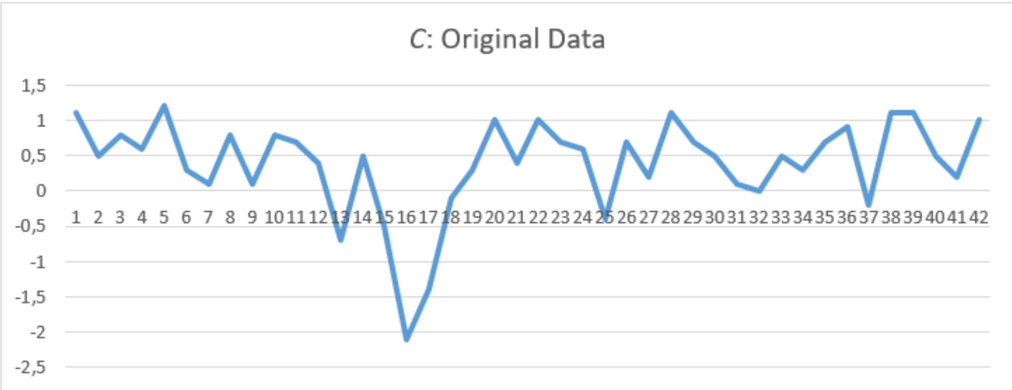

Figure A.3: Original data for GDP rate for US.

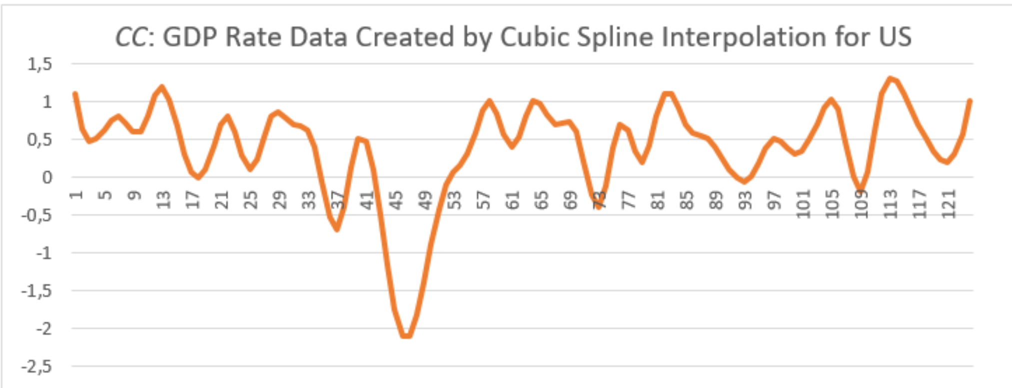

Figure A.4: GDP rate created by cubic spline interpolation for US.

Table A.4: Descriptive statistics of GDP rate for the US.

C: Original data	
Mean	0.383333333
Standard Error	0.103027159
Median	0.5
Mode	0.5
Standard Deviation	0.6676923
Sample Variance	0.445813008
Kurtosis	4.278986449
Skewness	-1.773606763
Range	3.3
Minimum	-2.1
Maximum	1.2
Sum	16.1
Count	42
Confidence Level (95%)	0.208067568

Table A.5: Descriptive statistics of GDP rate created by cubic spline interpolation for the US.

CC: Data created by CSI	
Mean	0.366271821
Standard Error	0.057228672
Median	0.5
Mode	0.5
Standard Deviation	0.637271524
Sample Variance	0.406114996
Kurtosis	4.463576276
Skewness	-1.857661612
Range	3.413092562
Minimum	-2.106952019
Maximum	1.306140542
Sum	45.41770582
Count	124
Confidence Level (95%)	0.113280648

GDP rate for US is distributed with mean of 0.383333 and the variance of 0.445813 within the range of -2.1 and 1.2 while GDP rate created by cubic spline interpolation for US is distributed with mean of 0.366272 and variance of 0.406115 within the range of -2.10695 and 1.306141.

t-Test: Two-Sample Assuming Unequal Variances

Table A.6: *t*-Test: Two-Sample Assuming Unequal Variances for GDP rate of the US.

	C	*CC*
Mean	0.383333	0.366271821
Variance	0.445813	0.406114996
Observations	42	124
Hypothesized Mean Difference	0	
df	68	
t Stat	0.144767	
$P(T \leq t)$ one-tail	0.442661	
t Critical one-tail	1.667572	
$P(T \leq t)$ two-tail	0.885323	
t Critical two-tail	1.995469	

In order to test the difference between data *C* and *CC* *t*-test is used. For this purpose, we state our hypothesis as,

H_0: $\mu_y - \mu_{yy} = 0$,

$p_{val} > \alpha$, = 0.05,

$t_{critical} = -1.995469 < t_{stat} = 0.144767 < t_{critical} = 1.995469.$

Therefore, we do not reject the null hypothesis.
The observed difference between the sample means is not convincing enough to say that the original data and the data created by cubic spline differ significantly. Data and the data created by cubic spline differ significantly.

A.3 GDP for EU

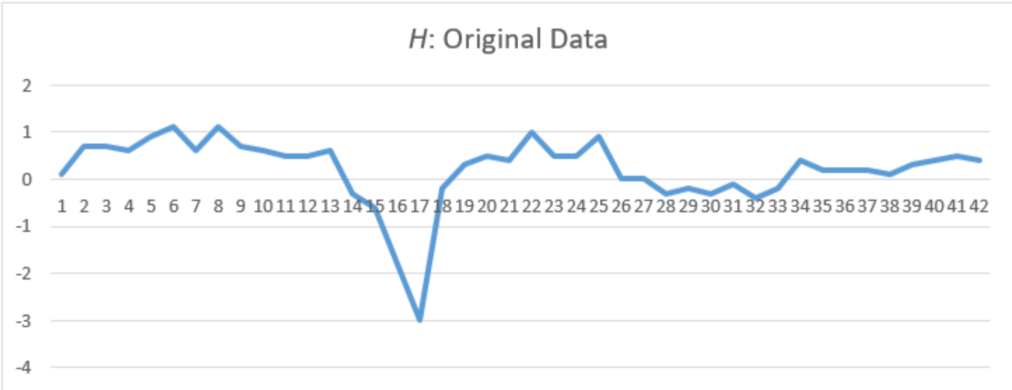

Figure A.5: Original data for GDP rate for the Eurozone.

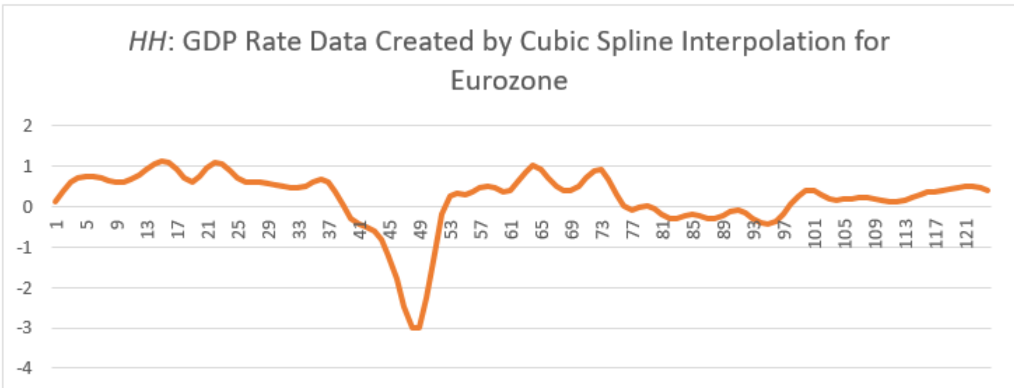

Figure A.6: GDP rate created by cubic spline interpolation for Eurozone.

Table A.7: Descriptive statistics of GDP rate for the Eurozone.

H: original data	
Mean	0.192857
Standard Error	0.11305
Median	0.4
Mode	0.5
Standard Deviation	0.732651
Sample Variance	0.536777
Kurtosis	8.833221
Skewness	-2.48455
Range	4.1
Minimum	-3
Maximum	1.1
Sum	8.1
Count	42
Confidence Level (95%)	0.22831

Table A.8: GDP rate created by cubic spline interpolation for the Eurozone.

HH: data created by cubic spline	
Mean	0.194374
Standard Error	0.064937
Median	0.378236
Mode	0.5
Standard Deviation	0.72311
Sample Variance	0.522888
Kurtosis	7.244153
Skewness	-2.34076
Range	4.136463
Minimum	-3.00415
Maximum	1.132308
Sum	24.10235
Count	124
Confidence Level (95%)	0.128539

GDP rate for Eurozone is distributed with mean of 0.192857 and the variance of 0.732651 within the range of -3 and 1.1 while GDP rate created by cubic spline interpolation for Eurozone is distributed with mean of 0.194374 and variance of 0.522888 within the range of -3.00415 and 1.132308.

t-Test: Two-Sample Assuming Unequal Variances

Table A.9: *t*-Test: Two-Sample Assuming Unequal Variances for GDP rate of the Eurozone.

	H	*HH*
Mean	0.192857	0.194373771
Variance	0.536777	0.522887632
Observations	42	124
Hypothesized Mean Difference	0	
df	70	
t Stat	-0.01163	
P(*T*≤*t*) one-tail	0.495376	
t Critical one-tail	1.666914	
P(*T*≤*t*) two-tail	0.990752	
t Critical two-tail	1.994437	

In order to test the difference between data *H* and *HH* *t*-test is used. For this purpose, we state our hypothesis as,

H_0: $\mu_y - \mu_{yy} = 0$,

$p_{val} > \alpha$ = 0.05,

$t_{critical} = -1.995469 < t_{stat} = -0.01163 < t_{critical} = 1.995469.$

Therefore, we do not reject the null hypothesis.
The observed difference between the sample means is not convincing enough to say that the original data and the data created by cubic spline differ significantly. Data and the data created by cubic spline differ significantly.

APPENDIX B

Input Variables Used for Modeling

Table B.1: Input Variables for USD/TRY Model.

Commodities	Metals	Stock Indices	Interest Rates	Macroeconomic Variables
Ethanol	Platinum	Dow Jones	US FED Target	Unemployment Rate TR
Soybean Oil	Silver	NASDAQ	TR LIBOR 3 months	Inflation Rate TR
Canola	Palladium	S&P 500	TR LIBOR 6 months	GDP Rate TR
Lumber	Gold	BIST All	US LIBOR 3 months	Unemployment Rate US
Wheat	Iron	BIST 100	US LIBOR 6 months	Inflation Rate US
Soybean Meal	ø	BIST 30	ø	History of USD/TRY Rate
Oat	ø	ø	ø	GDP Rate US
Lean Hogs	ø	ø	ø	ø
Gasoil	ø	ø	ø	ø
Coffee	ø	ø	ø	ø
Cocoa	ø	ø	ø	ø
Sugar	ø	ø	ø	ø
Rough Rice	ø	ø	ø	ø
Feeder Cattle	ø	ø	ø	ø
Corn	ø	ø	ø	ø
Soybeans	ø	ø	ø	ø

ø: not included in the model.

Table B.2: Input Variables for EUR/TRY Model.

Commodities	Metals	Stock Indices	Interest Rates	Macroeconomic Variables
Ethanol	Platinum	EUROSTOXX50	ECB Refinancing Rate	Unemployment Rate TR
Soybean Oil	Silver	BIST All	TR LIBOR 3 months	Inflation Rate TR
Canola	Palladium	BIST 100	TR LIBOR 6 months	GDP Rate TR
Lumber	Gold	BIST 30	EURIBOR 3 months	Unemployment Rate EU
Wheat	Iron	ø	EURIBOR 6 months	Inflation Rate EU
Soybean Meal	ø	ø	ø	History of EUR/TRY Rate
Oat	ø	ø	ø	GDP Rate EU
Lean Hogs	ø	ø	ø	ø
Gasoil	ø	ø	ø	ø
Coffee	ø	ø	ø	ø
Cocoa	ø	ø	ø	ø
Sugar	ø	ø	ø	ø
Rough Rice	ø	ø	ø	ø
Feeder Cattle	ø	ø	ø	ø
Corn	ø	ø	ø	ø
Soybeans	ø	ø	ø	ø

ø: not included in the model.

APPENDIX C

Descriptive Statistics of Data Used in Final Model

C.1 USD/TRY Data

Table C.1: Descriptive statistics of USD/TRY data.

	WHEAT	SUGAR	ROUGHRICE	IRON	GOLD
Mean	598.19	16.95	12.90	80.64	1080.66
Standard Error	15.13	0.54	0.29	3.73	35.90
Median	592.75	16.36	13.66	69.69	1141.52
Mode	664.25	13.93	15.62	36.63	869.95
Standard Deviation	168.43	6.04	3.25	41.49	399.81
Sample Variance	28367.00	36.53	10.57	1721.45	159848.90
Kurtosis	-0.48	-0.05	-0.36	-1.29	-1.13
Skewness	0.22	0.78	-0.12	0.29	0.01
Range	771.50	25.36	15.08	134.69	1406.30
Minimum	301.50	8.61	6.40	28.11	417.00
Maximum	1073.00	33.97	21.48	162.80	1823.30
Sum	74175.50	2101.45	1599.39	9998.84	134002.26
Count	124.00	124.00	124.00	124.00	124.00

Table C.2: Descriptive statistics of USD/TRY data (cont'd).

	GASOIL	FEEDERCATTLE	CORN	BIST100	SP 500 US
Mean	745.30	130.84	454.44	54309.28	1379.06
Standard Error	17.80	3.23	15.06	1602.21	27.42
Median	709.63	115.34	404.38	54451.35	1324.27
Mode	915.00	146.30	362.50	#N/A	#N/A
Standard Deviation	198.18	35.98	167.65	17841.45	305.37
Sample Variance	39276.01	1294.65	28107.84	318317161.36	93252.37
Kurtosis	-1.06	1.12	-0.91	-1.01	0.11
Skewness	0.18	1.34	0.38	0.12	0.63
Range	876.00	144.43	619.00	65354.18	1369.41
Minimum	386.50	91.00	187.50	23591.64	735.09
Maximum	1262.50	235.43	806.50	88945.82	2104.50
Sum	92417.00	16224.55	56350.50	6734351.33	171003.63
Count	124.00	124.00	124.00	124.00	124.00

Table C.3: Descriptive statistics of USD/TRY data (cont'd).

	6MOLIBORTR	6MOLIBORUS	UNEMPLOYMENT RATE TR	UNEMPLOYMENT RATE TR
Mean	12.16	2.03	9.91	6.90
Standard Error	0.42	0.18	0.14	0.17
Median	10.51	0.74	9.55	6.70
Mode	8.10	5.33	9.10	4.70
Standard Deviation	4.70	1.96	1.58	1.88
Sample Variance	22.10	3.86	2.50	3.53
Kurtosis	-1.39	-1.17	0.88	-1.48
Skewness	0.38	0.73	1.10	0.17
Range	16.93	5.27	7.50	5.60
Minimum	5.30	0.32	7.30	4.40
Maximum	22.23	5.59	14.80	10.00
Sum	1508.18	251.23	1229.00	855.90
Count	124.00	124.00	124.00	124.00

Table C.4: Descriptive statistics of USD/TRY data (cont'd).

	INFLATION IN USA	**INFLATION IN TR**	**MONTHLY GDP US**	**MONTHLY GDP TR**
Mean	2.22	0.67	0.37	1.01
Standard Error	0.14	0.07	0.06	0.18
Median	2.12	0.56	0.50	1.05
Mode	1.66	0.42	0.50	0.40
Standard Deviation	1.51	0.79	0.64	2.00
Sample Variance	2.28	0.62	0.41	4.01
Kurtosis	0.20	0.43	4.46	5.62
Skewness	-0.39	0.40	-1.86	-1.42
Range	7.70	4.70	3.41	13.20
Minimum	-2.10	-1.43	-2.11	-7.22
Maximum	5.60	3.27	1.31	5.98
Sum	275.08	83.00	45.42	125.11
Count	124.00	124.00	124.00	124.00

Table C.5: Correlation values of USD/TRY data.

	TRY/USD	WHEAT	SUGAR	ROUGHRICE	IRON	GOLD	GASOIL	FEEDERCATTLE
TRY/USD	1.00							
WHEAT	-0.27	1.00						
SUGAR	-0.39	0.29	1.00					
ROUGHRICE	-0.52	0.68	0.39	1.00				
IRON	-0.60	0.48	0.77	0.61	1.00			
GOLD	-0.75	0.56	0.69	0.65	0.88	1.00		
GASOIL	-0.40	0.70	0.37	0.75	0.64	0.66	1.00	
FEEDERCATTLE	-0.75	0.11	0.09	0.17	0.28	0.50	0.30	1.00
CORN	-0.44	0.80	0.56	0.73	0.75	0.82	0.78	0.19
BIST100	-0.74	0.40	0.38	0.37	0.59	0.73	0.49	0.78
SP 500 US	-0.55	0.14	-0.17	0.10	0.06	0.24	0.30	0.86
6MOLIBORTR	0.70	-0.28	-0.71	-0.44	-0.81	-0.82	-0.43	-0.46
6MOLIBORUS	0.79	-0.30	-0.64	-0.54	-0.79	-0.82	-0.41	-0.52
UNEMPLOYMENT RATE TR	-0.06	-0.24	0.08	-0.06	-0.08	-0.13	-0.48	-0.18
UNEMPLOYMENT RATE US	-0.41	0.24	0.79	0.43	0.77	0.67	0.23	-0.04
INFLATION IN US	0.48	0.03	-0.20	0.01	-0.15	-0.24	0.27	-0.21
INFLATION IN TR	-0.04	-0.04	-0.04	0.09	-0.05	0.02	0.06	-0.01
MONTHLY GDP TR	-0.03	-0.24	0.12	-0.28	0.08	0.07	0.10	0.27
MONTHLY GDP US	0.09	-0.11	0.22	-0.32	0.05	-0.01	-0.07	-0.01

Table C.6: Correlation values of USD/TRY data (cont'd).

	CORN	BIST100	SP 500 US	6MOLIBORTR	6MOLIBORTR	UNEMPLOYMENT RATE TR
TRY/USD						
WHEAT						
SUGAR						
ROUGHRICE						
IRON						
GOLD						
GASOIL						
FEEDERCATTLE						
CORN	1.00					
BIST100	0.48	1.00				
SP 500 US	0.08	0.74	1.00			
6MOLIBORTR	-0.52	-0.76	-0.23	1.00		
6MOLIBORUS	-0.53	-0.65	-0.20	0.88	1.00	
UNEMPLOYMENT RATE TR	-0.28	-0.18	-0.32	-0.12	-0.24	1.00
UNEMPLOYMENT RATE US	0.49	0.26	-0.37	-0.75	-0.77	0.35
INFLATION IN US	-0.01	-0.28	-0.04	0.43	0.49	-0.52
INFLATION IN TR	-0.01	0.01	0.07	-0.01	0.03	0.02
MONTHLY GDP TR	-0.14	0.31	0.31	-0.26	-0.08	-0.25
MONTHLY GDP US	-0.10	0.17	0.04	-0.28	-0.04	0.02

Table C.7: Correlation values of USD/TRY data (cont'd).

	UNEMPLOYMENT RATE US	INFLATION IN US	INFLATION IN TR	MONTHLY GDP US	MONTHLY GDP TR
TRY/USD					
WHEAT					
SUGAR					
ROUGHRICE					
IRON					
GOLD					
GASOIL					
FEEDERCATTLE					
CORN					
BIST100					
SP 500 US					
6MOLIBORTR					
6MOLIBORUS					
UNEMPLOYMENT RATE TR					
UNEMPLOYMENT RATE US	1.00				
INFLATION IN US	-0.43	1.00			
INFLATION IN TR	-0.08	0.15	1.00		
MONTHLY GDP TR	0.00	-0.02	0.06	1.00	
MONTHLY GDP US	0.14	-0.21	-0.03	0.52	1.00

C.2 EUR/TRY Data

Table C.8: Descriptive statistics of EUR/TRY data.

	WHEAT	SUGAR	ROUGHRICE	IRON	GOLD
Mean	598.19	16.95	12.90	80.64	1080.66
Standard Error	15.13	0.54	0.29	3.73	35.90
Median	592.75	16.36	13.66	69.69	1141.52
Mode	664.25	13.93	15.62	36.63	869.95
Standard Deviation	168.43	6.04	3.25	41.49	399.81
Sample Variance	28367.00	36.53	10.57	1721.45	159848.90
Kurtosis	-0.48	-0.05	-0.36	-1.29	-1.13
Skewness	0.22	0.78	-0.12	0.29	0.01
Range	771.50	25.36	15.08	134.69	1406.30
Minimum	301.50	8.61	6.40	28.11	417.00
Maximum	1073.00	33.97	21.48	162.80	1823.30
Sum	74175.50	2101.45	1599.39	9998.84	134002.26
Count	124.00	124.00	124.00	124.00	124.00

Table C.9: Descriptive statistics of EUR/TRY data (cont'd).

	GASOIL	EUROSTOXX50	EURIBOR3MO	CORN	BISTALL
Mean	745.30	3112.93	1.80	454.44	53734.93
Standard Error	17.80	56.17	0.14	15.06	1618.98
Median	709.63	3013.53	1.25	404.38	53873.29
Mode	915.00	#N/A	0.21	362.50	#N/A
Standard Deviation	198.18	625.48	1.60	167.65	18028.16
Sample Variance	39276.01	391221.25	2.56	28107.84	325014627.06
Kurtosis	-1.06	-0.53	-0.81	-0.91	-1.05
Skewness	0.18	0.54	0.73	0.38	0.12
Range	876.00	2536.42	5.28	619.00	66115.31
Minimum	386.50	1976.23	-0.01	187.50	22641.60
Maximum	1262.50	4512.65	5.28	806.50	88756.91
Sum	92417.00	386003.23	222.83	56350.50	6663130.98
Count	124.00	124.00	124.00	124.00	124.00

Table C.10: Descriptive statistics of EUR/TRY data (cont'd).

	3MOLIBORTR	UNEMPLOYMENT RATE TR	UNEMPLOYMENT RATE EU	INFLATION EU
Mean	11.88	9.91	21.83	0.15
Standard Error	0.42	0.14	0.29	0.04
Median	10.08	9.55	22.59	0.19
Mode	10.00	9.10	#N/A	0.11
Standard Deviation	4.71	1.58	3.18	0.48
Sample Variance	22.16	2.50	10.13	0.23
Kurtosis	-1.40	0.88	-1.08	1.34
Skewness	0.34	1.10	-0.32	-0.47
Range	17.14	7.50	10.49	2.89
Minimum	5.00	7.30	16.03	-1.54
Maximum	22.14	14.80	26.52	1.35
Sum	1473.08	1229.00	2707.23	18.14
Count	124.00	124.00	124.00	124.00

Table C.11: Descriptive statistics of EUR/TRY data (cont'd).

	INFLATION TR	MONTHLY GDP EU	MONTHLY GDP TR
Mean	0.67	0.19	1.01
Standard Error	0.07	0.06	0.18
Median	0.56	0.38	1.05
Mode	0.42	0.50	0.40
Standard Deviation	0.79	0.72	2.00
Sample Variance	0.62	0.52	4.01
Kurtosis	0.43	7.24	5.62
Skewness	0.40	-2.34	-1.42
Range	4.70	4.14	13.20
Minimum	-1.43	-3.00	-7.22
Maximum	3.27	1.13	5.98
Sum	83.00	24.10	125.11
Count	124.00	124.00	124.00

Table C.12: Correlation values of EUR/TRY data.

	TL/EURO	WHEAT	SUGAR	ROUGHRICE	IRON	GOLD
TL/EURO	1.00					
WHEAT	-0.27	1.00				
SUGAR	-0.39	0.29	1.00			
ROUGHRICE	-0.52	0.68	0.39	1.00		
IRON	-0.60	0.48	0.77	0.61	1.00	
GOLD	-0.75	0.56	0.69	0.65	0.88	1.00
GASOIL	-0.40	0.70	0.37	0.75	0.64	0.66
EUROSTOXX50	0.46	-0.09	-0.57	-0.35	-0.63	-0.63
EURIBOR3MO	0.71	0.05	-0.52	-0.09	-0.59	-0.65
CORN	-0.44	0.80	0.56	0.73	0.75	0.82
BISTALL	-0.76	0.39	0.39	0.38	0.61	0.74
3MOLIBORTR	0.69	-0.29	-0.71	-0.45	-0.81	-0.82
UNEMPLOYMENT RATE TR	-0.06	-0.24	0.08	-0.06	-0.08	-0.13
UNEMPLOYMENT RATE EU	-0.73	0.06	0.52	0.21	0.68	0.74
INFLATION IN EU	0.09	0.04	-0.06	0.02	0.01	-0.01
INFLATION IN TR	-0.04	-0.04	-0.04	0.09	-0.05	0.02
MONTHLY GDP EU	0.23	-0.22	0.01	-0.45	-0.14	-0.23
MONTHLY GDP TR	0.09	-0.11	0.22	-0.32	0.05	-0.01

Table C.13: Correlation values of EUR/TRY data (cont'd).

	GASOIL	EUROSTOXX50	EURIBOR3MO	CORN	BISTALL	3MOLIBORTR
TL/EURO						
WHEAT						
SUGAR						
ROUGHRICE						
IRON						
GOLD						
GASOIL	1.00					
EUROSTOXX50	-0.16	1.00				
EURIBOR3MO	-0.12	0.64	1.00			
CORN	0.78	-0.42	-0.24	1.00		
BISTALL	0.49	-0.17	-0.67	0.49	1.00	
3MOLIBORTR	-0.42	0.61	0.83	-0.53	-0.75	1.00
UNEMPLOYMENT RATE TR	-0.48	-0.23	-0.17	-0.28	-0.17	-0.13
UNEMPLOYMENT RATE EU	0.32	-0.64	-0.95	0.38	0.73	-0.83
INFLATION IN EU	0.11	0.11	0.08	0.06	-0.03	-0.01
INFLATION IN TR	0.06	0.10	0.06	-0.01	0.01	-0.01
MONTHLY GDP EU	-0.08	0.49	-0.07	-0.28	0.15	-0.05
MONTHLY GDP TR	-0.07	0.09	-0.27	-0.10	0.16	-0.27

Table C.14: Correlation values of EUR/TRY data (cont'd).

	UNEMPLOYMENT RATE TR	**UNEMPLOYMENT RATE EU**	**INFLATION IN EU**	**INFLATION IN TR**	**MONTHLY GDP EU**	**MONTHLY GDP TR**
TL/EURO						
WHEAT						
SUGAR						
ROUGHRICE						
IRON						
GOLD						
GASOIL						
EUROSTOXX50						
EURIBOR3MO						
CORN						
BISTALL						
3MOLIBORTR						
UNEMPLOYMENT RATE TR	1.00					
UNEMPLOYMENT RATE EU	-0.07	1.00				
INFLATION IN EU	-0.06	-0.07	1.00			
INFLATION IN TR	0.02	-0.04	0.07	1.00		
MONTHLY GDP EU	-0.22	0.04	0.13	0.02	1.00	
MONTHLY GDP TR	0.02	0.20	0.12	-0.03	0.73	1.00