VECTOR AUTOREGRESSIVE MODEL INCORPORATING NEW INFORMATION USING BAYESIAN APPROACH

Michael Ngungu Musyoki

A Thesis Submitted in Partial Fulfillment of the Requirements for the Award of the Degree of Doctor of Philosophy in Statistics of Masinde Muliro University of Science and Technology

November, 2023

TITLE PAGE

DECLARATION

This thesis is my original work prepared with no other than the indicated sources and support and has not been presented elsewhere for a degree or any other award.

Signature.. Date ....

Michael Ngungu Musyoki

SES/H/01-53241/2018

CERTIFICATION

The undersigned certify that they have read and hereby recommend for acceptance of Masinde Muliro University of Science and Technology a thesis entitled "Vector Autoregressive Model Incorporating New Information using Bayesian Approach."

Signature... Date ....

Dr. David Anekeya Alilah

Department of Mathematics

Masinde Muliro University of Science and Technology

Signature... Date ....

Dr. David Nyachae Angwenyi

Department of Mathematics

Masinde Muliro University of Science and Technology

Signature... Date ....

Prof. Samuel Musili Mwalili

School of Mathematical Sciences

Jomo Kenyatta University of Agriculture and Technology

COPYRIGHT

DEDICATION

I dedicate this work to my Dad, Mr. Bernard Musyoki Ngungu and my wife, Mrs. Diana Wayua who always kept on reminding me to keep pressing on.

ACKNOWLEDGEMENTS

First, I take this opportunity to thank the Almighty God for giving me a chance to pursue my studies. My sincere gratitude to my supervisors Dr. David Anekeya Alilah, Dr. David Nyachae Angwenyi and Prof. Samuel Musili Mwalili for their guidance, ideas, insightful assistance and discussions during the entire research period. Without your help and guidance, this thesis would not have been completed. Also, a big thank you to the entire staff in the Department of Mathematics of Masinde Muliro University of Science and Technology (MMUST) led by the Chair, Dr. Achiles Nyongesa Simiyu, for both material and moral support during the entire study period. Special thanks to Dr. Michael Onyango Ojiema, Dr. Frankline Kibiwott Tireito and Dr. Charles Muriuki Wachira who kept on encouraging and informing me that one day it shall be over. To my parents and the entire family, I say thank you so much for always encouraging me to keep on pressing. Almighty God bless you all.

ABSTRACT

The Vector Autoregressive (VAR) Models have been applied extensively in many fields ranging from finance, economics, machine learning among others. In fact, the VAR models are the mostly applied among the multivariate time series models since they have shown to perform well especially when forecasting is done. Many researchers have fitted the VAR models to the available data so as to come up with a model that explains the relationship between the variables involved. However, despite this fact that the VAR models have performed well, there is a concern of what one should do in the event that new information is received after the model has been fitted. In this study, an approach is provided of updating the VAR model instead of fitting a new model whenever new information is received where the fitted VAR model is treated as the prior, new information or measurements as the likelihood to get an updated VAR model, the posterior, using the Bayesian Approach. Thus, updated VAR models of order one, two and three are developed after which generalization is done to a VAR model of order p. The performance of the existing VAR model is compared with the updated VAR model from which it is observed that the model performs well based on the fairly low values of root mean square error (RMSE) obtained. Furthermore, estimation of parameters is done using the joint estimation which estimates both the states and the parameters simultaneously. In the estimation, the estimated parameters converge to the true parameter value as time evolves. An application is considered where a penta-variate VAR(1) model is fitted using data for the contribution of five main sub-sectors of the agriculture sector to the Kenyan economy. The data considered was obtained from the Kenya National Bureau of Statistics (KNBS) on Statistical abstract reports from 2000 - 2021. The model was then updated and after comparing with the initial model, the model was found to perform well based on the lower values of the RMSE. From the study, it is then concluded that the updated Vector Autoregressive model performs well based on the Root Mean Square Error (RMSE). Finally, recommendations are also given regarding future work of updating other multivariate time series models to assimilate new information obtained after model fitting is done.

TABLE OF CONTENTS

List of Algorithms

List of Figures

LIST OF SYMBOLS AND ABBREVIATIONS

p	:	Order of Autoregressive component
ACF	:	Autocorrelation function
AIC	:	Akaike Information Criterion
AR	:	Autoregressive
ARMA	:	Autoregressive Moving average
ARIMA	:	Autoregressive Integrated Moving average
ACVF	:	Autocovariance function
BVAR	:	Bayesian Vector Autoregressive
FPE	:	Final Prediction Error
HQC	:	Hannan-Quinn Criterion
MSE	:	Mean Square Error
NNAR	:	Neuro Network Autoregressive
OLS	:	Ordinary Least Squares
RMSE	:	Root Mean Square Error
SARIMA	:	Seasonal Autoregressive Integrated Moving average
SC	:	Schwartz Criterion
VAR	:	Vector Autoregressive
VARMA	:	Vector Autoregressive Moving Average
VECM	:	Vector Error Correction Model
VMA	:	Vector Moving Average

CHAPTER ONE

INTRODUCTION

1.1 Background Information

The univariate Autoregressive Moving Averages (ARMA) models have been used favorably in forecasting of time series data which has resulted to researchers extending to the multivariate case since inclusion of more information increases the precision of forecasting and even identifying the correlation between the series [27]. Multivariate time series models are widely applied in a number of fields such as financial, economic, stock market, earth science among others. This is because they explain not only the serial dependence within each component series $\{X_{ti}\}$ but also interdependence between the different component series $\{X_{ti}\}$ and $\{X_{tj}\}$, $i \neq j$. The multivariate time series models includes the Vector Autoregressive (VAR) model, Vector Moving Average (VMA) and Vector Autoregressive Moving Average (VARMA) models as discussed in [5] among others. [32] earlier on carried out a study on multivariate Vector Autoregressive Moving Average (VARMA) models. However, the specification and estimation of such models was much difficult as compared to the univariate case. The success of the Box-Jenkins univariate modeling methodology in the 1970's triggered more research into strategies of modeling in the multivariate case. An example of such research is the work done by Tiao and Box in 1981 where they studied on multivariate time series modeling [41]. Other developments also done were the inclusion of cointegration in the multivariate modeling approach done by [11]. More recent work on modeling and forecasting with VARMA models is given by [27].

The Vector Autoregressive (VAR) models were developed by the macroeconometrician Christopher Sims in 1980 where the main aim was to model the joint dynamics and causal relations among a set of macroeconomic variables and they dominate time series econometrics modeling [37]. The joint dynamics includes how each variable in the model is explained by the past history of every variable and how the innovations may be correlated. A vector autoregressive (VAR) model provides forecasts which are superior to those obtained from the univariate time series models [45]. Traditionally, VAR models are widely much useful in describing the dynamic nature of most economic and financial time series [45]. However, recently the vector autoregressive models have gained much application in a wide range of disciplines such as Medicine, Epidemiology, Economics, Biology and Macroeconomics among others. Indeed, VAR models are one of the models mostly used for modelling multivariate time series data [15, 45].

Although the vector autoregressive models have been applied extensively, it should be noted that the presence of excessive parameters has been one of its main demerits [10, 19]. These has resulted to researchers developing the Bayesian Vector Autoregressive (BVAR) models where the model parameters are treated as random variables with prior probabilities. This as seen in the works of [9, 17, 19, 26, 42].

As indicated earlier on, the Vector Autoregressive models have attracted much application in diverse fields. They have got a number of advantages such as easy implementation where the parameters or coefficients are estimated by using the Ordinary Least Squares (OLS) method in each equation individually, testing for Granger causality where it is checked whether one or more variables have predictive impact to forecast the variable(s) of interest [45]. Even though the VAR models are believed to perform well due to the superior forecasts produced, there is a concern of what happens to the fitted VAR model when new information such as data is available after the fitting is done. It is in the view of this study that such new information should not be ignored since it may be crucial in affecting the existing model. In addition, it is quite involving to go through the entire process of fitting the model again when new information is obtained. Therefore, this study has developed an updated vector autoregressive model by using the Bayesian approach which is able to incorporate new information to update the existing VAR models where new information (considered as measurements) is used as the likelihood to update the model. The performance of the updated model is then compared with the performance of classical VAR model. Estimation of the parameters for the modified model is done using joint estimation where the state and the vector of parameters are augmented to form an extended state space and then they are estimated simultaneously.

1.2 Statement of the Problem

The Vector Autoregressive (VAR) Models have been applied extensively in many fields ranging from finance, economics among others. Many researchers have fitted the VAR models to the available data so as to come up with a model that explains the relationship among the variables involved. However, despite the fact that the existing VAR models produce superior forecasts, they can not incorporate new information when it is received. Therefore, there is a concern of what one should do in the event that new information is received given that information at time t is available. This study solves this problem by proposing an approach to use new information as likelihood and the fitted model as the prior and then update the initial model to come up with an updated model, the posterior. This study therefore considered the existing VAR model by treating it as prior and new measurement as likelihood to develop an updated VAR model that becomes the posterior using the Bayesian Approach. The study first develops an updated VAR model of order one, two and three after which it then generalizes to VAR model of order p. Afterwards, the study compares the performance of the updated model with the performance of some existing VAR models. Estimation of parameters is done using the joint estimation approach where the states and the vector of parameters are augmented to form an extended state space model.

1.3 Objectives of the Study

1.3.1 Main Objective

The main objective of this study is to assimilate new information into the vector autoregressive (VAR) model using the Bayesian approach.

1.3.2 Specific Objectives

The specific objectives of this study are:

(i) To develop an updated vector autoregressive model of order p using the Bayesian approach.

(ii) To compare the performance of the updated VAR with the classical VAR model.

(iii) To apply the updated VAR model in estimation of model parameters.

(iv) To apply the new model in a real life problem.

1.4 Justification of the Study

Vector Autoregressive (VAR) time series models have been widely applied in many areas. It is argued that they are the simplest multivariate time series models to use. Once the model is fitted, it is then used for prediction. However, as time evolves, new information of the variables in the model is obtained. Such information could have some impact on the model. Thus, instead of discarding the fitted model, there is need to factor in new information to the model using the Bayesian approach where the fitted VAR model is treated as the prior, new measurements as the likelihood to update the model and get the updated VAR model, the posterior.

1.5 Significance of the Study

Basically, the major role of the time series models is to perform forecasting. The VAR models have proven to perform well due to superior forecasts produced. However, as time goes on, more information of the variables in a time series model is obtained which may have an effect on the fitted model. Thus this study is significant in the sense that the inclusion of new information or measurements in the model makes the model to be updated so that it can be used to forecast precisely. In addition, the study is important to the consumers of time series data, especially econometricians, in that there is need to involve new measurements so as to get updated models. A good example is like the fluctuation of prices in the stock market where data is received on daily basis and so instead of discarding the fitted model, it is just updated to incorporate the new data.

1.6 Software

In this research, MATLAB R2017b and RStudio softwares were used whose versions are 9.3 and 1.4.1106 respectively.

1.7 Definition of Terms

1.7.1 V-variate Time Series

A v-variate time series is the $(v \times 1)$ vector time series written as $\{\mathbf{Y}_t\}$. More formally the vector is given by

$$\mathbf{Y}_t = \begin{pmatrix} Y_{t1} \\ Y_{t2} \\ \vdots \\ Y_{tv} \end{pmatrix} \tag{1.1}$$

for any time t.

1.7.2 Stationary V-variate Series

The v-variate series $\{\mathbf{Y}_t\}$ is (weakly) stationary if

(i) $\mu_Y(t)$ is independent of t and

(ii) $\mathbf{\Gamma}_Y(t, t-h)$ is independent of t for each h.

The mean of the series and the covariance matrix at lag h are given by

$$\mu = E(\mathbf{Y}_t) = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_v \end{pmatrix} \tag{1.2}$$

and

$$\mathbf{\Gamma}(h) = E[(\mathbf{Y}_t - \mu)(\mathbf{Y}_{t-h} - \mu)'] = \begin{pmatrix} \gamma_{11}(h) & \cdots & \gamma_{1v}(h) \\ \vdots & \cdots & \vdots \\ \gamma_{v1}(h) & \cdots & \gamma_{vv}(h) \end{pmatrix} \tag{1.3}$$

respectively, see [6].

1.7.3 Autocovariances, Cross lag covariances, Autocorrelations and Cross lag correlations

The auto-covariances of y_{it} for $i = 1, 2, \cdots, v$ are given by

$$\gamma_{ii}(h) = cov(y_{i,t}, y_{i,t-h}) \tag{1.4}$$

On the other hand, the cross lag covariances between y_{it} and y_{jt} for $i, j = 1, 2, \cdots, v$ are given by

$$\gamma_{ij}(h) = cov(y_{it}, y_{j,t-h}) \tag{1.5}$$

However, it is worth noting that

$$\begin{aligned} \gamma_{ij}(h) &= cov(y_{i,t}, y_{j,t-h}) \\ &\neq cov(y_{i,t}, y_{j,t+h}) \\ &= cov(y_{jt}, y_{i,t-h}) = \gamma_{ij}(-h) \end{aligned} \tag{1.6}$$

The autocorrelations of y_{it} for $i = 1, 2, \cdots, v$ are given by

$$\rho_{ii}(h) = corr(y_{i,t}, y_{i,t-h}) = \frac{\gamma_{ii}(h)}{\sqrt{\gamma_{ii}(0)\gamma_{ii}(0)}} = \frac{\gamma_{ii}(h)}{\gamma_{ii}(0)} \tag{1.7}$$

while the cross lag correlations between y_{it} and y_{jt} for $i, j = 1, 2, \cdots, v$ are given by

$$\rho_{ij}(h) = corr(x_{i,t}, x_{j,t-h}) = \frac{\gamma_{ij}(h)}{\sqrt{\gamma_{ii}(0)\gamma_{jj}(0)}} \tag{1.8}$$

The correlation matrix is given by

$$\mathbf{R}(h) = \begin{pmatrix} \rho_{11}(h) & \cdots & \rho_{1v}(h) \\ \vdots & \ddots & \vdots \\ \rho_{v1}(h) & \cdots & \rho_{vv}(h) \end{pmatrix} \tag{1.9}$$

see [6].

1.7.4 Multivariate White Noise

Multivariate white noise is the simplest multivariate time series. It is the building block from which variety of multivariate time series can be constructed. The v-variate series $\{\mathbf{u}_t\}$ is called white noise with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$ written as $\{\mathbf{u}_t\} \sim WN(\mathbf{0}, \mathbf{\Sigma})$, if $\{\mathbf{u}_t\}$ is stationary with mean vector $\mathbf{0}$ and covariance matrix function

$$\mathbf{\Gamma}(h) = \begin{cases} \mathbf{\Sigma}, & \text{if } h = 0 \\ \mathbf{0}, & \text{otherwise} \end{cases}$$

see [6]. The v-variate series $\{\mathbf{u}_t\}$ is independent and identically distributed (iid) noise with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$ written $\{\mathbf{u}_t\} \sim iid(\mathbf{0}, \mathbf{\Sigma})$ if the random vectors $\{\mathbf{u}_t\}$ are independent and identically distributed with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$.

1.7.5 Root Mean Square Error (RMSE)

The Root Mean Square Error (RMSE) is a statistical measure that is used to evaluate the quality of forecasts. It is given by

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{1.10}$$

where y_i is the actual measurements, $\hat{y}_i$ is the forecasted value and n is the number of data points.

1.8 Methods

In order to achieve the objectives, the methods used in the study are:

(i) Bayesian approach for Model development.

To achieve the first specific objective, Bayesian approach is used which employs the concept of using the prior and likelihood to get the posterior. Therefore, to update the Vector Autoregressive (VAR) model, the existing VAR model is considered to be the prior, the measurements or new information as the likelihood while the updated VAR model is the posterior. This is done in two steps namely: the prediction step and the update step.

(ii) Comparing the performance of the updated VAR with the classical VAR model.

This is as indicated in the second specific objective. In doing this, the Root Mean Square Error (RMSE) is used as a tool for checking adequacy of the model. The errors considered are between the existing VAR model and the predicted and between the existing VAR model and the updated.

(iii) Dual estimation precisely joint estimation to estimate the parameters which involves augmenting the states and the vector of parameters to form the extended state space model and then estimate the parameters to check if there is convergence to the true parameter value as time evolves. This will lead to achieving the third specific objective.

(iv) VAR model development technique which involves model specification, estimation of model parameters and model checking to develop a classical model using secondary data. In model specification, the order of the VAR model is chosen by considering Schwartz criterion while estimation of parameters is done by the least squares approach. Diagnostic checking involves checking the adequacy of the model which is by checking whether the residuals are white noise, normally distributed and uncorrelated. The fitted model will then be updated using the developed Algorithm. This will help us to achieve our fourth objective.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter highlights the existing Vector Autoregressive (VAR) model, that is, its origin, and how it has been applied in a wide range of fields. In addition, the Bayesian VAR models are also highlighted. These are discussed in Sections 2.2, 2.3 and 2.4.

2.2 Existing VAR Model

The Vector Autoregressive (VAR) models were developed by the macroeconometrician Christopher Sims in 1980 where the main aim was to model the joint dynamics and causal relations among a set of macroeconomic variables and dominate time series econometrics modeling [37, 45]. A v-variate vector autoregressive time series model of order p, VAR(p), is given by

$$\mathbf{Y}_t = \mathbf{A_1 Y_{t-1}} + \mathbf{A_2 Y_{t-2}} + \cdots + \mathbf{A_p Y_{t-p}} + \mathbf{u_t} \tag{2.1}$$

where $\mathbf{Y}_t$ is a $(v \times 1)$ vector of time series variables and $\mathbf{u_t}$ is a $(v \times 1)$ vector of white noise process. In matrix form, equation 2.1 can be written as

$$\begin{aligned}\begin{pmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{v,t} \end{pmatrix} &= \begin{pmatrix} a_{11,1} & a_{12,1} & \cdots & a_{1v,1} \\ a_{21,1} & a_{22,1} & \cdots & a_{2v,1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{v1,1} & a_{v2,1} & \cdots & a_{vv,1} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{v,t-1} \end{pmatrix} + \begin{pmatrix} a_{11,2} & a_{12,2} & \cdots & a_{1v,2} \\ a_{21,2} & a_{22,2} & \cdots & a_{2v,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{v1,2} & a_{v2,2} & \cdots & a_{vv,2} \end{pmatrix} \begin{pmatrix} y_{1,t-2} \\ y_{2,t-2} \\ \vdots \\ y_{v,t-2} \end{pmatrix} \\ &+ \cdots + \begin{pmatrix} a_{11,p} & a_{12,p} & \cdots & a_{1v,p} \\ a_{21,p} & a_{22,p} & \cdots & a_{2v,p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{v1,p} & a_{v2,p} & \cdots & a_{vv,p} \end{pmatrix} \begin{pmatrix} y_{1,t-p} \\ y_{2,t-p} \\ \vdots \\ y_{v,t-p} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \\ \vdots \\ u_{v,t} \end{pmatrix} \end{aligned} \tag{2.2}$$

However, it should be noted that the VAR models can further be classified into: the reduced form (Equation 2.1) and the structural VAR model given by

$$\mathbf{A_0 Y_t} = \mathbf{A_1 Y_{t-1}} + \mathbf{A_2 Y_{t-2}} + \cdots + \mathbf{A_p Y_{t-p}} + \mathbf{u_t} \tag{2.3}$$

where the the main diagonal terms of matrix A_0 have been scaled to one [28]. The structural shocks (error terms) in Equation 2.3 have zero mean and are uncorrelated. In the reduced form VAR model, each variable is a function of its own past and the past values of the other variables. On the other hand, structural form is used when the error terms are uncorrelated and that the variables can have a contemporaneous impact on other variables [38].

The identification or fitting of a ordinary VAR model involves model specification, estimation of model parameters and model checking to test whether the model is adequate. The order, p, of VAR is chosen which minimizes the Schwartz and Hannan-Quinn criteria as outlined by [28]. The Schwartz criterion is given by

$$SC(p) = \ln|\widehat{\Sigma}_u(p)| + \frac{\ln T}{T} pv^2$$

On the other hand, the Hannan-Quinn criterion is given by

$$HQ(p) = \ln|\widehat{\Sigma}_u(p)| + \frac{2\ln\ln T}{T} pv^2$$

where, for both criteria, $\widehat{\Sigma}_u$ is the estimated white noise covariance matrix, T is the sample size and v is the number of time series components. The criteria compares the residuals of the models and estimates the relative information loss of representing the original data using each of the model. In addition, the criteria weighs the quality of fit (covariance of residuals) against the complexity (number of free parameters) and therefore the model with least criterion value is considered [34].

The parameters of a fitted VAR model can be estimated by ordinary least squares estimation method under the assumptions that error term has mean of zero, the variables are stationary and no outliers. The developed model is then subjected to diagnostic checking for its adequacy and this involves checking whether the residuals are white noise, normally distributed and uncorrelated. Afterwards, the model is used to forecast which is the main function of the VAR models. Apart from forecasting, the VAR models can be used to give the dynamics that are predicted by the models in addition to estimating the model's parameters which involves Granger-causality statistics, impulse response function and forecast error decomposition as given in [28]. Granger-causality involves testing whether one variable is statistically significant when predicting another variable while impulse response function traces the dynamic path of variables in the system to shocks to other variables in the system. On the other hand, forecast error decomposition separates the forecast error variance into proportions attributed to each variable in the model which enables understanding of how much of an impact one variable has on another variable in the VAR model [28, 45].

2.3 Some Developed Vector Autoregressive Models

Vector Autoregressive (VAR) models have been extensively applied in a number of fields to study the relationship between the variables of interest. For instance, [35] studied on the causal relationship between crop production index (CPI) and permanent cropland (PCL) in Nigeria. They used time series data on CPI and PCL as the variables. The study used unrestricted Vector Autoregression (VAR) modeling techniques to develop the model. The time series data showed an upward trend and so differencing was applied to achieve stationarity. A VAR model of order 3, VAR(3), was chosen as best model that fitted the data. The model developed is given by

$$\begin{pmatrix} CPI_t \\ PCL_t \end{pmatrix} = \begin{pmatrix} -0.721 & 7.054 \\ 0.007 & 1.134 \end{pmatrix} \begin{pmatrix} CPI_{t-1} \\ PCL_{t-1} \end{pmatrix} + \begin{pmatrix} 0.574 & -6.25 \\ 0.003 & -0.021 \end{pmatrix} \begin{pmatrix} CPI_{t-2} \\ PCL_{t-2} \end{pmatrix} + \begin{pmatrix} -0.568 & 4.905 \\ -0.001 & -0.136 \end{pmatrix} \begin{pmatrix} CPI_{t-3} \\ PCL_{t-3} \end{pmatrix} + \begin{pmatrix} -12.02 \\ 0.478 \end{pmatrix} \tag{2.4}$$

In addition, it was concluded that Nigeria's CPI can be predicted by Nigeria's PCL and vice versa.

Dynamic Modeling and forecasting of data export of agricultural commodity by Vector Autoregressive model has been done as seen in [15]. The study determined the best model that can be used to describe the relationship among the data export value of Indonesia's agricultural commodities namely: coffee beans, cacao beans and tobacco using monthly data from the year 2007 to 2018. The study applied VAR(1), VAR(2), VAR(3), VAR(4) and VAR(5) models but trivariate VAR(2) model given by

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{pmatrix} = \begin{pmatrix} -0.1002 & 0.0467 & -0.7534 \\ 0.0082 & -0.8053 & -0.90327 \\ -0.0223 & -0.007 & -0.522 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{pmatrix} + \begin{pmatrix} 0.2713 & 0.0146 & -0.052 \\ 0.1983 & -0.296 & -0.096 \\ -0.033 & -0.0067 & -0.4043 \end{pmatrix} \begin{pmatrix} y_{1,t-2} \\ y_{2,t-2} \\ y_{3,t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} \tag{2.5}$$

was selected to be the best model based on Akaike Information Criterion with correction, Akaike Information Criterion, Schwarz Bayesian Criterion and Hannan - Quinn Information Criterion. Afterwards, the model was then used to forecast for the next 10 months.

[21] did a study to analyze the relationships between Finnish and global milk markets using the Vector Autoregressive (VAR) model. The study found that the greatest forecasting power for the milk price in Finland are the VAR models with combination of three or four variables namely: lagged price of milk in Finland, the price of oil, the world feed price and the quantity of milk produced. The

study used time series data from January 2007 to December 2016. Forecasting of oil prices was done from January 2017 to August 2019 and compared with the observed series from which it was concluded that the model did well.

Another application of the VAR models is given by [34] which is on prediction of gross domestic product using autoregressive models. They constructed a vector autoregressive model of order 4, VAR(4), model by selecting few macroeconomic indicators and predicted the Gross domestic product. The study relied on extensive database of historical economic data by the Federal Reserve Bank of St.Louis and found that the results from the model matched with historical data an implication that the model predicted consistently.

[30] determined the econometric connection between agriculture and gross domestic product (GDP) in Morocco using the VAR modeling approach. The study considered the macroeconomic variables: GDP per capita, agricultural GDP, investment rate, money supply, and trade openness and developed a VAR(2) model. From the study it was found out that there is presence of bidirectional Granger causality between agriculture and GDP, implying a feedback relationship, and on the other hand a unidirectional causal relationships involving the other macroeconomic variables used in the VAR model. In this study, we developed an updated multivariate vector autoregressive (VAR) time series model and illustrated estimation of the model parameters by dual estimation approach to check convergence of the model parameters to the true parameters.

The Growth Domestic Product of Ghana has been modeled using VAR models as seen in [2]. The study considered two more selected macroeconomic variables (inflation and real exchange rate) for the period 1980 to 2013 where the data were taken from the World Bank's World Development Indicators and Bank of Ghana. Co-integration test and vector error correction models (VECM) were used to examine dynamic relationships among the variables in both the long run and short run. It was found that there is co-integration between the the macroeconomic variables and GDP indicating long run relationship. A VECM model of order 3 was appropriately identified as a suitable model.

[24] forecasted the spread of COVID-19 infection based on the vector autoregression model. They combined the time series data for the new number of cases and the number of deaths to obtain a joint forecasting model. The study developed a VAR(29) model and applied the model to predict the number of new cases and deaths in UAE, Saudi Arabia and Kuwait using out of sample forecast from which it was found that the model achieved high level of accuracy.

A research on analyzing the relationship between two time series namely: global monthly oil price and global monthly gold price in dollars using Vector autoregressive model as seen in [1]. The study used monthly data from January 2015 to June 2019 and identified VAR(7), VAR(8) and VAR(10) models as possible models. However, based on the Mean Square Error (MSE), VAR(10) model was selected as the suitable model which was then used to forecast for the period June 2019 to June 2021.

[7] forecasted the dynamics of output for the Romanian economy using the Bayesian VAR model. The study estimated several versions of Bayesian VARs and compared with the OLS and unrestricted VAR model. From the study, it was confirmed that the BVAR model outperformed the standard models (OLS and the unrestricted VAR). The best BVAR model was then used for forecasting.

[16] carried out a study on modeling and forecasting of climatic parameters where they considered both the seasonal autoregressive moving average (SARIMA) and vector autoregressive (VAR) models approach. They developed univariate SARIMA

and multivariate VAR models for monthly maximum and minimum temperatures, humidity and cloud coverage in Bangladesh and performed forecasting using both models. In the study, it was found that the VAR(9) model which was developed gave better forecasts than the univariate SARIMA model based on the forecast accuracy measures considered.

Furthermore, [43] forecasted prices of coffee seeds using vector autoregressive (VAR) time series model in India. The VAR model was applied to model and forecast monthly wholesale price of clean coffee seeds in different coffee consuming centers namely: Bengaluru, Chennai and Hyderabad. After achieving stationarity, model selection was done based on the Akaike Information Criterion and VAR(2) model was selected. The model was also compared with univariate ARIMA models after which the study concluded that the VAR models fitted better that the ARIMA models based on the forecast accuracy measures. In addition, the study argues that when the ARIMA models are not available, then the VAR model can be used which makes use of the information available from other series when the series are cointegrated.

[13] investigated the effect of export and import on real economic growth of Ethiopia using the VAR model. The study developed a VAR(2) model using yearly data for the period 1982 to 2015 from the national Bank of the country. From the results of the VAR analysis, lagged variables of both export and import have significant contributions in predicting the economic growth of the country.

So far, it should be noted that in the above discussed fitted models, there is data that has been obtained since they were developed and such information or data could have an impact on the model. Therefore, such data or information needs to be incorporated in the model to get an updated model instead of discarding the model and fitting a new one which is the main objective of this study.

2.4 Bayesian Vector Autoregressive Models

Although the vector autoregressive models have been applied extensively, the presence of excessive parameters has been one of its main demerits and leads to unstable inference and inaccurate out-of-sample forecasts particularly for models with many variables [10, 19]. These has resulted to researchers developing the Bayesian Vector Autoregressive (BVAR) models where the model parameters are treated as random variables with prior probabilities. Recent research has shown that Bayesian vector autoregression is an appropriate tool for modelling large data sets [45]. Given the limited length of standard macroeconomic data sets relative to the large number of parameters available, Bayesian methods have become an increasingly popular way of dealing with the problem of too many parameters. As the ratio of variables to observations increases, the role of prior probabilities becomes increasingly important. The general idea is to use informative priors to shrink the unrestricted model towards a parsimonious naive benchmark, thereby reducing parameter uncertainty and improving forecast accuracy. An example is the shrinkage prior, proposed by Robert Litterman and subsequently developed by other researchers which is known as the "Minnesota prior" [23, 40]. The Minnesota prior captures widely held beliefs about the long-run properties of the data, properties that are not readily apparent in the short samples typically used for estimation. Bayes theorem then provides the optimal way of combining these two sources of information leading to sharper inference and more precise forecasts [23]. In particular, the Minnesota prior assumes that each variable follows a random walk process and therefore consists of a normal prior on a set of parameters with fixed and known covariance matrix [40].

In Bayesian inference, if the posterior probability distribution and the prior probability distribution belong to the same type of distribution, then this is referred to as conjugate prior distribution [19]. Assuming that the parameters are conjugate prior distributions, the parameters can be obtained. The assumption of a conjugate prior distribution is advantageous in that it can greatly reduce the amount of calculations and much studies have shown that this assumption is reliable in many cases [19]. The Minnesota prior model is one of the conjugate prior models which solves the problem of too many parameters in the VAR model under the conjugate prior distribution and improves the prediction accuracy of the model [19]. Based on this, much BVAR models have been developed. For instance, [19] used the BVAR based on Minnesota prior to study on Application of Bayesian Vector Autoregressive Model in Regional Economic Forecast. From the study it was found that the prediction error of the BVAR model is very small and the prediction ability is very satisfactory. [42] did a study on using the BVAR model to forecast the quarterly GDP in Singapore. The study found out that the BVAR model does forecasting accurately based on the out-of-sample forecasting done. [26] used the BVAR model to forecast household credit in Kenya using the Sims-Zha prior. In the study, the model parameters were treated as random variables and then prior probabilities assigned to them. The results from the BVAR were compared from those of ARIMA model and concluded that the BVAR outperformed the ARIMA model.

[9] did a study to forecast the UK economy with a medium-scale Bayesian vector autoregressive model and assessed the performance of the model in forecasting GDP growth and consumer price index (CPI) inflation in real time relative to forecasts from Central Organising Model for Projection Analysis and Scenario Simulation (COMPASS), the Bank of England's dynamic stochastic general equilibrium (DSGE) model and other benchmarks. The BVAR outperformed COMPASS when forecasting both gdp and its expenditure components. The study opted to

use Bayesian since the relatively large number of variables and the limited sample size available make classical estimation of an unrestricted VAR difficult.

Another work on the use of the Bayesian Vector Autoregressive model is given by [17] which is on forecasting Chinese inflation and output. The study developed several BVAR models to forecast the price inflation and output growth in China. From the study, it was found that models with shrinkage and model selection priors that restrict some VAR coefficients to be close to zero performed better than models with normal prior.

From this literature review, it is true that the VAR models have been shown to perform well based on the superior forecasts produced. However, of interest is what happens to the developed VAR model when new information is available. In this study, we formulate an updated Vector Autoregressive model by using Bayesian approach where the existing VAR model is treated as the prior and new measurements as the likelihood to get an updated VAR model. This will be achieved in two steps namely: the prediction step and the update step. Therefore, we develop an updated Vector Autoregressive (VAR) time series model of order one, two and three, after which it is then generalized to modified Vector Autoregressive (VAR) time series model of order p. Furthermore, performance of the updated vector autoregressive model is tested by comparing its performance with performance of some corresponding VAR models where the root mean square error (RMSE) is used as a measure of adequacy. In addition, the study considers estimation of parameters by means of dual estimation approach, precisely joint estimation, where the states and the parameters are estimated simultaneously. This will involve checking if there is convergence of the parameters to the true parameter values as time evolves.

CHAPTER THREE

FORMULATION OF THE UPDATED VAR MODEL AND ITS PERFORMANCE

3.1 Introduction

In this chapter we discuss model formulation of the updated vector autoregresive (VAR) model using the Bayesian approach. First, the study updates the Vector Autoregressive model of order 1, 2 and 3 after which it is then generalized to VAR model of order p. In doing this, the existing model is treated as the prior while new information which we refer to it as measurements is treated as the likelihood. Afterwards, the performance of the updated model is compared with the performance of the classical model as given in Section 3.3.

3.2 Updated Vector Autoregressive VAR Model

In this section, the updated Vector Autoregressive model is discussed. First, the updated VAR(1), VAR(2) and VAR(3) models are discussed after which the updated VAR(p) model is given.

3.2.1 Updated VAR(1)

A v-variate VAR model of order 1 is given by

$$Y_t = A_1 Y_{t-1} + u_t \qquad\qquad u_t \sim \mathcal{N}(0, Q) \tag{3.1}$$

Now, let the relation between Y_t, which is assumed to be the state at time t, and X_t, the measurements at time t, be given by

$$X_t = P_t Y_t + \eta_t \qquad\qquad \eta_t \sim \mathcal{N}(0, R) \tag{3.2}$$

where P is a matrix that may depend on time t and η_t is the measurement error which is white noise. Equation (3.1) is a transition equation giving transition from state t to state $t+1$ while Equation (3.2) is known as measurement equation. Equations (3.1) and (3.2) now form a system of models referred to as state-space models given by

$$\begin{aligned} Y_t &= A_1 Y_{t-1} + u_t \qquad && u_t \sim \mathcal{N}(0, Q) \\ X_t &= P_t Y_t + \eta_t && \eta_t \sim \mathcal{N}(0, R) \end{aligned} \tag{3.3}$$

where: Y_t is an $v \times 1$ state vector, X_t is a $n \times 1$ vector of measurement/observable variables, P_t is a $n \times v$ measurement matrix, A_1 is a $v \times v$ state transition matrix which may be time dependent, u_t is a $v \times 1$ vector of transition equation errors and η_t is a $n \times 1$ vector of measurement errors.

The goal is to get the estimate of the state Y_t given the observations X_t for the representation given by Equation (3.3). To achieve this, we do it in two steps, namely; the prediction and the update step. In the prediction step, we assume that the previous belief $p(Y_{t-1}|X_{t-1})$ is known and we wish to get $p(Y_t|X_{t-1})$ given by

$$p(Y_t|X_{t-1}) = \int p(Y_t, Y_{t-1}|X_{t-1}) dY_{t-1} \tag{3.4}$$

From conditional probability we have that Equation 3.4 can be written as

$$p(Y_t|X_{t-1}) = \int p(Y_t|Y_{t-1}, X_{t-1}) p(Y_{t-1}|X_{t-1}) dY_{t-1}$$

or

$$p(Y_t|X_{t-1}) = \int p(Y_t|Y_{t-1}) p(Y_{t-1}|X_{t-1}) dY_{t-1} \tag{3.5}$$

where $p(Y_t|Y_{t-1}, X_{t-1}) = p(Y_t|Y_{t-1})$, that is, under the assumption that the future state Y_t is independent of the past given the present Y_{t-1}. The probability density functions $p(Y_{t-1}|X_{t-1})$ and $p(Y_t|Y_{t-1})$ are Gaussian, where

$$\begin{aligned} p(Y_{t-1}|X_{t-1}) &= \mathcal{N}\left(E[Y_{t-1}|X_{t-1}], Var[Y_{t-1}|X_{t-1}]\right) \\ &= \mathcal{N}\left(\hat{Y}_{t-1|t-1}, S_{t-1|t-1}\right) \end{aligned} \tag{3.6}$$

and

$$
\begin{aligned}
p(Y_t|Y_{t-1}) &= \mathcal{N}\left(E[Y_t|Y_{t-1}], Var[Y_t|Y_{t-1}]\right) \\
&= \mathcal{N}\left(A_{1,t-1}Y_{t-1}, Q\right) \qquad (3.7)
\end{aligned}
$$

Substituting Equations (3.6) and (3.7) in the prediction posterior, Equation (3.5), we have

$$
p(Y_t|X_{t-1}) = \int \mathcal{N}\left(A_{1,t-1}Y_{t-1}, Q\right)\mathcal{N}\left(\hat{Y}_{t-1|t-1}, S_{t-1|t-1}\right) dY_{t-1}
$$

which can then be given as

$$
\begin{aligned}
p(Y_t|X_{t-1}) &= \mathcal{N}\left(A_{1,t-1}\hat{Y}_{t-1}, S_{t|t-1}\right) \\
&= \mathcal{N}\left(\hat{Y}_{t|t-1}, S_{t|t-1}\right) \qquad (3.8)
\end{aligned}
$$

where the predicted mean in Equation (3.8) is given by

$$
\begin{aligned}
\hat{Y}_{t|t-1} &= E[Y_t|X_{t-1}] \\
&= E[A_{1,t-1}Y_{t-1} + u_t|X_{t-1}] \\
&= E[A_{1,t-1}Y_{t-1}|X_{t-1}] + E[u_t|X_{t-1}] \qquad (3.9)
\end{aligned}
$$

But since u_t are independent and identically distributed and not dependent on X_{t-1}, then Equation (3.9) becomes

$$
\begin{aligned}
\hat{Y}_{t|t-1} &= A_{1,t-1}E[Y_{t-1}|X_{t-1}] + E[u_t] \\
&= A_{1,t-1}\hat{Y}_{t-1|t-1} \qquad (3.10)
\end{aligned}
$$

since $E(u_t) = 0$. On the other hand, the predicted covariance $S_{t|t-1}$ is given by

$$
\begin{aligned}
S_{t|t-1} &= Var[Y_t|X_{t-1}] \\
&= Var[A_{1,t-1}Y_{t-1} + u_t|X_{t-1}] \\
&= Var[A_{1,t-1}Y_{t-1}|X_{t-1}] + Var[u_t|X_{t-1}] \qquad (3.11)
\end{aligned}
$$

But since u_t is independent of X_{t-1}, then Equation (3.11) becomes

$$\begin{aligned} &= A_{1,t-1} Var[Y_{t-1}|X_{t-1}] A_{1,t-1}^T + Var[u_t] \\ &= A_{1,t-1} S_{t-1|t-1} A_{1,t-1}^T + Q \end{aligned} \tag{3.12}$$

where $Var(u_t) = Q$. In the update step, new measurement X_t is used to obtain the posterior $p(Y_t|X_t)$. From Bayes' theorem,

$$\begin{aligned} p(Y_t|X_t) &= \frac{p(X_t|Y_t)p(Y_t)}{p(X_t)} \\ &= \frac{p(X_t, X_{t-1}|Y_t)p(Y_t)}{p(X_t, X_{t-1})} \\ &= \frac{p(X_t|X_{t-1}, Y_t)p(X_{t-1}|Y_t)p(Y_t)}{p(X_t|X_{t-1})p(X_{t-1})} \end{aligned} \tag{3.13}$$

But

$$p(X_{t-1}|Y_t) = \frac{p(X_{t-1}, Y_t)}{p(Y_t)} = \frac{p(Y_t, X_{t-1})}{p(Y_t)} = \frac{p(Y_t|X_{t-1})p(X_{t-1})}{p(Y_t)} \tag{3.14}$$

and therefore substituting Equation (3.14) in Equation (3.13) we have

$$\begin{aligned} p(Y_t|X_t) &= \frac{p(X_t|X_{t-1}, Y_t)p(Y_t|X_{t-1})p(X_{t-1})p(Y_t)}{p(X_t|X_{t-1})p(X_{t-1})p(Y_t)} \\ &= \frac{p(X_t|X_{t-1}, Y_t)p(Y_t|X_{t-1})}{p(X_t|X_{t-1})} \\ &= \frac{p(X_t|Y_t)p(Y_t|X_{t-1})}{p(X_t|X_{t-1})} \end{aligned} \tag{3.15}$$

Furthermore,

$$\begin{aligned} p(X_t|X_{t-1}) &= \int p(X_t, Y_t|X_{t-1})dY_t = \int p(X_t|Y_t, X_{t-1})p(Y_t|X_{t-1})dY_t \\ &= \int p(X_t|Y_t)p(Y_t|X_{t-1})dY_t \end{aligned} \tag{3.16}$$

Substituting Equation (3.16) in Equation (3.15) we have

$$p(Y_t|X_t) = \frac{p(X_t|Y_t)p(Y_t|X_{t-1})}{\int p(X_t|Y_t)p(Y_t|X_{t-1})dY_t} \tag{3.17}$$

From the measurement equation we have that $p(X_t|Y_t) = \mathcal{N}[P_tY_t, R]$ and since $p(Y_t|X_{t-1}) = \mathcal{N}[\hat{Y}_{t|t-1}, S_{t|t-1}]$, then Equation (3.17) becomes

$$\begin{aligned} p(Y_t|X_t) &= \frac{p(X_t|Y_t)p(Y_t|X_{t-1})}{\int p(X_t|Y_t)p(Y_t|X_{t-1})dY_t} \\ &= \frac{\mathcal{N}[P_tY_t, R]\mathcal{N}[\hat{Y}_{t|t-1}, S_{t|t-1}]}{\int \mathcal{N}[P_tY_t, R]\mathcal{N}[\hat{Y}_{t|t-1}, S_{t|t-1}]dY_t} \qquad (3.18) \end{aligned}$$

In the numerator to Equation (3.18), we have that

$$\begin{aligned} \mathcal{N}[P_tY_t, R]\mathcal{N}[\hat{Y}_{t|t-1}, S_{t|t-1}] &= \frac{1}{\sqrt{det(2\pi R)}} e^{-\frac{1}{2}(X_t-P_tY_t)^T R^{-1}(X_t-P_tY_t)} \times \\ & \quad \frac{1}{\sqrt{det(2\pi S_{t|t-1})}} e^{-\frac{1}{2}(Y_t-\hat{Y}_{t|t-1})^T S^{-1}_{t|t-1}(Y_t-\hat{Y}_{t|t-1})} \\ &= \frac{1}{2\pi\sqrt{det(R)det(S_{t|t-1})}} e^{-\frac{1}{2}[M]} \qquad (3.19) \end{aligned}$$

where $M = (X_t - P_tY_t)^T R^{-1}(X_t - P_tY_t) + (Y_t - \hat{Y}_{t|t-1})^T S^{-1}_{t|t-1}(Y_t - \hat{Y}_{t|t-1})$. But from [29] page 699, M can be written as

$$\begin{aligned} M &= (X_t - P_tY_t)^T R^{-1}(X_t - P_tY_t) + (Y_t - \hat{Y}_{t|t-1})^T S^{-1}_{t|t-1}(Y_t - \hat{Y}_{t|t-1}) \\ &= (X_t - P_t\hat{Y}_{t|t-1})^T(R + P_tS_{t|t-1}P_t^T)^{-1}(X_t - P_t\hat{Y}_{t|t-1}) \\ & \quad +(Y_t - \hat{Y}_{t|t})^T(S_{t|t-1} + P_t^T R^{-1}P_t)(Y_t - \hat{Y}_{t|t}) \qquad (3.20) \end{aligned}$$

From which

$$det(R) \times det(S_{t|t-1}) = det(R + P_tS_{t|t-1}P_t^T) \times det(S_{t|t-1} + P_t^T R^{-1}P_t) \qquad (3.21)$$

Substituting Equations (3.20) and (3.21) in Equation (3.19) we have

$$\begin{aligned} \mathcal{N}[P_tY_t, R]\mathcal{N}[\hat{Y}_{t|t-1}, S_{t|t-1}] &= \frac{1}{\sqrt{det(2\pi(R + P_tS_{t|t-1}P_t^T))}} \times \\ & \quad e^{-\frac{1}{2}(X_t-P_t\hat{Y}_{t|t-1})^T(R+P_tS_{t|t-1}P_t^T)^{-1}(X_t-P_t\hat{Y}_{t|t-1})} \times \\ & \quad \frac{1}{\sqrt{det(2\pi(S_{t|t-1} + P_t^T R^{-1}P_t)^{-1})}} \times \\ & \quad e^{-\frac{1}{2}(Y_t-\hat{Y}_{t|t})^T(S_{t|t-1}+P_t^T R^{-1}P_t)(Y_t-\hat{Y}_{t|t})} \\ &= \mathcal{N}[P_t\hat{Y}_{t|t-1}, R + P_tS_{t|t-1}P_t^T] \times \\ & \quad \mathcal{N}[\hat{Y}_{t|t}, (S_{t|t-1} + P_t^T R^{-1}P_t)^{-1}] \qquad (3.22) \end{aligned}$$

The denominator in Equation (3.18) can be expressed as

$$\begin{aligned}\int \mathcal{N}[P_t Y_t, R]\mathcal{N}[\hat{Y}_{t|t-1}, S_{t|t-1}]dY_t &= \int \mathcal{N}[P_t\hat{Y}_{t|t-1}, R + P_t S_{t|t-1} P_t^T] \times \\ &\quad \mathcal{N}[\hat{Y}_{t|t}, (S_{t|t-1} + P_t^T R^{-1} P_t)^{-1}]dY_t \\ &= \mathcal{N}[P_t\hat{Y}_{t|t-1}, R + P_t S_{t|t-1} P_t^T] \times \\ &\quad \int \mathcal{N}[\hat{Y}_{t|t}, (S_{t|t-1} + P_t^T R^{-1} P_t)^{-1}]dY_t \\ &= \mathcal{N}[P_t\hat{Y}_{t|t-1}, R + P_t S_{t|t-1} P_t^T] \qquad (3.23)\end{aligned}$$

where $\int \mathcal{N}[\hat{Y}_{t|t}, (S_{t|t-1} + P_t^T R^{-1} P_t)^{-1}]dY_t = 1$. Therefore, the updated posterior is given by

$$\begin{aligned}p(Y_t|X_t) &= \frac{\mathcal{N}[P_t\hat{Y}_{t|t-1}, R + P_t S_{t|t-1} P_t^T]\mathcal{N}[\hat{Y}_{t|t}, (S_{t|t-1} + P_t^T R^{-1} P_t)^{-1}]}{\mathcal{N}[P_t\hat{Y}_{t|t-1}, R + P_t S_{t|t-1} P_t^T]} \\ &= \mathcal{N}[\hat{Y}_{t|t}, (S_{t|t-1} + P_t^T R^{-1} P_t)^{-1}] \qquad (3.24)\end{aligned}$$

Defining the covariance of the update as

$$\hat{S}_{t|t}^{-1} = S_{t|t-1}^{-1} + P_t^T R^{-1} P_t \qquad (3.25)$$

then we have that

$$p(Y_t|X_t) = \mathcal{N}[\hat{Y}_{t|t}, \hat{S}_{t|t}] \qquad (3.26)$$

By definition, see [29],

$$\hat{S}_{t|t}^{-1}\hat{Y}_{t|t} = S_{t|t-1}^{-1}\hat{Y}_{t|t-1} + P_t^T R^{-1} X_t \qquad (3.27)$$

Thus to obtain $\hat{S}_{t|t}$, we apply the Woodbury matrix identity given as

$$(E + FGH)^{-1} = E^{-1} - E^{-1}F(G^{-1} + HE^{-1}F)^{-1}HE^{-1} \qquad (3.28)$$

see [29] page 702. Hence, applying Equation (3.28) to Equation (3.25) we have that

$$\begin{aligned}
[\hat{S}_{t|t}^{-1}]^{-1} = \hat{S}_{t|t} &= (S_{t|t-1}^{-1} + P_t^T R^{-1} P_t)^{-1} \\
&= S_{t|t-1} - S_{t|t-1} P_t^T (R + P_t S_{t|t-1} P_t^T)^{-1} P_t S_{t|t-1} \\
&= (I - S_{t|t-1} P_t^T (R + P_t S_{t|t-1} P_t^T)^{-1} P_t) S_{t|t-1} \\
&= (I - K_t P_t) S_{t|t-1} \qquad (3.29)
\end{aligned}$$

where $K_t = \frac{S_{t|t-1} P_t^T}{R + P_t S_{t|t-1} P_t^T}$. To obtain the updated state, suppose that (3.27) is multiplied by $\hat{S}_{t|t}$ so that we have

$$\hat{S}_{t|t} \hat{S}_{t|t}^{-1} \hat{Y}_{t|t} = (I - K_t P_t) S_{t|t-1} [S_{t|t-1}^{-1} \hat{Y}_{t|t-1} + P_t^T R^{-1} X_t] \qquad (3.30)$$

Thus

$$\begin{aligned}
\hat{Y}_{t|t} &= (I - K_t P_t) \hat{Y}_{t|t-1} + (I - K_t P_t) S_{t|t-1} P_t^T R^{-1} X_t \\
&= \hat{Y}_{t|t-1} - K_t P_t \hat{Y}_{t|t-1} + S_{t|t-1} P_t^T R^{-1} X_t - K_t P_t S_{t|t-1} P_t^T R^{-1} X_t \\
&= \hat{Y}_{t|t-1} + (S_{t|t-1} P_t^T (R + P_t S_{t|t-1} P_t^T)^{-1} (R + P_t S_{t|t-1} P_t^T) R^{-1} \\
&\quad - K_t P_t S_{t|t-1} P_t^T R^{-1}) X_t - K_t P_t \hat{Y}_{t|t-1} \\
&= \hat{Y}_{t|t-1} + (K_t (I + P_t S_{t|t-1} P_t^T R^{-1}) - K_t P_t S_{t|t-1} P_t^T R^{-1}) X_t - K_t P_t \hat{Y}_{t|t-1} \\
&= \hat{Y}_{t|t-1} + (K_t + K_t P_t S_{t|t-1} P_t^T R^{-1} - K_t P_t S_{t|t-1} P_t^T R^{-1}) X_t - K_t P_t \hat{Y}_{t|t-1} \\
&= \hat{Y}_{t|t-1} + K_t (X_t - P_t \hat{Y}_{t|t-1}) \\
&= A_{1,t-1} \hat{Y}_{t|t-1} + K_t (X_t - P_t \hat{Y}_{t|t-1}) \qquad (3.31)
\end{aligned}$$

Therefore, the algorithm equations for the modified VAR(1) model are given as

$$\begin{aligned}
\hat{Y}_{t|t-1} &= A_{1,t-1} \hat{Y}_{t-1|t-1} \\
S_{t|t-1} &= A_{1,t-1} S_{t-1} A_{1,t-1}^T + Q \\
K_t &= \frac{S_{t|t-1} P_t^T}{P_t S_{t|t-1} P_t^T + R} \qquad (3.32) \\
\hat{Y}_{t|t} &= A_{1,t-1} \hat{Y}_{t|t-1} + K_t \left(X_t - P_t \hat{Y}_{t|t-1} \right) \\
S_{t|t} &= S_{t|t-1} - K_t P_t S_{t|t-1}
\end{aligned}$$

The equation

$$K_t = \frac{S_{t|t-1}P_t^T}{P_tS_{t|t-1}P_t^T + R} \tag{3.33}$$

is known as the gain while the term $\left(X_t - P_t\hat{Y}_{t|t-1}\right)$ is referred to as the innovation, the residual in the measurement, which is equivalent to the measurement noise. The summarised Algorithm is given by:

Algorithm 1 Algorithm for Modified VAR(1) model

1: Predict the State: $\hat{Y}_{t|t-1} = A_{1,t-1}\hat{Y}_{t-1|t-1}$ and Error covariance: $\hat{S}_{t|t-1} = A_{1,t-1}S_{t-1}A_{1,t-1}^T + Q$
2: Compute the gain: $K_t = \frac{S_{t|t-1}P_t^T}{P_tS_{t|t-1}P_t^T+R}$
3: Update the state: $\hat{Y}_{t|t} = A_{1,t-1}\hat{Y}_{t|t-1} + K_t\left(X_t - P_t\hat{Y}_{t|t-1}\right)$
4: Update the error covariance: $\hat{S}_{t|t} = S_{t|t-1} - K_tP_tS_{t|t-1}$

3.2.2 Updated VAR(2)

Now, suppose we consider the VAR(2) model given by

$$Y_t = A_1Y_{t-1} + A_2Y_{t-2} + u_t \qquad u_t \sim \mathcal{N}(0, Q)$$

with the corresponding measurement equation given by

$$X_t = P_tY_t + \eta_t \qquad \eta_t \sim \mathcal{N}(0, R)$$

Then, we have model given by

$$\begin{aligned} Y_t &= A_1Y_{t-1} + A_2Y_{t-2} + u_t & u_t &\sim \mathcal{N}(0, Q) \\ X_t &= P_tY_t + \eta_t & \eta_t &\sim \mathcal{N}(0, R) \end{aligned} \tag{3.34}$$

As indicated earlier on, the main focus is to get $p(Y_t|X_t)$. Similarly, we do this in two steps, the prediction and the update steps. In the prediction step, we need to get $p(Y_t|X_{t-1}, X_{t-2})$ on the assumption that $p(Y_{t-1}|X_{t-1}, X_{t-2})$ is known. But it can be assumed that

$$p(Y_t|X_{t-1}, X_{t-2}) = p(Y_t|X_{t-1}) \tag{3.35}$$

see [18]. Likewise, from [18] we can assume that

$$p(Y_{t-1}|X_{t-1}, X_{t-2}) = p(Y_{t-1}|X_{t-1})$$

Thus equation 3.35 can be expressed as

$$\begin{aligned} p(Y_t|X_{t-1}, X_{t-2}) &= p(Y_t|X_{t-1}) \\ &= \int p(Y_t|Y_{t-1})p(Y_{t-1}|X_{t-1})dY_{t-1} \\ &= \mathcal{N}(\hat{Y}_{t|t-1}, S_{t|t-1}) \end{aligned} \tag{3.36}$$

where

$$\begin{aligned} \hat{Y}_{t|t-1} &= E[Y_t|X_{t-1}, X_{t-2}] \\ &= E[Y_t|X_{t-1}] \\ &= E[A_1 Y_{t-1} + A_2 Y_{t-2} + u_t|X_{t-1}] \\ &= E[A_1 Y_{t-1}|X_{t-1}] + E[A_2 Y_{t-2}|X_{t-1}] + E[u_t|X_{t-1}] \\ &= A_1 E[Y_{t-1}|X_{t-1}] + A_2 E[Y_{t-2}|X_{t-1}] + E[u_t] \\ &= A_{1,t-1}\hat{Y}_{t-1|t-1} + A_{2,t-2}\hat{Y}_{t-2|t-2} \end{aligned} \tag{3.37}$$

and

$$\begin{aligned} S_{t|t-1} &= Var(Y_t|X_{t-1}) \\ &= Var[A_1 Y_{t-1}|A_2 Y_{t-1} + u_t|X_{t-1}] \\ &= Var[A_1 Y_{t-1}|X_{t-1}] + Var[A_2 Y_{t-2}|X_{t-1}] + Var[u_t|X_{t-1}] \\ &= A_1 Var[Y_{t-1}|X_{t-1}]A_1^T + A_2 Var[Y_{t-2}|X_{t-1}]A_2^T + Var[u_t] \\ &= A_{1,t-1}S_{t-1|t-1}A_{1,t-1}^T + A_{2,t-2}S_{t-2|t-2}A_{2,t-2}^T + Q \end{aligned} \tag{3.38}$$

Therefore, the prediction equations for the VAR(2) model are

$$\hat{Y}_{t|t-1} = A_{1,t-1}\hat{Y}_{t-1|t-1} + A_{2,t-2}\hat{Y}_{t-2|t-2} \tag{3.39}$$

and

$$S_{t|t-1} = A_{1,t-1}S_{t-1|t-1}A_{1,t-1}^T + A_{2,t-2}S_{t-2|t-2}A_{2,t-2}^T + Q \tag{3.40}$$

respectively.

In the update step we need $p(Y_t|X_t)$, that is the new state after getting new information. Previously, it has been obtained that

$$\begin{aligned} p(Y_t|X_t) &= \frac{p(X_t|Y_t)p(Y_t)}{p(X_t)} \\ &= \frac{p(X_t|Y_t)p(Y_t|X_{t-1})}{\int p(X_t|Y_t)p(Y_t|X_{t-1})dY_t} \end{aligned}$$

which after simplification gives

$$p(Y_t|X_t) = \mathcal{N}[\hat{Y}_{t|t}, (S_{t|t-1} + P_t^T R^{-1} P_t)^{-1}]$$

Thus the updated state and the error covariance for the VAR(2) is given by

$$\hat{Y}_{t|t} = A_{1,t-1}\hat{Y}_{t|t-1} + K_t(X_t - P_t\hat{Y}_{t|t-1})$$

and

$$S_{t|t} = S_{t|t-1} - K_t P_t S_{t|t-1}$$

respectively where $K_t = \frac{S_{t|t-1}P_t^T}{P_t S_{t|t-1} P_t^T + R}$ is the gain.

3.2.3 Updated VAR(3)

Extending to VAR(3) model, when augmented with measurement equation, we have the model given by

$$\begin{aligned} Y_t = A_1Y_{t-1} + A_2Y_{t-2} + A_3Y_{t-3} + u_t \qquad & u_t \sim \mathcal{N}(0, Q) \\ X_t = P_tY_t + \eta_t \qquad & \eta_t \sim \mathcal{N}(0, R) \end{aligned}$$

where the predicted state and error covariance are given by

$$\hat{Y}_{t|t-1} = A_{1,t-1}\hat{Y}_{t-1|t-1} + A_{2,t-2}\hat{Y}_{t-2|t-2} + A_{3,t-3}\hat{Y}_{t-3|t-3} \tag{3.41}$$

and

$$S_{t|t-1} = A_{1,t-1}S_{t-1|t-1}A_{1,t-1}^T + A_{2,t-2}S_{t-2|t-2}A_{2,t-2}^T + A_{3,t-3}S_{t-3|t-3}A_{3,t-3}^T + Q \quad (3.42)$$

respectively. On the other hand, the updated state and the error covariance for the VAR(3) are

$$\hat{Y}_{t|t} = A_{1,t-1}\hat{Y}_{t|t-1} + K_t(X_t - P_t\hat{Y}_{t|t-1})$$

and

$$S_{t|t} = S_{t|t-1} - K_tP_tS_{t|t-1}$$

respectively.

3.2.4 Generalization to the Updated VAR(p) model

Having obtained the algorithm for the updated VAR(1), VAR(2) and VAR(3) models, then the same approach can be applied to a vector autoregressive model of order p, VAR(p). Therefore it can be generalized that the updated vector autoregressive model of order p, VAR(p) model, is

$$\hat{Y}_{t|t} = A_{1,t-1}\hat{Y}_{t|t-1} + K_t\left(X_t - P_t\hat{Y}_{t|t-1}\right)$$

and the corresponding covariance is

$$\hat{S}_{t|t} = S_{t|t-1} - K_tP_tS_{t|t-1} \quad (3.43)$$

This can be used to update the existing vector autoregressive model given the new information which is considered as the likelihood. Therefore, the algorithm for the updated generalized vector autoregressive model of order p is

Algorithm 2 Algorithm for Generalized updated VAR(p) model

1: Predict the state: $\hat{Y}_{t|t-1} = A_{1,t-1}\hat{Y}_{t-1|t-1} + \cdots + A_{p,t-p}\hat{Y}_{t-p|t-p}$
2: Predict the error covariance: $\hat{S}_{t|t-1} = A_{1,t-1}S_{t-1}A^T_{1,t-1} + \cdots + A_{p,t-p}S_{t-p}A^T_{p,t-p} + Q$
3: Compute the gain: $K_t = \frac{S_{t|t-1}P_t^T}{P_tS_{t|t-1}P_t^T + R}$
4: Update the state: $\hat{Y}_{t|t} = A_{1,t-1}\hat{Y}_{t|t-1} + K_t\left(X_t - P_t\hat{Y}_{t|t-1}\right)$
5: Update the error covariance: $\hat{S}_{t|t} = S_{t|t-1} - K_tP_tS_{t|t-1}$

3.3 Comparing Performance of the Updated Model with Classical Model

This section gives the discussion of the results on comparing performance of the updated model with that of the classical model. We begin with a 1×1 vector, that is, one-dimension which we consider to be the simplest and then later on proceed to dimensions two, three and then five respectively.

3.3.1 Case I: One Dimension

Here we begin by considering the model in scalar form up to lag 1. If the model is in one dimension, then A_1 is a scalar. We set $A_1 = 0.9999$, $P_t = 1$, $Q = 0.001$, $R = 0.001$ and $S_0 = 0.001$ so that the state space model becomes

$$\begin{aligned} Y_t &= 0.9999Y_{t-1} + u_t \\ X_t &= Y_t + \eta_t \end{aligned} \tag{3.44}$$

With this values and using Algorithm 2, we have the plot in Figure 3.1 where the first subplot represents the output for AR(1), modified AR(1) estimate and modified AR(1) prediction, denoted by the blue line, red line and the yellow line, respectively. The second subplot represents the RMSE in the estimate and prediction, denoted by the blue and the red lines, respectively. In Figure 3.1 subplot one, the AR(1), modified AR(1) estimate and modified AR(1) prediction are seen to move together as time evolves. On the other hand, subplot two shows that errors between AR(1) and the modified AR(1) estimate are less. Furthermore, the errors

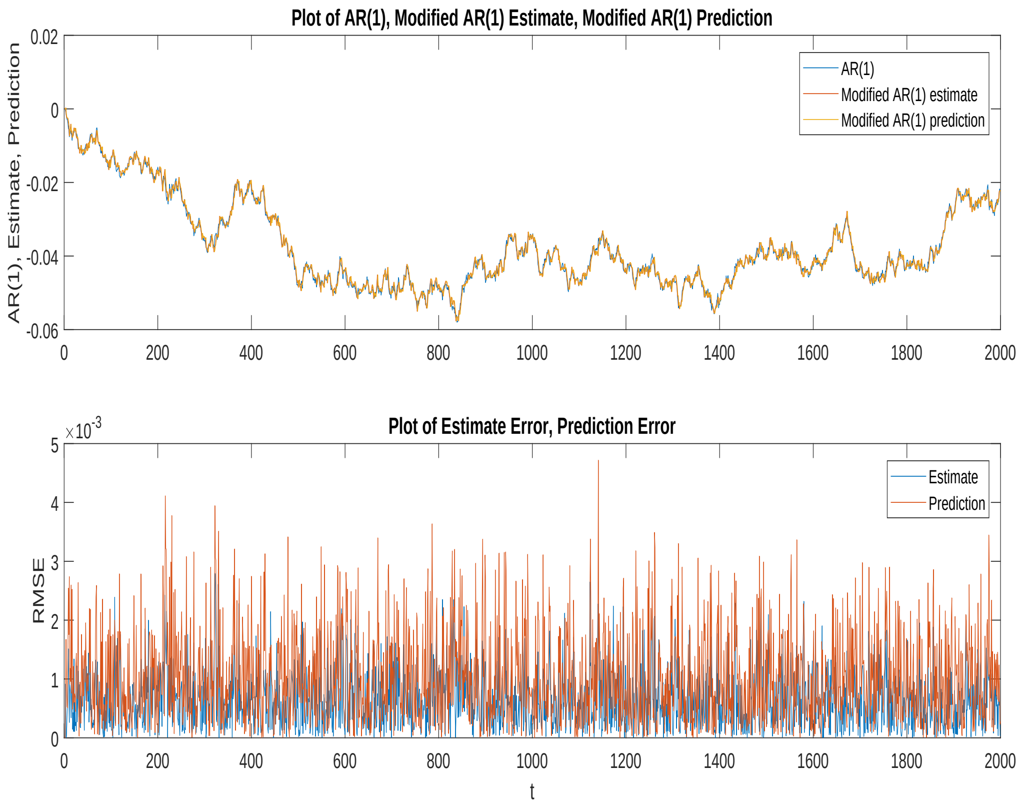


Figure 3.1: Univariate modified AR(1). Subplot one gives the comparison of the AR(1), modified AR(1) estimate and modified AR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot gives the errors between AR(1) and the modified AR(1) and between AR(1) and the modified AR(1) prediction.

between AR(1) and the modified AR(1) prediction are as well low. This indicates that the updated model performs well due to the small values of RMSE obtained in the estimate and in the prediction.

In the case of lag 2, the model becomes

$$\begin{aligned} Y_t &= A_1 Y_{t-1} + A_2 Y_{t-2} + u_t \\ X_t &= P_t Y_t + \eta_t \end{aligned} \tag{3.45}$$

where both u_t and η_t are white noise processes with covariance Q and R respectively as earlier on assumed. Suppose, under lag 2, that the model is

$$\begin{aligned} Y_t &= 0.5Y_{t-1} + 0.5Y_{t-2} + u_t \\ X_t &= 1Y_t + \eta_t \end{aligned} \tag{3.46}$$

where $A_1 = 0.5$, $A_2 = 0.5$, $P_t = 1$, $Q = 0.001$, $R = 0.001$ and $S_0 = 0$. Then the plot in Figure 3.2 is obtained.

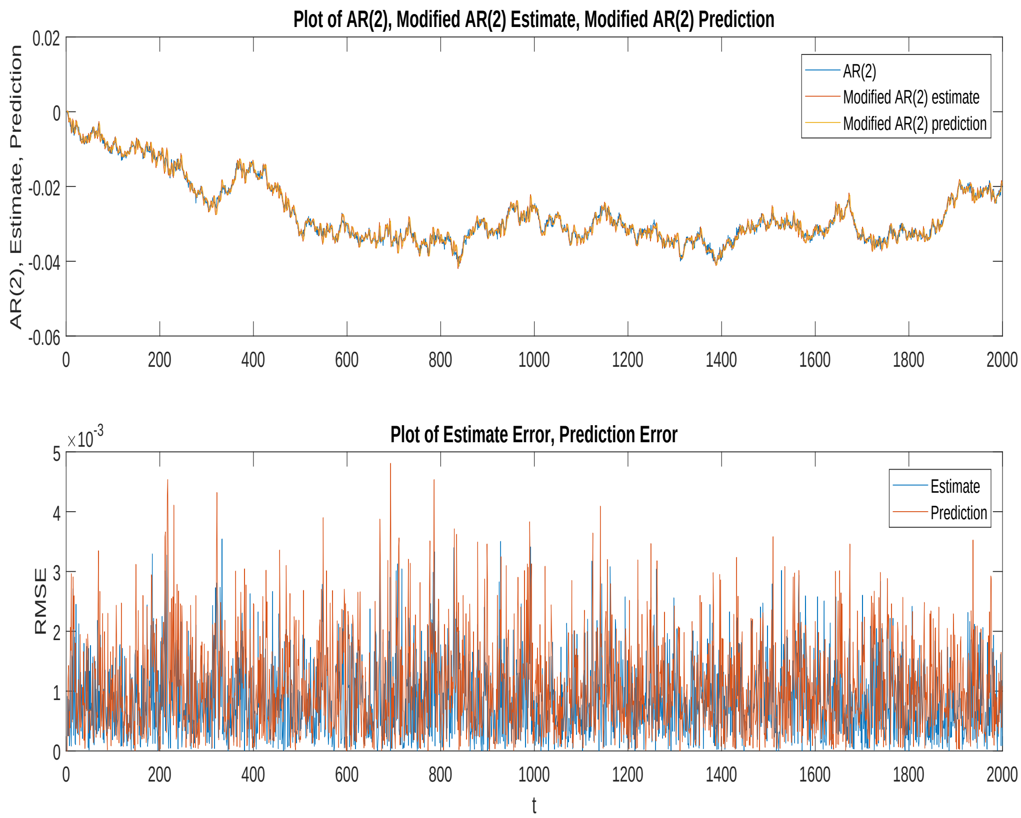


Figure 3.2: Univariate modified AR(2). In the first subplot, comparison of the AR(2), modified AR(2) estimate and modified AR(2) prediction, denoted by the blue line, red line and the yellow line, respectively is given while the second subplot gives the errors between AR(2) and the modified AR(2) and between AR(1) and the modified AR(1) prediction.

From Figure 3.2 subplot one, it can be observed that the output for AR(2), modified AR(2) estimate and modified AR(2) prediction, almost lie together. In addition, the errors between AR(2) and the modified AR(2) estimate and between

AR(2) and the modified AR(2) prediction are small an indication that the model performs well as seen in subplot two.

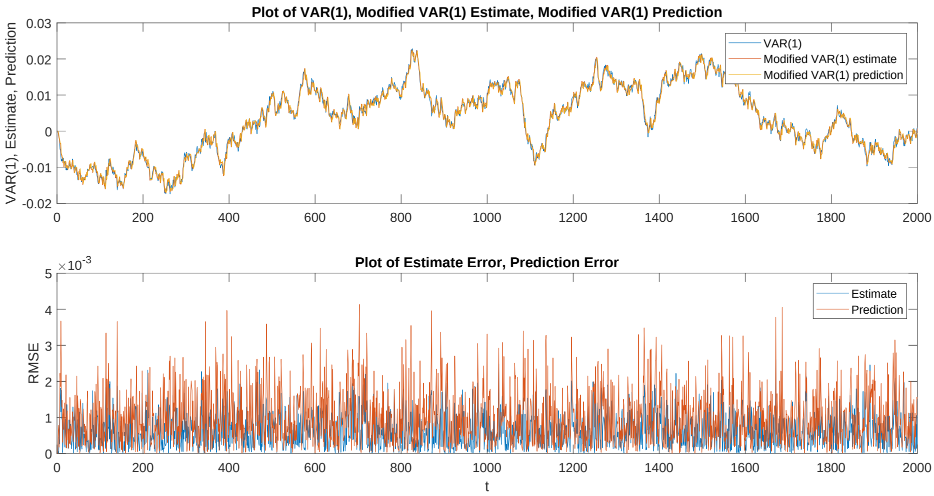
3.3.2 Case II: Two Dimensions

In two dimensions, consider the model given by

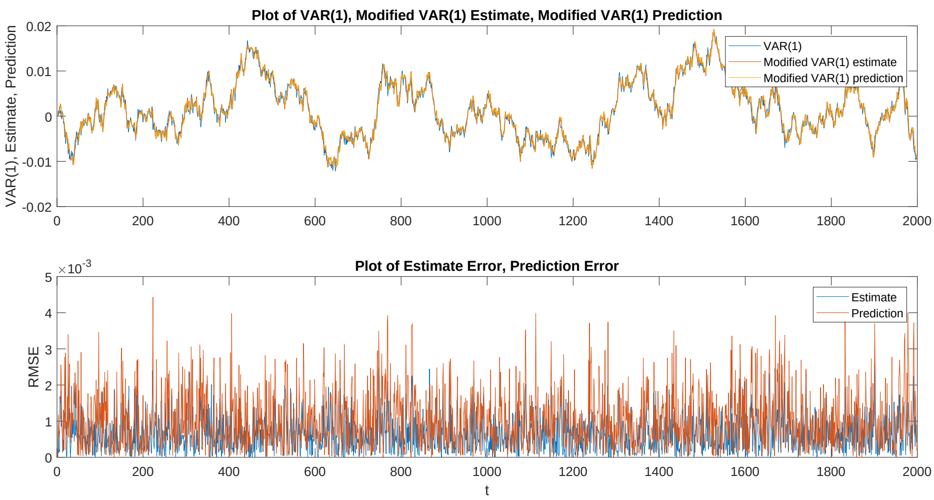
$$\begin{aligned} \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} &= \begin{pmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} \\ \begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} + \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \end{pmatrix} \end{aligned} \tag{3.47}$$

where setting $A_1 = \begin{pmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{pmatrix}$, $P_t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $Q = \begin{pmatrix} 0.001 & 0 \\ 0 & 0.001 \end{pmatrix}$, $R = \begin{pmatrix} 0.001 & 0 \\ 0 & 0.001 \end{pmatrix}$, $S_0 = \begin{pmatrix} 0.001 & 0 \\ 0 & 0.001 \end{pmatrix}$ and using MATLAB, we have the plots in Figure 3.3 where (a) and (b) represent the first and second variable respectively. The first subplot in (a) and (b) represents the output for VAR(1), modified VAR(1) estimate and modified VAR(1) prediction denoted by the blue line, red line and the yellow line, respectively. The second subplot in (a) and (b) represents the RMSE in the estimate and prediction denoted by the blue and the red lines, respectively.

From Figure 3.3 it can be observed that the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction move together as time grows. In addition, the errors between VAR(1) and the modified VAR(1) estimate and between VAR(1) and the modified VAR(1) prediction are small implying that the model can be considered adequate.



(a)



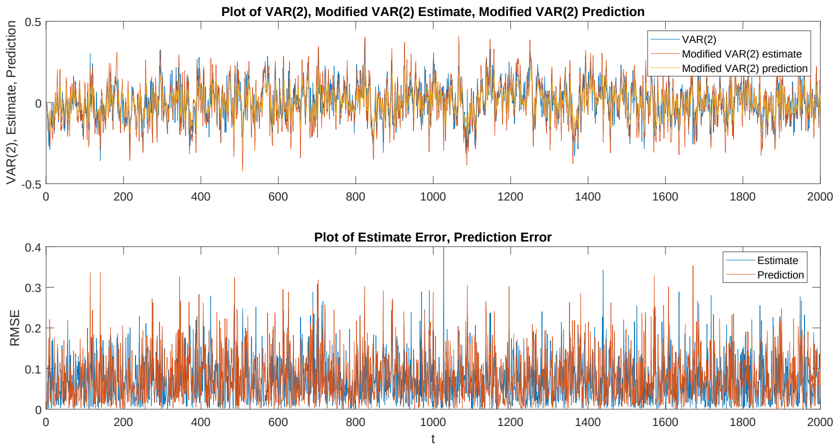
(b)

Figure 3.3: Bivariate modified VAR(1). The first subplot in (a) and (b) gives comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot in (a) and (b) gives the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction.

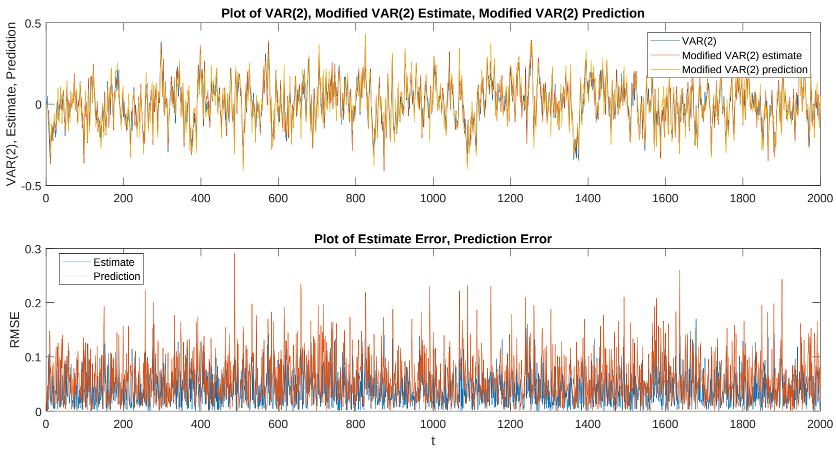
Furthermore, consider the Bivariate VAR(2) model given by

$$y_t = \nu + \begin{pmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{pmatrix} y_{t-1} + \begin{pmatrix} 0 & 0 \\ 0.25 & 0 \end{pmatrix} y_{t-2} + u_t \tag{3.48}$$

where it is assumed that $\Sigma_u = \begin{pmatrix} 0.09 & 0 \\ 0 & 0.04 \end{pmatrix}$ and ν is assumed to be a null matrix, see [28]. We use the model in Equation 3.48 to test the performance of the updated model given by Algorithm 2. Setting $A_1 = \begin{pmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & 0 \\ 0.25 & 0 \end{pmatrix}$, $P_t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $Q = \Sigma_u = \begin{pmatrix} 0.09 & 0 \\ 0 & 0.04 \end{pmatrix}$, $R = \begin{pmatrix} 0.09 & 0 \\ 0 & 0.04 \end{pmatrix}$ and $S_0 = \begin{pmatrix} 0.09 & 0 \\ 0 & 0.04 \end{pmatrix}$ we obtain the output in Figure 3.4 where (a) is for the first variable and (b) is for the second variable. From Figure 3.4, it can be observed that the updated model



(a)



(b)

Figure 3.4: Bivariate modified VAR(2). Here, subplot one in (a) and (b) shows the comparison of the VAR(2), modified VAR(2) estimate and modified VAR(2) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot in (a) and (b) shows the errors between VAR(2) and the modified VAR(2) and between VAR(2) and the modified VAR(2) prediction.

performs well due to the fairly small values of RMSE obtained.

3.3.3 Case III: Three Dimensions

In three dimensions, then A_1 and P_t are 3×3 matrices. Let the state space model be given by

$$\begin{aligned} \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{pmatrix} &= \begin{pmatrix} 0.19 & 0.38 & -0.74 \\ 0.38 & -0.21 & -0.14 \\ 0.01 & 0.05 & 0.99 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{pmatrix} \\ \begin{pmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{pmatrix} + \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \end{pmatrix} \end{aligned} \tag{3.49}$$

Upon setting $Q = \begin{pmatrix} 0.001 & 0 & 0 \\ 0 & 0.001 & 0 \\ 0 & 0 & 0.001 \end{pmatrix}$, $R = \begin{pmatrix} 0.001 & 0 & 0 \\ 0 & 0.001 & 0 \\ 0 & 0 & 0.001 \end{pmatrix}$ and $S_0 = \begin{pmatrix} 0.001 & 0 & 0 \\ 0 & 0.001 & 0 \\ 0 & 0 & 0.001 \end{pmatrix}$, we have the plots in Figures 3.5 - 3.7 which represent the first, second and the third variables respectively.

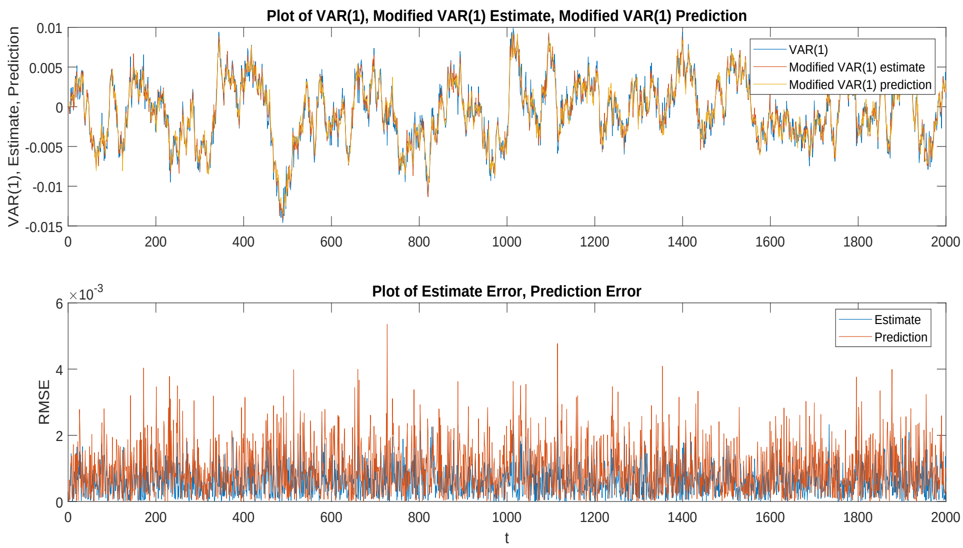


Figure 3.5: First Trivariate VAR(1) - Variable 1. The first subplot gives comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot shows the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction for variable 1.

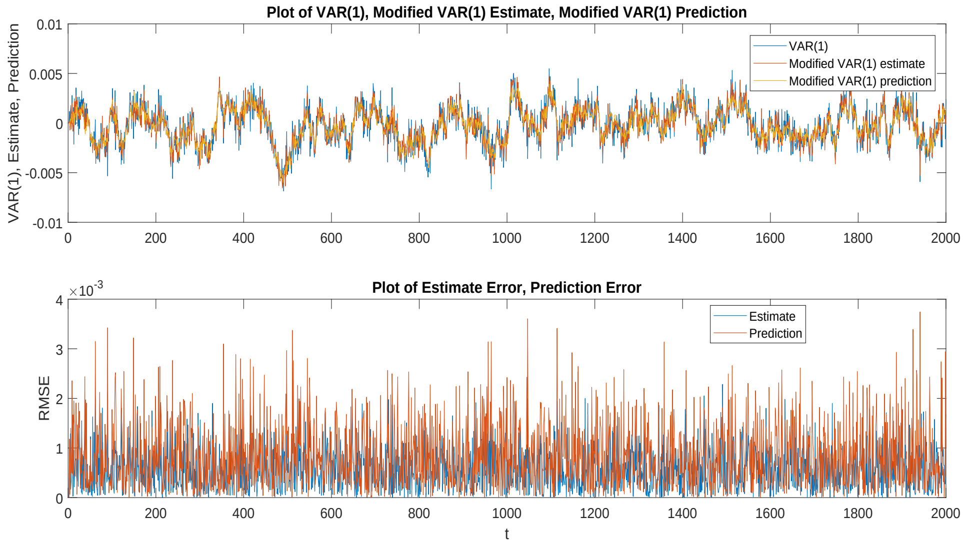


Figure 3.6: First Trivariate VAR(1) - Variable 2. Subplot one gives comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot gives the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction in the second variable.

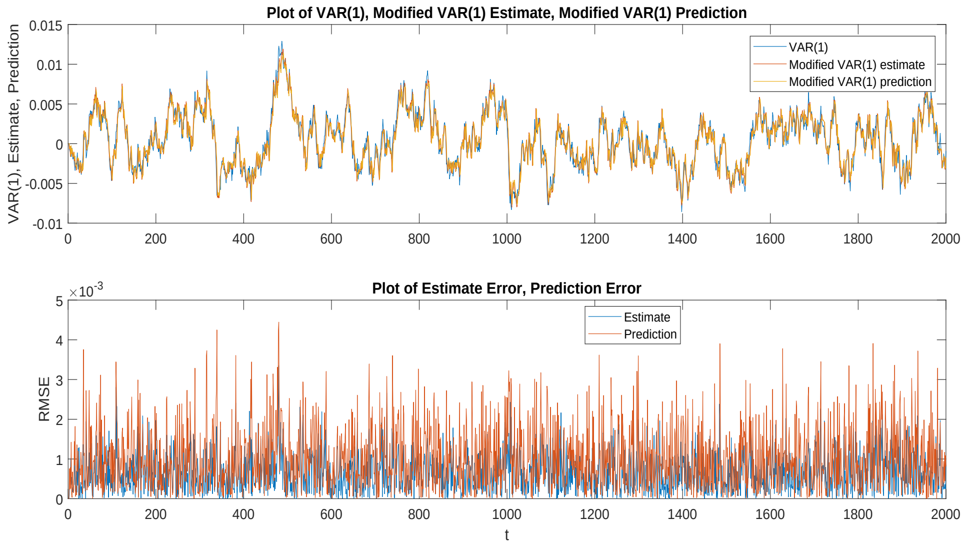


Figure 3.7: First Trivariate VAR(1) - Variable 3. In the first subplot, we have the comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot displays the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction in the third variable.

The plots in Figures 3.5 - 3.7 give small values of the RMSE between VAR(1) and the modified VAR(1) estimate and between VAR(1) and the modified VAR(1) prediction for each variable in the model implying good model performance.

Suppose we consider the tri-variate VAR(1) model given in [28] where

$$y_t = \nu + \begin{pmatrix} 0.5 & 0 & 0 \\ 0.1 & 0.1 & 0.3 \\ 0 & 0.2 & 0.3 \end{pmatrix} y_{t-1} + u_t \tag{3.50}$$

where we assume ν is a null matrix, $\Sigma_u = Q = \begin{pmatrix} 2.25 & 0 & 0 \\ 0 & 1 & 0.5 \\ 0 & 0.5 & 0.74 \end{pmatrix}$, $R = \begin{pmatrix} 2.25 & 0 & 0 \\ 0 & 1 & 0.5 \\ 0 & 0.5 & 0.74 \end{pmatrix}$ and $S_0 = \begin{pmatrix} 2.25 & 0 & 0 \\ 0 & 1 & 0.5 \\ 0 & 0.5 & 0.74 \end{pmatrix}$. We test the performance of the updated model under the model given by Equation 3.50 whose output is given in Figures 3.8 - 3.10 for the first, second and third variables respectively.

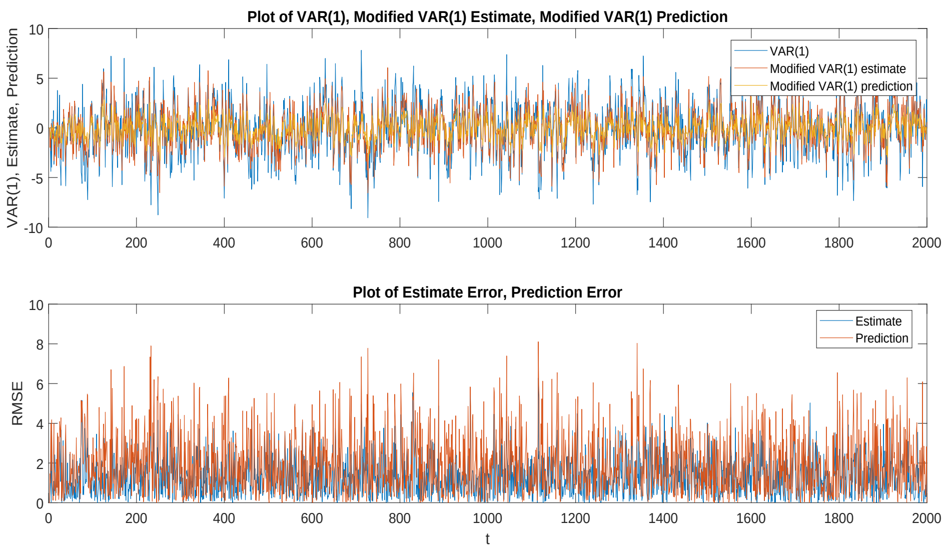


Figure 3.8: Second Trivariate VAR(1) - Variable 1. Subplot one gives comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot gives the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction in variable 1.

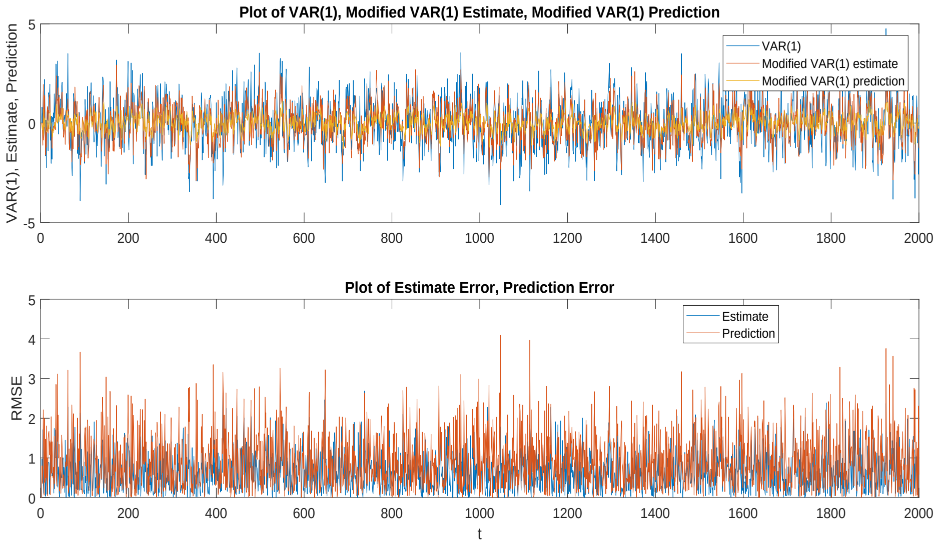


Figure 3.9: Second Trivariate VAR(1) - Variable 2. The first subplot gives comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot shows the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction for variable 2.

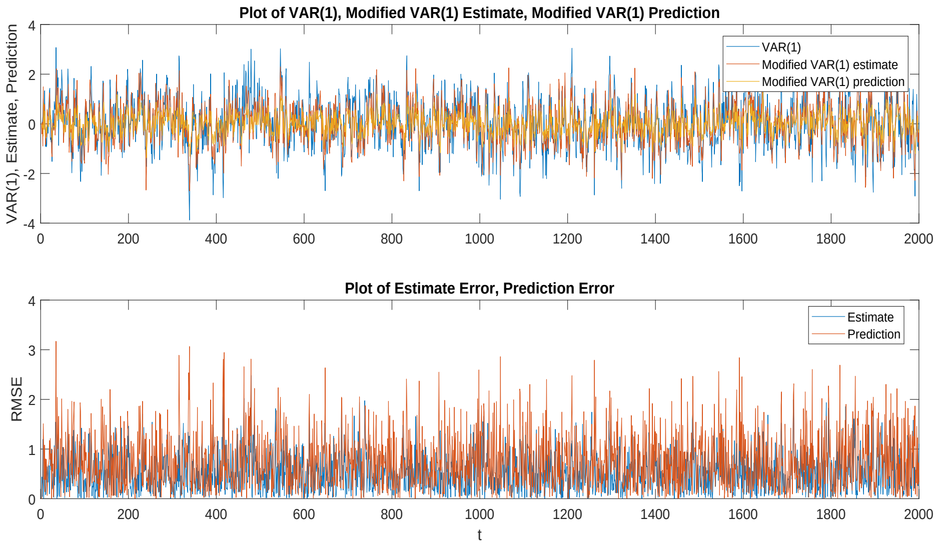


Figure 3.10: Second Trivariate VAR(1) - Variable 3. Here, the first subplot gives comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot shows the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction for variable 3.

From Figures 3.8 - 3.10 it can be seen that the updated model performs well due to the small RMSE values obtained between VAR(1) and the modified VAR(1) estimate and between VAR(1) and the modified VAR(1) prediction in each of the variables for the model.

Now, assume we consider the tri-variate VAR(2) model given by

$$y_t = \nu + \begin{pmatrix} 0.5 & 0 & 0 \\ 0.1 & 0.1 & 0.3 \\ 0 & 0.2 & 0.3 \end{pmatrix} y_{t-1} + \begin{pmatrix} 0.5 & 0 & 0 \\ 0.1 & 0.1 & 0.3 \\ 0 & 0.2 & 0.3 \end{pmatrix} y_{t-2} + u_t \tag{3.51}$$

where we assume ν is a null matrix, $\Sigma_u = Q = \begin{pmatrix} 2.25 & 0 & 0 \\ 0 & 1 & 0.5 \\ 0 & 0.5 & 0.74 \end{pmatrix}$, $R = \begin{pmatrix} 2.25 & 0 & 0 \\ 0 & 1 & 0.5 \\ 0 & 0.5 & 0.74 \end{pmatrix}$ and $S_0 = \begin{pmatrix} 2.25 & 0 & 0 \\ 0 & 1 & 0.5 \\ 0 & 0.5 & 0.74 \end{pmatrix}$. From Equation 3.50, we assume that the parameters for the matrix at lag 1 are the same as parameters for the matrix at lag 2. Running Algorithm 2 in MATLAB under Equation 3.51, we have the output in Figures 3.11 - 3.13.

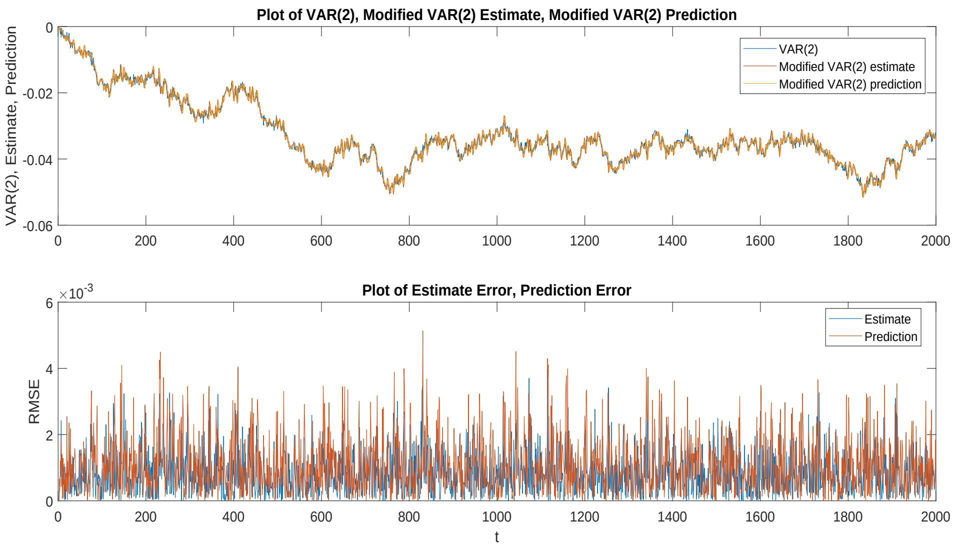


Figure 3.11: Trivariate VAR(2) - Variable 1. The first subplot gives comparison of the performance of classical VAR(2) and the updated VAR(2) while the second subplot gives the errors between classical VAR(2) and updated model.

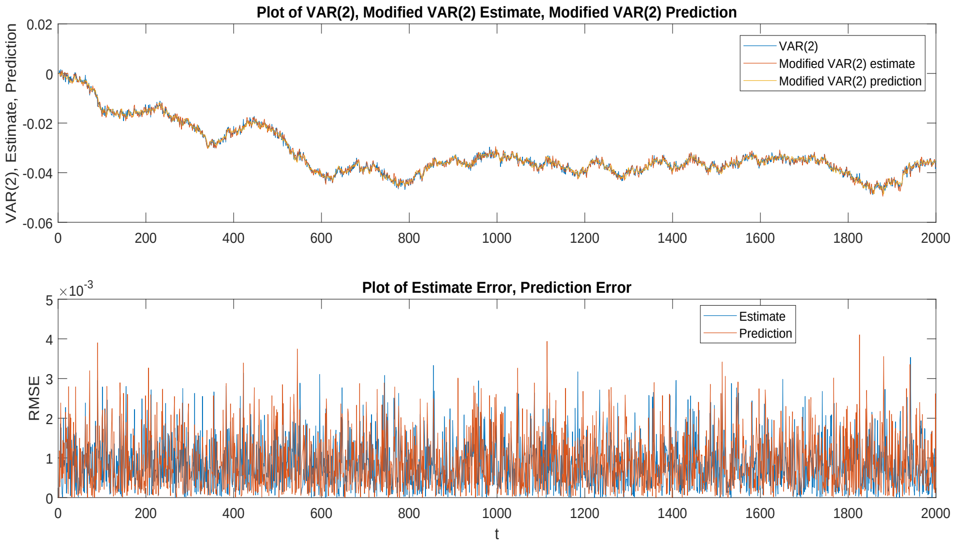


Figure 3.12: Trivariate VAR(2) - Variable 2. Subplot one gives comparison of the VAR(2), modified VAR(2) estimate and modified VAR(2) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot the errors between VAR(2) and the modified VAR(2) and between VAR(2) and the modified VAR(2) prediction as time grows.

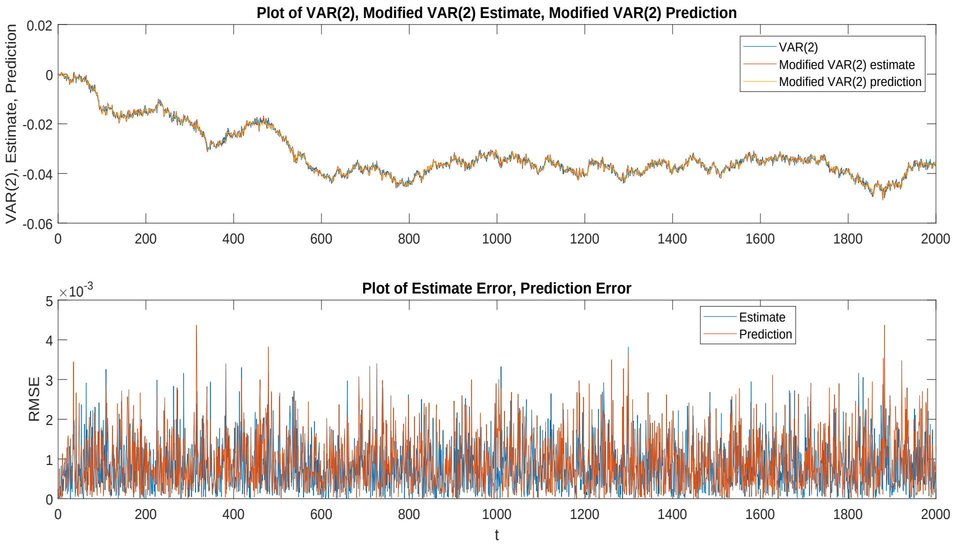


Figure 3.13: Trivariate VAR(2) - Variable 3. In the first subplot, we have comparison of the VAR(2), modified VAR(2) estimate and modified VAR(2) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot displays the errors between VAR(2) and the modified VAR(2) and between VAR(2) and the modified VAR(2) prediction.

Figures 3.11 - 3.13 shows that the updated model performs well since the VAR(2), the modified VAR(2) estimate and the modified VAR(2) prediction are observed to move together. This is supported further by the fact that small values of the root mean square error are obtained.

3.3.4 Case IV: Five Dimensions

Lastly, we check the performance of the updated model by considering the model in five dimensions. In five dimensions, then A_1 and P_t are 5×5 matrices. Suppose now that the state space model is given by

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \\ y_{5,t} \end{pmatrix} = \begin{pmatrix} 0.99 & 0 & 0 & 0 & 0 \\ 0 & 0.99 & 0 & 0 & 0 \\ 0 & 0 & 0.99 & 0 & 0 \\ 0 & 0 & 0 & 0.99 & 0 \\ 0 & 0 & 0 & 0 & 0.99 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \\ y_{4,t-1} \\ y_{5,t-1} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \\ u_{4,t} \\ u_{5,t} \end{pmatrix}$$
$$\begin{pmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \\ x_{4,t} \\ x_{5,t} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \\ y_{5,t} \end{pmatrix} + \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \\ \eta_{4,t} \\ \eta_{5,t} \end{pmatrix} \tag{3.52}$$

Upon setting $Q = \begin{pmatrix} 0.001 & 0 & 0 & 0 & 0 \\ 0 & 0.001 & 0 & 0 & 0 \\ 0 & 0 & 0.001 & 0 & 0 \\ 0 & 0 & 0 & 0.001 & 0 \\ 0 & 0 & 0 & 0 & 0.001 \end{pmatrix}$,

$$R = \begin{pmatrix} 0.001 & 0 & 0 & 0 & 0 \\ 0 & 0.001 & 0 & 0 & 0 \\ 0 & 0 & 0.001 & 0 & 0 \\ 0 & 0 & 0 & 0.001 & 0 \\ 0 & 0 & 0 & 0 & 0.001 \end{pmatrix}$$

and

$$S_0 = \begin{pmatrix} 0.001 & 0 & 0 & 0 & 0 \\ 0 & 0.001 & 0 & 0 & 0 \\ 0 & 0 & 0.001 & 0 & 0 \\ 0 & 0 & 0 & 0.001 & 0 \\ 0 & 0 & 0 & 0 & 0.001 \end{pmatrix}$$

we have the plots in Figures 3.14 - 3.18 which represent the first, second, third, fourth and fifth variables respectively.

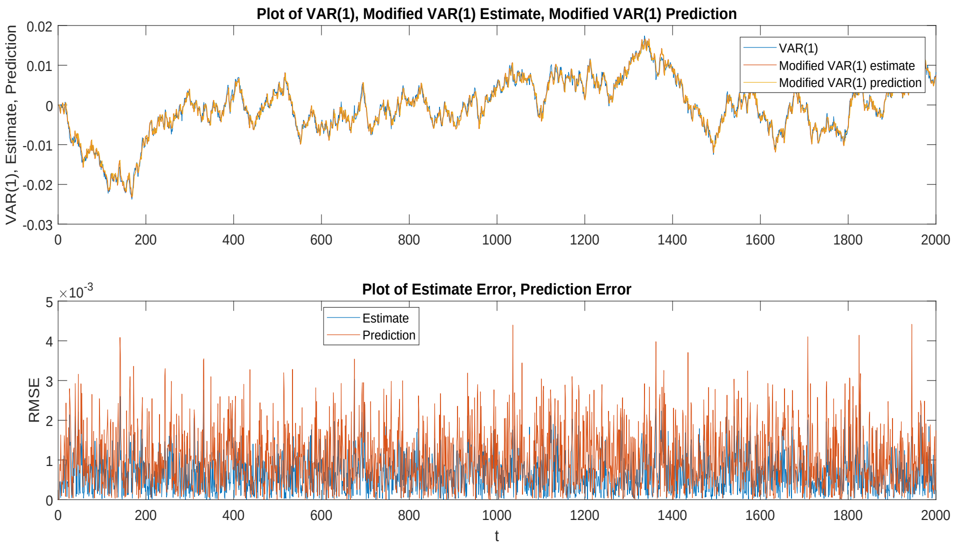


Figure 3.14: Pentavariate VAR(1) - Variable 1. The first subplot gives comparison of the performance o classical VAR(1) and updated model while the second subplot shows the errors between classical VAR(1) and updated model.

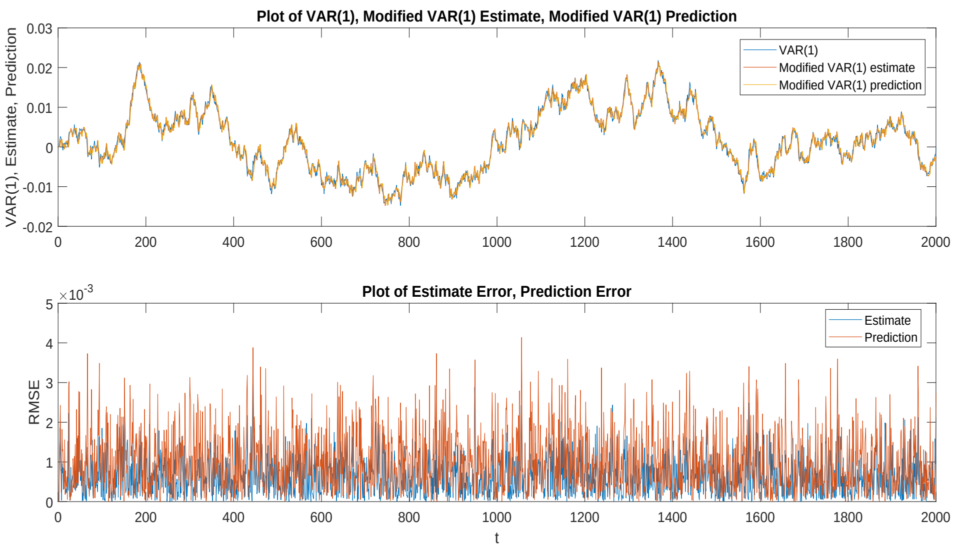


Figure 3.15: Pentavariate VAR(1) - Variable 2. In subplot one compares of the performance of the models while the second subplot gives the errors between the models.

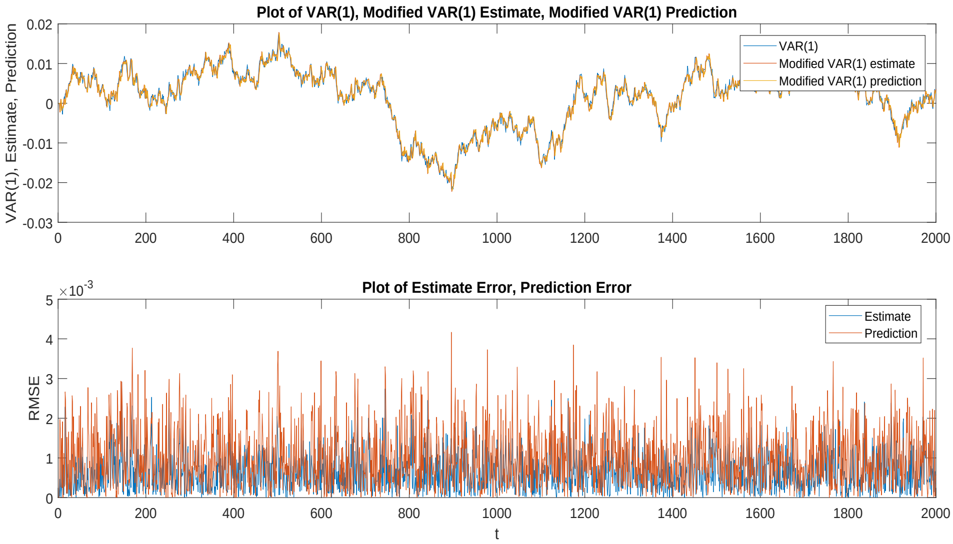


Figure 3.16: Pentavariate VAR(1) - Variable 3. Subplot one gives comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot displays the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction.

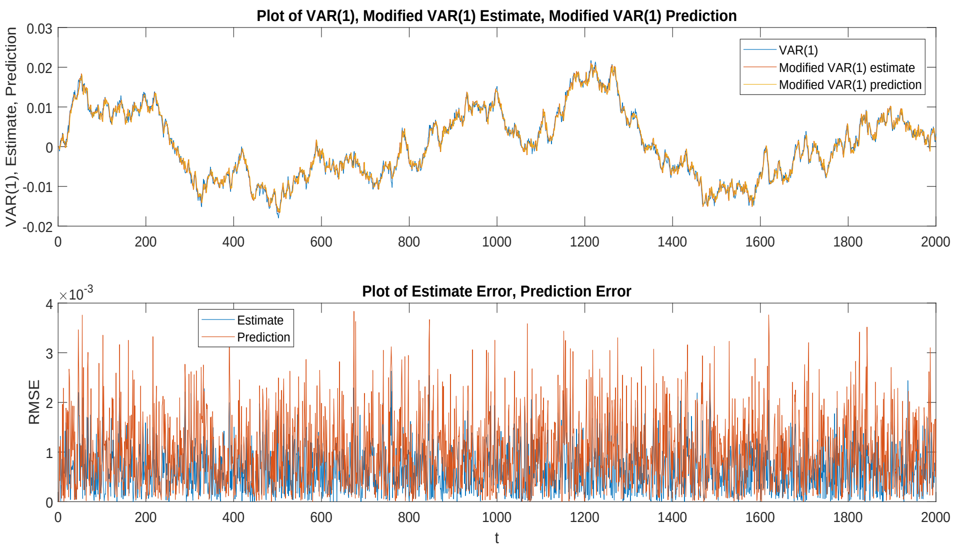


Figure 3.17: Pentavariate VAR(1) - Variable 4. In the first subplot, we have comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while in the second subplot the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction.

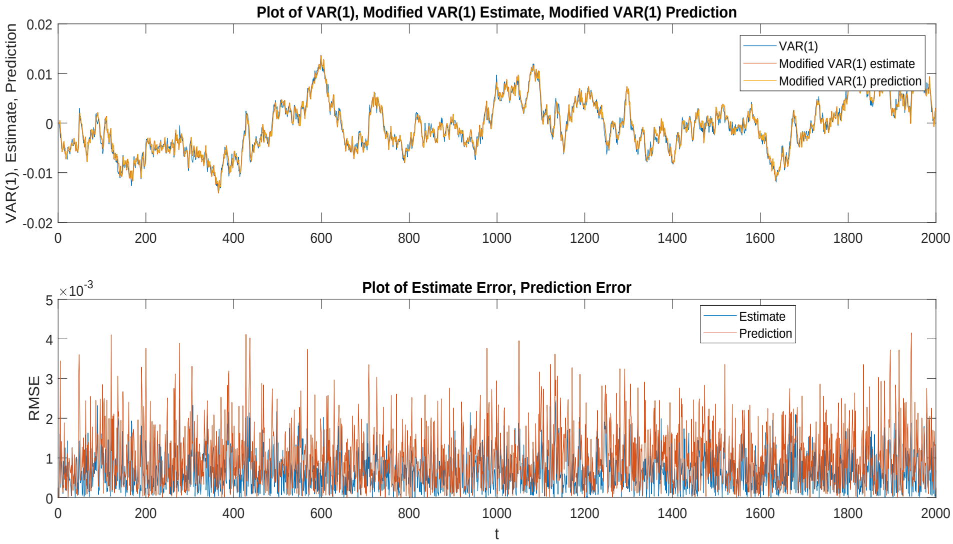


Figure 3.18: Pentavariate VAR(1) - Variable 5. Here, subplot one gives comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while subplot two shows the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction.

From Figures 3.14 - 3.18, it can be observed that the updated model gives precise estimates as seen from the small value of root mean square error between VAR(1) and the modified VAR(1) estimate and between VAR(1) and the modified VAR(1) prediction for each variable.

CHAPTER FOUR

PARAMETER ESTIMATION

4.1 Introduction

Many statistical models, be it deterministic or stochastic, usually contain a number of parameters that make up the model(s). Therefore, parameter estimation is a key step that cannot be avoided as far as modelling or model building is concerned. Commonly, the mostly applied parameter estimation methods include the maximum likelihood estimation, Bayesian estimation and the least squares methods among others. For instance, the maximum likelihood method is usually applied to a density function $f(X|\theta)$ which depends on a set of parameters θ and data set $x_1, x_2, \cdots, x_n$ that are independent and identically distributed (i.i.d) [14, 41]. The likelihood function is given by

$$L(\theta|X) = \prod_{i=1}^{n} fX|\theta$$

which is then maximised though mostly the log-likelihood function given by

$$l(\theta) = lnL(\theta|X)$$

is used [5, 14, 28]. However, this work attends not to these estimation procedures but by considering the use of derived algorithm in the estimation of state and parameter in the updated VAR models.

In state-space models, estimation of parameters and the state is key. Some researches have been done on the estimation of the parameters excluding the state while others have tried to estimate both the parameters and the states after deriving their algorithm. For instance, Kantas *et al* [22] did a study on comprehensive review on particle methods proposed to perform static parameter estimation in state space models. The work mainly focuses on the estimation of the parameters by use algorithms on particle filter. In addition [3] did a research about on-line parameter estimation in non-linear non-Gaussian state-space models with intention to estimate static parameters by point estimation.

[33] did a study on parameter and state estimation for state space models. In the work, a least squares parameter identification algorithm is derived which is then used to estimate the parameters. Thereafter, the estimated parameters are then used to compute the system states by incorporating input-output data. Indeed, it can be noted that the estimation of parameters and states in state space models has attracted interest for most researchers based on the algorithm derived depending on the nature of the state space model as seen in other works such as [8, 12, 31, 36].

In this Chapter, we consider estimation of the parameters by use of the dual estimation approach which estimates both the state and the parameters though our main focus is on the parameter estimation. The approach, dual estimation, involves the estimation of the state and the parameters simultaneously [4, 25]. However, dual estimation can be done in two ways namely: joint estimation and dual filter whereby joint estimation requires only one filter whereas dual filter requires two filters. Joint estimation is advantageous over the dual estimation in that the former allows for dependencies in parameters and states while the later assumes no autocorrelation, that is, the cross covariances are zero, [25]. Joint estimation involves augmenting the state vector with vector of parameters to form an extended state-space and then the algorithm is run forward in time to update both the state and the parameters with expectations of the algorithm converging to the optimal state and parameter values. In this work, joint estimation is considered.

4.2 Joint Estimation

As mentioned earlier, joint estimation involves augmenting the state vector with vector of parameters to form an extended state-space and then the algorithm is run forward in time to update both the state and the parameters with expectations of the algorithm converging to the optimal state and parameter values. Therefore, for the VAR(1) model, let

$$z_t = \varphi_t \tag{4.1}$$

where

$$z_t = \begin{pmatrix} Y_t \\ \theta_t \end{pmatrix} \qquad \textit{and} \qquad \varphi_t = \begin{pmatrix} A_1 Y_{t-1} + u_t \\ \theta_{t-1} \end{pmatrix}$$

On the other hand, for the VAR(p) model

$$z_t = \begin{pmatrix} Y_t \\ \theta_t \end{pmatrix} \qquad \textit{and} \qquad \varphi_t = \begin{pmatrix} A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + u_t \\ \theta_{t-1} \end{pmatrix}$$

Algorithm 2 is then applied on the joint state space system given in Equation 4.1 which estimates the states and the parameters simultaneously. Therefore, next we proceed to estimate the parameters of the models to check if their is convergence to the true parameter values. First, a case of one dimension is considered after which we consider the case of two dimension. As the estimation is done, it should be noted that we assume that the parameters are time-invariant, that is, they are static.

4.2.1 Case I: One Dimension

We consider the model given by

$$\begin{aligned} Y_t &= A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + u_t \\ X_t &= P_t Y_t + \eta_t \end{aligned} \tag{4.2}$$

where we assume it is in scalar form and proceed to estimate the parameters, $A_1, \cdots, A_p$ and the state, Y_t through joint estimation. We assume the initial state, Y_0, and the values of $A_1, \cdots, A_p$, P_t, u_t and η_t are given and then run the algorithm to investigate if it converges to the true parameter value. Suppose that we have the model given by

$$\begin{aligned} Y_t &= A_1 Y_{t-1} + u_t \\ X_t &= P_t Y_t + \eta_t \end{aligned} \tag{4.3}$$

We set $A_1 = -0.2$ so that the model can be written as

$$\begin{aligned} Y_t &= -0.2 Y_{t-1} + u_t \\ X_t &= Y_t + \eta_t \end{aligned} \tag{4.4}$$

where in Equation 4.4, $A_1 = a = -0.2$. Using Algorithm 2, we proceed to estimate the parameter a. Using MATLAB, and setting $Q = 0.01$ and $R = 0.001$ in Algorithm 2, we have the panels as given in Figure 4.1 which gives the estimates of the parameter a over time plus its corresponding Box-plot. From Figure 4.1 it can be observed that the algorithm yields converging results to the true parameter value as time evolves which is -0.2 with some margin of error. In addition, the accompanying Box-plot for parameter a in Figure 4.1 shows the dispersion in the results though with some outliers present on both the lower and upper sides of the Box-plot, with most of them on the lower side than on the upper side. Next, suppose that we have the model given by

$$\begin{aligned} Y_t &= A_1 Y_{t-1} + A_2 Y_{t-2} + u_t \\ X_t &= P_t Y_t + \eta_t \end{aligned} \tag{4.5}$$

Suppose we set $A_1 = -0.2$ and $A_2 = 0.2$. Then the model in Equation (4.5) can be written as

$$\begin{aligned} Y_t &= -0.2 Y_{t-1} + 0.2 Y_{t-2} + u_t \\ X_t &= Y_t + \eta_t \end{aligned} \tag{4.6}$$

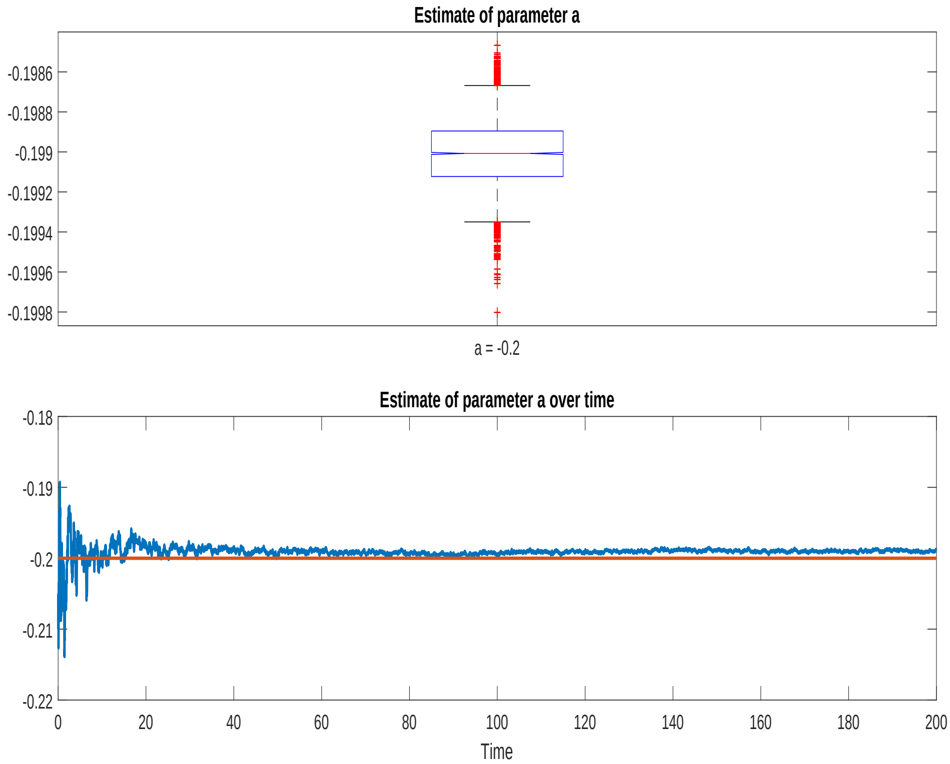


Figure 4.1: Parameter Estimation in AR(1) model. Subplot one gives the corresponding box plot showing the dispersion while the second subplot shows the convergence of the parameter a to the true parameter value -0.2 as time evolves.

where in Equation 4.6, $A_1 = a = -0.2$ and $A_2 = b = 0.2$. Using the Algorithm 2, we estimate parameters a and b. Setting $Q = 0.01$, $R = 0.001$, from MATLAB, we have the panels as given in Figure 4.2 which gives the estimates of the parameters a and b over time plus their corresponding Box-plots. From Figure 4.2 it is clear that the algorithm yields converging results to the true parameter values as time evolves which are -0.2 for a and 0.2 for b but with some margin of error. The accompanying Box-plots for parameters a and b in Figure 4.2 shows the dispersion in the results and it can be observed that there are few outliers present which appear on the lower side of the Box-plots for both parameters a and b.

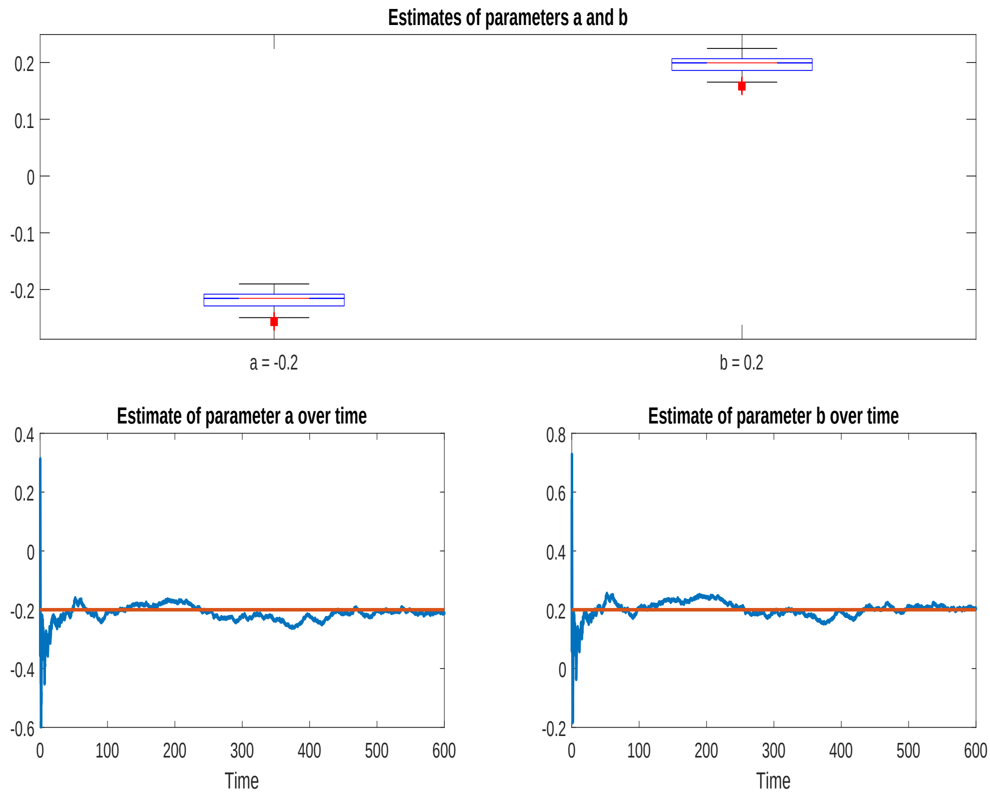


Figure 4.2: Parameter Estimation in AR(2) model. The first subplot gives the corresponding box plots showing the dispersion in both parameters a and b, while the second subplot shows the convergence of the parameters a and b to the true parameter values -0.2 and 0.2, respectively as time evolves.

4.2.2 Case II: Two Dimensions

Next we proceed to estimate the parameters in the two dimensions model. Consider the model given by

$$\begin{aligned} \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} &= \begin{pmatrix} -0.2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} \\ \begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} + \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \end{pmatrix} \end{aligned} \tag{4.7}$$

Again using Algorithm 2, we estimate the parameters (elements) of matrix $a = \begin{pmatrix} -0.2 & 0 \\ 0 & 0 \end{pmatrix}$ where $a_{11} = -0.2$, $a_{12} = 0$, $a_{21} = 0$ and $a_{22} = 0$. Using MATLAB, setting $Q = \begin{pmatrix} 0.001 & 0 \\ 0 & 0.001 \end{pmatrix}$ and $R = \begin{pmatrix} 0.0101 & 0 \\ 0 & 0.0101 \end{pmatrix}$, we have the panels as given in Figure 4.3 which gives the estimate of parameter a_{11} over time plus its corresponding Box-plot. From Figure 4.3 it is evident that the algorithm yields

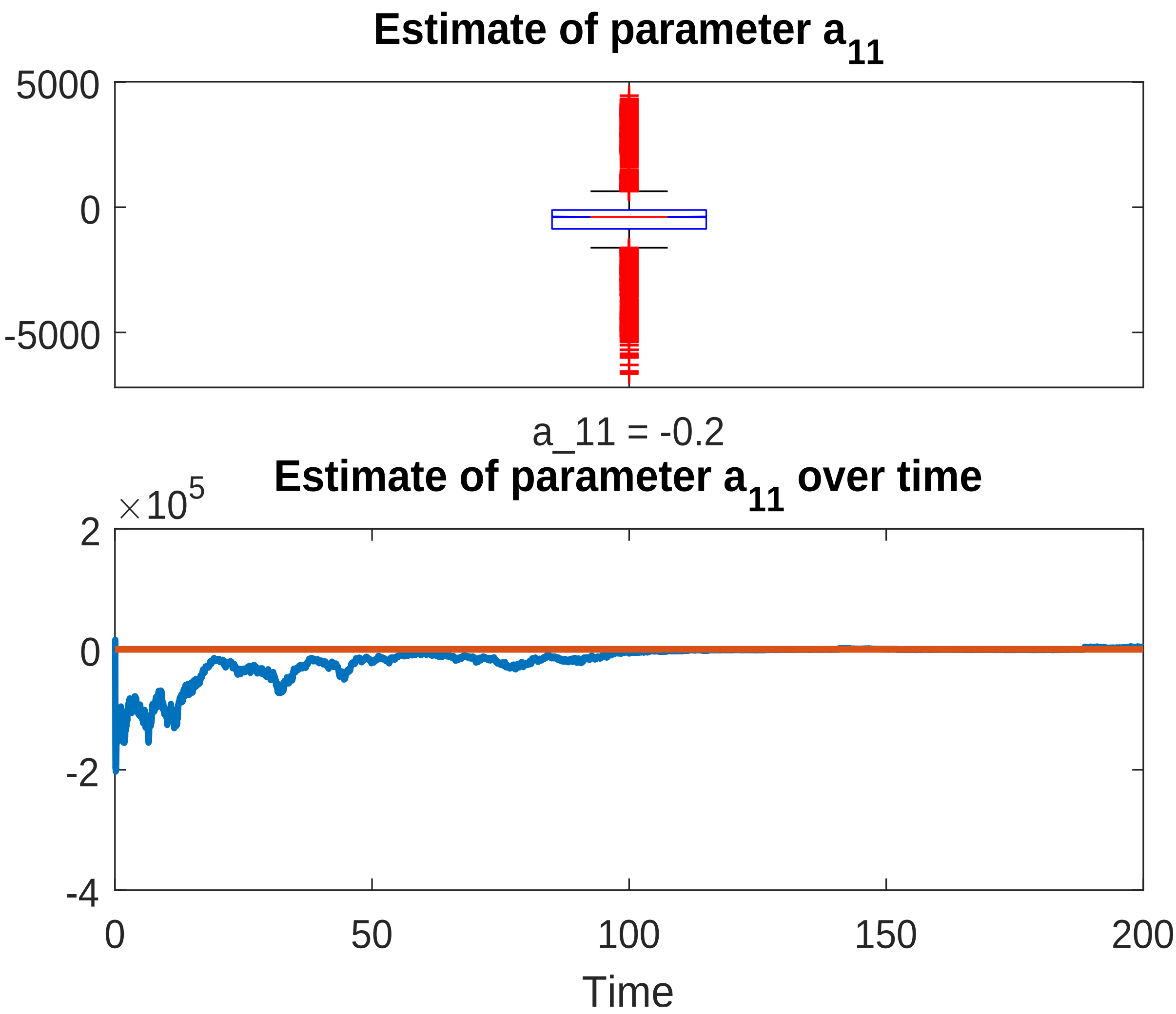


Figure 4.3: Parameter Estimation in VAR(1) model. Subplot one shows the corresponding box plot showing the dispersion while the second subplot shows the convergence of the parameter to the true parameter value as time evolves.

converging results to the true parameter value as time evolves which are -0.2. However, there is some margin of error present in the convergence. The accompanying Box-plot for parameter a_{11} in Figure 4.3 represents the dispersion in the results with outliers present on both the lower and upper sides of the Box-plot. However, the outliers appear more on the lower side as compared to the upper side.

CHAPTER FIVE

APPLICATION OF THE UPDATED VAR MODEL

5.1 Introduction

In this chapter, illustration is given regarding the application of the updated VAR model to some real data.

5.2 Modeling the Contribution of the Agricultural Sub-sectors to the National GDP using the Updated VAR Model

In this section, we consider secondary data obtained from the Kenya National Bureau of statistics, Statistical Abstract reports from 2000-2021. The data is on monetary value marketed at current prices (Ksh. Million) from crops, horticulture, livestock and related products, fisheries and forestry. We consider the contribution of the listed agricultural sub-sectors to the national gross domestic product due to the fact that the agriculture sector plays a key role as far as a country's economic GDP is concerned. The data was entered into Excel and saved under CSV format. It was then read in R software for analysis. First, a time series plot of the variables is done which is as given in Figure 5.1. From figure 5.1, we observe that almost all the variables depict an increasing trend implying that the GDP from the variables has been increasing steadily over the years. Thus the series are non-stationary as confirmed by Dickey-Fuller test. We make the series stationary by applying differencing twice to the log of the variables. This gives the plot in Figure 5.2.

From Figure 5.2, it is observed that the variables appear stationary and can be adopted. The Augmented Dickey Fuller test also shows that the series are stationary. Using the lagselect function, we find that the Akaike information criteria

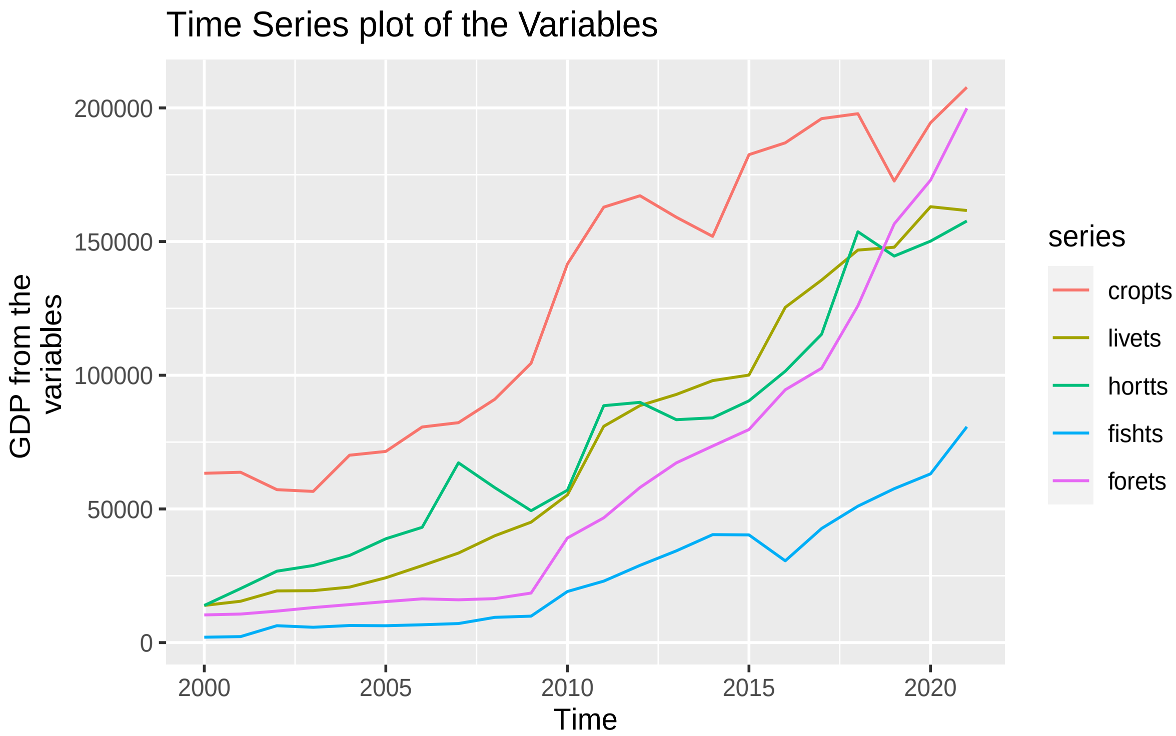


Figure 5.1: Time series plot of GDP from the variables, namely; crops, livestock and related products, horticulture, fisheries and forestry.

(AIC), Hannan-Quinn (HQ) criteria, Schwartz Criteria (SC) and Final Prediction Error (FPE) selects order of the model as 1 i.e $p = 1$ implying VAR(1) model is a suitable model. The model is given by

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \\ y_{5,t} \end{pmatrix} = \begin{pmatrix} -0.3930 & -0.0627 & -0.1585 & -0.0901 & 0.0852 \\ 0.2784 & -0.4275 & 0.1043 & -0.0419 & 0.2829 \\ 0.6599 & 0.0578 & -0.2950 & -0.0195 & 0.4020 \\ -0.2265 & 0.0063 & 0.1889 & -0.6139 & 0.1257 \\ 0.1563 & -0.2264 & -0.0798 & -0.0194 & -0.4631 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \\ y_{4,t-1} \\ y_{5,t-1} \end{pmatrix} + \begin{pmatrix} 0.0001487 \\ -0.001047 \\ -0.0008652 \\ 0.0001187 \\ -0.0000301 \end{pmatrix} + \begin{pmatrix} \omega_{1,t} \\ \omega_{2,t} \\ \omega_{3,t} \\ \omega_{4,t} \\ \omega_{5,t} \end{pmatrix} \tag{5.1}$$

The eigenvalues of the matrix of the penta-variate VAR model in equation 5.1 are obtained by solving for λ in the equation

$$det[\Phi_1 - \lambda I_5] = 0 \tag{5.2}$$

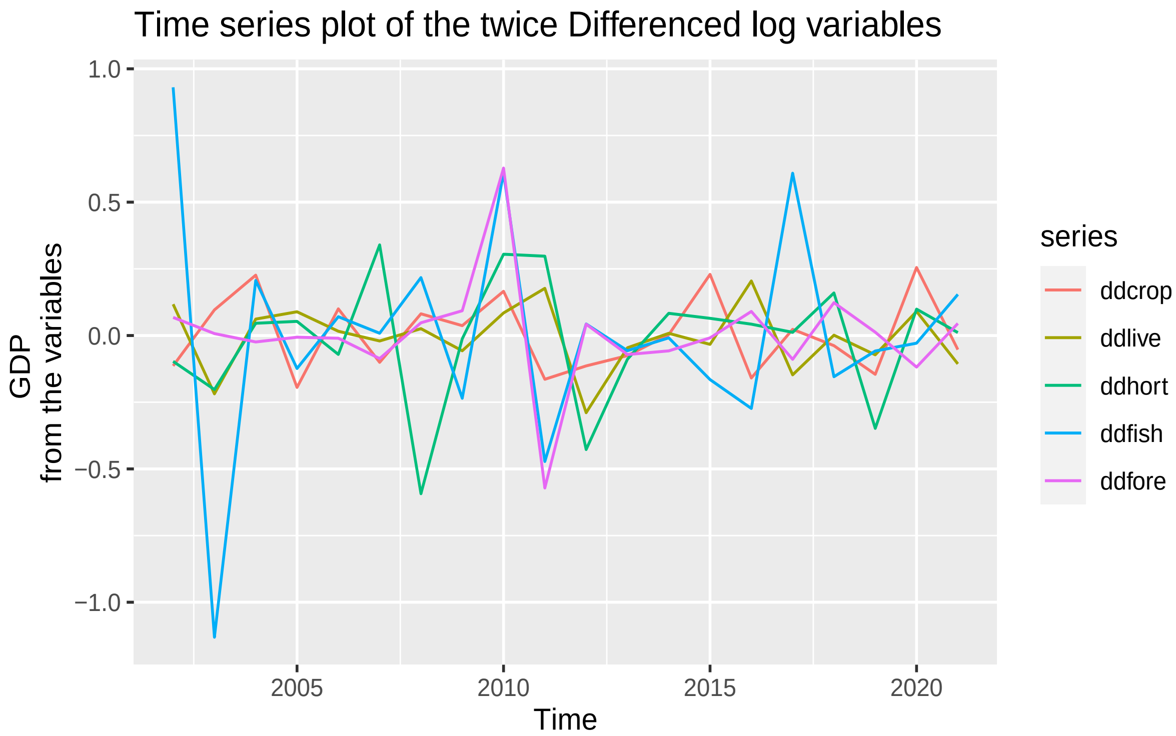


Figure 5.2: Time series plot of twice differenced log GDP from the Variables, that is, after applying second differencing to the log of the variables.

where

$$\Phi_1 = \begin{pmatrix} -0.3930 & -0.0627 & -0.1585 & -0.0901 & 0.0852 \\ 0.2784 & -0.4275 & 0.1043 & -0.0419 & 0.2829 \\ 0.6599 & 0.0578 & -0.2950 & -0.0195 & 0.4020 \\ -0.2265 & 0.0063 & 0.1889 & -0.6139 & 0.1257 \\ 0.1563 & -0.2264 & -0.0798 & -0.0194 & -0.4631 \end{pmatrix} \tag{5.3}$$

Equivalently, this is given by

$$det \begin{pmatrix} -0.3930-\lambda & -0.0627 & -0.1585 & -0.0901 & 0.0852 \\ 0.2784 & -0.4275-\lambda & 0.1043 & -0.0419 & 0.2829 \\ 0.6599 & 0.0578 & -0.2950-\lambda & -0.0195 & 0.4020 \\ -0.2265 & 0.0063 & 0.1889 & -0.6139-\lambda & 0.1257 \\ 0.1563 & -0.2264 & -0.0798 & -0.0194 & -0.4631-\lambda \end{pmatrix} = 0 \tag{5.4}$$

The eigenvalues are $\lambda_1 = -0.7078666 + 0i$, $\lambda_2 = -0.2623644 + 0.4698648i$, $\lambda_3 = -0.2623644 - 0.4698648i$, $\lambda_4 = -0.4799684 + 0.0770515i$ and $\lambda_5 = -0.4799684 - 0.0770515i$ whose moduli are 0.7078666, 0.5381524, 0.5381524, 0.4861138 and 0.4861138 respectively. All the eigenvalues have modulus less than one (lie within the complex unit circle) thus the model is stable.

We tested for Granger-Causality of the variables and found that Crops, Livestock, Horticulture, Fishing and Forestry do not Granger-cause the other variables. The test on normality of the residuals for the model in Equation 5.1 found that the residuals are normally distributed. Also, the residuals were found to be uncorrelated as seen from the ACF plots of the residuals (see appendices) of the developed model in Equation 5.1.

Suppose we now combine the fitted model given in equation 5.1 with the measurement equation so that we have,

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \\ y_{5,t} \end{pmatrix} = \begin{pmatrix} -0.3930 & -0.0627 & -0.1585 & -0.0901 & 0.0852 \\ 0.2784 & -0.4275 & 0.1043 & -0.0419 & 0.2829 \\ 0.6599 & 0.0578 & -0.2950 & -0.0195 & 0.4020 \\ -0.2265 & 0.0063 & 0.1889 & -0.6139 & 0.1257 \\ 0.1563 & -0.2264 & -0.0798 & -0.0194 & -0.4631 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \\ y_{4,t-1} \\ y_{5,t-1} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \\ u_{4,t} \\ u_{5,t} \end{pmatrix}$$
$$\begin{pmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \\ x_{4,t} \\ x_{5,t} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \\ y_{5,t} \end{pmatrix} + \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \\ \eta_{4,t} \\ \eta_{5,t} \end{pmatrix} \tag{5.5}$$

We then subject equation 5.5 to algorithm 2 to update the model. Setting

$$Q = \begin{pmatrix} 0.0190578 & 0.001970 & 0.005758 & 0.0002899 & 0.009302 \\ 0.0019701 & 0.008454 & 0.005498 & 0.0043966 & 0.001251 \\ 0.0057579 & 0.005498 & 0.047239 & 0.0244832 & 0.004615 \\ 0.0002899 & 0.004397 & 0.024483 & 0.0982311 & 0.029322 \\ 0.0093015 & 0.001251 & 0.004615 & 0.0293222 & 0.046527 \end{pmatrix},$$

$$R = \begin{pmatrix} 0.001 & 0 & 0 & 0 & 0 \\ 0 & 0.001 & 0 & 0 & 0 \\ 0 & 0 & 0.001 & 0 & 0 \\ 0 & 0 & 0 & 0.001 & 0 \\ 0 & 0 & 0 & 0 & 0.001 \end{pmatrix}$$

and

$$S_0 = \begin{pmatrix} 0.001 & 0 & 0 & 0 & 0 \\ 0 & 0.001 & 0 & 0 & 0 \\ 0 & 0 & 0.001 & 0 & 0 \\ 0 & 0 & 0 & 0.001 & 0 \\ 0 & 0 & 0 & 0 & 0.001 \end{pmatrix}$$

the plots in Figure 5.3 - Figure 5.7 are obtained for the five variables being referred to.

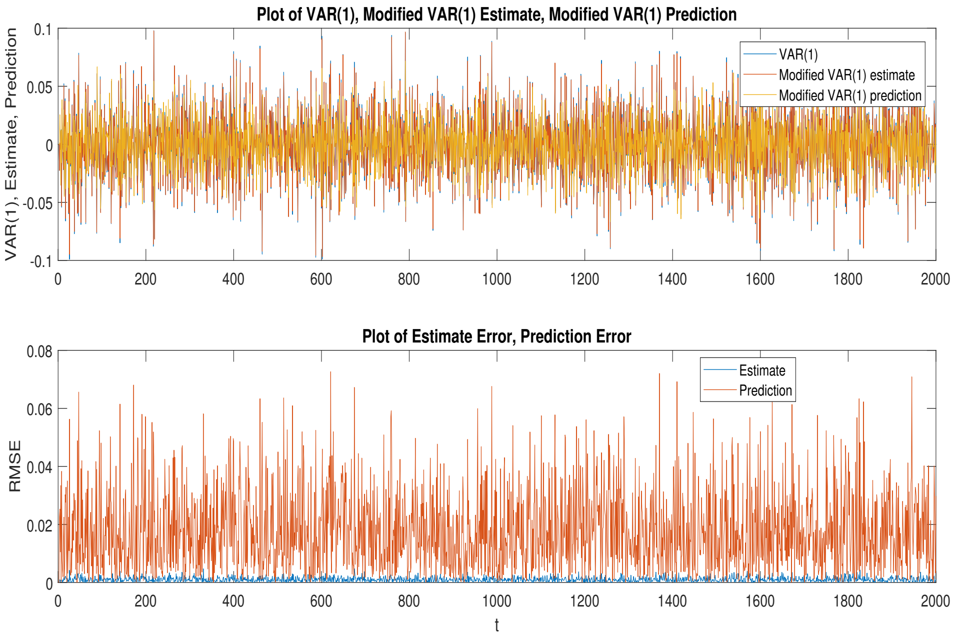


Figure 5.3: Fitted Pentavariate VAR(1) - Variable 1. Subplot one gives comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot shows the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction as time evolves.

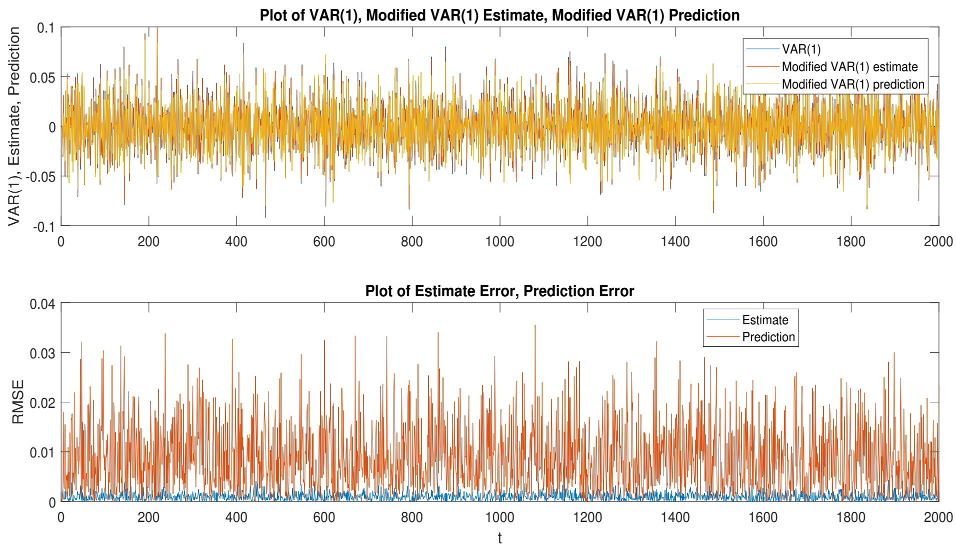


Figure 5.4: Fitted Pentavariate VAR(1) - Variable 2. The first subplot gives comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction.

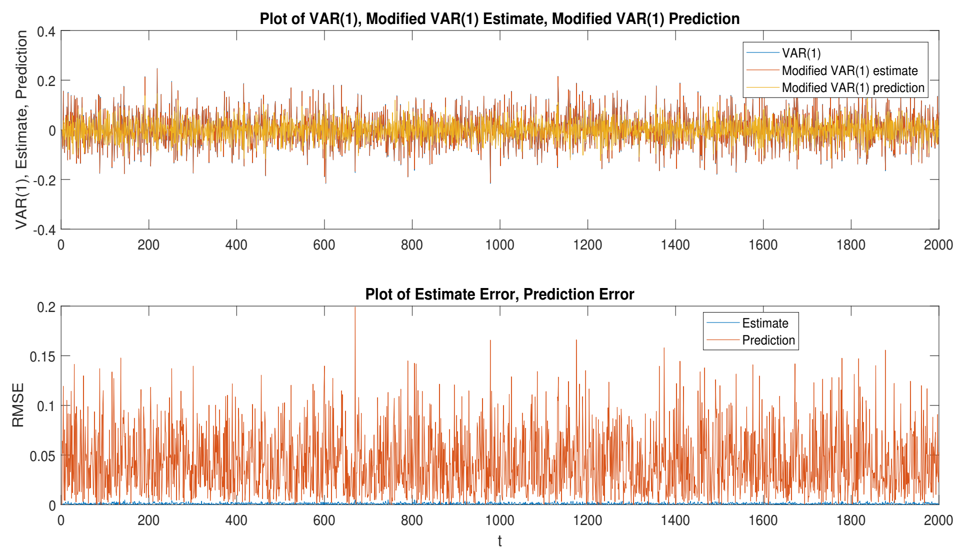


Figure 5.5: Fitted Pentavariate VAR(1) - Variable 3. In subplot one, comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively is shown, while the second subplot displays the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction.

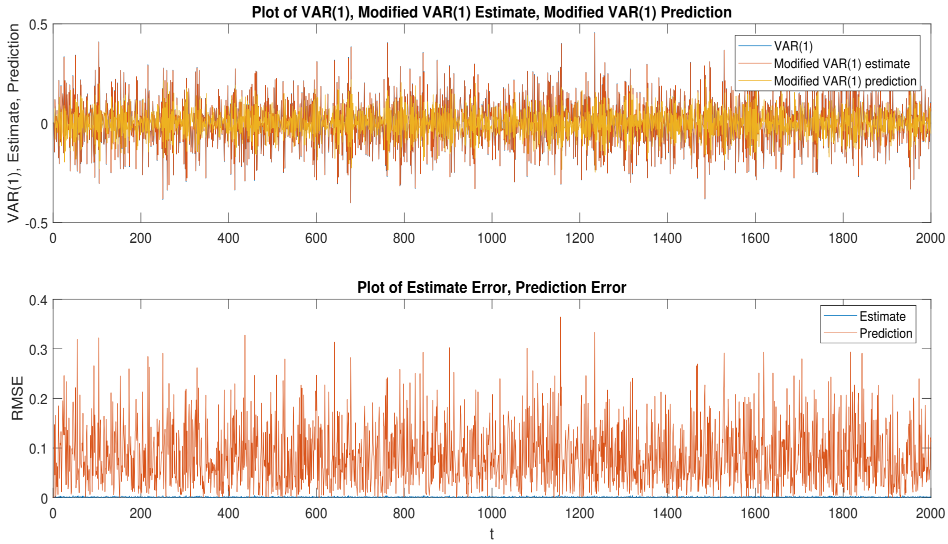


Figure 5.6: Fitted Pentavariate VAR(1) - Variable 4. Here, the first subplot gives comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while the second subplot shows the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction.

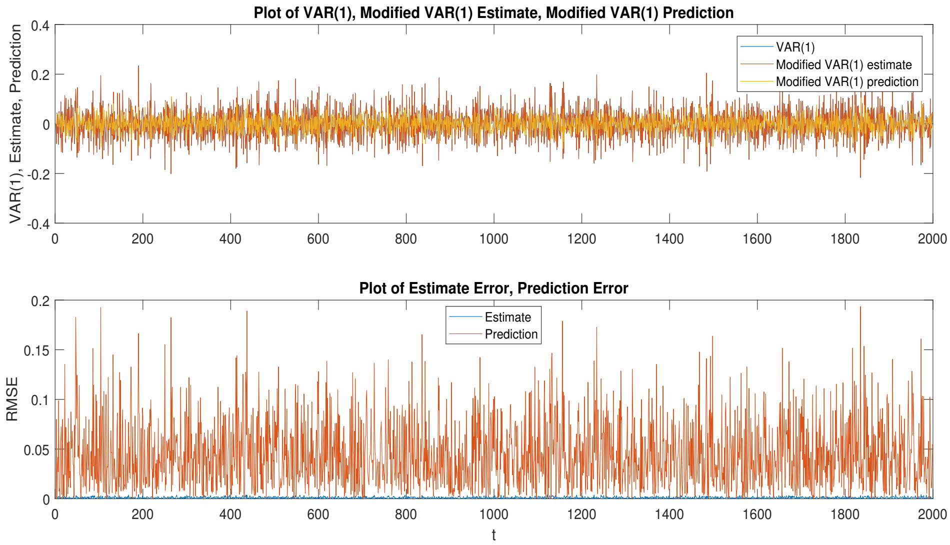


Figure 5.7: Fitted Pentavariate VAR(1) - Variable 5. In the first subplot, we have comparison of the VAR(1), modified VAR(1) estimate and modified VAR(1) prediction, denoted by the blue line, red line and the yellow line, respectively while in the second subplot, we have the errors between VAR(1) and the modified VAR(1) and between VAR(1) and the modified VAR(1) prediction.

From Figures 5.3 - 5.7, it can be observed that the values of the root mean square error are fairly small an indication that the updated model performs well. Figures 5.3 - 5.7 represent the first, second, third, fourth and fifth variables respectively. In addition, it is evident that if the model has too many parameters, its ability to perform well is lowered since the values of the RMSE are observed to be high in Figures 5.3 - 5.7 for the model with many parameters than in Figures 3.14 - 3.18 which is for the model with few parameters.

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

6.1 Introduction

In this chapter, conclusions for the findings from the research are given. In addition, recommendations regarding future work which may be done regarding other time series models are given.

6.2 Conclusions

This study developed an updated VAR model to incorporate new information. Using the Bayesian technique, an updated VAR model of order 1 is developed as seen in Algorithm 1. Afterwards, VAR models of order 2 and 3 are considered after which generalization to the VAR model of order p is done as given in Algorithm 2. The performance of the updated VAR model is compared with the performance of corresponding VAR models from which it is observed that the updated model performs well based on the low values of the root mean square error (RMSE) which is used as a tool for adequacy checking. From the analysis, the values of the root mean square error between the existing VAR model and the modified VAR estimate and between the existing VAR model and the modified VAR predicted were low an indication of good model performance. In addition, it is observed that as time goes on, the existing VAR model, modified VAR estimate and modified VAR predicted are seen not to diverge from each other an indication of good performance from the models. Furthermore, estimation of parameters for some VAR models is done using the joint estimation. In joint estimation, both the states and the parameters are estimated simultaneously using Algorithm 2 as time evolves where it is checked whether there is convergence to the true parameter values. From the results, by incorporating Algorithm 2, it is found that there is convergence to the true parameters as time goes on which signifies the betterment of the model as time goes on when new information is obtained.

6.3 Recommendations

In this study, an updated Vector Autoregressive (VAR) model has been formulated using the Bayesian approach where the existing VAR model is considered as the prior, the new information (the measurements) as the likelihood to update the VAR model and hence get the updated VAR model. However, other multivariate time series models such as the Vector Autoregressive Moving Average (VARMA), Vector Moving Average (VMA) and Factor Augmented Vector Autoregressive Moving Average (FAVARMA) models also need to be considered incase new information emerges after the model has been developed. Therefore, as recommendations for future work, we suggest that future work be done to update the models (other multivariate time series models) to cater for the new information which is obtained after the model has been developed.

REFERENCES

[1] Abdullah T. L. (2022) Forecasting time series using Vector Autoregressive Model, *Int. J. Nonlinear Anal. Appl.*, **13**(1), 499-511.

[2] Amoah E. (2015). *Modeling Gdp Using Vector Autoregressive (VAR) Models: An Empirical Evidence From Ghana.* [Masters dissertation, University of Ghana]. University of Ghana. http://ugspace.ug.edu.gh

[3] Andrieu C., Doucet A., and Tadic B. V. (2005) On-line Parameter Estimation in General State-Space Models. Available at: https://www.cs.ubc.ca/ arnaud/andrieu_doucet_tadic_onlineparameterestimationCDC2005.pdf

[4] Angwenyi N. D. (2019). *Time-Continuous State and Parameter Estimation with Application to Hyperbolic SPDEs* [Doctoral dissertation, University of Potsdam]. Institutional Repository of the University of Potsdam. https://doi.org/10.25932/publishup-43654

[5] Box G. E. P. and Jenkins G. M., Reinsel G. C. and Ljung G. M. (2015). *Time Series Analysis: Forecasting and Control*, 5th Edition, ISBN: 978-1-118-67502-1.

[6] Brockwell J. P. and Davis A. R. (2016). *Introduction to Time Series and Forecasting*, Springer-Verlag New York.

[7] Caraiani P. (2010) Forecasting Romanian GDP using a BVAR Model, *Journal of Economic Forecasting*, **0**(4), 76 - 87.

[8] Cui T., Ding F., Alsaedi A. and Hayet T. (2020) Data filtering-based parameter and state estimation algorithms for state-spacesystems disturbed by coloured noises, *International Journal Of Systems Science*, **51**(9), 1669–1684.

[9] Domit S., Monti F. and Sokol A. (2019) Forecasting the UK economy with a medium-scale Bayesian VAR. *International Journal of Forecasting*, **35**(4), 1669-1678.

[10] Giannone D., Lenza M. and Primiceri E. G. (2013) Prior Selection for Vector Autoregressions. Avalaible at: *https://faculty.wcas.northwestern.edu/gep575/Draft_GLP_V24.pdf*

[11] Granger C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods, *Econometrica*, **37**(3), 424–438.

[12] Gu Y., Zhu Q., Liu J., Zhu P. and Chou Y. (2020) Multi-Innovation Stochastic Gradient Parameter and State Estimation Algorithm for Dual-Rate State-Space Systems with d-Step Time Delay, *Complexity*, **2020**, 1 - 11.

[13] Gudeta O. D., Arero G. B. and Goshu T. A. (2017) Vector Autoregressive Modelling of some Economic Growth Indicators of Ethiopia, *American Journal of Economics*, **7**(1), 46-62.

[14] Halladay A. (2006) *Parameter Estimation in Linear State-Space Models using Particle Filters and the EM Algorithm* [Bachelors Honours, Dalhousie University]. Available at: https://www.mathstat.dal.ca/ amandah/HonoursThesis.pdf

[15] Hamzah L. M., Nabilah S. U., Usman M., Virginia E. and Wamiliana (2020). Dynamic Modeling and Forecasting of Data Export of Agricultural Commod-

ity by Vector Autoregressive Model, *Journal of Southwest Jiaotong University*, **55**(3), 1 - 10.

[16] Hossain, M., Al Amin, A. A., & Islam, A. H. M. S. (2018). Modeling and forecasting of climatic parameters: univariate SARIMA versus multivariate vector autoregression approach. *Journal of the Bangladesh Agricultural University*, **16**(1), 131–143.

[17] Huang Y-F. (2012) Forecasting Chinese Inflation and Output: A Bayesian Vector Autoregressive Approach. Available at: https://mpra.ub.uni-muenchen.de/41933/

[18] Jazwinski, A. (1970), Stochastic Processes and Filtering Theory, Academic Press, New York.

[19] Jinghao M., Yujie S. and Hongyan Z. (2021), Application of Bayesian Vector Autoregressive Model in Regional Economic Forecast, *Complexity*, **2021**, 1 - 10.

[20] Joyce, James, "Bayes' Theorem", The Stanford Encyclopedia of Philosophy (Fall 2021 Edition), Edward N. Zalta (ed.), URL = ¡https://plato.stanford.edu/archives/fall2021/entries/bayes-theorem/¿.

[21] Kalliovirta L., Niskanen O. and Heikkila A. M. (2019) Forecasting Milk prices with VAR Models - Application to Farm Gate Price in Finland. Available at SSRN: *https://ssrn.com/abstract=3473862* or *http://dx.doi.org/10.2139/ssrn.3473862*

[22] Kantas N., Doucet A., Singh S. S., Maciejowski J. and Chopin N. (2015) On Particle Methods for Parameter Estimation in State-Space Models, *Statistical Science*, **30**(3), 328–351.

[23] Karlsson S. (2013) Forecasting with Bayesian Vector Autoregression. In G. Elliott, A. Timmermann (Eds.), *Handbook of Economic Forecasting*, 791-897, Elsevier.

[24] Khairan R., Firuz K. and Aswani K. C. (2022) Forecasting COVID-19: Vector Autoregression-Based Model, *Arabian Journal for Science and Engineering* **47**, 6851–6860.

[25] Le, Cassidy, "Use of Kalman Filtering in State and Parameter Estimation of Diabetes Models" (2020). HMC Senior Theses. 232. https://scholarship.claremont.edu/hmc_theses/232

[26] Lidiema C., Waititu A., Mageto T. and Ngunyi A. (2018) Forecasting Household Credit In kenya Using Bayesian Vector Autoregressive (BVAR) Model. *American Journal of Applied Mathematics and Statistics* **6**(1), 17-24.

[27] Lutkepohl H. (2004) Forecasting with VARMA Models. *European University Institute* Badia Fiesolana I-50016 San Domenico (FI).

[28] Lutkepohl H. (2005) *New Introduction to Multiple Time Series Analysis*, Springer Berlin Heidelberg New York.

[29] Mahler R. P. S. (2007) *Statistical Multisource-multitarget Information Fusion.* Artech House, Inc.

[30] Ouahiba Elalaoui, Aziz Fadlaoui, Nassreddine Maatala, Abdelouafi Ibrahimy (2021). Agriculture and GDP Causality Nexus in Morocco: Empirical Evidence from a VAR Approach. *International Journal of Agricultural Economics*, **6**(4), 198-207.

[31] Ouambo T., Boum A., and Imano M. (2022) States and Parameters Estimation for InductionMotors Based on a New Adaptive Moving Horizon Estimation, *Journal of Electrical and Computer Engineering*, **2022**, 1 - 11.

[32] Quenouille M. H. (1957) *The Analysis of Multiple Time-Series*, Griffin, London

[33] Ruifeng D., Linfan Z., (2014) Parameter and State Estimator for State Space Models, *The Scientific World Journal*, **2014**, 1 - 10.

[34] Roush J., Siopes K. and Hu G. (2017) Predicting gross domestic product using autoregressive models, *IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA)*, 317-322

[35] Saheed B. A., Kola Y. K., Adewole O. G., Muideen O. A., Saheed O. J., Muhammed T. M., Ekundayo S. O., Olakiitan I. A., Magdalene P. and Oluwole J. O. (2021) Vector Autoregressive Modeling of Crop Production Index - Permanent Cropland Relationship in Nigeria, *Annals: Computer Science Series*, **19**(1), 92 - 97.

[36] Salhi H. and Kamoun S. (2016) Combined Parameter and State Estimation Algorithms for Multivariable Nonlinear Systems Using MIMO Wiener Models, *Journal of Control Science and Engineering*, **2016**, 1 - 12.

[37] Sims C. A. (1980), Macroeconomics and Reality, *Econometrica*, **48**(1), 1 - 48.

[38] Stock, J.H. and Watson M.W. (2001). “Vector Autoregressions,” *Journal of Economic Perspectives*, **15**(4), 101–115.

[39] Stylianides N. and Kontou E. (2020) Bayes Theorem and its Recent Applications, *MA3517 Mathematics Research Journal*, 1 - 7.

[40] Sugita, K. (2022), "Forecasting with Bayesian vector autoregressive models: comparison of direct and iterated multistep methods", *Asian Journal of Economics and Banking*, **6**(2), 142-154.

[41] Tiao G. C. and Box G. E. P. (1981). Modeling Multiple Time Series with Applications, *Journal of the American Statistical Association*, **76**(376) 802-816.

[42] Wong C. F. and Beng W. G. (1993) A bayesian Vector-autoregression model for forecasting quarterly GDP: the Singapore experience. *The Singapore Economic Review*, **38(1)**, 015-034.

[43] Yashavanth B. S., SinghK. N., Paul K. A. and Paul K. R. (2017) Forecasting prices of coffee seeds using Vector Autoregressive Time Series Model. *Indian Journal of Agricultural Sciences*, **87**(6), 754-758.

[44] Youssef Y. (2022) Bayes Theorem and Real-life Applications. Available at:https://www.researchgate.net/publication/361402449_Bayes_Theorem_and_Real-life_Applications

[45] Zivot E. and Wang J. (2003) *Modeling Financial Time Series with S-Plus.* Springer, New York

APPENDICES

(i) ACF Plots of the Residuals for the Fitted Pentavariate VAR(1) Model

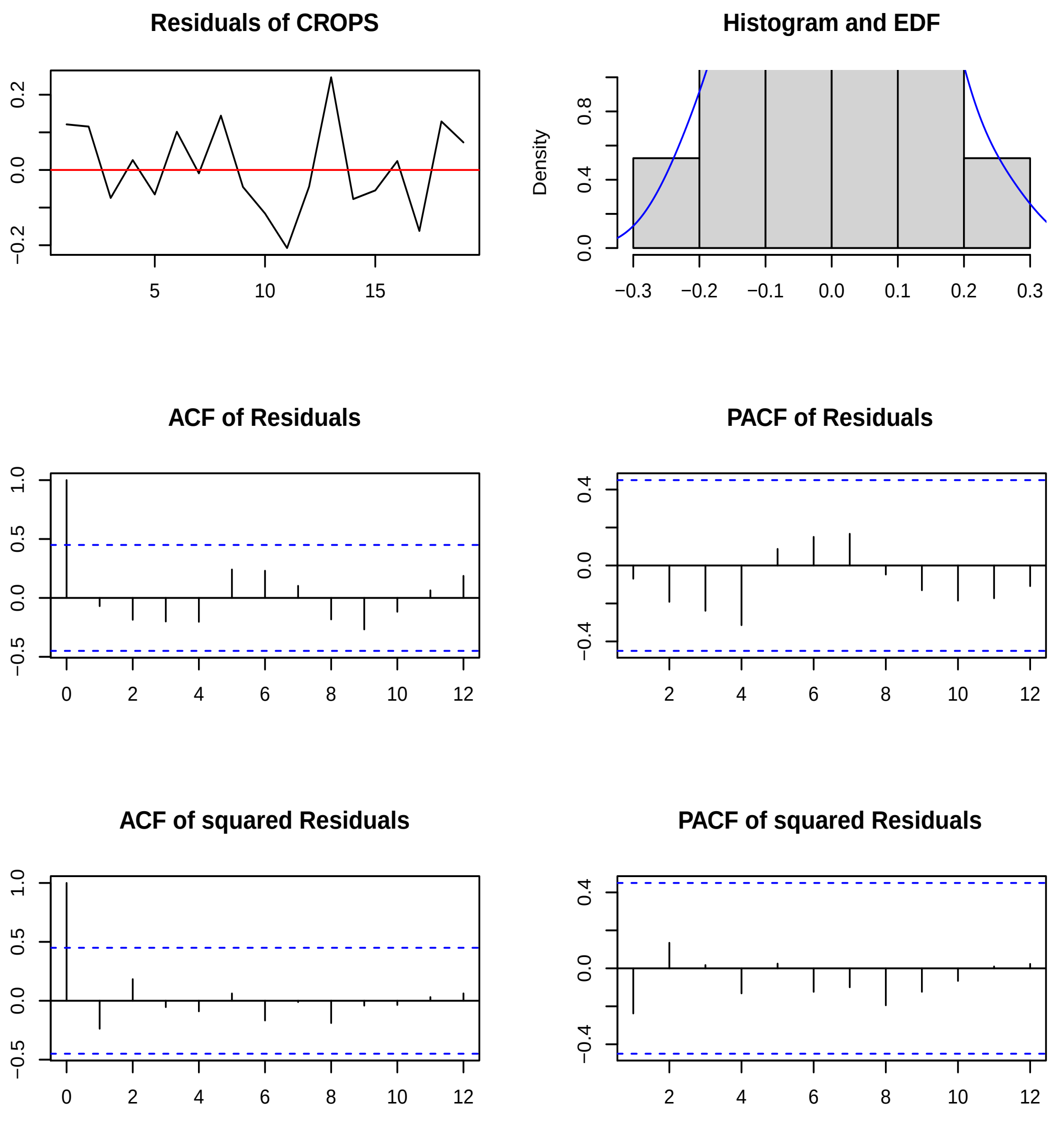


Figure 0.1: Residuals from Crops. The plot shows the residuals from the crops, the corresponding histogram, the autocorrelation function (ACF) plot of the residuals and the partial autocorrelation function (PACF) plot of the residuals. From the ACF plot, all the autocorrelation values are within the boundaries an indication that the residuals are uncorrelated.

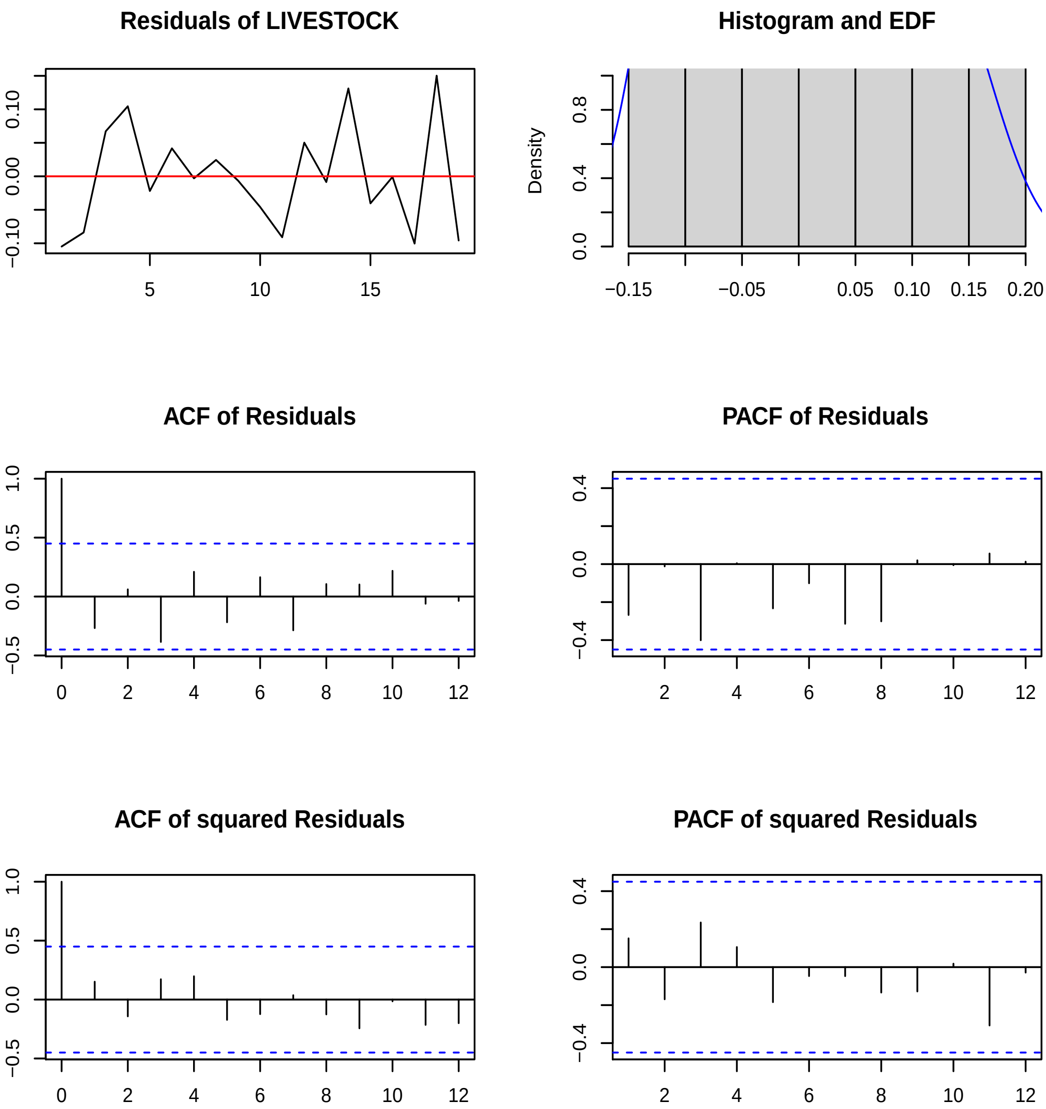


Figure 0.2: Residuals from Livestock. The plot shows the residuals from the livestock, the corresponding histogram, the autocorrelation function (ACF) plot of the residuals and the partial autocorrelation function (PACF) plot of the residuals. From the ACF plot, all the autocorrelation values are within the boundaries an indication that the residuals are uncorrelated.

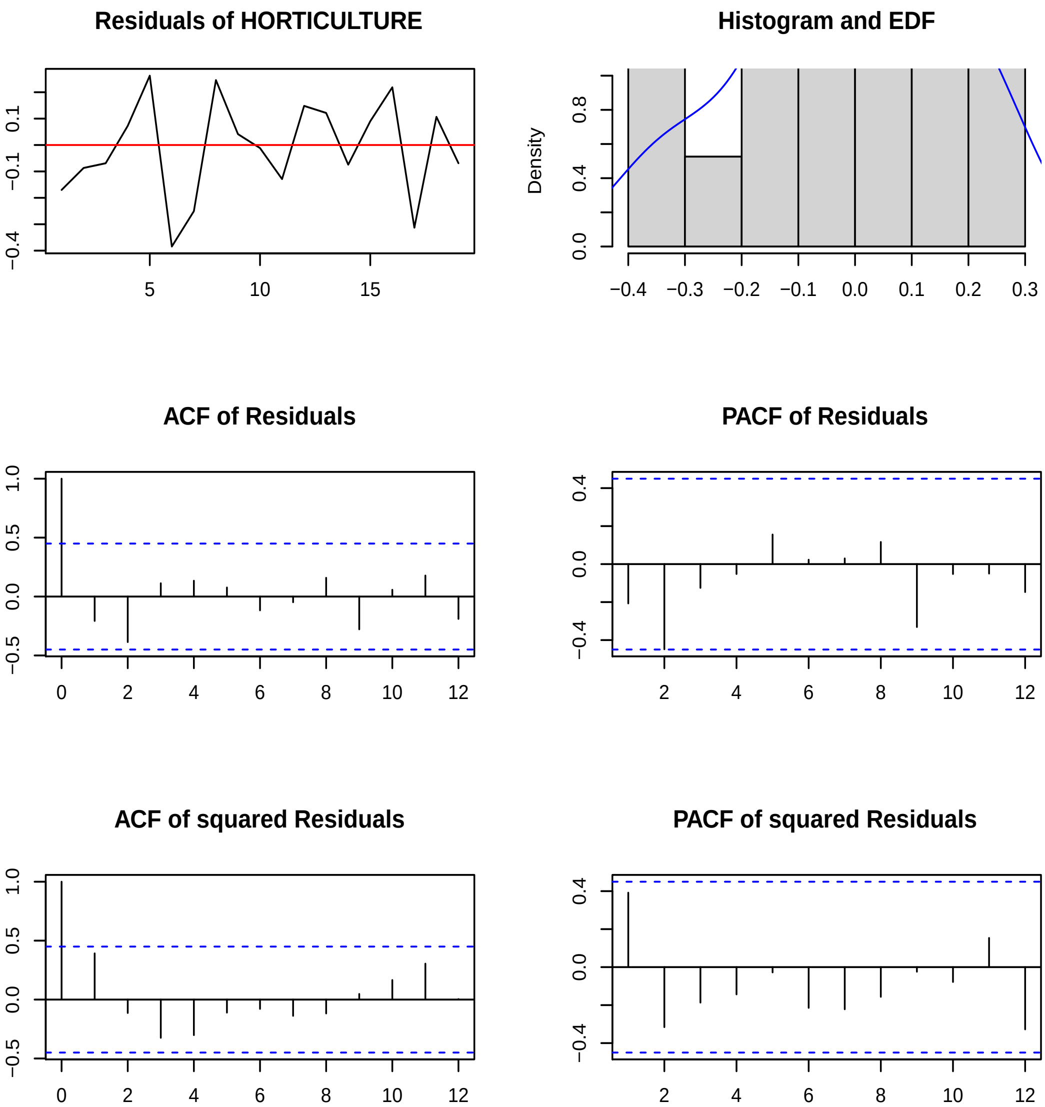


Figure 0.3: Residuals from Horticulture. The plot shows the residuals from the horticulture, the corresponding histogram, the autocorrelation function (ACF) plot of the residuals and the partial autocorrelation function (PACF) plot of the residuals. From the ACF plot, all the autocorrelation values are within the boundaries an indication that the residuals are uncorrelated.

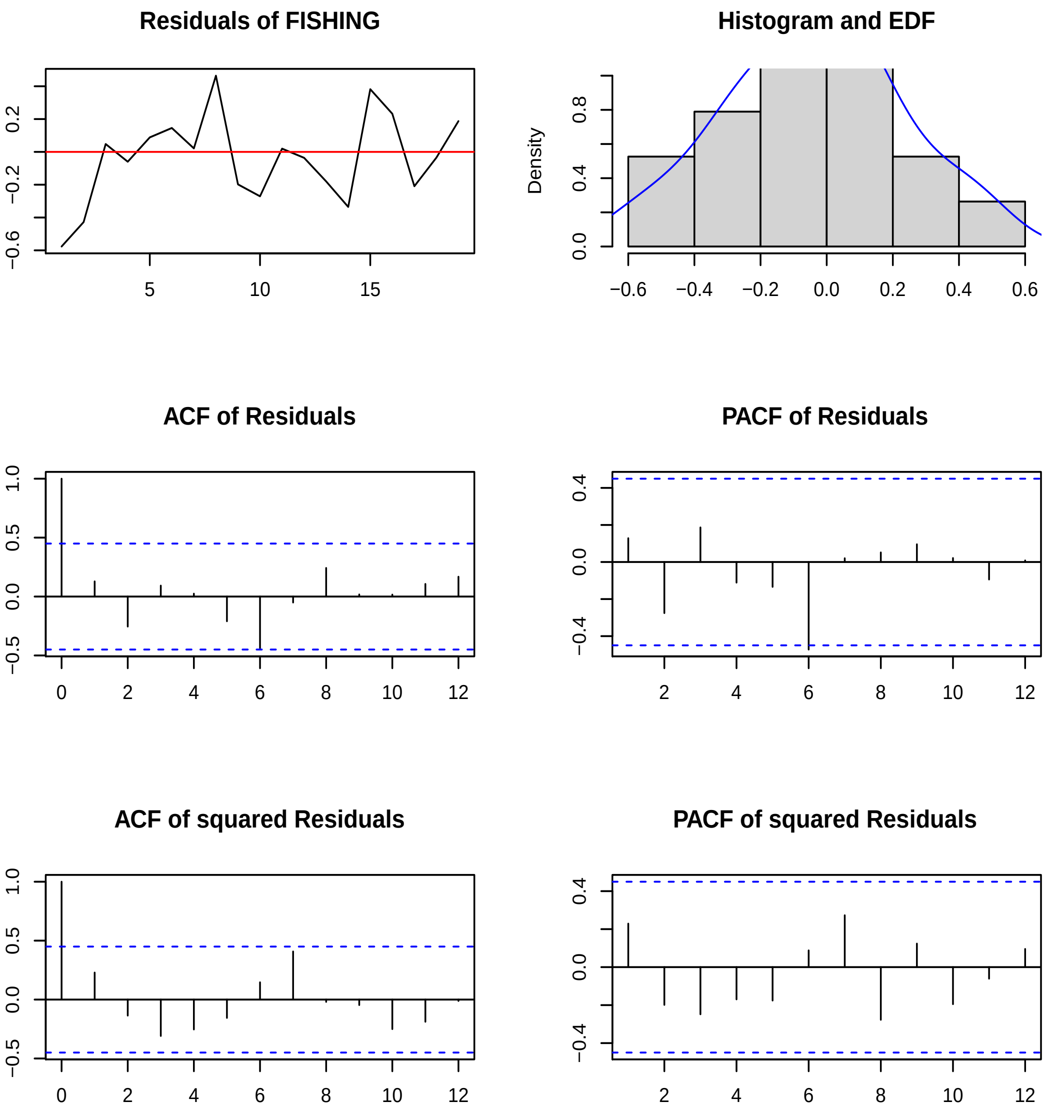


Figure 0.4: Residuals from Fishing. The plot shows the residuals from the fishing, the corresponding histogram, the autocorrelation function (ACF) plot of the residuals and the partial autocorrelation function (PACF) plot of the residuals. From the ACF plot, all the autocorrelation values are within the boundaries an indication that the residuals are uncorrelated.

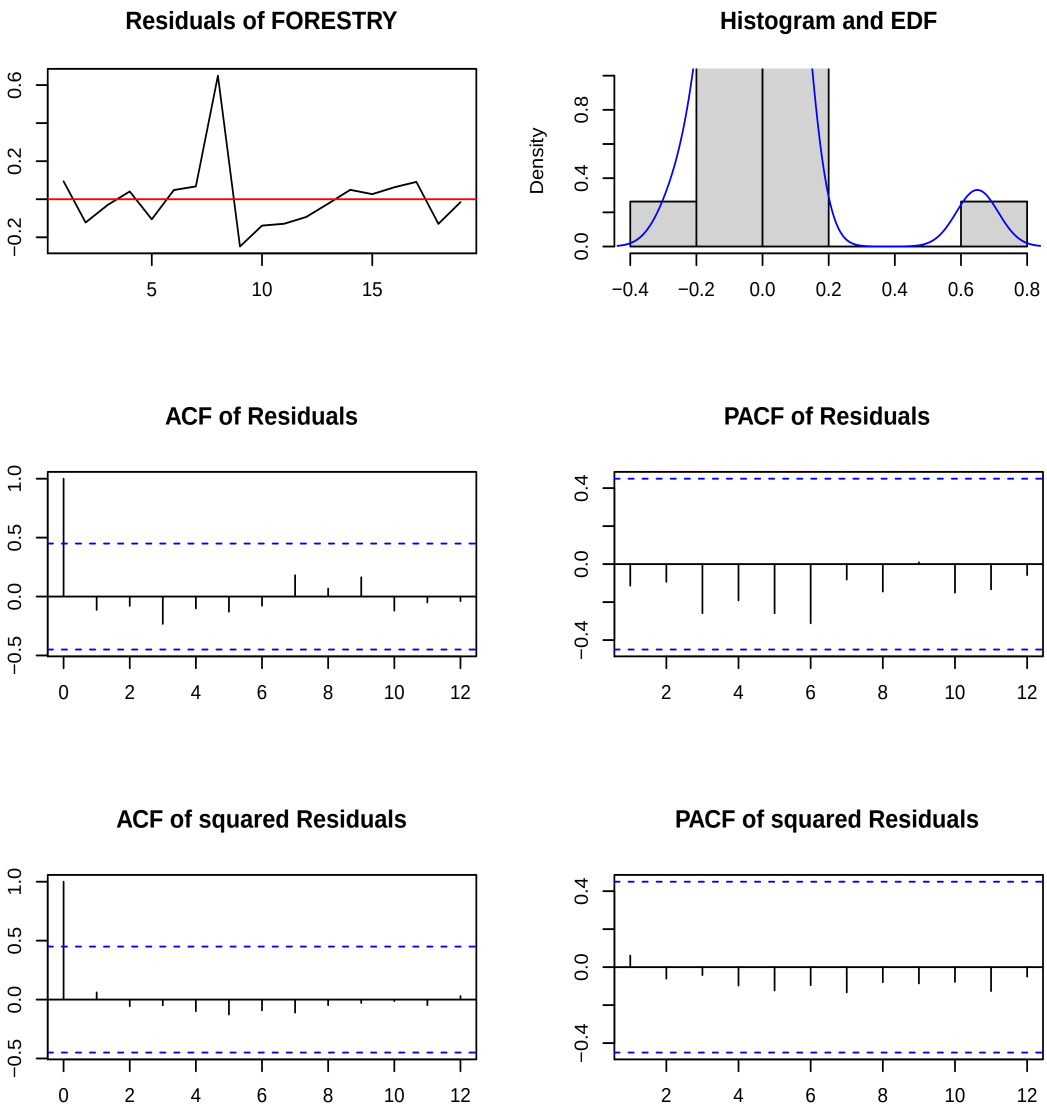


Figure 0.5: Residuals from Forestry. The plot shows the residuals from the forestry, the corresponding histogram, the autocorrelation function (ACF) plot of the residuals and the partial autocorrelation function (PACF) plot of the residuals. From the ACF plot, all the autocorrelation values are within the boundaries an indication that the residuals are uncorrelated.

(ii) Data

Year	Crops	Livestock	Horticulture	Fishing	Forestry
2000	63330.55	13910.61	13890	2027.2	10365
2001	63714.06	15467.78	20222	2248.1	10686
2002	57210.1	19339.78	26722	6323	11791
2003	56550.14	19432.46	28840	5739	13106
2004	70094.93	20766.7	32591	6403	14221
2005	71520.14	24258.92	38838.1	6313	15333
2006	80656.96	28795.62	43120.8	6679	16365
2007	82260.25	33490.76	67253.7	7127	16021
2008	91012.4	39969.15	57965.8	9450	16452
2009	104490.1	45049.95	49352.1	9903	18539
2010	141604.6	55260.16	56993	19111	39143
2011	162863	80899.53	88621.6	22999	46661
2012	167152.1	88670.7	89868	28902	58039
2013	159106.5	92827.4	83381.5	34315	67230
2014	151948.2	97989.2	84084.3	40387	73520
2015	182503.2	100071.4	90438.8	40300	79697
2016	186933.9	125401.9	101513.6	30595	94562
2017	195990.5	135600.5	115322.8	42687	102617
2018	197825.2	146822.3	153681.4	51034	125979
2019	172644.7	147879.2	144578.6	57572	156606
2020	194436.8	163037.1	150163.9	63112	172949
2021	207706	161619.6	157692.9	80715	199836

Figure 0.6: The secondary data that was used in the study. The data is yearly, 2000 to 2021, in million Kenyan shillings.