PAKISTAN GRAM PRODUCTION FORECASTING USING BAYESIAN TIME SERIES MODELING

By

ADIL FAIZAN

NATIONAL UNIVERSITY OF MODERN LANGUAGES,

ISLAMABAD.

December, 2022

Pakistan Gram Production Forecasting Using Bayesian Time Series Modeling

By

ADIL FAIZAN

BSIT, PMAS Arid Agriculture University, Rawalpindi, 2018.

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In **Software Engineering**

To

FACULTY OF ENGINEERING & COMPUTER SCIENCES

NATIONAL UNIVERSITY OF MODERN LANGUAGES ISLAMABAD

NATIONAL UNIVERSITY OF MODERN LANGUAGES FACULTY OF ENGINEERIING & COMPUTER SCIENCE

THESIS AND DEFENSE APPROVAL FORM

The undersigned certify that they have read the following thesis, examined the defense, are satisfied with overall exam performance, and recommend the thesis to the Faculty of Engineering and Computer Sciences for acceptance.

Thesis Title: **Pakistan Gram Production Forecasting Using Bayesian Time Series Modeling.**

Submitted by: Adil Faizan **Registration #:** 15 MSSE/lbd/S19

Master of Science in Software Engineering
Degree name in full

Software Engineering
Name of Discipline

Dr. Raheel Zafar
Name of Research Supervisor

Signature of Research Supervisor

Dr. Javvad Ur Rehman
Name of Research Co-Supervisor

Signature of Research Co-Supervisor

Dr. Basit Shahzad
Name of Dean (FE&CS)

Signature of Dean (FE&CS)

Prof. Dr. Aamir Ijaz
Name of Pro-Rector Academics

Signature of Pro-Rector Academics

December, 2022
Date

AUTHOR'S DECLARATION

I Adil Faizan

Son of Muhammad Munir

Registration # 15 MSSE/lbd/S19

Discipline Software Engineering

Candidate of **Master of Science in Software Engineering (MSSE)** at National University of Modern Languages do hereby declare that the thesis **Pakistan Gram Production Forecasting Using Bayesian Time Series Modeling** submitted through me in partial success of MSSE degree, is my unique work, and has now no longer been submitted or posted earlier. I additionally solemnly claim that it shall now no longer, with inside the future, be submitted through me for acquiring some other degree from this or some other college or institution. I additionally apprehend that if proof of plagiarism is determined in my thesis/dissertation at any stage, even after the award of a degree, the printings can be canceled and the degree revoked.

Signature of Candidate

Adil Faizan

Name of Candidate

December, 2022

Date

ABSTRACT

Pakistan Gram Production Forecasting Using Bayesian Time Series Modeling

The Bayesian approach/statistics, is a statistical decision approach that provides a tool for combining prior probabilities and their distribution about the nature of states. It provides tool to the people to modernize their views in the indication of fresh improved record or data. When working with such issue along time series models is that they too fit commonly when estimating models have large numbers of attributes above somewhat short length/time periods. In our case, this is not such a problem but possibly be when eyeing many attributes, these are common quite in economic prediction. One explanation to over fit problem is using a Bayesian approach, which opens a way to enforce specific priors on attributes. The aim of this thesis is to forecast the production of Gram, which include different attributes like gram cultivation area, production of Gram, the yield of a gram, the cost and prices of gram. For time collection or series data, ARIMA based state space modeling is used to forecast different future attributes of rabbi food crops of Pakistan including gram.

TABLE OF CONTENTS

LIST OF FIGURES

Figure NO. **TITLE** **PAGE**

LIST OF ABBRIVATIONS

SSM	State Space Model
SEM	State Equation Model
MEM	Measurement Equation Model
AR	Auto Regression
MA	Moving Average
ARMA	Auto Regressive Moving Average
ARIMA	Auto Regressive Integrated Moving Average
SARIMA	Seasonal Auto Regressive Integrated Moving Average
SMA	Simple Moving Average
VAR	Vector Auto Regression
VARMA	Vector Auto Regression Moving Average
ES	Exponential Smoothing
NN	Neural Network
BSS	Bayesian State Space

ACKNOWLEDGEMENTS

I would like to acknowledge and give my warmest thanks to my supervisor (Dr. Raheel Zafar) and co-supervisor (Dr. Javvad Ur Rehman) who made this work possible. Their guidance and advice carried me through all the stages of doing my research.

I would also like to give special thanks to my parents and my family as a whole for their continuous support and understanding when undertaking and writing my research. Your prayer for me was what sustained me this far.

Finally, I would like to thank Allah, for letting me through all the difficulties. I have experienced Allah's guidance day by day. Allah has let me finish my degree. I would keep on trusting Allah my whole life.

DEDICATION

I dedicate this work to God Almighty our originator, robust pillar, cause of encouragement, intelligence, knowledge and understanding. He has been the source of my strength for the duration of this program. I additionally dedicate this thesis to my loved mother and father who have supported me a lot. I dedicate my research work to my own circle of relatives and my teachers. A unique feeling of gratitude to my loving mother, father and teachers.

I dedicate this work and give special thanks to my inspiration my supervisor Dr. Raheel Zafar and co-supervisor Dr. Javvad Ur Rehman for being there for me throughout the entire Master's program.

CHAPTER 1

INTRODUCTION

1.1 Overview

Agriculture data can be used to understand crop production and to cater to the growing number of people in the world. Agriculture Statistics are useful for planning, monitoring, evaluation, and research & development purpose. Agriculture data makes it possible to achieve supply chain efficiency by offering tracking and optimization opportunities. Agriculture data-based decisions can improve resource utilization and conservation practices.

Time Series Data, time series data is a stepwise pool of statistical points, calculated normally over the sequential interval of time. In Linear System, if the difference between the outputs of the equation is consistent when using unknown variables, then the system is linear. Nonlinear System, if the difference between the outputs of the equation is inconsistent when using unknown variables, then the system is nonlinear.

Prediction is a procedure of making a decision based upon present and past data for the future and commonly by analyzing the trends. In this study, for the time-series data, Bayesian State Space Modeling is used for the prediction of Pakistan Rabi Crops Production. Thomas Bayes in the 1770s introduced the 'Bayes Theorem' [1]. Till-day, Bayesian Statistics importance hasn't gone. The Bayesian approach in general is a statistical decision approach that provides a tool for combining prior probabilities and their distribution about the nature of states. It provides a tool to modernize their opinions in the indication of new improved data. When working with such issues along time series models is that they too fit commonly when estimating models have large numbers of attributes above somewhat short length periods. When eyeing multiple attributes, these are common quite in economic prediction. One explanation for over fit problem is using a Bayesian approach, which opens a way for us to enforce specific priors on our attributes.

For Time collection or series data, State Space Modeling is used for forecasting different future attributes. The Bayesian approach is a statistical technique that is used for attribute selection, now-casting, time series forecasting, and other applications [1]. The benefits of Bayesian forecasting are the easiness, arranging, and understanding provided by the parameterization demonstration of a method are radical.

In commercial situations, the Bayesian approach is natural not only do people take an interest to send/receiving cumulative probabilities on parameters, but people are concerned about making decisions in ambiguity, which adds multiple beliefs regarding the upcoming with significant costs of choices taken for likely future consequences. It is stressful for the person ambiguity, through which the upcoming is still uncertain as of the future casual term but also the ambiguity on the values of the present parameters and the ability of the function to link present to the future. [2]

Now a day's intensive computer techniques have been developed, and the State Space model [3] is used for prediction in this thesis for the prediction of Gram. The state-space model is used for two main types of systems. Linear and nonlinear systems. An SSM state space model comprises (I) A SEM state equation model, which distributes undelaying time-based difference dynamic. (II) A MEM measurement equation model, that narrates the observations to SEM State Equation Model variables through noise issues [3].

1.2 Time Series

A time collection is a stepwise collection of statistical points, calculated normally over sequential times. Time series evaluation contains strategies for examining time series data with the purpose to mine significant records and different features of the data.

Time collection forecasting is the technique for analyzing time-series information at the stage of data and exhibiting to make forecasting and state strategic decisions makings. Time collection examination involves growing functions to benefit an understanding of the information to recognize the fundamental causes. The examination can explain the why at the back of the results you're sighting. Forecasting then takes the subsequent step of what to do with that data and the predictable extrapolations of what would possibly occur in the future.

Naturally, there are boundaries when managing the unpredictable and the unknown. Time series prediction is not dependable and is not suitable or beneficial for all states. Because there surely isn't any specific set of guidelines for the weather you have to or should no longer use forecasting, it miles as much up to the analysts and data groups to realize the limits of the analysis and what their functions can support. Time series evaluation indicates in what way data modifications over some time and accurate forecasting can pick out the course in which the record is altering.

1.3 Types of Domain Approach

There are two main types of Approaches, Time domain, and frequency domain approach.

1.3.1 **Time-domain approach:**

How does what occur these days affect what is going to occur in the future?

These tactics sight the research of lagged associations as maximum significant, e.g. autocorrelation analysis.

1.3.2 **Frequency-domain approach:**

What is the economic cycle through an interval of growth and recession? These procedures sight the research of cycles as maximum significant, e.g. spectral evaluation and wavelet evaluation.

The time-domain graph presents the adjustments in a signal over some time, and the frequency domain presents how plenty of the signal exists inside a given frequency band regarding a variety of frequencies.

1.4 Types of Data

There are two main types of datasets, micro dataset, and macro dataset.

1.4.1 Micro Datasets:

Micro record/data, that’s person-specific. It might not translate to a bigger scope directly, however together turns into a part of a macro data point. And micro datasets inclusive of the European Social Survey. Microdata examples (constructing blocks): A person's particulars (first name, last name, age, e-mail, location, etc.), each person’s preferences (SMS vs e-mail), each person's channel preference (Facebook-messenger vs e mail).

1.4.2 Macro Datasets:

Macro record/data is precious to a big-image degree however is not person-specific. Macro databanks consisting of the ones produced via way of means of World Bank and the United Nations, Macro record/data examples (things we build): A cluster of our consumers with e-mails that quit with @gmail.com, a series of orders wherein people requested for a spicy meal, a desk of clients that opt for Facebook-messenger over different channels, the pool of our transactions that use credit source paying cards as opposed to PayPal.

1.5 Components of Time Series

There are numerous components in time series collections, some of them are explained below:

1.5.1 Univariate vs Multivariate:

A time collection comprising statistics of one variable is named univariate, however, if statistics of a couple of variables are taken into consideration then it miles termed multivariate.

1.5.2 Linear vs Non-linear:

A time collection function is stated to be linear or non-linear relaying on whether or not the present value of the collection is a linear function or nonlinear function of previous observations.

1.5.3 Discrete vs Continuous:

In a continuous-time collection, observations are calculated at each occurrence of time, while a discrete-time collection includes observations calculated at a discrete occasion in time. Generally, a time collection is suffering from four (4) components or elements, i.e. trend, seasonal, cyclical, and irregular components.

1.6 Trend

The universal tendency of a time collection to upsurge, down-surge, or stagnate over a protracted length of time.

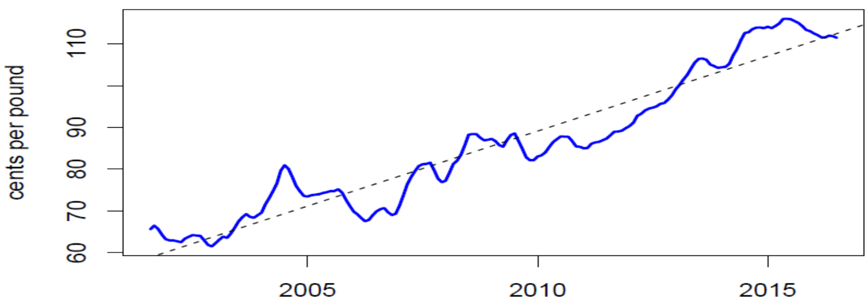


Figure 1. 1: Trend Representation [19]

1.6.1 Seasonal Variation:

This element clarifies variations inside a year all through the season, typically due to weather and climate conditions, conventional habits, customs, etc.

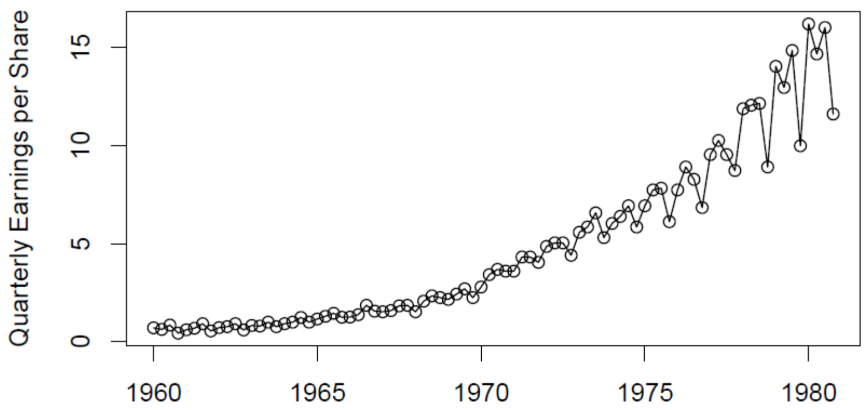


Figure 1. 2: Seasonal variation representation [19]

1.6.2 Cyclical variation:

This element describes the medium-time period modifications as a result of situations, that repeat in cycles. The length of a cycle ranges over an extended duration of time.

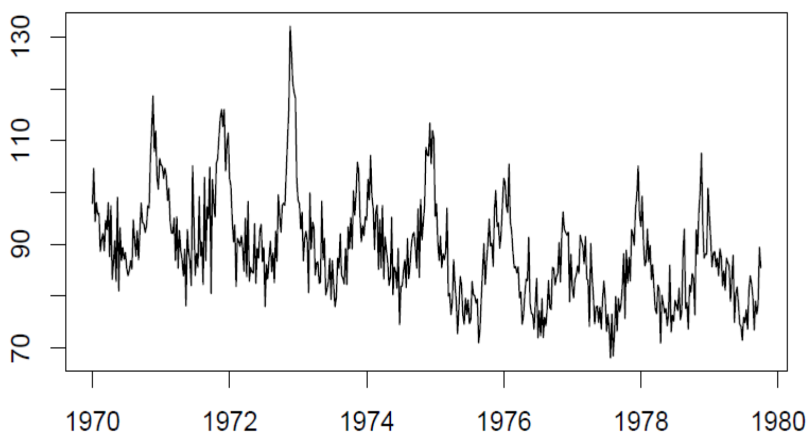


Figure 1. 3: Cyclical variation representation [19]

1.6.3 Irregular variation:

Irregular or random distortions in a time collection are resulting from irregular effects, which aren't ordinary and additionally do now no longer repeated in a specific pattern. These variations are resulting from incidences together with earthquakes, war, floods, strikes, revolution, etc.

There isn't any described statistical approach for measuring random fluctuations in a time collection.

1.7 Noises in Time Series

There are two kinds of noises in time series, White noise, and Gaussian White noise.

1.7.1 White Noise:

An easy time collection potencially be a set of uncorrelated casual variables, $\{w_t\}$, $\mu = 0$ with mean $= 0$ and σ_w^2 finite variance denoted such as $W_t \sim w_n\ (0;\ \sigma_w^2)$.

1.7.2 Gaussian White Noise:

A specific beneficial white noise is Gaussian white noise, in which the independent regular casual variables W_t , with mean $= 0$ and variance$= \sigma_w^2$, are denoted as $W_t \sim iid\ N\ (0;\ \sigma_w^2)$.

1.8 Personal Motivation

Data mining can be seen as the result of natural advances in data technology. The database system market has supported the following evolutionary directions in the development of capabilities such as data acquisition, database creation, data management and advanced data analytics.

Data mining has received a great deal of attention in the information market and society at large in recent years due to the vast amount of data widely available and the need to transform that data into useful data and knowledge. The information and insights obtained can be used for a variety of software, from industry analysis, fraud detection, prediction purposes and user engagement to production control and scientific research.

1.9 Why Bayesian Modeling?

Bayesian strategies are important when researchers don’t have a lot of information. By using a robust prior, a researcher could make affordable estimates from as low as one information point. Their end result is a chance distribution, in preference to a point estimate. Using an informative previous permit assuaging the various problems that plague classical significance testing. The rigorous technique to deal with statistical estimation problems. The Bayesian viewpoint is developed and dominant. Even in case you are not Bayesian, the Researcher might outline an “uninformative” prior and whole lot reduces to most likelihood estimation. It’s simple to impose constraining knowledge (as priors). It’s simple to mix information from exception record sources. Recursive estimators originate logically. You're posteriorly calculated at time t-1 due to the previous for time t. This is mixed with the likelihood at time t, and renormalized to get the posterior at time t. This new posterior turns into the prior for time t+1, and so on... Bayesian techniques are crucial while you don’t have a great deal of record/data. With the usage of a sturdy prior, you may make affordable estimates from as low as one record point.

1.10 Problem Statement

Increasing the accuracy of agricultural forecasting is always challenging, as agricultural output depends on many factors. There is constantly a lack to broaden advanced technique of mixing data from numerous sources of forecasting record in phases of the timing of the forecasts, use of the statistical technique as opposed to non-statistical sampling, and use of various listing areas [16].

Besides ARIMA models [17], other available methods have not been exploited in-depth, such as Bayesian State-Space modeling [18] or other models. Similarly, the use of big data analytics is expanding for agriculture forecasts. There is a need to review ARIMA-based Bayesian approaches and methods for the forecast of Rabi crops, which are major agriculture crops of Pakistan.

1.11 Research Questions

The accuracy of agricultural forecasting is always important as agricultural output depends on many factors. In this thesis, the Bayesian prediction model is used for the Production of Gram in Pakistan.

RQ 1: How algorithms and models combine to implement state-space modeling for gram production forecasting?

RQ 2: How to integrate the Bayesian parameter learning technique for the estimation of ARIMA unknown parameters?

1.12 Aim of Research

This thesis aims to understand the available data and generate and implement ARIMA-based Bayesian state-space modeling for that available dataset. Selection of feasible algorithms to model and introduce our model and codes for them in MATLAB for simulations and are expecting good simulated forecasted results for Rabi agriculture crops in Gram, which are important crops of Pakistan.

1.13 Research Objectives

The key objectives of current thesis are:

- To select and integrate appropriate algorithms and models to implement state-space modeling for Pakistan gram production forecasting.
- To incorporate the Bayesian parameter learning method for estimation of ARIMA unknown parameters.

1.14 Scope of Research Work

This thesis is covering the forecasting of Pakistan agriculture Rabi crops, which are wheat, potato, and gram. This thesis will forecast the production of crops in thousand tons, the yield of a crop in kilograms per hectare, the area of crops (kgs), the cost of production, and wholesale prices (tons), this thesis is using ARIMA-based Bayesian time series analysis [2] for the forecast of the Pakistani agriculture crop 'Gram', State Space Modeling is used for forecasting. This study is dealing with different variables related to the Rabi Crops which are prices, Cultivation area, production, and yield of crops.

1.15 Thesis Organization

The rest of the thesis is organized as follows:

Chapter 2 provides a literature review.

Chapter 3 will present the methodology.

Chapter 4 will provide Results and Analysis**.**

Chapter 5 will give a conclusion, future work, and summery

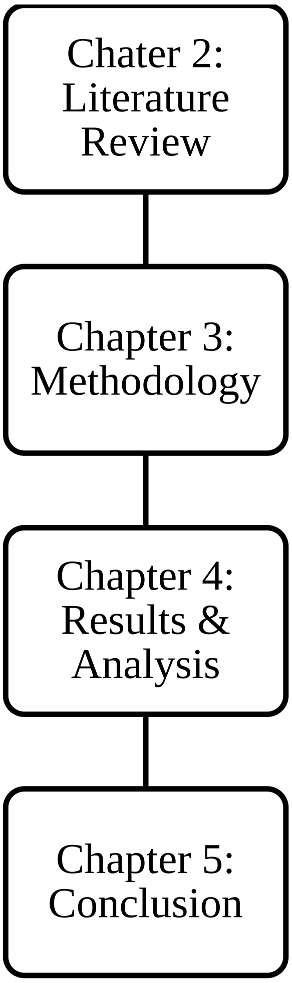


Figure 1. 4: Thesis Organization

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

This chapter covers material associated with literature, in the first section of this chapter, time series and their components are provided, then in the second phase different data sets of Rabi crops are discussed, and their datasets are also time series data. Then in the third phase, regressive models are discussed which include: autoregressive models AR, MA, SMA, ARMA, ARIMA, SARIMA, NN, VAR, VARMA, ES, Bayesian" in the third phase. Then in the fourth phase applications of other autoregressive models are discussed and references are provided in this section where those autoregressive models are used before and in the last or fifth phase Summary of Chapter 02 is provided.

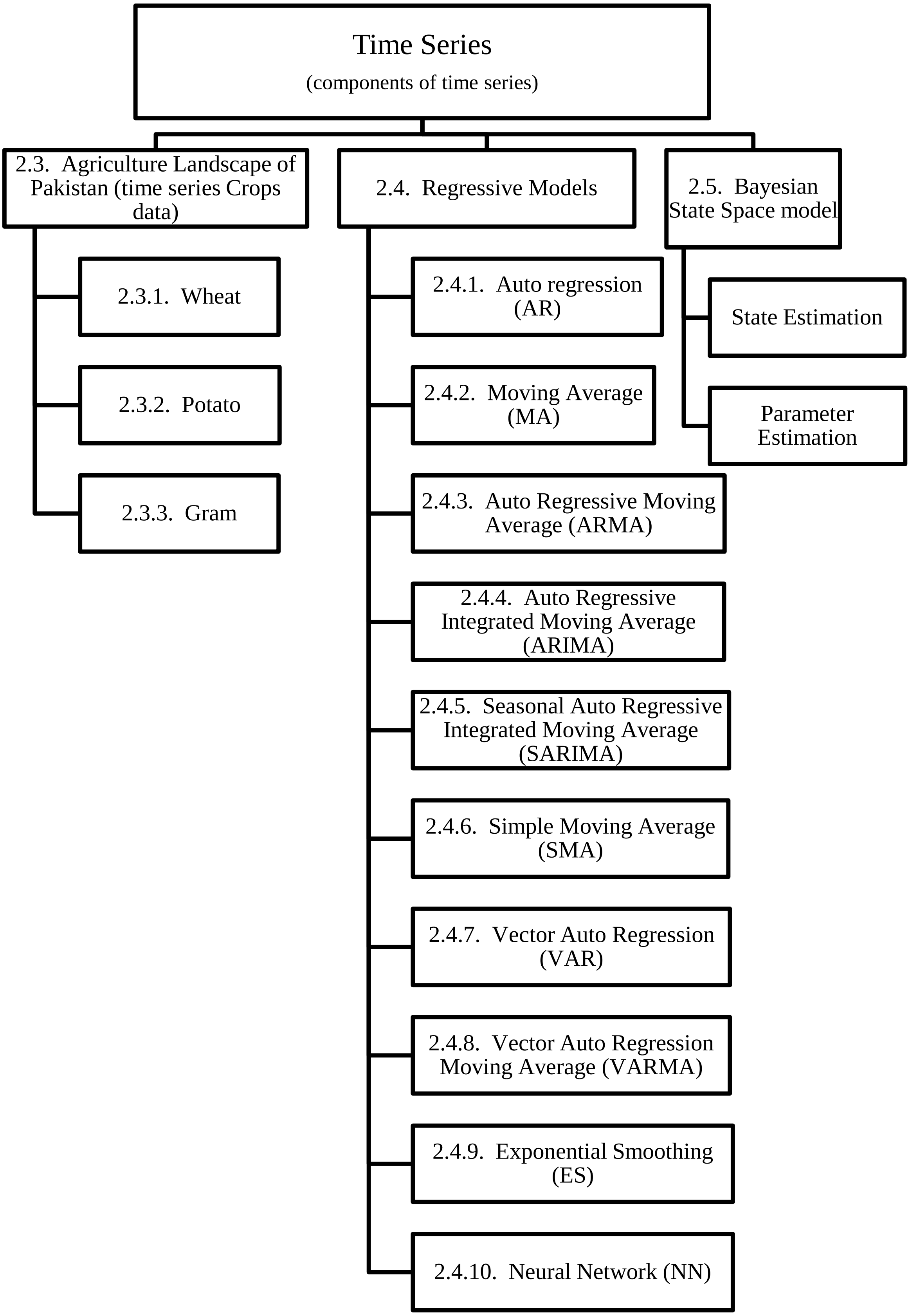


Figure 2. 1: Overview Diagram of chapter 02

2.2 An Overview of Agricultural Landscape in Pakistan

Agricultural improvement is one of the vital approaches to eliminating severe poverty, enhancing collective prosperity, and feeding a expected 9.7 billion humans within side the globe by 2050. The agriculture segment boom is two to four instances greater powerful in growing earnings of the poorest in compression to different sectors, because the World Bank estimations, sixty-five percent of poor make a residing over agriculture worldwide. Agriculture is likewise crucial to the financial boom and responsible for 1/3 of the world's GDP.

In 2018, countries such as Russia, Australia, Argentina and South Africa recorded negative growth in their agricultural sectors (Figure 2.2). South Asian countries such as Pakistan, India, and Bangladesh play an important role in the growth of domestic and global agriculture. These South Asian countries have the potential to significantly increase agricultural production, but face serious challenges of climate change and extreme weather events.

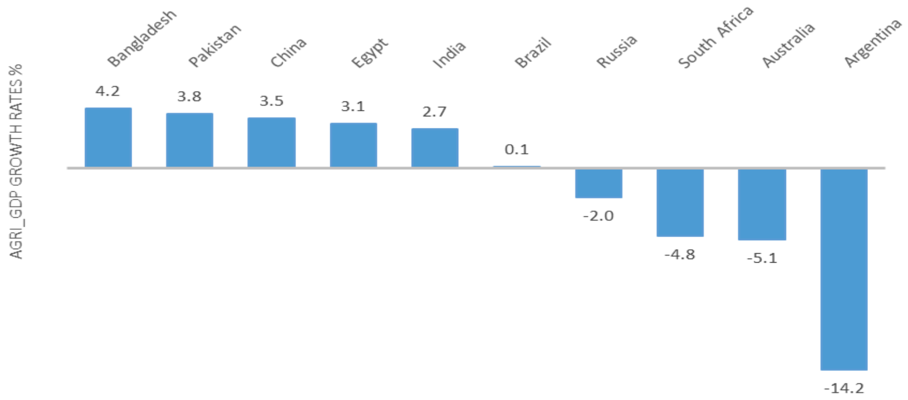


Figure 2. 2: Global growth in agriculture GDP (2018)

Pakistan has achieved significant increases in agricultural output over the past two decades (Figure 2.2), but the country faces a range of shocks in the future due to climate patterns, economic conditions and agricultural policies. The negative growth of Pakistan's agricultural sector in 2018-2019 may be partly due to unusual weather patterns. At the same time, the economic situation in the country is also putting pressure on the sector due to the devaluation of the currency and rising prices of inputs such as fertilizers and fuels.

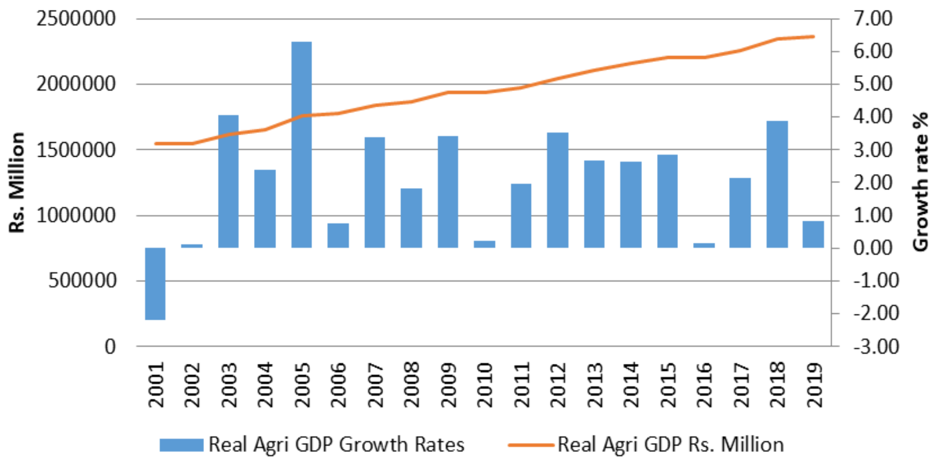


Figure 2. 3: Agriculture sector growth of Pakistan [20]

Figure 2.3 shows a substantial increase in agricultural employment in Pakistan. Agricultural employment increased from 16.7 million in 2002 to 23.7 million in 2018.

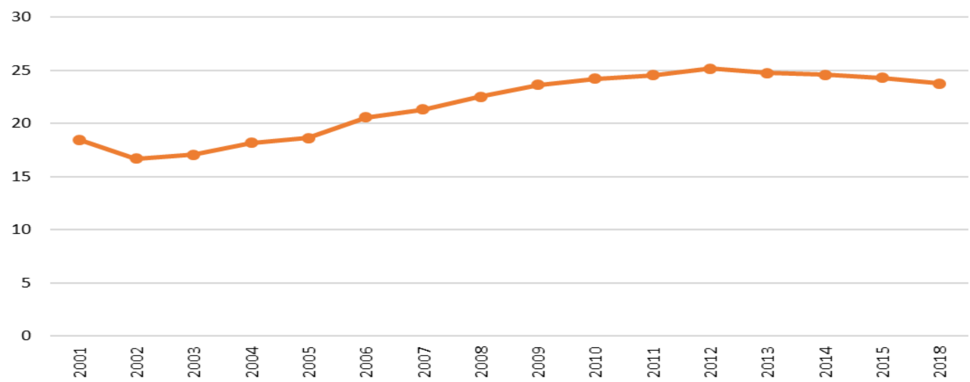


Figure 2. 4: Trends of employment in the agriculture sector-number of employed workers (in millions) [21]

Figure 2.4 shows the contribution of men and women agricultural workers in Pakistan. Women make a significant contribution to total economically active manpower in the agriculture sector as shown in figure 2.5. Men farmers are normally involved in crop production, water management, crop protection, transportation, and output marketing. Women actively contribute to a range of agricultural activities such as preparing manure, sowing plants, weeding, harvesting crops/vegetables, collecting fodder, growing vegetables, fruit processing, packing, and raising poultry and livestock.

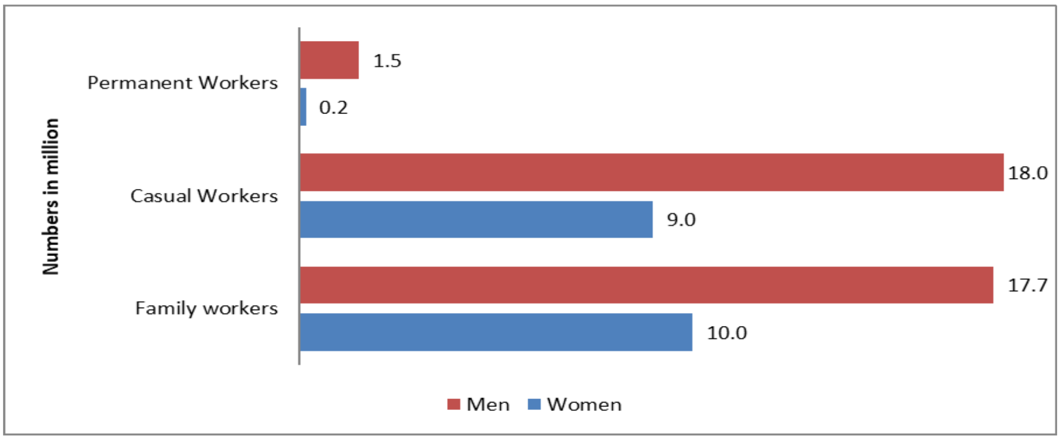


Figure 2. 5: Classification of agriculture workers [20],Landholding, Farm Size and Distribution of Farms in Pakistan.

Figure 2.6 shows that in the Rabi season, wheat occupies about 72 percent of the total cultivated area followed by the gram, fodder, and other crops. The area is under high value and nutritious crops such as vegetables and fruits are small. The crop sector in the Rabi season is not ecologically feasible, as it reveals less diversification and hence limited biodiversity and associated ecosystem services. The crop rotation generally relies on wheat such as wheat-rice, wheat cotton, and others. The rice-wheat rotation system is generally considered the main hurdle to improving crop yield. The rice-wheat crop rotation system also leads to high salinity and soil erosion. Hence, for sustainable agriculture development, there is a need for a highly diversified cropping system in Pakistan.

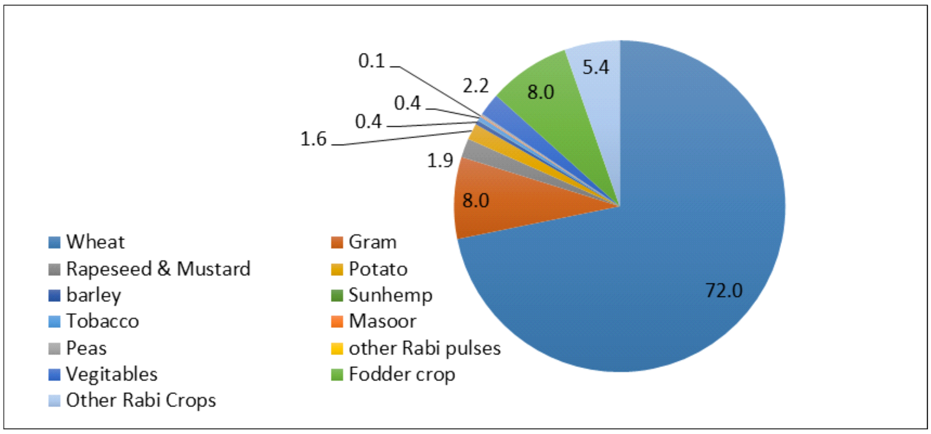


Figure 2. 6: Cultivated area under different crops in Rabi 2017-18 [20]

2.2.1 Wheat

Wheat is the main Rabi crop and staple food in Pakistan. It's the maximum vital crop, grown by 80% of farmers on about 40% of general cultivated land (72 percent of

Rabi cultivated area) in all four provinces of Pakistan. Wheat is responsible for 9.1 percent of the value brought in agriculture and 1.7 percent of GDP. It’s also the most important crop for achieving food security in the country, as it accounts for about 37 percent of both food energy and protein intake.

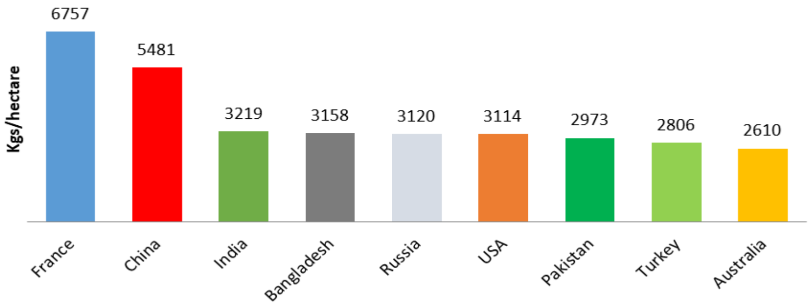


Figure 2. 7: Wheat yield of major wheat-producing countries of the world [22]

Figures 2.8, to 2.12 shows the trends in the area, yield, and production of wheat at the nationwide and provincial stages. The area under wheat has enlarged from 6.98 million hectares in 1981 to 8.66 million hectares in 2019, an increase of around 20 percent overall and 0.6 percent per annum. However, since 2008, there has only been a marginal increase in area, with high fluctuations between 8.45 to 9.2 million hectares, largely due to a leveling of the area expansion in Punjab. The largest wheat-producing province is Punjab, with 6.49 million hectares or 75 percent of the total wheat cropped area. Sindh follows with 1.05 million hectares (12.1 percent of the total), Khyber Pakhtunkhwa with 0.72 million hectares (8.4 percent), and Baluchistan with 0.39 million hectares (4.5 percent).

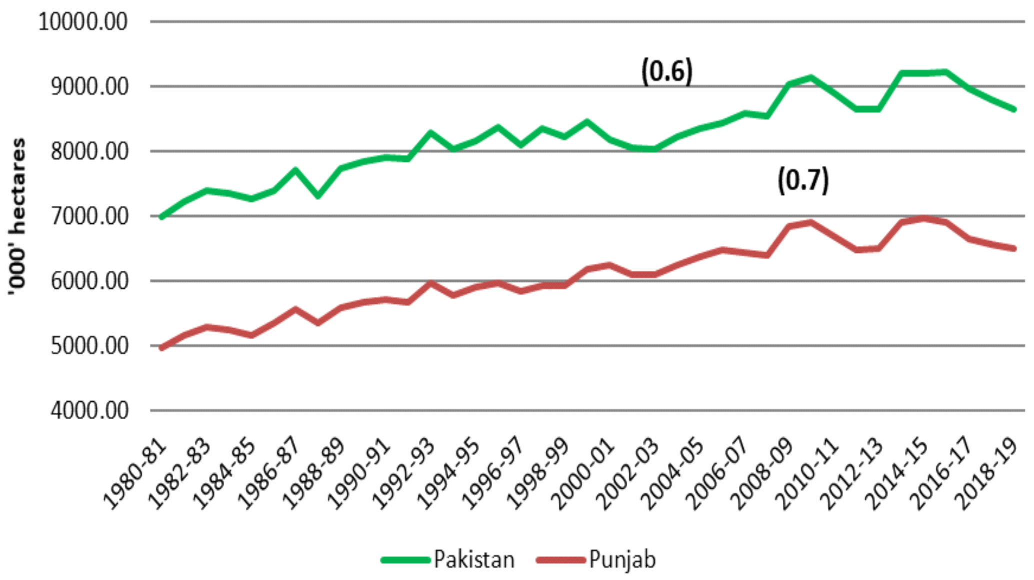


Figure 2. 8: Cultivated area of wheat in 000 hectares in Pakistan and Punjab [20]

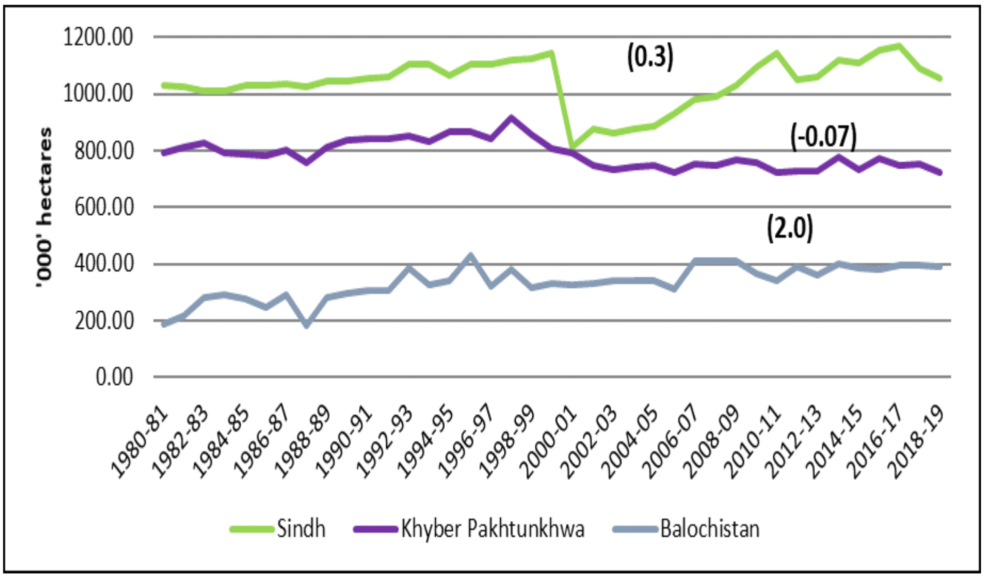


Figure 2. 9: Province-wise cultivated area of wheat in thousand hectares [20]

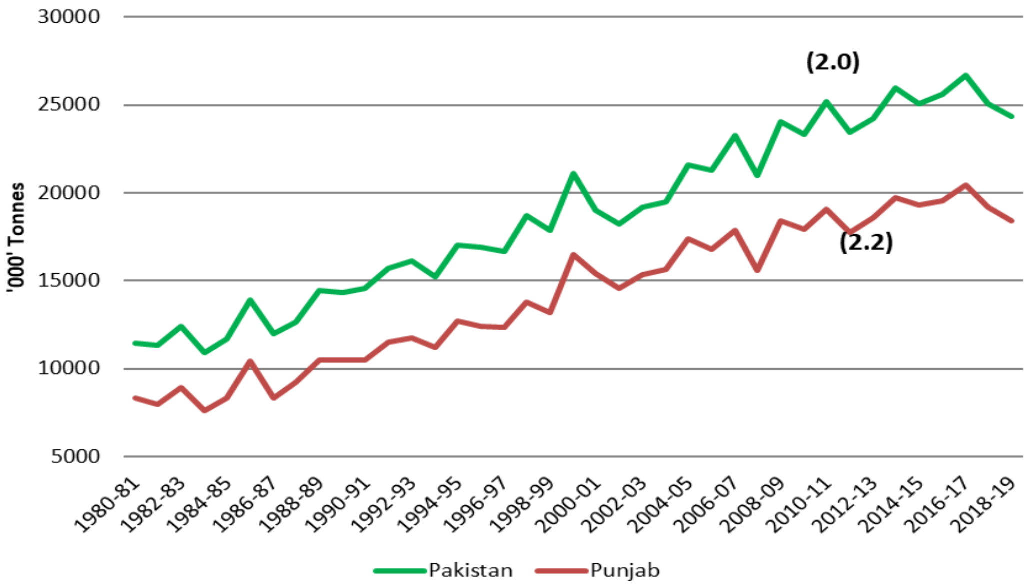


Figure 2. 10: Wheat production in thousand tones in Pakistan and Punjab [20]

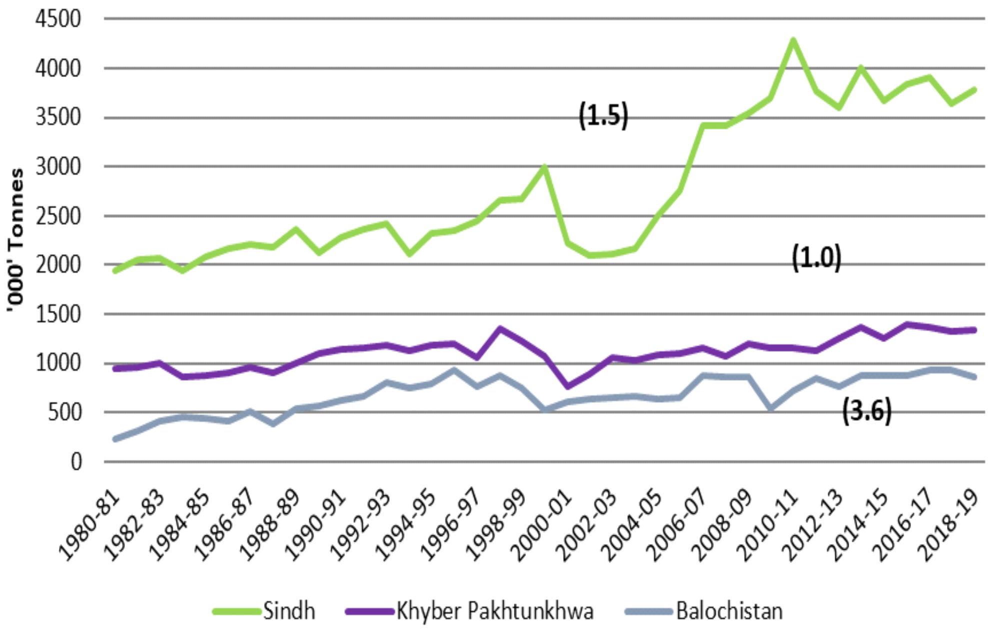


Figure 2. 11: Province-wise wheat production in thousand tones [20]

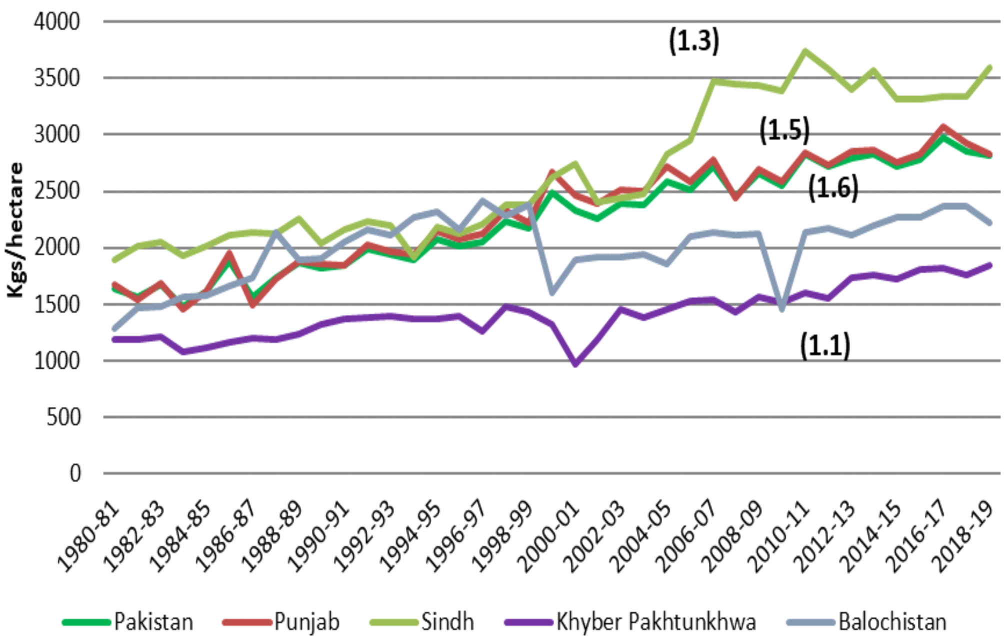


Figure 2. 12: Yield of wheat in kgs per hectare in Pakistan and provinces [20]

Figure 2.13 shows that the wheat prices in Pakistan are less volatile than international prices. This is due to Government interventions through price stabilization schemes. As a result, wholesale and support prices of wheat have been higher than international prices during the last five years, which has raised Government stocks and even, in some years, to more than 20 percent of total production.

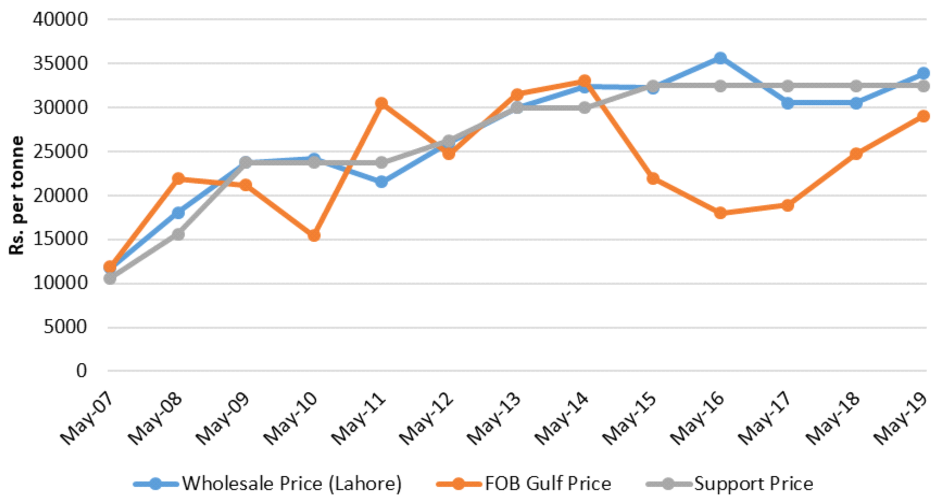


Figure 2. 13: International and local prices of wheat [20]

Figure 2.14 shows the wholesale prices in different markets of Pakistan. The government only purchases around 25 percent of total wheat production, while the remaining wheat is either consumed by farmers or sold in the market.

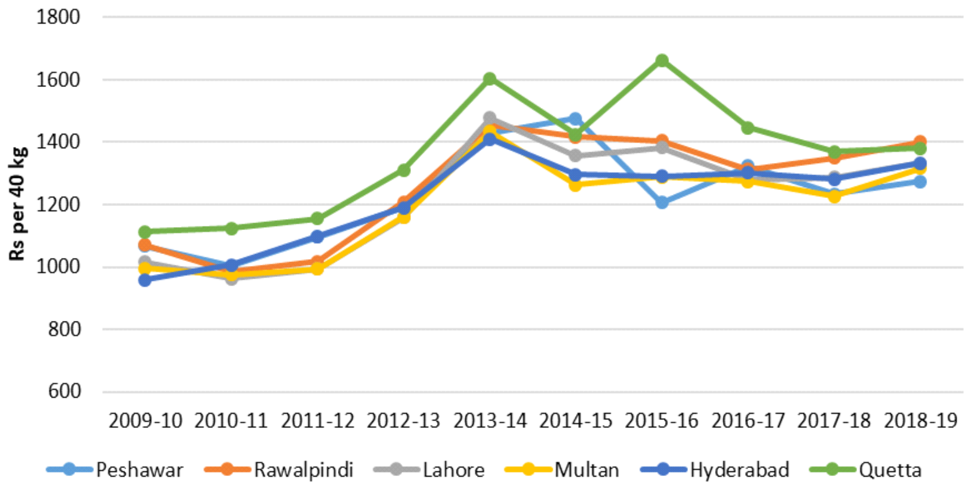


Figure 2. 14: Wholesale wheat prices in major markets [20]

Figure 2.15 shows the cost of production compared to support prices in the Punjab province, which has been variable over the past decade. The cost of production showed an increasing trend up to 2014-15, but it decreased after 2014-15, mainly due to a decrease in input prices and the stabilization of the currency.

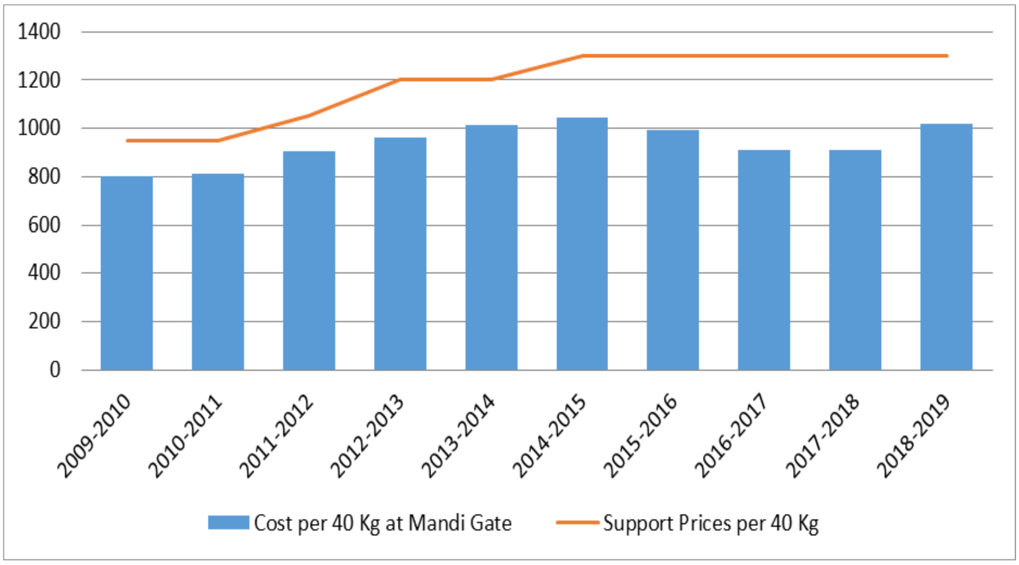


Figure 2. 15: Cost of production and prices of wheat (40kg) in Rs, [20]

2.2.2 Potato

Potato is an important crop for both farmers and consumers in Pakistan. It’s the fourth most consumed crop in the world after corn, rice, and wheat. Potatoes have the potential to be one of the most important cash crops in Pakistan for future food security and the generation of revenue through exports. Pakistan with 4.87 million tons contributes a small share to global potato production.

The per capita consumption and availability are very small as compared to other countries. The per capita production was 18 kgs in 2017 as compared to 69 kgs/capita/annum in China, 36 kgs/capita/annum in India, 43 kgs/capita/annum in South Africa and 62 kg/capita/annum in USA. The average consumption of potatoes remained between 12-15 kgs in the last decade which is much lower than the world average consumption of more than 35kg.

In terms of the average yield of potatoes, Pakistan is comparatively better than China, India, and Russia. However, it’s much less than the average yield of the USA, France, Australia, and Turkey (Figure 2.16). It’s evident from different surveys that some progressive farmers achieved a comparable yield of more than 40,000 kg/hectare in different districts of Punjab. Therefore, there is a huge potential for yield growth in Pakistan.

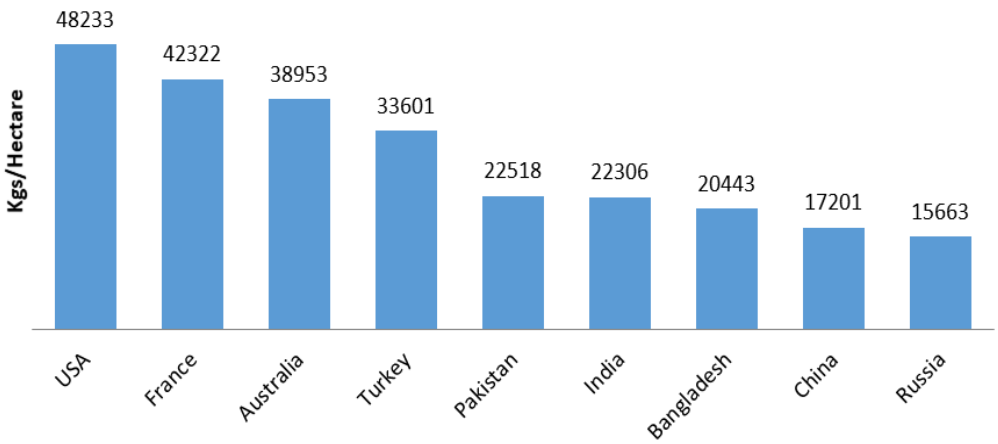


Figure 2. 16: Potato yield in major potato-producing countries of the world [22]

Pakistan is part of the Southern Hemisphere Cropping Cycle and has natural, geographic, and climatic advantages for potato production. Figures 2.17 to 2.21 show the trend in the area, yield, and production of potatoes at the national and provincial levels. Potato production continues to rise, owing to an increase in yield per hectare and the land devoted to potato production. Production increased by an annual growth rate of 2.3 percent from 0.394 million tons in 1981 to 4.87 million tons in 2018-19. The area under potatoes increased from 38,000 hectares in 1981 to 193,800 hectares in 2019-19 with an annual increase of 4.3 percent. Potato yield increased from 10,376 kgs/hectare in 1981 to 25,125 kgs/hectare in 2018-19 with an annual boom of 2.3 percent.

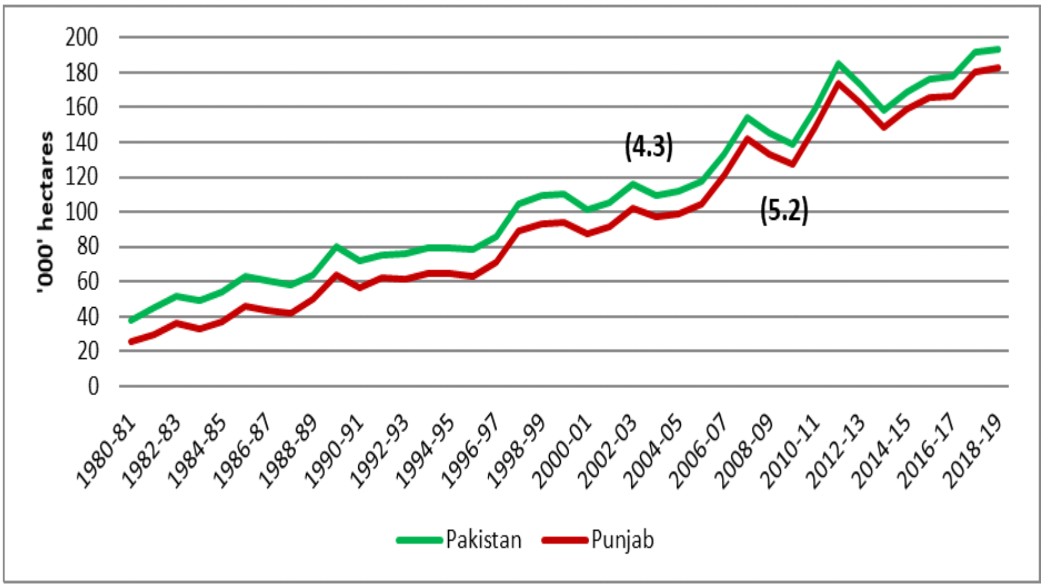


Figure 2. 17: Cultivated area of potato in 000 hectares in Pakistan and Punjab [20]

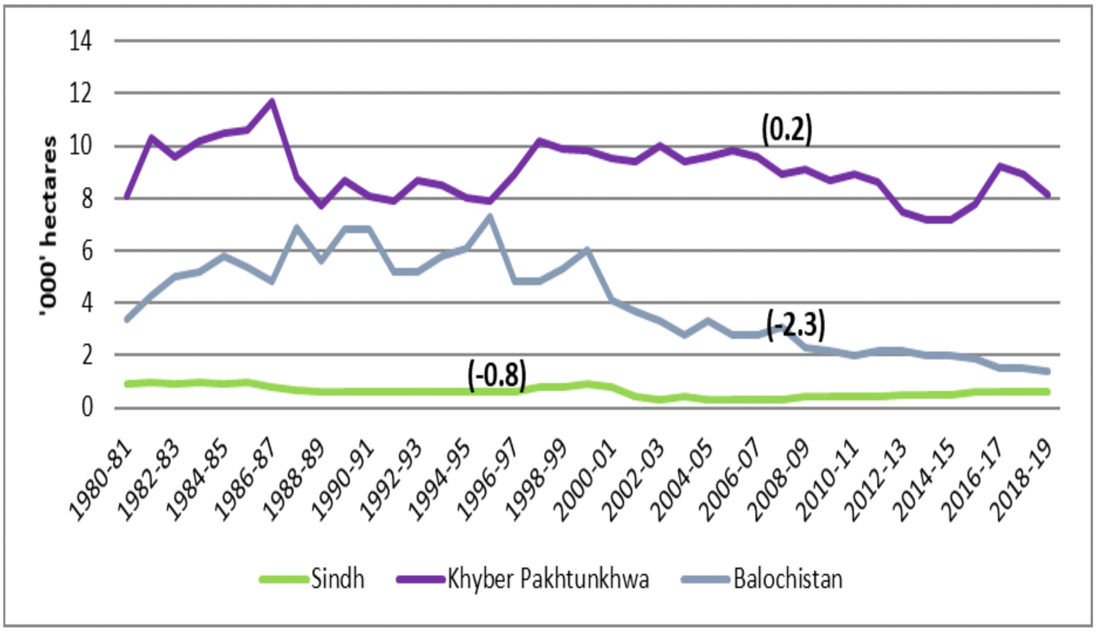


Figure 2. 18: Province-wise area of potato in thousand hectares [20]

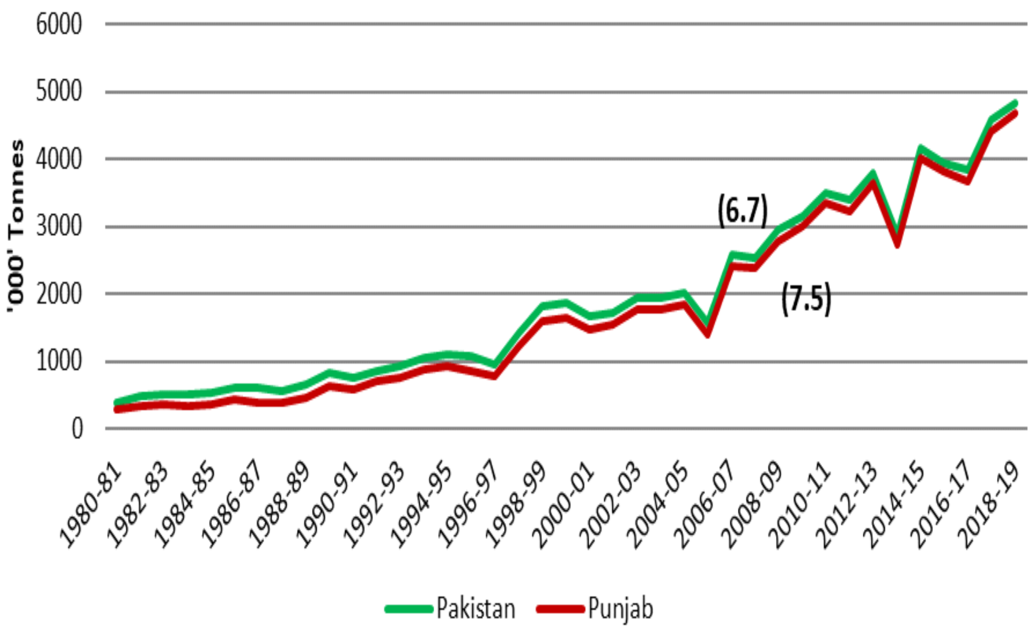


Figure 2. 19: Production of potato in thousand tones in Pakistan and Punjab [20]

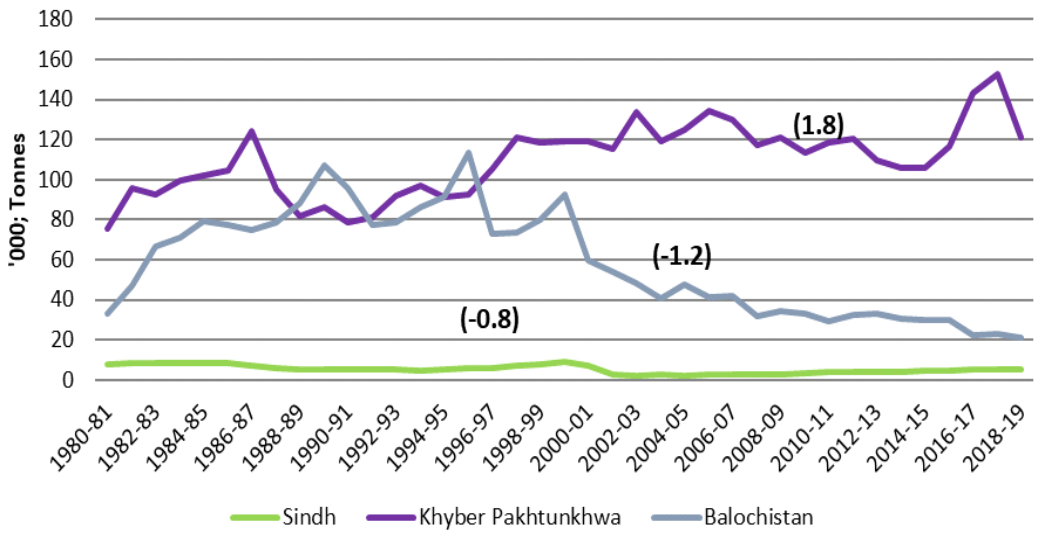


Figure 2. 20: Province-wise potato production in thousand tones [20]

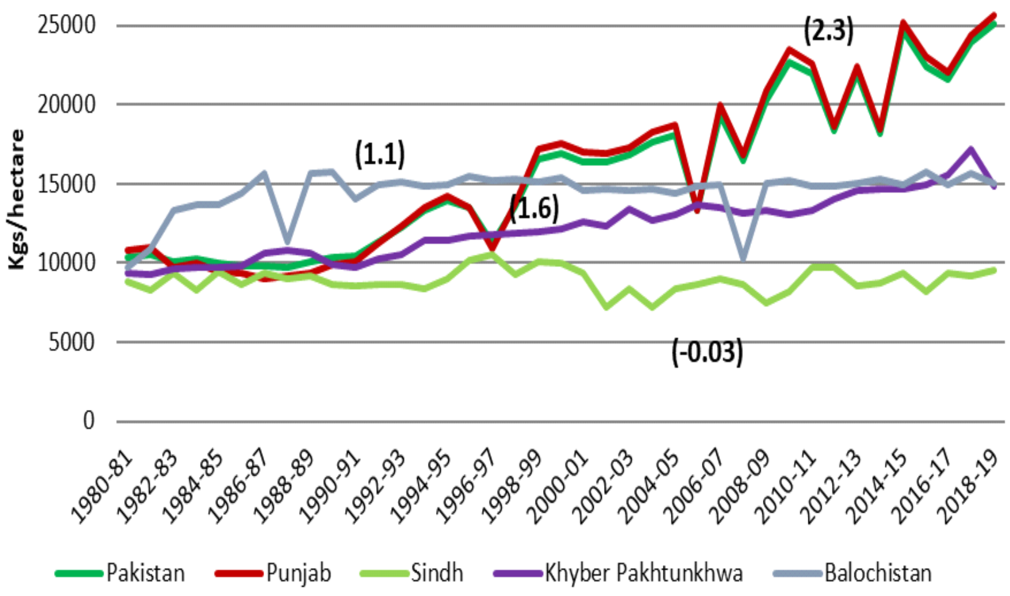


Figure 2. 21: Yield of potato in kgs per hectare in Pakistan and provinces [20]

Pakistan has a large potential to export potatoes as the country is part of the Southern Hemisphere Production Cycle. Around 90 percent of potatoes are harvested in winter (January to March), hence domestic prices tend to be lower during this period.

The exports of potatoes have been increasing in a highly variable pattern around a growth trend of 31000 tons per annum with an increase of 312 percent since 2006-07 (Figure 2.22). However, there is an increasing trend (90 percent) in potato export from 2014-15 to 2018-19. In the last two years, Pakistan's exports of potatoes showed impressive growth and were the highest in the last decade. For the first time, exports crossed 500,000 tons in 2017-18 and 600,000 tons in 2018-19. This increase in export may be due to currency depreciation resulting in high returns from exports.

Along with exports, Pakistan is also importing potato seeds from other countries. However, in 2013-14 and 2014-15, Pakistan imported potatoes in large quantities for domestic consumption to stabilize market prices in specific months.

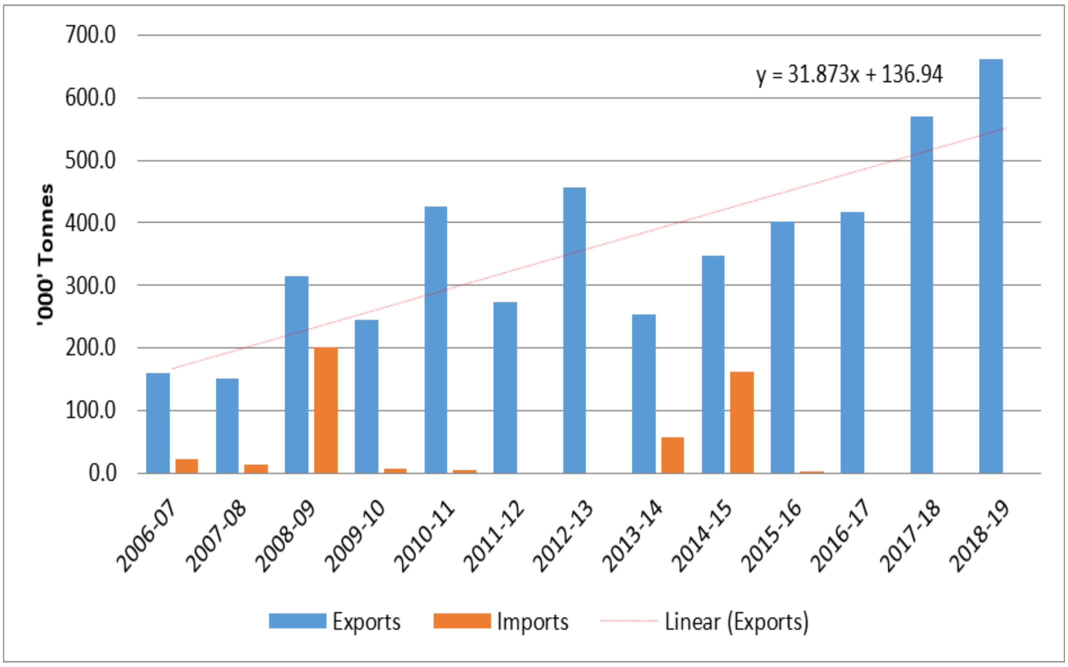


Figure 2. 22: Trade analysis of potato in Pakistan in thousand tones [23]

Potato Prices: In Lahore, the prices are at their lowest from January to April, however, the prices are comparatively high in Peshawar due to transportation costs. Figure 2.23 shows that prices are low at the time of potato harvest in Punjab (February to April) and comparatively high at the time of harvest in KP (September to November).

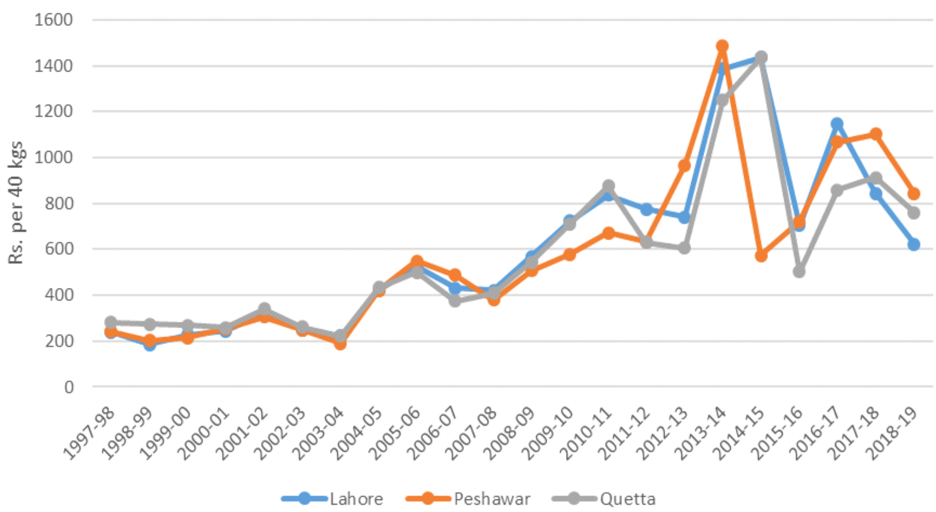


Figure 2. 23: Wholesale potato prices in major markets [20]

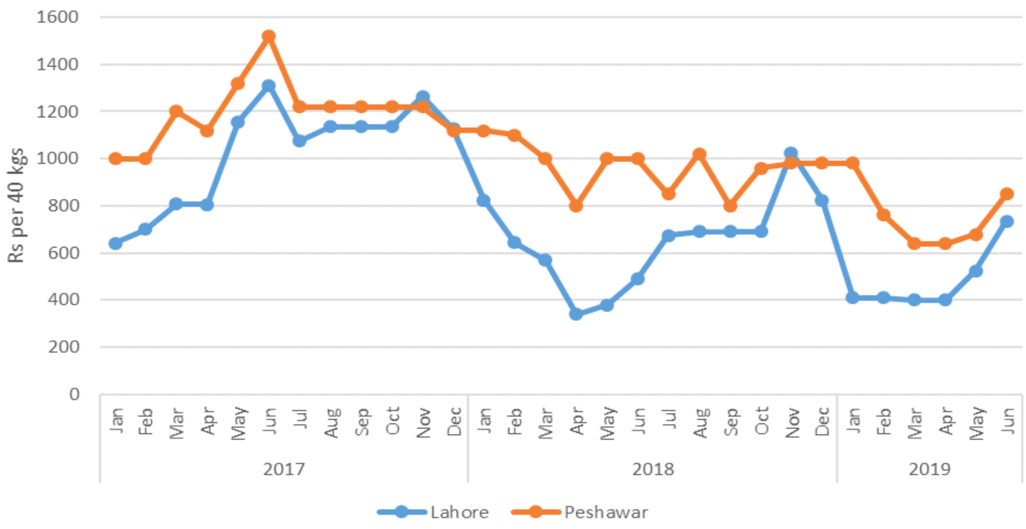


Figure 2. 24: Trends of monthly potato price

2.2.3 Gram

Gram, also known as chickpea, is an important pulse crop both in terms of production and consumption in Pakistan. There are two types of gram in Pakistan – Desi and Kabuli. Desi gram (black gram) is cultivated mainly in the Indo-Pak subcontinent while a majority of Kabuli variety is imported. Gram is mainly grown in Cholistan, Khushab, Bahawalpur, and Thal desert areas of Punjab, the southern districts of KP, and in the rice tract of Sindh. In Baluchistan, gram is grown in irrigated regions of Naseerabad and Jafferabad.

Figures 2.25 to 2.29 show the trends in the area, yield, and production of grams at the countrywide and provincial stages. As depicted in Figure 2.34, though Punjab accounts for more than 85 percent of the total area under gram cultivation, this area fluctuated between 780,000-990,000 hectares over the last two decades. Gram area in KP and Sindh has significantly decreased during this period mainly due to competition from cash crops and other high-value vegetables (Figure 30). Though the gram area in KP increased significantly from 1981 to 1991 and reached a record high of 123,000 hectares in 1991, it fell soon afterward, plummeting to a record low of 29,000 hectares in 2019. In Sindh, the area under gram has decreased significantly, from 131,000 hectares in 1982 to 19,000 hectares in 2019, at an annual rate of 4.6 percent. Interestingly, there is a slight increase in the area under gram cultivation in Baluchistan, and the province has now become the second-largest producer of gram in Pakistan.

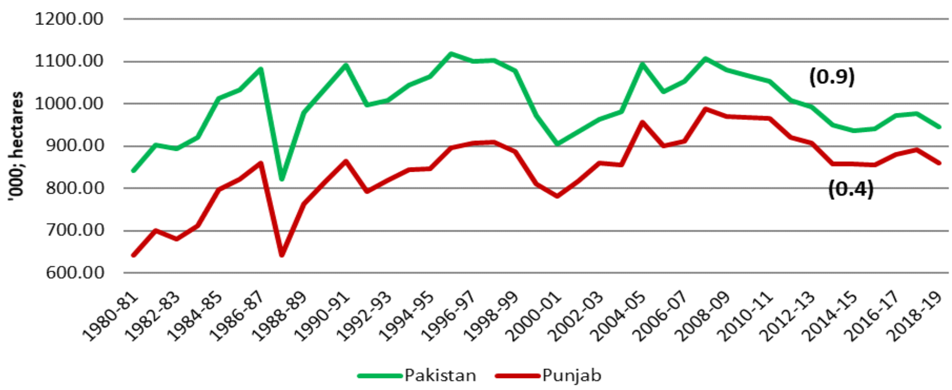


Figure 2. 25: Cultivated area of a gram in 000 hectares in Pakistan and Punjab [20]

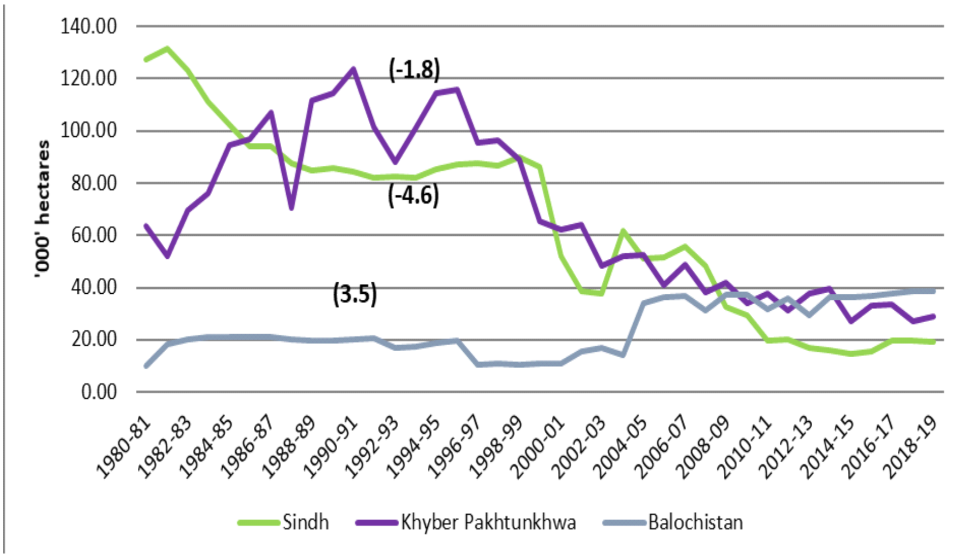


Figure 2. 26: Province-wise gram area in thousand hectares [20]

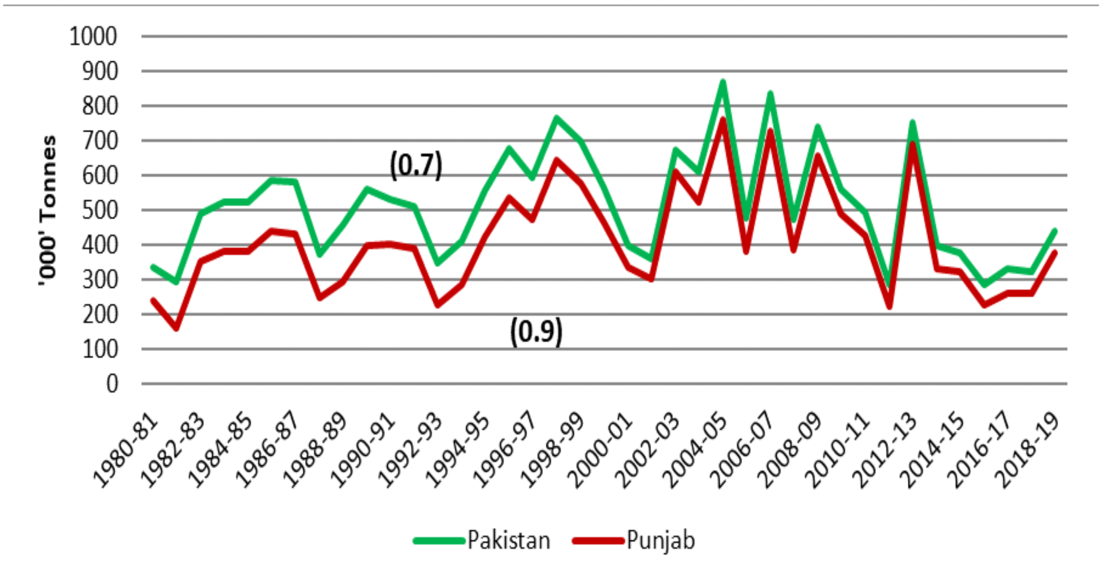


Figure 2. 27: Production of a gram in 000 tons for Pakistan and Punjab [20]

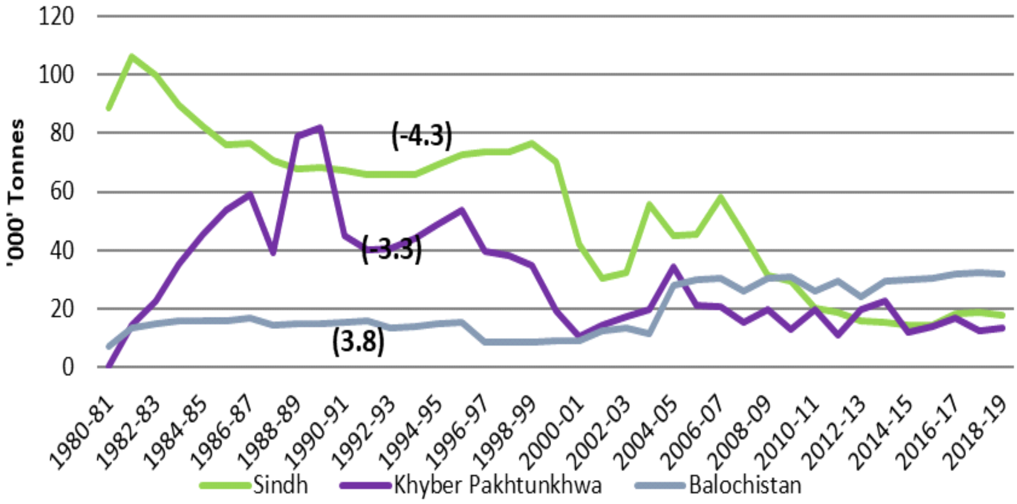


Figure 2. 28: Province-wise gram production in thousand tones [20]

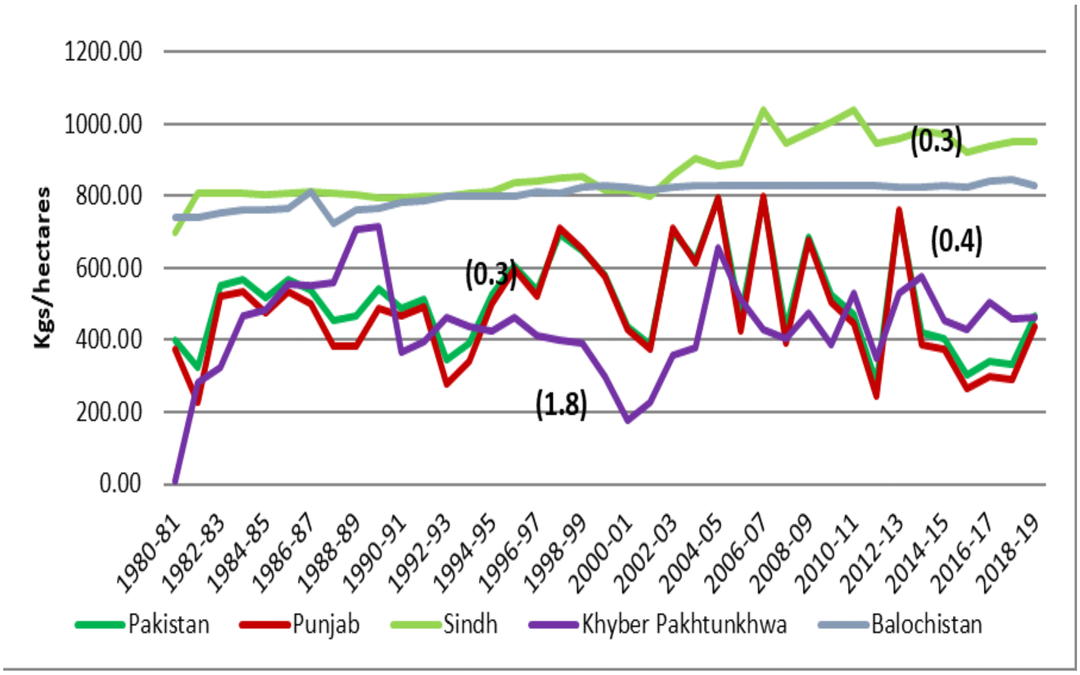


Figure 2. 29: Yield of a gram in kgs per hectare for Pakistan and provinces [20]

Cost of Production and Wholesale Prices of Gram: Though gram prices show high variations, mainly due to fluctuating domestic production, there has been a steady increase in the cost of production of a gram, reflecting the increase in input prices (Figure 2.31).

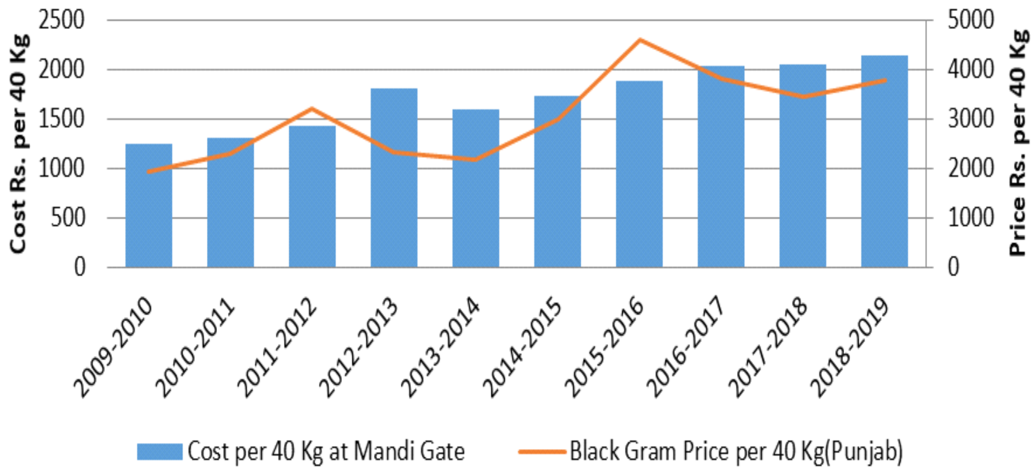


Figure 2. 30: Trends in the cost of production and prices [23]

The wholesale price reflects significant differences in different markets in Punjab since 2016. For instance, the wholesale prices in Rawalpindi and Lahore remained high in the last five years as compared to other markets of Multan and Sargodha. The average wholesale price in Rawalpindi in 2018-19 was Rs. 10,770 per 100 kg as compared to Rs. 8,715 per 100 kg in Sargodha. This could be attributed to local production in the Sargodha division. However, this difference is higher than the transportation cost.

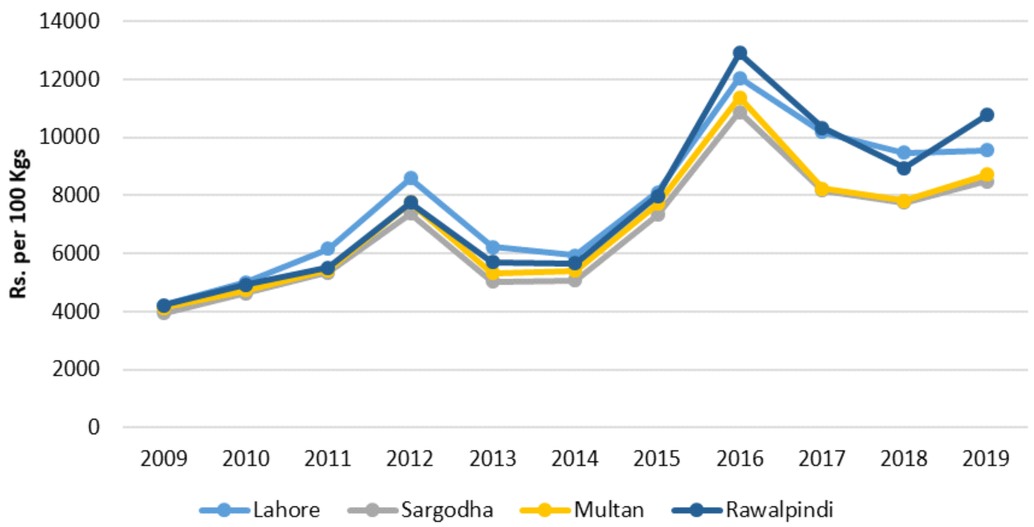


Figure 2. 31: Wholesale gram (black) prices in major markets in Punjab [23]

2.3 Regressive Time Series Models:

Forecasting is the methodology used in time collection or series, time-series data is studied to create forecasts and tell strategic decision-making by using statistics and modelings. It includes newly developed models to increase knowledge of data/records to identify underlying causes. Researchers can only answer WHY when the results come. When the forecasted results are analyzed then it will take subsequent steps to connect with that expertise and expectable assumptions of what may appear in the upcoming future.

Forecasting has many usages in many industries. It has many realistic applications in diverse industries which include weather, climate, economic, engineering, healthcare, retail, business, environmental studies, social studies, financial forecasting, and many more. Anybody who has regular historic records can examine those records with time series evaluation techniques after which model, forecasting, and predict sequentially based on that historical records.

Naturally, there are boundaries while managing the unpredictable and the unknown. Time collection prediction is not dependable and is not suitable or beneficial for all states/conditions. As there is no obvious collection of guidelines for when you must or must no longer use prediction, it is up to experts and data teams to distinguish the limits of the study and what their simulations can support.

Time series investigation illustrates how records variate over time, and top predictions can perceive the trend in which the record or data is varying. There are many models to deal with such time-series data and forecast differently according to available data, some of the time-series models are discussed below:

2.3.1 Auto-regression (AR)

AR function or model is additionally referred to as transition function, Markov function, or conditional function. An AR function predicts upcoming conduct hooked up mostly on previous behavior. Auto-Regressive AR (p) can model various forms of time series patterns because AR (p) is quite flexible. The function or model is a linear regression of the information inside the present series towards one or more prior values inside the same series. Generally, stationary time series data may be processed through

AR models. The representation for the Auto Regression model includes stating the order of the function p as Parameter. For example, AR (p), AR (1) is the primary order Auto Regression function. Which includes the range of the parameters phi (ϕ) [24].

An Auto-Regressive model of order p can be notated as AR (p). The AR (p) model can be represented as follow:

$$X_t = c + \sum_{i=1}^{p} \varphi X_{t-i} + \varepsilon_t$$

Where the parameters are $\varphi_1, \ldots, \varphi_p$, a constant is C , and white noise is ε_t.

Limitation:

The model is suitable for univariate time series, which doesn't involve trend and seasonal components.

Researchers are probably capable of expecting/predicting future trends quite properly with past data, however, they are by no means going to get one hundred percent accuracy AR functions also are referred to as Markov, conditional, or transition functions.

Standard Auto-Regressive functions carry out the most effective polynomial-time computation to compute the change or probability of the subsequent symbol. While it's attractive, which means they can't model distributions whose next symbol change or probability is difficult to calculate [25].

2.3.2 Moving Average (MA)

The moving average function (MA) model, additionally referred to as the moving average process, it's a mutual method for modeling univariate time series. The moving average function identifies that the output variable relies upon linearly at the current and numerous proceeding values of a stochastic imperfectly predictable duration. The finite MA function is constantly stationary opposing the AR function.

A moving average is a trend-following indicator and it's primarily based totally on previous prices. A moving average is calculated by selecting positive intervals after which dividing this variety of selected intervals. Moving averages can assist in smoothing out change action. Moving averages are used now no longer best to become aware of the trend however additionally for trade entry [5].

Rather than the usage of previous values of the prediction variable in a regression, a MA function makes use of previous forecast residual or errors in a regression alike function.

The representation MA (q) mentions back to the MA function of order q:

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta q \varepsilon_{t-q}$$

The μ = mean of the series, the $\theta_1, \ldots, \theta_q$ are the parameters of the model and hit noise error terms are $\varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \varepsilon_{t-q}$.

Limitation:

Requires retaining the records of various time durations for every predicted interval. Often overlooks complicated relationships stated within side the data. Does not reply to the variation that takes room for a reason, e.g. seasonal impacts, and cycles.

2.3.3 Auto regression and Moving Average (ARMA)

The ARMA functions are utilized in time collection evaluation to explain static time series collection. These functions constitute time series this is produced through passing white noise via a recursive and a no-recursive linear filter, repeatedly. In altered words, the ARMA function is a mixture of an AR and MA function. The notation ARMA (p, q) refers back to the version with p Auto-Regressive phrases and q Moving Average phrases. This function incorporates the AR ($\boldsymbol{p}$) and MA ($\boldsymbol{q}$) functions [6],

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

Auto-Regressive Moving Average function models as a linear feature of the observation and errors or residual at previous time steps. It combines AR and MA functions. The representation for the Moving Average feature includes stating the order of the AR (p) and MA (q) because of the Parameter of the ARMA model. For example ARMA (p, q) [26].

Limitation:

The version is appropriate for univariate time series, which does not involve trend and seasonal modules.

2.3.4 Autoregressive Integration Moving Average (ARIMA)

An ARIMA version may search to forecast a stock's future costs primarily founded totally upon its previous overall performance or predict a company's income primarily founded totally upon previous periods.

Autoregressive integrated moving average (ARIMA) functions are expecting future values primarily based totally upon past values. ARIMA uses lagged moving averages to smooth or clean the time-series collection of data. They are broadly utilized in technical evaluation to predict future safety prices. Auto-Regressive functions implicitly expect that the future will look like the previous or past. Therefore, they can show incorrect underneath certain market circumstances, such as economic crises or eras of quick technological alteration [8].

It's a statistical method that makes use of time-series data to forecast the future. Given time-series data X_t where an integer index is t and the real numbers are X_t, an ARMA (p', q) function is specified by

$$X_t - \alpha_1 X_{t-1} - \cdots - \alpha_{p'} X_{t-p'} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

Or

$$\left(1 - \sum_{i-1}^{p'} \alpha_i L^i\right) X_t = \left(1 + \sum_{i-1}^{q} \theta_i L^i\right) \varepsilon_t$$

Where L is the operator of lag, α_i is the parameters of the AR part of the model, θ_i is the parameters of the MA part and ε_t are the error terms. The ε_t error expressions are normally supposed to be independent, and identically distributed variables are sampled from a normal distribution with 0 means.

Suppose the polynomial $\left(1 - \sum_{i-1}^{p'} \alpha_i L^i\right)$ has a unit root or a factor $(1 - L)$ of multiplicity d. Then it may be redrafted as:

$$(1-\sum_{i-1}^{p'} \alpha_i L^i) = (1+\sum_{i-1}^{p'-d} \varphi_i L^i)(1-L)^d$$

This polynomial factorization property process expresses an ARIMA (*p*, *d*, *q*) with *p*=*p′*−*d* and is specified by:

$$\left(1-\sum_{i=1}^{p} \varphi_i L^i \right)(1-L)^d X_t = (1+\sum_{i=1}^{q} \theta_i L^i)\varepsilon_t$$

And for this reason may be the notation of a specific case of an ARMA ($p + d, q$) procedure having the AR polynomial with d unit-roots. Therefore, no process that is precisely defined by an ARIMA model with $d > 0$ is huge feel stationery.

The overhead equation may be generalized as:

$$\left(1-\sum_{i=1}^{p} \varphi_i L^i \right)(1-L)^d X_t = \delta + (1+\sum_{i=1}^{q} \theta_i L^i)\varepsilon_t$$

The Parameters may be described as:

p: The range of lag observations within the side of the model; additionally, referred to as lag order. d: The range of instances that which raw observations are differenced; additionally, referred to as the degree of difference. q: The scale of the MA window; additionally, also identified as the order of the MA.

AR (p) abbreviated as an autoregressive model, the p is an integer parameter that confirms what number of lagged collections or series goes for use of prediction periods. I (d) is differencing part, the d parameter tells what number of differencing orders are going for used to make the collection or series stationery [27].

MA model MA (q), q is the term for number of lagged forecast error in the prediction equation. SARIMA is seasonal ARIMA and its miles are used along with time collection or series with seasonality.

ARIMA version is appropriate for univariate time series or collections along with trends and without seasonal elements. An ARIMA version may be used to expand AR, MA, and ARMA models.

Limitation:

The core limitation of this model is that the parameters (p, d, q) want to be manually defined; therefore, locating the most correct match may be a long trial and error process.

ARIMA model relies upon extraordinarily at the reliability of ancient records and the differencing of the data/records. It's vital to make certain that records/data become amassed appropriately and over a protracted-time period so that the model affords correct results and forecasts [28].

Difference between ARMA and ARIMA:

The two functions proportion many similarities. The AR and MA functions are approximately identical in both models ARMA and ARIMA, by combining or merging an AR (p) function and MA (q) function. AR (p) makes forecasts by using previous values or records of the established variable. MA (q) makes predictions about the usage of the series/collection mean and former errors.

An ARMA version is a stationary function/model; if one version or model is not stationary, then one can gain stationarity via way of taking a chain of differences. The 'I' within the side of the ARIMA version/model stands for integrated; it's a degree of what number of non-seasonal differences are had to gain stationarity. If no difference is worried within the side of the version/model, then it turns into simply an ARMA.

A version or model with a d^{th} difference to fit and ARMA (p, q) version or model is referred to as ARIMA procedure of the order (p, d, q).

2.3.5 Seasonal Auto-Regressive Integrated Moving Average (SARIMA)

SARIMA, or Seasonal ARIMA, is an addition of ARIMA that cares for univariate time series records or data along with a seasonal constituent.

It provides 03 new hyper parameters to specify AR=Auto Regression, I=differencing, and MA=Moving Average for the seasonal element of the series or collection, In addition to a further parameter during the seasonality.

SARIMA functions are commonly denoted ARIMA (p, d, q) $(P, D, Q)_m$, in which m refers back to the wide variety of intervals or periods in every season, and the uppercase P, D, Q refers back to the Auto-Regressive, differencing, and Moving Average elements for the seasonal portion of the ARIMA version. And its representation is as follows:

$$ARIMA \quad \underbrace{(p, d, q)}_{\text{Non-seasonal part of model}} \quad \underbrace{(P, D, Q)_m}_{\text{Seasonal part of the model}}$$

Where the wide variety of observations according to a year = m. Researchers use the uppercase representation for the seasonal components of the function and lowercase representation for the non-seasonal components of the function. This technique is appropriate for univariate time collection or series along with trend and or seasonal components [8].

The distinction between ARIMA and SARIMA is all about the seasonality of the data. In case the data is seasonal like it takes place after a sure time frame. then it's going to use SARIMA. SARIMA will add one more term which is seasonal order $(p, d, q, period)$

p, q, d Values will stay identical. The length cost could be the cost after what time frame seasonality happens [29].

2.3.6 Simple moving average (SMA)

An SMA calculates the average of a specific verity of prices, typically closing prices, through the verity of intervals in that range. An SMA is a technical indicator that could useful resource in figuring out if an asset rate will retain or if it's going to oppose a bull or bear trend. An SMA may be better than an EMA (Exponential Moving Average) that is extra closely weighted on the latest price action [9].

The only sort of approach to forecasting is calculated with the aid of using adding up the last 'k' period's values and after which dividing the number by 'k'.

Moving Averages may be used to quickly pick out whether a trend is transforming into an uptrend or a downtrend relaying at the pattern captured with the aid of using the moving average. The components of Simple Moving Average:

$$\text{Simple moving Average} = \frac{A_1 + A_2 + \cdots + A_n}{n}$$

The SMA is straightforward to calculate and is the average stock price over a positive length based totally on a set of parameters. The moving average is calculated by including an inventory's expenses over a positive length or period and dividing the sum through the total range of periods [9].

Limitation:

Some investors consider that a Simple Moving Average offers an excessive amount of weight to old data and like to apply an exponential moving average instead. Simple Moving Average may not correctly reflect the maximum current trends.

2.3.7 Exponential Smoothing (ES)

Exponential Smoothing shortens as ES. ES of the time series data assigns exponentially lessening weights from most recent to oldest observations. In different words, the older the record or data, the data is given less priority; newer data is considered more relevant and is allocated higher weight. Smoothing parameters also known as smoothing constants are usually denoted with aid of using α to decide the weights for observations.

If the usage of the moving averages approaches forecast, researchers want to have q previous values. This is bulky if there are numerous objects for which forecasting/prediction is required. Because exponential smoothing is based on the handiest portion of data, the former period's real value and the forecasted value for the actual period, it reduces the data storing requirements [12].

ES is very quite simple in idea and smooth to understand. ES could be very effective due to its weighting process. ES is commonly used to make short-time period forecasts, as long-time period forecasts the usage of this approach may be pretty untrustworthy [12].

The handiest shape of an ES system is represented by:

$$s_t = \alpha x_t + (1 - \alpha)s_{t-1} = s_{t-1} + \alpha(x_t - s_{t-1})$$

In this formulae, it's the current observation x_t a simple weighted average smoothed statistic= s_t, Previous smoothed statistic= s_{t-1}, Smoothing factor of data = α where $0 < \alpha < 1$ and time period= t.

If the Smoothing aspect value is higher, then the smoothing level will reduce. The cost of α near 1 has much less of a smoothing impact and offers extra weight to recent adjustments within side the data, at the same time as the value/cost of α in the direction of 0 has an extra smoothing impact and is much less reactive to current changes or adjustments. This technique is used to provide a smoothed Time Series. ES is often a manner of smoothing out the data or facts with the aid of using getting rid of tons of noise also known as random variations from the data by giving a better forecast [12].

- **Simple Exponential Smoothing**:

  Simple or single exponential smoothing makes use of a weighted moving average with exponentially declining weights.

- **Holt's Exponential Smoothing:**

  Holt's trend corrected double exponential smoothing is typically greater dependable for handling data that suggest trends, in comparison to the solo procedure.

- **Winters' Three Parameter Linear and Seasonal Exponential Smoothing:**

  Triple exponential smoothing is additionally known as the Multiplicative Holt-Winters, commonly dependable for parabolic data or trends that suggest seasonality and trends.

Limitation:

ES will lag. In different words, the forecast can be behind, because the trend upsurges or drops over time.

ES is unsuccessful to account for the dynamic adjustments at work within side the actual world, and the prediction will continuously require updating to reply to new data.

2.4.8 Vector Auto Regression (VAR)

Vector Auto-Regressive version is implacable on multivariate time series. The shape is that every variable is a linear feature of previous lags of itself and previous lags of the further variables [10].

Vector Auto Regression function, this version is the simplification of AR to more than one Parallel time collection or series for example multivariate time collection or series. The representation for this version entails stating the order for the AR (p) functions parameters to a VAR characteristic for example. VAR (p). This version is appropriate for multivariate time collection or series without trend and seasonal elements [30].

Limitation:

Contemporaneous variables are not related to one another.

The error terms is interrelated across equations. This means researchers cannot consider what impacts individual shocks will have on the system.

2.4.9 Vector Auto Regression Moving Average (VARMA)

VARMA version is one of the statistical analyses regularly used for multivariate time collection or series data or information in business, finance, and economy. It’s used very much due to its simplicity. Moreover, the VARMA version can give us an explanation for the dynamic conduct of the connection between endogenous and exogenous variables or endogenous variables. It also can explain the effect of a variable or a collection of variables utilizing the impulse reaction function and Granger causality. Furthermore, it's able to be used to predict and forecast time series data [31].

Vector Auto Regression Moving Average approach version is the simplification of ARMA to more than one Parallel time collection or series for example multivariate time collection or series. The representation for this version entails stating the order for the AR (p) and MA (q) functions parameters to a VARMA function for example VARMA (p, q).VARMA version may be used to broaden VAR or VMA functions [32].

Limitation:

This version is appropriate for multivariate time collection or series without trend and seasonal elements.

2.4.10 Neural Network (NN)

The NN is a machine getting to know (learning) technique that functions the human mind and includes numerous artificial neurons, their capability to pick up by example makes them very elastic and powerful.

NN may be utilized in any type of enterprise and get benefitted due to its being very elastic and additionally does not require any algorithms. They are frequently used to function as parts of residing organisms and look at the inner mechanism of the mind [33].

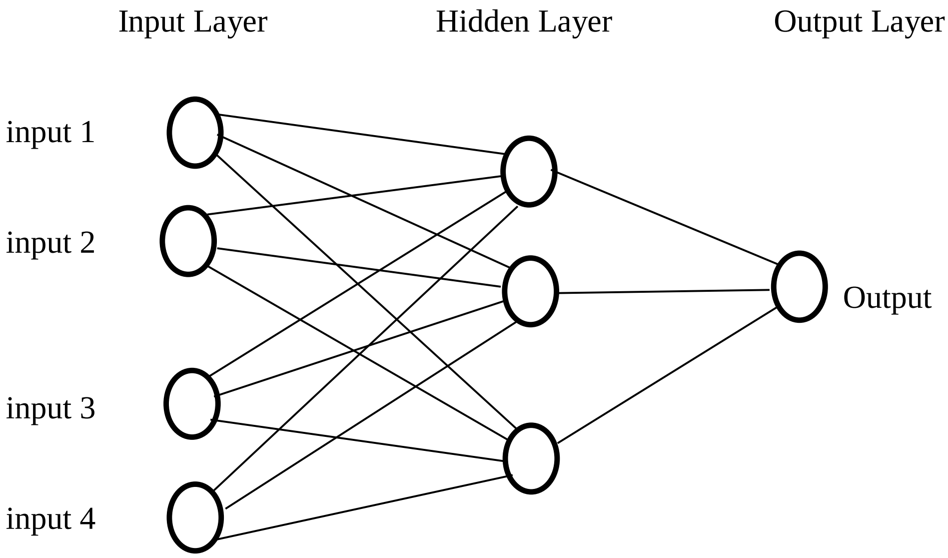


Figure 2. 32: Architectuere of nural network

Does Researcher have the proper information or data? For example, if researchers have a classification problem, they will want labeled information or data. Is the dataset researcher want publicly available, or are they able to create it? In this example, spam mails emails might be classified as spam mails, and the labels might allow the algorithm to plot from inputs to the classifications the researcher care about. The researcher cannot realize which researcher has the proper information or data till the researcher get hold of it. If a researcher is a data scientist working on a problem, you cannot everyone to inform the researcher whether or not the information or data is ideal enough. Only direct exploration of the information or data will give a solution to this question [13].

Limitation:

Neural networks commonly require a lot of extra statistics or data than conventional machine learning procedures, as a minimum slot if not tens of thousands and millions of categorized samples. This is not clean trouble to address and plenty of machine learning issues may be solved properly with much fewer statistics or data in case researchers use different algorithms [33].

2.5 Bayesian State Space

A time collection series is defined probabilistically through a stochastic process Y_t; where $t = 1, 2, \ldots$, that is, through an ordered arrangement of casual vectors with index time $= t$. For ease, researchers would consider equal interval time points (everyday data, month-to-month data, and so on). The goal is to do predictions approximately the value of the subsequent observation y_{n+1} having observed data up to time$n$, $(y_1 = y_1, \ldots, y_n)$] [2].

State Space Models ‘SSM’ are hidden variable functions which can analyze time collection/series data or information because of their elastic and general framework. The first unique case is Linear Gaussian SSM (LGSSM) where each state's densities and observation are Gaussian with linear relationships with the states. Another unique case is SSM with discrete state space, which is occasionally known as hidden Markov models.

Bayes model is simple to construct and specifically beneficial for extremely huge records or data collections. Simply, Naive Bayes is understood to beat even pretty state-of-the-art classification approaches [14].

Bayes theorem gives a manner of computing previous probability $P(c|x)$ from$P(c)$, $P(x)$, and $P(x|c)$. The below equation represent the idea:

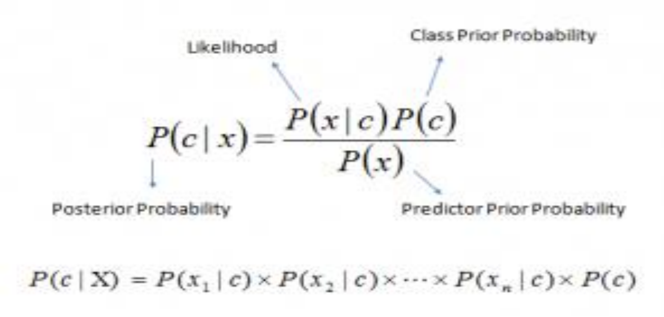


Figure 2. 33: Bayes theorem

The previous probability of class is $P(c|x)$ $(c, target)$ given predictor $(x, attributes)$. The previous probability of class is $P(c)$. The likelihood is $P(x|c)$ which the probability of the predictor given class is. The previous probability of predictor is $P(x)$.

This regularization technique help to reduce overfitting by penalizing when the parameter value gets large, another big reason researcher often prefers to use the Bayesian model is that it permits us to include uncertainty in parameter estimates that's beneficial when forecasting. Bayes theorem is constructed on the principle of conditional probability and lies within-side the coronary heart of Bayesian Inference [14].

$$\mathrm{P(A|B)} = \frac{\{\mathrm{P(B|A)x\ P(A)}\}}{\mathrm{P(B)}}$$

Bayesian Theorem is a mathematical system which describes that the way to replace the probabilities of hypothesis while early data is to be associated with the event. Figure 9 Bayes theorem provides a path to revise existing predictions or theories by giving new or additional information [34].

Bayes Theorem

Likelihood
Probability of collecting this
Data when our hypothesis is true

Prior
The Probability of the hypothesis being
true before collecting data

$$P(H|D) = \frac{P(D|H)\,P(H)}{P(D)}$$

Posterior
The probability of our hypothesis
Is true given the data collected

Marginal
what's the Probability of collecting
data under all possible hypotheses?

Figure 2. 34: Bayes Theorm

A random compression approach to lessen a big dimensional foreign marketplace data into a mile's minor matrix. Then the Bayesian model averaging (BMA) technique is haired to the weight of every random compressed Vector Autoregressive for accomplishing the first-rate prediction function. They determined that Bayesian compressed Vector Auto-Regressive (BCVAR) and time-varying Bayesian compressed Vector Auto-Regressive (TVP-BCVAR) have been capable of carrying the greatest forecasting at the forex. [1] DLM consists of a system of equations that has two main objectives: [2]

Observations of the randomly decided methods of a method depending on the modern method parameters,

In time, method parameters evolve, because of the inherent method dynamics and from random issues or shocks.

DLM Observation equation:

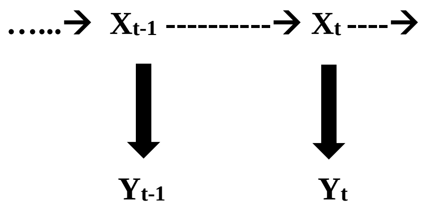
$$Y_t = F_t\theta_t + v_t, \{v_t \sim N(0, V_t)\}$$

System equation:

$$\theta_t = G\,\theta_{t-1} + \{\, w_t \sim N\,(0, W_t)\}$$

For revising or updating information concerning the parameter vectors, the Kalman result delivers recurrence relationships [34].

State Space Model: Described by two principals. 1st process X_t called state process (Markov Process). Present and past are autonomous conditions in the present X_t. 2nd process is called Observation Y_t are independent of Xt. Dependence between the observations is generated via the means of states.



Linear Gaussian Model: Employs can order one, P dimensional as:

$$X_t = \phi X_{t-1} + W_t$$

W_t are $P * 1$ identically distributed and unbiased, 0 mean regular vectors along with $Q =$ covariance matrix;

$$W_t \sim iid\ N_p\ (0, Q)$$

Assume start process X_o, such that

$$X_o \sim N_p\ (\mu_o, €_o)$$

Observation Equation is written as:

$$Y_t = A_t X_t + V_t$$

At $q * p$ measurement equation or observation equation Yt q-dimensional $>$ $p <$ (state dimension).

$$V_t \sim \text{iid}\ N_q(0, R)$$

Suppose researchers have $r * 1$ vector inputs μ_t

$$X_t = \phi X_{t-1} + r_{\mu t} + W_t$$

$$Y_t = A_t X_t + \dot{T}_{\mu t} + V_t$$

Both of these matrices can be 0, where r is $p * r$ and $\dot{T}$ is $q * r$.

Example: AR (1) function with observational noise:

Consider univariate SSM with noisy observation,

$$Y_t = X_t + V_t$$

Also, AR (1) process is a signal state.

$$X_t = \phi X_{t-1} + W_t$$

Where

$$V_t \sim \text{iid}\ N\ (0, \sigma^2 v_2),$$

$$W_t \sim \text{iid}\ N(0, \sigma^2 w)\ \ \text{And}$$

$$X_o \approx N\ [0, \frac{(\sigma_w^2)}{(1 - \phi^2)}];\ \{V_t\}, \{W_t\}$$

And X_o are independent and $t = 1,2,3,\ldots$

Auto-covariance of X_t is

$$r_x(h) = \left\{\frac{(\sigma_w^2)}{(1-\phi^2)}\right\} * \phi^h, h = 0,1,2,3,\ldots$$

Investigation: - How does observation noise affects the dynamics?

Assumed: - X_t is static

In this case, observation is likely stationary due to the fact Y_t is the amount of 2 autonomous static elements X_t and V_t.

We have

$$r_y\ (o) = \mathrm{var}(Y_t) = \mathrm{Var}\ (X_t + Y_t) = \left\{\frac{(\sigma_w^2)}{(1-\phi^2)}\right\} + \sigma_w^2\ , h \geq 1,$$

$$r_y\ (h) = \mathrm{cov}\ (Y_t, Y_{t-h}) = \mathrm{cov}\ (X_t + V_t, X_{t-h} + V_{t-h}) = r_x(h)$$

Therefore, for $h \geq 1$, the AFC of observation is

$$P_y(h) = \{\frac{r_y(h)}{r_y(o)}\} = \{1 + (\frac{\sigma_v^2}{\sigma_w^2}) * (1-\phi^2)\}^2\ \phi^h.$$

Observation Y_t , is not AR (1) except $\sigma_v^2 = 0$. The Auto-Correlation Arrangement of Y_t is similar to the Auto-Correlation of ARMA (1,1). ARMA (1,1) form: $Y_t = \phi Y_{t-1} +$ ⍰$u_{t-1} + \mu_t$, Where μ_t is a Gaussian white noise of variance σ_u^2 and with σ_u^2 suitably chosen.

2.6 Literature Viewed

Table 2.1: Literature and Accuracies

Year	Author	Title	Dataset and Model used.	Accuracy
14 March 2020	P. Chandra Shaker Reddy & A. Sureshbabu	An Applied Time Series Forecasting Model for Yield Prediction of Agricultural Crop	Two seasons of rice production data were collected from 2008 to 2014. SARIMA time series model used for predicting rice production. [26]	Mean absolute percentage error MAPE=0.3574 and root mean square error RMSE= 0.3574,
20 Sep 2021	Sourav Kumar Purohit, Sibarama Panigrahi, Prabira	Time Series Forecasting of Price of Agricultural Products Using Hybrid Methods	Predict the monthly retail & wholesale price of tomato, onion, and potato. Two additive hybrid methods and five multiplicative hybrid methods were used.	Additive-ARIMA-ANN provides the best RMSE= 1.9925, SMAPE= 6.6665,

	Kumar Sethy, Santi Kumari Behera.			and MAE= 1.5696. [27]
2014	M. Amin, M. Amanullah and A. Akbar	TIME SERIES MODELING FOR FORECASTING WHEAT PRODUCTION IN PAKISTAN	Wheat data from 1902 to 2005 was used. They developed time series models and the best model is identified for wheat production of Pakistan, is ARIMA [28]	The confidence level of ARIMA is higher than 95%.
24 Jan 2021	V. Jadhav, B.V. Chinnappa Ready, and M.G. Gaddi	Application of ARIMA model for forecasting agriculture prices	Forecast and validate agricultural commodity prices for major crops such as paddy, ragi, and maize in Karnataka using time series data from 2002 to 2016. We apply a univariate ARIMA method to generate price forecasts for major crops. [29]	The models ARIMA(1,1,1) for paddy, ARIMA(1,1,2) for ragi, and ARIMA(1,2,1) for maize are good candidates as they have the lowest AIC and SBC values. has been proven.
1983	Jon A. Brandt, David A. Bessler	Price forecasting and evaluation: An application in agriculture	US hog price records are used. They describe seven forecasting approaches and examine their performance over his 24 quarters from 1976 to 1981. Methods include exponential smoothing, autoregressive integrated moving average processes, econometric models, expert judgement, and compound forecasting approaches. [30]	The ARIMA model performed better and appeared to be the most accurate of the individual forecasting methods. The mean absolute percentage error (MAPE) is less than 75% for the naive method.
2019	Qing Wen, Yapeng Wang, Haodong Zhang & Zhen Li	Application of ARIMA and SVM mixed model in agricultural management under the background of intellectual agriculture	Jiangsu Province corn dataset 1995–2004 was used. Three different models are used ARIMA model, SVM principles, and ARIMA–SVM model. [31]	The prediction accuracy of ARIMA-SVM is greatly improved, along with certain advantages, with MSE= 1.32 and MAPE= 5.1

2020	Shakir Khan, Hela Alghulaiakh	ARIMA Model for Accurate Time Series Stocks Forecasting	ARIMA Model is used. Netflix stock historical dataset of five years was used. [32]	The ARIMA(1,1,33) model showed better accuracy. Accuracy was 99.74% and ARIMA(1,2,33) was 99.75%, which is almost the same.
2020	NAJEEB IQBAL, KHUDA BAKHSH, ASIF MAQBOOL, AND ABID SHOHAB AHMAD.	Use of the ARIMA Model for Forecasting Wheat Area and Production in Pakistan	The wheat dataset is used from 2002 to 2022 and the ARIMA model is used to predict wheat area and production in Pakistan. [33]	ARIMA(1,1,1) and ARIMA(2,1,2) were obtained 20 years ago and contain wheat area and yield forecasts with 95% confidence intervals.
2011	MA Awal, MAB Siddique	Rice Production in Bangladesh Employing by ARIMA Model	This study uses the Bangladesh Rice dataset and uses an ARIMA model to predict Bangladeshi rice production. [34]	This study revealed that ARIMA models are superior to the respective deterministic models. With RMSE= 205.5 and MAPE= 6.37
2013	D.P. Singh, Prafull Kumar, and K. Prabakaran	Application of ARIMA model for forecasting Paddy production in Bastar division of Chhattisgarh	Paddy area and production data in Bastar district, Chhattisgarh from 1974-75 to 2010-2011 were analyzed using the time series method. ARIMA(2,1,2) and ARIMA(2,1,0) models were used to predict rice area and production in his Bastar district of Chhattisgarh for the main four years. [35]	The selected ARIMA model is suitable. The residuals ACF and PACF also show that the model is "well-fitting". For RMSE=121.74 and MAPE=16.322.
20 July 2019	Anil KumarMahto, Ranjit Biswas & M.	Short Term Forecasting of Agriculture Commodity Price by Using ARIMA: Based	You have selected sunflower seed prices for the period 01/01/2011 to 12/31/2016. We used an autoregressive integrated moving average (ARIMA) model in time series analysis for forecasting. [36]	The ARIMA(1, 1, 2) MAPE and RMSEP for sunflower seeds are 2.30 and 3.44, respectively.

	Afshar Alam	on the Indian Market		
15 March 2016	Michael D.Johnson William W.Hsieh Alex J.Cannon Andrew Davidson Frédéric Bédard	Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods	Predict yield using various combinations of MODIS-NDVI, MODIS-EVI, and NOAA-NDVI as predictors using multiple linear regression (MLR) and nonlinear machine learning models, Bayesian Neural Networks (BNN) Did. Statistics Canada's Department of Agriculture collects detailed annual yield data for all crops across Canada. Yield data for barley, oilseed rape and spring wheat by CAR (kg ha−1) from 2000 to 2011 were obtained from Statistics Canada.	Different predictor sets and different methods of MLR, BNN, and MOB models were used to group the CARs with 95% confidence intervals from the bootstrap method. [37]
20 May 2014	K. MATSUMURA, C.F. GAITAN, K. SUGIMOTO, A.J. CANNON, and W.W. HSIEH	Maize yield forecasting by linear regression and artificial neural networks in Jilin, China	Maize data from Jilin, China, from 1962 to 2004. We investigated using multiple linear regression (MLR) and nonlinear artificial neural networks (ANN). [38]	The best retrospectively validated SS was 0.136 for ANN2 with HN=1, with a bootstrap 95% confidence interval. Predictive skill scores calculated by both cross-validation and retrospective validation showed that the ANN model significantly outperforms MLR.
12 February 2018	Cristanel Razafimandimby, d Valeria Loscri, Anna Maria Vegni, Alessandro Neri.	Efficient Bayesian Communication Approach for Smart Agriculture Applications	They present a Bayesian inference approach (BIA) that allows to avoid sending data that are spatio-temporally strongly correlated. Data from Taoyuan collected by close monitoring using low-power wireless mesh network technology will be used.	The best performances are for scenario s3 in Table 4, where we observe an error of 1.0 for 99.31% of the time.
2007	Jiejun Huang,	Development of a Data Mining	They focus on developing agricultural data mining	Test data set =200, correct=

	Yanbin Yuan, Wei Cui & Yunjun Zhan	Application for Agriculture Based on Bayesian Networks	applications based on Bayesian networks. The dataset contains 2000 cases, a portion of the data, and 6 variables containing status.. [39]	175, and Accuracy= 87.5%.
February 2020	A.Kocian, .Massa, S.Cannazzaro, L.Incrocci, S.Di Lonardo, P.Milazzo, S.Chessa.	Dynamic Bayesian network for crop growth prediction in greenhouses	A dynamic Bayesian network (DBN) is used in this study. Three lettuce growing days were used to test the performance of the proposed DBN. [40]	Van Henten's analytical model is accurate for the cycles used for calibration. Growing days is T = 15 and prediction lengths of q = 1 & q = 5 lead to prediction errors of about 6% and 23%, respectively.
03 November 2016	Niketa Gandhi, Leisa J. Armstrong, Owaiz Petkar.	PredictingRice crop yield using Bayesian networks	This paper reports a method for predicting rice yields in Maharashtra, India, using Bayesian networks. The data set contains parameters selected for study: precipitation, minimum temperature, average temperature, maximum temperature, evapotranspiration of reference crops, area, and production for the Kharif season from 1998 to 2002. quantity and yield. [41]	Classifier name = BayesNet, Precision = 97.53, Sensitivity = 96.31%, Specificity = 98.16%.
03 December 2018	Rashmi Priya, Dharavath Ramesh, Ekaansh Khosla.	Crop Prediction on the Region Belts of India: A Naïve Bayes MapReduce Precision Agricultural Model	Made from cotton, maize, rice and chili records. A model of simple Bayesian models is used for prediction. A precision agriculture model is presented to suggest to farmers which crops to plant depending on field conditions. [42]	Using Naive Bayes makes the model very efficient computationally.

By reviewing the above literature and their accuracies, we concluded that the accuracy of the Bayesian framework used is higher than that of the ARIMA framework. The accuracy of the Bayesian framework is 95% to 99.31% and prediction errors are 6% to 23%, respectively. This is the main reason for the selection of the Bayesian framework because the accuracy of the Bayesian framework is relatively higher than of other models. By combining both ARIMA and Bayesian modeling we get an ARIMA-based Bayesian framework, the combination of both models will help us to achieve effectiveness in our results along with higher prediction accuracy and errors or residuals is further reduced by combining both models. And the accuracy of any model is very important as prediction outcomes depend on them and will give outcomes near to or approximately to the actual outcome.

2.7 Summary

This chapter took a brief look at the time series and also discussed Pakistan's agricultural landscape for Rabi food crops, then took a brief look at different data related to the potato, wheat, and gram of different years and different locations observed. Then for time series prediction or forecasting different regressive models were looked through, which include AR, MA, ARMA, ARIMA, SARIMA, VAR, VARMA, Exponential smoothing, NN, and at last Bayesian State-Space modeling were discussed in brief.

CHAPTER 3

METHODOLOGY

3.1 Overview

The explanation of the operational framework and research design and development are described in this chapter. In the Bayesian framework, this chapter will determine how the system will work and which functionalities is used to achieve research objectives. In the Bayesian framework section Kalman filtering is explained and then the metropolis Hastings Algorithm and how it is followed for a multivariate linear system. This chapter also presented the research design and development phases, in the manner of how work is performed in this thesis.

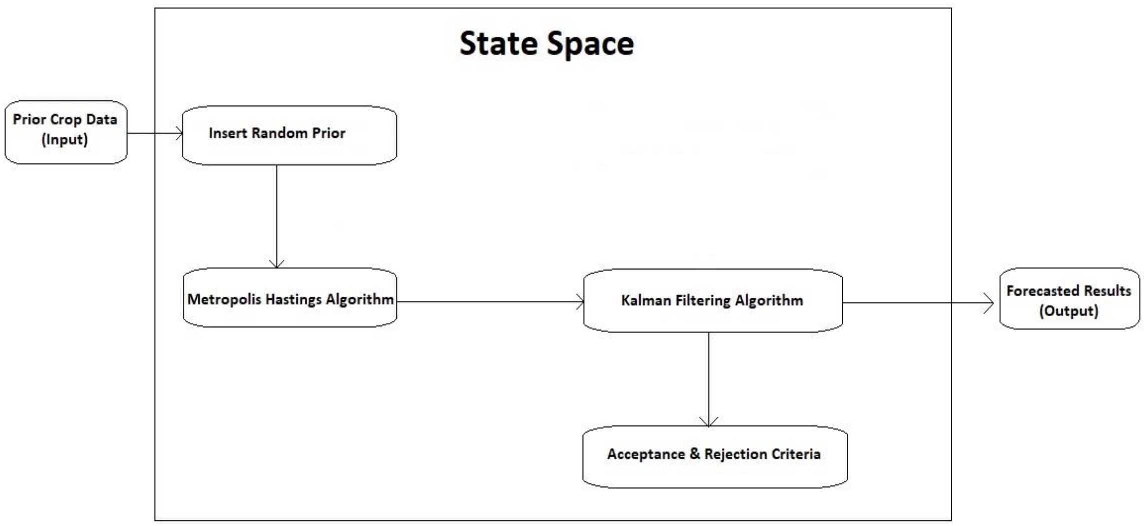


Figure 3. 1: Methodology

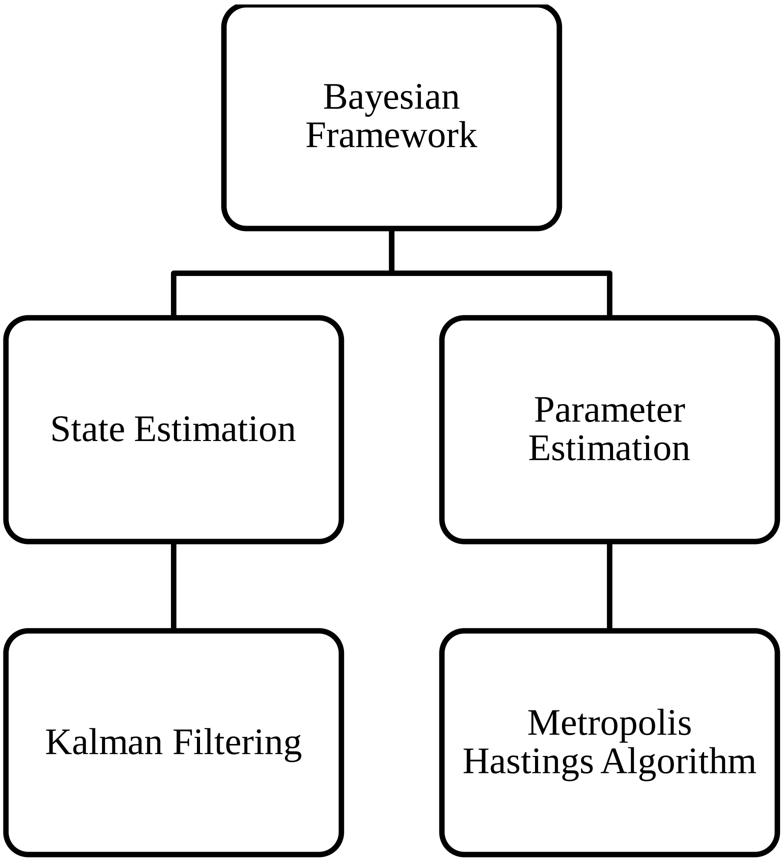


Figure 3. 2: Models parameters

3.2 Research Design and Development

This research has been designed in a very simple manner; this thesis consists of five phases as can be seen in Figure 3.03. First of all, literature has been studied related to the rabbi crops and autoregressive models. Then in the second phase, started collecting data of rabbi crops which include wheat data, potato data, and gram data, the dataset extracted consists of area, yield, production, and prices. In the third phase after data collection, now model was selected according to the data collected, after analysis, the Supervised learning model is selected, and in which, select multiple linear regression model. For this model, Kalman filtering is used and the metropolis algorithm is used to extract the required results, and all this simulation and their code generation is performed in MatLab. Then in the fourth phase, the results is discussed and analyzed and a comparison is done between different results. And then in the fifth and final phase, the results is concluded, and limitations and future work is identified.

Literature review	Rabbi Crops
	Regressive models
Data Collection	Data collection phase (Area, Production, Yield and Prices)
Model Identification and Algorithms Selection	Selection of Machine Learning model
	Supervised Learning Model
	Multiple linear regression
	Kalman Filtering and MCMC based Metropolis Hstings
Model Generation, and Implimentation	MATLAB code generation
	Simulations (MATLAB)
Results and Analysis	Results type: Graphs (Histogram, etc)
	Results will be analysed
Conclustion, limitations and Future work	Conclude our results
	limitation are set
	Future work will be proposed

Figure 3. 3: Research design and development phases

3.3 Bayesian Framework

A sampling of collected data is utilized in many fields and disciplines for system identification. We may follow two significant steps: Characterization and Parameter Estimation. Firstly, characterization of the system is to know whether it's linear or non-linear and secondly, parameter estimation of the agreed method depends on input and output calculations. This literature used the Bayesian technique for filtering that requires a state space approach [35]. Modeling of time series in measurements focuses consideration on the state vector of a function. To explain the execution of the function, the state vector comprises the necessary material for the process. On the other hand, kinematic properties of the target can be modeled by utilizing these state vectors, and so on [36]. . The measuring vector represents a visual observation associated with a regional vector. In most cases, the measurement vector is partially smaller than the state vector. The State-space tactic offers significant benefit over time collection strategies and also multivariate data management and non-Gaussian or non-linear processes [37]. Consequently analyzing and interpreting the dynamical arrangement, it's essential to enterprise two systems, the first system model describes the emergence of the situation over time, and the second is associated with the sound state scale [38].

In this thesis, formulate a linear state-space model which is:

$$X_n = F_n x_{n-1} + G_n w_n$$

$$Y_n = H_n x_n + v_n$$

$x_n \epsilon\, R^k$is the State,

$y_n \epsilon\, R^p$is the Measurement,

$G_n \epsilon\, R^{k*T}$, $F_n \epsilon\, R^{k*k}$and $H_n \epsilon\, R^{p*k}$ are the system Matrices,

And $w_n \epsilon\, R^T$and $v_n \epsilon\, R^p$ denote the process and measurement residual or noises that are white Gaussian with known covariance and zero mean, i.e., $w_n \sim N(0, Q_n)$ and $v_n \sim N(0, R_n)$. It's assumed that $w_n, v_n\ and\ x_0 \sim N(\hat{x}_0, P_0)$ are pair-wise uncorrelated at each sampling instant.

Bayesian estimation is divided into two main sections 1st part includes state estimation and 2nd part includes Parameter estimation: for state estimation filtering is performed, as in the current thesis Kalman filtering is used for state estimation. And for the parameter estimation phase Metropolis Algorithm is used for parameter estimation

3.3.1 State Estimation

For state estimation in current thesis Kalman filtering is used. The proposed filter is explained below:

- **Kalman filtering:**

Multiple linear regression is a familiar model in statistical mathematics, and it's also a mathematical approach to addressing multivariate relations. In financial problems, a variable is frequently tormented by more than one variable, so the maximum direct manner to perform such examination is the way of more than one or multiple linear regression system. While the relation is linear among dependent variables and independent variables, the overall shape of the multiple linear regression model may be stated:

$$\mathrm{Y(x) = Y_i(x) = X_i\beta + \varepsilon_i}$$

KF is an easy linear prediction technique [39] and offers one of the best solutions when the state and measurement equations are linear with Gaussian residual or noise arrangements. Kalman filter stays one of the maximum normally used model-based state estimators because of its robustness, simplicity, stability, and optimality [38], [40]. Depending on an essential linear version of machine dynamics. The Kalman filter makes use of input-output dimensions and probabilistic motives of uncertainty to produce a minimal mean squared optimal estimate of a system's inner state vector because it evolves in time.

Kalman filter clear cut out is a Recursive records/information processing algorithm and produces an optimum estimation of favored portions specified in the set of measurements. KF clear cut out is ideal for linear systems and Gaussian residuals or errors, KF clear cut out is the state-of-the-art estimate mostly established totally on all previous measurements. K F is recursive and does not want to save all preceding measurements and reprocess all records or data every time step.

Basic idea:

Based on the preceding record or data, make a prediction: $\hat{y}^{-}$,σ^{-}

⬇

Take measurement: z_k, σ_z

⬇

(ŷ) Optimal estimate = prediction + (kalman gain) * (measurement-prediction)

The variance of estimate = variance of prediction * (1 – Kalman gain)

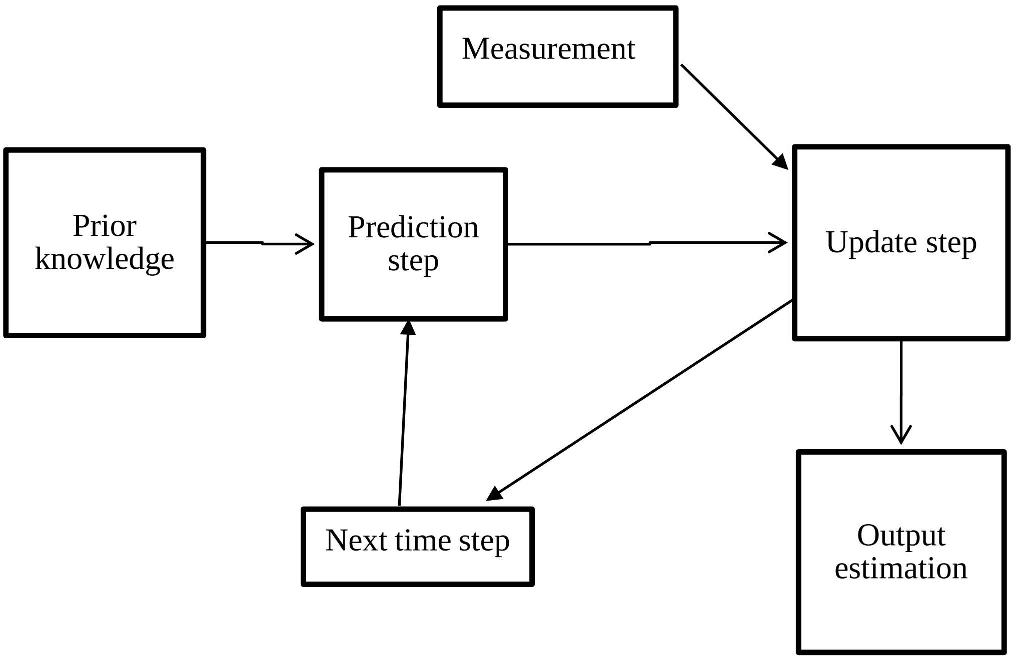


Figure 3. 4: Kalman filtering

KF algorithm is fragmented into 2 parts: 1st prediction and 2nd update.

Part 1 (Prediction). The state of the present time (k) is anticipated along with the previous estimation of the posterior time ($k-1$), and the earlier estimation of time k is acquired.

Part 2 (Update). Utilize the measured figure or value on the current time (k) to accurately the measured figure or value on the forecast degree and get the previously calculated figure or value on the current time (k).

Primary circumstances ($\hat{y}_{k-1}$ and σ_{k-1})

Prediction ($\hat{y}_k$, σ_k^-)

- Use the primary circumstances and model (e.g., constant velocity) to make a forecast or prediction.

Measurement (z_k)

- Takings measurement

Correction ($\hat{y}k$, σ_k)

- Use measurement to an accurate forecast by molding forecast and errors or residual – permanently a case of merging 2 Gaussians
- Optimal estimate along with minor variance

KF clear cut out may be divided right into a time renewal equation and a measurement renewal equation that is additionally referred to as a prediction equation and a correction equation. So KF clear-cut-out algorithm is a recursive predicting-correcting method. The center of the KF includes two (2) functions or processes and five (5) formulas.

The process to be estimated:

$y_k = Ay_{k-1} + Bu_k + w_{k-1}$ Process Noise (w) with covariance Q

$z_k = Hy_k + v_k$ Measurement Noise (v) with covariance R

1) The prediction manner is described as follows:

   $\hat{y}_k^-$ Is an estimate primarily based totally on measurements at preceding time steps [41].

$$\hat{y}_k^- = Ay_{k-1} + Bu_k$$

$$P_k^- = AP_{k-1}A^T + Q$$

2) The update process is defined as follows:

   $\hat{y}_k$ has extra information – the measurement at time k

$$\hat{y}_k = \hat{y}_k^- + K(z_k - H\hat{y}_k^-)$$

$$K = P_k^- H^T (HP_k^- H^T + R)^{-1}$$

$$P_k = (I - KH)P_k^-$$

$\hat{y}_{k-1}$ And $\hat{y}_k^-$ are they up-to-date results of filtering at $k-1$ and k. $\hat{y}_k$ is the previous state estimation at time k and $\hat{y}_{k-1}$is forecasted in keeping with the foremost estimation on the preceding time $k-1$. $\hat{y}_k^-$ is the result of the prediction method, P_k Is previous estimation covariance at time k that is an intermediate result of the filtering.

H_k Is a transformation matrix from a few state values to measurement variables. z_k are measured values (a vector of witnessed values), which might be input data or statistics of a filter. H_k is a filtering gain matrix, that's an intermediate calculation result, additionally known as Kalman gain or Kalman coefficient. A_k is a transition matrix to switch state vectors from one state to other, that's an estimation model for the target state transition. Q_k is technique excitation residual or noise covariance 'covariance of system process, that's used to symbolize errors among a state transition matrix and a real process. Researchers can't look at process indicators directly, so it's mile hard to decide Q_k value. Q_k is used to approximate state variables of the discrete-time procedure, which is likewise called noise or residual as a result of a forecast model or function. It's a covariance matrix of state transition [42].

3.3.2 State Estimation Model

Auto-Regressive Integrated Moving Average States

Non-seasonal ARIMA is characterized as; ARIMA $(\mathrm{p}, \mathrm{d}, \mathrm{q})$ model where

p: Autoregressive part

d: differences part needed for stationarity

q: Forecast lagged errors in the forecast equation

The forecast equation is created as follows, first, let x represent the d^th difference of X, [49] which means:

$If\ d = 0 : X_t = x_t$

$If\ d = 1 : X_t = x_t - x_{t-1}$

$If\ d = 2 : X_t = (x_t - x_{t-1}) - (x_{t-1} - x_{t-2}) = x_t - 2x_{t-1} + x_{t-2}$

In terms of x, the general forecasting equation is

$$\hat{x}_t = \mu + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$$

ARIMA (0,1,0)

ARIMA (0,1,0) is:

$$\hat{x}_t = x_{t-1} + \mu$$

And it is a simple model with residuals, the so-called random walk model. ARIMA(0,1,0) is a random walk model with a constant trend, written as Also called random walk with drift.

$$\hat{x}_t = x_{t-1} + \mu + c$$

ARIMA(0,0,0) is $y_t = \mu$ or white noise. ARIMA(p,0,0) is an autoregressive model and ARIMA(0,0,q) is a moving average model.ARIMA (1,1,0)

Differenced 1st-order Auto-Regressive model, If the error terms are auto correlated to a random walk model, perhaps the condition can be solved by adding 1 lag to the dependent variable to the forecast equation, such that, by regressing the 1st difference of X on the situation lagged by 1-time period. This can generate the forecast equation as:

$$\hat{X}_t - x_{t-1} = \theta_1(x_{t-1} - x_{t-2}) + \mu$$

$$\hat{X}_t - x_{t-1} = \mu$$

Rearrange

$$\hat{X}_t = x_{t-1} + \theta_1(x_{t-1} - x_{t-2}) + \mu$$

ARIMA (0,1,1)

SES along with Growth, an ARIMA (0,1,1) is:

$$\hat{X}_t = x_{t-1} + \alpha\epsilon_{t-1}$$

$$\epsilon_{t-1} = x_{t-1} - \hat{x}_{t-1}$$

by definition, it iscome:

$$\hat{X}_t = x_{t-1} + \alpha(x_{t-1} - \hat{x}_{t-1})$$

for $t > 1$. If $\alpha - 1 = \theta$ then this equation iscome:

$$\hat{X}_t = x_{t-1} - (1 - \alpha)\epsilon_{t-1}$$

$$\hat{X}_t = x_{t-1} - \theta_1\epsilon_{t-1}$$

ARIMA (1,1,1)

A model with an AR(1) term, an initial difference term, and an MA(1) term will have orders (1,1,1). A model ARIMA(1,1,1) with AR and MA terms is applied to the variables:

$$\mathrm{Z_t = x_t - x_{t-1}}$$

The first difference explains the linear trend of the data. Differential order refers to the first consecutive difference [50].

ARIMA (2,1,0)

ARMA (2,0) version, which includes 2 parameters in it, might be transformed into an AR version of order 2, or AR (2) model, considering that q is zero in MA. Therefore, this function is listed as:

$$X_t = \mu + (\theta_1 * (x_{t-1} - \mu)) + (\theta_2 * (x_{t-2} - \mu)) + \varepsilon_t$$

Where X_t = stationary time series (studying), μ = mean of time series X_t, ε_t = white noise with mean $= 0$ and constant variance, and θ_1 and θ_2 are parameters to be estimated [51].

ARIMA (1,0,0)

First order autoregressive model. If your series data are stationary and autocorrelated, you can predict as multiples of previous values by adding a constant. The forecasting equation is:

$$\vec{X}_t = \vec{\theta} x_{t-1} + \mu_{t-1}$$

The standard multiple linear regression State model is represented as follows:

$$X_t^i = \theta_0 + \theta_1 x_{t-1}^1 + \theta_2 x_{t-1}^2 + \cdots + \theta_4 x_{t-1}^4 + \mu_{t-1}$$

Think about the regression equation:

$$X_t^1 = \theta_0 + \theta_1 x_{t-1}^1 + \theta_2 x_{t-1}^2 + \cdots + \theta_4 x_{t-1}^4 + \mu_{t-1}^1$$
$$X_t^2 = \theta_0 + \theta_1 x_{t-1}^1 + \theta_2 x_{t-1}^2 + \cdots + \theta_4 x_{t-1}^4 + \mu_{t-1}^2$$
$$X_t^3 = \theta_0 + \theta_1 x_{t-1}^1 + \theta_2 x_{t-1}^2 + \cdots + \theta_4 x_{t-1}^4 + \mu_{t-1}^3$$
$$X_t^4 = \theta_0 + \theta_1 x_{t-1}^1 + \theta_2 x_{t-1}^2 + \cdots + \theta_4 x_{t-1}^4 + \mu_{t-1}^4$$

Identifying vectors and matrixes:

$$\vec{x} = \begin{bmatrix} x_{t-1}^1 \\ x_{t-1}^2 \\ x_{t-1}^3 \\ x_{t-1}^4 \end{bmatrix} \text{ and } \vec{\mu} = \begin{bmatrix} \mu_{t-1}^1 \\ \mu_{t-1}^1 \\ \mu_{t-1}^1 \\ \mu_{t-1}^1 \end{bmatrix}$$

$$\vec{x}_t = (x_{t-1}^0, x_{t-1}^1, x_{t-1}^2, \dots, x_{t-1}^4)$$

$$\vec{\theta} = (\theta_0, \theta_1, \theta_2, \dots, \theta_4)$$

The variable associated with observation can be written as:

$$X_t = \begin{bmatrix} x_{t-1}^0 & x_{t-1}^1 & \cdots & x_{t-1}^4 \\ x_{t-1}^0 & x_{t-1}^1 & \cdots & x_{t-1}^4 \\ x_{t-1}^0 & x_{t-1}^1 & \cdots & x_{t-1}^4 \\ x_{t-1}^0 & x_{t-1}^1 & \cdots & x_{t-1}^4 \end{bmatrix} = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vec{x}_3 \\ \vec{x}_4 \end{bmatrix}$$

By taking the dot product observation $1-4$ vector form for each observation

$$X_t^1 = \vec{\theta}.\vec{x}_t^1 + \mu_{t-1}^1$$

$$X_t^2 = \vec{\theta}.\vec{x}_t^2 + \mu_{t-1}^2$$

$$X_t^3 = \vec{\theta}.\vec{x}_t^3 + \mu_{t-1}^3$$

$$X_t^4 = \vec{\theta}.\vec{x}_t^4 + \mu_{t-1}^4$$

The variable associated with State estimation can also be written as:

$$X_t\,.\vec{\theta} = \begin{bmatrix}\vec{x}_1\\ \vec{x}_2\\ \vec{x}_3\\ \vec{x}_4\end{bmatrix}\begin{bmatrix}\theta_0\\ \theta_1\\ \vdots\\ \theta_4\end{bmatrix}$$

We can write an entire set of regression equations in matrix form:

$$\begin{bmatrix}X_t^1\\ X_t^2\\ \vdots\\ y_4\end{bmatrix} = \begin{bmatrix}x_{t-1}^0 & x_{t-1}^1 & \cdots & x_{t-1}^4\\ x_{t-1}^0 & x_{t-1}^1 & \cdots & x_{t-1}^4\\ x_{t-1}^0 & x_{t-1}^1 & \cdots & x_{t-1}^4\\ x_{t-1}^0 & x_{t-1}^1 & \cdots & x_{t-1}^4\end{bmatrix}\begin{bmatrix}\theta_0\\ \theta_1\\ \theta_2\\ \vdots\\ \theta_4\end{bmatrix} + \begin{bmatrix}\hat{\mu}_1\\ \hat{\mu}_2\\ \hat{\mu}_3\\ \hat{\mu}_4\end{bmatrix}$$

Or

$$\vec{X}_t = \vec{\theta}x_{t-1} + \mu_{t-1}$$

3.3.3 Kalman Filter Algorithm

Algorithm Kalman filter**:** **(**u_{t-1}, $\sum_{t-1}$, , $\boldsymbol{u_t}$, $\boldsymbol{z_t}$)

Prediction:

$$\vec{u}_t = A_t u_{t-1} + B_t u_t$$

$$\vec{\sum}_t = A_t \vec{\sum}_{t-1} A_t^T + R_t$$

Correction:

$$K = \vec{\sum}_t C_t^T (C_t \vec{\sum}_t C_t^T + Q_t)^{-1}$$

$$u_t = \vec{u}_t + K_t(z_t - C_t\vec{u}_t)$$

$$\vec{\sum}_t = (I - k_t C_t)\vec{\sum}_t$$

Return: u_t, $\sum_t$

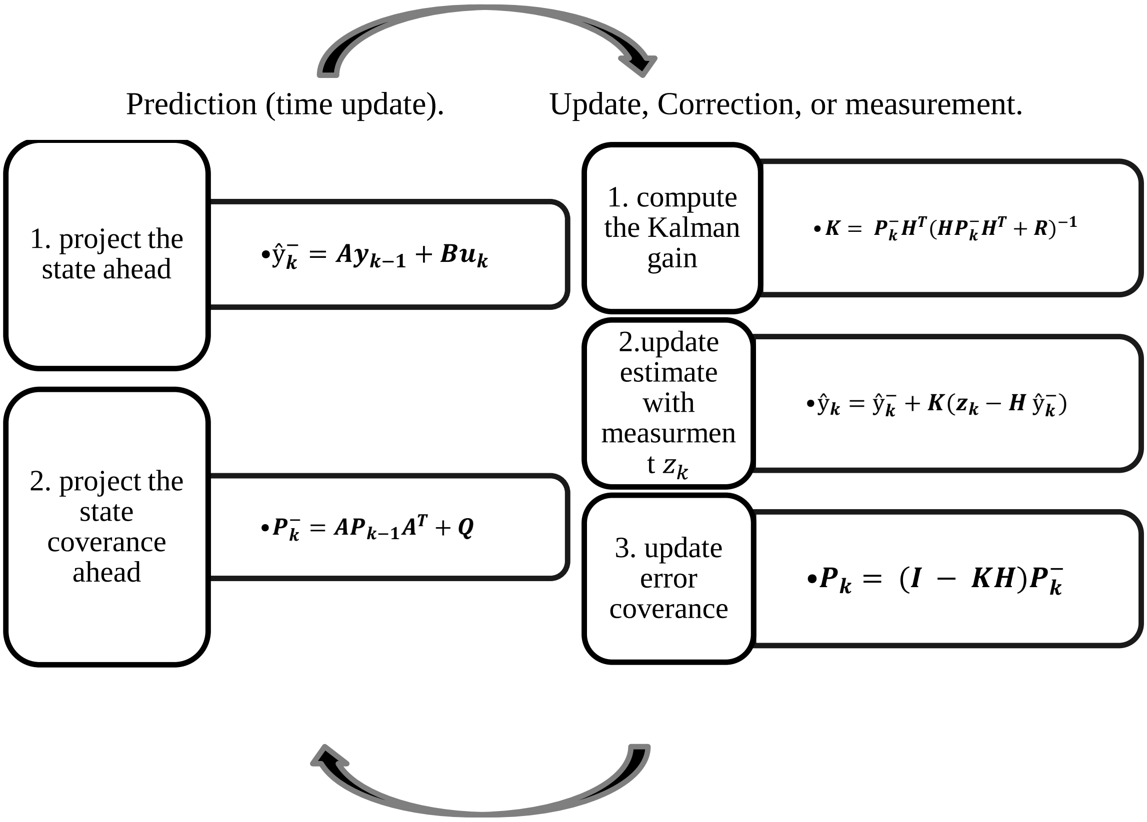


Figure 3. 5: Kalman Filtering Algorithm

3.3.4 Parameter Estimation:

Parameter estimation is a commonly used method for calculating the probability density function of random variables through unknown parameters. For parameter estimation in current thesis MCMC (Markov Chain Monte Carlo) and Metropolis Hasting algorithm is used. The proposed algorithm is explained below.

- **MCMC Markov Chain Monte Carlo**

Bayesian Filter Monte Carlo AMCMC for Adaptive Markov Chain, Estimation of Unknown Parameters $\theta \in R^d$ in its stochastic differential equations (SDEs),

$$dx = f(x,t;\ \theta)dt + L(x,t;\ \theta)dW,$$

wherein $x(t) \in R^n$, and $W(t)$ is an *n*-dimensional vector of independent standard Wiener processes. Above, $f(\cdot)$ is the non-linear drift function and $L(\cdot)$ is the dispersion matrix of the SDE.

The primary difficulty in the parameter estimation problem is that the transition density of the SDE in the above equation cannot be evaluated in closed form. There exists an extensive variety of strategies for estimating parameters of the SDE models, which keep away from this problem. These consist of simulated most likelihood-based strategies or methods, Markov Chain Monte Carlo-based strategies or methods, in addition to Exact Algorithm-based strategies or methods. It is likewise feasible to immediately approximate the answers or solutions of the Fokker Planck Kolmogorov 'FPK' equation with the aid of using well-known numerical strategies or methods for fixing or solving PDEs.

In this article, observe some other elegance of methods, that are based totally on forming a Gaussian (process) approximation to the parameter conditioned diffusion process. This technique is attached to the Taylor collection or series approximations used withinside the extended Kalman filter [43] that's an extensively used technique or approach in control, guidance, goal tracking, and different applications. Given the Gaussian approximation, it's miles viable to assess the corresponding marginal likelihood or probability of the parameters and similarly, the approximate un normalized marginal posterior density through the use of so-referred to as prediction error or residual decomposition. Similar approximations have additionally been proposed withinside the context of linear noise approximations 'LNA' of sour Markov methods or processes and the associated master or grasp equations.

The use of MCMC or Markov chain Monte Carlo strategies rather than the ML estimates withinside the EKF primarily based totally on the SDE parameter estimation framework become currently investigated [44]. To enhance the overall performance of the simple Metropolis-Hastings based totally on MCMC sampling, [44] proposed using Hamiltonian Monte Carlo 'HMC' method. The maximum probability of likelihood is the estimation in SDE functions, and the use of Gauss Hermite quadrature approximations of diffusions,

In this article, the purpose is to observe the accuracy and computational necessities of Gaussian approximation primarily based totally on parameter estimation methods while they may be combined with MCMC Markov chain Monte Carlo strategies or methods, and in particular, with adaptive MCMC Markov chain Monte Carlo strategies or methods. More specifically, we observe using these days developed Gaussian quadrature 'sigma point' primarily based totally on Gaussian approximations developed for continuous and discrete filtering and smoothing withinside the SDE parameter estimation problem. [45]

3.3.5 Observation Estimation Model

The standard multiple linear regression model is represented as follows:

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i,1} + \hat{\beta}_2 x_{i,2} + \cdots + \hat{\beta}_4 x_{i,4} + \hat{\mu}_i$$

Such that $i = 1,2,\dots,4$.

$$x_{i,j} \qquad\qquad j = 0,1,\dots,4.$$

$\hat{\mu}_i$ Associated with i.

$$\hat{\beta}_j$$

$$\hat{\beta}_0 = \hat{\beta}_0 x_{i,0}$$

$$x_{i,0} = 1$$

$$y_i = \hat{\beta}_0 x_{i,0} + \hat{\beta}_1 x_{i,1} + \hat{\beta}_2 x_{i,2} + \cdots + \hat{\beta}_4 x_{i,4} + \hat{\mu}_i = \hat{y}_i + \hat{\mu}_i$$

$$\hat{\mu}_i = y_i - \hat{y}_i$$

Actual output values in the dataset

$$y_i = \hat{y}_i + \hat{\mu}_i$$

Predicted values along the regression line

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Residual or error at each observation

$$\hat{\mu}_i = y_i - \hat{y}_i$$

Think about the regression equation as a whole

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i,1} + \hat{\beta}_2 x_{i,2} + \cdots + \hat{\beta}_4 x_{i,4} + \hat{\mu}_i$$

$$y_1 = \hat{\beta}_0 x_{1,0} + \hat{\beta}_1 x_{1,1} + \hat{\beta}_2 x_{1,2} + \cdots + \hat{\beta}_4 x_{1,4} + \hat{\mu}_1$$

$$y_2 = \hat{\beta}_0 x_{2,0} + \hat{\beta}_1 x_{2,1} + \hat{\beta}_2 x_{2,2} + \cdots + \hat{\beta}_4 x_{2,4} + \hat{\mu}_2$$

$$y_3 = \hat{\beta}_0 x_{3,0} + \hat{\beta}_1 x_{3,1} + \hat{\beta}_2 x_{3,2} + \cdots + \hat{\beta}_4 x_{3,4} + \hat{\mu}_3$$

$$y_4 = \hat{\beta}_0 x_{4,0} + \hat{\beta}_1 x_{4,1} + \hat{\beta}_2 x_{4,2} + \cdots + \hat{\beta}_4 x_{4,4} + \hat{\mu}_4$$

Identifying vector and matrixes: values on the left-hand side and residual or errors on the right-hand side, $4 * 1$ dimensional (column) vectors given by:

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \qquad \text{And } \vec{\mu} = \begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{\mu}_3 \\ \hat{\mu}_4 \end{bmatrix}$$

The variable associated with observation can be written as:

$$X = \begin{bmatrix} x_{1,0} & x_{1,1} & \cdots & x_{1,4} \\ x_{2,0} & x_{2,1} & \cdots & x_{2,4} \\ x_{3,0} & x_{3,1} & \cdots & x_{3,4} \\ x_{4,0} & x_{4,1} & \cdots & x_{4,4} \end{bmatrix} = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vec{x}_3 \\ \vec{x}_4 \end{bmatrix}$$

$$\vec{x}_1 = (x_{1,0}, x_{1,1}, x_{1,2}, \ldots, x_{1,4})$$

$$\vec{\beta}_1 = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_4)$$

By taking the dot product observation $1 - 4$ vectors from for each observation

$$y_1 = \vec{x}_1 . \vec{\beta} + \hat{\mu}_1$$

$$y_2 = \vec{x}_2 . \vec{\beta} + \hat{\mu}_2$$

$$y_3 = \vec{x}_3 . \vec{\beta} + \hat{\mu}_3$$

$$y_4 = \vec{x}_4 . \vec{\beta} + \hat{\mu}_4$$

Multiplying a matrix by a vector

$$X\vec{\beta} = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vec{x}_3 \\ \vec{x}_4 \end{bmatrix} \vec{\beta} = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vec{x}_3 \\ \vec{x}_4 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_4 \end{bmatrix}$$

We can write an entire set of regression equations in matrix form:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_4 \end{bmatrix} = \begin{bmatrix} x_{1,0} & x_{1,1} & \cdots & x_{1,4} \\ x_{2,0} & x_{2,1} & \cdots & x_{2,4} \\ x_{3,0} & x_{3,1} & \cdots & x_{3,4} \\ x_{4,0} & x_{4,1} & \cdots & x_{4,4} \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_4 \end{bmatrix} + \begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{\mu}_3 \\ \hat{\mu}_4 \end{bmatrix}$$

Or

$$\vec{y} = X\vec{\beta} + \vec{\mu}$$

3.3.6 Metropolis-Hastings Algorithm

MH and additional MCMC procedures are commonly used for sampling from multivariate distributions, in particular, while the wide variety of dimensions is high.

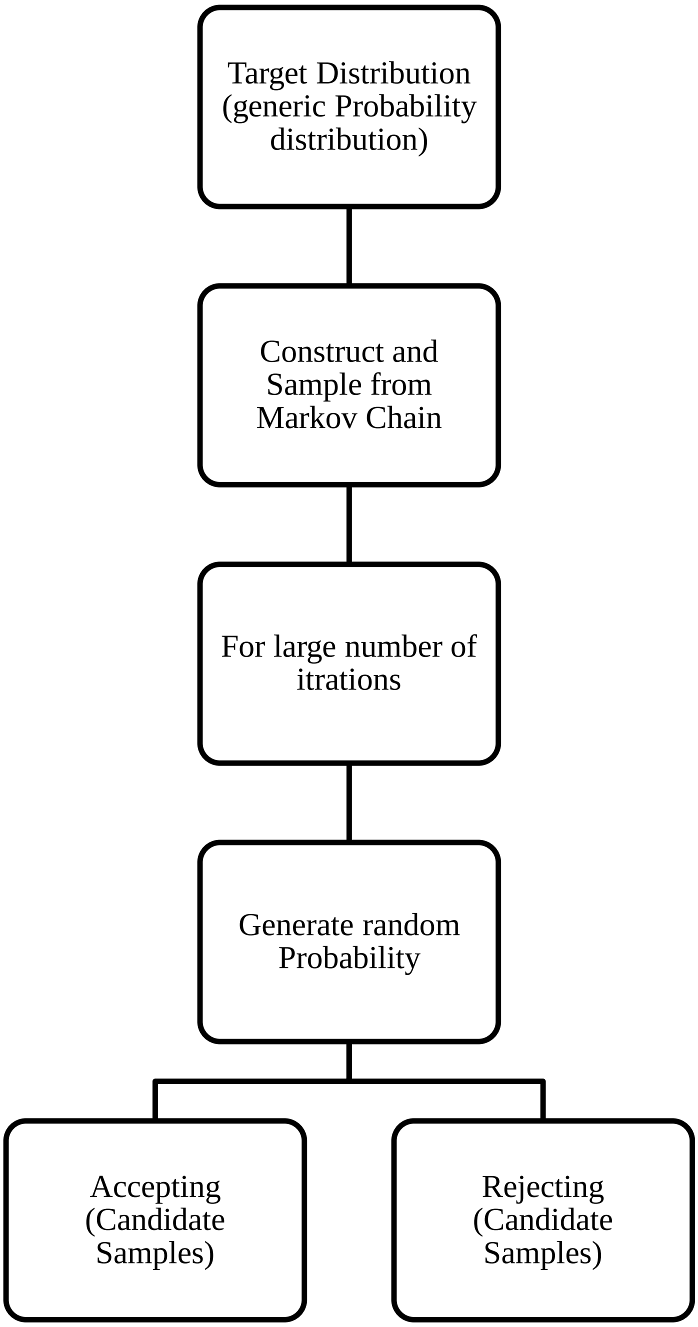


Figure 3. 6: Metropolis-Hastings

Target state model:

$$x_k = f_{k-1}^{\theta}(x_{k-1}) + v_{k-1}, k = 0, \ldots, M$$

Whereas the measurement state model is:

$$z_k = h_k^{\theta}(x_k) + w_k, k = 0, \ldots, M$$

f_{k-1}^{θ} and h_k^{θ} are linear functions in the equation, v_{k-1} and w_k are state noise and measurement noise respectively, and θ is the unidentified parameter to be anticipated withinside the state space arrangement.

MCMC approach is the effective approach for approximate computation of expectation integrals. The MH algorithm is an MCMC approach [45]. This pool may be used to calculate integrals like

$$E\{g(\theta)|\ Zk\} = \int R^d\ g(\theta)\ P(\theta|Zk)\ d\theta$$

or approximate the distribution $P(\theta|Z_k) \propto P(Z\_k\ |\theta)P(\theta)$ is understood simplest as much as a normalization constant.

Step 01.	Let $P(Z_{0:M}|\theta)$is the un-normalized target density and $\theta^{(n)}$ is the current value and $q(\theta|\theta^{(n)})$ be a proposed distribution then
Step 02.	
	do
	Start $\theta^{(0)}{\sim}\pi_0$
	At time step n assume $\theta^{(n)}$ is given
	At time step $n+1$ generate sample $\theta^{*}{\sim}q(\theta|\theta^{(n)})$.
	Evaluate the probability of Acceptance
	$\alpha = \min\{1, \frac{P(\theta^{*}|Z_{0:M})q(\theta^{(n)}|\theta^{*})}{P(\theta^{(n)}|Z_{0:M})q(\theta^{*}|\theta^{(n)})}\}$
	Set $\theta^{(n+1)} = \theta^{*}$with probability α
	else $\theta^{(n+1)} = \theta^{(n)}$with probability $1-\alpha$
	if $n \to \infty$ we will have the true posterior $\theta^{(n)} \sim P(\theta|Z_{0:M})$

Figure 3. 7: Metropolis-Hastings Algorithm

The previous distribution is commonly clean to estimate, however, but the difficult component in parameter estimation is the assessment of the marginal likelihood of posterior inference on $P(Z_k|\theta)$, wherein states are being included out. The use of particle clear cut-out filter for approximation of the likelihood method is also a collective relevant methodology [46]. The excessive computational sources required for this technique are the principal disadvantage. In precept following technique is in widespread used for efficient or green computing of the marginal likelihood $P(Zk|\theta)$. Computation of the prediction densities $P(x_k|Z_{k-1},\theta)$ and filtering densities $P(x_k|Z_k,\theta)$ for $k = 1,\ldots,M$ through the subsequent algorithm recursively.

Step 01.	Start from the prior $P(x_0\|\theta) \equiv P(x_0\|Z_0,\theta)\forall z_k$ for k=1,...,M
Step 02.	
	do
	Prediction Step: prediction density $P(x_k\|Z_{k-1},\theta)$ could be evaluated afterintegration of Chapman equation from initial condition $P(x_{k-1}\|Z_{k-1},\theta)$ to time t_k
	Update Step: For computing the filtering density of time step k use Bays rule:

Figure 3. 8: Filtering in Parameter Estimation

For the development of marginal likelihood representation, filtering outcomes may be utilized as below. Factors of marginal likelihood can be similar to

$$P(Z_k|\theta) = \prod P(z_k|Z_{k-1},\theta)$$

Recursive calculation of the term specified in the product could be primarily based totally on a filtering solution.

$$P(z_k|Z_{k-1},\theta) = \int P(z_k|x_k,\theta)P(x_k|Z_{k-1},\theta)dxk$$

$$\approx \frac{1}{N}\left(\sum_{j=1}^{N} P(Z_k|x_k^j\right)$$

Where $x_k^j \sim P(x_k|Z_{k-1})$ is filtering.

To implement Markov Chain Monte Carlo, sampling researchers seek formulas for marginal probability functions. Determining the posterior distribution and normalization constant is straightforward.

3.4 Posterior Inference

In the evaluation of parameters with MCMC techniques, we need to discover a manner to assess the marginal posterior opportunity probability density of the parameters $p(\theta|Y_M) \propto p(Y_M|\theta)p(\theta)$ as much as a normalization constant. The previous is normally clean to calculate, however, withinside the case of Stochastic Differential Equation parameter estimation the thought component is the assessment of the marginal likelihood chance $p(Y_M|\theta)$, wherein the states were incorporated out. One typically relevant method is to approximate his likelihood chance via way of means of the use of particle filtering, which results in so-known as PMCMC Particle Markov chain Monte Carlo methods [47].

The marginal likelihood probability $p(Y_M|\theta)$ can be, in principle, computed effectively as follows. The classical filtering theory [48] states that it will calculate the prediction or forecasting densities $p(x(t_k)|Y_{k-1},\theta)$ and the filtering densities $p(x(t_k)|Y_k,\theta)$ for $k = 1,\ldots,T$ through the subsequent recursive algorithm:

Start from the earlier posterior $p(x(t_0)|\theta) \equiv p(x(t_0)|Y_0,\theta)$.

For every dimension y_k for $k = 1,\ldots,M$ do:

Prediction step: Integrate the FPK equation from the preliminary circumstance $p(x(t_{k-1})|Y_{k-1},\theta)$ to time t_k , which ends up resulting withinside the anticipated or predicted density $p(x(t_k)|Y_{k-1},\theta)$.

Update step: Use Bayes' rule for calculating the filtering density for k step:

$$p(x(t_k)|Y_k,\theta) = \frac{p\big(y_k\big|x(t_k),\theta\big)p\big(x(t_k)\big|Y_{k-1},\theta\big)}{\int p\big(y_k\big|x(t_k),\theta\big)p\big(x(t_k)\big|Y_{k-1},\theta\big)dx(t_k)},$$

The filtering consequences or results may be used for building an expression for the marginal likelihood probability as follows. The marginal likelihood probability may be factored as:

$$p(YM|\theta) = \prod_{k=1}^{M} p(yk|Yk-1,\theta).$$

Given the filtering solution, the phrases withinside the product may be calculated recursively as

$$p(y_k|Y_{k-1},\theta) = \int p(y_k|x(t_k),\theta)p(x(t_k)|Y_{k-1},\theta)dx(t_k)\,.$$

Once we've acquired the expression for the marginal likelihood probability, we will effectively compare the posterior distribution as much as the normalization constant, that's all it wants for the execution of MCMC sampling.

3.5 Summary

This chapter has discussed the methodology of this thesis which components is used how they is used and when these components is used. This chapter contains material related to our Bayesian modeling i.e. state estimation model, and observation estimation model for our model to function, it also contains our formulae and equations for forecasting results. This thesis discussed in detail how different algorithms will combine and work, and will provide us with the required outputs. This chapter also discussed the design and development phases in this thesis and also provide a table where they can be understood easily.

CHAPTER 4

RESULTS AND ANALYSIS

4.1 Overview

This chapter includes clarifying unclear points and presenting data-driven or represent results with a short discussion easily explaining them at the end for the target audience. This chapter also provides negative results, stability, accuracy, and other values and provides justification of the method used and their technical explanation.

4.2 Practical Implementation

MATLAB stands for matrix laboratory. MATLAB is a multi-paradigm programming platform and numeric computing environment developed by Math-Works especially for engineers and scientists to design and analyze matrix manipulations, algorithm development, plotting of functions and data, modeling, simulation, and prototyping, data analysis, visualization, scientific and engineering graphical user interface (GUI) building, and additional products that transform our world. In another argument, we can say that MATLAB is used for high-performance technical computing, it inaugurates programming, computation, and visualization in an easy and comfortable environment [52].

A MATLAB system consists of five main parts: (1) the MATLAB language, (2) the MATLAB work environment, (3) the MATLAB math function library, (4) handle graphics, and (5) the MATLAB application program interface (API).

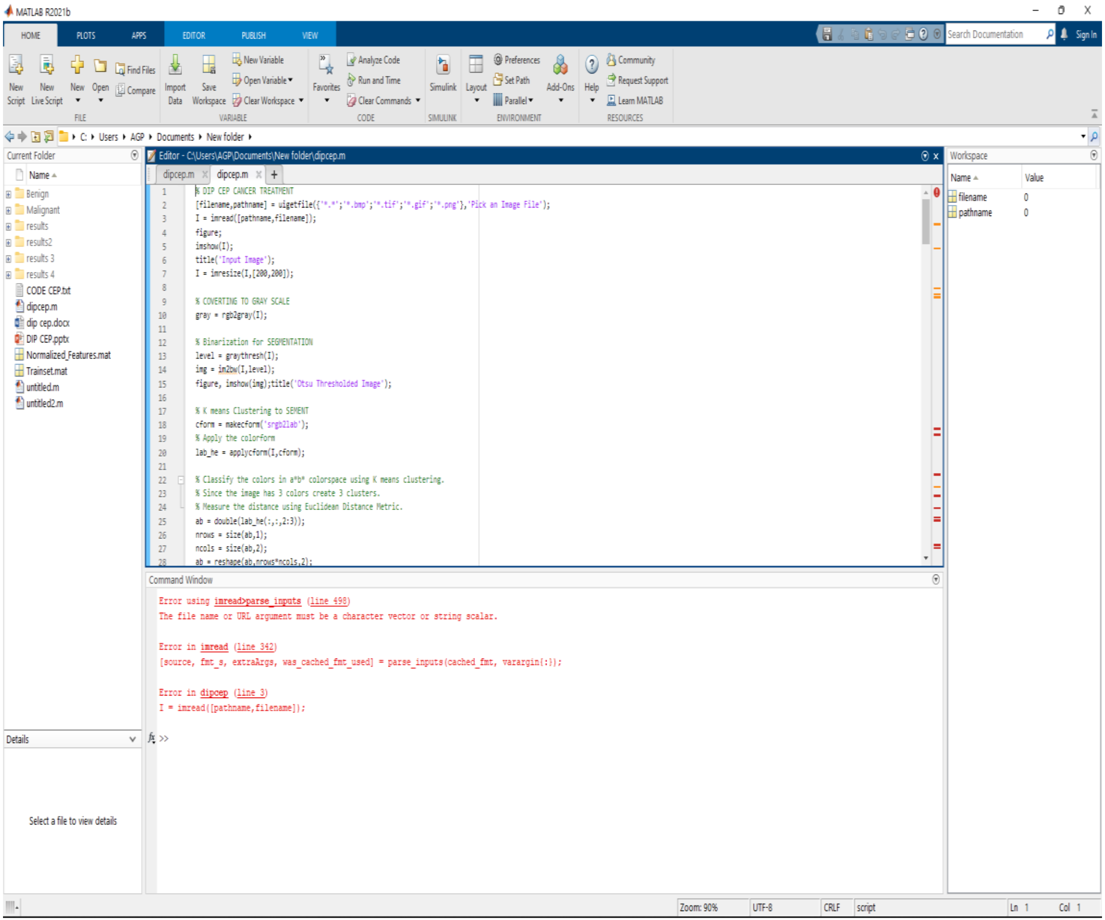


Figure 4. 1: MATLAB

Code generated for simulations on MATLAB can be seen in the below Figures 4.2-4.4, it involves the MATLAB language, the mathematical functions or linear regression functions, and matrix manipulations, by using MATLAB-API and its working environment forecasting is done through the below-mentioned code and the results are obtained in Graphical form (histograms). Results are attached and discussed in the next section.

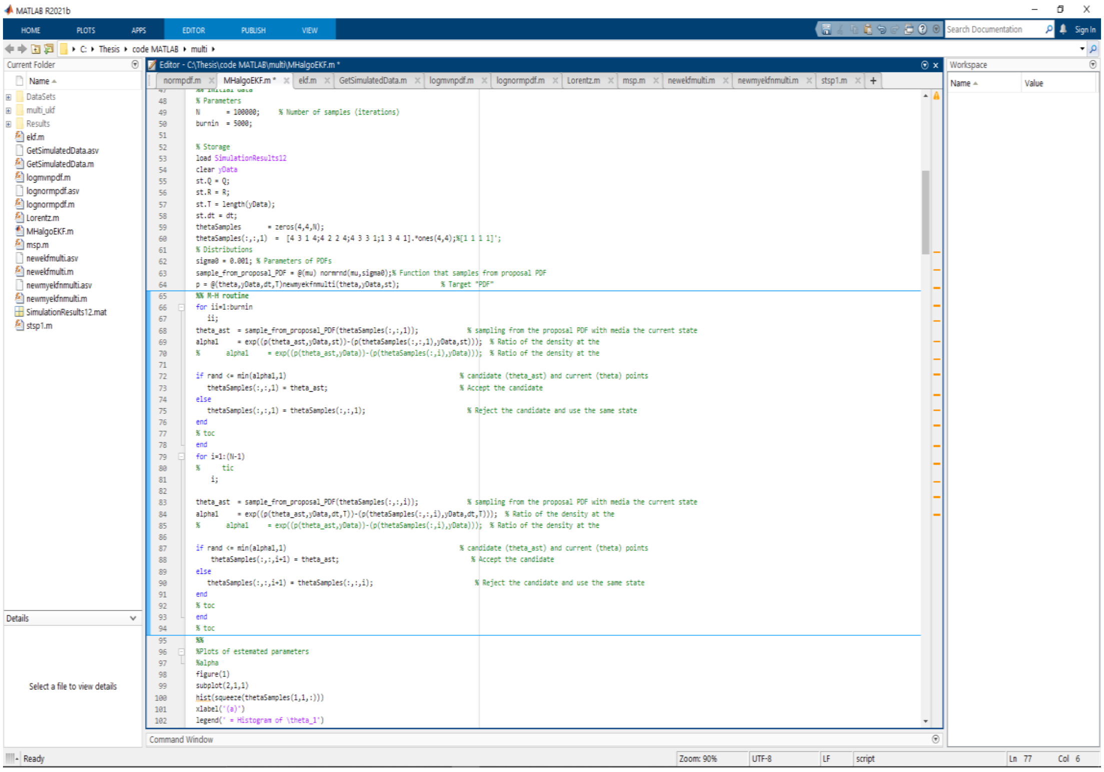


Figure 4. 2: Function Metropolis-Hastings

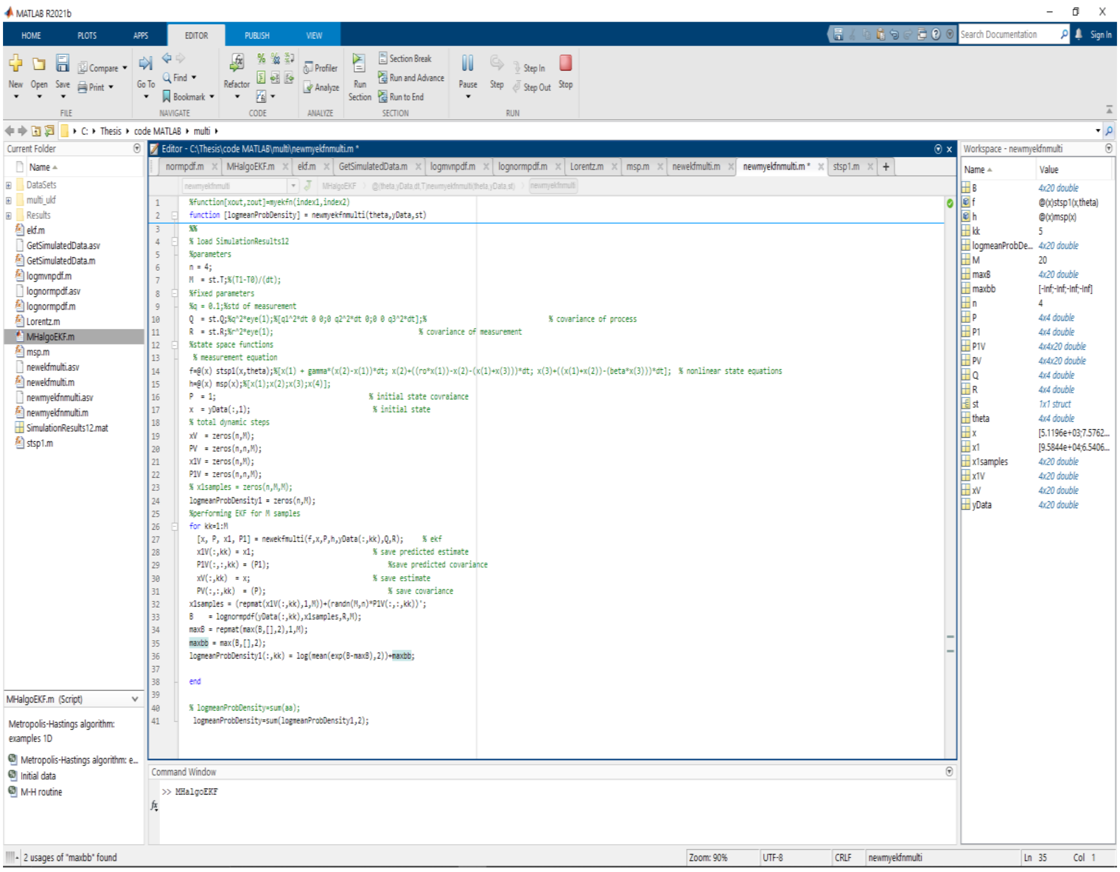


Figure 4. 3: Function Kalman Filter

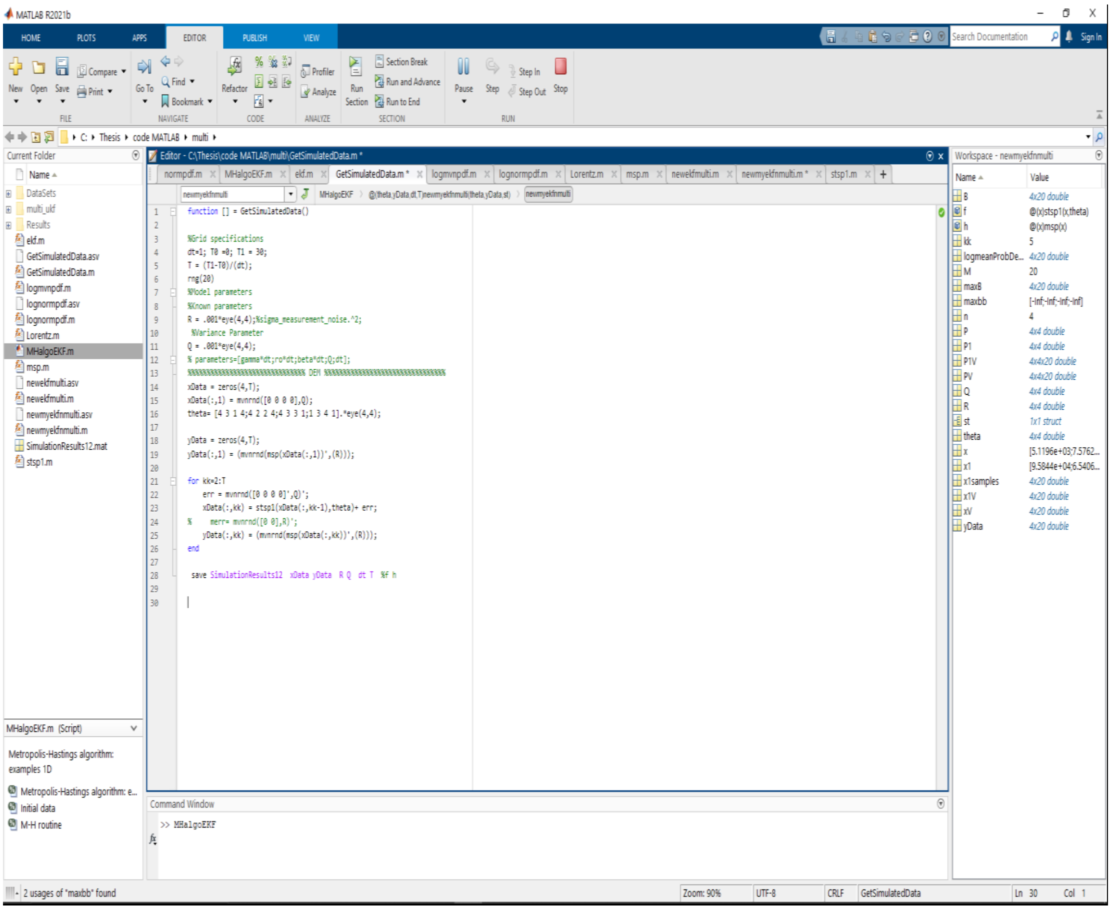


Figure 4. 4: Function Get Simulated Data

4.3 Results and discussion

Agriculture improvement is one of the vital approaches to eliminating severe poverty and enhancing shared prosperity and feed. In Pakistan Rabbi food crops we have selected three main and important crops of the country, which are Wheat, Potato, and Gram. Pakistan achieved a substantial increase in agriculture production in the last two decades, however, going forward the country is facing various shocks due to climatic patterns, economic conditions, and agricultural policy.

Statistics show that wheat occupies about 72 percent of the total cultivated area followed by the gram which is 8 percent, and other crops. From Pakistan Rabbi food crops we have selected three main crops of the country and collected data from those three crops. We have selected four main attributes of the Rabbi crops which are cultivation area, production of crops, Yield of crops, and prices of crops. While viewing data of all three of our crops, it has been noticed that the wheat and potato graphs are not that much different in sighting their graphical representation, both wheat and potato cultivated area, production, yield, and prices are going uptrend along with time and linearity is in the data representation of both of these crops.

Gram, also known as chickpea, is an important pulse crop both in terms of production and consumption in Pakistan. While viewing Gram data, its cultivated area, production, yield, and prices, there are too many random variations in the data, especially in the production and yield dataset there are too many variations.

The purpose is now to use the Bayesian state-space Approach and try to deal with that random variation, especially in the gram data, and try to figure out the best-predicted outcome for our all of crops but our main objective is to deal with gram and present the best gram predicted outcomes for future use.

A multivariate linear regression model is used to estimate the Determinants of Rabbi Food Crops (Wheat, Potato, and Gram) Area, Production, yield, and Prices in Pakistan. The result of this regression is given below in below figures and tables

Gram production in Pakistan was analyzed using a multivariate ARMA-based Bayesian State-Space model. Where multivariate linear regression integrated along with Bayesian State-Space framework is used to forecast Pakistan rabbi food crops production with regards to multiple variables which include Prices, Cultivation Area, Production and yield data of crops., to estimate best results number of samples (iterations) or $N = 100000$ and $burnin = 5000$ which means if set burnin to a high enough value that researchers believe that the Markov Chain approximately reaches stationarity after burnin samples, which helps in forecasts based on these high number of samples iterations close enough to the actual values.

The main data for this study was rabbi food crops prices Rs. per ‘000’ kgs, cultivation area ‘000’ hectares, production ‘000’ tons, and yield ‘000’ haters data from 1980 to 2019 data were collected from the publications of Bureau of Statistics and Agriculture Statistics of Pakistan, Ministry of National Food Security and Research (MNSF&R).

As the first step, the monthly time series data from 1980 to 2019 was decomposed into different components to study the temporal pattern. And then with the help of multivariate linear regression, these variables forecast would be analyzed concerning each other's impact on each other for all three Rabbi crops including Gram. Results are attached in the below section one by one and their attributes impact on each other and what impact they will put by combing their impact on crops.

Due to four multiple variables, $4 * 4 = 16$ times they will forecast the forecasted histogram or results because of a multiple linear regression $4 * 4$ matrix was formed for our model so the forecasted results is also generated 16 times and are represented by θ. Each theta represents different variables' impact on each other, the below Table 4.1 put downlight on it.

Table 4.1: Theta θ representation

Theta $\boldsymbol{\theta}$ Representation	Price	Cultivation Area	Production	Yield
Price	$\boldsymbol{\theta_1}$ Price vs the previous price	$\boldsymbol{\theta_2}$ Price vs Cultivation Area	$\boldsymbol{\theta_3}$ Price vs Production	$\boldsymbol{\theta_4}$ Price vs Yield
Cultivation Area	$\boldsymbol{\theta_5}$ Cultivation Area vs Price	$\boldsymbol{\theta_6}$ Cultivation Area vs previous Cultivation Area	$\boldsymbol{\theta_7}$ Cultivation Area vs Production	$\boldsymbol{\theta_8}$ Cultivation vs Area Yield
Production	$\boldsymbol{\theta_9}$ Production vs Price	$\boldsymbol{\theta_{10}}$ Production vs Cultivation Area	$\boldsymbol{\theta_{11}}$ Production vs the previous Production	$\boldsymbol{\theta_{12}}$ Production vs Yield
Yield	$\boldsymbol{\theta_{13}}$ Yield vs Price	$\boldsymbol{\theta_{14}}$ Yield vs Cultivation Area	$\boldsymbol{\theta_{15}}$ Yield vs Production	$\boldsymbol{\theta_{16}}$ Yield vs the previous Yield

4.3.1 Wheat

Given the importance of wheat as the leading crop cultivated during the rabbi season and as the stable food grain consumed throughout the year, we forecast wheat at the country-level

With Respect to Price Forecasted Results:

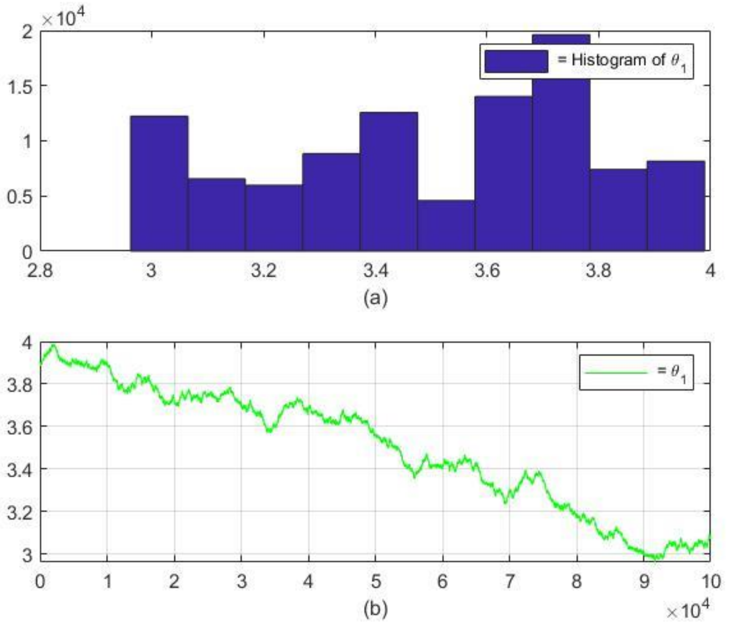


Figure 4. 5: θ_1 (Wheat Price vs Previous Price)

Price at $X_{1|t}$ $(t = current\ time)$ concerning the previous price at x_{t-1} $(t - 1 = Previous\ time\ step)$ and the convergence rate is 3.7.

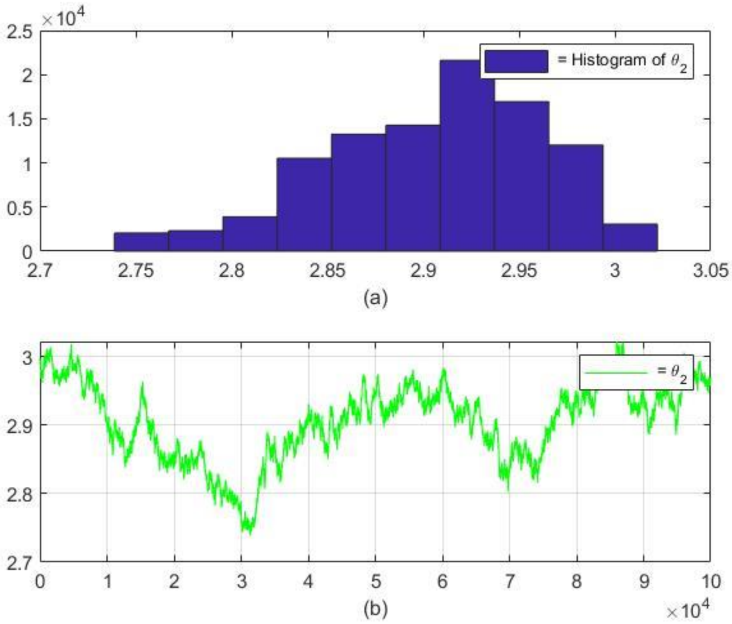


Figure 4. 6: θ_2 (Wheat Price vs Area)

Current Price at $X_{1|t}$ $(t = current\ time)$ concerning the Current Cultivation Area at x_t $(t = current\ time\ step)$ and the convergence rate is 2.93.

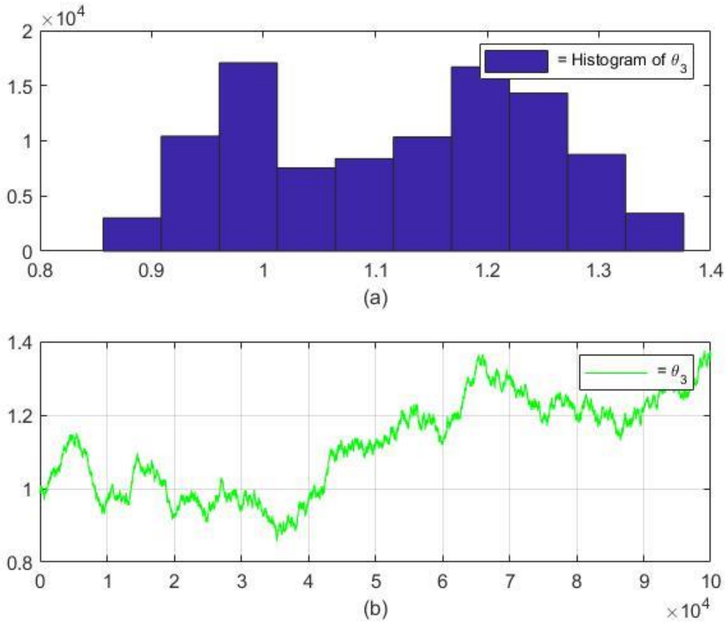


Figure 4. 7: θ_3 (Wheat Price vs production)

Current Price at $X_{1|t}$ ($t = current\ time$) concerning Current Production at x_t ($t = current\ time\ step$) and the convergence rate is 1.

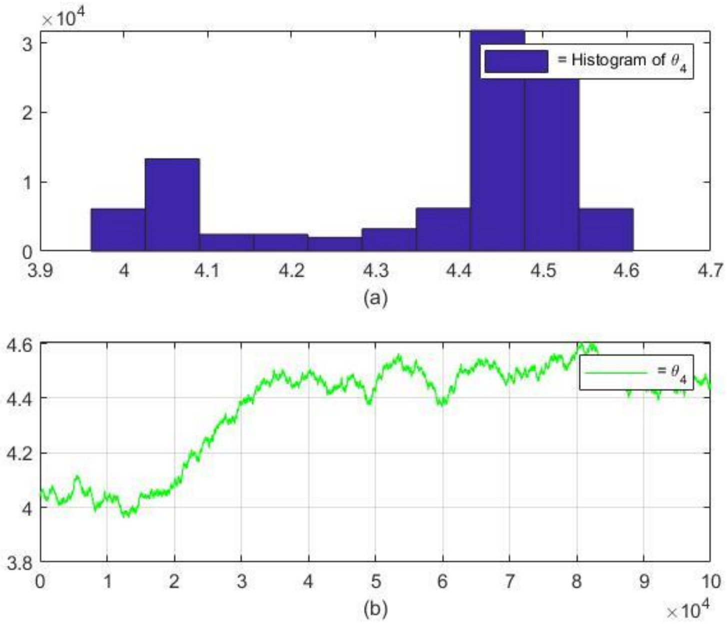


Figure 4. 8: θ_4 (Wheat Price vs Yield)

Current Price at $X_{1|t}$ ($t = current\ time$) concerning the Current yield at x_t ($t = current\ time\ step$) and the convergence rate is 4.4.

With Respect to Cultivation Area Forecasted Results:

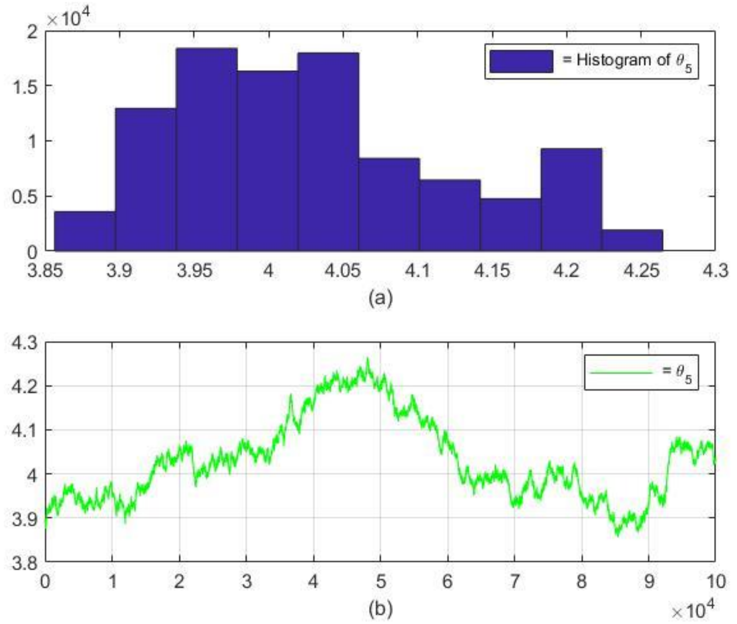


Figure 4. 9: θ_5 (Wheat Area vs Price)

Current Cultivation Area $X_{2|t}$ ($t = current\ time$) concerning the Current Price at x_t ($t = current\ time\ step$) and the convergence rate is 3.95.

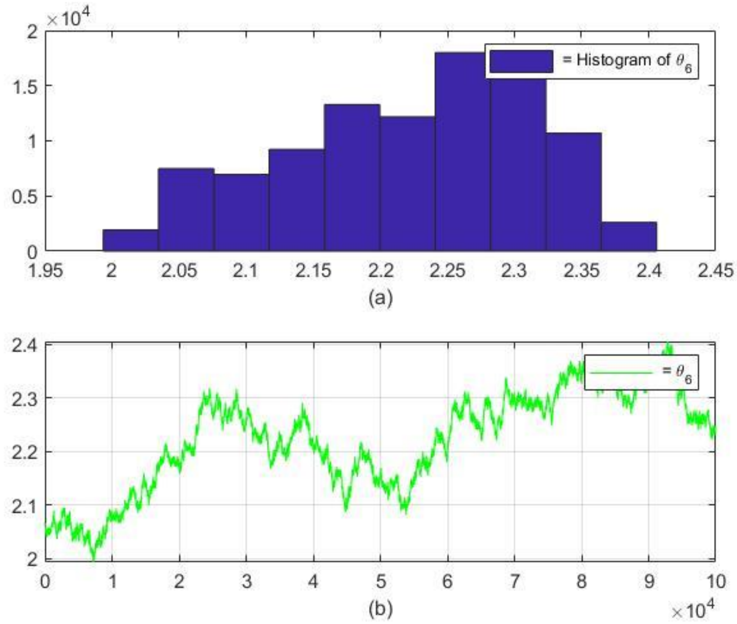


Figure 4. 10: θ_6 (Wheat Area vs Previous Area)

Current Cultivation Area $X_{2|t}$ ($t = current\ time$) concerning the Previous Cultivation Area at x_{t-1} ($t - 1 = previous\ time\ step$) and the convergence rate is 2.25.

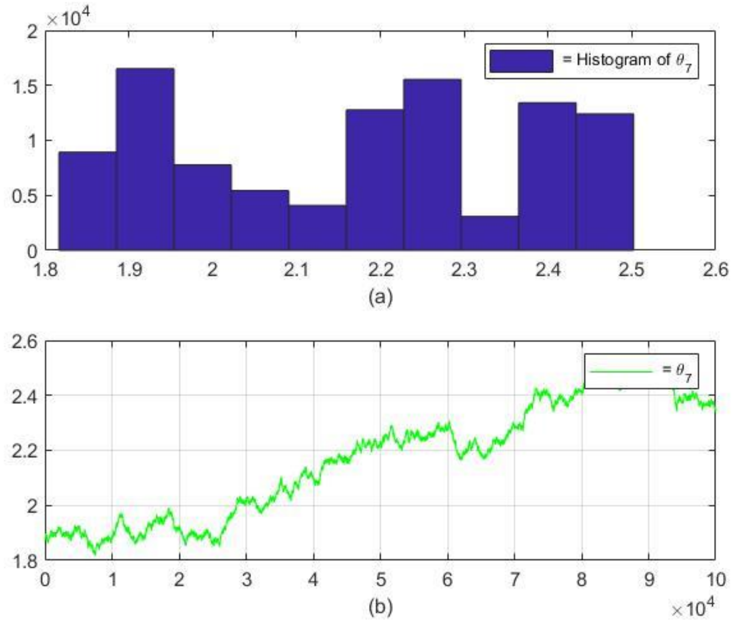


Figure 4. 11: θ_7 (Wheat Area vs production)

Current Cultivation Area $X_{2|t}$ $(t = current\ time)$ concerning Current Production at x_t $(t = current\ time\ step)$ and the convergence rate is 2.2.

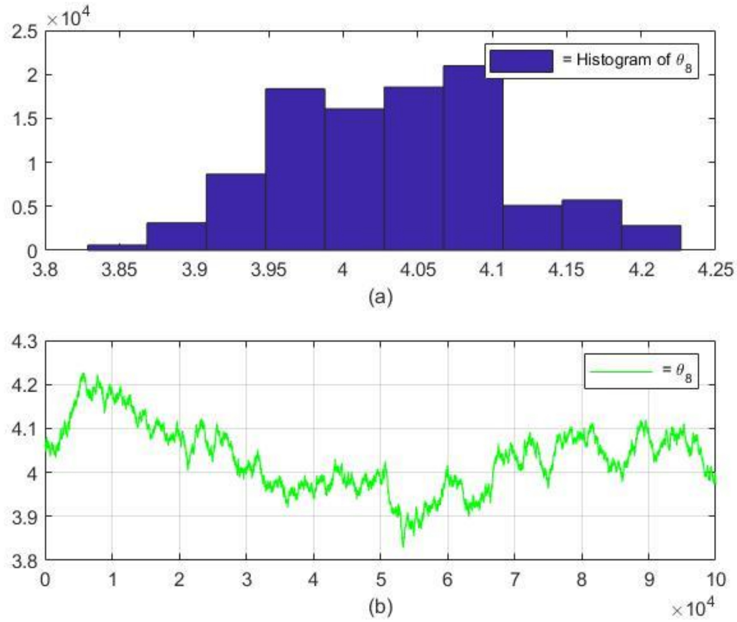


Figure 4. 12: θ_8 (Wheat Area vs Yield)

Current Cultivation Area $X_{2|t}$ $(t = current\ time)$ concerning the Current Yield at x_t $(t = current\ time\ step)$ and the convergence rate is 4.1.

With Respect to Production Forecasted Results:

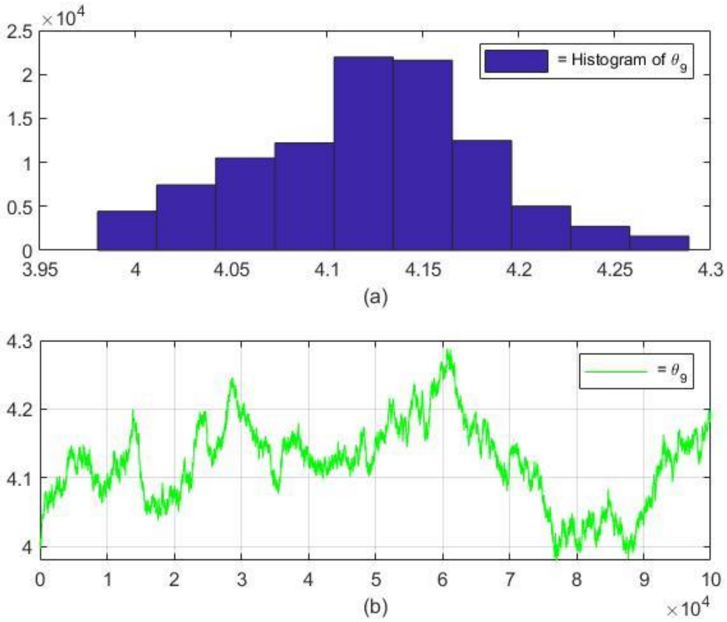


Figure 4. 13: θ_9 (Wheat Production vs Price)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning the Current Price at x_t $(t = current\ time\ step)$ and the convergence rate is 4.1.

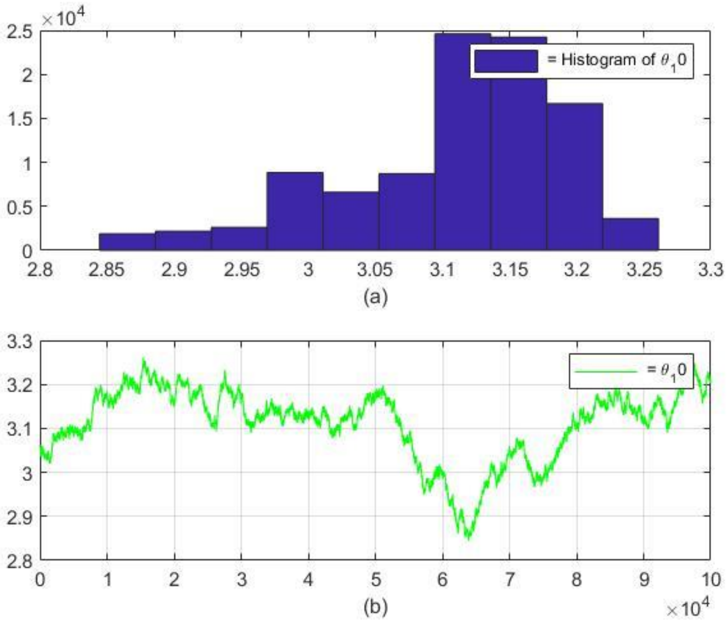


Figure 4. 14: θ_{10} (Wheat Production vs Area)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning the Current Cultivation Area at x_t $(t = current\ time\ step)$ and the convergence rate is 3.1.

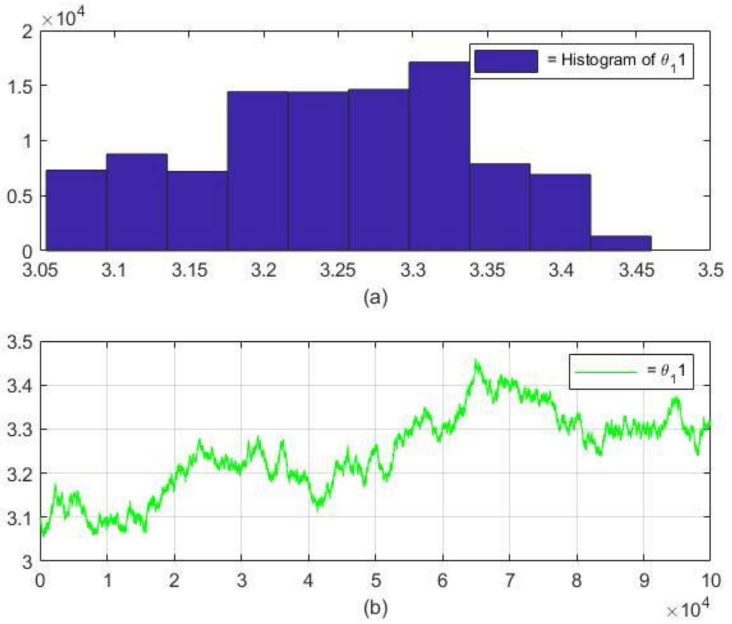


Figure 4. 15: θ_{11} (Wheat Production vs Previous Production)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning the Previous Production at x_{t-1} $(t - 1 = previous\ time\ step)$ and the convergence rate is 3.3.

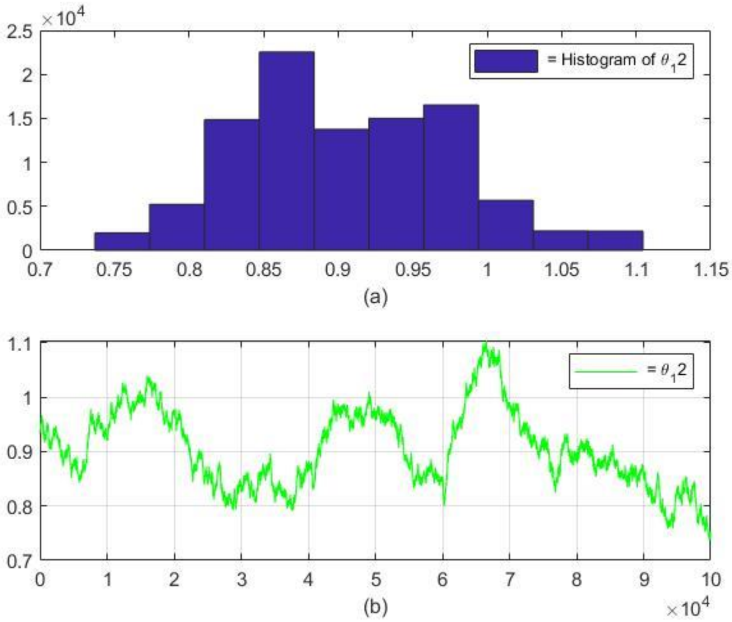


Figure 4. 16: θ_{12} (Wheat Production vs Yield)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning the Current Yield at x_t $(t = current\ time\ step)$ and the convergence rate is 0.85.

With Respect to Yield Forecasted Results:

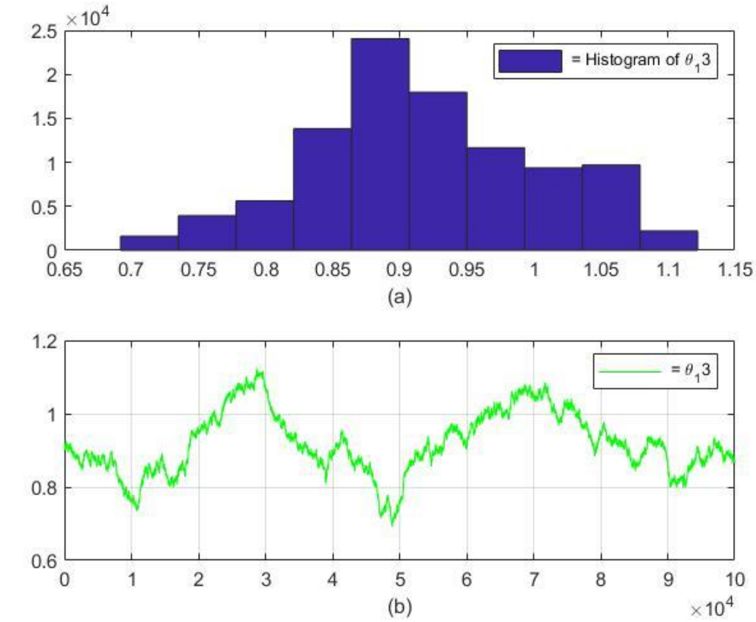


Figure 4. 17: θ_{13} (Wheat Yield vs Price)

Current Yield$X_{4|t}$ $(t = current\ time)$ concerning the Current Price at x_t $(t = current\ time\ step)$ and the convergence rate is 0.85.

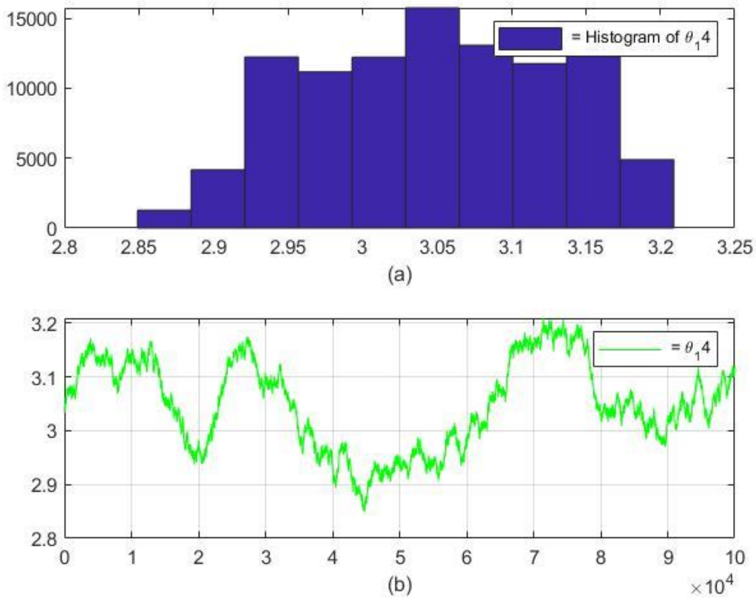


Figure 4. 18: θ_{14} (Wheat Yield vs Area)

Current Yield $X_{4|t}$ $(t = current\ time)$ concerning the Current Cultivation Area at x_t $(t = current\ time\ step)$ and the convergence rate is 3.05.

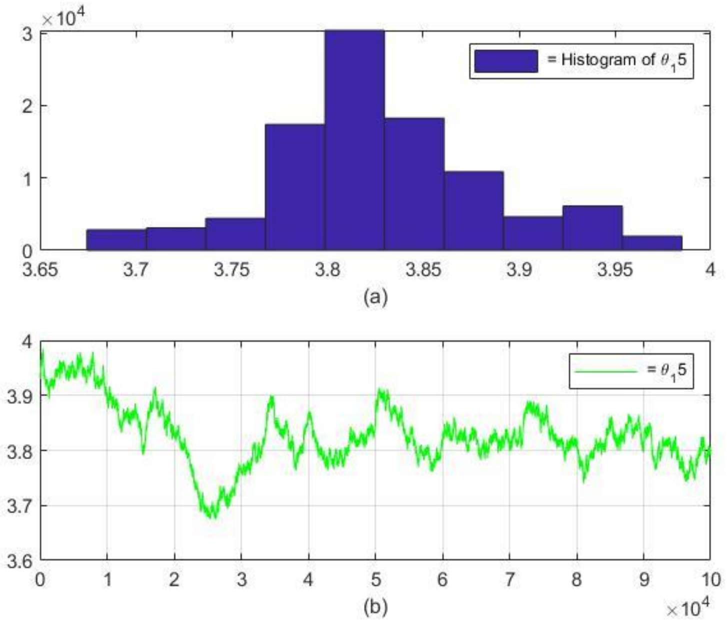


Figure 4. 19: θ_{15} (Wheat Yield vs Production)

Current Yield $X_{4|t}$ $(t = current\ time)$ concerning Current Production at x_t $(t = current\ time\ step)$ and the convergence rate is 3.8.

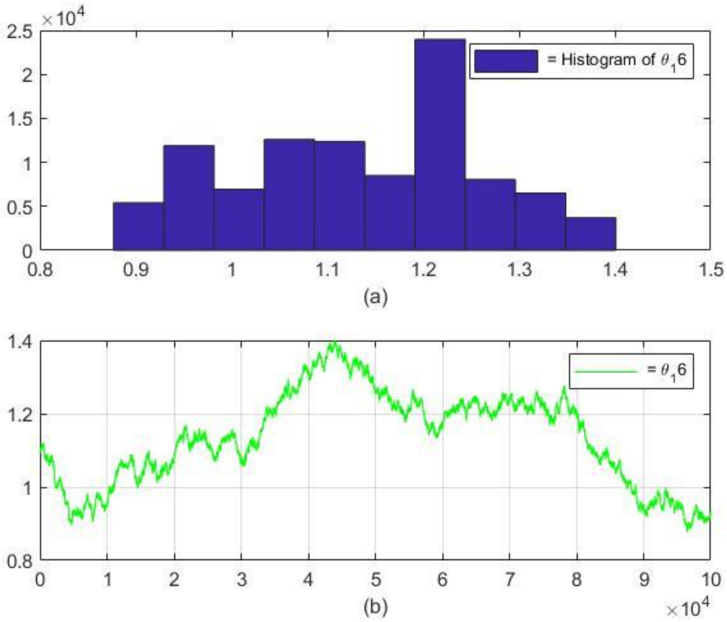


Figure 4. 20: θ_{16} (Wheat Yield vs Previous Yield)

Current Yield $X_{4|t}$ $(t = current\ time)$ concerning the Previous Yield at x_{t-1} $(t - 1 = previous\ time\ step)$ and the convergence rate is 1.2.

Discussion for Forecasted Wheat for Pakistan:

According to the results of the forecast, the total wheat area in Pakistan is forecasted to increase by 2.5 percent. The forecast for wheat yield indicates an increase of 8 percent. The considerable increase in area and yield leads to an increase of 8.7 percent in wheat Production. The increase (forecasted) in wheat production is likely to have benefited from the availability of irrigation water. The government release of wheat stocks significantly affects the wheat prices in the market.

Factors Influencing Area, Yield, Production, and Prices of Wheat:

Several factors have affected the area and yield of Pakistan's wheat crop. The area devoted to wheat depends on support prices, the availability of irrigation water, competition from other crops, and prices of inputs, such as fertilizer and energy. For yield, the major influencing factors are the availability of skilled labor, seed quality, weather patterns, soil fertility, fertilizer use, and degree of mechanization.

4.3.2 Potato

Potato is an important crop for both farmers and consumers in Pakistan. Pakistan with 4.87 million tons contributes a small share to global potato production.

With Respect to Price forecasted Results:

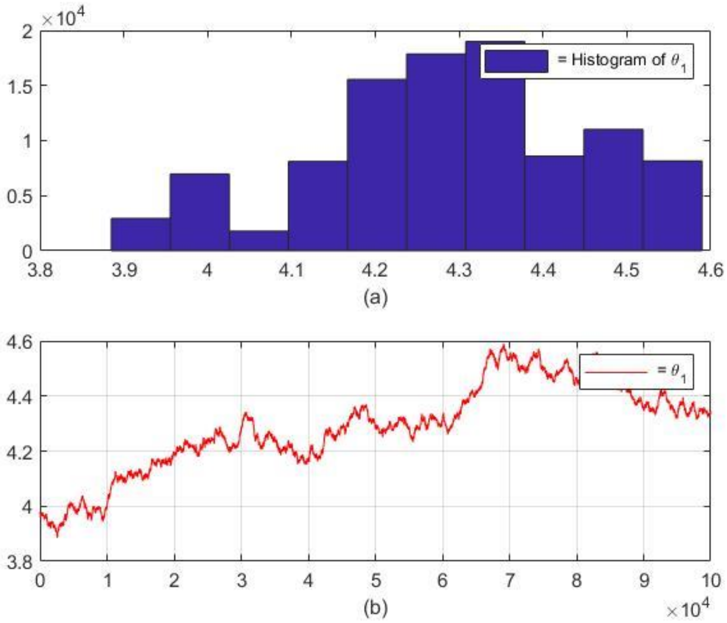


Figure 4. 21: θ_1 (Potato Price vs Previous Price)

Current Price $X_{1|t}$ $(t = current\ time)$ concerning the Previous Price at x_{t-1} $(t - 1 = previous\ time\ step)$ and the convergence rate is 4.3.

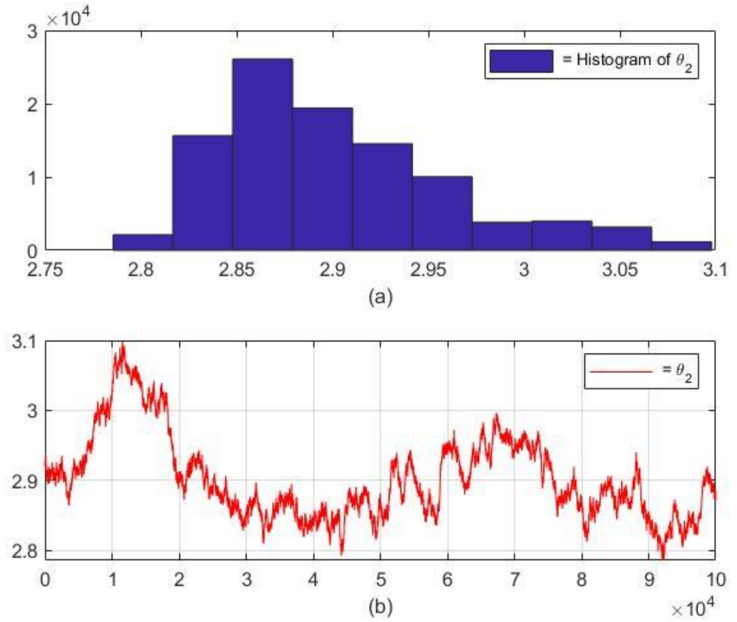


Figure 4. 22: θ_2(Potato Price vs Area)

Current Price $X_{1|t}$ $(t = current\ time)$ concerning the Current Cultivation Area at x_t $(t = current\ time\ step)$ and the convergence rate is 2.85.

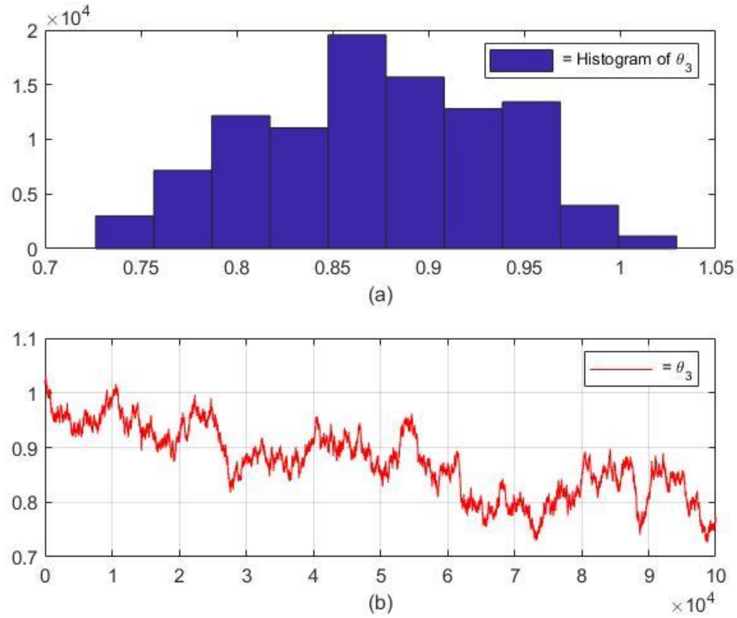


Figure 4. 23: θ_3(Potato Price vs Production)

Current Price $X_{1|t}$ $(t = current\ time)$ concerning Current Production at x_t $(t = current\ time\ step)$ and the convergence rate is 0.85.

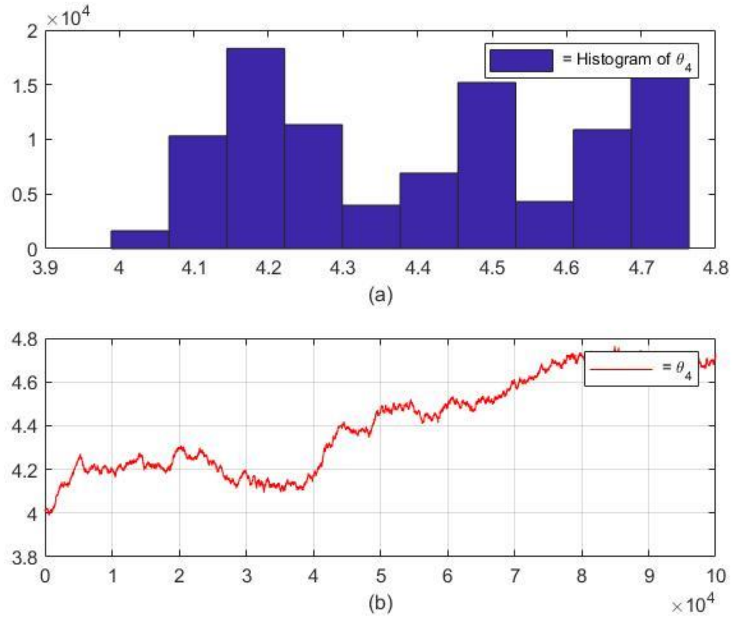


Figure 4. 24: θ_4(Potato Price vs Yield)

Current Price $X_{1|t}$ $(t = current\ time)$ concerning the Current Yield at x_t $(t = current\ time\ step)$ and the convergence rate is 4.2.

With Respect to Cultivation Area Forecasted Results:

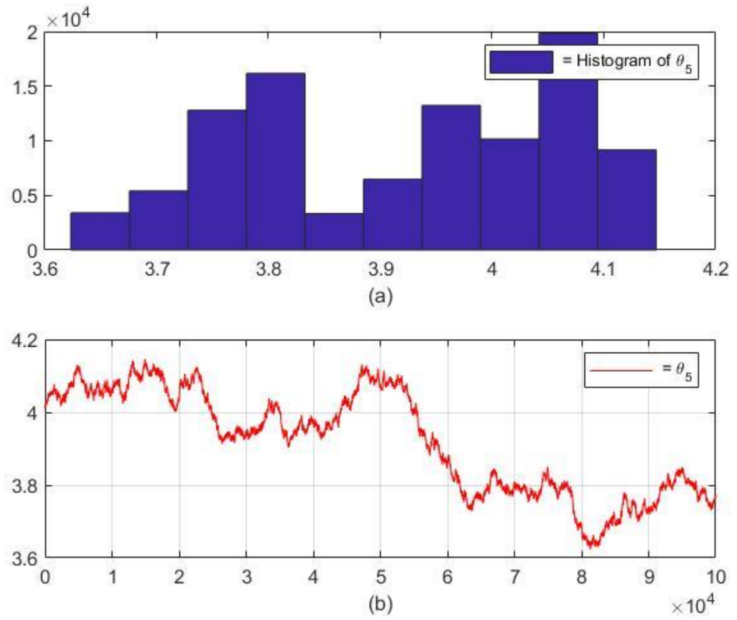


Figure 4. 25: θ_5(Potato Area vs Price)

Current Cultivation Area $X_{2|t}$ $(t = current\ time)$ concerning the Current Price at x_t $(t = current\ time\ step)$ and the convergence rate is 4.05.

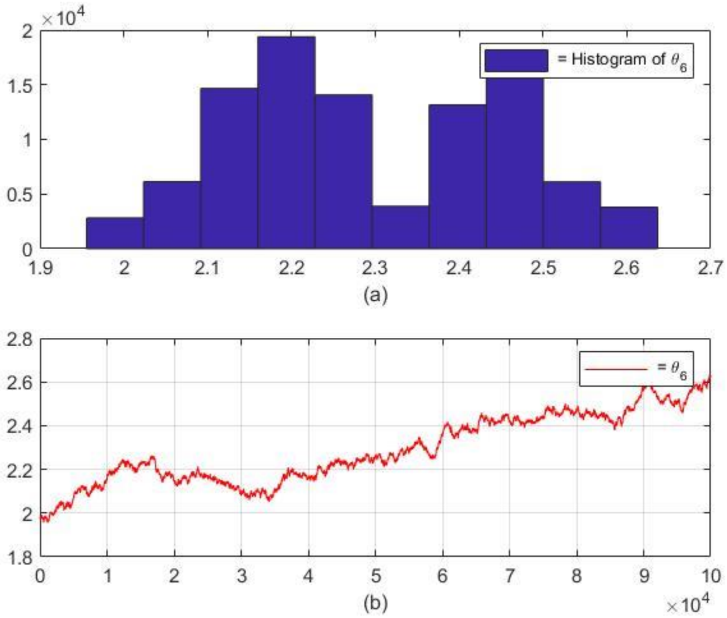


Figure 4. 26: θ_6(Potato Area vs Previous Area)

Current Cultivation Area $X_{2|t}$ $(t = current\ time)$ concerning the Previous Cultivation Area at x_{t-1} $(t - 1 = Previous\ time\ step)$ and the convergence rate is 2.2.

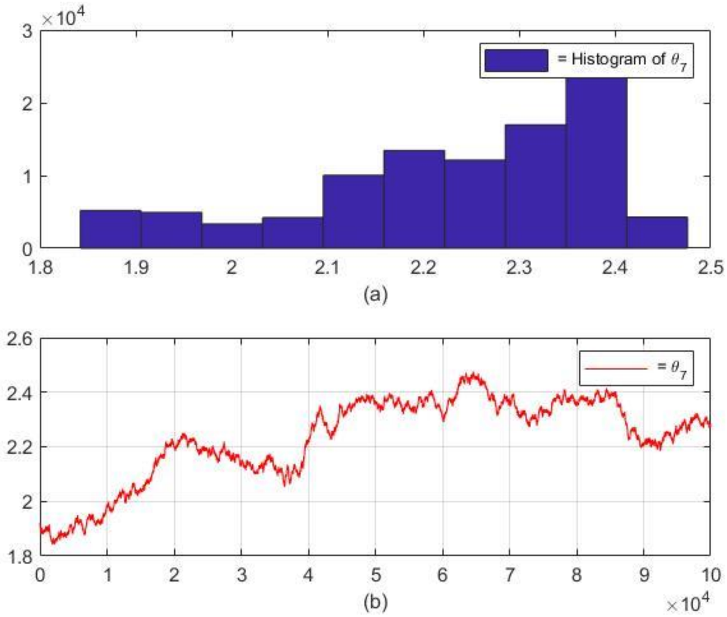


Figure 4. 27: θ_7(Potato Area vs Production)

Current Cultivation Area $X_{2|t}$ $(t = current\ time)$ concerning Current Production at x_t $(t = current\ time\ step)$ and the convergence rate is 2.4.

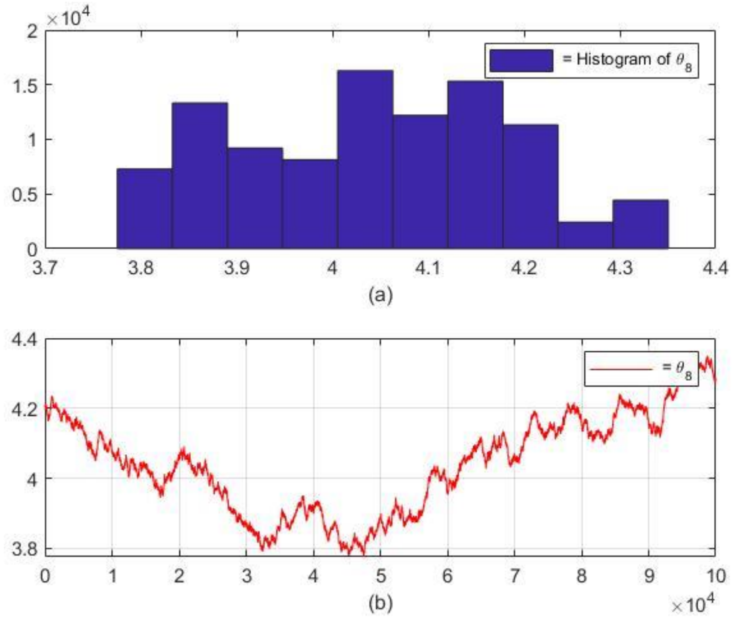


Figure 4. 28: θ_8(Potato Area vs Yield)

Current Cultivation Area $X_{2|t}$ $(t = current\ time)$ concerning the Current Yield at x_t $(t = current\ time\ step)$ and the convergence rate is 4.

With Respect to Production Forecasted Results:

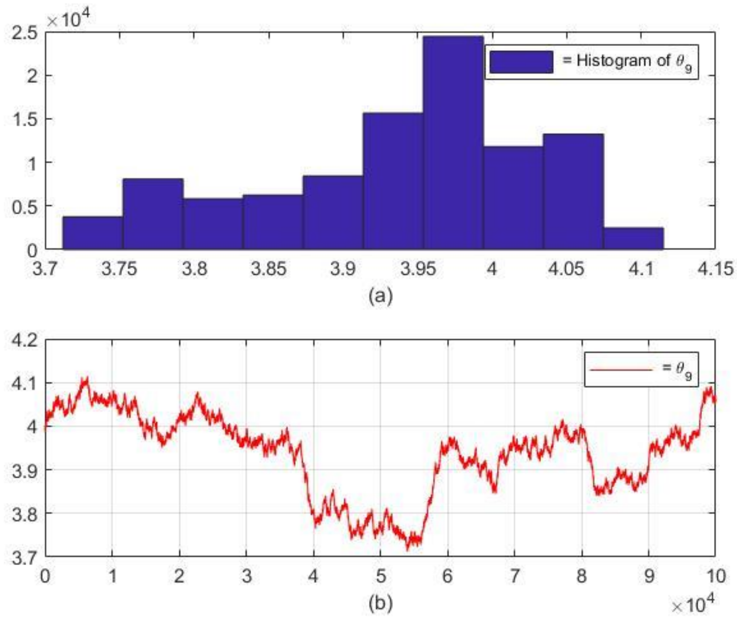


Figure 4. 29: θ_9 (Potato Production vs Prices)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning the Current Price at x_t $(t = current\ time\ step)$ and the convergence rate is 3.95.

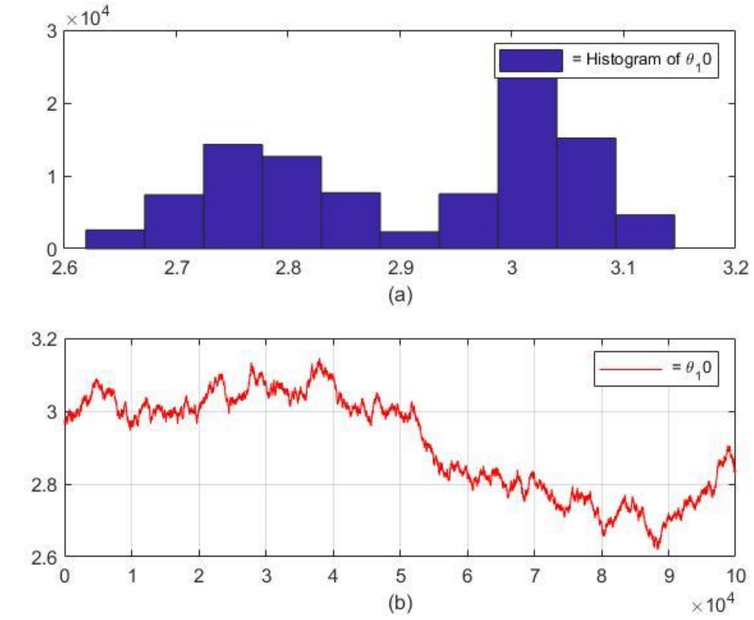


Figure 4. 30: θ_{10}(Potato Production vs Area)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning the Current Cultivation Area at x_t $(t = current\ time\ step)$ and the convergence rate is 3.

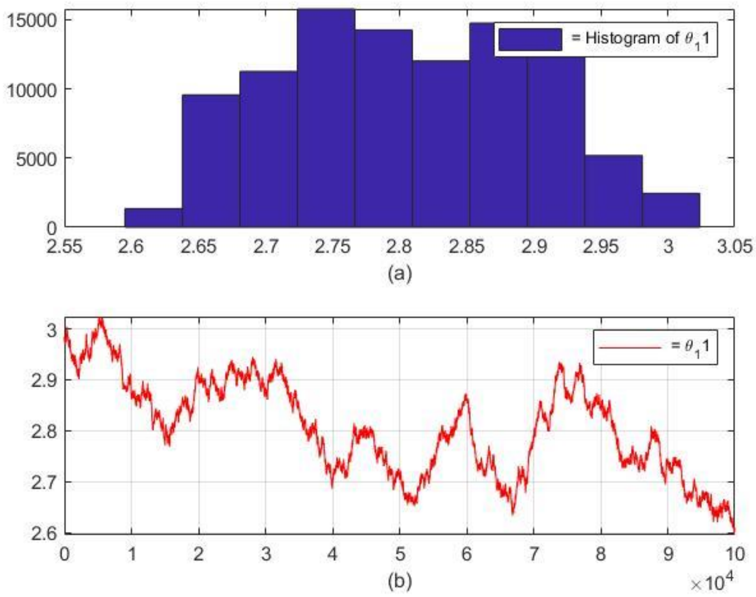


Figure 4. 31: θ_{11}(Potato Production vs Previous Production)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning Previous Production at x_{t-1} $(t - 1 = Previous\ time\ step)$ and the convergence rate is 2.75.

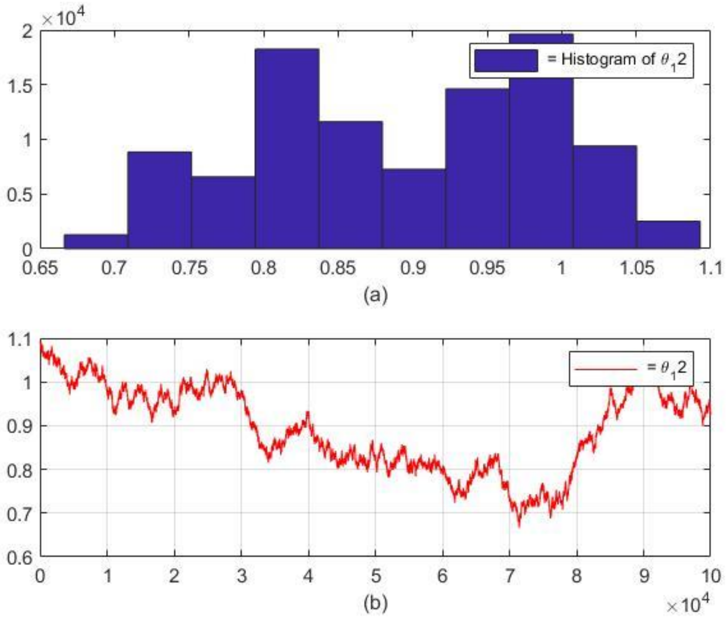


Figure 4. 32: θ_{12}(Potato Production vs Yield)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning the Current Yield at x_t $(t = current\ time\ step)$ and the convergence rate is 1.

With Respect to Yield, Forecasted Results:

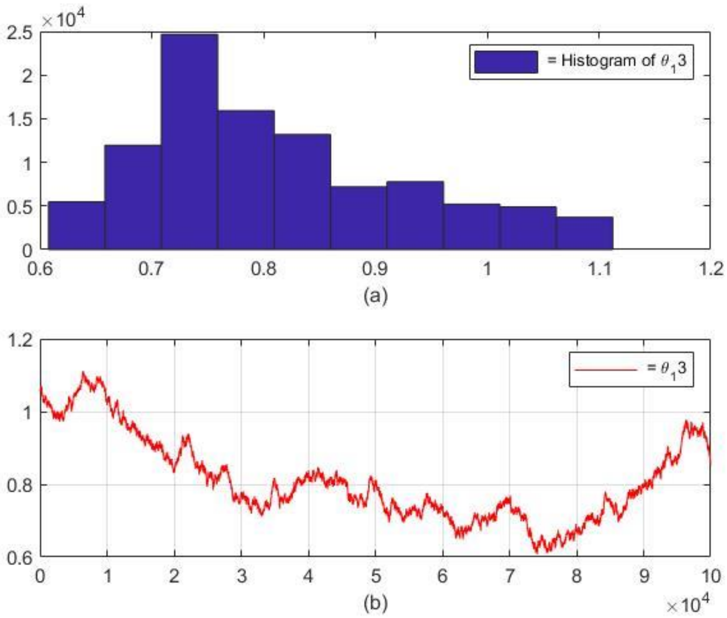


Figure 4. 33: θ_{13} (Potato Yield vs Prices)

Current Yield $X_{4|t}$ $(t = current\ time)$ concerning the Current Price at x_t $(t = current\ time\ step)$ and the convergence rate is 0.7.

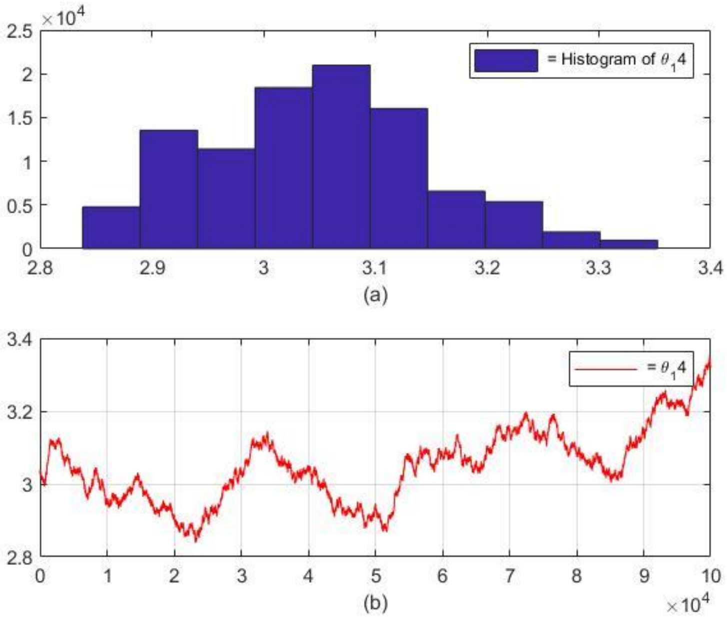


Figure 4. 34: θ_{14}(Potato Yield vs Area)

Current Yield $X_{4|t}$ $(t = current\ time)$ concerning the Current Cultivation Area at x_t $(t = current\ time\ step)$ and the convergence rate is 3.05.

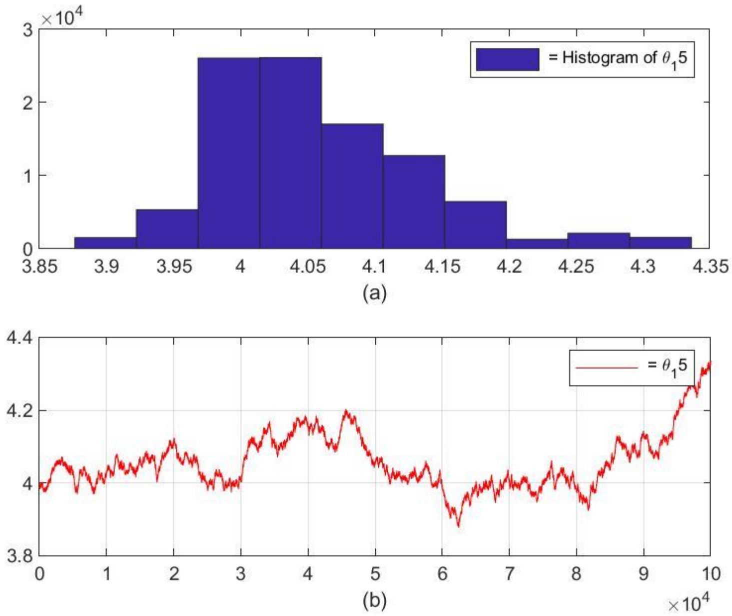


Figure 4. 35: θ_{15}(Potato Yield vs Production)

Current Yield $X_{4|t}$ $(t = current\ time)$ concerning Current Production at x_t $(t = current\ time\ step)$ and the convergence rate is 4.

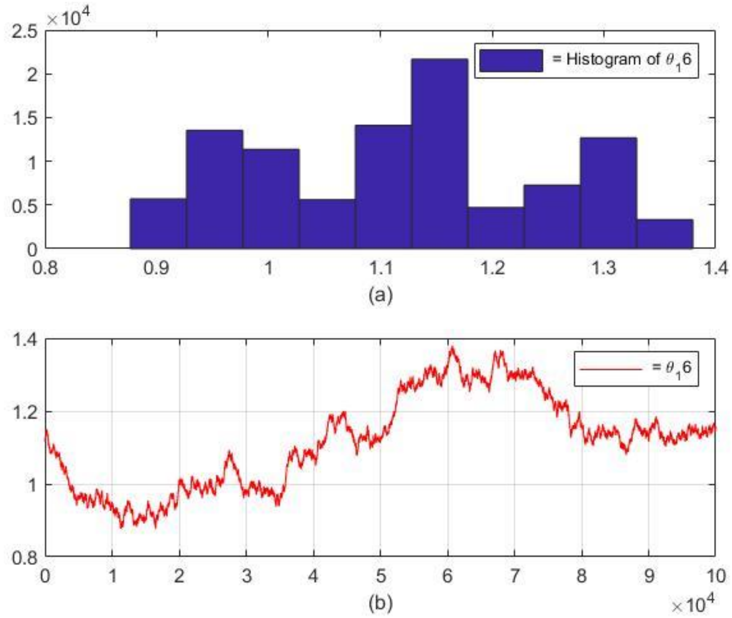


Figure 4. 36: θ_{16}(Potato Yield vs Previous Yield)

Current Yield $X_{4|t}$ $(t = current\ time)$ concerning the Previous Yield at x_{t-1} $(t - 1 = Previous\ time\ step)$ and the convergence rate is 1.15.

Factors Influencing Area, Yield, and Production of Potato:

The area, yield, and production of potatoes are highly responsive to several factors, which include prices, climatic factors, diseases, and export policies. The area devoted to potatoes by farmers highly depends on last year's prices. The yield of potatoes depends on the availability of inputs and climatic conditions. Due to the increase in energy prices, transportation cost has increased significantly which prevents farmers to sell their produce in distant markets. Similarly, the increase in prices of potato seeds also prevents farmers to devote a large area to potato cultivation. The climatic factors also affect potato yield.

Forecast Analysis for Potato in Pakistan:

According to the quantitative forecast, the potato area is forecasted to decrease by 1 percent. Potato yield is forecasted to remain the same as in 2018-19, so total production is expected to decrease by 1 percent due to the decrease in the cultivated area. Potato cultivation mainly depends on last year's prices.

4.3.3 Gram

Gram, also known as chickpea, is an important pulse crop both in terms of production and consumption in Pakistan.

With Respect to Prices, forecasted Results:

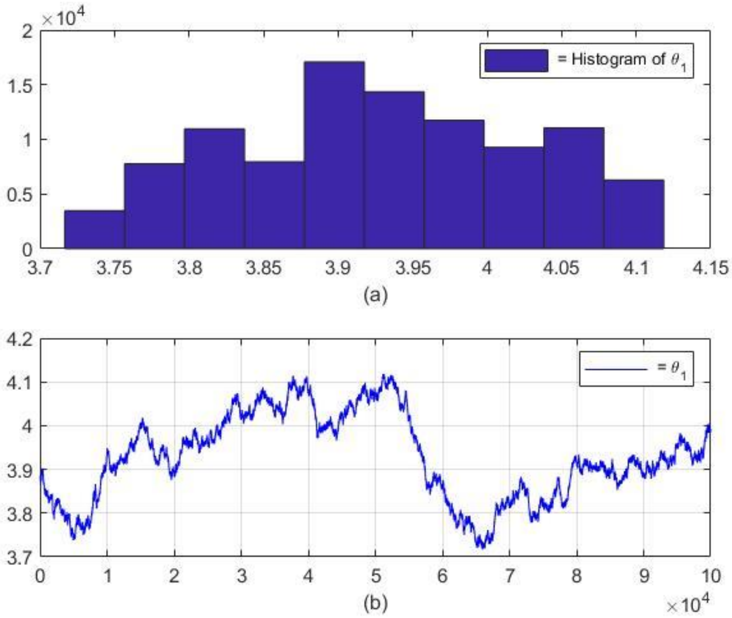


Figure 4. 37: θ_1(Gram Prices vs Previous Prices)

Current Price $X_{1|t}$ ($t = current\ time$) concerning the Previous Price at x_{t-1} ($t - 1 = Previous\ time\ step$) and the convergence rate is 3.9.

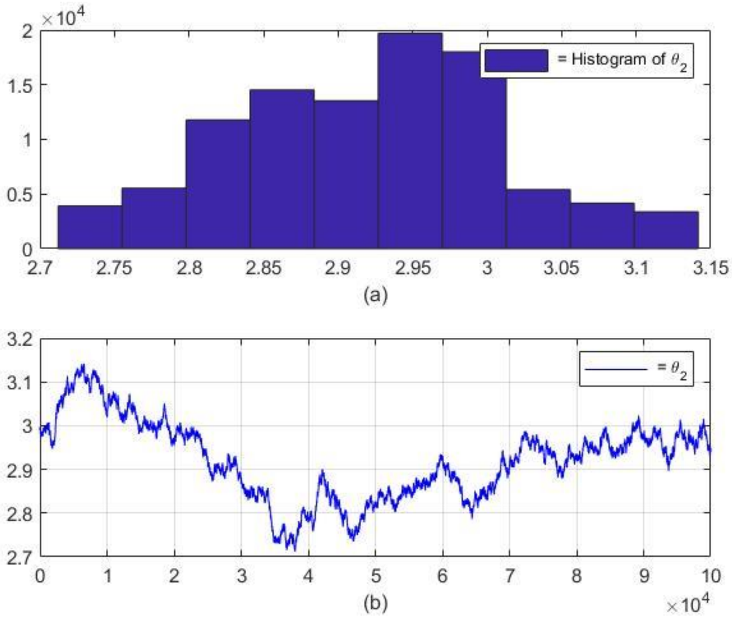


Figure 4. 38: θ_2(Gram Prices vs Area)

Current Price $X_{1|t}$ ($t = current\ time$) concerning the Current Cultivation Area at x_t ($t = current\ time\ step$) and the convergence rate is 2.95.

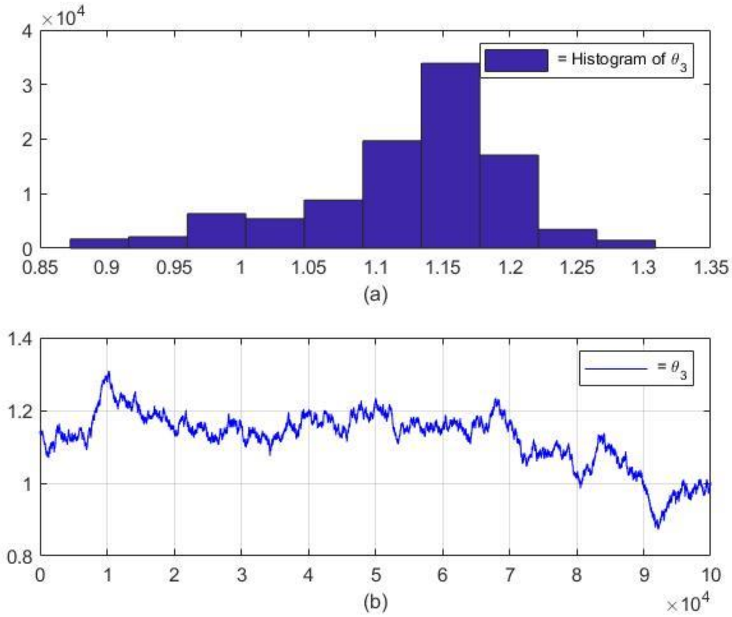


Figure 4. 39: θ_3(Gram Prices vs Production)

Current Price $X_{1|t}$ ($t = current\ time$) concerning Current Production at x_t ($t = current\ time\ step$) and the convergence rate is 1.15.

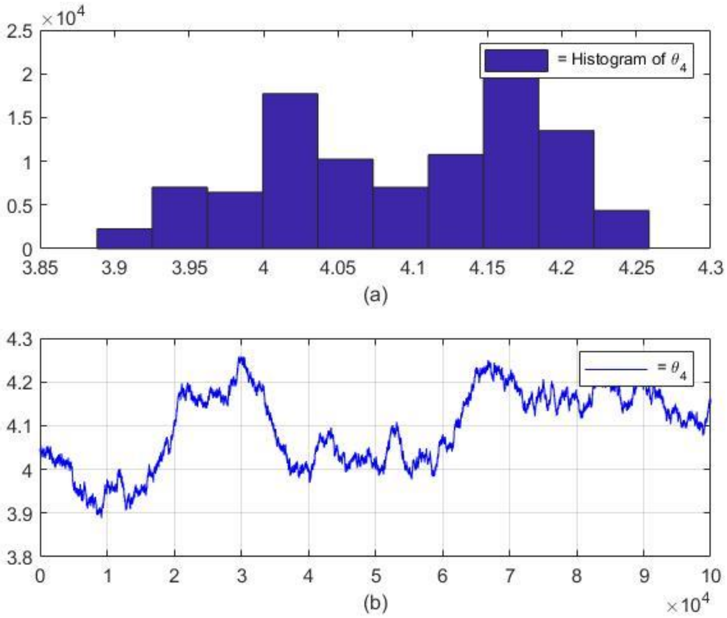


Figure 4. 40: θ_4(Gram Prices vs Yield)

Current Price $X_{1|t}$ $(t = current\ time)$ concerning the Current Yield at x_t $(t = current\ time\ step)$ and the convergence rate is 4.15.

With Respect to Cultivation Area, Forecasted Results:

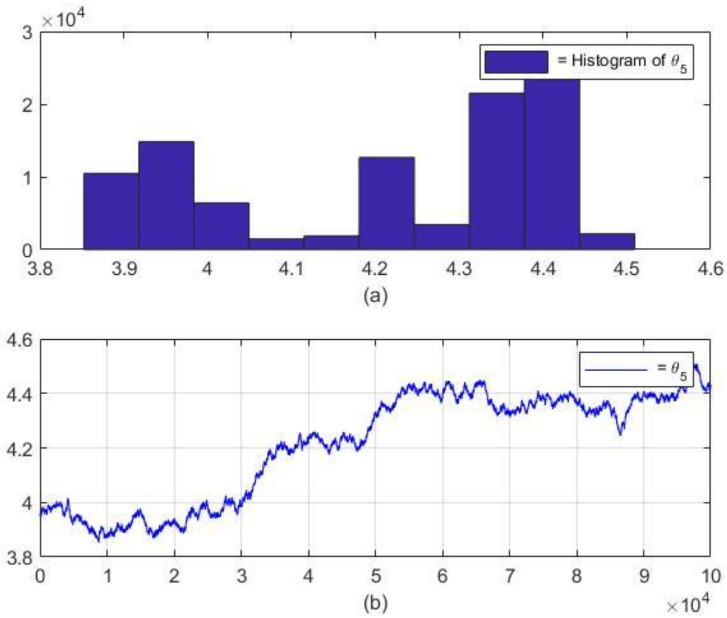


Figure 4. 41: θ_5(Gram Area vs Prices)

Current Cultivation Area $X_{2|t}$ $(t = current\ time)$ concerning the Current Price at x_t $(t = current\ time\ step)$ and the convergence rate is 4.4.

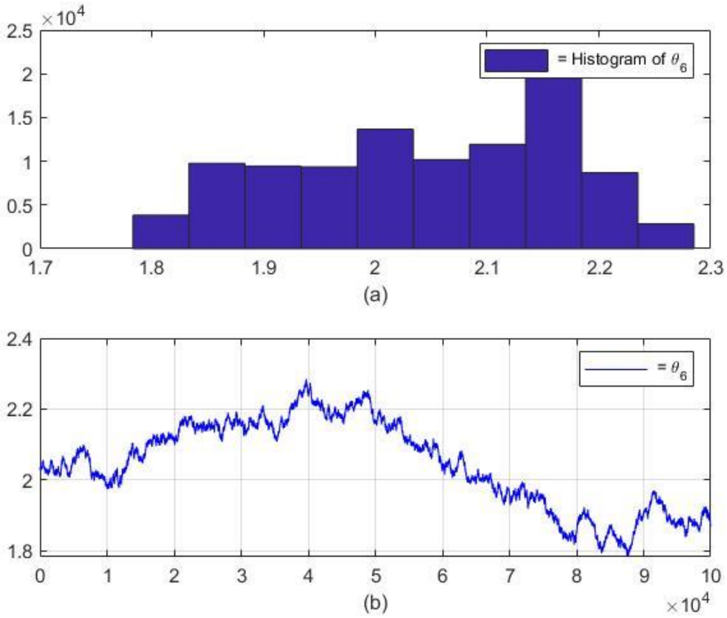


Figure 4. 42: θ_6(Gram Area vs Previous Area)

Current Cultivation Area $X_{2|t}$ $(t = current\ time)$ concerning the Previous Cultivation Area at x_{t-1} $(t-1 = Previous\ time\ step)$ and the convergence rate is 2.15.

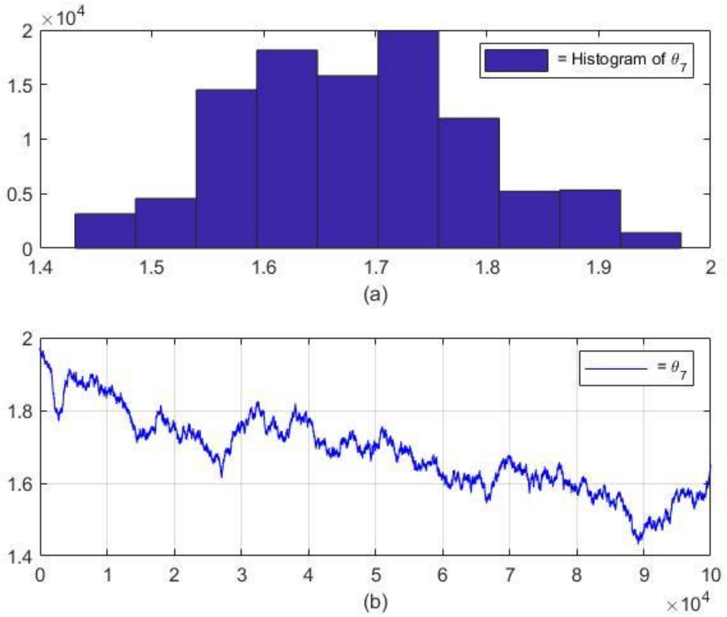


Figure 4. 43: θ_7(Gram Area vs Production)

Current Cultivation Area $X_{2|t}$ $(t = current\ time)$ concerning Current Production at x_t $(t = current\ time\ step)$ and the convergence rate is 1.7.

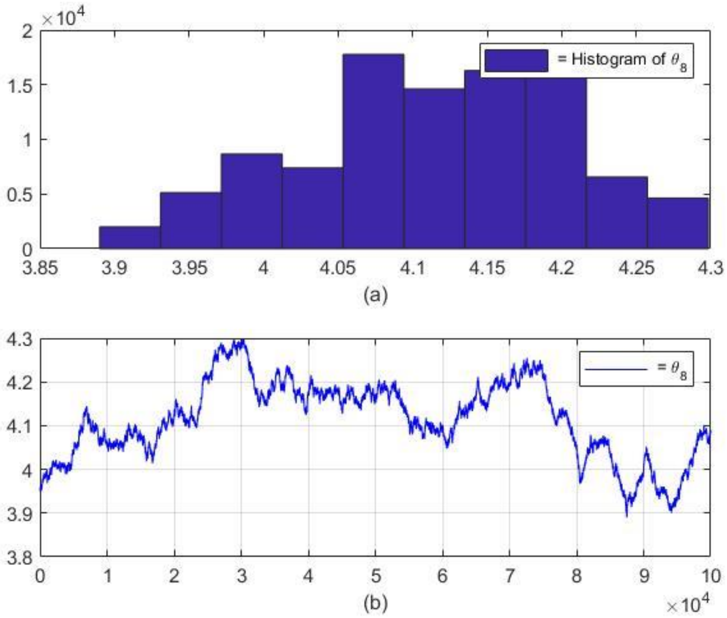


Figure 4. 44: θ_8 (Gram Area vs Yield)

Current Cultivation Area $X_{2|t}$ $(t = current\ time)$ concerning the Current Yield at x_t $(t = current\ time\ step)$ and the convergence rate is 4.05.

With Respect to Production Forecasted Results:

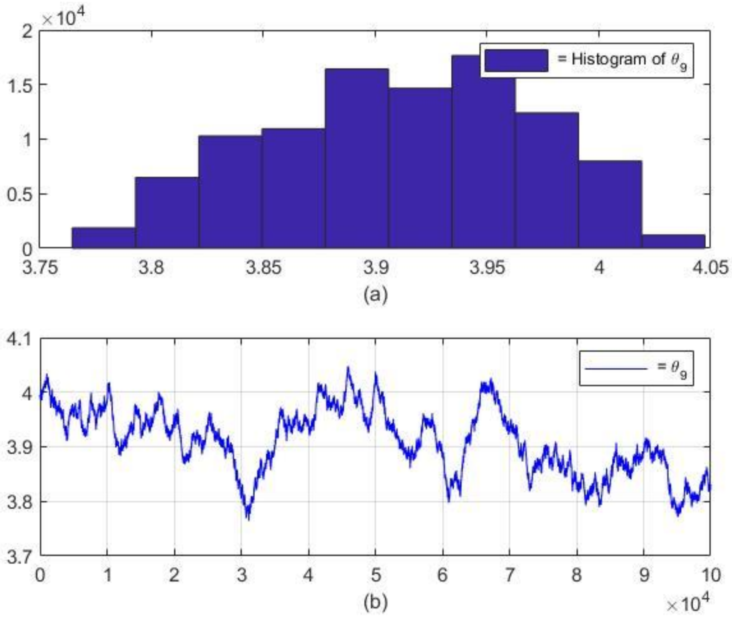


Figure 4. 45: θ_9 (Gram Production vs Prices)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning the Current Price at x_t $(t = current\ time\ step)$ and the convergence rate is 3.95.

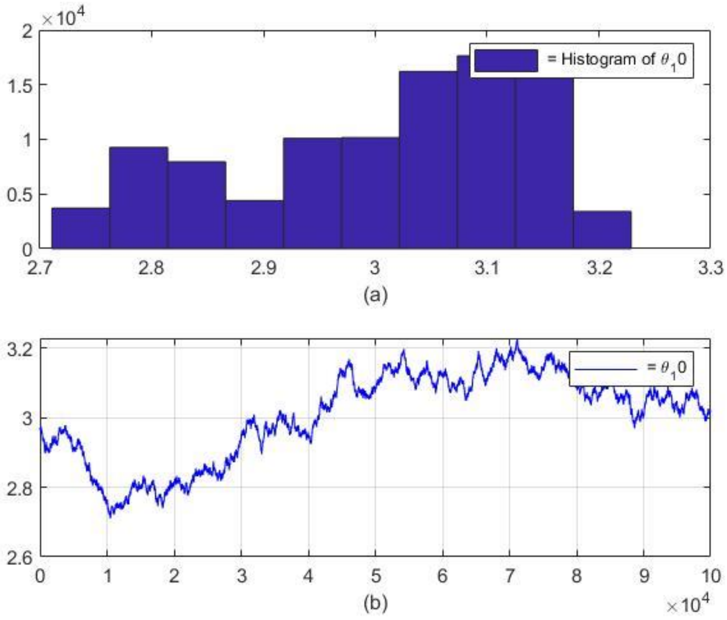


Figure 4. 46: θ_{10}(Gram Production vs Area)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning the Current Cultivation Area at x_t $(t = current\ time\ step)$ and the convergence rate is 3.1.

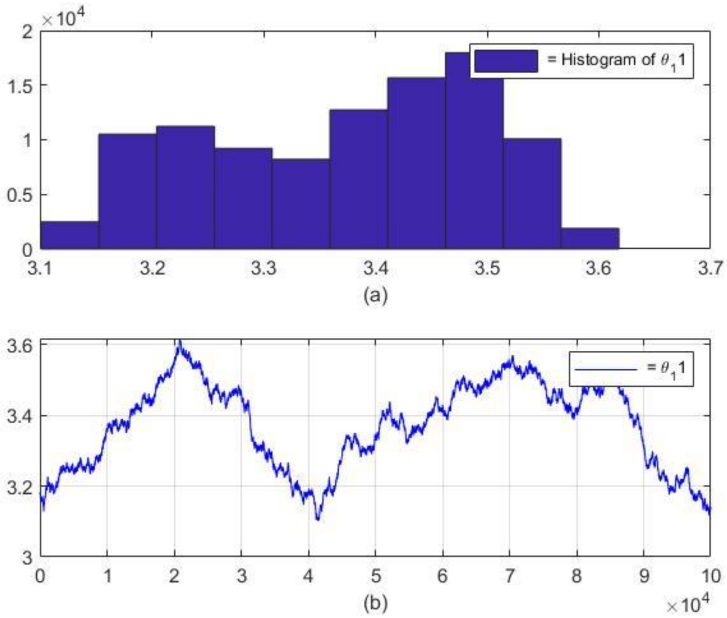


Figure 4. 47: θ_{11} (Gram Production vs Previous Production)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning Previous Production at x_{t-1} $(t - 1 = Previous\ time\ step)$ and the convergence rate is 3.5.

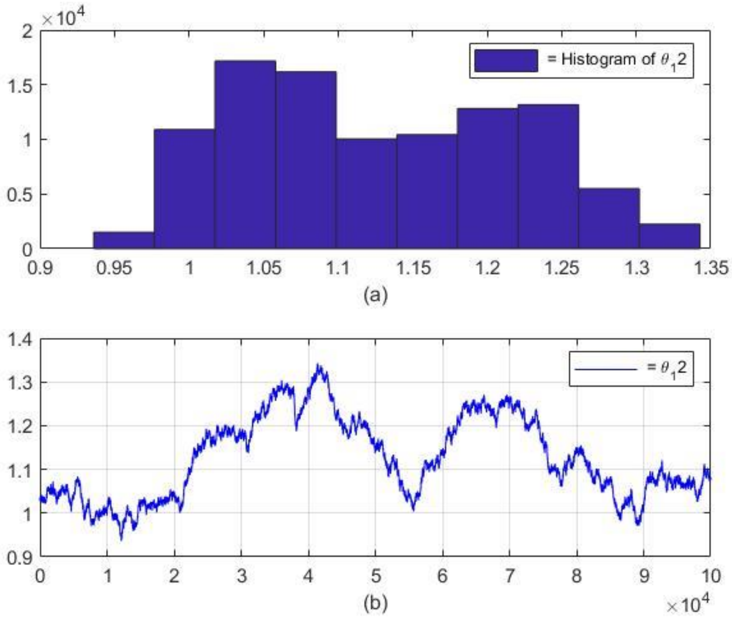


Figure 4. 48: θ_{12} (Gram Production vs Yield)

Current Production $X_{3|t}$ $(t = current\ time)$ concerning the Current Yield at x_t $(t = current\ time\ step)$ and the convergence rate is 1.03.

With Respect to Yield, Forecasted Results:

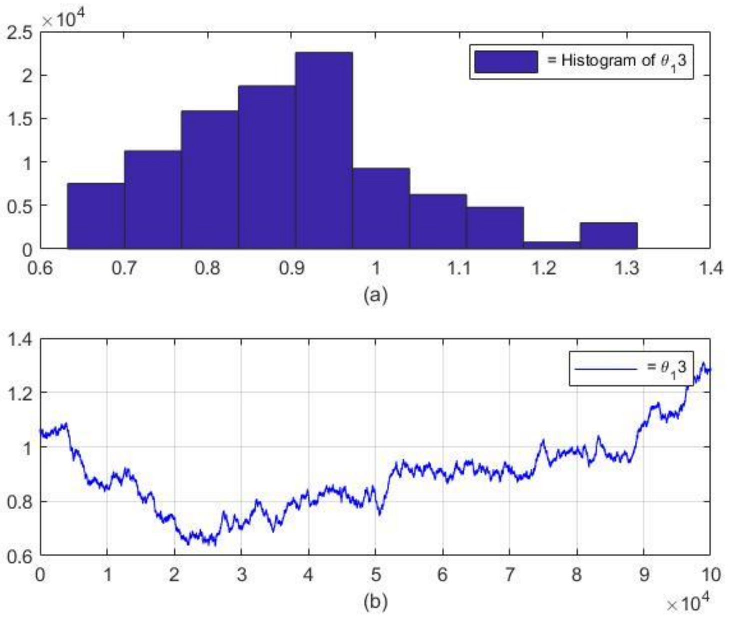


Figure 4. 49: θ_{13} (Gram Yield vs Prices)

Current Yield $X_{4|t}$ $(t = current\ time)$ concerning the Current Price at x_t $(t = current\ time\ step)$ and the convergence rate is 0.9.

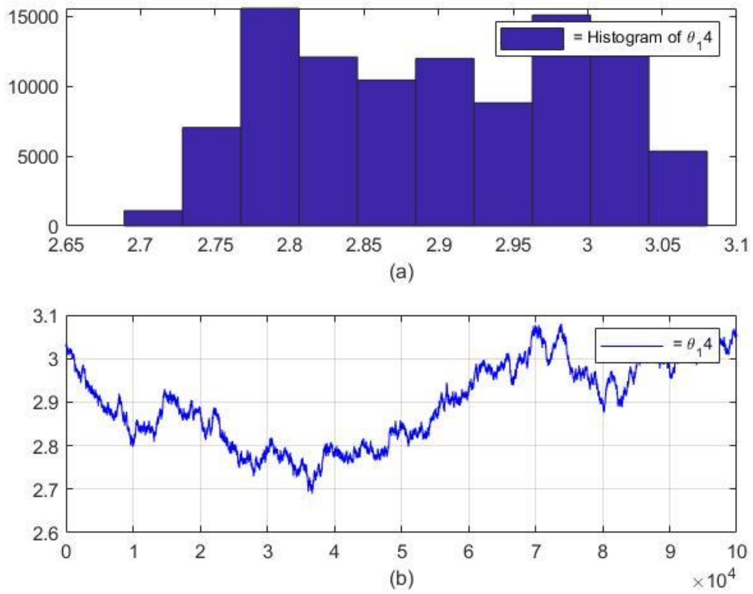


Figure 4. 50: θ_{14}(Gram Yield vs Area)

Current Yield $X_{4|t}$ $(t = current\ time)$ concerning the Current Cultivation Area at x_t $(t = current\ time\ step)$ and the convergence rate is 2.76.

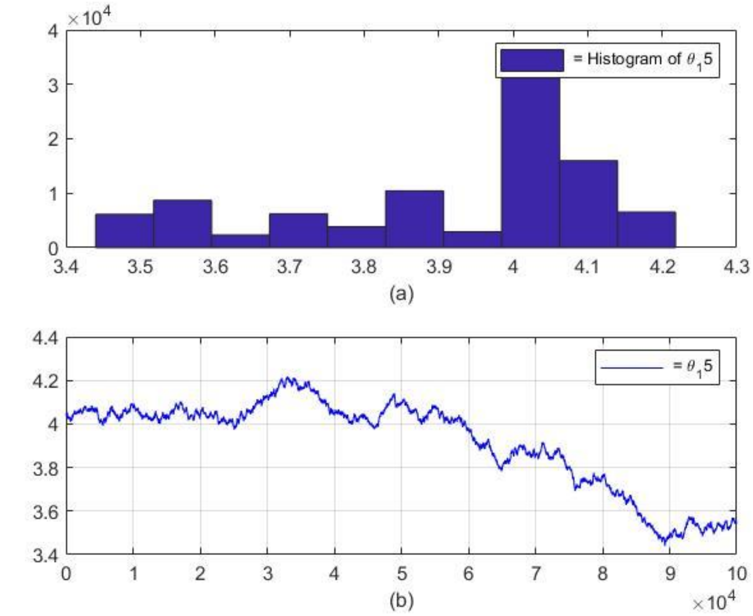


Figure 4. 51: θ_{15} (Gram Yield vs Production)

Current Yield $X_{4|t}$ $(t = current\ time)$ concerning Current Production at x_t $(t = current\ time\ step)$ and the convergence rate is 4.

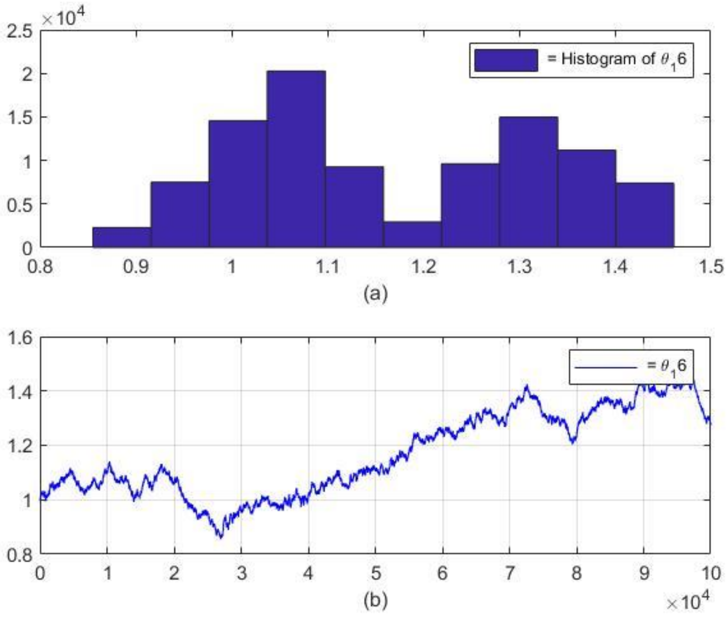


Figure 4. 52: θ_{16} (Gram Yield vs Previous Yield)

Figure 4.52

Current Yield $X_{4|t}$ ($t = current\ time$) concerning the Previous Yield at x_{t-1} ($t - 1 = Previous\ time\ step$) and the convergence rate is 1.05.

Factors Affecting Yield, Area, and Production of Gram

The production of Gram reveals severe fluctuations that can be attributed to several factors. As gram is mostly cultivated in the rain-fed regions of Pakistan, climatic patterns greatly affect production. Harsh weather patterns like drought and erratic rainfall are major obstacles to increasing gram crop production. Heavy rainfall at the time of flowering and harvesting may also damage the crop. Similarly, persistent drought events could destroy the product, and could even result in farmers having no seed to sow for the next year.

Forecast Analysis of Gram for Pakistan:

As per the econometric analysis, the forecast for the gram crop is a 2.5 percent increase in the cultivated area for the next year's Rabi season. However, the forecast predicts a marginal decrease in gram yield. Due to the forecasted increase in cultivated area, the production of the gram is expected to increase by 2 percent for the next year. The price for a gram is estimated by first forecasting the real price. In this case, nominal prices of a gram are forecasted to increase by 14 percent for next year.

Table 4.2: Forecasted result and comperision

Variables	Values in 2018-2019	ARIMA Forecast Values for 2019-2020	Bayesian State Space Forecast Values for 2019-2020	Actual Vales in 2019-2020	Bayesian State Space Forecast Values for 2020-2021	Bayesian State Space Forecast Values for 2021-2022
Area	945.90	965.63	940	950	873	866.6
Yield	465.93	465.18	470	465	276	368
Production	440.72	449.19	545	547	261	355
Prices	9028	10294	10500	10670	17000	21500

In Table 4.2 it is clear that the forecasted results for the production of Gram is more accurate then of ARIMA forecasted results. The actual values for 2019-2020 is more approximate to the Bayesian State Space forecasted values for 2019-2020. which prove the accuracy of our model. Which means the Bayesian state space model is more accurate then of Autoregression integrated moving average (ARIMA). So Bayesian state space is more suitable then other regression models.

4.4 Summary

This chapter presented the complete details related to our ARIMA-based Bayesian State-Space model i.e. Observation Equating Model and State Estimation Model and their Mathematical Equations. The rest of the chapter includes the findings or results of the conducted forecast of rabbi food crops i.e. Wheat, Potato, and especially Gram. As the result, the best objective has been attained and the undertaking of the RQs accomplished. However, the basic purpose of this chapter was to make our Multivariate Linear (ARIMA-based) Bayesian State-Space model for our data and forecast the results of Crops including Gram. In the Results and Discussion section, the objectives of this study alongside forecasted results are discussed alongside graphs and tables.

CHAPTER 5

CONCLUSION AND LIMITATIONS

5.1 Overview

This chapter accounts for the discussion of the contributions of this study and the conclusion drawn from this thesis. However, this chapter provides the achievement, especially in the context of research questions i.e. RQ1, RQ2, and Research Objectives. Furthermore, the major goal of this study was to analyze Pakistan's Gram Production Forecasting using the Bayesian Estimation Model. In this same context, this chapter comprehends the complete discussion in terms of their contributions, limitations, and future work respectively. And at the last of this chapter References are provided which are cited for this study.

5.2 Contributions

To meet the straightforward goals and objectives of this manner of thesis, this study holds two research questions i.e. RQ1 and RQ2, and their two objectives.

To meet the objective of the RQ1, this objective deals with the Forecasting and Analysis of Pakistan's Rabbi Crop Gram using Bayesian Estimation. To meet the objective of the RQ2, Multivariate Linear ARIMA-based Bayesian State Space Model using Kalman Filtering and Metropolis-Hastings algorithms provided in previous chapters are used to derive forecasts. Simulations Code is developed for the identified model in MATLAB, and simulations are performed on MATLAB using Rabbi Food Crops including Gram/chickpeas data which consists of previous twenty years' data, which includes Area, Production, Yield, and Prices of crops.

5.3 Limitations and Future Work

The Gaussian approximation-based method presented in this article can be easily extended to nonlinear measurement models by simply replacing the update step with a nonlinear Gaussian filter update step. [53] [54].

The bias in the dataset used is possible because data is collected from online sources and is not 100 percent accurate due to the challenging environment of the country the data is possibly average for the whole years but we assumed the data is correct. And is always challenging to obtain fully correct data all time.

The number of variables can be increased, to increase from four to too many more variables including factors like labor cost, weather coefficient, and prices of fertilizer etc. which will help the researcher to increase the accuracy of the forecast, so it can be illustrated if we increase the number of variables for this model in the future form more than four variables it will increase the accuracy of the forecasts.

5.4 Conclusion

The Bayesian approach/statistics, in general, is a statistical decision approach that provides a tool for combining prior probabilities and their distribution about the nature of states. When working with such models along with time-series data, they too fit commonly when estimating models having multiple (large numbers) attributes above somewhat short length periods. This study aimed to forecast the production of Gram, which included different attributes of gram including cultivation area, production of Gram, the yield of a gram, and the prices of a gram. For time collection or series data, ARIMA-based State-Space modeling is used for forecasting different future attributes for rabbi food crops.

This thesis has dealt with ARIMA-based State-Space form used on different rabbi food crops data to forecast future outcomes. For state-space modeling there are two main models or sets of equations are required to fulfill Bayesian modeling, state estimation model and observation estimation model, for state estimation modeling, Auto-Regressive Integrated Moving Average modeling is used to model the states estimations of the Bayesian modeling and then is further processed by Kalman filtering before giving us the data or statistics for a set of our state equations and often the resulting data is represented by $X(t)$.

Then this output data is further processed by Another part of Bayesian Modeling which is the observation Estimation model, in our case we have a linear Regression equation sort of observation estimation model represented commonly by $Y(t)$, this part of Bayesian state-space modeling is responsible for the further processing of outcome of the state estimation by taking their outcome as input which is X_t and then it processes it by converting it into matrix form then this data is further processed by taking it as a multivariate linear regression having 4 linear attributes which are production, Area, yield, and prices then for each attribute different statistics are withdrawn by further processing it by using Markov Chain Minto Carlo and metropolis Hastings Algorithm.

After completing this complex forecasting process for wheat, Potato, and gram their results the results show the forecasted impact of variables on each other, in which came upon an understanding of their relationship and effect proportionality and their convergence upon histograms, where one's can understand the impact of each variable with regards to the other variables, their discussion is provided in Chapter 4's results and discussion section.

Gram is a major Rabi crop. The variables of Gram have irregular variations in its data trends and for its, forecasting ARIMA-based Bayesian model has been applied. The variables prices, cultivation area, production, and yield of a gram have the foremost effect on the production of Grams. Though, other variables can be incorporated, aimed at better accuracy of the forecasting of Gram on the proposed model in this study. Moreover, division-wise data is also available for Gram whose impact can be calculated by using additional variables to increase the accuracy of the proposed model. Thus, detailed analysis and forecasting of non-linear produce can be done in detail.

We have discussed the state-space form of ARIMA with 4 different variables and showed that it over-fit the conventional ARIMA. This study has shown that the ARIMA-based State-Space model makes things easier for Gram forecasting and can even be used on data with a short-time history. In addition, the ARIMA-based Bayesian State-Space form allows comparing different multiple variables with impact on each other directly using the proposed model, they can be initialized with zero period, so that models of different orders have equal sample sizes.

Using the proposed model and algorithms on gram data to test an ARIMA-based state-space framework that outperforms traditional ARIMA package runs for prediction in MATLAB in terms of accuracy, and We have shown that it works quickly. It seems to work especially well for seasonal data [53].

An advantage of approaches built with Gaussian approximations over many others is that the family of models is fairly general, but the use of Gaussian state approximations keeps the required computations light. The computation is orders of magnitude easier than, for example, the particle MCMC method. The main weakness of the method is that it uses a Gaussian approximation to the state, so it may not give accurate results when the back state is highly non-Gaussian. However, as our experiments show, the results can be accurate if the model is highly linear.

REFERENCES

[1] N. A. (nattapol.aun@mfu.ac.th), "Bayesian Compressed Vector Autoregression for Financial Time-Series Analysis and Forecasting," *IEEE Access,* 2019.

[2] P. H. a. C. Stevens, "Bayesian Forecasting," *Journal of the ROYAL STATISTICAL SOCIETY,* vol. 38, 1976.

[3] J. G. &. H. Tanizaki, "Markov chain Monte Carlo methods for nonlinear and non-gaussian state-space models," *Communications in Statistics - Simulations and Computations,* vol. 28, no. 4, p. 28, 1999.

[4] G. A. H. Sr., "Timely and Accurate Crop Yield Forecasting and Estimation History and Initial Gap Analysis," 28 Octuber 2010.

[5] H. S. C. H. P. Z. a. J. S. Chenghao LIU, "Singapore Management University, chhoi@smu.edu.sg," January 2016.

[6] J. G. &. H. Tanizaki, "Markov chain Monte Carlo methods for nonlinear and non-gaussian state-space models," *Communications in Statistics - Simulation and Computation,* vol. 28, no. 4, p. 28, 1999.

[7] S. Glen., "Autoregressive Model: Definition & The AR Process," *StatisticsHowTo.com: Elementary Statistics for the rest of us!,* 19th August, 2015.

[8] M. Peixeiro, "Time Series Forecasting with Autoregressive Processes," Towards Data Science, 11th June, 2020.

[9] R. J. H. a. G. Athanasopoulos, ""8.4 Moving average models," in Forecasting: Principles and Practice (2nd ed), OTexts: Melbourne, Australia. OTexts.com/fpp2.," 2018.

[10] M. P. M. Guidolin, "Chapter 2 - Autoregressive Moving Average (ARMA) Models and Their Practical Applications,"in Essentials of Time Series for Financial Applications,," 2018.

[11] P. Bruce, "Autoregression and Moving Average (ARMA) Models," statistics.com The Institute for Statistics education, 2002.

[12] R. J. H. a. G. Athanasopoulos, ""8.9 Seasonal ARIMA models," in Forecasting: Principles and Practice (2nd ed), OTexts: Melbourne, Australia. OTexts.com/fpp2.,," 2018.

[13] C. S. K. M. ADAM HAYES, "Autoregressive Integrated Moving Average (ARIMA)," Investopedia, 12 October 12, 2021.

[14] "Autoregressive Integrated Moving Average (ARIMA)," Corporate Finance institute, 2015.

[15] J. Brownlee, "A Gentle Introduction to SARIMA for Time Series Forecasting in Python," Machine Learning Mastery, Making develpors Awesome at Machine Learning, 17 August 2018.

[16] A. HAYES, "investopedia.com," 30 Aprail 2021.

[17] R. J. H. a. G. Athanasopoulos, "Chapter 7 Exponential smoothing," Forecasting: Principles and Practice (2nd ed), OTexts: Melbourne, Australia. OTexts.com/fpp2., 2018.

[18] R. J. H. a. G. Athanasopoulos, "11.2 Vector autoregressions," in Forecasting: Principles and Practice (2nd ed), OTexts: Melbourne, Australia. OTexts.com/fpp2.," 2018.

[19] "Vector Autoregressive models VAR(p) models, Applied time series Analysis,," Department of Statistics, PennState Eberly College of Science.

[20] E. R. W. W. M. U. Warsono, "Modeling and Forecasting by the Vector Autoregressive Moving Average Model for Export of Coal and Oil Data (Case Study from Indonesia over the Years 2002-2017), International Journal of Energy Economics and Policy," 25 January 2019.

[21] W. S. a. M. Deistler, "Chapter 6 - Vector autoregressive moving average models," in *Handbook of Statistics*, Volume 41, 2019., pp. 145-191.

[22] C. Nicholson, "A Beginner's Guide to Neural Networks and Deep Learning," PathMind,

[23] R. J. H. a. G. Athanasopoulos, "11.3 Neural network models," Forecasting: Principles and Practice (2nd ed), OTexts: Melbourne, Australia. OTexts.com/fpp2., 2018.

[24] S. D. R. K. B. K. K. M. M. G. T. M. V. A. G. D. V. J. W. &. C. Y. R. v. d. Schoot, "Bayesian statistics and modeling," *Nature Reviews Methods Primers,* p. 26, 14 January 2021.

[25] S. Ray, "6 Easy Steps to Learn Naive Bayes Algorithm with codes in Python and R," Analytics Vidhya, 11 September 2017.

[26] P. C. S. R. &. A. Sureshbabu, "An Applied Time Series Forecasting Model for Yield Prediction of Agricultural Crop," in *SpringerLink, Soft Computing and Signal Processing*, Singapore, 2020.

[27] S. P. P. K. S. S. K. B. Sourav Kumar Purohit, "Time Series Forecasting of Price of Agricultural Products Using Hybrid Methods," *Applied Artificial Intelligence* , vol. 35, no. 15, pp. 1388-1406 , 2021.

[28] M. A. a. A. A. M. Amin, "TIME SERIES MODELING FOR FORECASTING WHEAT PRODUCTION IN PAKISTAN," *Journal of Animal and Plant Sciences,* vol. 24, no. 5, pp. 1444-1451, 2014.

[29] J. V. |. C. R. B. |. G. G.M., "APPLICATION OF ARIMA MODEL FOR FORECASTING AGRICULTURAL PRICES," *JOURNAL OF AGRICULTURAL SCIENCE AND TECHNOLOGY,* vol. 19, no. 4, pp. 981-992, 2021.

[30] D. A. B. Jon A. Brandt, "Price forecasting and evaluation: An application in agriculture," *Journal of Forecasting,* vol. 2, no. 3, pp. 237-248, 1983.

[31] Q. W. Y. Z. H. e. a. Wen, "Application of ARIMA and SVM mixed model in agricultural management under the background of intellectual agriculture.," *Springerlink, Cluster Computing,* vol. 22, no. 5, p. pages14349–14358, 2019.

[32] H. A. Shakir Khan, "ARIMA Model for Accurate Time Series Stocks Forecasting," *(IJACSA) International Journal of Advanced Computer Science and Applications,,* vol. 11, no. 7, 2020.

[33] K. B. A. M. A. A. S. A. NAJEEB IQBAL, "Use of the ARIMA Model for Forecasting Wheat Area and Production in Pakistan," *JOURNAL OF AGRICULTURE & SOCIAL SCIENCES* , vol. 120, no. 122, p. 1813–2235, 2020.

[34] M. S. MA Awal, "Rice Production in Bangladesh Employing by ARIMA Model," *Bangladesh Journal of Agricultural Research,* vol. 36, no. 1, p. 51–62, 2011.

[35] P. K. a. K. P. D.P. Singh, "Application of ARIMA model for forecasting Paddy production in Bastar division of Chhattisgarh," *American International Journal of Research in Science, Technology, Engineering & Mathematics* , pp. 2328-3491, 2013.

[36] R. B. &. M. A. A. Anil KumarMahto, "Short Term Forecasting of Agriculture Commodity Price by Using ARIMA: Based on Indian Market," in *Springer, Advances in Computing and Data Sciences*, Singapore, 2019.

[37] M. D. W. W. A. J. A. D. F. Bédard, "Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods," *Agricultural and Forest Meteorology,* Vols. 218-219, pp. 74-84, 2016.

[38] C. F. G. K. S. A. C. a. W. H. K. MATSUMURA, "Maize yield forecasting by linear regression and artificial neural networks in Jilin, China," *The Journal of Agriculture Science*, vol. 153, no. 3, pp. 399 - 410, 2014.

[39] Y. Y. W. C. &. Y. Z. Jiejun Huang, "Development of a Data Mining Application for Agriculture Based on Bayesian Networks," in *Springerlink, Computer And Computing Technologies In Agriculture*, 2007.

[40] A. LonardodP.MilazzoaS.Chessaa, "Dynamic Bayesian network for crop growth prediction in greenhouses," *Science Direct, Computers and Electronics in Agriculture,* vol. 169, 2020.

[41] N. Gandhi, L. J. Armstrong and O. Petkar, "PredictingRice crop yield using Bayesian networks," in *IEEE Explore*, Jaipur, India, 2016.

[42] R. Priya, D. Ramesh and E. Khosla, "Crop Prediction on the Region Belts of India: A Naïve Bayes MapReduce Precision Agricultural Model," in *IEEE Explore*, Bangalore, India, 2018.

[43] C.-J. K. a. C. R. Nelson, "State-space models with regime switching: classical and Gibbs-sampling approaches with applications," *MIT Press Books,* vol. 1, 1999.

[44] B. Arulampalam, "Beyond the Kalman Filter: Particle Filters for Tracking Applications," Artech House.

[45] G. F. Welch, "Kalman filter," in *Computer Vision: A Reference Guide*, 2014.

[46] M. H. M. I. a. J. R. S. Ghauri, "Performance analysis of ldpc codes on different channels, in Next Generation Mobile Apps, Services and Technologies (NGMAST)," in *Eighth International Conference*, September, 2014.

[47] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Fluids Engineering*, vol. 82, no. 1, p. 35–45, 1960.

[48] G. F. Welch, "Kalman filter," *Computer Vision: A Reference Guide,* p. 435–437, 2014.

[49] William Regli, "Machine Learning, Linear Regression Kalman Filters,"

[50] H. D. S. H. Xiaohui Li, "Multiple Linear Regression with Kalman Filter for Predicting End Prices of Online,"

[51] R. Nau, "Statistical forecasting: notes on regression and time series analysis," Fuqua School of Business Duke University.

[52] P. S. W. C. D. o. Statistics, "STAT 510: Applied Time Series Analysis.," PennState Eberly College of Science.

[53] K. M. a. A. Madhu, "AN APPLICATION OF TIME SERIES ARIMA FORECASTING MODEL FOR PREDICTING SUGARCANE PRODUCTION IN INDIA," 2014.

[54] F. a. R. KONRADREIF, "An EKF-Based Nonlinear Observer with a Prescribed Degree of Stability," in *Automatica*, vol. 34, Science Direct, September 1998, pp. 1047-1158.

[55] S. S. &. H. H. Isambi S. Mbalawata, "Parameter estimation in stochastic differential equations with Markov chain Monte Carlo and non-linear Kalman filtering," *Springer,* pp. 1195-1223, Published: 03 August 2012.

[56] J. H. I. S. M. a. H. H. S. S¨arkk¨a, "Posterior inference on parameters of stochastic differential equations via nonlinear gaussian filtering and adaptive mcmc," *Statistics and Computing,* p. 1–11, 2013.

[57] S. S. a. H. H. I. S. Mbalawata, "Parameter estimation in stochastic differential equations with markov chain monte carlo and nonlinear kalman filtering," *Computational Statistics,* vol. 28, no. 3, p. 1195–1223, 2013.

[58] C. A. A. D. Nick Whiteley, "Efficient Bayesian Inference for Switching State-Space Models using Discrete Particle Markov Chain Monte Carlo Methods," *arXivLabs Cornell University*, 10 Nov 2010.

[59] Y. Y. a. T. Xu, "An Adaptive Kalman Filter Based on Sage Windowing Weights and Variance Components," *The Journal of Navigation,* vol. Volume 56, no. Issue 2, pp. pp. 231 - 240, 13 May 2003.

[60] "MathWorks," © 1994-2022 The MathWorks, Inc., 1994-2022.

[61] K. I. a. K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Transactions on Automatic Control,* vol. 45, no. 5, pp. 910-927, May, 2000.

[62] S. S. a. J. Sarmavuorib, "Gaussian filtering and smoothing for continuous-discrete dynamic systems," *Signal Processing (Science Direct),* vol. 93, no. 2, pp. Pages 500-510, February, 2013.

[63] I. S. a. J. E. Boylan, "State-space ARIMA for supply-chain forecasting," *International Journal of Production,* vol. 54, no. 3, 04 Apr 2019.

[64] S. Glen., "Statistics HowTo : Elementary Statistics for the rest of us!," 19 August 2015.

[65] R. J. H. a. G. Athanasopoulos, ""8.5 Non-seasonal ARIMA models," in Forecasting: Principles and Practice (2nd ed), OTexts: Melbourne, Australia. OTexts.com/fpp2.,," 2018.

[66] W. S. a. M. Deistler, "Chapter 6 - Vector autoregressive moving average models," *Handbook of Statistics,* vol. 41, pp. 145-191, 2019.

[67] G. D’Agostini, "A multidimensional unfolding method based on Bayes’ theorem," *Nuclear Instruments and Methods in Physics Research,* p. 12, 2 March 1995.

[68] M. Zhang, "University of Pittsburgh," 23, October 2018.

[69] I. P. Statistics House, "Pakistan Bureau of Statistics,," Statistics House, 21-Mauve Area, G-9/1, Islamabad, Pakistan.

[70] G. o. P. Finance Division, "Economic Survay of Pakistan," Economic Adviser’s Wing, Finance Division, Government of Pakistan, Islamabad.

[71] U. Nations, "Food and Agriculture Organization of the United Nations," Viale delle Terme di Caracalla, 00153 Rome, Italy,T: (+39) 06 570 55303.

[72] G. o. Pakistan, "Agriculture Marketing Information System (AMIS)," Directrate of Agriculture Punjab, Lahore .