YANGON UNIVERSITY OF ECONOMICS
DEPARTMENT OF STATISTICS
Ph.D. PROGRAMME

MODELLING THE IMPACT OF CLIMATE CHANGE ON RICE PRODUCTION IN AYEYAWADY REGION

THET MAR LWIN

AUGUST, 2023

YANGON UNIVERSITY OF ECONOMICS
DEPARTMENT OF STATISTICS
Ph.D. PROGRAMME

MODELLING THE IMPACT OF CLIMATE CHANGE ON RICE PRODUCTION IN AYEYAWADY REGION

Submitted in Partial Fulfillment of the Requirement for the Degree of Doctor of Philosophy (Ph.D.) of Statistics, Yangon University of Economics

Supervised by:

Dr. Mya Thandar
Pro-Rector
Yangon University of Economics

Submitted by:

Thet Mar Lwin
4 Paragu Ah-1

AUGUST, 2023

YANGON UNIVERSITY OF ECONOMICS
DEPARTMENT OF STATISTICS
Ph.D. PROGRAMME

This is to certify that this dissertation entitled "**Modelling the Impact of Climate Change on Rice Production in Ayeyawady Region**" submitted as the requirement for the Degree of Doctor of Philosophy (Ph.D.) in Statistics has been accepted by the Board of Examiners.

BOARD OF EXAMINERS

Prof. Dr. Tin Tin Htwe
(Chairman)
Rector
Yangon University of Economics

Dr. Soe Win
(External Examiner)
Permanent Secretary
Department of Higher Education
Ministry of Education

Prof. Dr. Lay Kyi
(Referee)
Pro-Rector (Retired)
Yangon University of Economics

Prof. Dr. San Kyi
(Member)
Professor (Retired)
Yangon University of Economics

Prof. Dr. Khin May Than
(Member)
Professor / Head (Retired)
Yangon University of Economics

Prof. Dr. Swe Swe Zin
(Internal Examiner)
Professor / Head
Department of Shipping Management
Myanmar Maritime University

Prof. Dr. Aye Thida
(Member)
Professor / Head
Department of Statistics
Yangon University of Economics

Prof. Dr. Mya Thandar
(Promoter)
Pro-Rector
Yangon University of Economics

AUGUST, 2023

CERTIFICATION

I hereby certify that content of this dissertation is wholly my own work unless otherwise referenced of acknowledged. Information from sources is referenced with original comments and ideas from the writer herself.

Thet Mar Lwin
4 Paragu Ah-1

ABSTRACT

This study provides a comprehensive analysis of the impact of climate change on the rice yield in the Ayeyawady Region. The rainfall, temperature (maximum and minimum), and relative humidity at 9 AM, and 6 PM are considered as the climatic variables in this study. The secondary data were collected from Pathein, Hinthada, Maubin and Myaungmya Districts for the period from 1992-1993 to 2020-2021 focusing on monsoon (May to October) and summer (November to April). The Multiple Linear Regression (MLR), Seasonal Autoregressive Integrated Moving Average with Predictors (SARIMAX), Vector Autoregressive (VAR) and Artificial Neural Network (ANN) models were used to analyze the impact of climatic variables on rice yield. The findings reveal that maximum temperature and rainfall have negative effects, whereas minimum temperature and humidity have positive effects on rice yield in all districts. The ANN model was the most appropriate model for forecasting the rice yield. The actual and forecast values were found to be quite close, and the yield of summer rice was higher than that of monsoon rice. It was further recognized that the rice yield could be increased by fostering sustainable agricultural practices, planting climate-resilient rice crop varieties, implementing water management strategies, composting crop residues, providing timely weather forecasts for farmers and improving farmer’s adaption to climate change. A further study could be conducted to examine the impact of climatic variables on other types of crop in different States and Regions. It was also recommended to include the various socioeconomic variables in order to analyze the changes in rice yield.

ACKNOWLEDGEMENTS

I am grateful to many people who in different ways contributed to the successful completion of this thesis.

First and foremost, I wish to extend my sincere gratitude to thanks to Prof. Dr. Tin Tin Htwe, Rector, Yangon University of Economics; Prof. Dr. Mya Thandar, Pro-Rector; Prof. Dr. Khin Thida Nyein, Pro-Rector; Prof. Dr. Cho Cho Thein, Pro-Rector; and Prof. Dr. Tin Tin Wai, Pro-Rector, Yangon University of Economics for their kind permission to submit this Ph.D. dissertation.

Secondly, my grateful thanks are extended to Prof. Dr. Soe Win, Permanent Secretary, Ministry of Education, for his invaluable guidance and advice. I deeply thank to Prof. Dr. Lay Kyi, Pro-Rector (Rtd.), Yangon University of Economics because I would never have been able to finish my dissertation without his guidance and support for the documents needed.

Also, I would like to express my gratitude to Prof. Dr. San Kyi (Rtd.), Department of Statistics, Yangon University of Economics, who gave me her valuable guidance, scholarly inputs and consistent encouragement throughout this research work. I would like to express my special appreciation and thanks to Pro. Dr. Khin May Than, Head of Department (Rtd.), Department of Statistics, Yangon University of Economics, for her sympathetic attitude and guidance for this dissertation.

Moreover, I am indebted to Prof. Dr. Swe Swe Zin, Head of Department, Department of Shipping Management, Myanmar Maritime University, for valuable suggestions and help to complete this dissertation. I would like to express my special appreciation and thanks to Prof. Dr. Aye Thida, Head of Department, Department of Statistics, Yangon University of Economics, for her valuable advice and encouragement.

I would like to express my deep and sincere gratitude to my supervisor, Prof. Dr. Mya Thandar, Pro-Rector Yangon University of Economics, for her support to carry out this thesis. She continuously encouraged, guided and helped me throughout research work, and she has constantly forced me to remain focused on achieving my goal.

Thanks must go to Daw Khin Myo Aye, Associate Professor, Head of Department, Department of English, Monywa University of economics, who patiently checked the language in my writing and shared her experience. I am also appreciative of the Department of Agriculture in the Ayeyawady Region and the Department of Meteorology and Hydrology, Myanmar, for granting me permission to utilize the data for the thesis.

I would like to thank to all Professors, Associate Professors, Lecturers, Assistant Lecturers and Tutors from Department of Statistics, Yangon University of economics, who help me to complete this dissertation.

Finally, I would like to convey my gratitude to my family for their understanding, patience, and support in helping me successfully complete my dissertation.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACF	Autocorrelation Function
ADB	Asian Development Bank
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
ANN	Artificial Neural Network
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
ARIMAX	Autoregressive Integrated Moving Average with Predictors
ARMA	Autoregressive Moving Average
BIC	Bayesian Information Criterion
CO_2	Carbondine Oxide
DOA	Department of Agriculture
ERH	Evening Relative Humidity
ET	Evapotranspiration
FAO	Food and Agriculture Organization
FEVD	Forecast Error Variance Decomposition
GDP	Gross Domestic Product
GHG	Greenhouse Gases
GLS	Generalized Least Squares
HQ	Hannan-Quinn
IRRI	International Rice Research Institute
IWR	Irrigation Water Requirement
Kg	Kilogram
MA	Moving Average
MAPE	Mean Absolute Percentage Error
MATLAB	Matrix Laboratory
MaxT	Maximum Temperature
MCCA	Myanmar Climate Change Alliance
MCR	Myanmar Climate Report
MCSA	Myanmar Climate-Smart-Agriculture
ME	Mixed Effect

MET	Meterology
MinT	Minimum Temperature
MLE	Maximum Likelihood Estimator
MLP	Multilayer Perceptron
MLR	Multiple Linear Regression
MR	Multiple Regression
MRH	Morning Relative Humidity
MSDP	Myanmar Sustainable Development Plan (MSDP)
MSE	Mean Squared Error
MT	Metric Tons
NNs	Neural Networks
OLS	Ordinary Least Squares
PACF	Partial Autocorrelation Function
PPR	Projection Pursuit Regression
RCP	Representative Concentration Pathway
RF	Rainfall
RMSE	Root Means Squared Error
SARIMA	Seasonal Autoregressive Integrated Moving Average
SARIMAX	Seasonal Autoregressive Integrated Moving Average with Predictors
SC	Schwarz Criterion
SDG	Sustainable Development Goals
SEP	Standard Error of Prediction
SMLR	Stepwise Multiple Linear Regression
SSE	Sum of Squares Error
SSR	Sum of Squares Regression
SST	Sum of Squares Total
UNDP	United Nations Development Programme
UNEP	United Nations Environment Programme
UN-Habitat	United Nations Human Settlement Programme
USDA	United States Department of Agriculture
VAR	Vector Autoregressive
VIF	Variance Inflation Factor
WMO	World Meteorological Organization

CHAPTER I
INTRODUCTION

Agriculture is one of the most climate dependent human activities, as it is highly sensitive to different hydro-climatic conditions, such as rising of temperature and changes in rainfall. Myanmar is the second country in the world most affected by climate change from 1993 to 2014 (Lar et al., 2018). It will remain one of the most vulnerable countries, given the projected changes in extreme weather and climate events, being to be agricultural productivity less, and sea-level rise (UN Environment, 2019).

Rice is one of the most essential cereal crops. It is consume to more than three billion people, especially half of the world's population. A total of 95 percent of the world's rice crop is eaten by humans. Rice production in Myanmar faces several challenges, including the rice sector's vulnerability to climate conditions like higher temperatures, drought, flooding and other climatic stresses. The sector is also challenged by its demand for water, land, fertilizer and pesticides and its own environmental impact, as it has a significant contribution to greenhouse gas emissions (Oo, 2020).

The Ayeyawady Region is the rice bowl of Myanmar, but it is highly vulnerable to the climate change. Water drainage, salt intrusion and flood protection are major destructions for this region. Farmers in the Ayeyawady Region usually encounter the adverse impact of climate change (World Bank, 2016). The Cyclone Nargis severely hit the Ayeyawady Region in 2008. Therefore, the impact of climate change on rice production in this region is needed to be explored and it is considered as a major issue in this study.

1.1 Rationale of the Study

The Republic of the Union of Myanmar is geographically situated in Southeast Asia with latitudes between 09° 32' N and 28° 31' N, and longitudes between 92° 10' E and 101° 11' E. The total area of Myanmar is 261,228 square miles (677,000 square kilometers). It stretches for 582 miles (936 kilometers) from east to west and 1,275 miles (2,051 kilometers) from north to south. The climate of Myanmar is determined by its geographical position. Myanmar is bordered on the north and northeast by the People's Republic of China, on the east and southeast by the Lao People's Democratic Republic and the Kingdom of Thailand, on the south by the Andaman Sea and the Bay of Bengal and on the west by the People's Republic of Bangladesh and the Republic of India (Myanmar in Brief, 2018).

It is separated from neighboring countries by high mountain walls and also situated in the tropical climate region. It has three seasons, the winter from October to February with average temperature lying between 68˚F and 75˚F (20˚C - 24˚C), a summer season in March to May with average temperature lying between 86˚ F and 95˚F (30˚C - 35˚C) and a rainy season from June to September with average temperatures lying between 77˚F and 86˚F (25˚C - 30˚C). There are also three main agro-ecological zones such as Delta and Coastal zone, Central dry zone, and Hilly zone (Aung, Zin & Theingi, 2017).

Myanmar is one of the countries that are most vulnerable to the impact of climate change. Today, Myanmar has experienced with meteorological, hydrological and seismic hazards such as cyclones (strong winds), floods (tidal surges), intense rains, extremely high temperature, droughts and the rise of sea level (Myanmar National Study, 2015). The Great Sittwe Cyclone of 1968, the Pathein Cyclone of 1975, the Gwa Cyclone of 1982, the Maungdaw Cyclone of 1994, the Cyclone Mala of 2006, the Cyclone Nargis of May 2008, the effect of the Cyclone Koman and the historical floods encountered in year 2004, 2010 and 2015 were all extreme meteorological and hydrological events (Aung, Zin & Theingi, 2017).

The uncontrollable natures of climatic factors have changed over time and affected agricultural output, fisheries, livestock, forest, water resources, biodiversity, energy, industry, transport, human settlements, and cities and public health in Myanmar. Among them, climate change has the largest and the most significant impact on agriculture. The impact of climate change on agricultural production varies from country to country with different economic conditions, region to region, and from time to time (Alam et al., 2014).

Since climate variability is the major factor that significantly influences the agriculture productivity, the impact of climate change on agriculture has become an important issue for the countries with an agriculture-based economy. Agriculture production is highly dependent on climate and it is also adversely affected by increasing climate variability (Parekh & Suryanarayana, 2012).

The economy and society of Myanmar is still largely dependent on agriculture, which relies heavily on the rain. Thus, if there is too much rain or less rain, it would have significant impact on agriculture. Too much rainfall will cause floods and less or no rain will cause droughts, both of which will affect the productivity of agricultural products. The highly productive deltaic and low-lying coastal rice cultivation areas are usually exposed to increase salinity, coastal erosion and inundation. Particularly, the salinity and inundation are the most destructive to the agricultural land causing lower production of agriculture products. Decreased income from the agricultural sector can have many adverse effects on the income of individual farmers and also the economy of the country (Myanmar Climate Change Alliance, 2017).

With an agricultural sector that weighs so heavily on the livelihood of Myanmar’s people, climate changes will have a disproportionately negative impact on different areas of the country. An increase in extremely high temperatures has already creating problems such as the severe drought occurred in 2009, which affected major cereal crops in the dry zone including Sagaing, Mandalay and Magway Regions. Moreover, in 2010, severe drought diminished village water sources across the country and destroyed agricultural yields of peas, beans, pulses, sugar cane, tomato and rice, the main cash crops of Myanmar. Also the Zawgyi River floods in October 2006 caused extensive crop damage. In 2007, extensively record-breaking flooding resulted in the inundation of 809,284 hectares of crop land and more than 50 percent of crops were damaged. Again during July and October 2011 heavy rain and flooding in the Ayeyawady and Bago Regions, Mon and Rakhine States, resulted in losses of approximately 1.7 million tons of rice. The excessive sedimentation in the Rakhine State in 2010 damaged rice seedlings and reduced harvests (Myanmar Climate Change Alliance, 2017).

Agriculture is the backbone of the Myanmar economy because it contributes to the economic and social well-being of the entire nation and it influences the gross domestic product (GDP) and employment. Moreover, the agriculture sector is one of the most important concerns for the country’s economy; agricultural goods are Myanmar’s second largest export commodity. The agriculture sector generally contributes about 38% of GDP, accounts for 20% to 30% of total export earnings and employs more than 70% of the workforce. Out of 67.6 million hectares of land in Myanmar, 12.8 million hectares are cultivated land. In Myanmar, 70% of the country's population lives in rural areas and their livelihood drive the agriculture sector as an important growth engine of rural development. Myanmar's top agricultural exports include rice, maize, black gram, green gram, pigeon pea, chick pea, sesame, onion, tamarind, raw rubber, vegetables, and fruits. Among them, rice (paddy) is the major crop for both economy and food security of the country. Rice is the country's primary agricultural product, which accounts for nearly 43% of the total agricultural production value almost every year (Agriculture Overview, 2018).

Myanmar has a long tradition of rice production. In the years immediately before World War II, Myanmar has been the largest rice-producing nation in the world. The production of rice area declined during the post-war era and has since failed to reach the levels achieved during the pre-war period (Naing et al., 2008). Rice (paddy) is sown on 15658 thousand acres of lands (48% of net sown land) and is the most common crop choice for farmers (Agriculture Guide, 2020).

According to the Department of Agriculture (DOA) Statistics, national average yields for monsoon paddy (rice) were about 3.8 MT/Ha and for summer paddy were 4.6 MT/Ha in 2016- 2017. In the early 1960s, annual exports were in the range of 1.3 to 1.7 million MT (USDA data from World Rice Statistics). In recent years, exports have dropped below 1 million MT per annum, as population growth has outpaced productivity improvement. The national average per capita consumption of rice is ranged between 160 and 208 kg in Myanmar. This places Myanmar as one of the highest rice consumers in the world on a per capita basis (Aung, 2018).

The major rice producing regions of Myanmar include Ayeyawady, Bago and Sagaing Regions, making up almost all of the country's harvested rice area. According to the Department of Agriculture, the Ayeyawady Region covers about 28% of total paddy (rice) production, followed by Bago Region at about 17% and the Sagaing Region at about 12% in 2017/18 (Aung, 2018). Moreover, the Ayeyawady region is known as the "rice bowl" of the country. Thus, it has come to be as the lifeline of Myanmar's economy (Zaw et al., 2011).

Myanmar's paddy fields can be found mostly in the delta and central dry zone areas (the Ayeyawady and Sittoung River Basins). Paddy production has been increased by dry season paddy cultivation, which has followed rainy season paddy cultivation since 1992. Ninety percent of the annual rainfall in different regions of Myanmar is received during the rainy season. In addition to the rainy season, there may be rain due to low air pressure in other seasons, such as Thingyan rain in April. The dry season paddy is mostly cultivated in lower Myanmar using irrigation. The government of Myanmar strongly supported summer rice (Naing, 2005).

Ayeyawady Region experienced the destruction of Cyclone Nargis in May 2008 due to the adverse effect of climate change. Because of Cyclone Nargis, agricultural land, livestock and fisheries had been extensively damaged (Kyi, 2016). Impact analysis from Cyclone Nargis indicated that approximately 1.75 million hectares of rice producing land was inundated with salt water, and remains flooded at that time (USDA, 2008). In addition, Cyclone Nargis jeopardized the country's food security and export of agriculture products to many foreign countries (International Rice Research Institute, IRRI, 2020).

These conditions are likely to contribute substantially to food insecurity in the future, by increasing food prices, and reducing food production. It was also found that extreme weather events, associated with climate change may cause sudden reductions in agricultural productivity, leading to rapid price increases. The rising prices forced growing numbers of local people into poverty, providing a sobering demonstration of how the influence of climate change can result in food insecurity (Myanmar Climate-Smart-Agriculture, MCSA, 2019).

Therefore, this study intends to investigate the impact of climate change on rice production in Ayeyawady Region. The rice production is measured by rice yield (yield per acre) in this study. Accordingly, the study attempted to identify the major climatic conditions which have the impact on ice production and to choose the most appropriate model for assessing such as impact of climate change on rice yield of the Ayeyawady Region in Myanmar. In addition, the forecast values of the future rice yield are obtained in the study.

1.2 Objectives of the Study

The main objective of this study is to analyze the changes in rice yield due to the impact of climate change often encountered in the Ayeyawady Region.

The specific objectives of the study are:

(1) to investigate the relationship between rice yield and the climate change over time in Ayeyawady Region,

(2) to identify the most suitable statistical model for assessing the impact of climate change on rice yield in Ayeyawady Region and
(3) to forecast the rice yield of Ayeyawady Region using the most suitable model.

1.3 Method of Study

In this study, the secondary data were used to analyze the changes in rice yield due to the climatic variables. The characteristics of the climatic variables and rice yield were summarized by the descriptive statistics. The Multiple Linear Regression (MLR) model was first used to find the relations of the rice yield by assessing the complex connections with the independent climatic variables. Moreover, the Seasonal Autoregressive Integrated Moving Average with predictor variables (SARIMAX) model was used for modeling and forecasting rice yield. Furthermore, the study developed a Vector Auto-Regressive (VAR) model to provide useful heuristics for understanding the empirical causal relationships between climate variables and rice yield.

In addition, the present study is planned to investigate the potential for applying an Artificial Neural Network (ANN) model to forecast rice yield with climatic variables in the study area. The ANN has been widely used in time series predictions because of their characteristics of robustness, fault tolerance, and adaptive learning ability. Finally, a comparison of the models developed in this study was carried out to make sure the accuracy and reliability.

1.4 Scope and Limitations of the Study

In this study, the rice production in term of yield per acre was analyzed based on the impact of climate change only. The main focus of this study is Ayeyawady Region which owns the highest rice yield in Myanmar. However, this study includes only five districts which have the weather stations in Ayeyawady Region such as Pathein, Hinthada, Maubin, Myaungmya and Phyarpon. Therefore, the monsoon and summer rice yield data as well as the agro-meteorological data including Rainfall, Maximum Temperature, Minimum Temperature, Relative Humidity at 9:30 AM (Morning RH) and Relative Humidity at 6:30 PM (Evening RH) were collected from those five districts.

In more detail, the average rice yield per acre in basket (Bsk/Ac) were collected for the period from 1992-1993 to 2020-2021 through the Department of Agriculture in Ayeyawady Region, whereas the meteorological data were collected at the monthly scale for the period from 1992 until 2021 through the Department of Meteorology and Hydrology, Myanmar. Unfortunately, the meteorological data of Phyarpon district were available only for the period from 2001 to 2020. Moreover, the farmers' adaptations of climate change and other factors impacted on rice yield have not been considered in this study because the crop management systems are frequently changed over time and it is quite difficult to take into account those things.

1.5 Organization of the Study

This study is structured into five main chapters. Chapter I is the introduction which includes the rationale of the study, objectives of the study, scope, and limitations of the study, method of study, and organization of the study. Chapter II mentions a literature review relating to the effect of climate change on rice yield. In Chapter III, the methodology is presented that helps to identify the models for assessing the impact of climate change on rice yield. Chapter IV analyzes the various models being used in fitting the impact of climate change on rice yield and the most appropriate one will then be selected by comparing those stated models so that the future value of rice yield could be forecasted. After that, the conclusion with findings and suggestions for further research studies are presented in Chapter V.

CHAPTER II
LITERATURE REVIEW

The climate conditions can have a significant impact on crop production in different parts of the country. The review highlighted the works of many researchers who analyzed the impact of climate change on crop production.

2.1 Global Climate Change

The Food and Agriculture Organization (FAO) of the United Nations (2008) mentioned that climate refers to the characteristic conditions of the earth's lower surface atmosphere at a specific location; weather refers to the day-to-day fluctuations in these conditions at the same location. The variables commonly used by meteorologists to measure daily weather phenomena are air temperature, precipitation, atmospheric pressure and humidity, wind, sunshine and cloud cover.

Global climate is the average temperature of the earth's surface and the atmosphere in contact with it, and is measured by analyzing thousands of temperature records collected from stations all over the world, both on land and at sea. Most current projections of climate change refer to global climate, but climate can also be described at other scales, based on records for weather variables collected from stations in the zones concerned.

Sengar & Sengar (2015) pointed out that climate change is looming large for humanity in the coming decades. Agriculture also produces significant effects on climate change as a possible contributor of greenhouse gases to the atmosphere and as an industry that is highly sensitive to climate change. Global warming and climate change are often interchangeably used and understood, but these terms are not identical. Climate change includes both warming and cooling conditions, while global warming pertains only to climatic changes related to increases in temperatures. The climatic system is a complex interactive system consisting of the atmosphere, land surface, snow and ice, oceans, and other bodies of water and living things. The atmospheric component of the climatic system most obviously characterizes climate. It is often defined as ‘average weather." The climate is usually described in terms of the mean and variability of temperature, precipitation, and wind over a period ranging from months to millions of years.

All over the world, there has been a slow but steady rise in temperature over the last few decades. Moreover, alongside this warming, the globe has also been subject to a general decline in rainfall since the first half of the nineteenth century. While one may be inclined to think only in terms of more dramatic weather events such as floods, droughts, storms, and hurricanes, adversely affecting agricultural production, it is important to note that even little climate change could feasibly have substantial effects, particularly if countries do not have the necessary technology and endowments to deal with these. Indeed, agronomic models of climate sensitivity suggest that climate changes in most developing countries are likely to be harmful and can make less productive agricultural areas.

Climate change will have a profound impact on human and ecosystems in the coming decades through variations in global average temperature and rainfall. Climate change poses unprecedented challenges to human society and ecosystems in the twenty-first century, particularly in the developing nations in the tropics. The accelerating pace of climate change combined with global population and income growth threatens food security. Populations in the developing world which are already vulnerable and food insecure are likely to be more seriously affected. The impact of climate change will persist. This will affect the basic elements of life around the world such as access to water, food production, healthcare and the environment. Millions of people could suffer from hunger, water shortage and coastal flooding as the world gets warmer. There are certain regions, sectors, ecosystems and social groups which will be affected the most by climate change and the consequences of economic globalization. Managing the impact of climate change, therefore, poses a challenge to governments and societies.

Horton et al. (2017) described that the global climate is changing. Scientists from around the world have come to a consensus that global temperatures are rising and that human activities which emit greenhouse gases into the atmosphere are causing many of these changes. Additionally, 15 out of the 16 warmest years have occurred since 2000. Changes in greenhouse gas concentrations not only increase the global temperature but also have far-reaching effects on the climate. These effects are projected to include shifts in precipitation patterns, sea levels, heat extremes, storms, monsoon cycles, ocean currents, sea surface temperatures, land ice mass, and river flows. Natural and man-made systems alike will continue to be significantly affected by these changes. Species and ecosystems may experience climate conditions outside of those under which they have evolved to survive. Human settlements and economies will be impacted across the globe. Farmers will face increasing droughts and floods, and cities will have to respond to more frequent extreme events that will affect infrastructure and health. Coastal communities on every continent will face inundation from floodwaters reaching further inland as a result of sea level rise. Local and regional decision-makers need to be aware of the changes that will directly affect the areas they govern. These changes can be met with measures that improve the resiliency of human and natural systems to new climatic patterns and extreme events. By understanding local risks, action can be taken now to prevent the worst impacts of climate change in communities in Asia and around the globe.

World Meteorological Organization (WMO), (2020) expressed that climate describes the average weather conditions for a particular location and over a long period of time. Its variations and extremes, and its influences on a variety of activities including human health, safety and welfare are studied to support evidence-based decision-making on how to best adapt to a changing climate. WMO helps its Members to monitor the Earth's climate on a global scale so that reliable information is available to support evidence-based decision-making on how to best adapt to a changing climate and manage risks associated with climate variability and extremes. Climate information is essential for monitoring the success of efforts to reduce greenhouse gas emissions that contribute to climate change, as well as for promoting efforts to increase energy efficiency and to transition to a carbon-neutral economy.

The FAO of the United Nations (2020) described that climate change threatens the ability to ensure global food security, eradicate poverty and achieve sustainable development. Greenhouse gas (GHG) emissions from human activity and livestock are significant driver of climate change, trapping heat in the earth's atmosphere and triggering global warming. The consequences of climate change including changing rainfall patterns, drought, flooding and the geographical redistribution of pests and diseases that has both direct and indirect effects on agricultural productivity. The vast amounts of CO_2 absorbed by the oceans causes acidification, influencing the health of our oceans and those whose livelihoods and nutrition depend on them.

2.2 Climate Change in Myanmar

Myanmar usually experiences a tropical-monsoon climate with three dominant seasons: summer from March to May, rainy from June to October and winter from November to February. Myanmar consists of eight major physiographic regions: The Ayeyawady Delta, Central Dry Zone, Northern Hilly Region, Rakhine Coastal Region, Eastern Hilly Region, Southern Coastal Region, Yangon Deltaic Region, and Southern Interior Region.

Figure (2.1): Physiographic Regions of Myanmar

Source: Myanmar's National Adaptation Programme of Action (2012)

Myanmar receives most of its rainfall during the rainy season. In the summer and winter seasons, there is little rainfall, especially in the winter, yielding very little rainfall in all regions. There are pronounced regional differences in climate. The Central Dry Zone is a large inland swath of the country that is prone to extreme heat events and drought. The rainy coasts, such as the Rakhine, Southern Coastal, and Yangon Deltaic areas, are slightly cooler in annual average temperature but are prone to flooding.

Further inland are the cooler Northern and Eastern Hilly regions, which experience heat waves, droughts, and floods. The Yangon Deltaic Region has the highest mean temperature. Because of its higher elevation, the Northern Hilly Region has the lowest mean and maximum annual temperature. This pattern remains consistent for seasonal temperatures, such that the Yangon Deltaic has the highest mean and maximum annual temperatures for the summer, winter, and rainy seasons. The pattern for the Northern Hilly Region is similar, with the lowest mean and maximum annual temperature for the summer and winter seasons, with only one exception the rainy season has the same mean annual temperature as the Eastern Hilly Region.

In the summer and winter, the Southern Coastal Region receives the most rainfall, and it is observed that the second-highest rainfall is in the Northern Hilly Region (summer season) and Ayeyawady Delta (winter season). The highest annual precipitation happens in the Rakhine Coastal Region, followed by the Ayeyawady Delta, with the same pattern in the rainy season. It is also observed that the lowest annual precipitation is in the Eastern Hilly Region, followed by the Northern Hilly Region. These regions also receive the lowest wet-season precipitation, with the Eastern Hilly Region receiving the lowest, followed by the Northern Hilly Region. The rainy conditions are the same pattern during the wet season. In the summer, the lowest precipitation occurs in the Eastern Hilly Region, followed by the Southern Interior Region. In the winter, the Southern Interior Region receives the least rainfall, followed by the Yangon Deltaic Region.

Myanmar has already experienced climate change over recent decades. Although climate change trends that span only a few decades are often statistically weak at individual weather stations, a robust signal emerges when considering many weather stations at once. National average daily temperatures based on 19 weather stations across Myanmar increased by about 0.25°C per decade from 1981 to 2010, and daily maximum temperatures have risen at a slightly faster rate of 0.4°C per decade over the same period. These rates are similar to global averages for the same period.

The effects of climate change in Myanmar are already being felt and will increase in the coming decades, challenging a vulnerable population highly centered on climate-dependent livelihoods and ecosystem services. This vulnerability can be expected to increase in the future, as climate models project rising sea levels that would have devastating effects on the coastline, increased temperatures that will challenge agriculture productivity and affect human health through more frequent extreme hot days, and changing monsoon rainfall patterns that will affect agricultural livelihoods nationwide (Horton et al., 2017).

Slagle (2014) investigated that the climate has already been changing in recent decades with higher temperatures and an altered monsoon season interrupting traditional rainfall patterns. A shorter monsoon season will prolong the dry periods in the country, while warmer air and sea surface temperatures will cause increased rainfall intensity in the shorter wet season. Agreement among climate models and the consistency of their results make it clear that Myanmar is vulnerable to climate change.

According to Myanmar Climate Smart Agriculture (MCSA) Strategy (2015), Myanmar has various ecological zones with rice as the main crop. Myanmar also suffers from the adverse effects of climate change such as scarcity of rainfall, irregular rainfall, heat stress, droughts, flooding, sea water intrusion, land degradation, desertification, deforestation, and other natural disasters. The long term effects of climate change will have a huge impact on the agriculture.

Myanmar has experienced regular extreme weather events since 2005. The most significant example was Cyclone Nargis which happened in 2008 and killed more than 138,000 people. In July 2015, the Climate Asia study was conducted in Myanmar, and it can be found that Cyclone Komen brought strong winds and heavy rainfall to the country. This resulted in some of Myanmar's worst floods for decades. Not everyone has heard of climate change, but climate changes affect everyone (Colquhoun et al, 2016).

By the Myanmar Climate Change Adaptation (MCCA, 2017), in Myanmar where climate change has already resulted in increasing severity and frequency of: (i) extreme events such as tropical cyclones, heavy rains and flooding, heat waves, and drought events; and (ii) coastal hazards such as reverse storm surges, among other. The climate system is complex and consists of five major components; the atmosphere, hydrosphere, cryosphere, land surface, biosphere, and interactions between them. Weather refers to a condition of the atmosphere at a definite time and location described by meteorological variables such as temperature, rainfall, wind, humidity, atmospheric pressure, and cloudiness. Climate means the average weather conditions in a specific location at a given time of the year. In the future, the average annual and daily maximum temperatures in Myanmar will not rise. Predicting future rainfall patterns is difficult. Nevertheless, the projection indicates that Myanmar will experience more intense rainfall in the future, particularly during the wet season. This, in turn, could exacerbate wet-season flooding in some regions.

According to Myanmar Climate Report (Aung, Zin & Theingi, 2017), Myanmar is situated in a tropical climate region that is highly vulnerable to the impact of climate change. Therefore, information about climate change in Myanmar is in high demand. Due to climate change, it is observed that a decreasing amount of rainfall occurs during the monsoon period while the maximum temperature increased and the minimum temperature decreased in Myanmar from 1981 to 2010.

United Nations Human Settlement Programme (UN-Habitat, 2020), Myanmar is the 2nd most affected country by climate extreme events in the period of 1999-2018 as per the Global Climate Change Risk Index of 2020, Climate change projections in Myanmar such as higher temperature, rainfall variability, sea level rise etc., indicate increased risk of flood, cyclone, drought etc., which can severely impact the lives and livelihoods of people in Myanmar and hinder the socio-economic development goals. MCCA supported Government of Myanmar to reflect climate change in development agenda of Myanmar and will continue providing support on climate change issues and help Government of Myanmar to achieve the targets of Myanmar Sustainable Development Plan (MSDP) and Sustainable Development Goals (SDG).

2.3 Impact of Climate Change on Rice Production in Myanmar

Win (1991) emphasized that the rice production data reveal three distinct and significant growth trends generated by various forces at different times. The first growth period occurred in 1885-1910, after the final annexation of the country by the British. The second growth period came in 1955-65, a few years after the country gained independence. The third growth period took place in 1975-85, when technology development and transfer systems provided clear dividends. Forces that generated rice production growth differed with the period, creating differential impacts on the population. Rice production, however, cannot be taken as a single aggregate factor influencing the population. The long-term rice production trend is an important socioeconomic indicator. Rice production in Myanmar depends on many factors within and outside the control of the government.

In pre-British days, the Burmese King adopted a restrictive commercial policy prohibiting the export of many products, including rice. This restriction on rice exports discouraged farmers from growing more than that was required for their own consumption. In addition, the low returns on rice offered very little incentive for farmers to produce more than they needed for food, seed, and taxes. The Burmese ruler imposed many duties and restrictions on the merchants who traded in the country. The British colonized the country in three stages. The Arakan and Tenasserim coastal strips were colonized in 1826, Pegu and Martaban in 1852, and the rest of the country in 1885. Rice development was initiated by the British after the second colonization stage by putting the Irrawaddy deltaic area under rice cultivation. This deltaic area offered a favorable rice environment in terms of both weather and soil. With a view to exporting rice to Europe, the British government encouraged increased production in every possible way. Land, labor, and capital are the three main resources necessary for rice production. Rice production rose sharply between 1885 and 1910. The rice-sown area of 1.5 million ha in 1885 increased to 4 million ha in 1910. While the yield remained almost the same, the rapid area expansion increased production from 2 million to 6 million tons. As a result, rice exports also rose from a few hundred thousand to 1.5 million tons. Production stability during the period was striking, in spite of the fact that rice cultivation relied totally on the weather. The colonial government that induced rapid rice production through the expansion of area finally created political instability (Win, 1991).

The second rice production growth period occurred between 1955 and 1965. It started when the country gained independence from the British. The rice-sown area increased from 4 million to 5 million ha, raising production from 6 million to 8 million tons. There was a slight increase in yield due to the use of improved varieties and a small amount of chemical fertilizer. The country experienced poor weather conditions in 1957 and 1961, significantly reducing rice production. After World War II, food shortages in many countries offered good opportunities for the rice export trade. The third period of growth, which occurred between 1975 and 1985, was generated by science. The time was most appropriate for practicing scientific methods. Regarding equity of production, higher rice production during this period benefited both the individual farmer and the country (Win, 1991).

Shrestha, Thin & Deb (2014) analyzed the impacts of climate change on irrigation water requirement (IWR) and yield for rain-fed rice and irrigated paddy at Ngamoeyeik Irrigation Project in Myanmar by using the Statistical DownScaling Model. The analysis shows a decreasing trend of IWR observed for irrigated paddy under the three scenarios indicating that small irrigation schemes are suitable to meet the requirements. An increasing trend in the yield of rain-fed paddy was estimated under climate change demonstrating increased food security in the region.

Weather determines the average rice production in a particular year. Myanmar's economy will continue to rely on the agricultural sector in the foreseeable future. Rice, which occupies a prominent position in that sector and will certainly shape the economic viability, political stability, and social status of the country. Internal rice consumption will rise with the increasing population. Other cereals are not expected to assume a significant share of the dietary staple. The demand for rice in the international market will also rise because of increasing populations in the rice-consuming countries of Asia, Africa, and South America. It is, therefore, imperative that rice production in the country be increased. Rice has been cultivated in Myanmar since prehistoric times. Before World War II, Myanmar became the largest rice exporter in the world. Rice area and production declined during the post-war era. It is grown mainly during the monsoon season as a single crop. In 1992, summer rice was introduced to regions across the country where irrigation facilities were available. In general, the sowing time of monsoon rice and summer rice are from May to October and November to March, respectively. This varies from region depending on geographic and climatic conditions (Myanmar National Study, 2015).

By Myanmar Agriculture at a Glance (2018), Myanmar is an agrarian economy with almost two-third of its estimated 60 million people, approximately 40 million, dependent upon agriculture in rural areas. Since Myanmar is rich in natural resources and diverse in agro-ecological conditions, opportunities for doing business in agriculture are abundant along every segment of supply chain of various agricultural products. Myanmar is a major rice export country and exported about 3 million Tons between 1921 and 1941, in colonial period.

Myint (2018) presented that the trend in cultivated area, yield, and total rice production constantly increased during the Second World War. But the rice exports have not reached the level attained during the British colonial government era. In the early 1940s, Myanmar became the world's largest rice-exporting country. In Myanmar, most of paddy producing areas were in ecological zones, such as the delta, dry zone, coastal zone, and mountainous areas. The delta region is the largest cultivated in both monsoon and summer. It includes the Ayeyawady, Bago and Yangon Regions. Regarding the destination of rice export, it was mainly to China via border trade which was operated by high proportion of about 51.62% of the total trade volume in 2017-2018.

Lar et al. (2018) assessed the effect of climate change on rice yield by using the environmental Policy Integrated Climate model under climate change. The study found that rice yield reduction will be significantly higher under the Representative Concentration Pathway (RCP) 8.5 than under the RCP 4.5 for both rice. Yield reductions are attributed to increases in the mean of maximum and minimum temperatures and variation in rainfall pattern. The model result suggests that changing the sowing date is a good option for compensating for future rice yield reduction. The other adaptations that offset the rice yield response to climate change include providing farming machines, and irrigation facilities, improving infrastructure and in cultivars that resist disease, pests, and drought, better weather forecasts, and extension systems.

According to Agriculture Guide (2019), paddy is sown on 48% of net sown land is the most common crop choice for farmers. However, paddy output fall due to the 2015 floods. Aung (2018) analyzed that rice production is predicted to increase by 4.4 percent in 2017-2018 due to expectations of more favorable weather, high price in centives, and increased use of farm machinery. Myanmar’s rice exports are forecasted to decrease by 10 percent in anticipation of lower old crop supply.

The United Nations Environment (UN-Environment, 2019) announced that rice production in Myanmar faces several challenges, including the rice sector’s vulnerability to climate change impacts like higher temperatures, drought, flooding, and other stresses.

2.4 Background of Ayeyawady Region

Ayeyawady is a region of Myanmar, occupying the delta region of the Ayeyawady River. It is bordered by Bago Region to the north, Bago Region and Yangon Region to the east and the Bay of Bengal to the south and west. It is contiguous with the Rakhine State in the northwest. The region lies between approximately latitude 15° 40' and 18° 30' north and between longitude 94˚ 15' and 96˚ 15' east. The population is more than 6.5 million, making it the most populous of states and regions. According to the 2014 population census, the total population of the Ayeyawady Region was 6,184,829. Ayeyawady Region is flanked by the Rakhine Yoma (Arakan Mountains) range in the west. Large areas have been cleared for paddy cultivation, leading to its preeminent position as the main rice producer in the country, a position it has retained into the 21st century. It has also a number of lakes. Of the rivers branching out from the mighty Ayeyawady, Ngawun, Pathein and Toe are famous. The capital city of Ayeyarwady division is Pathein. Chaungtha Beach and Ngwesaung Beach are popular resorts for both foreigners and Myanmar citizens. They are in the west of the Ayeyarwady Region, an hour from Pathein city and four hours from Yangon city by road (Ayeyawady Region, 2022).

Ayeyawady Region is heavily forested and wood products are an important component of its economy. The principal crop of the Ayeyawady Region is rice, and the division is called the “granary of Myanmar.” In addition to rice, other crops include maize, sesame, groundnut, sunflower, beans, pulses, and jute. Despite the importance of agriculture to the region, landlessness is high in rural households. Most farms are small; nearly half are under 5 acres. Paddy agriculture is dominant during the monsoon but irrigation is limited, especially in smaller farms, during the dry season. Seeds are sourced from own reserves rather than from specialized traders. Yields from farms, an average of 3.3 tons per hectare, are lower than other Asian countries (Ayeyawady Region, 2022).

Located at an elevation of 5.13 meters (16.83 feet) above sea level, Ayeyawady has a Tropical monsoon climate. The city’s yearly temperature is 29.31ºC (84.76ºF) and it is 2.29% higher than Myanmar’s averages. Ayeyawady typically receives about 72.89 millimeters (2.87 inches) of precipitation and has 115.72 rainy days (31.7% of the time) annually. Ayeyawady Region was the site of heavy devastation when Cyclone Nargis made landfall in early May 2008. The cyclone made landfall on the town of Wagon near Haigyi Island. Labutta Township was most heavily struck with around 80,000 deaths. The cyclone's path devasted the low-lying delta regions going through south-central Ayeyawady Region and Bogale before entering neighbouring Yangon Region. Ayeyawady Region consists of six districts:

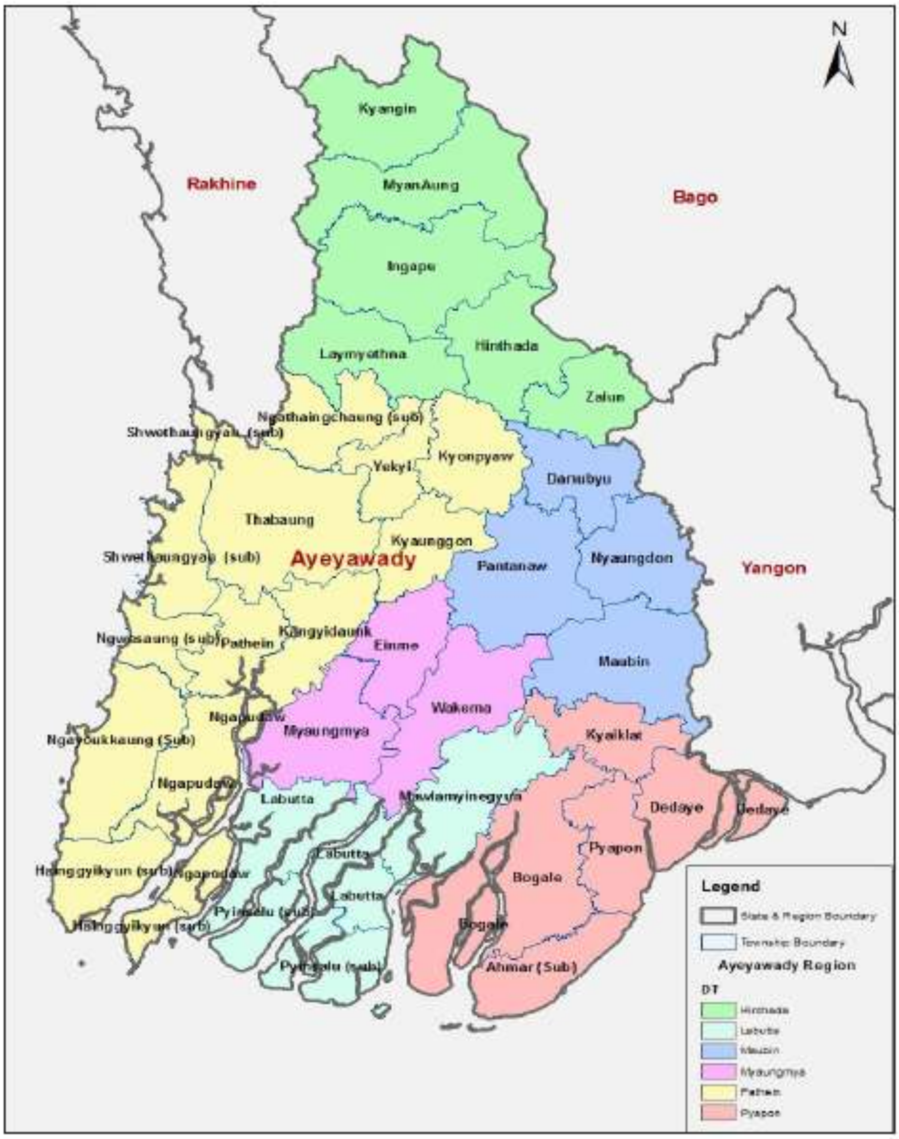


Figure (2.2) Map of Ayeyawady Region by Districts and Townships

Source: Ayeyawady Region Census Report (2015)

Pathein District is a district of Ayeyawady Region, Myanmar. It is located around and includes the urban area of the city of Pathein. The area of the Pathein District is 10,899.8 km^2 (4,208.4 sq mi). Its population was 1,630,716 in 2014.

Hinthada District (formerly **Henzada District**) is a district of Ayeyawady Region, Myanmar. By 2010 it had grown to 170,312. The trade of locally grown rice and grain goes through the port of Hinthada.

Myaungmya is a district of the Ayeyawady Region in south western Myanmar. It consists of 5 cities. Myaungmya is the chief town of the district. Myaungmya district was formed in 1893 out of a portion of Bassein district, and reconstituted until 1903. It has an area of 2,663 square miles (6,900 km^2) and a population of 280,000 and density of 104 inhabitants to the square mile. The district is a deltaic tract, bordering south on the sea and traversed by many tidal creeks. Rice cultivation and fishing occupy practically all the inhabitants of the district. The district contains three townships: Myaungmya Township, Einme Township, and Wakema Township. In the Townships, there are 50 wards, 489 village tracts and 2557 villages.

Maubin District is a district in Ayeyawady Region, Myanmar. It consists of 39 wards, 235 village tracts and 1642 villages. The district lies on a flat plain, cut by many streams, which is 1,362 feet above sea level. It has an area of 1651.49 square miles (1,056,952 acres). The majority of the population are Burmese and Kayin nationals.

Phyarpon District is a district of the Ayeyawady Region in south western Myanmar. It consists 4 cities. They are Pyapon, Bogalay, Kyaiklat and Dedaye.

Labutta District or **Latputta District** is a district in Ayeyawady Region, Myanmar. Labutta District was established in 2008 after the region was hit by Cyclone Nargis in May 2008. The administrative seat is the town of Labutta (Ayeyawady Region, 2022).

2.5 Studies Related to the Impact of Climate Change on Crop Production

Drummond et al., (2003) evaluated and compared the predictive accuracies of various approximation techniques, including feed forward neural networks (NN), projection pursuit regression (PPR), and stepwise multiple linear regression (SMLR), in relating crop yields to topography and soil parameters. Yield estimation within individual site-years was carried out through the use of a 5-fold cross validation technique. SMLR, PPR and NN methods were investigated on each of ten individual site-year data sets and multiple site-year data sets of climatological variables including temperatures and rainfall. This study found that NN methods produce the minimal SEP (Standard Error of Prediction) among all the methods used.

Mall et al. (2007) mentioned that crop simulation models have been extensively used to study the impact of climate change on agricultural production and food security. The output provided by the simulation models can be used to make appropriate crop management decisions and to provide farmers and others with alternative options for their farming system. It is expected that in the coming decades with the increased use of computers, the use of simulation models by farmers and professionals, as well as policy and decision makers, will increase. In India, substantial work has been done in last decade aimed to understanding the nature and magnitude of change in yield of different crops due to projected climate change. It found out an overview of the stage of knowledge of the possible effects of the climate variability and change on food grain production in India.

Joshi, Maharjan & Piya (2011) analyzed the impact of current climate trend on six main food crops (rice, wheat, maize, millet, barley and potato) in Nepal. These food crops are divided into two groups based on their growing seasons, namely, summer and winter season crops. The impact for each crop based on the growing season of respective crop is assessed, and the effects of observed climate variables on yield of these major food crops are analyzed by using the regression model. The results showed that climate variables like temperature and precipitation are the important determinants of crop yields. Moreover, climate variables showed some influences on the yield of these major food crops. The increase in summer rain and maximum temperature has positively contributed to rice yield.

Awal & Siddique (2011) carried out to estimate growth pattern of rice production through choosing the best ARIMA model. The study examined the efficiency of those models in forecasting the rice production in Bangladesh and revealed that the best models were ARIMA (4,1,4), ARIMA (2,1,1) and ARIMA (2,2,3). Moreover, it is further indicated that short-term forecasts were more efficient for ARIMA models compared to the deterministic models.

Laxmi & Kumar (2011) applied that the NNs for crop yields forecasting using Multi-Layer Perceptron (MLP) architecture with different learning algorithm and considered that yields of crop at district as output variables and indices of weather variables (Maximum temperature, Minimum temperature, Rainfall and Morning Relative Humidity) as input variables. The study found that MLP performed better than linear regression.

Stastny, Koneey & Trenz (2011) contributed the implementation of the multi-layer neural network for the prediction of crop yield, and the comparison of the accuracy of this approach with the accuracy of the well-known regression model designed for the prediction of empirical data. The study showed that the use of a multi-layer neural network has proven to be more accurate in the case of the given task than the previously published regression model.

Parekh & Suryanarayana (2012) carried out the study to determine the predominance of various meteorological data on yield of wheat, using Neural fitting tool of ANN. It can be evidently concluded from the study that yield of a crop is very much depended on maximum and minimum temperatures and relative humidity.

Socio-economic factors were also taken into account in Ghodsi et al. (2012), where rainfall, guaranteed purchasing price, area under cultivation, subsidy, insured area, inventory, import, population and value-added agricultural production were used as predictors for wheat production. The comparison of real wheat production with ANN output in the last five years of the study showed that the proposed ANN model is a suitable way of predicting wheat production.

Sharma (2012) compared the performance of three models (Linear Regression (LR), Autoregressive Integrated Moving Average (ARIMA), and Artificial Neural Network (ANN) to forecast weather parameters. A comparative study of the existing and proposed weather forecasting models was performed to identify the precise and reliable weather forecasting models. The study divulges that Hybrid MLR-ANN model is an appropriate forecasting tool to estimate the weather parameters, in contrast to the MLR, ARIMA, ANN and hybrid MLR_ARIMA models.

Wenjiao, Fulu & Zhao (2013) reviewed the progress of identifying contributions of climate change to crop yields based on statistical models and improved the theory of the effects of climate change on agriculture. The study observed that the correlation exists between extreme temperatures and mean temperature.

Somah (2013) suggested that critical impact asymmetries due to climatic factors that affected subsistence crops in the Sudano-sahel of Cameroon. Furthermore, the study results indicated incidences of droughts; with the Multilinear Regression (MLR) models showing temperature and rainfall to an extent determined agricultural crop productivity in the Study area. However, other factors such as population growth have undoubtedly caused enormous impacts on the agricultural system as seen in remote sensing analyses.

Matsumura et al. (2014) forecasted the maize yield with climate conditions and fertilizer as predictors by using Multiple Linear Regression (MLR) and non-linear Artificial Neural Network (ANN) models. The ANN model behaves very differently from the MLR model during extrapolation. The results showed that the ANN extrapolates more gently than a linear function.

Lamba and Dhaka (2014) represented the forecasting techniques in the field of the wheat crop. The study shows all the past research development of forecasting in all areas. In the field of agricultural yield, the major forecasting models are statistical, metrological, simulation, agronomic, remote satellite sensed, synthetic and mathematical. Moreover, it shows a compact combination of all these models and shows why Neural Network Model is important compared to other models for nonlinear data behavior system like wheat crop yield prediction. Then the study presented a compact combination of all major forecasting models and showed why Neural Network Model is important compared to other models for nonlinear data behavior systems like wheat yield prediction.

Khairunniza-Bejo, Muataffha & Wan Ismail (2014) showed that the application of plays a crucial role in the future evaluation of the concept of precision agriculture as a sustainable means of meeting world's food demands. However, further research associated with the ANN impacts on crop yield production must be conducted to ensure sustainability of future food needs. It has been also shown that ANN provides a better interpretation of crop variability compared to the other methods.

Ranjeet and Armstrong (2014) examined the application of artificial neural networks (ANNs) for predicting crop yields for an agricultural region in Nepal. The neural network algorithm has become an effective data mining tool and the outcome produced by this algorithm is considered to be less error prone than other computer science techniques. The experiment shows that the trained neural network produced a minimum error which indicated that the test model is capable of predicting crops yield in Nepal.

Chowdhury & Khan (2015) undertook to examine the potential impact of climate change on the yield of three different rice crops. A multiple regression analysis using the OLS method was performed to assess the climate-crop yield interrelations on the basis of country level time series data. This study found that all the climate variables (maximum temperature, minimum temperature, rainfall, and humidity) had a significant impact on rice yield over the period under study, but these effects varied among the three different rice yields. It also suggested that sustainable agricultural development may play a vital role in mitigating adverse climate change effects.

Farook & Kannan (2015) examined the relationship between the yield of rice crop and three main climate variables (maximum temperature, minimum temperature, rainfall) based on Vector AutoRegression (VAR) model. The results showed that average maximum temperature and total rainfall have negative effects on yield, whereas average minimum temperature affect yields positively.

Aboukarima, Elsoury & Menyawi (2015) investigated an ANN model in the prediction of cotton leaf area. ANN model performance was tested successfully to describe the relationship between measured and predicted cotton leaf area. The developed ANN model produced a satisfactory correlation between measured and predicted values and a minimum inspection error. According to the study, the neural network approach is promising for rapid collection of cotton leaf area information in an effective manner without cost.

Hilal et al. (2016) attempted to optimize the oil palm yield amount by studying parameters of land quality and climate, determines which of them is distinctly effective on oil palm yield amount, develops ANN model and simulation of Oil Palm production by using MATLAB software and Design Expert software, conducted an experiment to determine the effect of the number of neurons and the number of hidden layers in the network ANN is used. The results of the simulation showed that the average accuracy percentage of the simulation was 0.9867% and the MSE was 0.0513%. The climatic changes that influenced the simulation are very high, where the relative humidity recorded at a proportion of impact of up to 100%, while the recorded rainy days, which is ranked as the second influential factor, were almost 90% and the effect of temperature was up to 70%. The influence of several climatic changes that decrease the quantity of rainfall, Rainy days, Temperature rise, Evaporation and increasing Humidity, reduces the productivity of oil palm plantations for 2.35 tons/ha/year. This research concludes that ANN can be used effectively to predict the production of palm oil based on the quality of land and local climate.

Mathieu & Aires (2016) used mixed-effect (ME) models, linear models, and neural network models to forecast corn production over the United States. Weather-sensitivity assessments and seasonal-forecast applications are discussed. The results show that, for the particular application, state spatial scale is a good trade-off: it allows specializing to local conditions while keeping enough data to calibrate the linear model or the neural-network model. Even if in theory the nonlinearity of the neural-network model allows it to specialize to local conditions, the group information used in ME models is more direct information than what could be inferred by the NN on the basis of weather-input data alone. The ME model with county classification can predict county corn-yield anomalies with a 50% correlation between forecast and observed values. In more weather-sensitive regions, this correlation rises to 60%: this result means that 40% of the variance can be explained by monthly weather information.

Aryal et al. (2016) employed a multivariate regression analysis to estimate the empirical relationships between crop yield and climate change variables (rainfall, maximum summer temperature and minimum summer temperature) and described climate change is threatening the agriculture sector, especially present and future food security in low income countries. The results showed that climate change variables are significant effects on crop yield.

Syeda (2017) worked out to qualify the long-term effect of climate change on wheat production in Dinajpur District using multiple regression analysis techniques, taking several climate variables into account. The approximately significant effects were found for the climatic variables of average minimum temperature, average dry temperature, and total rainfall on wheat production.

Sitienei, Juma & Opere (2017) developed a multiple linear model to predict tea yield using climate variables and found that the climatic variables during some months in both the concurrent year and the previous year were positively correlated with the tea yield. However, there was an inverse relationship between maximum temperature and rainfall.

Wiah & Twumasi-Ankrah (2017) analyzed the impact of climate change on the yield of cocoa in Ghana using the VAR model. It is indicated that the direction of causality is from maximum temperature, minimum temperature and precipitation to yield. However, there was no causation from the number of rainy days to yield. Maximum temperature, numbers of rainy days have negative effects on yield, whereas minimum temperature and precipitation affect yield positively.

Abdullahi & Elkiran (2017) explored the potentiality of black box modeling of ANN in determining the nonlinear correlation between evapotranspiration (ET) and climate variables in the study region. The results revealed that ANN can predict the climate impact on ET magnificently in the study area using limited parameters, such as minimum and maximum temperature, but the prediction precision was higher when more climate parameters were supplied. The developed ANN model produced significant correlation between measured and predicted values with minimum inspection error.

Inconsistency in climate regimes of rainfall and temperature is a source of biotic and abiotic stresses in agricultural systems worldwide. This variability is a cause of poor yield potential and crop failure. Rainfall has shifted with an increasing trend during monsoon and almost static during other seasons. Climate change is unswervingly influencing the human survival through its agricultural impacts by higher temperatures, droughts, floods, soil erosion and rainfall variations affecting the food security of the globe (Rahman et al., 2017).

Hossain, Al-Amin & Islam (2018) highlighted univariate seasonal autoregressive integrated moving average (SARIMA) and multivariate vector autoregressive (VAR) for modeling and comparing the forecasting abilities of the climatic data. The researchers revealed that VAR (9) model gives a more appropriate forecast than the univariate SARIMA model and the selected climate variables are important in future because the study explained the future variation. The monthly forecasts using the appropriate model reveal that maximum temperature, minimum temperature, and humidity are slightly increasing while the cloud coverage is decreasing minimally.

Mathieu & Aires (2018) presented a statistical method for forecasting extreme corn yield losses caused by weather extremes. It is shown that the Neural Networks (NNs) is able to find and exploit the simultaneous combination of high heat and low moisture that is devastating for the crop yields. Park, Das & Park (2018) developed ANN based localized models to estimate rice production within South Korea and found that increasing temperature at a higher rate may have a positive effect on future rice production in this study area.

Raj, Ramesh & Rajkumar (2019) measured the relations between meteorological factors and crop yield variability by using statistical models (step-wise multiple regression (SMLR), seasonal autoregressive integrated moving average (SARIMAX), artificial neural network (ANN) and vector autoregressive (VAR) model. It is shows that the multivariate time series models are better suited for capturing the non-linear short-term fluctuations and long-term variations.

Climate variability is one of the factors that directly and greatly affect cropping system and plant yield. It is therefore very important to obtain a good understanding about climate variability, or changes in the climate and the effect of these changes to clearly understand the vulnerability of food crops as well as its agronomic impacts to create and implement adaptive strategies to mitigate its negative effects. The study showed overall decreasing trends for both minimum temperature and relative humidity and increasing trends for rainfall and annual mean temperature. There are significant correlations between rice yield the all the climate variables in both irrigated and rainfed farming ecosystem types. For irrigated ecotype, rainfall and relative humidity have negative correlation to yield while both positive correlation for mean temperature and minimum temperature. On the other, for rainfed ecotype, rainfall, minimum temperature, and relative humidity have positive correlation to yield while negative correlation for mean temperature alone (Enovejas et al., 2021).

Moreover, the temperature rise is one of issues of climate change that has the profound effect on rice production. In addition, it would make the age of rice the shorter and decrease the rice yields. The early rice production was significantly influenced by high temperatures, with a yield decrease of 8.1% per 1 °C increase of rice-growing season temperatures. This is mainly because daily maximum temperatures more frequently exceed the threshold for the reproductive growth stages of early rice. Late-rice yields were, on the other hand, only slightly affected by increasing temperatures. Then the negative influence of high temperature on rice production will likely be more serious, making rice production more vulnerable (Song, Chunyi & Wang, 2022).

A number of the international researchers have been analyzed the effect of climate change on crop production by using the MLR, SARIMAX or ARIMAX, VAR and ANN models. However, there are no previous studies about the impact of climate change on rice production using these models in Myanmar. Hence, in this study, MLR, SARIMAX or ARIMAX, VAR and ANN models were used to fit the rice yield, and a comparative analysis was done to identify the most appropriate model to forecast the rice yield by using the time series data.

2.6 Measurable Variables for Climate Change

In the long run, the climatic changes could affect agriculture in several ways, such as quantity and quality of crops in terms of productivity, growth rates, photosynthesis and transpiration rates, moisture availability, etc. Among the crops, rice is a sensitive crop that depends highly on weather conditions. If water is not the limiting factor, the most important weather parameters are temperature and solar radiation. Rainfall is the direct critical weather parameter in rainfed ecologies. Most of the world's rice is grown in the tropics, and the critical determinant for its growth is temperature. Relative humidity is a function of temperature, and moisture in the atmosphere is invariably much higher in the morning than in the afternoon. Rice that is cultivated in standing water builds up an environment with high relative humidity (Ray, 2016).

Temperature and rainfall have been known as the key determinant factors that affect rice production, and it is found that non-climate factors such as fertilizers, water, cultivars, and soil fertility cause 40% variation to rice yields, whereas the remaining 60% can be influenced by climate variability (Pheakdey, Xuan & Khanh, 2017).

The current global climate change is considered difficult to control. The major contributions of climate change, such as temperature, water pressure, humidity, and rainfall, are continuously encountered in all countries. Climate change has a direct and significant impact on agricultural production, especially paddy. A number of previous studies highlighted the impact of climate change on the production of rice. The climate factors that have the most significant influence on the production of paddy are temperature and humidity (Faradiba, 2020). According to Molla et al., (2020), the rainfall variable explained 69% of the rice yield variability.

2.7 Conceptual Framework

Based on the previous studies, the conceptual framework of the present study can be depicted in Figure (2.3). As shown in this figure, climate change is represented by five variables such as Maximum Temperature, Minimum Temperature, Rainfall, and Relative Humidity (Morning & Evening), which are used as the predictor variables while rice yield (Monsoon and Summer) is the research's response variable. Accordingly, this study is intended to analyze how the fluctuation of rice yield will be affected by those predictor variables.

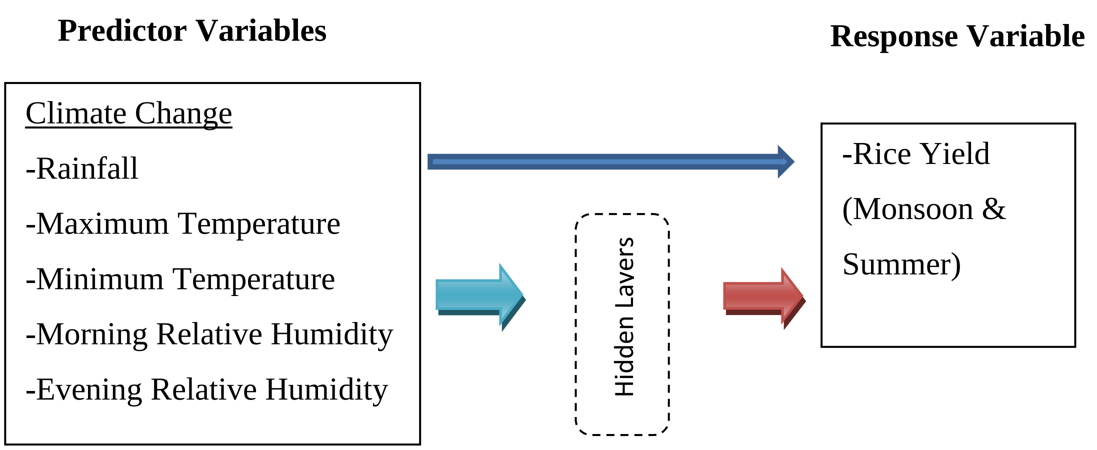


Figure: (2.3) Conceptual Framework

Source: Own Compilation

The Hidden Layers, expressed in conceptual frame work were technically used in computing ANN model only. The definitions of the response and predictor variables are described as follows.

Rice Yield

According to the Department of Agriculture in Myanmar, rice production was measured by the standard units such as average rice yield per acre (Bsk/Ac) in this study.

Rainfall

Rainfall is a measurement of how much water falls as rain in a certain period of time. It is measured by millimeter as unit.

Maximum Temperature

It is the highest temperature recorded during a specified period of time. The most common reference is to the daily maximum temperature, or "high". In this study, daytime temperature is usually measured during 12:00 to 14:00, for identifying the hottest temperature in a day.

Minimum Temperature

It is the lowest temperature attained during a specified period. In this study, nighttime temperature is usually measured during 18:30 to 06:30 in the morning to determine the coolest temperature in a day.

Morning and Evening Relative Humidity

Relative humidity is the amount of water vapors present in air expressed as a percentage of the amount needed for saturation at the same temperature. It is measured two times in a day, in the morning (at 9:30 AM) and evening (at 6:30 PM) respectively.

CHAPTER III
METHODOLOGY

The methodology of statistical modelling applied to forecast the yield of rice based on measurable variables of climate change were explained in this chapter. These include Multiple Linear Regression (MLR), Seasonal Autoregressive Integrated Moving Average with predictors (SARIMAX), Vector Auto-Regressive (VAR) and Artificial Neural Network (ANN).

3.1 Multiple Linear Regression (MLR)

Regression analysis is a statistical technique for modelling and investigating the relationships between an outcome or response variable and one or more predictor or regressor. A regression analysis is often to generate a model that can be used to forecast or predict future values of the response variable, given specified values of the predictor variables. It is widely used in crop yield prediction.

Multiple linear regression (MLR) involves two or more independent (predictor) variables. The linear regression model such as the dependent (response) variable Y_t is expressed as a linear function of the k predictor variables is

$$Y_t = \beta_0 + \sum_{i=1}^{k} \beta_i X_{it} + e_t \quad , t = 1, 2, \dots, n \tag{3.1}$$

where β_i are the regression coefficients determined by fitting the straight-line to the data, i.e. minimizing the mean squared error (MSE) between the modeled estimate and the observed Y_t value and e_t is the stochastic error term which is distributed as normal with mean zero and these errors are not autocorrelated.

The unknown parameters $\beta_0, \beta_1, \dots, \beta_k$ in a regression model are typically estimated using the method of least squares. This is an important application of regression models in forecasting. Thus the least squares estimator of $\widehat{\boldsymbol{\beta}}$ by matrix notation is

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{Y} \tag{3.2}$$

where $\boldsymbol{Y} = (y_1, y_{2,\dots,} y_n)'$,

$\widehat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\beta}_{2_{,\dots,}} \hat{\beta}_k)'$ and

$\boldsymbol{X} = [x_{ti}]$ with $x_{t0} = 1$ for all t.

The **X** matrix has n rows reflecting the number of observations and k+1 columns reflecting the intercept which is represented by the column of ones plus the number of predictors. The fitted values of the response variable from the regression model are computed from

$$\widehat{\boldsymbol{Y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}} \tag{3.3}$$

The difference between the actual observation y_t and the corresponding fitted values is the residual $e_t = y_t - \hat{y}_t, t = 1,2, \dots, n$. The n residuals can be written as an (n×1) vector denoted by

$$\boldsymbol{e} = \boldsymbol{Y} - \widehat{\boldsymbol{Y}} \quad , \quad SSE = (\boldsymbol{Y} - \widehat{\boldsymbol{Y}})'(\boldsymbol{Y} - \widehat{\boldsymbol{Y}}) \tag{3.4}$$

In addition, it is also necessary to estimate the variance of the model errors. The residual variance is estimated using

$$\hat{\sigma}^2 = \frac{SSE}{n - k - 1} \ . \tag{3.5}$$

3.1.1 Assumptions of Multiple Linear Regression Model

Assumptions of the regression analysis are required to provide valid results. If a regression analysis fails to meet the assumptions, regression analysis can provide invalid results. These assumptions are:

Linearity: Multiple linear regression requires the relationship between the independent and dependent variables to be linear.

Normality: Multiple linear regression analysis requires that the errors between observed and predicted values should be normally distributed.

Multicollinearity: Multiple linear regression assumes that there is no multicollinearity in the predictor variables. Collinearity is a data issue that arises if two independent variables are highly correlated.

Homoscedasticity: There should be **homoscedasticity** or equal variance in errors.

Autocorrelation: One of the critical assumptions of multiple linear regression is that there should be no autocorrelation in errors. When the residuals are dependent on each other, there is autocorrelation.

3.1.2 Tests for Significance and Fitness of Model

The test for significance of regression is a test to determine whether there is a linear relationship between the response variable y and a subset of the predictor variables. The appropriate hypotheses are

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$$
$$H_1: at\ least\ one\ \beta_i \neq 0.$$

Rejection of the null hypothesis H_0 implies that at least one of the predictor variables contributes significantly to the model. The test procedure involves an analysis of variance partitioning of the total sum of squares (SST),

$$SST = \sum_{t=1}^{n} (Y_t - \bar{Y})^2 = \boldsymbol{Y'Y} - n\bar{Y}^2 \tag{3.6}$$

into a sum of squares due to the model (SSR) and a sum of squares due to residual (SSE),

$$SST = SSR + SSE \tag{3.7}$$

$$SST = \left(\boldsymbol{\hat{\beta}'X'Y} - n\bar{Y}^2\right) + (\boldsymbol{Y'Y} - \boldsymbol{\hat{\beta}'X'Y}) \tag{3.8}$$

Now if the null hypothesis is true and the model errors are normally and independently distributed with constant variance as assumed then the test statistic for significance of regression is

$$F = \frac{SSR/k}{SSE/(n-k-1)} \tag{3.9}$$

and one rejects H_0 if the test statistic F exceeds the upper tail point of the F distribution with $F_{\alpha,k,n-k-1}$. Alternately, the P-value approach could be used to hypothesis testing and thus reject the null hypothesis if the P-value for the statistic F is less than α. Alpha is also known as the level of significance. This represents the probability of obtaining the results due to chance. The chance that reject the null hypothesis when in reality, it should fail to reject the null hypothesis.

Then, hypotheses tests on the individual regression coefficients would be useful in determining the value or contribution of each predictor variable in the regression model. The hypotheses for testing the significance of any individual regression coefficient, β_i, are

$$H_0: \beta_i = 0$$
$$H_1: \beta_i \neq 0$$

If the null hypothesis $H_0 : \beta_i = 0$ is not rejected, then this indicates that the predictor variable can be deleted from the model. The test statistic for this hypothesis is

$$t = \frac{\hat{\beta}_i - 0}{Se(\hat{\beta}_i)}, \; Se(\hat{\beta}_i) = \sqrt{\hat{\sigma}^2 C_{ii}}. \tag{3.10}$$

where C_{ii} is the diagonal element of the $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ matrix corresponding to the regression coefficient $\hat{\beta}_i$. This t- test measures the contribution of a variable while the remaining variables are included in the model. Moreover, the coefficient of determination is a measure of usefulness of regression model which is denoted R^2. It is computed by

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \tag{3.11}$$

R-squared is always between 0 and 100%. If percentage of R-squared is 0% indicates that the model explains none of the variability of the response data around its mean. If percentage of R-squared is 100% indicates that the model explains all the variability of the response data around its mean. However, a large value of R^2 does not necessarily imply that the regression model is a good one. Adding a variable to the model will never cause a decrease in R^2, even in situations where the additional variable is not statistically significant.

In almost cases, when a variable is added to the regression model R^2 increases. If too many predictors putting in the model, a measure of model adequacy often results in overfitting. In general, the adjusted R^2 will not always increase as variables are added to the model. If unnecessary predictors are added, the value of the adjusted R^2 will often decrease. Consequently, models with a large value of the adjusted R^2 are usually considered good regression model. And then, the adjusted R^2,

$$R^2_{adj} = 1 - \left[\frac{(1 - R^2)(n - 1)}{n - k - 1}\right] , \tag{3.12}$$

tells that the percentage of variation explained by only the predictor variables that actually affect the response variable (Montgomery, Jennings & Kulahci, 2015).

3.2 Autoregressive Integrated Moving Average (ARIMA)

In statistical modeling, a major assumption that often provides relief in modeling efforts is the linearity assumption. A linear filter is a linear operation from one time series x_t to another time series y_t,

$$Y_t = L(x_t) = \sum_{i=-\infty}^{+\infty} \Psi_i X_{t-1} \quad , t = \cdots, -1,0,1, \ldots. \tag{3.13}$$

In that regard the linear filter can be seen as a "process" that converts the input, X_t , into an output, y_t , and that conversion is not instantaneous but involves all (present, past, and future) values of the input in the form of a summation with different "weight", $\{\Psi_i\}$, on each X_t. Furthermore, this linear filter have the following properties:

1. **Time-invariant** as the coefficients $\{\Psi_i\}$ do not depend on time.
2. **Physically realizable** if $\{\Psi_i\} = 0$ for i $<$ 0; that is, the output y_t is a linear function of the current and past values of the input: $y_t = \Psi_0 x_t + \Psi_1 x_{t-1} + \cdots$.
3. **Stable** if $\sum_{i=-\infty}^{+\infty} |\Psi_i| < \infty$.

In linear filter, under certain conditions, some properties such as stationarity of the input time series are also reflected in the output. The stationarity of a time series is related to its statistical properties in time. It exhibits similar "statistical behavior" in time and this is often characterized as a constant probability distribution in time. However, it is usually satisfactory to consider the first two moments of the time series and define stationarity as follows:

(1) The expected value of the time series does not depend on time and

(2) The autocovariance function defined as $Cov(y_t, y_{t+k})$ for any lag k is only a function of k and not time; that is, $\gamma_y(k) = Cov(y_t, y_{t+k})$.

Crudely, the stationarity of a time series can be determined by taking arbitrary "snapshots" of the process at different points in time and observing the general behavior of the time series. If it exhibits "similar" behavior, one can proceed with the modeling efforts under the assumption of stationarity. Further preliminary tests also involve observing the behavior of the autocorrelation function. A strong and slowly dying ACF will also suggest deviations from stationarity. Better and more methodological tests of stationarity also exist.

3.2.1 Stationary Time Series

For a time-invariant and stable linear filter and a stationary input time series x_t with $\mu_x = E(x_t)$ and $\gamma_x(k) = Cov(x_t, x_{t+k})$, the output time series y_t is also a stationary time series with

$$E(y_t) = \mu_y = \sum_{-\infty}^{+\infty} \Psi_i \mu_x \tag{3.14}$$

And

$$Cov(y_t, y_{t+k}) = \gamma_y(k) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \Psi_i \Psi_j \Upsilon_x(i - j + k) \tag{3.15}$$

It is then easy to show the following stable linear process with white noise time series, ε_t, is also stationary:

$$y_t = \mu + \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i} \tag{3.16}$$

with $E(\varepsilon_t)=0$, and

$$\gamma_\varepsilon(h) = \begin{cases} \sigma^2 & if\ h = 0 \\ 0 & if\ h \neq 0 \end{cases} \tag{3.17}$$

So for the autocovariance function of y_t is

$$\gamma_y(k) = \sigma^2 \sum_{i=0}^{\infty} \Psi_i \Psi_{i+k} \tag{3.18}$$

The linear process of Y_t can be rewritten in terms of the backshift operator, B, as

$$Y_t = \mu + \left(\sum_{i=0}^{\infty} \Psi_i B^i \right) \varepsilon_t \tag{3.19}$$

where $\left(\sum_{i=0}^{\infty} \Psi_i B^i\right) = \Psi(B)$, $\{\Psi_i\}$ satisfy $\sum_{i=0}^{\infty} {\Psi_i}^2 < \infty$. A more intuitive interpretation of this theorem is that a stationary time series can be seen as the weighted sum of the present and past random “disturbances”. And the correlation between Y_t and Y_{t+k} as

$$\rho_k = \frac{Cov(y_t, y_{t+k})}{\sqrt{Var(y_t)}\sqrt{Var(y_{t+k})}} = \frac{\gamma_k}{\gamma_0} \tag{3.20}$$

where $Var(y_t)=Var(y_{t+k})= \gamma_0$. As function of k, γ_k is called the autocovariance function and ρ_k is called the autocorrelation function (ACF) because these functions represent the covariance and autocorrelation between y_t and y_{t+k} from the same process, separated only by k time lags. For a stationary process, γ_k and ρ_k have the following properties:

1. $\gamma_0 = Var(y_t)$; $\rho_0 = 1$.
2. $|\gamma_k| \leq \gamma_0$; $|\rho_k| \leq 1$.
3. $\gamma_k = \gamma_{-k}$ and $\rho_k = \rho_{-k}$ for all k, γ_k and ρ_k are even functions and hence symmetric about the lag k=0.
4. Another important property of the γ_k and ρ_k is that these functions are positive semidefinite in the sense that

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \gamma_{|t_i - t_j|} \geq 0 \tag{3.21}$$

and

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \rho_{|t_i - t_j|} \geq 0 \tag{3.22}$$

for any set of time points $t_1, t_2, \ldots, t_n$ and real numbers $\alpha_{1,} \alpha_2, \ldots, \alpha_n$. The ACF of an pth order autoregressive AR(p) process satisfies the Yule-Walker equations

$$\rho_k = \sum_{i=1}^{p} \phi_i \rho(k-i); k = 1,2, \ldots \tag{3.23}$$

It is implying that ACF for an AR(p) model can be a mixture of exponential decay and damped sinusoid expressions.

The Partial Autocorrelation Function (PACF) between y_t and y_{t+k} is the autocorrelation between y_t and y_{t-k} after adjusting for y_{t-1}, y_{t-2}, …, y_{t-k+1}. For any given k, the PACF of the process at lag k, ϕ_{kk} *or* P_k is

$$\mathrm{P}_k = \frac{Cov[(y_t - \hat{y}_t)(y_{t+k} - \hat{y}_{t+k})]}{\sqrt{Var(y_t - \hat{y}_t)}\sqrt{Var(y_{t+k} - \hat{y}_{t+k})}} \quad . \tag{3.24}$$

The sample estimate of ϕ_{kk}, $\hat{\phi}_{kk}$, is obtained by using the sample ACF, $\gamma_{(k)}$. In a number of observations (n) from an AR(p) process, $\hat{\phi}_{kk}$ for k>p is normally distributed.

$$E(\hat{\phi}_{kk}) \approx 0 \quad \text{and} \quad Var(\hat{\phi}_{kk}) \approx \frac{1}{n}. \tag{3.25}$$

3.2.2 Nonstationary Time Series

It is often the case that while the processes may not have a constant level, these process exhibit homogeneous behavior over time. The homogenous nonstationary time series can be reduced to a stationary time series by taking a proper degree of differencing. Firstly, non-seasonal ARIMA model is considered in this section.

The autoregressive integrated moving average (ARIMA) model is a stochastic model. The equally spaced univariate time series data are analyzed by using the Autoregressive Integrated Moving Average (ARIMA). This model predicts a value in a response time series as a linear combination of its own past values, past errors (shocks or innovation), and current and past values of their time series. The ARIMA models are often called as Box-Jenkins models.

In general, an ARIMA model is designed by the notation ARIMA (p,d,q) where p,d and q denote orders of autoregression (AR), I (d) integration (differencing) and moving average (MA) lags respectively. The ARIMA model is written as;

$$\emptyset_p(B)(1-B)^d Y_t = \theta_0 + \theta_q(B)a_t \tag{3.26}$$

where the stationary operator $\emptyset_p(B) = (1 - \emptyset_1 B - \cdots - \emptyset_p B^p)$

the invertible operator $\theta_q(B) = (1 - \theta_1 B - \cdots - \theta_q B^q)$

the response series $W_t = (1-B)^d Y_t$

t = indexes time, B = the backshift operator

d = the required number of differences to become a series stationary

a_t = the independence disturbance (random error term)

when d = 0, the original process is stationary and θ_0 is related to the mean of the process, i.e. $\theta_0 = \mu(1 - \emptyset_1 - \cdots - \emptyset_p)$.When d ≥ 1, θ_0 is called the deterministic trend term. Box and Jenkins introduced the ARIMA model consists of three iterative steps: identification, estimation and diagnostic checking.

3.2.3 Model Identification

The foremost step in the process of modeling is to check for the stationary nature of the series, as the estimation procedures are available only for stationary series. If the underlying generating process is based on a constant mean and variance with its autocorrelation function (ACF) essentially through time, this time series is said to be stationary. A statistical test for stationary is the most widely used Dickey-Fuller test. Another way of checking for stationary is to look at the graph of the data. The structure of autocorrelation and partial correlation coefficients may provide clues about the presence of stationary.

If the model is found to be non-stationary, stationarity could be achieved mostly by differencing the series. Stationary variance could be achieved by some models of transformation, say, the log transformation. This is applicable for both seasonal and non-seasonal stationarity. The next step in the identification process is to find the initial values for the orders of parameters, p and q. These values could be obtained by looking for significant autocorrelation and partial autocorrelation coefficients. The maximum number of useful autocorrelations is roughly n/4, where n is the number of periods upon which information on y_t is available. The general characteristics of theoretical autocorrelation functions (ACFs) and partial autocorrelation functions (PACFs) are as follows:

Table (3.1): Models with ACF and PACF for Stationary Processes

Process	ACF	PACF
AR	Tails off as exponential decay or damped sine wave	Cuts off after lag p
MA	Cuts off after lag q	Tails off as exponential decay or damped sine wave
ARMA	Tails off after lag(q-p)	Tails off after lag(p-q)

Source: Wei (2006)

Still these process can be used as initial values while the final models are achieved after going through the stages repeatedly.

3.2.4 Parameter Estimation

At the identification stage one or more models are tentatively chosen by adequately representations of the data. The next step is to estimate the parameters in the model. The main approaches to fit Box–Jenkins models are method of moments, least squares and maximum likelihood estimation. The Maximum likelihood estimation is generally the preferred technique. The conditional and unconditional maximum likelihood estimations are used to estimate the parameters $(\boldsymbol{\phi}, \boldsymbol{\theta}, \mu, \sigma_a^2)$.

In ARIMA (p,d,q) model , let $\mathbf{Y}=(Y_1,Y_2,\ldots,Y_n)'$ assume that the initial conditions $\mathbf{Y}_*=(Y_{1-p},\ldots,Y_1,Y_0)'$ and $\mathbf{a}_*=(a_{1-q},\ldots,a_{-1},a_0)'$ are known. The conditional likelihood function and the conditional log-likelihood function are

$$L(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \sigma_a^2 | \boldsymbol{Y}_*, \boldsymbol{a}_*) = (2\pi\sigma_a^2)^{-n/2} exp\left[-\left(\frac{1}{2\sigma_a^2}\right) S_*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta})\right] \qquad (3.27)$$

and

$$lnL_*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \sigma_a^2) = -\frac{n}{2} ln2\pi\sigma_a^2 - \frac{S_*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta})}{2\sigma_a^2} \qquad (3.28)$$

where $S_*(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}) = \sum_{t=1}^{n} a_t^2\,(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \sigma_a^2 | Y_*, \boldsymbol{a}_*)$ is the conditional sum of squares function and the quantities $\widehat{\boldsymbol{\phi}}, \widehat{\boldsymbol{\theta}}, \hat{\mu}$ are the conditional maximum likelihood estimators. The unconditional log likelihood function is (3.29)

$$lnL(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \sigma_a^2) = -\frac{n}{2} ln2\pi\sigma_a^2 - \frac{S(\boldsymbol{\phi}, \mu, \boldsymbol{\theta})}{2\sigma_a^2}$$

with $S(\boldsymbol{\phi}, \mu, \boldsymbol{\theta}) = \sum_{t=-\infty}^{n}[E(a_t | \boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \boldsymbol{Y})]^2$, the unconditional sum of squares function. It can be replaced by the finite form,

$$\sum_{t=-M}^{n} [E(a_t | \boldsymbol{\phi}, \mu, \boldsymbol{\theta}, \boldsymbol{Y})]^2 \tag{3.30}$$

where M is a sufficiently large integer.

3.2.5 Diagnostic Checking

After parameter estimation, the tentative model is checked with the following diagnostics:

(i) **Chi-square Goodness of Fit Test**: To check whether the errors are normally distributed, one can construct a histogram of the standardized residuals $\hat{a}_t / \hat{\sigma}_a$ and compare it with the standard normal distribution using Chi-square goodness of fit test or even Tukey's simple five-number summary.

(ii) **Residual Plots**: To check whether the variance is constant, it can be examined by the plots of residuals.

(iii) **ACF and PACF**: To check whether the residuals are approximately white noise, it can be computed using the sample ACF and PACF of the residuals to see whether they do not form any pattern and are all statistically insignificant. i.e., within two standard deviations if α=0.05.

(iv) **Portmanteau Test** (Q tests based on Chi square statistics)-**Box Pierce tests**: To check whether the residuals are white noise, the sample ACF and PACF of the residuals are needed to see whether they are non-significance. The approximate distribution is Chi squares computed as $Q = n \sum r_k^2$ where summation extends from 1 to K with k as the maximum lag considered, n is the number of observations in the series, r_k is the estimated autocorrelation at lag k; k can be any positive integer and is usually around 20. Q approximately follows Chi-square with (K-m) degrees of freedom where m is the number of parameters estimated in the model. The Q statistic (the Lung-box statistic) is given by

$$Q = n(n+2)\sum_{k=1}^{K}(n-k)^{-1}\hat{\rho}_k^2 \tag{3.31}$$

It is compared to the critical values from Chi-square distribution. If model is correctly specified, residuals should be uncorrelated and Q should be small (the probability value should be large). A significant value indicates that the chosen model does not fit well. All these stages require considerable care and work and they themselves are not exhaustive (Lamba & Dhaka, 2014).

To determine the best model with the appropriate lag value, the criteria are set as follows: relatively small Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC). The AIC is an important and learning statistics which can determine the order of the model. It takes into account both how well the model fits the observed series and the number of parameters to be used in the fit. These criteria take the form

$$AIC = -2\mathrm{L} + 2p. \tag{3.32}$$

$$BIC = 2L - p\ln N_p \tag{3.33}$$

where $L = \ln(\ell)$, ℓ is the likelihood, p is the number of coefficients or weights in the considered model.

3.2.6 Forecasting with ARIMA Models

One of the most important objectives in the analysis of a time series is to forecast its future values. ARIMA models are insightful when used to generate prediction interval for l period forecasts. In the general ARIMA (p,d,q) model,

$$\emptyset_p(B)(1-B)^d Y_t = \theta_q(B)a_t, \tag{3.34}$$

where the deterministic trend parameter θ_0 is omitted for simplicity but no loss of generality. This model is one of the most commonly used models in forecasting applications. When the series is known to time n, the optimal forecast of Y_{n+l} is given by its conditions expectation $E(Y_{n+l}|Y_n, Y_{n-l}, \dots)$.

That is, $\hat{Y}_n(l) = E(Y_{n+l}|Y_n, Y_{n-l}, \dots)$. The model at time $t+l$ can be written as in AR representation as follows

$$\pi(B)Y_{t+l} = a_{t+l} \tag{3.35}$$

where

$$\pi(B) = \frac{\emptyset_p(B)(1-B)^d}{\theta_q(B)} = \sum_{j=0}^{\alpha}\pi_j B^j, \tag{3.36}$$

or equivalently,

$$Y_{t+l} = \sum_{j=1}^{\infty} \pi_j Y_{t+l-j} + a_{t+l} \tag{3.37}$$

for $t \le n$, Y_t is $\hat{Y}_n(l) = E(Y_{n+l} | Y_t, t \le n) = \sum_{j=1}^{\infty} \pi_j^{(l)} Y_{t-j+1}$ for $j > 0$. The ψ_j weights can be calculated recursively from the π_j weights as follows:

$$\psi_j = \sum_{i=0}^{j-1} \pi_{j-i} \psi_i \;, \qquad j = 1, \ldots, l-1. \tag{3.38}$$

For a normal process, the (1-α)100% forecast limits are

$$\hat{Y}_n(l) \pm N\alpha_{/2} \left[1 + \sum_{j=1}^{l-1} \psi_j^2 \right]^{1/2} \sigma_a, \tag{3.39}$$

where $N\alpha_{/2}$ is the standard normal deviate such that $P\left(N > N\alpha_{/2}\right) = \frac{\alpha}{2}$.

3.2.7 Seasonal Process of ARIMA

Seasonality in a time series is a regular pattern of changes that repeats over S time periods, where S defines the number of time periods until the pattern repeats again. The ARIMA model can be extended to account for seasonal fluctuations, with the expression $(p, d, q)(P, D, Q)_S$. In a seasonal ARIMA model, seasonal AR and MA terms predict Y_t using data values and errors at times with lags that are multiples of *s* (the span of the seasonality). The Box-Jenkins multiplicative seasonal ARIMA model is

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D Y_t = \theta_q(B)\Theta_Q(B^s)a_t \tag{3.40}$$

where

$\Theta_Q(B^s) = 1 + \Theta_1 B^s + \cdots + \Theta_Q B^{Qs}$ and

$\Phi_P(B^s) = 1 - \Phi_1 B^s - \cdots - \Phi_P B^{Ps}$ are seasonal polynomial functions of order P and Q,

D = the number of seasonal differences.

It may be necessary to examine differenced data when we have seasonality. Seasonality usually causes the series to be nonstationary because the average values at some particular times within the seasonal span may be different than the average values at other times. Seasonal differencing is defined as a difference between a value and a value with lag that is a multiple of *s*. Seasonal differencing removes seasonal trend and can also get rid of a seasonal random walk type of non-stationarity.

Identifying a Seasonal Model

Step-1: **Do a time series plot of the data**: Examine it for features such as trend and seasonality. The gathered seasonal data look at the pattern across those time units to see if there is indeed a seasonal pattern.

Step-2: **Do any necessary differencing**. If there is seasonality and no trend, then take a difference of lag S. Seasonality will appear in the ACF by tapering slowly at multiples of S. Seasonal differences are supported in the ACF/PACF of the original data. If there is linear trend and no obvious seasonality, then take a first difference.

If there is a curved trend, consider a transformation of the data before differencing. If there are both trend and seasonality, apply a seasonal difference to the data and then re-evaluate the trend. If a trend remains, then take first differences. If there is neither obvious trend nor seasonality, do not take any differences.

Step-3**: Examine the ACF and PACF of the differenced data** (if differencing is necessary): Examine the patterns across lags that are multiples of *S*.

Step- 4**: Estimate the model(s)** that might be reasonable on the basis of step 3.

Step-5**: Examine the residuals** (with ACF, Box-Pierce, and any other means) to see if the model seems good.

And then, the model parameter estimation, diagnostic checking and forecasting for the seasonal ARIMA model follow the same general methods expressed in the non-seasonal ARIMA model.

3.2.8 AutoRegressive Integrated Moving Average with Exogenous (ARIMAX) Model and Seasonal (SARIMAX) Model

The autoregressive moving average model including exogenous variables (predictors), ARMAX(p,q), extends the ARMA(p,q) model by including the linear effect that one or more exogenous series has on the stationary response series y_t. The general form of the ARMAX (p,q) model is

$$Y_t = \sum_{i=1}^{p} \emptyset_i Y_{t-i} + \sum_{k=1}^{r} \beta_k X_{tk} + a_t + \sum_{j=1}^{q} \theta_j a_{t-j}, \tag{3.41}$$

and it has the following condensed form in lag operator notation:

$$\phi(B)X_t = c + X_t'\beta + \theta(B)a_{t.} \tag{3.42}$$

In Equation (3.42), the vector X_t' holds the values of the r exogenous, time-varying predictors at time t, with coefficients denoted β.

ARMAX models have the same stationarity requirements as ARMA models. Specially, the response series is stable if the roots of the homogeneous characteristic equation of $\emptyset(B) = B^p - \emptyset_1 B^{p-1} - \cdots - \emptyset_p B^p = 0$ lie outside of the unit circle. If the response series y_t is not stable, then it can be differed to form a stationary ARIMAX model. Therefore, the exogenous variables enter a model with a stationary response, the ARIMAX (p,d,q) model is

$$\phi(B)Y_t = c^* + X'_t\beta + \theta^*(B)a_{t.} \tag{3.43}$$

where $c^* = c/(1-B)^d$ and $\theta^*(B) = \theta(B)/(1-B)^d$.

Subsequently, the interpretation of β has changed to the expected effect a unit increase in the predictor has on the difference between current and lagged values of the response. The maximum likelihood estimation was used for conditional mean models, ARIMAX models.

An ARIMAX model includes the seasonal components which creates a SARIMAX (p,d,q)(P,D,Q)$_s$ model. Assuming that the response series Y_t is stationary, the model has the form

$$Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{k=1}^{r} \beta_k X_{tk} + \sum_{j=1}^{q} \theta_j a_{t-j} + \sum_{i=1}^{P} \Phi_i Y_{t-si} + \sum_{j=1}^{Q} \Theta_j a_{t-sj} + a_t\,, \tag{3.44}$$

and it was written with the backshift operator:

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D Y_t = X'_t\beta + \theta_q(B)\Theta_Q(B^s)a_t \tag{3.45}$$

where

- s = seasonal lag,
- ϕ = coefficient for AR process,
- Φ = coefficient for seasonal AR process,
- θ = coefficient for MA process,
- Θ = coefficient for seasonal MA process,
- X'_t = the predictor variables at time t,
- β = the coefficients for predictor variables.

where $\Phi(B)$ and $\Theta(B)$ are the seasonal lag polynomials. If Y_t is not stationary, the degree of non-seasonal or seasonal integration can be specified (Artley, 2022).

3.3 Vector Auto-Regression (VAR)

Multivariate time series analysis considers simultaneously multiple time series. It is a branch of multivariate statistical analysis but deals specifically with dependent variable. The most commonly used multivariate time series model is the vector autoregressive (VAR) model. One can use the least-squares (LS) method, the maximum likelihood (ML) method, or Bayesian method. All three estimation methods have closed-form solutions.

For a VAR model, the least-squares estimates are asymptotically equivalent to the ML estimates and the ordinary least-squares (OLS) estimates are the same as the generalized least-squares (GLS) estimates. And then VAR models are widely used in multivariate statistical analysis.

The Vector Autoregression models are the better alternative to traditional dynamic simultaneous-equation models to examine the dynamic interactions among the interrelated time series data. The VAR model is one of the most successful, flexible, and easy models for the analysis of multivariate time series.

The VAR models are the multivariate extensions of the univariate AR models to the multivariate case and they explain and predict the values of a set of variables at any given point in time. In terms of the use of VAR model, it required applying stationary condition criteria. A *p*th-order VAR is denoted "VAR(*p*)" and sometimes called "a VAR with *p* lags".

The time series y_t follows a VAR (p) model, if it satisfies

$$\boldsymbol{Y}_t = \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \mathbf{Y}_{t-1} + \cdots + \boldsymbol{\Phi}_p \mathbf{Y}_{t-p} + \boldsymbol{a}_t\,, \quad p > 0 \tag{3.46}$$

where, $\boldsymbol{Y_t}$ is a vector of the dependent variable; there are n equations.

$\boldsymbol{\Phi}_0$ is a n-dimensional vector of constants.

p is the order of the VAR.

$\boldsymbol{\Phi}'$s are $(n \times n)$ coefficient matrices and

$\boldsymbol{a}_t$ is a sequence of serially uncorrelated random vectors with mean zero and covariance matrix $\boldsymbol{\Sigma}$ such that

1. E($\boldsymbol{a}_t$)= 0. Every error term has a mean of zero.
2. E($\boldsymbol{a}_t \boldsymbol{a}_t'$)= $\boldsymbol{\Sigma}$. The contemporaneous covariance matrix of error terms is a $n \times n$ positive-semidefinite matrix denoted $\boldsymbol{\Sigma}$.
3. E($\boldsymbol{a}_t \boldsymbol{a}_{t-k}'$)=0 for any non-zero k. There is no correlation across time. In particular, there is no serial correlation in individual error terms.

The process of choosing the maximum lag p in the VAR model requires special attention because inference is dependent on correctness of the selected lag order.

In lag operator notation, the VAR(p) is written as

$$\boldsymbol{\Phi}(B)\boldsymbol{Y_t} = \boldsymbol{\Phi_0} + \boldsymbol{a}_t \ , \tag{3.47}$$

where $\boldsymbol{\Phi}(B) = \left(1 - \Phi_1 B - \cdots - \Phi_p B^p\right)$ is a matrix of polynomials in the lag operator.

All variables are served as endogenous variables. In other words, each endogenous variable is explained by its lagged or past values and the lagged values of all other endogenous variables in the model.

3.3.1 Estimation of VAR

Since the ***VAR(p)*** may be written as a system of equations with the same sets of explanatory variables, its coefficients can be efficiently and consistently estimated by estimating each of the components using the Ordinary Least Squares (OLS) method. Under standard assumptions regarding the behavior of stationary and ergodic VAR models the estimators of the coefficients are asymptotically normally distributed.

The estimation of the parameters is starting from concise matrix notation:

$$Y = BX + A \tag{3.48}$$

the multivariate least squares (MLS) approach for estimating B yields:

$$\hat{B} = Y\acute{X}(\acute{X}X)^{-1} \tag{3.49}$$

This can be written alternatively as:

$$Vec\left(\hat{B}\right) = ((X'X)^{-1}X \otimes I_k)Vec(Y), \tag{3.50}$$

where $\otimes$ denotes the Kronecker product and Vec is the vectorization of the indicated matrix.

This estimator is consistent and asymptotically efficient. It is furthermore equal to the conditional maximum likelihood estimator. As the predictor variables are the same in each equation, the multivariate least squares estimator is equivalent to the ordinary least squares estimator applied to each equation separately. The covariance matrix of the errors is estimated by the maximum likelihood estimator (MLE) of the covariance matrix differs from the ordinary least squares (OLS) estimator. In matrix notation:

$$\hat{\Sigma} = \frac{1}{T - np - 1}\left(Y - \hat{B}X\right)\left(Y - \hat{B}X\right)'. \tag{3.51}$$

The covariance matrix of the parameters can be estimated as

$$\widehat{Cov}\left(Vec(\hat{B})\right) = (X'X)^{-1} \otimes \hat{\Sigma}\,. \tag{3.52}$$

After estimating VAR model, it is of crucial interest to see whether the residuals satisfy the model's assumptions such as, serial correlation, heteroscedasticity and normality. Portmanteau test is applied for testing the lack of serial correlation as well as the multivariate normality test.

3.3.2 Lag Length Selection

The biggest practical challenge in VAR modeling is to choose the appropriate lag length. For VAR models, model specification is to select the order *p*. Several methods have been proposed in the literature to select the VAR order. In this study, the information criteria approach is used.

For normal distribution, the maximized likelihood is equivalent to the determinant of the covariance matrix of the innovations. This determinant is known as the *generalized variance* in multivariate analysis. The selection of the penalty, on the other hand, is relatively subjective. Different penalties result in different information criteria. Three criteria functions are commonly used to determine VAR order. Under the normality assumption, these three criteria for a VAR(p) model are

$$AIC = ln\left|\hat{\Sigma}_p\right| + \frac{2}{\mathrm{T}} pn^2 \tag{3.53}$$

$$BIC\ or\ SC = ln\left|\hat{\Sigma}_p\right| + \frac{\ln(\mathrm{T})}{\mathrm{T}} pn^2, \tag{3.54}$$

$$HQ = ln\left|\hat{\Sigma}_p\right| + \frac{2\ln\,[\ln(\mathrm{T})]}{\mathrm{T}} pn^2\ , \tag{3.55}$$

where pn^2 is the number of estimated parameters, T is the sample size, $\hat{\Sigma}_p$ is the ML estimate of $\boldsymbol{\Sigma}_a$, AIC is the Akaike information criterion, BIC (SC) stands for Bayesian information criterion, and HQ(ℓ) is the Hannan and Quinn Criterion. The AIC penalizes each parameter by a factor of 2. BIC and HQ, on the other hand, employ penalties that depend on the sample size. Although, the lag order has the lowest value of these information criteria, it take into account that the reason for the poor forecasting may be due to the over-parameterization of VAR models (Tsay, (2014)).

3.3.3 Granger Causality Test

The importance tool in VAR model is to perform the Granger causality testing to examine the direction of causality among the variables. It is a technique for determining whether one time series is useful in forecasting another. Consider a bivariate series and the *h*-step ahead forecast. In this case, the VAR model and univariate models for individual components can be used to produce forecasts.

The accuracy of a forecast is measured by the variance of its forecast error. In other words, under Granger's framework, Y_t causes X_t if the past information of Y_t improves the forecast of X_t. Tests of the restrictions can be based on simple F tests in the single equations of the VAR model. The causality equation is

$$Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{p} \beta_j X_{t-j} + \varepsilon_t \qquad (3.56)$$

It is postulates that the current Y_t is related to its past values as well as that of X_t and vice versa. Unidirectional causality from X_t and Y_t is indicated if the estimated coefficient on the lagged X_t are statistically different from zero as a group ($\sum \beta_j \neq 0$).

If lagged values of a variable Y_t have no explanatory power for any of the variables in a system, then it would be viewed as weakly exogenous to the system. For testing Granger causality in the VAR model, only the restricted equations are relevant. Based on this insight it is straightforward to test the null hypothesis that the second variable does not cause the first one within the VAR(p) context: H_0 : $\{Y_t\}$ does not cause $\{X_t\}$. The alternative hypothesis is that the null hypothesis is violated. The hypothesis can be tested using the likelihood ratio statistic. This test is wedded to the normal distribution limits its generality. The Wald test or its transformation to an approximate F statistic is an alternative that should be more generally applicable. The test statistic is given by

$$F = \frac{(RSS_R - RSS_{UR})/m}{RSS_{UR}/(n-k)} \qquad (3.57)$$

where RSS_R, restricted residual sum of squares, RSS_{UR}, unrestricted residual sum of squares, *m*, number of lagged X terms, and *k*, number of parameters estimated in the unrestricted regression. The test statistic follows the F-distribution with *m* and *(n-k)* degrees of freedom.

The causality tests are predicated on a model that may be missing either intervening variables or additional lagged effects that should be present but are not.

3.3.4 Forecasting VAR

If the fitted model is adequate, then it can be used to obtain forecasts. For forecasting, same techniques in the univariate analysis can be applied. Let *h be the forecast origin,* $\ell > 0$ be the forecast horizon, and F_h be the information available at time h (inclusive).

For a VAR (p) model, the l -step ahead forecast at the time origin h is:

$$Y_h(\ell) = E(y_{h+\ell}|F_h) \tag{3.58}$$

$$= \phi_0 + \sum_{i=1}^{p} \phi_i y_h(\ell - i), \tag{3.59}$$

where it is understood that $y_h(j) = y_{h+j}$ for $j \leq 0$. Thus, the point forecast of a VAR(p) model can be computed recursively. The associated forecast error is $e_h = a_{h+l}$. The covariance matrix of the forecast error is Σ. If Y_t is weakly stationary, then the l -step ahead forecast $Y_h(l)$ converges to its mean vector μ as the forecast horizon (l) increases.

Turn to forecast errors. For ℓ-step ahead forecast, the forecast error is

$$e_h(\ell) = y_{h+l} - y_h(\ell). \tag{3.60}$$

To study this forecast error, it is most convenient to use the MA representation of the VAR(p) model,

$$Y_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i}, \tag{3.61}$$

where $\mu = [\phi(1)]^{-1}\phi_0, \psi_0 = \mathrm{I}_k$, and ψ_i can be obtained recursively. The ℓ-step ahead forecast error is

$$e_h(\ell) = a_{h+\ell} + \psi_1 a_{h+l-1} + \cdots + \psi_{\ell-1} a_{h+1}. \tag{3.62}$$

Consequently, the covariance matrix of the forecast error is

$$Cov[e_h(\ell)] = \Sigma_a + \sum_{i=1}^{\ell-1} \psi_i \Sigma_a \acute{\psi}_i. \tag{3.63}$$

In practice, the parameters of a VAR(p) model are unknown, and one would like to take into account the parameter uncertainty in forecasting. For simplicity and similar to real-world applications, we assume that the parameters are estimated using the information available at the forecast origin $t=h$. That is, estimation is carried out based on the available information in F_h. Under this assumption, parameter estimates are functions of F_h and, hence, the ℓ-step ahead minimum mean squared error (MSE) forecast of $y_{h+\ell}$ with estimated parameters is

$$\hat{Y}_h(\ell) = \hat{\phi}_0 + \sum_{i=1}^{p} \hat{\phi}_i \hat{Y}_h(\ell - i), \tag{3.64}$$

where, as before, $\hat{Y}_h(j) = Y_{h+j}$ for $j \leq 0$. Together with forecasts, impulse response analysis and forecast error variance decomposition are other tools for investigating the dynamic relationships.

3.3.5 Impulse Response Functions

After an adequate VAR model is obtained, the two important tools such as Impulse Response Functions and Variance Decompositions are used to interpret the parameters. Impulse response functions provided by VAR models are used to know where the impact of change in one variable can be found through all the other variables. They exhibit the current and lagged effects over time of changes in error terms on the endogenous variables (Y_{1t} ,Y_{2t} ,…, Y_{kt}). When the VAR process of order 'p' is stable, the error term $\boldsymbol{a}_{1t}$ has immediate effects and all have lagged effects on y_{1t}.

In multiplier analysis, assume that $E(\boldsymbol{Y_t}) = \boldsymbol{0}$ because the mean does not affect the pattern of the response of $\boldsymbol{y_t}$ to any shock. To study the effects of changes in y_{1t} on $\boldsymbol{Y}_{t+j}$ for $j > 0$ while holding other quantities unchanged, assume that t=0, $\boldsymbol{Y}_t = \boldsymbol{0}$ for $t \leq 0$, and $\boldsymbol{a}_0 = (1,0, ...)'$. To this end, it can be traced out $\boldsymbol{Y_t}$ for t=1, 2, ..., assuming $\boldsymbol{a}_t = \boldsymbol{0}$ for $t > 0$. Using the MA representation of a VAR(p) model with coefficient matrix $\boldsymbol{\psi}_\ell = [\psi_{\ell,ij}]$, it becomes

$$\boldsymbol{Y}_0 = \boldsymbol{a}_0 = \begin{bmatrix} 1 \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix}, \boldsymbol{Y}_1 = \boldsymbol{\psi_1 a_0} = \begin{bmatrix} \psi_{1,11} \\ \psi_{1,21} \\ . \\ . \\ . \\ \psi_{1,k1} \end{bmatrix}, \boldsymbol{Y}_2 = \boldsymbol{\psi_2 a_0} = \begin{bmatrix} \psi_{2,11} \\ \psi_{2,21} \\ . \\ . \\ . \\ \psi_{2,k1} \end{bmatrix}, \ldots . \tag{3.65}$$

The results are simply the first columns of the coefficient matrices $\boldsymbol{\psi_i}$. Similarly, to study the effect on y_{t+j} by increasing the ith series y_{it} by 1, it have $a_0 = e_i$, where e_i is the ith unit vector. Hence,

$$\boldsymbol{Y}_0 = e_i, \quad \boldsymbol{Y}_1 = \boldsymbol{\psi}_{1,.i}, \quad \boldsymbol{Y}_2 = \boldsymbol{\psi}_{2,.i}, \ldots . \tag{3.66}$$

They are the i^th^ columns of the coefficient matrices ψ_i of the MA representation of y_t. For this reason, the coefficient matrix ψ_i of the MA representation of a VAR(p) model is referred to as the coefficients of impulse response functions.

3.3.6 Forecast Error Variance Decomposition

The variance decomposition analysis is typically performed by VAR models, which supplements impulse response function analysis. It shows that how much the variance of the forecast errors of each variable can be explained by exogenous shocks to the other variables in the VAR. Thus, it provides information about the comparative magnitude of each random innovation. Using the MA representation of a VAR(p) model, the ℓ-step ahead forecast error of $y_{h+\ell}$ at the forecast orign t=h can be written as

$$e_h(\ell) = \boldsymbol{\psi}_0\eta_{h+\ell} + \boldsymbol{\psi}_1\eta_{h+\ell-1} + \cdots + \boldsymbol{\psi}_{\ell-1}\eta_{h+1}, \tag{3.67}$$

and the covariance matrix of the forecast error is

$$Cov[e_h(\ell)] = \sum_{\upsilon=0}^{\ell-1} \boldsymbol{\psi_\upsilon}\boldsymbol{\acute{\psi}_\upsilon}\,. \tag{3.68}$$

The variance of the forecast error $e_{h,i}(\ell)$, which is the ith component of $e_h(\ell)$, is

$$Var\left[e_{h,i}(\ell)\right] = \sum_{\upsilon=0}^{\ell-1}\sum_{\boldsymbol{j=1}}^{\boldsymbol{k}} \boldsymbol{\psi}^2{}_{\upsilon,ij} \tag{3.69}$$

where

$$w_{ij}(\ell) = \sum_{\upsilon=0}^{\ell-1} \boldsymbol{\psi}^2{}_{\upsilon,ij}. \tag{3.70}$$

Therefore, the quantity $w_{ij}(\ell)$ can be interpreted as the contribution of the jth shock η_{jt} to the variance of the ℓ-step ahead forecast error of y_{it}. Then the Forecast Error Variance Decomposition (FEVD) is obtained from the orthogonal impulse response coefficient matrices;

$$Var\left[e_{h,i}(\ell)\right] = \sum_{j=1}^{k} w_{ij}(\ell)\,. \tag{3.71}$$

In particular, $w_{ij}(\ell)/Var\left[e_{h,i}(\ell)\right]$ is the percentage of contribution from the shock η_{jt} (Tsay,(2014)).

3.4 Artificial Neural Networks (ANNs)

The ANN was pioneered more than 40 years ago and nowadays, there has been a great interest in neural networks since an artificial network shares some of the physical and behavioral aspects of a biological one. The ANN structure, a parallel system is based on the human brain's biological neural process used to solve complex problems, which it tries to imitate into mathematical models (Khairunniza-Bejo, Muataffha, & Wan, 2014).

The first step toward artificial neural networks came in 1943, when Warren McCulloch, a neurophysiologist, and a young mathematician, Walter Pitts, wrote a paper on how neurons might work. They modeled a simple neural network with electrical circuits (Greeshma, 2015). Neural Network (NN) is a vast domain of technology where one can implement "human brain decision-making power" into computer programs on the basis of error and approximation. Also, a lot of research and development has been made in the field of artificial intelligence with the help of neural networks (Karsoliya, 2012).

An Artificial Neural Network (ANN) is also called a multilayer perceptron. Several units representing neurons are interconnected to form an ANN. The synapses, the tips of an axon, are represented by a modifiable weight. Each unit receives an input integrated with the weight, a floating point number, and transfers other units. Each input unit multiplies with its associated weight on the connection, and all weighted inputs are added to get a quantity called the total input. The output of the network is highly influenced by the input-output function of individual units and the associated weights. The connection between the units and other units must be established for a neural network to carry out a certain task (Ranjeet & Armstrong, 2014).

Neural Networks have been widely used for time series prediction because the structure is flexible to train time series data. NNs are highly robust with respect to underlying data distributions and no assumptions are made about relationships between parameters. This method is able to successfully predict the outcome of a process by using pairs of input and output data in a learning procedure (Karkalos et al., 2019).

A prerequisite of intelligent system has brought artificial neural network (ANN) to become a new technology which provides assorted solution for the complex problems in agriculture researches. Since it can solve many problems that linear system is incapable to resolve, ANN becomes crucial especially in innovating and developing better products for society.

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. A neural network can be trained to perform a particular function by adjusting the values of the connections (weights) between the elements. Commonly neural networks are adjusted, or trained, so that a particular input leads to a specific target output. Figure (3.1) shows such a situation.

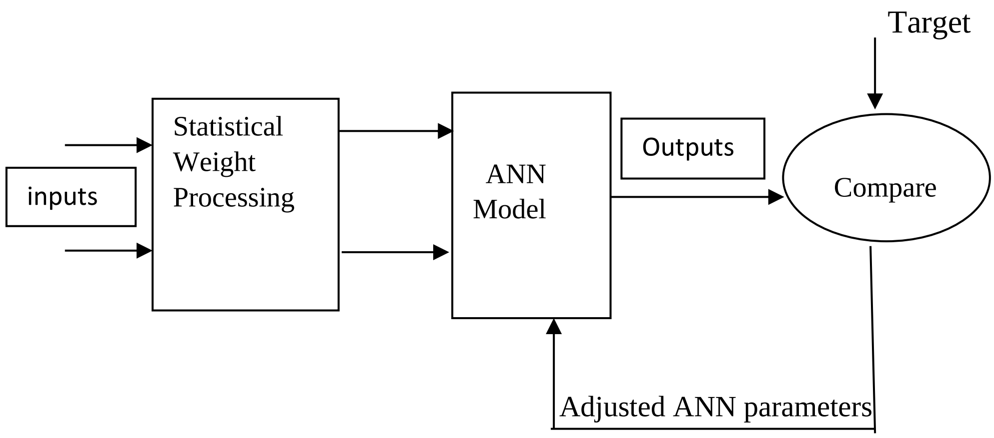


Figure (3.1): Basic Principle of ANNs

Source: Nanda et al. (2013)

Here, the network is adjusted, based on a comparison of the output and the target, until the sum of square differences between the target and output values becomes the minimum (Nanda et al., 2013).

3.4.1 An Artificial Neuron

Basic building block of every artificial neural network is artificial neuron. A neuron is an information-processing unit that is fundamental to the operation of a neural network. Moreover, an ANN is a mathematical model that tries to simulate the structure and functionalities of biological neural networks. (Lamba & Dhaka, 2014). Figure (3.2) shows basic representation of an artificial neuron.

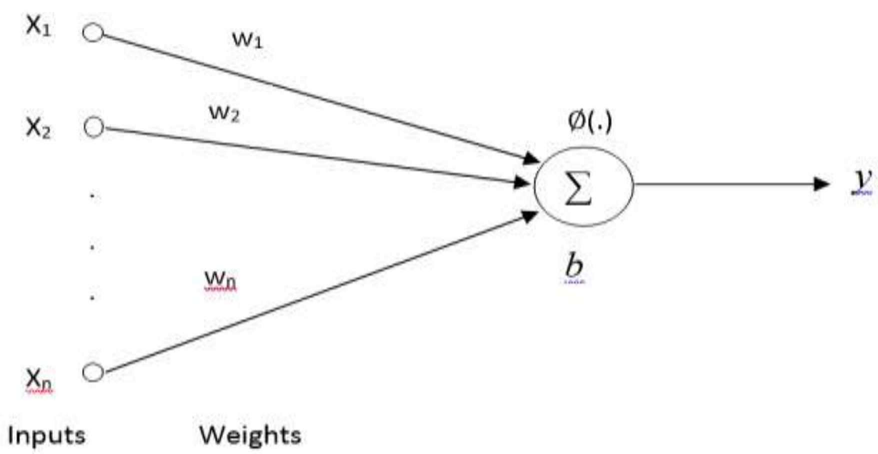


Figure (3.2): A Basic Artificial Neuron

Source: Bangal (2009)

In Figure (3.2), various inputs to the network are represented by the mathematical symbol, x_n. Each of these inputs is multiplied by a connection weight. The weights are represented by w_n. In the simplest case, these products are summed, fed to a transfer function (activation function) to generate a result, and this result is sent as output. This is also possible with other network structures, which utilize different summing functions as well as different transfer functions.

Moreover, Figure (3.2) illustrates the simple neuron model. The output of the neuron is given by

$$Y = \phi(net), \qquad net = \sum_{i=1}^{n} w_i X_i + b = \boldsymbol{w}^T \boldsymbol{X} + b \tag{3.72}$$

where X_i is the ith input, w_i is the link weight from the ith input, $\boldsymbol{w} = (w_1,\ldots,w_n)^T$, $X=(X_1,\ldots,X_n)^T$, b is a threshold or bias, and n is the number of inputs. The activation unction $\phi(.)$ is usually some continuous or discontinuous function.

3.4.2 Weighting Factors and Summation Function

A neuron usually receives many simultaneous inputs. Each input has a relative weight, which gives it the influence it requires on the summing function of the processing unit. Some inputs are more important than others to have a greater effect on the processing element as they combine to produce a neural response. Weights are adaptive coefficients that determine the intensity of the input signal as registered by the artificial neuron. They are a measure of an input's connection strength. These strengths can be modified in response to various training sets and according to a network's specific topology or its learning rules.

The quantity of neurons in the hidden layer is used to capture the complexity of the data. The connections between neurons should be limited when using neural networks for practical reasons. This is done by fixing some of the weights to zero so that they can be dropped from the calculations, the working principal for subsequent adjustment of weight is in accordance with the error propagation in the network. If increasing a given weight leads to more error, the weights are adjusted by downwards and if increasing a given weight leads to less error, the weights are adjusted by upwards (Sharma, 2012).

The inputs and corresponding weights are vectors which can be represented as (X_1, X_2 . . . X_n) and (w_1, w_2 . . . w_n). The total input signal is the dot product of these two vectors. The result; $(X_1 * w_1) + (X_2 * w_2) + \ldots\ldots + (X_n * w_n)$; is a single number. The summation function can be more complex than just weight sum of products. The input and weighting coefficients can be combined in many different ways before passing on to the transfer function.

In addition to summing, the summation function can select the minimum, maximum, majority, product, or several normalizing algorithms. The specific algorithm is determined for combining neural inputs is determined by the chosen network architecture and paradigm. Some summation functions have an additional 'activation function' applied to the result before it passes on to the transfer function to allow the summation output to vary the concerning time.

3.4.3 Activation (Transfer) Function

The activation function is also called the transfer function. It determines the relationship between the inputs and outputs of a node and a network (Zhang, Eddy & Patuwo, 1998). For neural networks to be able to rectify errors, this function must be differential and thus continuous. The local gradient must be calculated using the transfer function's derivative (Bangal, 2009). It can take several forms. The type of activation function is indicated by the situation of the neuron within the network. In the majority of cases, input-layer neurons do not have an activation, as their role is to transfer the inputs to the hidden layer.

In practice, only a small number of activation functions are used. These include:

(i) The sigmoid function: $f(x) = \frac{1}{1+\exp(-x)}$

(ii) The hyperbolic tangent (tanh) function: $f(x) = \frac{\exp(x)-\exp(-x)}{\exp(x)+\exp(-x)}$

(iii) The identity function: $f(x) = x$

The target output values usually need to be normalized to match the range of actual outputs from the network since the output node with a sigmoid or hyperbolic tangent function has a typical range of [0,1] or [-1,1], respectively. The sigmoid activation functions are for classification problems that involve learning about average behavior, and the hyperbolic tangent functions are used if the problem involves learning about deviations from the average, such as a forecasting problem. The most popular transfer (activation) function is the sigmoid function for the hidden and output. The input layer commonly uses a linear transfer function to pass the information to hidden layers. The sigmoid function is a non-linear, curved, S-shaped function. It strictly increases function by nature (Khairunniza-Bejo, Muataffha & Wan, 2013).

3.4.4 ANN Architecture

In the theory of artificial neural networks, several inputs are used as inputs, and a specified nonlinear function produces the output. All neurons in the neural network model are divided into an input layer, a hidden layer, and an output layer depending on their function, and each layer is functionally connected (Karsoliya, 2012). Though there are many types of ANN, the study only presented the most commonly used type, the multilayer feed-forward network.

A feed forward neural network begins with an input layer. This input layer is connected to a hidden layer. This hidden layer is connected with other hidden layer if any other one is there otherwise it is connected to the output layer. In common most of the neural networks have at least one hidden layer, and it is scare to have more than two hidden layer.

(i) Input Layer:

The Input layer is a layer which communicates with the external environment that presents a pattern to the neural network. Once a pattern is presented to the input layer, the output layer will produce another pattern. In essence, this is all the neural network does. The input layer should represent the condition for which we are training the neural network. Every input neuron should represent some independent variable that has an influence over the output of the neural network.

(ii) Output Layer:

The output layer of the neural network is what presents a pattern to the external environment. The pattern presented by the output layer can be directly traced back to the input layer. The number of output neurons should be directly related to the type of work that the neural network is to perform. To determine the number of neurons to use in the output layer, first, consider the intended use of the neural network. If the neural network is to be used to classify items into groups, then it is often preferable to have one output neuron for each group that input items are to be assigned into. If the neural network is to perform noise reduction on a signal, then it is likely that the number of input neurons will match the number of output neurons.

(iii) Hidden Layer:

A hidden layer is a group of neurons that serves as both an intermediary layer between the input layer and the output layer and has an activation function applied to it. Many researches had been made in evaluating the number of neurons in the hidden layer but still none were accurate (Karsoliya, 2012).

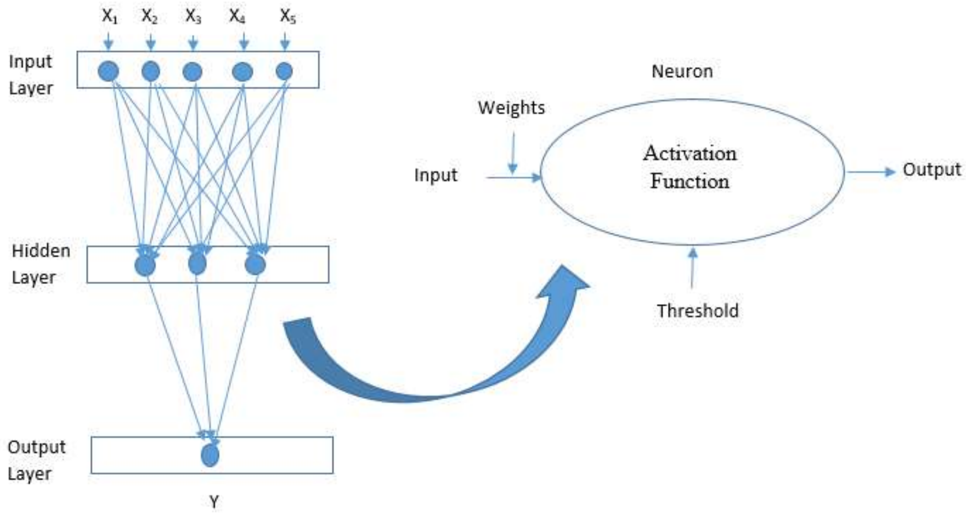


Figure (3.3): Architecture of ANN

Source: Panchal & Panchal (2014)

In Figure (3.3), the input layer connects the external input mode and is transmitted in units of hidden layers according to the input unit. Here, the hidden layer is the inner processing unit layer of the neural network and the output layer is used to generate the output mode. The input layer is used to distribute the inputs to a number of hidden layers, and the output of which is connected to an output layer, where the outputs of units are connected to the inputs of the next via connection weight (Khairunniza-Bejo, Muataffha & Wan, 2014).

Another concern is how many hidden nodes are required to solve a complex problem. It is necessary to determine how many neurons should be preserved in the hidden layer. "Underfitting" may happen if the number of neurons is lower than the complexity of the problem data. When there are not enough neurons in the hidden layers to properly detect the signals in a complex data set, underfitting takes place. "Overfitting" may happen if there are unnecessarily more neurons than necessary in the network. It happens when the network contains extra neurons that are not necessary.

To avoid overfitting and underfitting the data, a neural network is usually trained on a subset of inputs and outputs to determine weights and subsequently validated on the remaining data to measure the accurate prediction. There are some various approaches to finding the number of hidden nodes in a hidden layer. The layer approximation of a neural network is created by Maximum Developer using a "trial and error" approach (Karsoliya, 2012).

3.4.5 Partition of Data and Training Method

The first issue here is the division of the data into the training and test sets. The selection of the training and test samples may affect the performance of ANNs. Most authors select them based on the rule of 90% vs. 10%, 80% vs. 20% or 70% vs. 30%, etc (Zhang, Eddy & Patuwo, 1998). The training set is the largest set and is used by neural network to learn patterns present in the data. The testing set is used to evaluate the generalization ability of a supposedly trained network (Laxmi & Kumar, 2011).

Data normalization is often performed before the training process begins. It is needed to improve the performance of numerical computation and obtain better neural network output results by avoiding the influence of one attribute over another. When nonlinear transfer functions are used at the output nodes, the desired output value must be transformed to the range of the actual outputs of the network. The normalization formula is frequently used:

$$z_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{3.73}$$

where z and x represent the normalized and original data; x_{max} *and* x_{min} are the maximum and minimum of the original dataset. Moreover, standardization consists of subtracting a quantity related to a measure of localization or distance and dividing by a measure of the scale**.**

In most learning networks the difference between the current output and the desired output is calculated as an error which is then transformed by the error function to match particular network architecture. The error is propagated backwards to a previous layer. The algorithm for evaluating the derivative of the error function is called backpropagation, because it propagates the errors backward through the network. The backpropagation algorithm uses supervised learning. Supervised training involves a mechanism of providing the network with the desired output either by manually "grading" the network's performance or by providing the desired outputs with the inputs.

The idea of the backpropagation algorithm is to reduce the error (difference between actual and expected results), until the ANN learns the training data. Once a network has been structured for a particular application, it is ready for training. Basically, training is the process of determining the weights which are the key elements of an ANN. The knowledge learned by a network is stored in the nodes in the form of weights and node biases. The training starts with random weights. The goal is to adjust them so that the error will be minimal (Gershenson, 2003).

3.4.6 Assumptions of ANN

Certain assumptions on the flow of the information from one layer to another make neural network simpler. But, these assumptions might not need to follow. Those assumptions are

(i) Artificial Neurons are arranged in layers, which are sequentially arranged.
(ii) Neurons within the same layer do not interact or communicate to each other.
(iii) All inputs enter into the network through the input layer and passes through the output layer.
(iv) All hidden layers at same level should have same activation function.
(v) Artificial neuron at consecutive layers are densely connected.
(vi) Every inter-connected neural network has its own weight and biased associated with it (Santosh, 2020).

3.5 Model Validation

Finding the best forecast time series model it is pivotal to trace out the forecast accuracy of test data. There are many analytical methods for the evaluation and inter comparison of different models. The root means squared error (RMSE), which measures between fitted and observed values, was calculated to evaluate the systematic bias of the model. The smaller of the RMSE, the better the model is for forecasting:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \ . \tag{3.73}$$

Moreover, the statistical model validated based on the statistical significance (i.e. p-value) where linear regression was applied to compare observed and calibrated data and the model's explanatory power as measured by the coefficient of determination (R^2) for each model. It is computed by

$$R^2 = \frac{\sum_{t=1}^{n}(\hat{y}_t - \bar{y})^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2}. \tag{3.74}$$

A high of this value indicates the best model performance in capturing the observed crop yield response to climate.

Furthermore, bias and accuracy of models was measured through the mean absolute percentage error (MAPE) using the formula:

$$MAPE = \left(\frac{1}{n}\sum_{t=1}^{n}\left|\frac{(y_t - \hat{y}_t)}{y_t}\right|\right)100\% \tag{3.75}$$

where n is the total number of samples, Y_t and $\hat{Y}_t$ are the actual and predicted values respectively for t = 1,2,…,n. In terms of the MAPE, the good model will have the smallest possible value of MAPE (Raj, Ramesh & Rajkumar, 2019).

By using the above methodology for each model; MLR, ARIMAX, VAR and ANN, the rice production will be studied based on the weather parameters. Among of those models, the appropriated model will be chosen on the model validation criteria as RMSE, R^2 and MAPE. The parts of data analysis were presented in the next section.

CHAPTER IV
DATA ANALYSIS

This chapter attempted to identify the fitted model for forecasting rice production on which was effected by the climate change through assessing the several different approaches such as MLR, SARIMAX, VAR and ANN. The data analysis was conducted by four parts. The first one describes the sources of data. Secondly, the data used in study are summarized. Thirdly, the fitted model is evaluated. Finally, the forecasting is done by the fitted model acquired.

4.1 Sources of Data

The secondary data for the production and sown acreage of rice (monsoon and summer) were collected for the period from 1992-1993 to 2020-2021 from the Department of Agriculture in Ayeyawady Region. The rice production was used as rice yield in the analysis, and rice yield is calculated that the annual rice production is divided by the annual sown acreage.

The secondary data for the climatic variables at the monthly scale were collected for only those districts having weather stations such as Pathein, Hinthada, Maubin, and Myaungmya during the period from 1992–2021 from the Department of Meteorology and Hydrology in Myanmar. But, the climatic data for Phyarpon Station were not available for the period between 1992 and 2000. Therefore, the weather data at Phyarpon station were collected only for the period from 2001 until 2020. In this study, the climatic variables used for the analysis were Rainfall in millimeters (mm), Maximum Temperature in degrees centigrade (MaxT˚C), Minimum Temperature in degrees centigrade (MinT˚C), Morning Relative Humidity at 9:30 a.m. (MRH) in percentage, and Evening Relative Humidity at 6:30 p.m. (ERH) in percentage.

4.2 Trend of Rice Production in Ayeyawady Region

In Ayeyawady Region, rice is a predominant crop and there are different methods of cultivation depending on the availability of water resources. Almost the entire that region grows rice on the farmland. Rice is cultivated in two seasons, monsoon and summer. Monsoon rice cultivation refers to the practice of growing rice solely with rainwater as the primary water source. Summer rice cultivation is well-suited for areas with access to irrigation systems, such as canals, water reservoirs, or tank wells. Those two kinds of rice are grown from May to October and November to April, respectively. The trend of the production of two types of rice in Ayeyawady Region from 1992-1993 to 2020-2021 is shown in Figure (4.1).

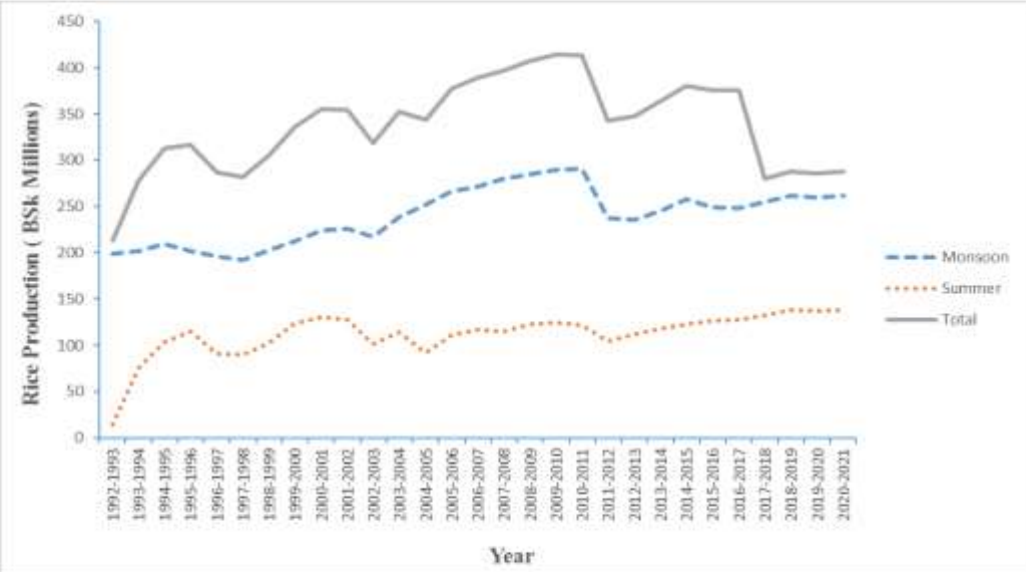


Figure (4.1): Trends of Rice Production in Ayeyawady Region

Source: Department of Agriculture in Ayeyawady Region

The trends of rice production for both monsoon and summer are presented in Figure (4.1) which indicates that monsoon rice contributes a large percentage of the total rice production of the Ayeyawady Region, whereas summer rice contributes less. But, both the monsoon and summer rice are considered in the study. Rice production data were expressed in standard units such as average rice yield per acre in basket (Bsk/Ac).

Moreover, the trends for both rice production by each of the five districts in Ayeyawady Region are depicted in Figure (4.2). The trends of rice yield in all districts show a slight fluctuation, but there is no appearing upward or downward trend in the data over the whole period. Besides, rice yields are high in summer and low in monsoon period. The difference in rice yields between these seasons can be attributed to several factors. During the summer season, the region receives limited rainfall, and farmers rely on irrigation to provide water to the paddy fields. This controlled irrigation allows for consistent water supply, which is beneficial for rice growth and development. This region experiences high temperatures and longer days, providing favorable conditions for rice cultivation. In contrast, the monsoon period, typically spanning from June to October, brings heavy rainfall. While water is plentiful during this time, excessive rainfall can lead to flooding and waterlogging, which negatively impacts rice plants, inhibits their growth, and reduces yields. Moreover, it is characterized by cloudy weather and shorter days, which can limit the photosynthetic activity and productivity of rice plants.

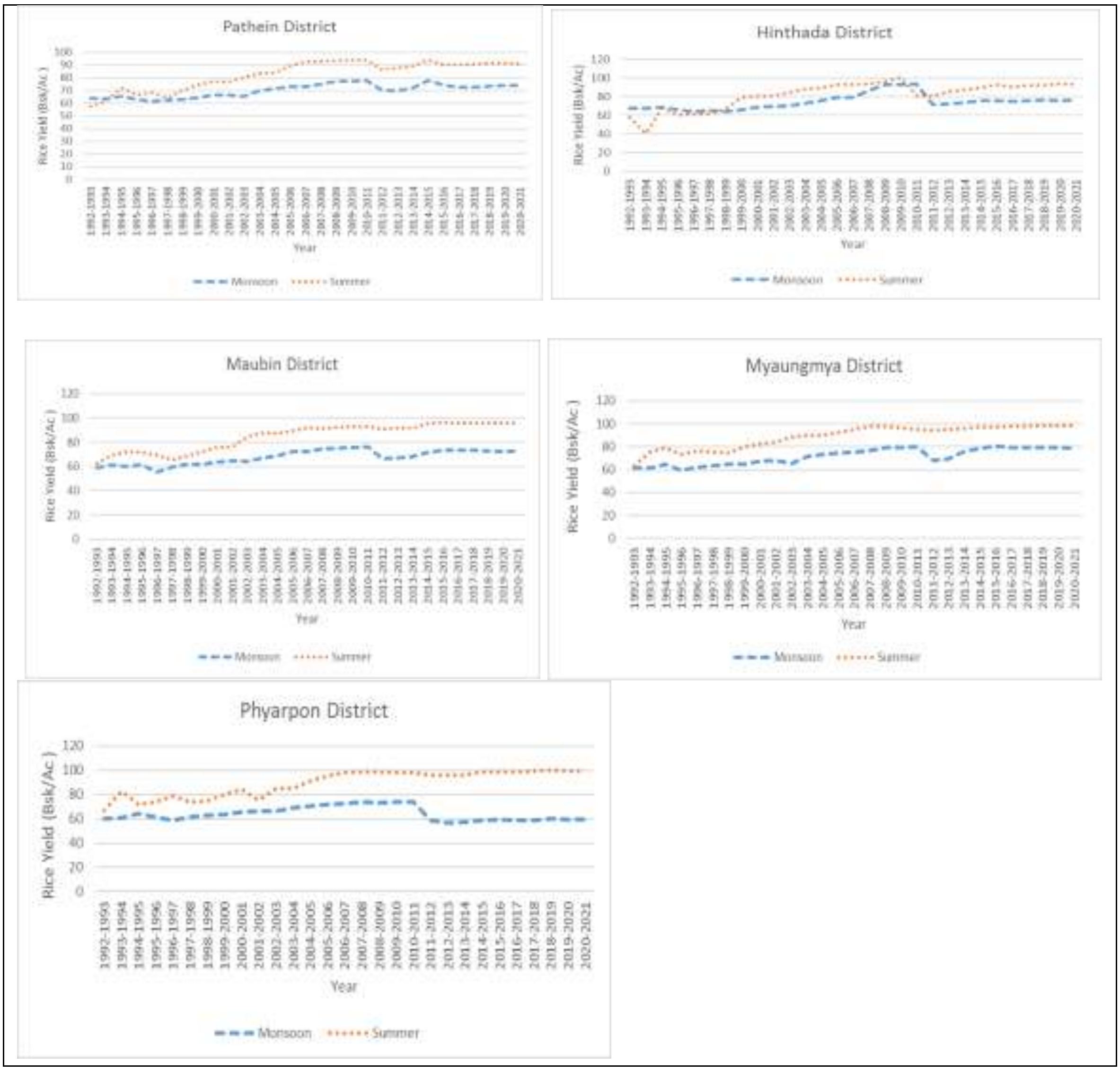


Figure (4.2): Trends of Monsoon and Summer Rice Yield for Five Districts in Ayeyawady Region

Source: Department of Agriculture (Ayeyawady Region)

4.3 Patterns of Climatic Variables

The patterns of each of the climatic variables (Rainfall, Maximum Temperature, Minimum Temperature, Morning Relative Humidity and Evening Relative Humidity) in all districts are described in Figures (4.3) to (4.7) respectively. The standard unit of these climatic variables was rainfall in millimeter (mm), the temperature in degree centigrade (˚C), and the relative humidity in percentage (%). In this study, the monthly weather data were converted into the average of May to October for monsoon and an average of November to April for summer.

Figure (4.3): Patterns of Rainfall in Ayeyawady Region

Source: Department of Meteorology and Hydrology, Myanmar.

As shown in Figures (4.3) to (4.7), the rainfall fluctuates in both the monsoon and summer, but Myaungmya District received the significant highest rainfall in June, July, and August in 2001-2002. Similarly, there were quite high fluctuations in the maximum temperature and morning relative humidity. The annual minimum temperature dropped from 1994–1995 to 2000–2001 in Myaungmya District, from 2001–2002 to 2003–2004 in Phyarpon District and from 2005–2006 to 2006–2007 in Maubin District, but the rest of the years are usually fluctuating. Moreover, the evening relative humidity was almost nonexistent and fluctuated from 2009-2010 to 2019-2020.

Figure (4.4): Patterns of Maximum Temperature in Ayeyawady Region

Source: Department of Meteorology and Hydrology, Myanmar

There were quite high fluctuations of maximum temperature in all districts. Whereas, there were a slight fluctuation of minimum temperature over the study period in Pathein, Hinthada and Phyarpon districts comparing with Maubin and Myaungmya Districts.

Figure (4.5): Patterns of Minimum Temperature in Ayeyawady Region

Source: Department of Meteorology and Hydrology, Myanmar

In Figure (4.5), it seems there might be some confusion in terms of seasons and temperature patterns. In many regions, summer is typically the warmest season, characterized by higher temperatures. Conversely, winter is generally the coldest season, with lower temperatures. On the other hand, the monsoon season usually occurs in certain regions with distinct wet and dry periods. It is characterized by heavy rainfall, and the associated cloud cover and moisture often lead to lower temperatures compared to the summer season. However, monsoon seasons can vary in timing and duration depending on the location. In this study, the coolest periods would indeed be during December, January and February, which aligns with the summer season. During this time, the minimum temperatures are likely to be lower compared to other months of the year.

Figure (4.6): Patterns of Morning Relative Humidity in Ayeyawady Region

Source: Department of Meteorology and Hydrology, Myanmar

Figure (4.6) also shows that there were little fluctuations of morning relative humidity in all districts except Pathein and Phyarpon. By the Figure (4.7), there were almost similar trend of evening relative humidity without having any significant fluctuations in monsoon in all the districts over the study period. However, there were fairly distinct fluctuations of evening relative humidity in summer in almost all of the districts.

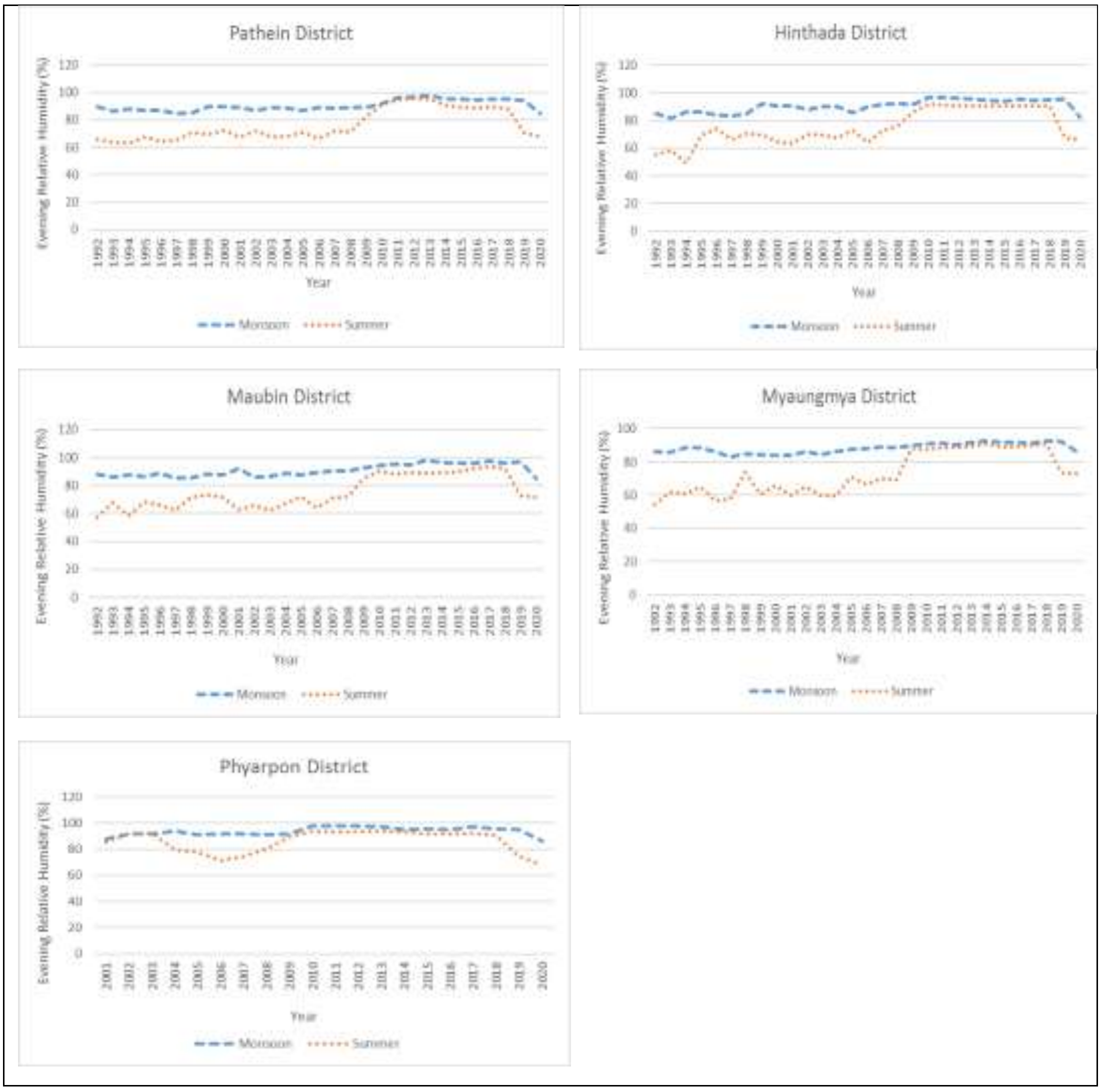


Figure (4.7): Patterns of Evening Relative Humidity in Ayeyawady Region

Source: Department of Meteorology and Hydrology, Myanmar

Overall, rainfall in all districts falls between 0 and 100 millimeters during the monsoon and between 300 and 500 millimeters throughout the summer. In general, maximum temperatures in all districts falls between 30˚C and 35˚C during the monsoon and the summer. The minimum temperatures in all districts falls between 10˚C and 20˚C during the monsoon and the summer. Around 80% of the air is humid. These conditions are typical.

4.4 Descriptive Statistics

The descriptive statistics including maximum, minimum, mean, median, range, standard deviation (to explore the fluctuation of the variable), coefficient of skewness and kurtosis (to explain about the shape of the curve) were computed for all the selected climatic variables and rice yield from each of the five districts. The descriptive statistics for four districts such as Pathein, Hinthada, Maubin and Myaungmya were presented for the 29 year periods starting from 1992-1993 to 2020-2021. For Phyarpon District, the descriptive statistics were described for the 20 years, from 2001-2002 to 2020-2021. These values are presented in Table (4.1). Each of these districts and variables for both monsoon and summer are interpreted as follows:

The average rainfall in Pathein District is 478.16 mm, with a minimum of 301.33 mm and a maximum of 586.83 mm. The range indicates significant variability in rainfall. The maximum and minimum temperatures in Pathein District have a mean of 31.65 ˚C and 23.94 ˚C, respectively. The maximum temperature ranges from 30.36˚C to 33.00˚C, while the minimum temperature ranges from 17.88 ˚C to 22.38 ˚C. The mean morning relative humidity (MRH) is 87.49%, while the mean evening rate humidity (ERH) is 75.40%. The relative humidity ranges from 68.46% to 90.50% for MRH and from 63.41% to 80.05% for ERH. The average rice yield is 70.00 (Bsk/Ac), with a minimum of 57.73 (Bsk/Ac) and a maximum of 94.22 (Bsk/Ac). The data generally show slightly negative skewness and moderate kurtosis, indicating that the distributions are somewhat close to normal but have slightly heavy tails. Based on this results, it's evident that Pathein District experiences a range of climatic conditions, including variable rainfall and temperature, along with fluctuations in humidity. The distribution of data is slightly skewed.

Table (4.1): Descriptive Statistics for Each Variable by District

Pathein District												
	Rainfall(mm)		**MaxTemp(˚C)**		**MinTemp(˚C)**		**MRH(%)**		**ERH(%)**		**Riceyield(Bsk/Ac)**	
	MS	**SM**	**MS**	**SM**	**MS**	**SM**	**MS**	**SM**	**MS**	**SM**	**MS**	**SM**
Min	301.33	0.17	30.36	32.23	20.19	17.88	82.98	68.46	84.67	63.41	61.57	57.73
Median	487.00	19.00	31.52	34.00	24.33	20.83	87.67	75.32	88.99	71.08	71.50	87.79
Mean	478.16	20.72	31.65	34.05	23.94	20.65	87.49	75.40	90.34	75.90	70.00	82.63
Max	586.83	64.50	33.00	35.62	25.33	22.38	90.50	80.05	97.67	95.67	77.65	94.22
Std	69.37	16.95	0.63	0.96	1.32	0.94	1.80	3.14	3.91	11.25	5.19	11.22
range	285.50	64.33	2.64	3.39	5.15	4.50	7.52	11.60	13.00	32.25	16.08	36.49
Skewness	-0.60	0.68	0.11	-0.18	-1.65	-1.33	-0.48	-0.47	0.44	0.63	-0.15	-0.76
Kurtosis	2.87	2.83	2.86	2.10	4.99	4.77	2.83	2.50	1.91	1.74	1.69	2.23
Hinthada District												
	Rainfall(mm)		MaxTemp(˚C)		MinTemp(˚C)		MRH(%)		ERH(%)		Riceyield(Bsk/Ac)	
	MS	SM	MS	SM	MS	SM	MS	SM	MS	SM	MS	SM
Min	262.33	1.00	31.27	33.07	21.32	16.87	80.56	61.78	81.79	49.14	63.79	40.37
Median	360.83	11.17	32.18	34.35	24.17	17.78	86.44	70.60	90.76	70.45	73.33	87.60
Mean	360.97	13.94	32.25	34.33	23.95	17.97	86.16	70.47	90.28	74.85	74.34	81.82
Max	486.00	73.17	33.59	36.19	25.05	20.70	90.79	78.74	96.50	91.75	93.59	100.42
Std	57.99	14.96	0.60	0.73	0.85	0.82	2.51	3.93	4.65	12.60	8.35	14.42
range	223.67	72.17	2.32	3.12	3.72	3.83	10.23	16.97	14.71	42.61	29.80	60.05
Skewness	0.23	2.27	0.47	0.13	-1.21	1.29	-0.40	-0.19	-0.33	0.04	1.05	-1.16
Kurtosis	2.39	9.45	2.66	2.90	4.34	5.31	2.81	3.05	1.83	1.92	3.47	3.61

Source: Department of Meteorology and Hydrology, Myanmar and Department of Agriculture (Ayeyawady Region)

Table (4.1): Descriptive Statistics for Each Variable by District (Contd.)

Maubin District												
	Rainfall(mm)		**MaxTemp(˚C)**		**MinTemp(˚C)**		**MRH(%)**		**ERH(%)**		**Riceyield(Bsk/Ac)**	
	MS	SM	**MS**	**SM**	**MS**	**SM**	**MS**	**SM**	**MS**	**SM**	**MS**	**SM**
Min	261.00	0.33	26.81	31.95	15.73	11.90	70.50	71.83	71.00	57.17	55.67	62.56
Median	410.83	15.34	31.40	33.05	23.92	19.23	87.83	77.33	89.33	71.80	67.99	90.60
Mean	409.98	18.01	31.28	33.29	23.28	18.77	87.01	77.30	90.26	75.05	67.87	84.88
Max	521.17	59.00	32.85	35.98	25.18	21.18	93.50	83.83	97.93	93.58	76.29	96.23
Std	62.34	15.19	1.06	1.06	2.20	2.01	3.99	3.22	5.65	11.71	6.02	11.27
range	207.97	58.67	2.84	4.04	9.50	9.28	9.83	12.00	13.01	36.41	20.62	33.67
Skewness	-0.26	1.10	-2.6	0.83	-1.91	-1.82	-2.37	0.18	-1.12	0.31	-0.29	-0.62
Kurtosis	2.51	3.51	12.39	2.95	6.40	6.40	11.27	2.26	5.60	1.64	1.80	1.80
Myaungmya District												
	Rainfall(mm)		MaxTemp(˚C)		MinTemp(˚C)		MRH(%)		ERH(%)		Riceyield(Bsk/Ac)	
	MS	SM	MS	SM	MS	SM	MS	SM	MS	SM	MS	SM
Min	320.50	0.00	30.58	27.62	15.97	12.55	78.39	52.08	83.05	50.36	59.61	63.56
Median	457.67	11.67	31.49	33.34	22.80	19.13	85.47	70.55	88.33	70.03	73.19	94.20
Mean	470.03	15.58	31.69	33.05	21.58	18.11	84.26	72.12	88.22	72.63	71.72	88.85
Max	1143.83	46.50	33.10	34.42	24.33	21.99	87.30	86.39	92.73	90.78	80.68	98.15
Std	144.02	12.29	0.72	1.25	2.65	2.59	2.49	7.31	2.99	13.37	7.08	10.04
range	823.33	46.50	2.52	2.55	8.36	9.44	8.91	24.75	9.68	36.11	21.07	34.59
Skewness	3.60	1.10	0.42	-2.91	-0.88	-0.80	-0.96	-0.35	-0.01	0.17	-0.23	-0.84
Kurtosis	17.86	3.26	2.11	13.65	2.29	2.48	2.64	3.31	1.69	1.56	1.51	2.49

Source: Department of Meteorology and Hydrology, Myanmar and Department of Agriculture (Ayeyawady Region)

Table (4.1): Descriptive Statistics for Each Variable By District (Contd.)

Phyarpon District												
	Rainfall(mm)		**MaxTemp(˚C)**		**MinTemp(˚C)**		**MRH(%)**		**ERH(%)**		**Riceyield(Bsk/Ac)**	
	MS	**SM**	**MS**	**SM**	**MS**	**SM**	**MS**	**SM**	**MS**	**SM**	**MS**	**SM**
Min	365.33	0.17	30.29	30.64	17.64	13.46	85.17	67.55	86.33	67.98	56.26	75.56
Median	508.92	15.25	30.77	33.13	24.61	21.12	87.99	75.50	94.28	90.99	62.98	98.11
Mean	500.34	21.67	31.02	33.03	23.39	19.92	88.05	75.48	93.64	85.92	64.86	95.25
Max	605.83	72.50	31.99	33.88	25.63	22.94	90.34	86.33	98.04	93.82	73.88	99.82
Std	58.80	19.88	0.57	0.77	2.56	2.79	1.53	3.71	3.30	8.67	6.81	6.43
range	240.50	72.34	1.70	3.24	7.99	9.48	5.17	18.78	11.71	25.84	17.62	24.26
Skewness	-0.42	1.05	0.36	-1.57	-1.47	-1.24	-0.20	0.74	-0.45	-0.77	0.21	-1.91
Kurtosis	3.09	3.19	1.60	5.72	3.70	3.31	2.14	5.48	2.49	2.06	1.30	5.74

Source: Department of Meteorology and Hydrology, Myanmar and Department of Agriculture (Ayeyawady Region)

In Hinthada District, the mean values of RF, MaxT, MinT, MRH, ERH and rice yield in both seasons are 360.97mm and 13.94mm, 32.25˚C and 34.33˚C, 23.95˚C and 17.97˚C, 86.16% and 70.47%, 90.28% and 74.85%, 74.34 (Bsk/Ac) and 81.82 (Bsk/Ac). The range of rainfall values is spanning 223.67 mm from the minimum to the maximum. This district also exhibits fluctuations in temperature, with the MaxT from 31.27˚C to 33.59˚C and the MinT ranging from 21.32˚C to 25.05˚C. The relative humidity levels at sunrise and sunset show variability, with the median MRH at around 86% and the median ERH at approximately 91%. The MaxT and ERH, show slightly positively skewed distributions. Additionally, some variables have higher kurtosis values, indicating peaked distributions.

Maubin District experiences varying rainfall patterns, with the median and mean rainfall around 410 mm. The highest recorded rainfall is 521 mm, and the lowest is 261 mm. The fluctuating temperature and humidity levels indicate dynamic climatic patterns that can influence agricultural practices and crop growth. The Higher MaxT suggest warm daytime conditions, while lower MinT imply cooler nights. The median value of ERH is approximately 71%, suggesting lower humidity levels during the evenings compared to mornings. The variation in relative humidity between morning and evening reflects diurnal changes in atmospheric moisture, which can impact plant growth and water availability. Rice yield data shows significant variation in rice production in the district, with yields ranging from 55.67 to 96.23 (Bsk/Ac). There are some observations with extreme rice yields, and the distribution may not be as spread out as a normal distribution.

In Maungmya District, the mean values of RF, MaxT, MinT, MRH, ERH and rice yield in both seasons are 457.67mm and 15.58mm, 31.69˚C and 33.05 ˚C, 21.58˚C and 18.11˚C, 84.26% and 72.12%, 88.22% and 72.63%, 71.72 (Bsk/Ac) and 88.85 (Bsk/Ac). The rainfall data shows a wide range of variability. This range suggests that diverse weather patterns and potential implications for agriculture and water management. It is experienced that the variations in both maximum and minimum temperatures as well as fluctuations in relative humidity levels during morning and evening. The rice yield data shows varying rice production levels, ranging from 59.61 to 98.15 (Bsk/Ac). The rainfall data indicates the occasional occurrence of heavy rainfall events, contributing to the overall variability in the rainfall distribution.

In Phyarpon District, the mean values of RF, MaxT, MinT, MRH, ERH and rice yield in both seasons are 500.34mm and 21.67mm, 31.02˚C and 33.03 ˚C, 23.39˚C and 19.92˚C, 88.05% and 75.48%, 90.64% and 85.92%, 64.86 (Bsk/Ac) and 95.25 (Bsk/Ac). It experiences varying levels of rainfall, temperature, and relative humidity, which can impact agricultural practices and crop growth. There are moderate variation in all variables. The rice yield indicates different levels of rice productivity in the district. The distributions have skewed and peaked than a normal distribution.

4.5 Fitting the Multiple Linear Regression (MLR) Model

The rice yield forecasting model based on multiple linear regression approach has been used in this study was specified as in Equation (3.1). The OLS method is employed to identify the impacts of climate variation on the rice yield (Y_t). The results of the rice yield model for each district were presented in Table (4.2).

The variation of rice yield against changes in climate variables is shown by the sign of the coefficients in Table (4.2). Even though there are significant correlations of rice yield with a few number of climate variables in the regression results, these coefficients can be used to determine the impact of climatic variables on the change in rice yield that was taken into account for this study.

The regression model of Pathein District is a good fit of the data because the F-statistic is significant. All variables are statistically significant except Minimum Temperature (MinT). There is a linear relationship between the response variable of rice yield and each of the predictor variables such as RF, MaxT, MinT, MRH and ERH. Moreover, 74.08% of the variation in the rice yield is explained by predictors that actually affect the rice yield. Continuously, the assumptions are checked which are met or not. The residuals are normally distributed because the closer the dots lie to the diagonal line in Normal Q-Q plot of regression standardized residual and Shapiro-Wilk test is not significant. Since the scatterplot of regression standardized residual and regression standardized predicted value generally appears more random than funneled and the Breusch Pagan test is not significant, the variance of the residuals is constant. The residuals are positively correlated. The predictors have a slight multicollinearity problem because VIF scores are above 10 in Rainfall (RF) and Morning Relative Humidity (MRH), and their tolerance scores are below 0.2.

Table (4.2): Relationship between Rice Yield and Climate Variables for Each District

Districts	Pathein		Hinthada		Maubin		Myaungmya		Phyarpon	
	Coef	**VIF**	**Coef**	**VIF**	**Coef**	**VIF**	**Coef**	**VIF**	**Coef**	**VIF**
(Constant)	-198.65***		100.067		-0.664		-34.57		168.27**	
Rf	-0.027**	11.12	-0.007	12.86	-0.049***	9.82	-0.029***	2.99	-0.075***	8.16
MaxT	5.721***	3.73	0.054	6.81	1.369	2.23	3.296***	1.65	-2.297	3.99
MinT	-0.983	2.94	0.579	10.05	-1.278***	2.48	0.423	1.79	0.869	2.06
MRH	1.192***	10.66	-1.252**	14.47	0.268	4.87	-0.321	4.23	0.066	8.36
ERH	0.223**	2.78	0.769***	3.16	0.584***	2.14	0.386**	3.59	-0.216	1.42
Adj-R^2	0.7408		0.3064		0.7121		0.5689		0.8082	
Durbin-Watson	1.3493***		0.7580***		1.0805***		0.9153***		2.5709*	
F-statistic	33.58***		6.036***		29.2***		16.05***		33.88***	

*** Significant at the 0.01 level, ** significant at the 0.05 level, and *significant at the 0.1 level

Source: Own Computation

It indicated that 30.64% of the variation of the model is explained by some predictors that actually affect the rice yield in Hinthada District. The Morning and Evening Relative Humidity variables are statistically significant. There is a linear relationship between the response variable and each of the predictor variables. The residuals are correlated and not normally distributed. But, the variance of the residuals is constant. The predictors have a slight multicollinearity. But, the regression model is a good fit of the data because the F-statistic is significant.

In Maubin District, the regression model is a good fit of the data because the F-statistic is significant. Moreover, 71.21% of the variation of the rice yield is explained by some predictors such as RF, MinT, and ERH that actually affect the rice yield. The RF, MinT and ERH variables are statistically significant. The residuals are normally distributed and have a serial correlation. Since the scatterplot of the regression standardized residuals and the regression standardized predicted values generally appears random, the variation in the residuals is roughly similar. The predictors are not highly correlated because there is no multicollinearity problem.

In Myaungmya District, the RF, MaxT and ERH variables are statistically significant. Moreover, 56.89% of the variation of the model is explained by some predictors including RF, MaxT and ERH that actually affect the rice yield. The residuals are approximately normal distributed and the variance is not constant. The predictors are not highly correlated. However, the regression model is a good fit for the data.

In Phyarpon District, 80.82% of the variation of the model is explained by some predictors that actually affect the rice yield. But, all variables are not statistically significant except RF. The residuals are normal distributed. Moreover, the variance of the residuals is not constant. The predictors are not highly correlated, and the regression model fits the data.

Overall, it is seen that some of the predictors do not have a significant effect in all districts and these are correlated in some models. The residuals in some districts are not normally distributed and are not independent. That is, some of the assumptions are not met in some districts.

Besides, the multiple linear regression models with the significant predictor variables are fitted for each district. The results are described in Appendix A-VI. By these results, the values of R^2, Adjusted R^2, RMSE and MAPE seem to have barely changed and the signs of the effect on rice yield have not been changed. Moreover, some assumptions are met.

Continuously, the Autoregressive Integrated Moving Average with predictors (ARIMAX) and Seasonal ARIMAX (SARIMAX) model as the alternative way are fitted. Firstly, the variables used in this study are checked to be the stationary or not.

4.6 Stationarity Test

Each variable included in the study should be required to validate a zero degree of integration because the current study focuses on a model to examine the effects of climate variation on rice yield. If the variables characterize distinct levels of integration, they cannot be used for correlation, causation, or OLS estimations. To ensure that all results are accurate and all estimations are consistent, it is first necessary to ensure that the data series is stationary. Stationary refers to the property of a time series where its statistical properties, such as mean and variance, remain constant over time. The Augmented Dickey-Fuller (ADF) test is a statistical test used to assess the stationarity of time series data. The findings of the (ADF) test are shown in Table (4.3).

Table (4.3): Results of Augmented Dickey-Fuller Test

Variable	Integrated Order				
	Pathein	**Hinthada**	**Maubin**	**Myaungmya**	**Phyarpon**
Rice Yield	I(1)	I(1)	I(1)	I(1)	I(1)
Rainfall	I(0)	I(0)	I(0)	I(0)	I(1)
Maximum Temperature	I(0)	I(1)	I(1)	I(0)	I(1)
Minimum Temperature	I(1)	I(0)	I(0)	I(1)	I(1)
Morning Relative Humidity	I(1)	I(0)	I(1)	I(1)	I(1)
Evening Relative Humidity	I(1)	I(1)	I(1)	I(1)	I(1)

Source: Department of Meteorology and Hydrology, Myanmar and Department of Agriculture (Ayeyawady Region)

As shown in Table (4.3), rice yield and evening relative humidity in all districts exhibit an integrated order of I(1), indicating that it requires differencing once to achieve stationarity. The remaining climate variables are integrated with orders of zero and one, or I(0) and I(1), respectively. An integrated order of I(0), indicating that it is already stationary. At 5%, each variable is significant. After stationarity in all variables is confirmed, comprehensive regression models are running in the next sections.

4.7 Fitting the Autoregressive Integrated Moving Average with Predictors (ARIMAX) Model and Seasonal Autoregressive Integrated Moving Average with Predictors (SARIMAX) Model

The seasonal autoregressive integrated moving average with predictors (SARMAX) model given in Equation (3.41) and the autoregressive integrated moving average with predictors (ARMAX) model given in Equation (3.44) are fitted to represent the rice yield in five districts of Ayeyawady Region.

This study used the ADF test and a correlogram with ACF and PACF to determine whether a particular series is stationary and to identify the proposed models. In order to construct the appropriate SARIMAX model for rice yield in each district, the autoregressive (p) and moving average (q) parameters have to be effectively determined for an effective model. This appropriated model has a relatively small Akaike Information Criterion (AIC). In addition, the Ljung-Box (Q) statistic and residuals correlogram were used to assess the model's suitability and compliance with the assumptions.

In Pathein district, the series of rice yield from 1992 to 2020 with the correlogram is described in Figure (4.8).

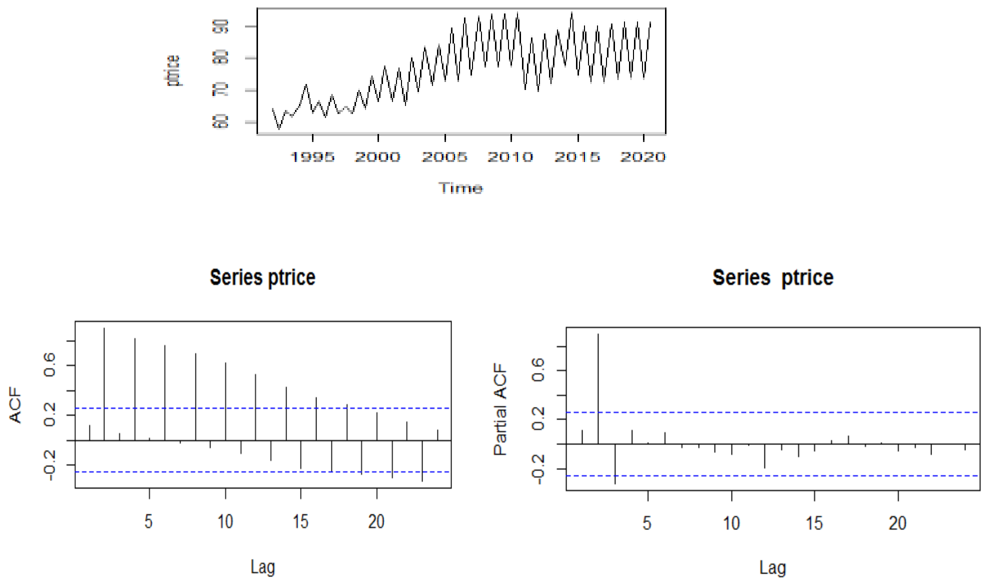


Figure (4.8): Original Pattern of Rice Yield with Correlogram in Pathein District

Source: Own Computation

The series is nonstationary, with trend and seasonality. Firstly, the data are taking the first seasonal differencing to make the data stationary. After the data are seasonal differentiated, it does not have trend and it turns out the data are stationary. . The best model is reported as ARIMAX (1,1,1) with AIC is 303.7 by the auto.arima method. This model seems to have the lowest AIC compared to the other models. The correlogram of the stationary series is described in Figure (4.9).

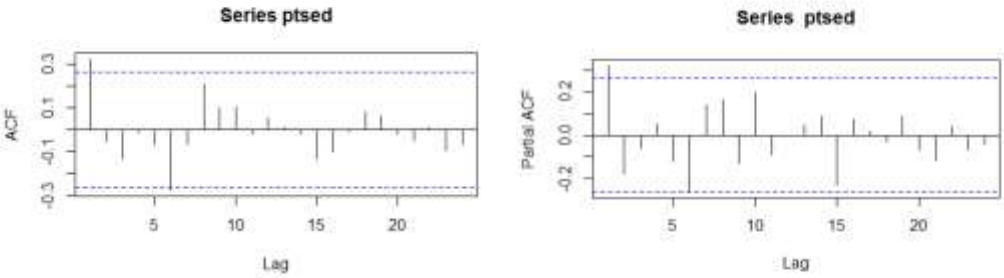


Figure (4.9): The Correlogram with ACF and PACF of the Stationary Series

Source: Own Computation

This is an ARIMA process with predictors because both the sample ACF and PACF have a similar pattern. In Figure (4.9), both the sample ACF and sample PACF have spike at lag1 and seasonal at lag3. Thus, the tentative model is SARIMAX(1,0,1)(3,1,3)$_2$ with AIC is 286.88. But, some of the parameters in this model are not significant. Therefore, the SARIMAX(0,0,1)(1,1,3)$_2$, the tentative model which the parameters are significant, is selected to represent the data series. The parameter estimation of the model is described in Table (4.4) with an AIC is 281.55.

Table (4.4): Estimation of SARIMAX (0,0,1)(1,1,3)$_2$ Model in Pathein District

Parameter	Estimate	S.E.	z-value	sig
c	9.2265	8.4798	1.0881	0.2765
θ_1	0.4510***	0.1230	3.6675	0.0245e^{-02}
Φ_1	-0.6414***	0.1539	-4.1667	0.3090e^{-04}
Θ_1	1.1389***	0.2248	5.0669	0.4044e^{-06}
Θ_2	0.5523**	0.2793	1.9772	0.0480
Θ_3	-0.3092*	0.1857	-1.6649	0.0959
β_1(RF)	-0.0015	0.0029	-0.5020	0.6156
β_2(MaxT)	-0.5703*	0.2975	-1.9173	0.0552
β_3(MinT)	0.0358	0.1614	0.2217	0.8246
β_4(MRH)	-0.3351**	0.1638	-2.0460	0.0408
β_5(ERH)	0.0494	0.1021	0.4837	0.6286

Note: ***,**,* represent 1%,5% and 10% level of significance

Source: Own Computation

In Table (4.4), coefficients of the non-seasonal and seasonal moving average terms at lag 1 and the seasonal autoregressive term at lag 1 are statistically significant at 1%. Moreover, the seasonal moving average terms at lag 2 and 3 are also statistically significant at 5%and 1%. Besides, MaxT and MRH variables have statistically significant negative effects on rice yield at 10% and 5%, respectively. It can be seen that rice yield decreases by 0.5703 unit and 0.3351unit, respectively when MaxT and MRH increase by 1 unit. At the same time, RF has negative effects; and MinT and ERH have positive effects on rice yield (Hamjah, 2014). Continuously, this model is checked to determine whether the assumptions are met or not.

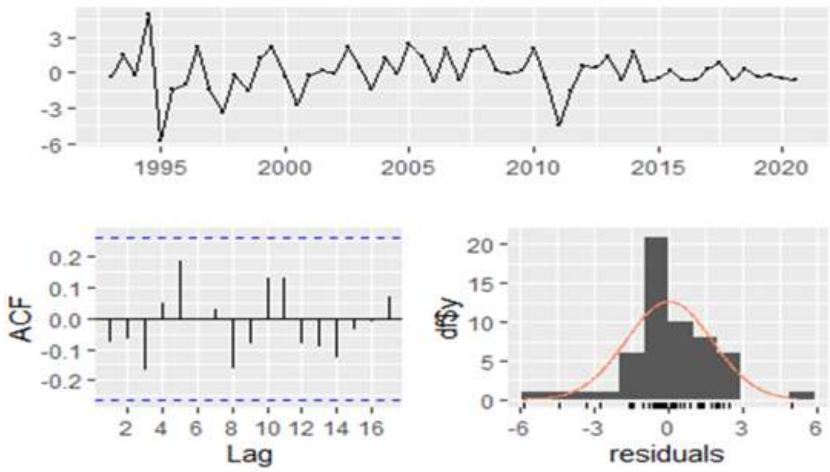


Figure (4.10): The Correlogram of the Residuals for SARIMAX(0,0,1)(1,1,3)$_2$

Source: Own Computation

The Q-statistic and correlogram of the residuals showed no significant pattern left in the ACFs and partial autocorrelation functions (PACFs) of the residuals, which implies that the residual of the selected model is white noise and normality. Moreover, the p-value of Ljung Box statistic is 0.0199. Therefore, the tentative model SARIMAX (0,0,1)(1,1,3)$_2$ is adequate. Continuously, the tentative model SARIMAX (0,0,1)(1,1,3)$_2$ with the predictors such as the significant estimated coefficients was fitted. The error value (AIC) and R squared of this model are 276.23 and 0.9571.

In Hinthada district, the rice yield from 1992 to 2020 with the correlogram is described in Figure (4.11).

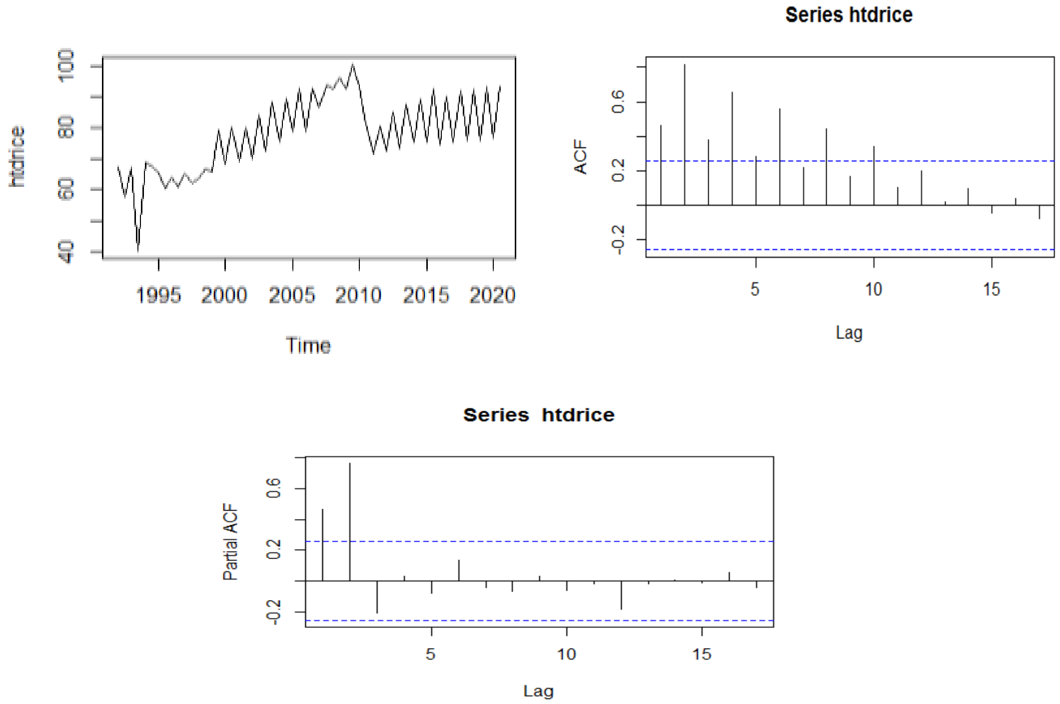


Figure (4.11): Original Pattern of Rice Yield with Correlogram in Hinthada District

Source: Own Computation

Since the rice yield data are nonstationary with seasonality, it was taking the seasonal differencing to become the data stationary. After these data are seasonal differentiated, it turns out the data are stationary. The best model is reported as ARIMAX (1,1,1) with AIC is 381.32 by the auto.arima method. This model seems to have the lowest AIC compared to the other models. The stationary series with correlogram is described in Figure (4.12).

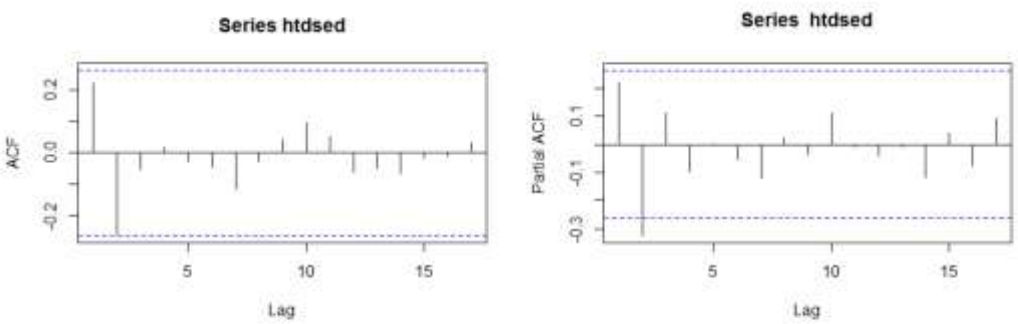


Figure (4.12): The Correlogram with ACF and PACF of the Stationary Series

Source: Own Computation

This is an ARIMA process with predictors because both ACF and PACF have a similar pattern. The sample PACF of the series has only one significant spike at lag2. Therefore, the tentative model ARIMAX $(2,0,0)(0,1,0)_2$ is selected to represent the data series. The parameter estimation of the model is described in Table (4.5) with an AIC is 370.96.

Table (4.5): Estimation of ARIMAX $(2,0,0)(0,1,0)_2$ Model in Hinthada District

Parameter	Estimate	S.E.	z-value	sig
c	29.0431	60.4338	0.4806	0.6308
ϕ_1	0.2758**	0.1249	2.2079	0.0273
ϕ_2	-0.4306***	0.1358	-3.1703	0.0015
β_1(RF)	0.0106	0.0144	0.4806	0.4626
β_2(MaxT)	-0.2623	1.5283	-0.1716	0.8637
β_3(MinT)	-0.6864	0.7306	-0.9395	0.3475
β_4(MRH)	-0.0032	0.2716	-0.0117	0.9907
β_5(ERH)	-0.0814	0.0954	-0.8528	0.3937

Note: ***,**,* represent 1%,5% and 10% level of significance

Source: Own Computation

The autoregressive terms have a p-value that is less than the significance level of 0.01 and 0.05. Thus, the coefficient for these terms is statistically significant. But, the coefficients of predictor variables do not have a statistically significant effect on rice yield. It is because the summers are short, the winters are warm and wet, and it experiences extreme seasonal variation in the perceived humidity. But, all coefficients of predictors have negative effects on rice yield except RF (Hamjah, 2014). This model is continually checked to determine whether the assumptions are met or not.

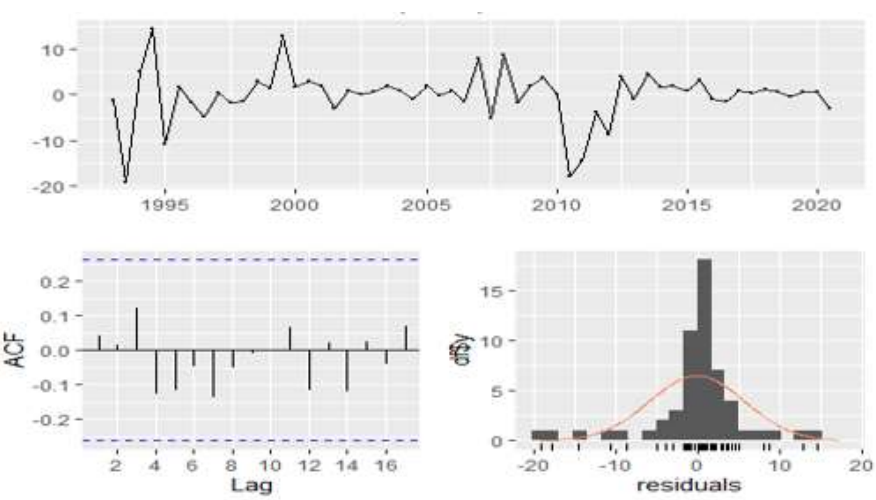
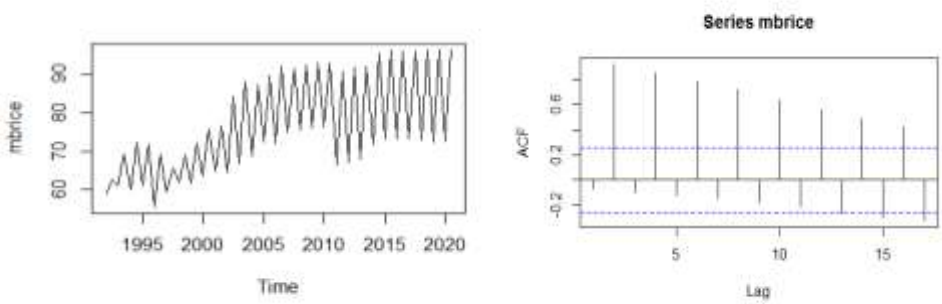


Figure (4.13): The Correlogram of the Residuals for ARIMAX (2,0,0)(0,1,0)$_2$

Source: Own Computation

The Q-statistic and correlogram of the residuals showed no significant pattern left in the ACFs and partial autocorrelation functions (PACFs) of the residuals, which implies that the residual of the selected model is white noise. Moreover, the p-value of Ljung Box statistic is 0.1926. It can be concluded that the model meets the assumptions. Therefore, the tentative model ARIMAX (2,0,0)(0,1,0)$_2$ is adequate. Continuously, the tentative model ARIMAX (2,0,0)(0,1,0)$_2$ with the predictors such as the significant estimated coefficients was fitted. The error value (AIC) and R squared of this model are 366.46 and 0.7808.

In Maubin district, the series of rice yield from 1992 to 2020 with the correlogram is described in Figure (4.14).



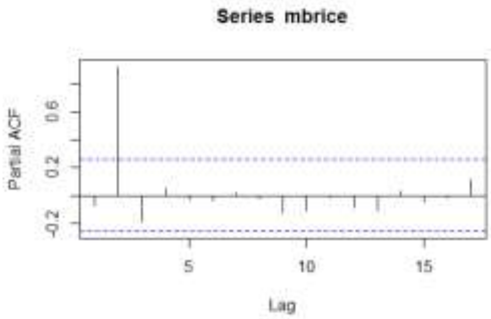


Figure (4.14): Original Pattern of Rice Yield with Correlogram in Maubin District

Source: Own Computation

Since the rice yield data are nonstationary with trend and seasonality, the data are taking the first differencing to make the data stationary. Firstly, the data are taking the first difference to make the data stationary. After the data are differentiated, it still have seasonality. Thus, it was taking seasonal differences and then the series seemed to be stationary. The best model is reported as ARIMAX (1,1,0)(1,0,0)$_2$ with AIC is 290.23 by the auto.arima method. This model seems to have the lowest AIC compared to the other models. The stationary series with correlogram is described in Figure (4.15).

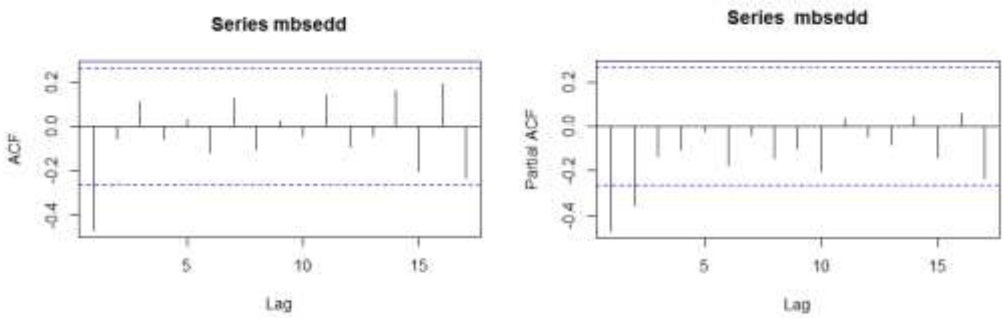


Figure (4.15): The Correlogram with ACF and PACF of the Stationary Series

Source: Own Computation

In figure (4.15), the sample PACF has spike at lag2 and the sample ACF cuts off after lag 1. Thus, the tentative model is SARIMAX (2,0,1)(0,1,0)$_2$ with AIC is 280.65. But, some of the parameters in this model are not significant. Therefore, the SARIMAX(0,1,1)(0,1,0)$_2$, the tentative model which the moving average parameter is significant, is selected to represent the data series. The parameter estimation of the model is described in Table (4.4) with an AIC is 277.45.

Table (4.6): Estimation of the ARIMAX $(0,1,1)(0,1,0)_2$ Model in Maubin District

Parameter	Estimate	S.E.	z-value	sig
c	-0.9763	1.7168	-0.5687	0.5696
θ_1	-0.9999^{***}	0.0524	-19.0710	$0.0002e^{-13}$
β_1(RF)	-0.0022	0.0042	-0.5144	0.6069
β_2(MaxT)	0.0712	0.2990	0.2381	0.8118
β_3(MinT)	0.0692	0.0534	1.2949	0.1954
β_4(MRH)	-0.0214	0.0975	-0.2193	0.8264
β_5(ERH)	-0.0266	0.0531	-0.5016	0.6159

Note: ***,**,* represent 1%,5% and 10% level of significance

Source: Own Computation

The p-value for the moving average term is under the 0.01 level of significance. As a result, the term's coefficient has a statistically significant impact. The predictor variables, however, do not statistically significant anything because the dry season is humid, partially cloudy, and hot, but the wet season is oppressive and overcast. However, MaxT and MinT have positive effects on rice yield while RF, MRH, and ERH have negative effects. This model is continually checked to determine whether the assumptions are met or not.

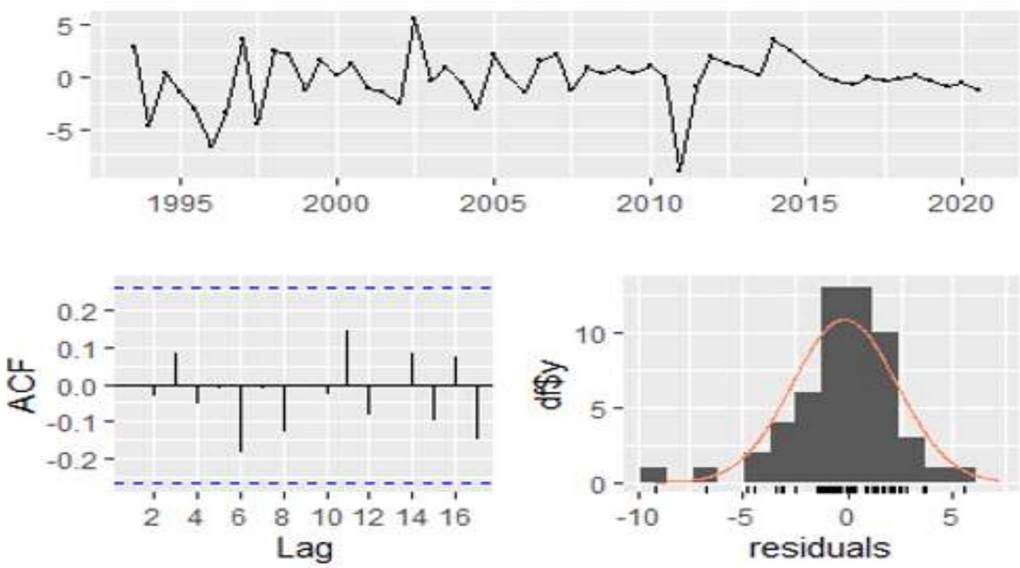


Figure (4.16): The correlogram of the residuals for ARIMAX$(0,1,1)(0,1,0)_2$

Source: Own Computation

The correlogram of the residuals showed no significant pattern left in the auto correlation functions (ACFs) and partial autocorrelation functions (PACFs) of the residuals imply that the residual of the selected model is white noise. Moreover, the model is adequate because the p-value of Ljung Box statistic is 0.1601. It can be concluded that the model meets the assumptions. Therefore, the tentative model ARIMAX $(0,1,1)(0,1,0)_2$ is adequate. Continuously, the tentative model ARIMAX $(0,1,1)(0,1,0)_2$ with the predictors such as the significant estimated coefficient was fitted. The error value (AIC) and R squared of this model are 271.49 and 0.9553.

In Myaungmya district, the rice yield from 1992 to 2020 with the correlogram is described in Figure (4.17).

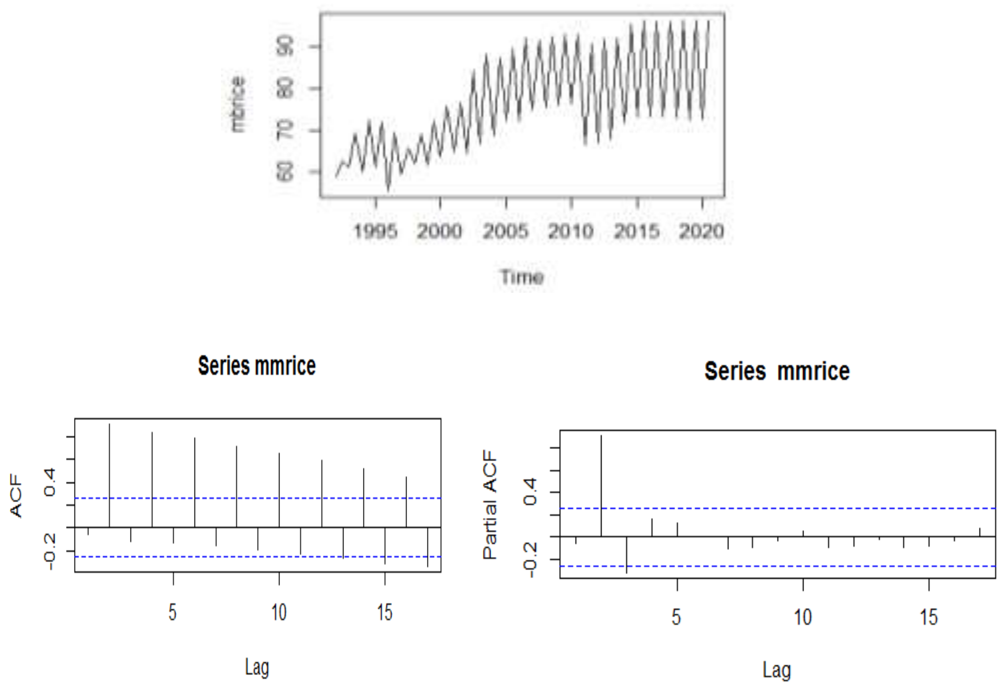


Figure (4.17): Original Pattern of Rice Yield with Correlogram in Myaungmya District

Source: Own Computation

Because of trend and seasonality, the rice yield data are nonstationary, hence, the first differencing is used to make them stationary. After the data are differentiated, it still have seasonality. Thus, it was taking seasonal differences and then the series seemed to be stationary. The best model is reported as ARIMAX (1,1,0) with AIC is 306.55 by the auto.arima method. This model seems to have the lowest AIC compared to the other models. The stationary series with correlogram is described in Figure (4.18).

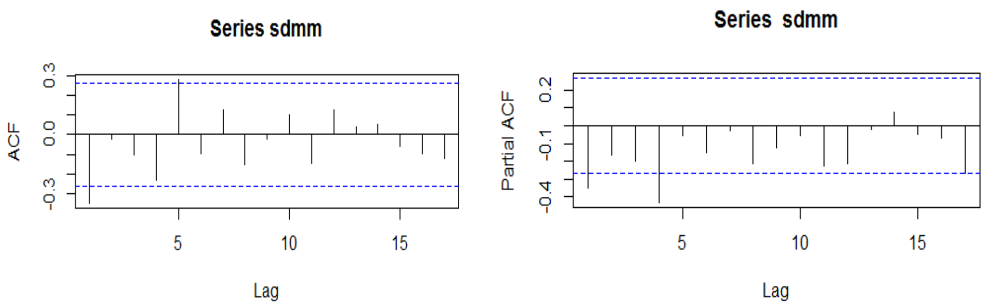


Figure (4.18): The Correlogram with ACF and PACF of the Stationary Series

Source: Own Computation

In Figure (4.18), the sample ACF has a significance spike at lag1 and the sample PACF has a spike at lag1 and seasonal spike at lag2. Thus, the tentative model is SARIMAX $(1,1,1)(2,1,0)_2$ with AIC is 294.72. But, some of the estimated parameters in this model are not significant. Therefore, the tentative model ARIMAX $(1,1,1)(0,1,0)_2$ is selected to represent the data series. The parameter estimation of the model is described in Table (4.7) with an AIC is 294.11.

Table (4.7): Estimation of the ARIMAX $(1,1,1)(0,1,0)_2$ Model in Myaungmya District

Parameter	Estimate	S.E.	z-value	sig
c	0.2041	12.4529	0.0164	0.9869
ϕ_1	0.2299*	0.1334	1.7232	0.0848
θ_1	-0.9999***	0.0986	-10.1388	0.0002e^{-13}
β_1(RF)	0.0034*	0.0019	1.8170	0.0692
β_2(MaxT)	0.2717	0.3534	0.7688	0.4420
β_3(MinT)	-0.3278	0.2446	-1.3401	0.1802
β_4(MRH)	-0.0901	0.0967	-0.9317	0.3515
β_5(ERH)	0.0391	0.0877	0.4460	0.6556

Note: ***,**,* represent 1%,5% and 10% level of significance

Source: Own Computation

The moving average term and the autoregressive term have p-values that are smaller than the significance levels of 0.1 and 0.01. Thus, the coefficients for these terms have a statistically significant effect. Moreover, the coefficient of RF is statistically significant effect on rice yield at the 10% level. It can be seen that the rice yield increases by 0.0034 unit when the RF increases by 1 unit. Particularly, this district is located in a major rainfed rice production area, and also includes the highest irrigated rice area in the Ayeyawady Region (Lar et al., 2018). At the same time, MaxT and ERH have positive effects, and MinT and MRH have negative effects on rice yield. This model is continually checked to determine whether the assumptions are met or not.

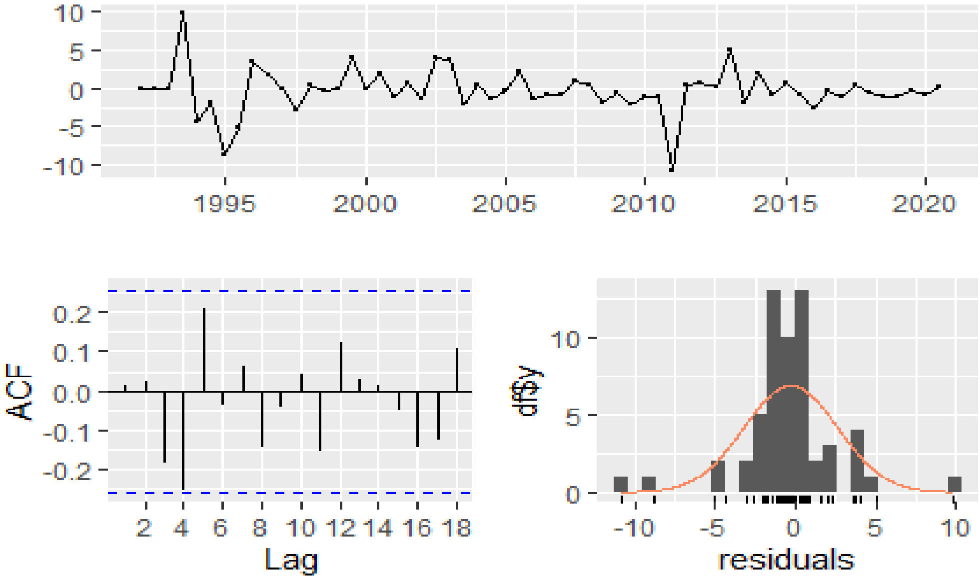


Figure (4.19): The Correlogram of the Residuals for ARIMAX(1,1,1)(0,1,0)$_2$

Source: Own Computation

The correlogram of the residuals showed no significant pattern left in the ACFs, and partial autocorrelation functions (PACFs) of the residuals imply that the residual of the selected model is white noise. Moreover, the model is adequate because the p-value of the Ljung Box statistic (0.01011) is greater than 0.01. It can be concluded that the model meets the assumptions. Therefore, the tentative model ARIMAX (1,1,1)(0,1,0)$_2$ is adequate. Continuously, the tentative model ARIMAX (1,1,1)(0,1,0)2 with the predictors such as the significant estimated coefficients such as the moving average, autoregressive and rainfall was fitted. The error value (AIC) and R squared of this model are 288.94 and 0.9411.

In Phyarpon district, the rice yield from 2001 to 2020 with the correlogram is described in Figure (4.20).

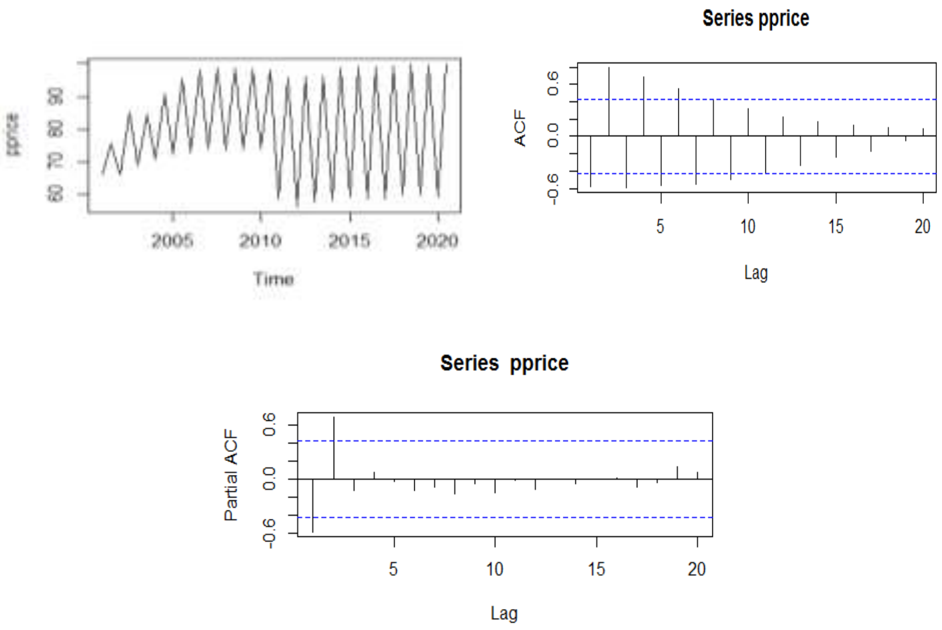


Figure (4.20): Original Pattern of Rice Yield with Correlogram in Phyarpon District

Source: Own Computation

The rice yield data are nonstationary, with trend and seasonality. Firstly, the data are taking the first differencing to make the data stationary. After the data are differentiated, it still has seasonality. Secondly, it was taking seasonal differences. Then, it does not have seasonality and it turns out the data are stationary. The best model is reported as ARIMAX (0,0,0)(0,1,0) with AIC is 209.46 by the auto.arima method. This model seems to have the lowest AIC compared to the other models. The correlogram of the stationary series is described in Figure (4.21).

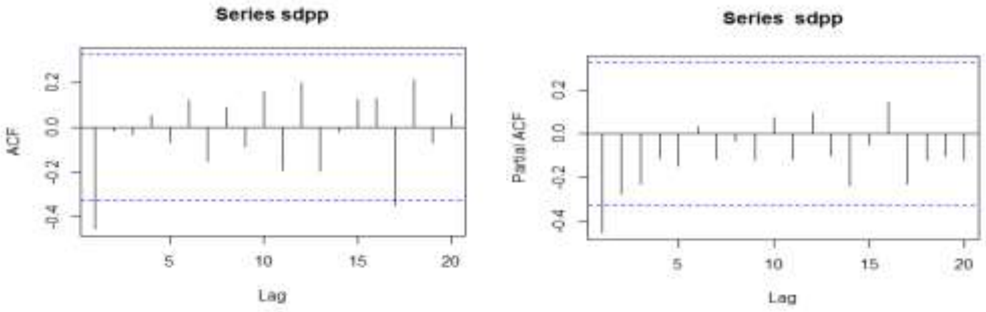


Figure (4.21): The Correlogram with ACF and PACF of the Stationary Series

Source: Own Computation

In Figure (4.21), the sample PACFs of the data series are exponentially decaying and the sample ACFs of the series have a significant spike for non-seasonal at lag1. Therefore, the tentative model ARIMAX $(0,1,1)(0,1,0)_2$ is selected to represent the data series. The parameter estimation of the model is described in Table (4.8) with AIC is 185.58.

Table (4.8): Estimation of ARIMAX$(0,1,1)(0,1,0)_2$ Model in Phyarpon District

Parameter	Estimate	S.E.	z-value	sig
c	-54.9142^{***}	20.0300	-2.7416	0.0061
θ_1	-0.9997^{***}	0.0708	-14.1255	$0.0022e^{-14}$
β_1(RF)	-0.0090^{***}	0.0035	-2.6008	0.0093
β_2(MaxT)	0.2681	0.3681	0.7284	0.4664
β_3(MinT)	-0.2055^{***}	0.0785	-2.6192	0.0088
β_4(MRH)	0.6130^{***}	0.1834	3.3425	0.0008
β_5(ERH)	0.0336	0.0376	0.8943	0.3712

Note: ***,**,* represent 1%,5% and 10% level of significance

Source: Own Computation

The non-seasonal moving average term has a p-value that is smaller than the significance level of 0.01. Consequently, the coefficient for the moving average term is statistically significant. Additionally, the predictor variables such as rainfall, minimum temperature, and morning relative humidity have a statistically significant effect at 1%. While RF and MinT have negative effects, and MRH has positive effects on rice yield. Both MaxT and ERH have positive effects on rice yield, although there are no significant effects. Besides, the deterministic term of the model is statistically significance at 1% level. This model is continually checked to determine whether the assumptions are met or not.

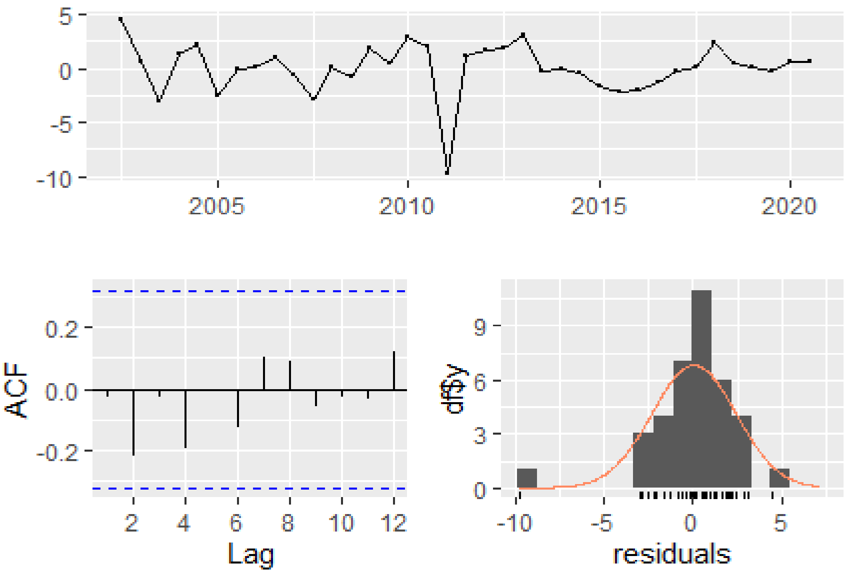


Figure (4.22): The Correlogram of the Residuals for ARIMAX (0,1,1)(0,1,0)$_2$

Source: Own Computation

The correlogram of the residuals revealed no discernible pattern in the residuals' ACFs and PACFs, indicating that the residual of the chosen model is white noise. The model is also sufficient because the Ljung Box statistic's p-value (0.292) exceeds the significance threshold of 5%. As a result, the tentative model ARIMAX (0,1,1)(0,1,0)$_2$ is adequate. Continuously, the tentative model ARIMAX (0,1,1)(0,1,0)$_2$ with the predictors such as the significant estimated coefficients such as the moving average, rainfall, minimum temperature and morning relative humidity was fitted. The error value (AIC) and R squared of this model are 187.97 and 0.6004. But, the AIC of the model included not significant estimated parameters has the lowest. Thus, this model is selected to represent the Phyarpon District.

Overall, the some of the predictors have a significant effect on rice yield in Pathein, Myaungmya and Phyarpon Districts. In the Hinthada and Maubin Districts, although none of the climatic variables significantly affect rice yield, it's essential to keep in mind that the rice yield may be subject to analyze the effect of these climate variables. Moreover, the assumptions are met in all districts such as the residuals in all models are white noise and normally distributed. Moreover, the VAR models are designed to analyze as the alternative way of assessing the effect of predictors.

4.8 Fitting the Vector Autoregressive (VAR) Model

The VAR approach sidesteps the need for structural modeling by treating every endogenous variable in the system as a function of the lagged values of all of the endogenous variables in the system. In the study, general equation including climatic variables and rice yield is:

Rice yield= f (RF, MaxT, MinT, MHR, ERH)

Johansen's procedure of multivariate cointegration requires the existence of sufficient number of time lags because the accuracy of forecasts from VAR models highly depends on selecting the true lag lengths. The choice of an appropriate order p for the estimates of the VAR model was determined using the AIC, BIC and HQ criterion.

Pathein District

Firstly, VAR model for rice yield of Pathein district was fitted in the study. The AIC, BIC (SC) and HQ criterion are computed to choose the appropriate model. Based on the result of Table (4.9), the lowest value of HQ criterion at lag 1 is 31.8318, the lowest value of SC criterion at lag 1 is 32.7822, and the lowest value of the AIC criterion at lag 1 is 31.2352. The VAR model of order p=1 is the appropriate model to represent the rice yield of Pathein District because the criteria are the lowest at lag 1. Therefore, the Vector Autoregressive model of order one, VAR(1) to the six variables was fitted and the estimation results are presented in Table (4.10).

Table (4.9): VAR Lags Order Selection Criteria for Pathein District

Lag	AIC	SC	HQ
0	37.4567	37.6777	37.5419
1	31.2352*	32.7822*	31.8318*
2	31.4417	34.3147	32.5497
3	31.3923	35.5912	33.0117

* indicates lag order selected by the criterion

Source: Own Computation

Table (4.10): VAR Estimation Results for Pathein District

	D(RY)	RF	MAXT	D(MINT)	D(MRH)	D(ERH)
D(RY(-1))	-0.7882***	5.1357	-0.0277***	0.0038	-0.0101	-0.0307
RF(-1)	0.0102	-0.0995***	0.0069*	-0.0046	-0.0095	-0.0125
MAXT(-1)	-0.5899	6.6479	0.4863***	0.0455	0.1073	-0.4797
D(MINT(-1))	0.3991	4.1897	-0.0661***	-0.2658*	-0.0169	0.4037
D(MRH(-1))	-0.0559	-11.1019***	-0.0233	-0.0817	-0.7722***	-0.0727
D(ERH(-1))	-0.0529	-0.7883	0.0252	-0.0151	-0.0190	-0.8761***
C	17.5823	54.0938	15.2047	-0.3379	-1.1943	18.8684
R-squared	0.9619	0.9509	0.8007	0.8959	0.9669	0.9412
Adj. R-squared	0.9572	0.9449	0.7763	0.8832	0.9629	0.9340
F-statistic	206.1146	158.3149	32.8098	70.3151	238.5933	130.8027
Akaike AIC	5.1214	10.9816	2.1949	3.3125	4.6937	5.8112
Schwarz SC	5.3746	11.2348	2.4480	3.5658	4.9468	6.0644
DW Statistic	1.29	1.98	2.18	1.97	1.82	1.39

Note: *** Significant at the 0.01 level, ** significant at the 0.05 level, and *significant at the 0.1 level

Source: Own Computation

From Table (4.10), the results of VAR model estimation revealed that though t statistic of some variables are significant and some are not significant for rice at conventional level of significance. However, F-statistic value for the rice yield is very high and also statistically significant, so it can be said that the model is said to fit. The coefficient of determination R-square for rice yield is 0.9618, and the adjusted coefficient of determination R-square are 0.9572 for rice yield, respectively. Both the values are between 0 and 1, and those values are very high, so it shows the goodness of the fit of the overall model. Therefore, it can be said that 96.18% variation can be explained by the climatic variables for rice yield. Thus, the impacts of climatic variables are significant for rice yield. But, the Jarque-Bera test, chi-square LM test and Joint test show that the residuals of the model are not constant, have serial correlation, and are not normally distributed. In addition, the selected model was checked to satisfy the stationary condition as Figure (4.23).

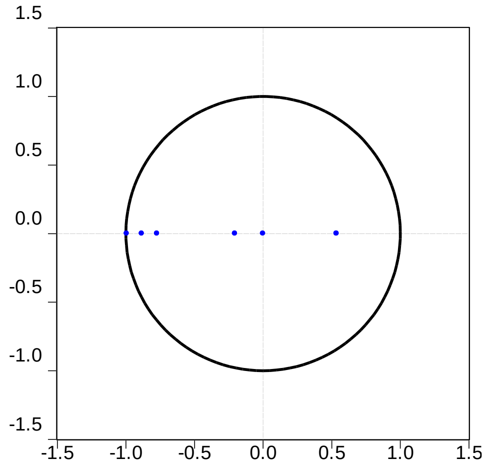


Figure (4.23): Inverse Roots of AR Characteristic Polynomial for Pathein District

Source: Own Computation

From Figure (4.23), no root lies outside the unit circle. So, VAR satisfies the stability condition. After that to trace out the endogeneity of the variables the Granger Causality test used and the results are shown in Table (4.11).

Table (4.11): Results of Granger Causality Tests for Pathein District

Null Hypothesis:	F-Statistic	P-value
RF does not Granger Cause RICEYIELD	8.2104***	0.006
MAXT does not Granger Cause RICEYIELD	4.2147**	0.045
D(MINT) does not Granger Cause RICEYIELD	6.1783**	0.0161
D(MRH) does not Granger Cause RICEYIELD	5.7342**	0.0202
D(ERH) does not Granger Cause RICEYIELD	3.7648*	0.0577

Note: *** Significant at the 0.01 level, ** significant at the 0.05 level, and *significant at the 0.1 level

Source: Own Computation

By the results of Table (4.11), it is suggested that the direction of causality from RF, MaxT, MinT, MRH and ERH to rice yield is observed since F-statistic for these variables are significant. Thus, it is indicated that there is causal relationship between climate variables and rice yield. Moreover, the impulse response function and the variance decomposition of rice yield are described in Table (4.12).

Table (4.12): Impulse Response and Variance Decomposition for Rice Yield in Pathein District

Variance Decomposition						
Period	**D(RICEYIELD)**	**RF**	**MAXT**	**D(MINT)**	**D(MRH)**	**D(ERH)**
1	100.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	95.6013	2.5060	0.2137	1.1967	0.2780	0.2042
3	92.8952	4.7783	0.1681	1.6175	0.2022	0.3385
4	90.6564	6.3939	0.2910	1.8398	0.3667	0.4521
5	88.3406	8.0289	0.3652	1.9640	0.7765	0.5247
6	85.9670	9.4663	0.5048	2.0327	1.4629	0.5662
7	83.5767	10.8317	0.6373	2.0650	2.3101	0.5792
8	81.2077	12.0686	0.7920	2.0727	3.2867	0.5721
9	78.9026	13.2101	0.9459	2.0636	4.3262	0.5514
10	76.6834	14.2466	1.1051	2.0430	5.3986	0.5231
Impulse Response Function						
Period	D(RICEYIELD)	RF	MAXT	D(MINT)	D(MRH)	D(ERH)
1	2.9551	0.0000	0.0000	0.0000	0.0000	0.0000
2	-2.5543	0.6324	-0.1846	0.4370	-0.2106	-0.1805
3	2.2917	-0.8093	0.0549	-0.4076	0.0165	0.2053
4	-2.0986	0.8379	-0.2069	0.3853	0.2369	-0.2224
5	1.9295	-0.9195	0.1959	-0.3620	-0.3885	0.2141
6	-1.7961	0.9518	-0.2621	0.3416	0.5389	-0.1993
7	1.6824	-0.9938	0.2782	-0.3221	-0.6455	0.1749
8	-1.5893	1.0173	-0.3137	0.3045	0.7399	-0.1476
9	1.5100	-1.0401	0.3324	-0.2882	-0.8120	0.1175
10	-1.4434	1.0546	-0.3546	0.2735	0.8728	-0.0873

Source: Own Computation

Impulse response function traces the effect of one time shock on one of the innovation on current and future values of endogenous variable. Table (4.12) traces out the responses of RF, MaxT, MinT, MRH and ERH. Increasing RF and MinT in the current period has a favorable impact on yield until period two, according to the impulse response functions from climate variables to yield. The other variables including

MaxT, MRH and ERH are decreasing negative effect on rice yield in current period till period two after that MaxT and MRH have decreasing positive and increasing negative effect.

Table (4.12) gives the variance decomposition values of rice yield for Pathein district. During the changes of yield, its own affect is 100% in first period and then gradually declines to 76.68%. The volatility of yield from 0% to 14.25% fluctuations can be explained by RF, 0% ~ 1.11%, 0% ~ 2.04%, 0% ~ 5.40% and 0 ~ 0.52% fluctuations can be explained by MaxT, MinT, MRH and ERH respectively.

Hinthada District

Based on the result of Table (4.13), the lowest value of HQ criterion at lag 1 is 33.1603, the lowest value of SC criterion at lag 1 is 34.1004, and the lowest value of the AIC criterion at lag 1 is 32.5675. The VAR model of order p=1 is the appropriate model to represent the rice yield of Hinthada District because the criteria are the lowest at lag 1. Therefore, the Vector Autoregressive model of order one, VAR (1) to the six variables was fitted, and the results are presented in Table (4.13).

Table (4.13): VAR Lags Order Selection Criteria for Hinthada District

Lag	AIC	SC	HQ
0	38.7847	39.0037	38.8694
1	32.5675*	34.1004*	33.1603*
2	32.8797	35.7265	33.9806

* indicates lag order selected by the criterion

Source: Own Computation

Table (4.14): VAR Estimation Results for Hinthada District

	D(RY)	RF	D(MAXT)	MINT	MRH	D(ERH)
D(RY(-1))	-0.7306	0.0440	0.0081	0.0132	0.0601	-0.0878
RF(-1)	0.0102	-0.4269^{***}	0.0081^{***}	-0.0139^{***}	-0.0312^{***}	-0.0138
D(MAXT(-1))	-0.2539	0.9899	-0.3594^{***}	0.1082	0.5911	0.1924
MINT(-1)	0.1382	-26.8441^{***}	-0.1785^{**}	-0.0318	0.2939	1.0563
MRH(-1)	0.1869	-0.6868	0.1026^{***}	-0.0006	0.0273	-0.4689^{*}
D(ERH(-1))	-0.1318	-0.4591	-0.0210^{**}	-0.0028	-0.0917	-0.7800^{***}
Constant	-18.7046	886.6511^{***}	-5.7891^{**}	24.3545^{***}	75.8567^{***}	17.3531
R-squared	0.7696	0.9308	0.9243	0.9140	0.8025	0.9081
Adj.R^2	0.7414	0.9223	0.9151	0.9035	0.7783	0.8968
F-statistic	27.2835	109.7858	99.7621	86.8284	33.1738	80.7113

Note: *** Significant at the 0.01 level, ** significant at the 0.05 level, and *significant at the 0.1 level

Source: Own Computation

From Table (4.14), the results of VAR model estimation revealed that though t statistic of some variables are significant while some are not significant for rice yield at conventional level of significance. The coefficient of determination R^2 for rice yield is 0.7696, and the adjusted R^2 are 0.7414 for rice yield, respectively. Both the values are between 0 and 1, and both the values are very high and shows the goodness of the fit of the overall model. Therefore, it can be said that 77% variation can be explained by the climatic variables for rice yield. So, there are effects of climatic variables on rice yield. Moreover, the residuals of the model are constant, have serial correlation, and are not normally distributed (Appendix C-II). In addition, the selected model was checked to satisfy the stationary condition as Figure (4.24).

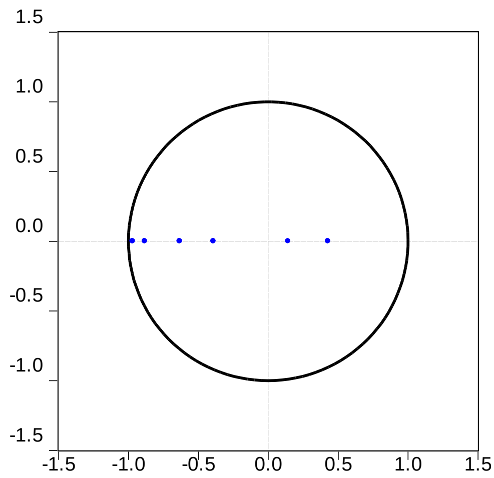


Figure (4.24): Inverse Roots of AR Characteristic Polynomial for Hinthada District

Source: Own Computation

From Figure (4.24), no root lies outside the unit circle. So VAR(1) satisfies the stationary condition. After that to trace out the endogeneity of the variables the granger causality test used as in Table (4.15).

Table (4.15): Results of Granger Causality Tests for Hinthada District

Null Hypothesis:	F-Statistic	P-value
RF does not Granger Cause D(RICEYIELD)	1.7114	0.1908
D(MAXT) does not Granger Cause D(RICEYIELD)	1.6001	0.2118
MINT does not Granger Cause D(RICEYIELD)	2.0855	0.1347
MRH does not Granger Cause D(RICEYIELD)	2.0638	0.1374
D(ERH) does not Granger Cause D(RICEYIELD)	0.4313	0.6520

Source: Own Computation

The table (4.15) presents the results of Granger causality tests for Hinthada District. Granger causality is a statistical concept used to determine whether one variable can be used to predict another variable. This test aims to assess whether certain variables can be used to predict changes in the rice yield. The null hypothesis for each test is that the variable mentioned does not Granger cause the rice yield. It suggests that there are no statistically significant evidence to reject the null hypothesis. Based on the results, none of the climatic variables (RF, MAXT, MINT, MRH, and ERH) are found to be significant predictors of changes in rice yield. Therefore, the lack of statistical significance suggests that these variables do not have a significant causal relationship with rice yield. In addition, Table (4.16) describes the impulse response function and the variance decomposition of rice yield.

Table (4.16): Impulse Response and Variance Decomposition for Rice Yield in Hinthada District

Variance Decomposition						
Period	D(RY)	RF	D(MAXT)	MINT	MRH	D(ERH)
1	100.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	98.3156	0.9307	0.1336	0.0272	0.0768	0.5158
3	96.2194	2.1917	0.2225	0.0823	0.0754	1.2083
4	93.7731	3.6156	0.2817	0.2634	0.1284	1.9374
5	91.2397	5.1468	0.3187	0.4943	0.2218	2.5784
6	88.7867	6.6849	0.3433	0.7772	0.3182	3.0894
7	86.5012	8.2006	0.3610	1.0710	0.4003	3.4655
8	84.4117	9.6642	0.3751	1.3650	0.4602	3.7236
9	82.5171	11.0647	0.3870	1.6459	0.4986	3.8864
10	80.8016	12.3946	0.3978	1.9096	0.5192	3.9769
Impulse Response Function						
Period	D(RY)	RF	D(MAXT)	MINT	MRH	D(ERH)
1	6.4133	0.0000	0.0000	0.0000	0.0000	0.0000
2	-4.9218	0.7865	-0.2980	0.1346	0.2260	-0.5855
3	3.7633	-1.0920	0.3083	-0.2235	0.1060	0.8097
4	-2.9310	1.2593	-0.2842	0.4236	-0.2416	-0.9065
5	2.3613	-1.3738	0.2500	-0.5100	0.3272	0.9094
6	-1.9540	1.4340	-0.2232	0.5881	-0.3487	-0.8632
7	1.6652	-1.4715	0.2037	-0.6229	0.3385	0.7903
8	-1.4522	1.4884	-0.1907	0.6434	-0.3068	-0.7064
9	1.2920	-1.4935	0.1819	-0.6479	0.2656	0.6199
10	-1.1671	1.4892	-0.1759	0.6449	-0.2205	-0.5362

Source: Own Computation

The impulse response function shows the response of rice yield to shocks or innovations in each variable (RF, D(MAXT), MINT, MRH, and D(ERH)) over different periods. In the first period, a shock to rice yield itself has a positive effect of 6.4133. In the second period, a shock to RF has a positive effect on rice yield, while MaxT, MinT and MRH have mixed positive and negative effects. But, ERH has a negative effect. The subsequent periods continue to show mixed effects of shocks in the variables on rice yield.

The variance decomposition indicates the proportion of the variation in rice yield that can be attributed to each variable over the specified periods. Table (4.16) gives the variance decomposition values of rice yield for Hinthada district. During the changes of yield, its own affect is 100% in first period and then gradually declines to 80.8016%. The volatility of yield from 0% to 12.40% , 0% ~ 0.40%, 0% ~ 1.91%, 0% ~ 0.52% and 0 ~ 3.98% fluctuations can be explained by RF, MaxT, MinT, MRH and ERH respectively.

It is important to note that specific agricultural practices and conditions in Hinthada District may vary across different townships and villages within the district. Local factors such as topography, access to irrigation, and availability of infrastructure can influence farming practices and agricultural productivity in different areas.

Maubin District

Based on the result of fitting VAR model for rice yield of Maubin District Table (4.17), the lowest value of HQ criterion at lag 1 is 34.4975, the lowest value of SC criterion at lag 1 is 35.4376, and the lowest value of the AIC criterion at lag 6 is 33.9047. The VAR model of order p=1 is the appropriate model to represent the rice yield of Maubin District because the criteria are the lowest at lag 1. Therefore, the Vector Autoregressive model of order one, VAR (1) to the six variables was fitted, and the results are presented in Table (4.18).

Table (4.17): VAR Lags Order Selection Criteria for Maubin District

Lag	AIC	SC	HQ
0	40.4280	40.6469	40.5126
1	33.9047*	35.4376*	34.4975*
2	34.1427	36.9895	35.2436

* indicates lag order selected by the criterion

Source: Own Computation

Table (4.18): VAR Estimation Results for Maubin District

	D(RY)	RF	D(MaxT)	MINT	D(MRH)	D(ERH)
D(RY(-1))	-0.9170***	3.9065***	0.0026	0.0782**	0.0636	-0.0689
RF(-1)	0.0099	-0.2459	0.0065*	-0.0053	-0.0157	-0.0115
D(MaxT(-1))	-0.0817	9.9916**	-0.6217***	0.1648	0.2648	0.3006
MINT(-1)	-0.1789	-5.2265	-0.0660	0.7106***	-0.1080	-0.5001
D(MRH(-1))	0.0327	-1.5291	-0.0577	0.0229**	-0.5781***	-0.1571
D(ERH(-1))	-0.0291	-2.0890**	0.0127	-0.0914**	0.0617	-0.6815***
Constant	2.4725	374.0985***	0.0386	7.2072	5.6165	13.1788
R-squared	0.9801	0.9413	0.7943	0.7689	0.8948	0.8979
Adj. R^2	0.9776	0.9341	0.7691	0.7406	0.8820	0.8854
F-statistic	401.1194	131.0503	31.5420	27.1807	69.5280	71.8879
AIC	4.9524	10.85145	3.3112	3.8720	5.5546	6.4710
BIC	5.2056	11.10462	3.5643	4.1252	5.8078	6.7242

Note: *** Significant at the 0.01 level, ** significant at the 0.05 level, and *significant at the 0.1 level

Source: Own Computation

From Table (4.18), the results of the VAR(1) model estimation mentioned that F-statistic value for the rice yield is very high and also statistically significant, so the model is said to fit. The coefficient of determination R-square for rice yield is 0.9801, and the adjusted coefficient of determination R-square are 0.9776 for rice yield, respectively. Therefore, it can be said that 98% of variation can be explained by the climatic variables for rice yield. So, the impacts of climatic variables are significant for rice yield. But the residuals of the model have no serial correlation, and are not normally distributed and its variance is not constant (Appendix C.III). In addition, the selected model was checked to satisfy the stationary condition as Figure (4.25).

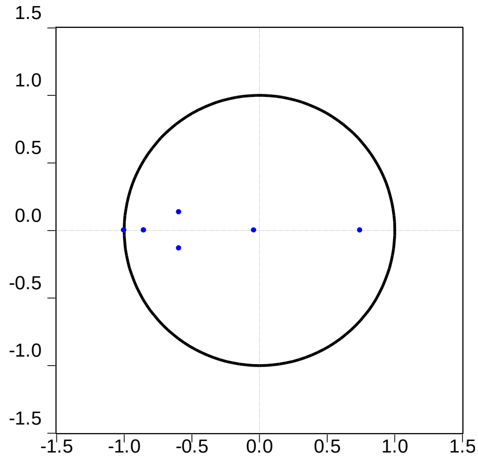


Figure (4.25): Inverse Roots of AR Characteristic Polynomial for Maubin District

Source: Own Computation

From Figure (4.25), no root lies outside the unit circle. So VAR(1) satisfies the stationary condition. After that to trace out the endogeneity of the variables the granger causality test used as in Table (4.19).

Table (4.19): Results of Granger Causality Tests for Maubin District

Null Hypothesis:	F-Statistic	Prob.
RF does not Granger Cause D(RICEYIELD)	1.7543	0.1835
D(MAXT) does not Granger Cause D(RICEYIELD)	1.8282	0.1713
MINT does not Granger Cause D(RICEYIELD)	0.4831	0.6197
D(MRH) does not Granger Cause D(RICEYIELD)	0.8763	0.4226
D(ERH) does not Granger Cause D(RICEYIELD)	2.0690	0.1370

Source: Own Computation

The table (4.19) presents the results of Granger causality tests for Maubin District. This test aims to assess whether certain variables can be used to predict changes in the riceyield. The null hypothesis for each test is that the variable mentioned does not Granger cause the riceyield variable. It suggests that there are no statistically significant evidence to reject the null hypothesis. Based on the results, none of the climatic variables (RF, MAXT, MINT, MRH, and ERH) are found to be significant predictors of changes in riceyield. Therefore, the lack of statistical significance suggests that these variables do not have a significant causal relationship with riceyield. In addition, Table (4.20) describes the impulse response function and the variance decomposition of rice yield.

Table (4.20): Impulse Response and Variance Decomposition for Rice Yield in Maubin District

Variance Decomposition						
Period	**D(RICEYIELD)**	**RF**	**D(MAXT)**	**MINT**	**D(MRH)**	**D(ERH)**
1	100.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	98.1771	1.0257	0.1850	0.4822	0.0024	0.1272
3	96.8350	2.1668	0.3828	0.3240	0.0178	0.2734
4	95.5678	2.9074	0.8325	0.3590	0.0462	0.2869
5	94.4959	3.6365	1.2167	0.2929	0.0690	0.2887
6	93.5098	4.1802	1.6722	0.2839	0.0950	0.2585
7	92.6532	4.6997	2.0503	0.2485	0.1191	0.2291
8	91.8652	5.1267	2.4276	0.2330	0.1461	0.2012
9	91.1594	5.5326	2.7437	0.2105	0.1735	0.1799
10	90.5054	5.8870	3.0410	0.1961	0.2034	0.1669
Impulse Response Function						
Period	D(RICEYIELD)	RF	D(MAXT)	MINT	D(MRH)	D(ERH)
1	2.7157	0.0000	0.0000	0.0000	0.0000	0.0000
2	-2.6623	0.3887	-0.1651	-0.2665	0.0190	-0.1369
3	2.6135	-0.5704	0.2385	0.0146	-0.0596	0.2034
4	-2.5636	0.6093	-0.3982	-0.1828	0.0977	-0.1534
5	2.5348	-0.6869	0.4458	0.0403	-0.1076	0.1452
6	-2.5037	0.7025	-0.5327	-0.1297	0.1271	-0.0859
7	2.4834	-0.7463	0.5556	0.0465	-0.1371	0.0582
8	-2.4617	0.7587	-0.6007	-0.0925	0.1546	-0.0082
9	2.4464	-0.7881	0.6115	0.0432	-0.1670	-0.0197
10	-2.4308	0.8001	-0.6358	-0.0662	0.1830	0.0571

Source: Own Computation

Table (4.20) traces out the responses of RF, MaxT, MinT, MRH, and ERH. The positive change in riceyield has a positive impact on itself in the next period. In subsequent periods, the response fluctuates between positive and negative values, suggesting a temporary effect. The positive change in RF in the second period, leads to a positive response in riceyield in the next period. The response then becomes negative in the third period and remains negative in the other periods. A negative change in MaxT in the second period, leads to a negative response in riceyield in the next period.

The response becomes positive in the third period and alternates between positive and negative values in subsequent periods. A negative change in MINT In the second period, leads to a negative response in riceyield in the next period. The response becomes positive in the third period and fluctuates between positive and negative values thereafter. A positive change in MRH in the second period, leads to a positive response in riceyield in the next period. The response becomes negative in the third period and remains negative in the other periods. A negative change in ERH in the second period, leads to a negative response in riceyield in the next period. The response becomes positive in the third period and fluctuates between positive and negative values thereafter.

The table shows the variance decomposition of rice yield in Maubin District for different time periods. In period 1, riceyield alone explains 100% of the variance, indicating that it is the only variable influencing rice yield during that specific time period. The other variables do not contribute to the variance of rice yield in period 1. This suggests that any changes or fluctuations in rice yield during this period are solely driven by the variable riceyield itself.

After the second period, riceyield remains the dominant contributor throughout these periods, but its contribution gradually decreases over time. This means that riceyield continues to have a significant influence on rice yield, but its relative importance decreases compared to the other variables. These variables start to contribute to the variance of rice yield in these periods. Initially, their contributions are relatively small, but as the time period progresses, their impact becomes more pronounced. During the changes in rice yield, its own affect is 100% in first period and then gradually declines to 36.74%. The fluctuations from 0% to 4.11%, 0% ~ 11.52%, 0% ~ 2.64%, 0% ~ 23.18% and 0 ~ 21.81% can be explained by the volatility of RF, MaxT, MinT, MRH and ERH respectively.

Myaungmya District

Based on the result of fitting VAR model for rice yield of Myaungmya District of Table (4.21), the lowest value of HQ criterion at lag 1 is 36.4635, the lowest value of SC criterion at lag 1 is 37.4036, and the lowest value of the AIC criterion at lag 1 is 35.8707. The VAR model of order p=1 is the appropriate model to represent the rice yield of Myaungmya District because the criteria are the lowest at lag 1. Therefore, the

Vector Autoregressive model of order one, VAR(1) to the six variables was fitted, and the results are presented in Table (4.22).

Table (4.21): VAR Lags Order Selection Criteria for Myaungmya District

Lag	AIC	SC	HQ
0	42.0197	42.2387	42.1044
1	35.8707*	37.4036*	36.4635*
2	35.8723	38.7191	36.9732

* indicates lag order selected by the criterion

Source: Own Computation

Table (4.22): VAR Estimation Results for Myaungmya District

	D(RY)	RF	MAXT	D(MINT)	D(MRH)	D(ERH)
D(RY(-1))	-0.8911***	9.2414***	-0.0257	0.1352***	0.1338	-0.1401
RF(-1)	0.0023	-0.0841	0.0008	0.0018	-0.0065	-0.0099
MAXT(-1)	-0.8023	2.5755	0.2573*	0.4991**	1.0131	1.4896*
D(MINT(-1))	0.2864	2.3778	0.0956*	0.7764***	0.7794***	0.5420*
D(MRH(-1))	-0.0484	-2.0666	-0.0114	-0.0213	0.1985	-0.7195***
D(ERH(-1))	0.0418	-2.4060	0.0104	-0.0405**	-0.2582***	-0.6829***
Constant	24.4023*	291.7586	22.8924***	-10.5236	16.1843	-0.0629
R-squared	0.9733	0.8213	0.4413	0.7795	0.7330	0.9252
Adj. R^2	0.9700	0.7994	0.3729	0.7526	0.7003	0.9160
F-statistic	297.7085	37.5407	6.4527	28.8778	22.4184	101.0744
AIC	5.1956	12.4023	2.8891	3.8490	5.9435	6.3529
BIC	5.4487	12.6555	3.1423	4.1021	6.1967	6.6061
DW Stats	1.7073	1.9157	1.8804	1.6375	2.0807	2.0653

Note: *** Significant at the 0.01 level, ** significant at the 0.05 level, and *significant at the 0.1 level

Source: Own Computation

From Table (4.24), the results of the VAR(1) model estimation said that though t statistic of some variables are significant while some are not significant for rice yield at conventional level of significance. However, F-statistic value for the rice yield is very high and also statistically significant, so the model is said to fit. The coefficient of determination R-square for rice yield is 0.9733, and the adjusted coefficient of determination R-square are 0.9700 for rice yield, respectively. It also shows the goodness of the fit of the overall model and it can be said that 97.33% of variation can be explained by the climatic variables for rice yield. The residuals of the model have no serial correlation, and are not normally distributed but the variance of residuals is constant. In addition, the selected model was checked to satisfy the stationary condition as Figure (4.26).

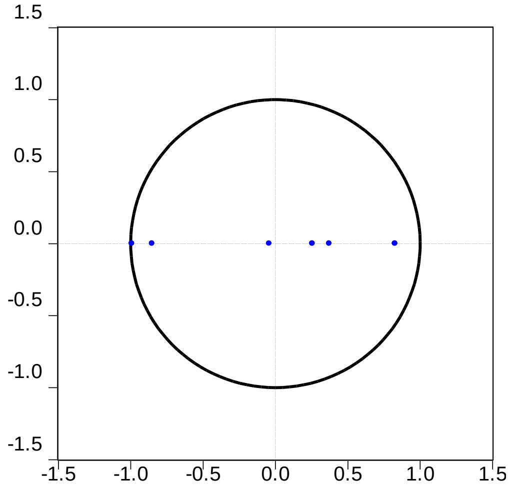


Figure (4.26): Inverse Roots of AR Characteristic Polynomial for Myaungmya District

Source: Own Computation

From Figure (4.26), no root lies outside the unit circle. So, VAR(1) satisfies the stability condition. After that to trace out the endogeneity of the variables the granger causality test used as in Table (4.23).

Table (4.23): Results of Granger Causality Tests for Myaungmya District

Null Hypothesis:	F-Statistic	Prob.
RF does not Granger Cause D(RICEYIELD)	2.4065	0.1268
MAXT does not Granger Cause D(RICEYIELD)	3.1438*	0.0820
D(MINT) does not Granger Cause D(RICEYIELD)	1.6258	0.2078
D(MRH) does not Granger Cause D(RICEYIELD)	0.1016	0.7512
D(ERH) does not Granger Cause D(RICEYIELD)	2.5570	0.1158

Note: *** Significant at the 0.01 level, ** significant at the 0.05 level, and *significant at the 0.1 level

Source: Own Computation

By the results of Table (4.23), it is suggested the direction of causality is from maximum temperature to rice yield since F-statistic for this variables is significant at 10%. Therefore, maximum temperature has a granger cause for rice yield. Then, the impulse response function and the variance decomposition of rice yield are in Table (4.24). The above results show that the effect of rainfall, MinT and ERH are positive, whereas MaxT and MRH have a negative influence on rice yield.

Table (4.24): Impulse Response and Variance Decomposition for Rice Yield in Myaungmya District

Variance Decomposition						
Period	**D(RICEYIELD)**	**RF**	**MAXT**	**D(MINT)**	**D(MRH)**	**D(ERH)**
1	100.0000	0.000000	0.000000	0.000000	0.000000	0.000000
2	95.20149	0.111561	3.531434	0.136834	0.098361	0.920325
3	93.15456	0.078950	2.532597	2.265192	1.315442	0.653261
4	89.83689	1.563602	2.317060	4.661074	1.031511	0.589862
5	88.37051	2.278612	1.999816	5.965814	0.851224	0.534022
6	87.13893	2.509309	1.956217	7.183708	0.724063	0.487769
7	84.51298	2.541310	2.345181	8.828809	1.240905	0.530818
8	82.35388	2.544289	2.806392	9.755475	1.732463	0.807501
9	79.92490	2.450216	3.810278	10.51120	2.149876	1.153531
10	78.27057	2.444134	4.452371	10.92392	2.633231	1.275777
Impulse Response Function						
Period	**D(RICEYIELD)**	**RF**	**MAXT**	**D(MINT)**	**D(MRH)**	**D(ERH)**
1	3.0668	0.0000	0.0000	0.0000	0.0000	0.0000
2	-2.8060	0.3996	-0.7660	0.5083	-0.1331	0.1993
3	2.7381	-0.3829	0.5074	-0.2896	-0.1036	-0.3761
4	-2.6624	0.4907	-0.4564	0.4009	0.2106	0.5115
5	2.5784	-0.5181	0.4191	-0.3048	-0.3337	-0.6362
6	-2.5246	0.5844	-0.3460	0.3680	0.4181	0.7316
7	2.4569	-0.6045	0.3264	-0.3034	-0.5031	-0.8180
8	-2.4124	0.6494	-0.2723	0.3453	0.5621	0.8841
9	2.3583	-0.6625	0.2576	-0.3001	-0.6213	-0.9433
10	-2.3201	0.6925	-0.2187	0.3278	0.6618	0.9881

Source: Own Computation

According to the Table (4.24) about impulse response functions from climate variables to yield, increasing of RF, ERH and MinT has positive effect on riceyield in the current period until the future. The increasing of MRH has positive effect on rice yield in three period after that of increasing has positive up to tenth period. The other variable, decreasing of MaxT has negative effect and that of increasing has negative effect on rice yield in the future.

Table (4.26) gives the variance decomposition values of rice yield for Myaungmya District. During the changes of yield, its own affect is 100% in first period and then gradually declines to 78.27%. The fluctuations from 0% to 2.44%, 0% ~ 2.45%, 0% ~ 10.92%, 0% ~ 2.63% and 0 ~ 1.28% can be explained by the volatility of RF, MaxT, MinT, MRH and ERH respectively.

Phyarpon District

Based on the result of fitting VAR model for rice yield of Phyarpon District Table (4.25), the lowest value of HQ criterion at lag 1 is 33.1411, the lowest value of SC criterion at lag 1 is 34.3071, and the lowest value of the AIC criterion at lag1 is 32.4972. The VAR model of order p=1 is the appropriate model to represent the rice yield of Phyarpon District because the criteria are the lowest at lag 1. Therefore, the Vector Autoregressive model of order one, VAR (1) to the six variables was fitted, and the results are presented in Table (4.25).

Table (4.25): VAR Lags Order Selection Criteria for Phyarpon District

Lag	AIC	SC	HQ
0	39.1281	39.3867	39.2201
1	32.4972*	34.3071*	33.1411*

* indicates lag order selected by the criterion

Source: Own Computation

Table (4.26): VAR Estimation Results for Phyarpon District

	D(RY)	D(RF)	D(MAXT)	D(MINT)	D(MRH)	D(ERH)
D(RY(-1))	-0.9236***	2.6141**	0.0057	-0.0106	-0.1138**	-0.0543
D(RF(-1))	0.0063	-0.4319**	0.0030	-0.0071*	-0.0254***	-0.0097
D(MAXT(-1))	0.7333	-11.1815	-0.5024***	0.2903	0.0667	1.9628
D(MINT(-1))	0.5033	-12.2387	0.0372	0.2005	0.1143	-0.4251
D(MRH(-1))	0.0074	-12.9105**	-0.0188	-0.0595	-0.3576	0.5145
D(ERH(-1))	-0.0714	-0.3298	-0.0362*	0.0248	0.1015	-0.7905***
Constant	0.3199	4.1543	0.0029	0.2102	0.4969	-0.4709
R-squared	0.9914	0.9861	0.9208	0.8923	0.9648	0.8162
Adj.R^2	0.9898	0.9834	0.9055	0.8714	0.9580	0.7807
F-statistic	600.2646	368.1348	60.1068	42.8219	141.7957	22.9575

Source: Own Computation

From Table (4.26), the results of the VAR(1) model estimation revealed that F-statistic value for the riceyield is very high and also statistically significant, so the model is said to be fit. The coefficient of determination R-square for rice yield is 0.9914, and the adjusted coefficient of determination R-square are 0.9898 for rice yield, respectively and it also shows the goodness of the fit of the overall model. Therefore, it can be said that 99.14% of variation can be explained by the climatic variables for rice yield. So, the impacts of climatic variables are significant for rice yield. The residuals of the model have serial correlation, and are normally distributed but the variance of it is constant. In addition, the selected model was checked to satisfy the stationary condition as Figure (4.27).

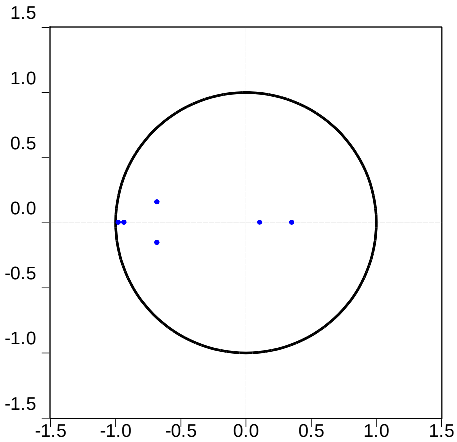


Figure (4.27): Inverse Roots of AR Characteristic Polynomial for Phyarpon District

Source: Own Computation

From Figure (4.27), no root lies outside the unit circle. So, VAR(1) satisfies the stationary condition. After that to trace out the endogeneity of the variables the granger causality test used as in Table (4.27).

Table (4.27): Results of Granger Causality Tests for Phyarpon District

Null Hypothesis:	F-Statistic	P-value
RF does not Granger Cause RICEYIELD	3.4764**	0.0430
MAXT does not Granger Cause RICEYIELD	2.3554	0.1111
MINT does not Granger Cause RICEYIELD	3.3697**	0.0470
MRH does not Granger Cause RICEYIELD	4.1380**	0.0252
ERH does not Granger Cause RICEYIELD	0.2525	0.7784

*** Significant at the 0.01 level, ** significant at the 0.05 level, and *significant at the 0.1 level

Source: Own Computation

By the results of Table (4.27), it is suggested that the direction of causality is from RF, MinT and MRH to rice yield since F-statistic for these variables are significant at 5%. However, there are no causation from MaxT and ERH to rice yield because of the F-statistic is statistically insignificant. The impulse response function and the variance decomposition of rice yield are described in Table (4.28).

Table (4.28): Impulse Response and Variance Decomposition for Rice Yield in Phyarpon District

Variance Decomposition						
Period	**D(RY)**	**D(RF)**	**D(MAXT**	**D(MINT)**	**D(MRH)**	**D(ERH)**
1	100.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	95.3915	1.6277	0.8047	1.7022	0.0002	0.4736
3	92.2341	3.3405	1.2627	1.7715	0.3481	1.0429
4	88.4879	5.5012	1.6297	1.7456	1.0320	1.6033
5	84.4215	7.9956	1.8800	1.6643	1.9333	2.1050
6	80.1763	10.7332	2.0290	1.5738	2.9540	2.5334
7	75.8874	13.6180	2.0951	1.4824	4.0233	2.8936
8	71.6503	16.5645	2.0991	1.3949	5.0958	3.1951
9	67.5365	19.5002	2.0596	1.3124	6.1422	3.4488
10	63.5960	22.3679	1.9919	1.2353	7.1449	3.6638
Impulse Response Function						
Period	**D(RY)**	**D(RF)**	**D(MAXT**	**D(MINT)**	**D(MRH)**	**D(ERH)**
1	3.3072	0.0000	0.0000	0.0000	0.0000	0.0000
2	-3.1220	0.5941	0.4177	0.6075	0.0023	-0.3204
3	2.9614	-0.8449	-0.4782	-0.4434	-0.3334	0.4799
4	-2.7982	1.1185	0.5324	0.4120	0.5689	-0.5852
5	2.6368	-1.3681	-0.5463	-0.3696	-0.7603	0.6537
6	-2.4809	1.5988	0.5386	0.3421	0.9184	-0.6994
7	2.3317	-1.8090	-0.5157	-0.3184	-1.0503	0.7324
8	-2.1903	1.9993	0.4845	0.2995	1.1627	-0.7581
9	2.0571	-2.1705	-0.4492	-0.2829	-1.2597	0.7800
10	-1.9321	2.3233	0.4128	0.2682	1.3444	-0.7997

Source: Own Computation

In the first period, the variable rice yield shows a positive response of 3.3072 to the shock, while all other variables have a response of 0.0000. This implies that a shock in the variable represented by rice yield leads to a significant positive response in itself, but does not immediately affect the other variables. In the second period, rice yield shows a negative response of -3.1220, suggesting that the previous shock has a dampening effect in the second period. Additionally, the variables RF and MaxT exhibit positive responses, while MinT, MRH, and ERH also display small positive or negative responses. The subsequent periods continue to show the cumulative response of each variable to the shock. Some variables continue to display positive or negative responses, indicating a persistent impact, while others fluctuate between positive and negative responses.

The variance decomposition values of rice yield for Phyarpon District is also shown in Table (4.28). In the first period, the variable rice yield has a variance decomposition of 100%, indicating that all the forecast error variance in rice yield can be attributed to its own shocks. The other variables have a variance decomposition of 0% since they are not affected by their own shocks in this period. In subsequent periods, the variance decomposition of rice yield gradually decreases, indicating that the shocks from other variables start to contribute to the forecast error variance of rice yield. Conversely, the variance decomposition of the other variables increases, suggesting that their own shocks become more influential. In the second period, rice yield has a variance decomposition of 95.3915%, implying that 95.3915% of the forecast error variance in rice yield is due to its own shocks, while the remaining percentages are attributable to the shocks from other variables. The increasing values in variance decomposition for the other variables indicate their growing impact on the forecast error variance of rice yield.

During the changes of yield, its own affect is 100% in first period and then gradually declines to 63.59%. The fluctuations from 0% to 22.37%, 0% ~ 1.99%, 0% ~ 1.24 %, 0% ~ 7.15% and 0 ~ 3.66% can be explained by the volatility of RF, MaxT, MinT, MRH and ERH respectively.

Overall, it is seen that almost all of the predictors have insignificant effect on rice yield in all Districts. But, all of the predictors have causality effects on rice yield in Pathein District. Although, there are no causal effects in Hinthada and Maubin Districts, maximum temperature has causal effect on rice yield in Myaungmya District and rainfall, minimum temperature and morning RH have the causal effect on rice yield in Phyarpon District. Moreover, the assumptions for the normality, heteroskedasticity and residual serial correlation, are not met in some districts. Then, the Artificial Neural Networks (ANNs) model as the alternative way was fitted.

4.9 Fitting the Artificial Neural Networks (ANNs) Model

In this study, the algorithm of the Artificial Neural Networks (ANNs) was investigated to test whether it can be effectively utilized in the context of the rice yield and climate change in Ayeyawady Region. This study attempts to use some of the influencing climatic factors for predicting rice yield at five districts in the Ayeyawady Region.

Table (4.29): Neural Network Configuration for All Districts

		N	Percent
Sample	Training	47	81%
	Testing	11	19%
Valid		58	100%
Excluded		0	
Total		58	

Source: Own Computation

Table (4.29) depicts information about the neural network configuration. Data collected over 58 observations are used in the network in which automatic architecture selection has chosen 81% of the total available data for training purpose and 19% of data for testing purpose for all districts. But, 40 observations are used in the network in which automatic architecture selection has chosen 75% of the total available data for training purpose and 25% of data for testing purpose for Phyarpon District. Predictions are made based on the past and current information.

Pathein District

A single hidden layer with two neurons in each hidden layer was used in this study. Several hidden layers, and several neurons, are determined by conducting the trial and error method iteratively. This method ensured that the chosen structure for the network was adequate for this study.

Table (4.30): Network Information of Pathein District

Input Layer	**Covariates 1**	**RF**
	2	MaxT
	3	MinT
	4	MRH
	5	ERH
	Number of Units	5
	Rescaling Method for Covariates	Standardized
Hidden Layer(s)	Number of Hidden Layers	1
	Number of Units in Hidden Layer 1	2
	Activation Function	Hyperbolic tangent
Output Layer	Dependent Variables	Rice Yield
	Number of Units	1
	Rescaling Method for Scale Dependents	Standardized
	Activation Function	Identity
	Error Function	Sum of Squares

Source: Own Computation

Table (4.30) shows the network information that contains five covariates, excluding a bias unit, which are used as input units. Single hidden layer with 2 neurons in each layer are used in the network. The standardized method is used for rescaling method for covariates. Rice yield is a variable used as an output of this network. Automatic architecture selection was applied to select the activation function for both the input and output layer. Sum of Square error is used in the output layer as an error function which the network tries to minimize during training.

Figure (4.28) depicts three layers of information, including an input, one hidden, and an output layer. The input layer consists of six units, including a bias unit. The hidden layer possesses two neurons and one bias unit. The hyperbolic tangent and identity activation functions that are appropriate for the input unit and output unit of this network are found automatically through architectural selection.

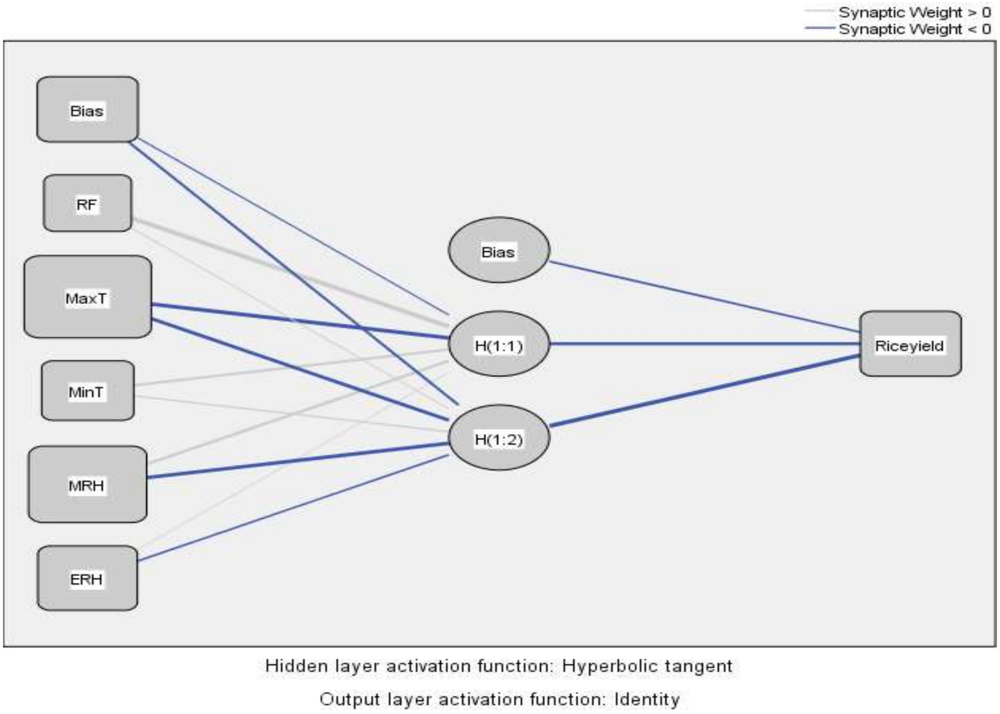


Figure (4.28): Neural Network of Pathein District

Source: Own Computation

It also shows the interaction between various nodes in the network. The connection between nodes is indicated either by a grey or blue line. The grey line indicates that the synaptic weight is high and implies the strength of communication between nodes is strong (positive). The blue lines represent weak (negative) node-to-node connectivity. It is seen that a higher positive value indicates a stronger bond. The first node in the hidden layer has the strongest bond of rainfall, on the other hand, evening relative humidity has the weakest positive bond with the first node in the hidden layer. The second node in the hidden layer has the highest bond of minimum temperature. Contrast, morning relative humidity has the weakest bond of the second node.

Table (4.31): Weight Values of the ANN in Pathein District

Predictor		Hidden Layer 1	
		H(1:1)	H(1:2)
Input Layer	(Bias)	-.302	-.694
	Rainfall	1.294	.054
	Maximum Temperature	-1.135	-1.049
	Minimum Temperature	.755	.333
	Morning Relative Humidity	.757	-1.115
	Evening Relative Humidity	.046	-.490
		Output Layer (RiceYield)	
Hidden Layer	(Bias)	-.556	
	H(1:1)	-.786	
	H(1:2)	-1.178	

Source: Own Computation

Table (4.31) shows weight values used in the network in the input and hidden layers. It is seen that maximum temperature has negative effects, and rainfall, minimum temperature, morning and evening relative humidity have positive effects on rice yield in Hidden layer (1:1). In addition, maximum temperature, morning and evening relative humidity have negative effects, and rainfall and minimum temperature have positive effects on rice yield in Hidden layer (1:2).

Table (4.32): Model Summary of Pathein District

Training	Sum of Squares Error	3.683
	Relative Error	.160
	Stopping Rule Used	1 consecutive step(s) with no decrease in error
Testing	Training Time	0:00:00.02
	Sum of Squares Error	.535
	Relative Error	.118

Source: Own Computation

The relative error and sum of square error throughout the training and testing phases are shown in Table (4.32). The network tries to minimize the sum of square error during the training phase. This error is displayed when the output layer has scale-dependent variables. The model returned the sum of squares error 3.683 and 0.535 in the training and testing phase, respectively. The relative error is 0.160 and 0.118 in the training and testing phase, respectively. The sum of square error with other error values, in this model relative error, are used to compute for the rescaled values of the dependent variable, rice yield.

The neural network prediction process stopped when the error did not decrease. These points are evidence of the accuracy of the model. The neural network output shows that the model is suitable for predicting rice yield in Pathein. Moreover, the importance of the independent variables (predictors) was described in Table (4.33) and Figure (4.29).

Table (4.33): Importance of the Predictors in Model for Pathein District

Predictor	Importance	Normalized Importance
Rainfall	0.086	25.0%
MaxT	0.345	100.0%
MinT	0.118	34.1%
MRH	0.287	83.3%
ERH	0.164	47.7%

Source: Own Computation

According to Table (4.33) and Figure (4.28) for the Pathein district, the importance of the rainfall variable is almost 25%, Maximum temperature is 100%, Minimum temperature is 34.1%, morning relative humidity is 83.3%, and evening relative humidity is 47.7%. It is found that the influence of maximum temperature on the change of rice yield in Pathein District is the most, and the influence of morning relative humidity is not as much as the influence of maximum temperature. In addition, the influence of evening relative humidity on the change in rice yield was found to be half a percent. However, only a small percentage of the changes in rice yield are found to be influenced by rainfall and minimum temperature. It can be obviously seen in Figure (4.29).

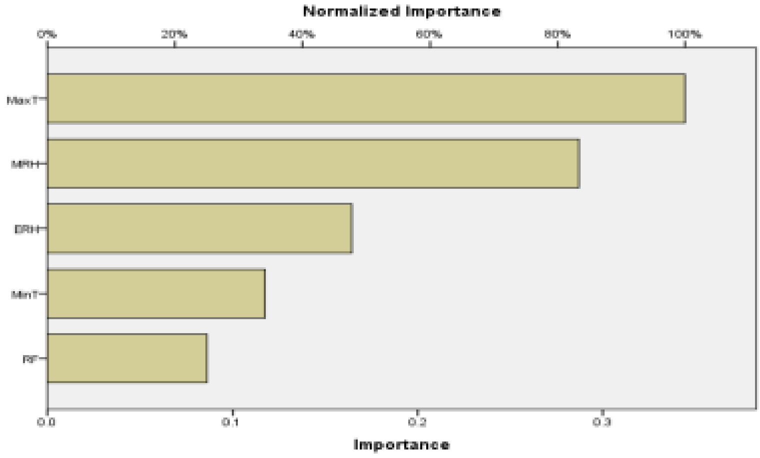


Figure:(4.29): Normalized Importance of the Predictors in Pathein district

Source: Own Computation

Hinthada District

In this district, single hidden layer with one neuron were employed. The network information that contains five covariates, excluding a bias unit, which are used as input units is shown in Table (4.34).

Table (4.34): Network Information of Hinthada District

Input Layer	**Covariates** 1	RF
	2	MaxT
	3	MinT
	4	MRH
	5	ERH
	Number of Units	5
	Rescaling Method for Covariates	Standardized
Hidden Layer(s)	Number of Hidden Layers	1
	Number of Units in Hidden Layer 1	2
	Activation Function	Hyperbolic tangent
Output Layer	Dependent Variables	1
	Number of Units	1
	Rescaling Method for Scale Dependents	Standardized
	Activation Function	Identity
	Error Function	Sum of Squares

Source: Own Computation

Single hidden layers with two neurons in this layer are used in the network. Rice yield is a variable used as an output of this network. Automatic architecture selection was applied to select the activation function for both the hidden and output layer. The sum of Square error is used in the output layer as an error function which the network tries to minimize during training.

Figure (4.30) shows three layers of information that include an input, one hidden and output layer. The input layer consists of six units including a bias unit. In hidden layer possesses with two neuron and one bias unit. The automatic architectural selection indicates that the input unit and output unit of this network are most effectively supported by the hyperbolic tangent and identity activation functions, respectively. In addition, Table (4.35) presents the weight values for this network.

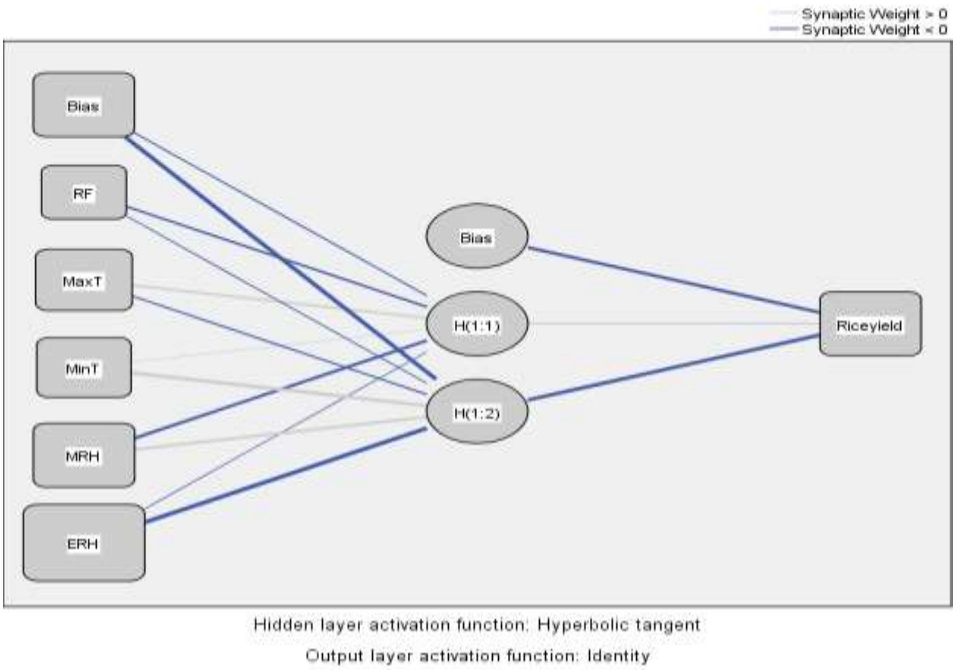


Figure (4.30): Neural Network of Hinthada District

Source: Own Computation

Table (4.35) shows weight values used in the network in the input and hidden layers. The positive and negative values by a grey and blue line are shown in Figure (4.30). The first node in the hidden layer has highly strong bond with maximum temperature. Conversely, the morning relative humidity has weakest bond with the first node in the hidden layer. The second node in the hidden layer has highly strong bond with minimum temperature. Contrary, the evening relative humidity variable has weakest bond with the second node in the hidden layer. Rice yield has positive bond with the first node and negative bond with the second node in the hidden layer.

Table (4.35): Weight Values of the ANN in Hinthada District

Predictor		**Hidden Layer 1**	
		H(1:1)	**H(1:2)**
Input Layer	(Bias)	-0.340	-1.706
	Rainfall	-0.510	-0.161
	Maximum Temperature	0.566	-0.442
	Minimum Temperature	0.027	1.090
	Morning Relative Humidity	-0.735	0.626
	Evening Relative Humidity	-0.045	-2.316
		Output Layer (Rice Yield)	
Hidden Layer	(Bias)	-0.818	
	H(1:1)	0.477	
	H(1:2)	-1.321	

Source: Own Computation

In Hidden layer (1:1), it is seen that rainfall, morning and evening relative humidity have negative effects, and maximum and minimum temperature have positive effects on rice yield. In addition, rainfall, maximum temperature and evening relative humidity have negative effects, and minimum temperature and morning relative humidity have positive effects on rice yield in Hidden layer (1:2).

Table (4.36): Model Summary of Hinthada District

Training	Sum of Squares Error	12.616
	Relative Error	.549
	Stopping Rule Used	1 consecutive step(s) with no decrease in error
Testing	Training Time	0:00:00.05
	Sum of Squares Error	.662
	Relative Error	.205

Source: Own Computation

Table (4.36) shows the sum of square error and relative error in the training and testing phase. The network tries to minimize the sum of square error during the training phase. This error is displayed when the output layer has scale-dependent variables. The model returned the sum of squares error 12.616 and 0.662 in the training and testing phase, respectively. The relative error is 0.549 and 0.205 in the training and testing phase, respectively. The sum of square error with other error values, in this model relative error, is used to compute for the rescaled values of the dependent variables, rice yield.

The neural network output shows that the model is suitable for predicting rice yield in Hinthada. The neural network prediction process stopped when the error did not decrease. These facts show how precise the model is. Moreover, the importance of the independent variables (predictors) was covered in Table (4.37) and Figure (4.31).

Table (4.37): Importance of the Predictors in Model for Hinthada District

Predictor	Importance	Normalized Importance
Rainfall	0.040	9.4%
MaxT	0.168	38.9%
MinT	0.142	33.0%
MRH	0.217	50.2%
ERH	0.432	100.0%

Source: Own Computation

In Table (4.37) and Figure (4.31), the importance of the rainfall variable is almost 9.4%, that of maximum temperature is 38.9%, and that of minimum temperature is 34%, that of morning relative humidity is 50.2% and that of evening relative humidity is 100% respectively in Hinthada District. It is found that the influence of evening relative humidity on the change of rice yield in the Hinthada District is the most. Additionally, it was discovered that the morning relative humidity had almost 50 percent impact on the change in rice yield. Furthermore, variations in minimum and maximum temperatures only have a minor impact on rice yield. However, rainfall has the least impact on changes in rice yield.

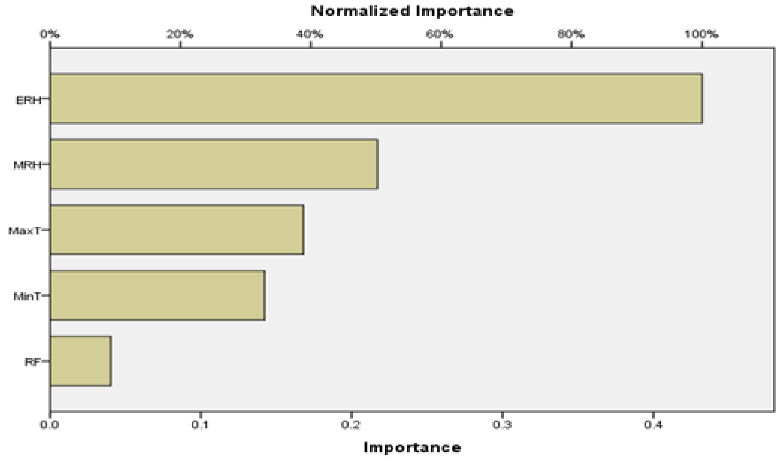


Figure (4.31): Normalized Importance of the Predictors in Hinthada district

Source: Own Computation

Maubin District

In this district, a single hidden layer with one neuron was employed. Table (4.38) shows the network information contains five covariates excluding a bias unit used as an input unit.

Table (4.38): Network Information of Maubin District

Input Layer	Covariates	1,2,3,4,5
	Number of units	5
	Rescaling Method for Covariates	Standardized
Hidden Layer(s)	Number of Hidden Layers	1
	Number of Units in Hidden Layers 1	1
	Activation Function	Hyperbolic tangent
Output Layer	Dependent Variables 1	Riceyield
	Number of Units	1
	Rescaling Method for Scale Dependents	Standardized
	Activation Function	Identity
	Error Function	Sum of Squares

Source: Own Computation

Single hidden layer with one neuron in the network in Maubin district was used in the network. Rice yield is a variable used as an output of this network. This network information was shown in Figure (4.32).

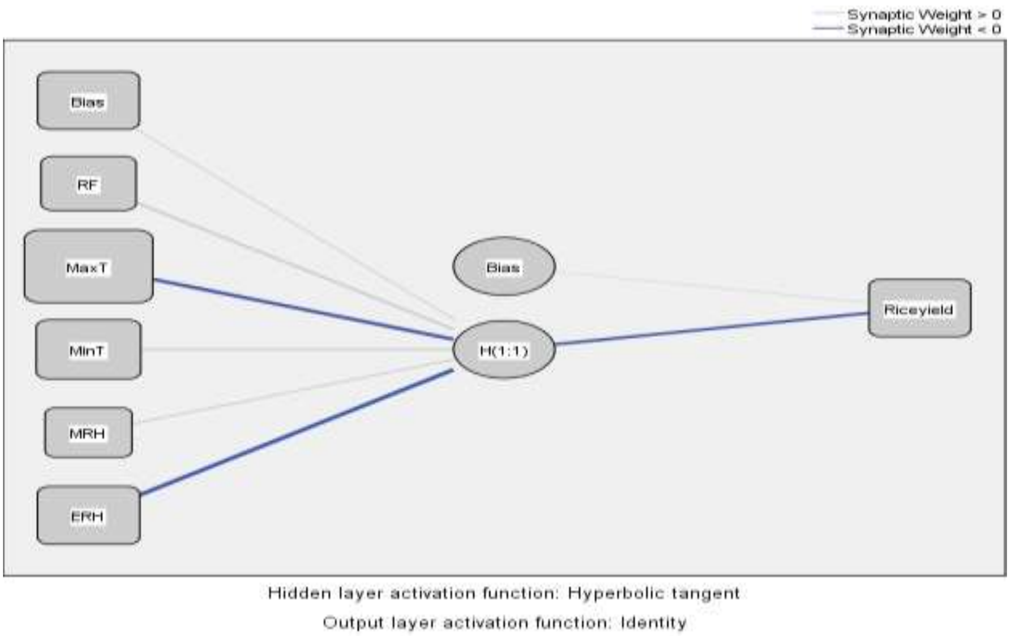


Figure (4.32): Neural Network of Maubin District

Source: Own Computation

Figure (4.32) depicts three layers of information, including an input, one hidden, and an output layer. The input layer consists of six units, including a bias unit. The hidden layer possesses one neuron and one bias unit. The identity activation function and the hyperbolic tangent, respectively, are the best options for this network's input unit and output unit, according to the automatic architectural selection. It is seen that the node in the hidden layer has highly strong bond with rainfall. On the other hand, the relationship between the node in the concealed layer and evening relative humidity is the weakest. The hidden layer's node and rice yield are negatively bonded.

Table (4.39): Weight Values of the ANN in Maubin District

Predictor		**Hidden Layer**
		H(1:1)
Input Layer	(Bias)	0.284
	Rainfall	0.671
	Maximum Temperature	-1.113
	Minimum Temperature	0.640
	Morning Relative Humidity	0.293
	Evening Relative Humidity	-1.232
		Output Layer
Hidden Layer	(Bias)	0.213
	H(1:1)	-1.103

Source: Own Computation

Table (4.39) shows that maximum temperature and evening relative humidity have negative effects, and rainfall, minimum temperature and morning relative humidity have positive effects on rice yield in Hidden layer.

Table (4.40): Model Summary of Maubin District

Training	Sum of Squares Error	6.990
	Relative Error	0.304
	Stopping Rule Used	1 consecutive step(s) with no decrease in error
Testing	Training Time	00:00.02
	Sum of Squares Error	0.273
	Relative Error	0.062

Source: Own Computation

Table (4.40) shows that the model returned the sum of squares error 6.990 and 0.273 in the training and testing phase, respectively. The relative error is 0.304 and 0.062 in the training and testing phase, respectively. The neural network prediction process stopped when the error did not decrease. These points are evidence the accuracy of the model. Moreover, the importance of the predictor variables is described in Table (4.41) and Figure (4.33).

Table (4.41): Importance of the Predictors in Maubin District

Predictor	Importance	Normalized Importance
Rainfall	0.141	37.8%
MaxT	0.374	100.0%
MinT	0.207	55.3%
MRH	0.085	22.6%
ERH	0.192	51.3%

Source: Own Computation

In the Table (4.41) and Figure (4.33), the importance of rainfall variable is 37.8%, that of maximum temperature is 100%, and that of minimum temperature is 55.3%, that of morning relative humidity is 22.6% and that of evening relative humidity is 51.3% respectively in Maubin district. It is found that the influence of maximum temperature on the change of rice yield in Maubin District is the most. In addition, the influence of minimum temperature and evening relative humidity on the change in rice yield was found to be half a percent. However, the influence of rainfall and morning relative humidity on the change in rice yield is found to be only a few percent.

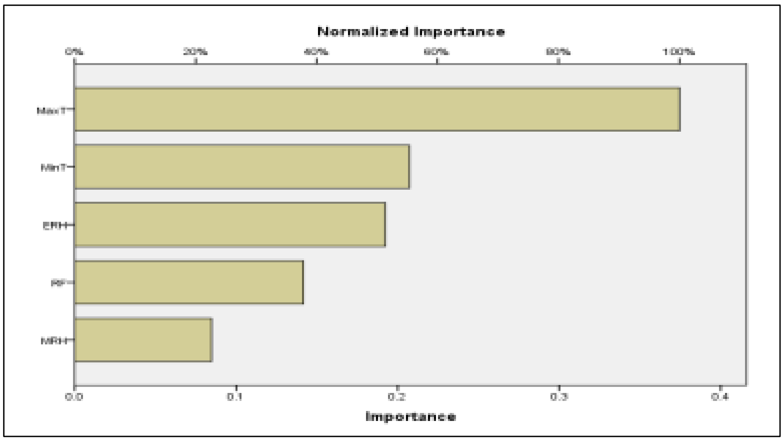


Figure (4.33): Normalized Importance of the Predictors in Maubin District

Source: Own Computation

Myaungmya District

In this district, single hidden layer with three neurons were employed. The network information that contains five covariates, excluding a bias unit, which are used as input units is shown in Table (4.42).

Table (4.42): Network Information of Myaungmya District

Input Layer	Covariates	1,2,3,4,5
	Number of units	5
	Rescaling Method for Covariates	Standardized
Hidden Layer(s)	Number of Hidden Layers	1
	Number of Units in Hidden Layers 1	3
	Activation Function	Hyperbolic tangent
Output Layer	Dependent Variables 1	Riceyield
	Number of Units	1
	Rescaling Method for Scale Dependents	Standardized
	Activation Function	Identity
	Error Function	Sum of Squares

Source: Own Computation

Single hidden layer with two neurons in the network in Myaungmya district is used in the network. Rice yield is a variable used as an output of this network. This network information was shown in Figure (4.34).

Three layers of information are shown in Figure (4.34), comprising an input, a hidden layer and an output layer. The input layer contains a bias unit and a total of six units. The hidden layer contains one bias unit and three neurons. It was discovered that the input unit and output unit for this network is suitable for the input unit and identity activation functions of the network, respectively, by applying automatic architectural selection.

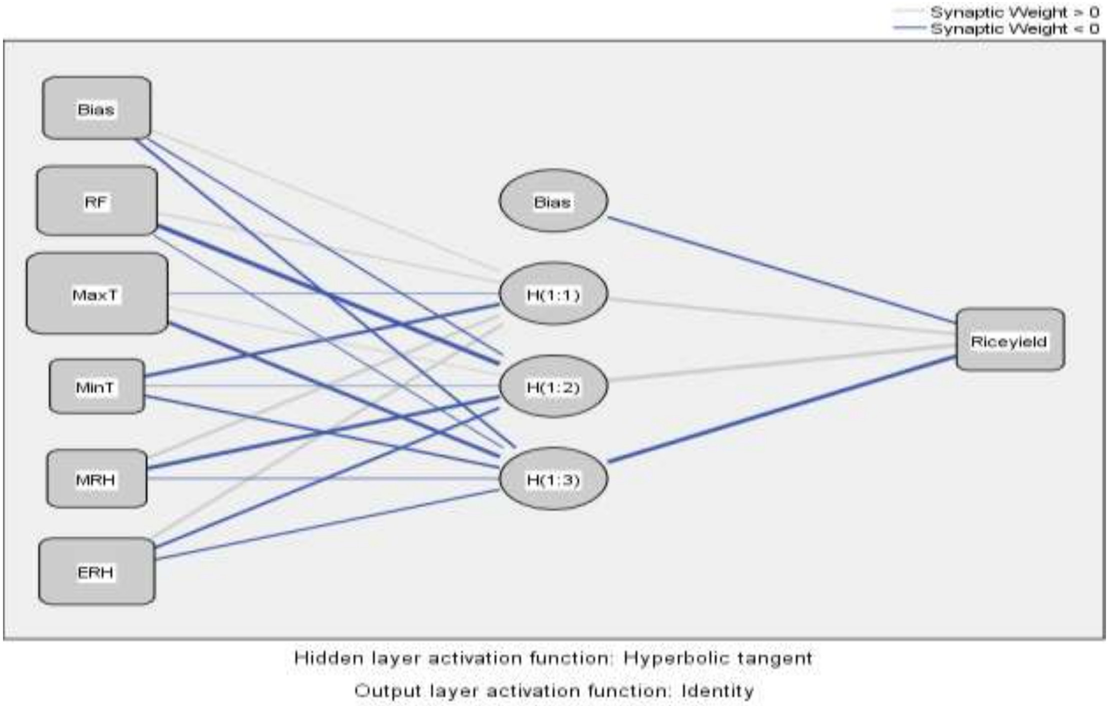


Figure (4.34): Neural Network of Myaungmya District

Source: Own Computation

It is seen that the first node in the hidden layer has highly strong bond with morning relative humidity. Contrary, rainfall has weakest bond with the second node in the hidden layer. Rice yield has a strong positive bond with the second node and a negative bond with the third node in the hidden layer.

Table (4.43): Weight Values of the ANN in Myaungmya District

Predictor		**Hidden Layer 1**		
		H(1:1)	H(1:2)	H(1:3)
Input Layer	(Bias)	.275	-0.288	-0.346
	Rainfall	.294	-1.711	-0.142
	Maximum Temperature	-0.042	0.233	-0.685
	Minimum Temperature	-0.660	-0.120	-0.370
	Morning Relative Humidity	0.524	-0.735	-0.099
	Evening Relative Humidity	0.409	-0.401	-0.318
		Output Layer (Rice Yield)		
Hidden Layer	(Bias)	-0.364		
	H(1:1)	0.550		
	H(1:2)	1.068		
	H(1:3)	-0.959		

Source: Own Computation

Table (4.43) shows that in hidden layer (1:1), maximum and minimum temperature have negative effects, and rainfall, morning and evening relative humidity have positive effects on rice yield. In hidden layer (1:2), rainfall, minimum temperature, morning and evening relative humidity have negative effects, and maximum temperature has positive effects on rice yield. All input variables in hidden layer (1:3), have negative effects on rice yield.

Table (4.44): Model Summary of Myaungmya District

Training	Sum of Squares Error	5.093
	Relative Error	0.221
	Stopping Rule Used	1 consecutive step(s) with no decrease in error
Testing	Training Time	00:00.02
	Sum of Squares Error	1.051
	Relative Error	0.281

Source: Own Computation

Table (4.44) shows that the model returned the sum of squares error 5.093 and 1.051in the training and testing phase respectively. The relative error is 0.221 and 0.281 in the training and testing phase respectively. The neural network output shows that the model is suitable to predict rice yield in Myaungmya district. The neural network prediction process stopped when the error did not decrease. These points are evidence the accuracy of the model. Moreover, the importance of the independent variables (predictors) is described in Table (4.45) and Figure (4.35).

Table (4.45): Importance of the Predictors in Myaungmya District

Predictor	Importance	Normalized Importance
Rainfall	0.230	68.7%
MaxT	0.335	100.0%
MinT	0.102	30.3%
MRH	0.127	37.8%
ERH	0.206	61.6%

Source: Own Computation

In the Table (4.45) and Figure (4.35), the importance of rainfall variable is 68.7%, that of maximum temperature is 100%, and that of minimum temperature is 30.3%, that of morning relative humidity is 37.8% and that of evening relative humidity is 61.6% respectively in Myaungmya district. It is found that the influence of maximum temperature on the change of rice yield in Myaungmya District is the most, and the influence of rainfall and evening relative humidity is not as much as the influence of maximum temperature. However, the influence of morning relative humidity and minimum temperature on the change in rice yield is found to be only a few percent.

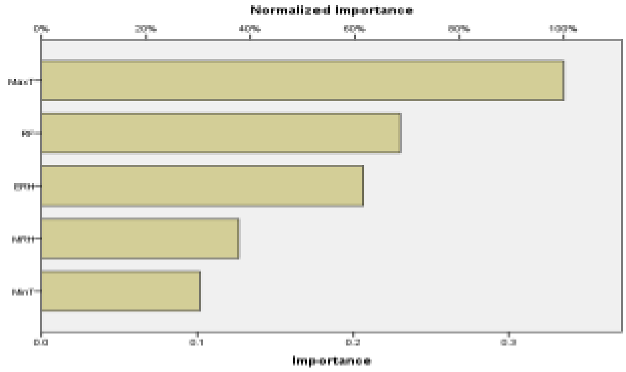


Figure (4.35): Normalized Importance of Predictors in Myaungmya District

Source: Own Computation

Phyarpon District

In this district, single hidden layer with seven neurons were employed. The network information that contains five covariates, excluding a bias unit, which are used as input units is shown in Table (4.46) and Figure (4.36).

Table (4.46): Network Information of Phyarpon District

Input Layer	Covariates	1,2,3,4,5
	Number of units	5
	Rescaling Method for Covariates	Standardized
Hidden Layer(s)	Number of Hidden Layers	1
	Number of Units in Hidden Layers 1	7
	Activation Function	Hyperbolic tangent
Output Layer	Dependent Variables 1	Riceyield
	Number of Units	1
	Rescaling Method for Scale Dependents	Standardized
	Activation Function	Identity
	Error Function	Sum of Squares

Source: Own Computation

Figure (4.36) depicts three layers of information including an input, one hidden, and an output layer. The input layer consists of six units, including a bias unit. The hidden layer possesses with seven neurons and one bias unit.

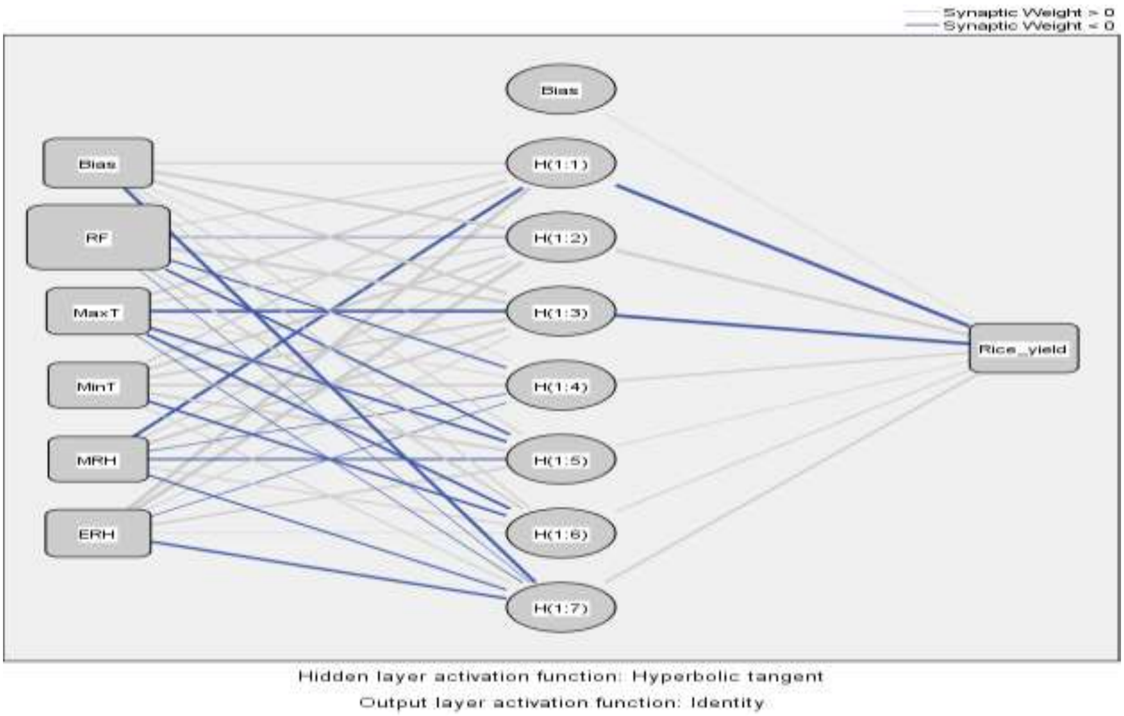


Figure (4.36): Neural Network of Phyarpon District

Source: Own Computation

It is seen that the second node in the hidden layer has the strongest bond with ERH. Contrast, the maximum temperature has weak bond with the third node in the hidden layer. Besides, rice yield has a negative bond with the first node and the third node in the hidden layer and a negative bond with the rest of nodes in the hidden layer. Moreover, the weight values associated with the ANN (5-7-1) model are described in Table (4.47).

Table (4.47) shows the maximum and minimum temperature, as well as the relative humidity in the evening, have a positive impact on rice yield and rainfall, whereas the relative humidity in the morning has a negative effect on both in hidden layer (1:1). Minimum and maximum temperatures, as well as morning and evening relative humidity, have a positive impact in the hidden layer (1:2) of rice yield, whereas rainfall has the opposite effect. In hidden layer (1:3), maximum temperature has negative effects on rice yield, despite the positive effects of the other input variables.

Table (4.47): Weight Values of the ANN in Phyarpon District

Predictor		Hidden Layer 1							Output Layer
		H(1:1)	**H(1:2)**	**H(1:3)**	**H(1:4)**	**H(1:5)**	**H(1:6)**	**H(1:7)**	**Riceyield**
Input Layer	(Bias)	.258	.885	.686	.031	.096	.117	-.387	
	RF	.143	-.086	1.297	-.249	-.311	.287	-.005	
	MaxT	.481	.196	-.803	.094	-.399	-.337	-.037	
	MinT	.271	.571	.487	.415	.469	-.314	.133	
	MRH	-.457	.207	.746	-.085	-.279	.197	-.208	
	ERH	1.006	1.387	.135	-.067	.328	.013	-.355	
Hidden Layer 1	(Bias)								.029
	H(1:1)								-.869
	H(1:2)								.825
	H(1:3)								-1.259
	H(1:4)								.305
	H(1:5)								.080
	H(1:6)								.172
	H(1:7)								.290

Source: Own Computation

Rice yield is negatively impacted by rainfall, morning and evening relative humidity, as well as maximum and minimum temperatures in hidden layer (1:4). Rainfall, the maximum temperature, and the morning relative humidity all have a detrimental impact on the yield of rice in hidden layer (1:5), whereas the minimum temperature and the evening relative humidity have a positive impact. The maximum and minimum temperatures have negative effects on rice yield in hidden layer (1:6), while the other input variables have positive effects. In hidden layer (1:7), only the minimum temperature affects rice yield favorably; all other input variables have an adverse effect.

Table (4.48): Model Summary of Phyarpon District

Training	Sum of Squares Error	0.895
	Relative Error	0.062
	Stopping Rule Used	1 consecutive step(s) with no decrease in error
Testing	Training Time	00:00.00
	Sum of Squares Error	0.058
	Relative Error	0.013

Source: Own Computation

Table (4.48) shows that the model returned the sum of squares error 0.895 and 0.058 in the training and testing phase, respectively. The relative error is 0.062 and 0.013 in the training and testing phase, respectively. The neural network output shows that the model is suitable for predicting rice yield in Phyarpon District. These points give evidence for the accuracy of the model. In addition, the importance of the independent variables (predictors) is described in Table (4.49) and Figure (4.37) respectively.

Table (4.49): Importance of the Predictors in Phyarpon District

Predictor	Importance	Normalized Importance
Rainfall	0.386	100.0%
MaxT	0.161	41.7%
MinT	0.142	36.7%
MRH	0.139	36.1%
ERH	0.172	44.5%

Source: Own Computation

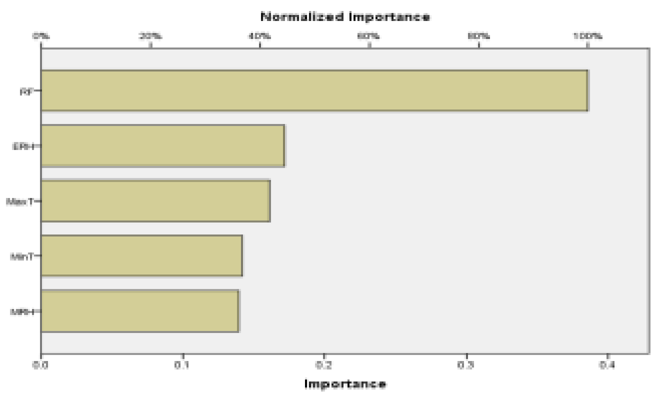


Figure (4.37): Normalized Importance of Predictors in Phyarpon District

Source: Own Computation

In Table (4.49) and Figure (4.37), the importance of rainfall variable is 100%, that of maximum temperature is 41.7%, which of minimum temperature is 36.7%, that of morning relative humidity is 36.1% and that of evening relative humidity is 44.5% respectively in Phyarpon District. It is found that the influence of rainfall on the change of rice yield in Phyarpon District is the most. The influence of morning relative humidity, evening relative humidity, maximum temperature and minimum temperature are not as much as the influence of rainfall. It can be obviously seen in Figure (4.37). Continuously, the most appropriate model of four models was chosen, as in Section (4.10).

4.10 Identification of Appropriate Model

In order to select the most appropriate model for forecasting rice yield on climatic variables, the forecast accuracy measures such as Root Mean Squared Error (RMSE) and Mean Absolute Percent Error (MAPE) and Coefficient of Determination (R^2) are used in the study. Generally, a relatively low in RMSE, MAPE and relatively high in R^2 if the response variable is well correlated with the predictor variables. These values are also presented in Table (4.50).

In Pathein District, the RMSE measure for the multiple linear regression (MLR) model is 5.3097, for the autoregressive integrated moving average with predictors (ARIMAX) model is 2.2671, for the vector autoregressive (VAR) model is 6.8721 and lowest at 0.2697 for the artificial neural network (ANN) model. Moreover, the MAPE obtained for the series from the MLR model is 5.4103, the ARIMAX model is 2.1721, the VAR model is 7.4806 and through the ANN model is 0.4793 which is smallest representing, and it is an appropriate model. Further, the coefficient of determination (R^2) is observed to be highest for the artificial neural network at 0.9994 in comparison to 0.7521, 0.9571, and 0.9618 for multiple linear regression, for autoregressive integrated moving average with predictor variables (ARIMAX) and vector autoregressive (VAR) models respectively. By these facts, the ANN model was chosen as the most appropriated model among the four models proposed.

Table (4.50): Comparison of the Performance of Forecasting Models for Rice Yield

District	Accuracy Measures	MLR	ARIMAX	VAR	ANN
Pathein	RMSE	5.3097	2.2671	6.8721	0.2697
	MAPE	5.4103	2.1721	7.4806	0.4793
	R^2	0.7521	0.9571	0.9618	0.9994
Hinthada	RMSE	9.6976	5.9009	9.7720	0.4785
	MAPE	9.7629	4.8566	10.5759	1.3000
	R^2	0.3649	0.7808	0.7696	0.9985
Maubin	RMSE	6.5572	2.6280	7.6431	0.3539
	MAPE	6.8024	2.4602	8.2780	0.6310
	R^2	0.7155	0.9553	0.9800	0.9992
Myaungmya	RMSE	7.5454	2.9877	7.5606	0.3254
	MAPE	8.2258	2.4769	8.2434	0.8655
	R^2	0.5896	0.9411	0.9733	0.9993
Phyarpon	RMSE	7.3877	2.3409	8.5721	0.1544
	MAPE	8.6401	1.9473	10.8318	0.1875
	R^2	0.7997	0.6929	0.9915	0.9992

Source: Own Computation

The RMSE for the series from the MLR, ARIMAX, and VAR models in the Hinthada District is 9.6976, 5.9009, and 9.7720, respectively, while the RMSE for the ANN model is 0.4785 and the lowest. The MAPE measure is also 9.7629 for the MLR model, 4.8566 for the ARIMAX model, 10.5759 for the VAR model, and 1.3000 for the ANN model, which is the lowest. Additionally, the ANN model's R^2 is seen to be the highest, coming in at 0.9985, compared to the ARIMAX and VAR models' respective MLR values of 0.3649, 0.7808, and 0.7696. As a result, when compared to the other models, the ANN model performs the best.

In the Maubin District, it can be observed that the RMSE for the series is lesser for artificial neural network than the values obtained from MLR, ARIMAX, and VAR because the values for MLR, ARIMAX, VAR, and ANN are 6.5572, 2.6280, 5.6142 and 0.3539. Besides, the MAPE obtained for the series from the MLR model is 7.6431, the ARIMAX model is 2.4837, the VAR model is 5.8609, and the ANN model is 0.6310 and the smallest. In addition, R^2 values for ARIMAX, VAR, and MLR models are

0.9553, 0.9800, and 0.7155, respectively, while the highest value, 0.9992, for the ANN model designates it as the preferable model.

In the Myaungmya District, it can be seen that the RMSE for the series is 0.3254 for the artificial neural network model is the lowest among all the others compared to 7.7454 for MLR, 2.9877 for ARIMAX, and 7.5606 for VAR models. In addition, the MAPE for the MLR model is 8.2258, for the ARIMAX model is 2.4769, and for the VAR model is 8.2434 while the measurement is noticeably lower than 0.8655 for the ANN model. Moreover, R^2 is observed to be highest for the ANN model at 0.9993 in comparison to 0.5896, 0.9411, and 0.9733 for MLR, for ARIMAX and VAR models, respectively. It indicates the ANN model as the preferred model.

In the Phyarpon District, the RMSE measure for the multiple linear regression (MLR) models is 7.3877, for the autoregressive integrated moving average with predictors (ARIMAX) model is 2.3409, for the vector autoregressive (VAR) model is 8.5721 and lowest at 0.1544 for the ANN model. Moreover, the MAPE obtained for the series from the MLR model is 8.6401, the ARIMAX model is 1.9473, the VAR model is 10.8318, and the ANN model is 0.1875 is the smallest. Further, R^2 is observed to be highest for the ANN model at 0.9992 in comparison to 0.7997, 0.6929, and 0.9915 for MLR, for ARIMAX and for VAR models, respectively. Based on the results, the artificial neural network model has a better performance and can be preferred as a reliable rice yield forecasting tool in all districts. Consequently, the rice yield is forecasted using the appropriate model, the artificial neural network, in the next section.

Generally, the main contributing variables of MLR, SARIMAX and VAR models and that of the ANN model are the similar in some districts. Additionally, in almost districts, minimum temperature and relative humidity have positive effect on rice yield, while rainfall has negative effect on rice yield and maximum temperature has not only the negative effect but also positive effect in some districts.

The rainfall and evening relative humidity have slight positive effect and slight negative effect on rice yield respectively in Pathein District. In Maubin District, although maximum temperature positively effects on rice yield in MLR, SARIMAX and VAR models, it negatively effects on rice yield in ANN model. This means that the raising the temperature can lower rice yield. Additionally, it can be seen that these results are consistent with those of other studies such as Farook & Kannan (2015), Rahman et.al. (2017) and Chowdhury & Khan (2015).

4.11 Forecasting the Rice Yield

In this section, the forecasting for rice yield was conducted based on the most appropriated model, ANN model, during the periods from 2021-2022 in monsoon to 2023-2024 in summer. When the forecasted values for rice yield are computed by the ANN model, input variables as rainfall, maximum temperature, minimum temperature, morning relative humidity and evening relative humidity are assumed as the random values ranging between mean or minimum and maximum values in monsoon and summer. The forecasted results are depicted by districts in Figures (4.38).

When the rice yield are forecasted by the ANN model in the Pathein District, input variables are presumptively random values between the mean and maximum values during monsoon and summer. Because neither the monsoon nor the summer input variables include extreme values. Thus, rainfall lies between 478 mm and 586 mm, maximum temperature lie between 31˚C and 33˚C, the minimum temperature lies between 23˚C and 25˚C, morning relative humidity lies between 89% and 90%, and evening relative humidity lies between 90% and 97% as monsoon and rainfall lie between 20 mm and 64 mm, maximum temperature lie between 34˚C and 36˚C, minimum temperature lie between 20˚C and 22˚C, morning relative humidity lie between 75% and 80% and evening relative humidity lies between 76% and 96% as summer. The forecasted rice yield (Bsk/Ac) from 2021-2022 to 2023-2024 is 73.68, 86.86, 74.04, 88.65, 71.3, and 92.29. Moreover, it can be seen that the forecasted values for the rice yield fluctuated down in monsoon and up in summer because there is irrigated paddy cultivation in summer and due to lack of heavy rain.

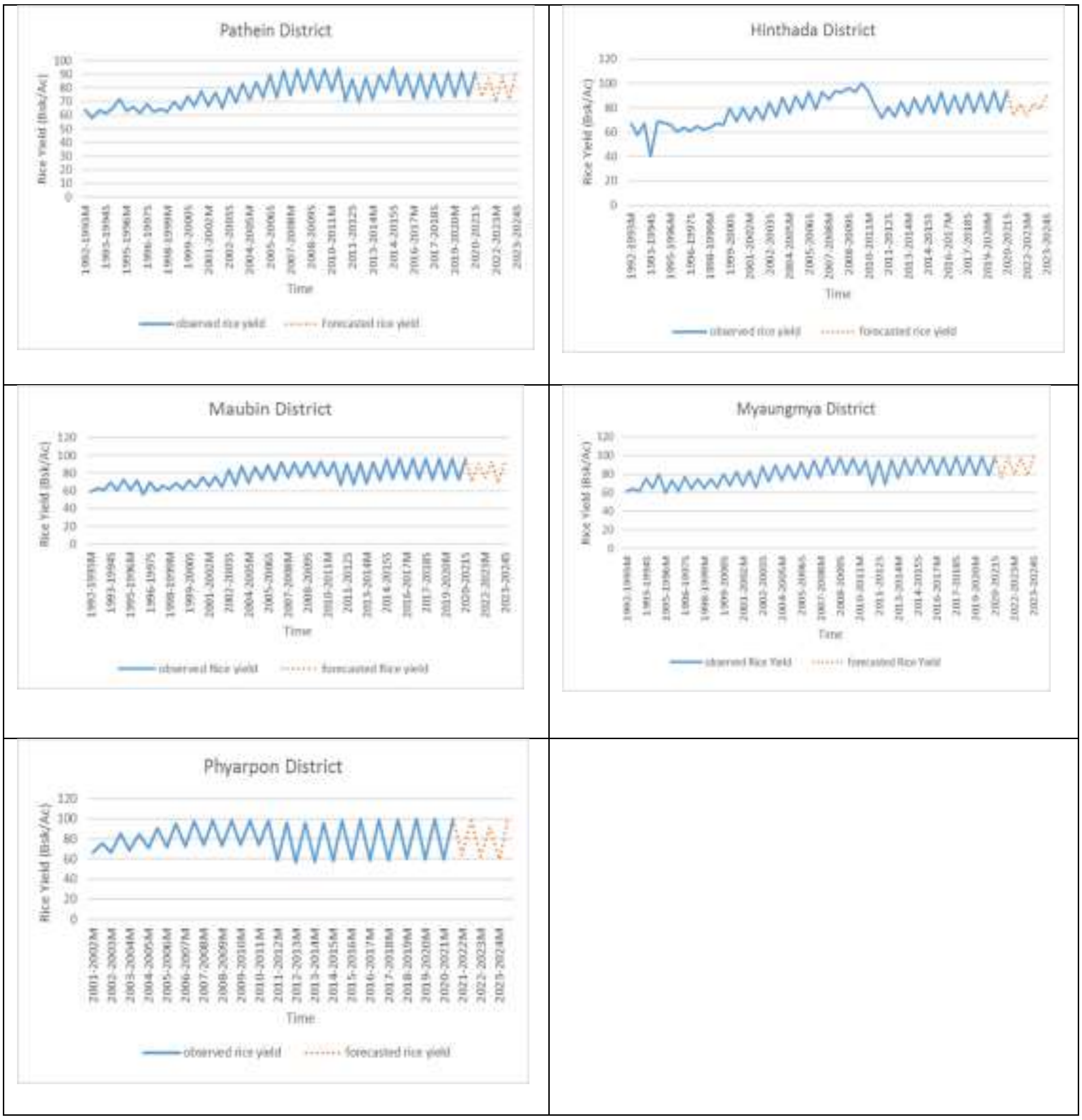


Figure (4.38): Forecasting Results of ANN Model by Each District

Source: Own Computation

When the forecasted values for rice yield in Hinthada District are computed using the ANN model, input variables include rainfall between 360 mm and 486 mm, maximum temperature lie between 32˚C and 34˚C, minimum temperature between 24˚C and 25˚C, morning relative humidity between 86% and 90% and evening relative humidity between 90% and 96% as monsoon and rainfall between 14 mm and 73 mm, maximum temperature between 34˚C and 36˚C, minimum temperature lie between 24˚C and 25˚C, morning relative humidity lie between 70% and 78% and evening relative humidity lie between 74% and 91% as summer. From 2021-2022 to 2023-2024, the predicted rice yield (Bsk/Ac) is expected to be 73.48, 83.26, 73.47, 83.96, 78.64, and

89.29. Furthermore, it is clear that due to the irrigated paddy cultivation in the summer and the lack of heavy rain, the forecasted values for the rice yield fluctuated upward in the summer and downward in the monsoon.

In Maubin District, input variables as rainfall, maximum temperature, minimum temperature, morning relative humidity and evening relative humidity are assumed the random values between mean and maximum values in monsoon and summer when the forecasted values for rice yield are computed by the ANN model. Because of there are not included extreme values in the input variables for both monsoon and summer. Thus, rainfall lie between 410 mm and 521 mm, maximum temperature lie between 26˚C and 32˚C, minimum temperature lie between 16˚C and 25˚C, morning relative humidity lie between 80% and 90% and evening relative humidity lie between 81% and 96% as monsoon and rainfall lie between 18 mm and 59 mm, maximum temperature lie between 31˚C and 35˚C, minimum temperature lie between 11˚C and 21˚C, morning relative humidity lie between 61% and 78% and evening relative humidity lie between 74% and 91% as summer. The forecasted rice yield (Bsk/Ac) from 2021-2022 to 2023-2024 are 70.56, 90.05, 74.89, 93.37, 68.84 and 92.31. Moreover, it can be seen that the forecasted values for the rice yield fluctuated down in monsoon and up in summer because there are irrigated paddy cultivation in summer and due to lack of heavy rain.

In Myaungmya District, input variables as rainfall, maximum temperature, minimum temperature, morning relative humidity and evening relative humidity are assumed the random values between mean and maximum values in monsoon and summer when the forecasted values for rice yield are computed by the ANN model. Because the input variables for the monsoon and summer seasons do not include extreme values. Thus, rainfall lie between 457 mm and 526 mm, maximum temperature lie between 30˚C and 34˚C, minimum temperature lie between 15˚C and 24˚C, morning relative humidity lie between 78% and 87% and evening relative humidity lie between 83% and 92% as monsoon and rainfall lie between 15 mm and 46 mm, maximum temperature lie between 27˚C and 34˚C, minimum temperature lie between 12˚C and 22˚C, morning relative humidity lie between 75% and 80% and evening relative humidity lie between 50% and 90% as summer. The forecasted rice yield (Bsk/Ac) from 2021-2022 to 2023-2024 are 76.26, 89.44, 77.27, 97.92, 68.12 and 90.99. Moreover, it can be seen that the forecasted values for the rice yield fluctuated down in monsoon and up in summer because there is irrigated paddy cultivation in summer and due to lack of heavy rain.

Input variables such as rainfall, maximum temperature, minimum temperature, morning relative humidity, and evening relative humidity are assumed to be the random values between mean and maximum values in monsoon and summer in the Phyarpon district when the forecasted values for rice yield are calculated by the ANN model. Because the input variables don't contain any extreme value for both monsoon and summer. Thus, rainfall lie between 500 mm and 605 mm, maximum temperature lie between 30˚C and 32˚C, minimum temperature lie between 17˚C and 26˚C, morning relative humidity lie between 85% and 90% and evening relative humidity lie between 86% and 98% as monsoon and rainfall lie between 21 mm and 72 mm, maximum temperature lie between 30˚C and 34˚C, minimum temperature lie between 13˚C and 23˚C, morning relative humidity lie between 67% and 86% and evening relative humidity lie between 67% and 93% as summer. The forecasted rice yield (Bsk/Ac) from 2021-2022 to 2023-2024 are 60.96, 97.95, 61.87, 98.64, 59.35 and 101.74. It can be seen that the forecasted values for the rice yield fluctuated down in monsoon and up in summer because there are irrigated paddy cultivation in summer and due to lack of heavy rain.

Continuously, the forecasted values for 2021-2022 and 2022-2023 in monsoon and summer are compared the actual values for those periods. But, the actual values for 2023-2024 are not available because the paddy does not harvest during the study periods. The compared results are described in Table (4.51).

Table (4.51): Comparing Results of the Forecast and Actual Values by Districts

Year	**2021-2022**				**2022-2023**			
Season	**Monsoon**		**Summer**		**Monsoon**		**Summer**	
Riceyield (Bsk/Ac)	Forecast	Actual	Forecast	Actual	Forecast	Actual	Forecast	Actual
Pathein	73.68	74.65	86.86	94.60	74.04	76.84	88.65	98.76
Hinthada	73.48	77.8	83.26	95.54	73.47	77.39	83.96	NA
Maubin	70.56	73.49	90.05	96.70	74.89	75.23	93.37	99.89
Myaungmya	76.26	78.92	89.44	98.19	77.27	82.65	97.92	101.53
Phyarpon	60.96	59.58	97.95	99.92	61.87	60.84	98.64	103.20

NA: not available

Source: Own Computation and Department of Agriculture (Ayeyawady Region)

As the results of Table (4.51), it can be seen that there is not much difference between the forecast rice yield and the actual rice yield in both monsoon and summer in 2021-2022 and 2022-2023. In detail, the results are close in monsoon and slightly underestimation in summer but fall within the 95% confidence interval. Therefore, it can be concluded that there are few variations in the forecast and the actual by using the Artificial Neural Network (ANN) model.

Ayeyawady Region is divided into eight districts after the 2021-2022 summer period. Therefore, Pathein District is divided into Kyonpyaw District and Hintatha District is divided into Myanaung District. The yield of summer rice is higher than that of monsoon rice. This is because a high percentage of successful seeds can only be obtained in the period of full temperature during the physiological stage of reproduction and ripening stage.

Although rice production fluctuates slightly from year to year, it is seen that the average yield per acre is high. The reasons are that some problems may be slightly reduced, such as the costs associated with rice farming inputs, including seeds, fertilizers, labor, rice farming has become more arduous for farmers due to higher input costs, diminishing their yields' profitability. In addition, the lack of access to resources, credit, and modern farming techniques that limit the ability to achieve high rice yields has also been slightly solved (USAID, 2023).

CHAPTER V
CONCLUSION

This chapter covers the findings, suggestions, and the need for further studies. Firstly, the findings are presented based on the results gained from a number analyses. Then, the relevant facts are suggested. Finally, the important considerations are presented for the further studies provided through critically analyzing the findings.

5.1 Findings and Discussions

In this study, Multiple Linear Regression (MLR), the Seasonal Autoregressive Integrated Moving Average with predictor variables (SARIMAX), Vector Auto-Regressive (VAR) model and Artificial Neural Network (ANN) model are used to analyze the impact of climate change on rice production in Ayeyawady Region. Climatic variables considered in this study are rainfall (RF), maximum temperature (MaxT), minimum temperature (MinT), morning relative humidity (MRH) and evening relative humidity (ERH). Five districts that have a Meteorology and Hydrology Department in the Ayeyawady region such as Pathein, Hinthada, Maubin, Myaungmya and Phyarpon are included in this study. The study period is from 1992-1993 to 2020-2021, focused on monsoon (May-October) and summer (November-April) for all districts except the data for Phyarpon in which the study period is from 2001-2002 to 2020-2021. The required data are collected from the Department of Agriculture in the Ayeyawady Region and the Department of Meteorology and Hydrology, Myanmar.

The descriptive statistics are computed for each of the climatic variables and rice yield for all districts. The observations are not widely spread out in all variables for both seasons but there is small variation in rainfall for Pathein District (summer), Maubin District (summer) and Myaungmya District (monsoon). In addition, the proposed four models are developed to choose the most appropriate model for the study.

By the results, since some of the predictor variables in MLR model does not have the significant effect and the assumptions are not met, SARIMAX model was explored continuously. Although the assumptions are met in this model, some of the predictor variables does not have the significant effect. Then, some of the predictors has not only the significant effect but also the causality effect in VAR model which was conducted to overcome this problem. Thus, the ANN model was revealed to prevail over these problems. Continuously, the detail results are presented.

Firstly, the accuracy of MLR model for Pathein District is 75.21% and the regression model has a good fit. All predictor variables are statistically significant except MinT. RF has significantly negative effect and MaxT, MRH and ERH have significantly positive effects on rice yield. Although MinT is not significant, there is negative effect on rice yield.

The accuracy The MLR model for the Hinthada District is 36.49%, and the regression model fits the data well. MRH greatly reduces rice production when ERH significantly increases it since both variables are statistically significant. The rice yield is simultaneously impacted negatively by RF and favorably by MaxT and MinT. The accuracy MLR model for the Maubin District is 71.55%, and the regression model fits the data well. RF, MinT, and ERH are statistically significant, and as a result, RF and MinT have a considerable negative effect on rice yield, whereas ERH has a significantly positive impact.

For Myaungmya District, the accuracy of MLR model is 58.96% and the regression model has a good fit. Since RF, MaxT and ERH have statistically significant, RF has significantly negative effects and MaxT and ERH have significantly positive effects on rice yield. For Phyarpon District, the accuracy of MLR model is 79.97% and the regression model has a good fit. Since RF is statistically significant, it have significantly negative effects on rice yield.

Secondly, the Seasonal AutoRegressive Integrted Moving Average with predictor variables (SARIMAX) model is developed for all districts. The data series for the Pathein District is chosen to be represented by the SARIMAX $(0,0,1)(1,1,3)_2$ model. MaxT and MRH dramatically reduce rice yield by 10% and 5% level of significance, respectively. The data series for the Hinthada District is chosen to be represented by the ARIMAX $(2,0,0)(0,1,0)_2$ model. Even though none of the predictor variables are statistically significant, all of the predictor coefficients except RF have negative effect on rice yield. The data series for the Maubin District are chosen to be represented by the ARIMAX $(0,1,1)(0,1,0)_2$ model. The coefficients of all predictor variables are not statistically significant. But, RF, MRH and ERH have negative effects and MaxT and MinT have positive effects on rice yield. The data series for the Myaungmya District is chosen to be represented by the ARIMAX $(1,1,1)(0,1,0)_2$ model. The coefficient of RF is statistically significant effect on rice yield at 10% level. The ARIMAX $(0,1,1)(0,1,0)_2$ model is chosen to represent the data series for Phyarpon District. The predictor variables in this model, including the RF, MinT, and MRH variables, are statistically significant at 1%. The MRH has positive effects on rice yield, whereas the RF and MinT have negative effects.

The VAR(1) models for all districts using parsimonious are selected to study the impact of climatic variables on rice yield. In Pathein District, there is causal relationship. Besides, 96.18% variation can be explained by the climatic variables for rice yield. The RF and MinT have effect on rice yield positively. The other variables such as MaxT, MRH and ERH are negative effect on rice yield. The volatility of yield from 0% to 14.25% fluctuations at highest can be explained by rainfall. In Hinthada District, there is no causal relationship and 76.96% variation can be explained by the climatic variables for rice yield. It reveals that RF, MinT and MRH have a positive effects on rice yield and the other variables such as MaxT and ERH have negative effects on rice yield. The volatility of yield from 0% to 12.40% fluctuations at highest can be explained by RF.

There is no causal relationship in Maubin District. But, 98% variation can be explained by the climatic variables for rice yield. The climate variables such as RF and MRH have a positive effect and MaxT, MinT and ERH have negative effects on rice yield. The volatility of yield from 0% to 23.18% fluctuations at the highest can be explained by MRH. In Myaungmya District, there is causal relationship the rice yield and the maximum temperature. But, 97.33% variation can be explained by the climatic variables for rice yield. The climatic variables such as RF, MinT and ERH have a positive effect and MaxT and MRH and ERH have negative effects on rice yield. The volatility of yield from 0% to 10.92% fluctuations at the highest can be explained by minimum temperature. In Phyarpon District, there is causal relationship the rice yield and RF, MinT and MRH. Moreover, 99.15% variation can be explained by the climatic variables for rice yield. The climatic variables except ERH have a positive effects on rice yield. The volatility of yield from 0% to 22.37% fluctuations at the highest can be explained by rainfall.

Moreover, the neural network configuration is 81% of the total available data for training purpose and 19% of data for testing purpose for all districts except Phyarpon District in which there is 75% of the total available data for training purpose and 25% of the data for testing purpose. Rice yield is a variable used as an output of this network. The network information that contains five covariates such as rainfall, maximum temperature, minimum temperature, morning relative humidity and evening humidity, excluding a bias unit, which are used as input units in all districts. The standardized method is used for rescaling method for covariates. Automatic architecture selection was applied to select the activation function for both the input and output layer. It is identified that the hyperbolic tangent and identity activation functions are suitable for these layers.

In Pathein District, single hidden layer with 2 neurons in each layer are used in the network. The mainly positive impact on rice yield is due to the changes of maximum temperature and morning relative humidity. Thus, the rice yield changes when the MaxT and MRH rise and other predictor variables fall. In Hinthada District, ANN model is single hidden layer with 2 neurons and the mainly negative impact on rice yield is due to the changes of evening and morning relative humidity. Thus, the rice yield changes when the MRH and ERH fall and other predictor variables rise. In Maubin District, ANN model is single hidden layer with one neurons and the mainly negative impact on rice yield is due to the changes of maximum. Thus, the rice yield changes when the MaxT fall and other predictor variables rise. In Myaungmya District, ANN model is single hidden layer with 2 neurons and the mainly negative impact on rice yield is due to the changes of maximum temperature. Thus, the rice yield changes when the MaxT fall. In Phyarpon District, ANN model is single hidden layer with 7 neurons and the main contribution on rice yield is due to the changes of rainfall. Thus, the rice yield changes when rainfall rise.

According to the results of MLR, SARIMAX, VAR and ANN, it is revealed that the rice yield is mainly dependent on the changes of RF, MaxT and MRH in Pathein District. The MRH and ERH have mainly effects on the changes of rice yield in Hinthada District. The rice yield can vary due to the changes of MInT and RF in Maubin district. The RF, MinT and MaxT have mainly effects on rice yield in Myaungmya District. Moreover, the effect of RF and MRH is on the changes of rice yield in Phyarpon District.

Additionally, the findings showed that the ANN model is chosen as the most appropriate model having a relatively low RMSE, MAPE and relatively high R^2, for all districts. Finally, the rice yield of the periods 2021-2022 and 2022-2023 are forecasted by using ANN model in all those districts. The forecast values for monsoon rice yield

(Bsk/Ac) are between 70.04 and 73.68 in Pathein, between 73.48 and 78.64 in Hinthada district, between 68.84 and 74.89 in Maubin District, between 67.27 and 76.26 in Myaungmya District and between 59.35 and 63.96 in Phyarpon District respectively. In contrast, the forecast values for summer rice yield (Bsk/Ac) are between 86.86 and 92.29 in Pathein District, between 83.26 and 89.29 in Hinthada District, between 90.05 and 93.37 in Maubin District, between 89.44 and 95.92 in Myaungmya District, between 91.64 and 101.74 in Phyarpon District respectively.

Generally, it can be seen that summer rice is naturally more productive than monsoon rice. Meanwhile, it is more cloudless in the summer than in the monsoon season. Summer rice is more solar energy available for the rice plant. The rice plants make food for the plant with the solar energy and it grow and blossom. Thus, summer rice, is cultivated in the summer season, is cloudless and it means getting more solar and producing more rice yield. Monsoon rice is grown by rainfall, while summer rice is grown by dam or river water, lake water cultivated with underground irrigation water. Summer rice has a longer lifespan than monsoon rice. The reason is that summer rice starting the plant growth at a lower temperature.

After analyzing the results, it can be concluded that the Artificial Neural Network (ANN) has well predicted the rice yield based on the climatic variables as proposed. The contribution for this study deals not only with the investigation of the impact of climatic variables on rice yield but also with the implementation of the multi-layer neural network for the forecasting of the rice yield. Generally, it can be seen that maximum temperature and rainfall have negative effect, and minimum temperature and humidity have positive effect on rice yield in all districts. This also implies that, rising in MaxT and RF could have reduced on the rice yield, and rising the MinT and humidity could have increased on rice yield.

The results of the present study are similar to the results of the study of Farook and Kannan (2015). It analyzed the impact of climate change on rice yield using aggregate level time series data. Three climate variables (MaxT, MinT and Rainfall) have significant effects on the rice yield of Kharif (June-December) and Rabi (January-May) crops. MaxT and RF have negative effect on yield, whereas MinT affect yield positively. But, this study only used the Vector Autoregressive method.

Among the climatic factors, temperature and moisture was identified to be the most important factor determining the crop yield uncertainty than rainfall variability. Moreover, the ANN model produced the best fits over MLR, SARIMAX, and VAR, probably indicating that the time series models may be better suited for capturing the long-term and short-term variation in the time series (Raj, Ramesh and Rajkumar, 2019).

5.2 Suggestions

One of the most climate-sensitive agro-ecosystems is the cultivation of rice. The assessment of the impact of climate change on rice yield using the most appropriate statistical model is essential for formulating the vision and goals of future agricultural performance. Specifically, it can provide useful information for long term agricultural development plan for each district. Based on the findings of the effects of current and future climate change on rice yield in Ayeyawady Region, a number of suggestions are made as follows.

Pathein District has a hot and humid climate with distinct monsoon and summer seasons. Therefore, the rice yield decrease by increasing maximum temperature and humidity. At the same time, the rice yield increase when the rainfall and maximum temperature rising. The monsoon rains are essential for agriculture but can sometimes lead to flooding and other related issues. Due to excessive heat, the grain of rice often becomes sterile. Therefore, it is necessary to balance the planting time to coincide with the most suitable temperatures for each growth stage of the rice crop plant.

Summers in Hinthada District are hot and dry, with high temperatures and limited rainfall. This season is associated with sweltering heat and can be challenging for agriculture due to water scarcity. Monsoon season is characterized by warm temperatures, high humidity, and heavy rainfall. Monsoons are vital for agriculture, and provide water for rice cultivation. The rice yield decrease by raising the humidity since Due to excess moisture, the biological processes of plants often have negative effects according to their growth stages. Due to excess moisture, the biological processes of plants often have negative effects according to their growth stages such as the rate of germination decreases, slow growth of young seedlings, discoloration of leaves and crop damage irregular ripening time.

In Maubin District, the monsoon season is oppressive and overcast, the summer season is muggy and partly cloudy, and it is hot year round. The rice yield decrease by rising maximum temperature (daytime) and evening humidity. But, the rice yield increase by the rising of the minimum temperature (nighttime). As it is slightly dry, the rice yield will increase as the rainfall increases. Therefore, these facts are the main factors that can change the growth and yield of the rice plant.

Myaungmya District is located in the deltaic area, there are numerous streams. These stream networks also support irrigated water for summer paddy cultivation. Rice yields will decline as a result of the extraordinarily high temperatures, heavy rainfall, and high humidity levels. For a short ripening time, the process of filling grains may be impacted by the rising maximum temperature. The cause for the decrease in rice yield is this circumstance. New rice crop varieties with a higher per-day yield potential should be chosen and sown on a daily basis in order to reduce the decline in rice production caused by high temperatures during rice crop growth.

In Phyarpon District, the monsoon season is overcast, the summer season is partly cloudy, and it is hot and oppressive. Heavy rains were responsible for the changes in rice yield, and the increase in maximum temperature was responsible for the minor reduction in rice yield. Rice-growing regions are losing agricultural land as a result of saltwater seeping in from the sea due to severe rainfall. As a result, it is best to develop the rice crop during the dry season rather than during the periods of heavy rain.

In addition, the tropical monsoon climate of Ayeyawady plays a crucial role in supporting agriculture, as it provides the necessary water for rice cultivation. However, it also poses challenges due to the risk of flooding, especially in low-lying coastal areas, which can be exacerbated by rising sea levels and increased intensity of rainfall associated with climate change. Proper water management, flood mitigation, and adaptation strategies are essential to cope with the impacts of the changing climate in the region. In order to mitigate the effects of climate change and increase rice yields, the following actions should be taken.

First, Encouraging farmers to adopt sustainable agricultural practices can significantly reduce the environmental impact of rice cultivation. This includes practices such as conservation tillage, crop rotation, agroforestry, and integrated pest management. These techniques help preserve soil health, reduce water usage, and minimize greenhouse gas emissions.

Second, agricultural researchers and farmers can collaborate to develop and implement climate-resilient rice crop varieties that will enhance food security and sustainability in the face of a changing climate such as drought, floods, and heat stress. Climate-resilient varieties can ensure stable yields even under adverse conditions. Planting drought-resistant crop varieties can help crops survive and produce yields even in limited water conditions.

Third, creating small ponds or reservoirs in low-lying areas allows for the collection of rainwater during periods of rainfall. This water can then be used for irrigation during dry spells, helping to supplement water needs during droughts.

Fourth, after harvest, instead of burning or disposing of crop residues, they can be collected and composted. Crop residues include stalks, leaves, husks, and other parts of plants that remain after harvesting. Composting is a natural process where organic materials decompose, creating a nutrient-rich soil amendment. While implementing soil protection techniques like composting crop residues is essential for enriching soil fertility, the judicious use of nitrogen and potassium fertilizers can complement these practices to ensure optimal plant growth and maximize crop yields.

Fifth, timely and accurate weather forecasts are essential for farmers to make informed decisions regarding planting, irrigation, pest control, and harvesting. Improvements in weather prediction technologies, such as the use of satellite data, advanced modeling techniques, and real-time monitoring, can provide more precise and location-specific weather information.

Moreover, the method of rice cultivation in a particular region depends largely on the situation of land, type of soils, irrigation facilities, availability of labourers' intensity, and distribution of rainfalls (Farmers' Portal, 2021). There are various technologies strategically used in today's agriculture sector such as the rice productivity technology (Saud & Wang, 2022). The government should provide better access to credit for smallholder farmers because Myanmar's smallholders using high input-use farming technologies will benefit from climate change-induced paddy yield changes towards the end of the century (Jensen, Keogh-Brown & Tarp, 2021).

The government needs to encourage crop diversification, which improves the farm income of smallholder farmers. In order to properly use agricultural inputs such as fertilizers and pesticides, relevant organizations need to provide knowledge and awareness. Policies expected to increase crop production should focus on climate change adaptation in Myanmar (Htoo, 2021).

Overall, this study suggested that the changes in rice yield generally depend on rainfall, temperature, and soil moisture. However, it can also be influenced by other factors such as farm size, type of fertilizer, pesticide, labor, wind speed, and geographical location. Accordingly, the government should lead in building modern farmland, providing paddy production facilities, breeding high-yield and stress-resistant rice varieties, and supporting modern rice farming techniques.

5.3 Needs for Further Studies

Further research requires a deeper understanding of the statistical methods applied to analyze the impact of climate change. The rice yield can be influenced by various socio-economic factors such as market dynamics, access to markets, rural infrastructure, nutrition policies, and socioeconomic disparities. Moreover, farmers' adaptation to climate change, type of soil, and rice cultivation are needed to be considered in studying the rice yield of future research. It is also necessary to obtain the data required for the study from different potential sources for complete and more comprehensive results.

The researchers should engage farmers, local communities, agriculture experts, and policymakers in the research process whose perspectives and knowledge could provide valuable insights and ensure that the recommendations align with the ground realities and address the actual needs of the affected communities. Besides, the inclusion of indigenous knowledge and practices should be considered, which have demonstrated resilience to environmental changes over generations.

Future studies should be carried out to analyze how rice production is being impacted by other factors, apart from climatic factors. These may include farm size, types of fertilizer used, labor, wind speed, prices of rice, irrigation, rice production technology, and geographical location.

Additionally, the models explored in this research could be applied to the comparison of climatic effects on different crop productions or different States and Regions. These different regions may have unique ago-ecological characteristics, socio-economic contexts, and government structures that can significantly influence the effectiveness and applicability of the proposed interventions.

REFERENCES

Abdullahi, J., & Elkiran, G. (2017). Prediction of the Future of Climate Change on Reference Evapotranspiration in Cyprus using Artificial Neural Network. *Scientific Committee of the 9th international Conference on Theory and Application of Soft Compurting, Computing with Words and Perception* (pp. 276-283). Budapest, Hungary: Elsevier B V.

Aboukarima, A. M., Elsoury, H. A., & Menyawi, M. (2015). Artificial Neural Network Model for the Prediction of the Cotton Crop Leaf Area. *Plant & Soil Science*, Vol.4, 1-13.

Adedeji, I. A., Tiku, N. E., Waziri-Ugwa, P. R., & Sanusi, S. O. (2017). The Effect of Climate Change on Rice Production in Adamawa State, Nigeria. *Agroeconomia Ctoatica 7, (1)*, 1-13.

Agriculture Guide. (2020). *European Chamber of Commerce in Myanmar.* Yangon.

Agriculture Overview. (2018). *Burma Country Commercial Guide.* https//www.export.gov.

Alam, M. M., Siwar, C., Talib, B., & Toriman, M. E. (2014). Impact of Climate Changes on Paddy Production in Malaysis: Micero Study on IADA at North West Selangor. *Research journal of Environmental and Earth Sciences*, Vol.6(5), pp.251-258.

Artley, B. (2022). *Time Series Forecasting with ARIMA , SARIMA and SARIMAX.* Retrieved from https://towardsdatascience.com.

Aryal, M., Regmi, P. P., Thapa, R. B., Pande, K. R., & Pant, K. P. (2016). Impact of Climate Variables To Major Food Crops Yield in Midhills of Western Development Region, Nepal. *Journal of Agricure and Environment*, Vol.17, 65-72.

Aung, L. L., Zin, E. E., & Theingi, P. (2017). *Myanmar Climate Report.* MET Report.

Aung, S. M. (2018). *Global Agriculture Information Network.* Gain and Feed Annual Report.

Awal, M., & Siddique, M. (2011). Rice Production In Bangladesh Employing By Arima Model. *Bangladesh Journal of Agricultural Research, 36(1), 51–62.*, 36(1), 51-62.

Ayeyawady Region. (2022). *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Ayeyarwady_Region

Ayeyawady Region Census Report. (2015). Ministry of Immigration and Population, Department of Population, Nay Pyi Taw.

Ayinde, O. E., Ojehomon, V., Daramola, F. S., & Falaki, A. A. (2013). Evaluation of the Effects of Climate Change on Rice Production in Niger State, Nigeria. *Ethiopian Jounal of Environmental Studies and Management, Vol.6*, 763-773.

Bangal, B. C. (2009). Artificial Neural Networks. In *"Automatic generation Contril of Interconnected Power Systems Using Artificial Neural Network Techniques"* (pp. 74-100).

Chowdhury, I. U., & Khan, M. A. (2015). The Impact of Climate Change on Rice Yield in Bangladesh: A Time Series Analysis. *Faculty of Business Administration, BGC Trust University, Bangladesh*, 1-28.

Climate Smart Agriculture. (2019). *Sustainable Cropland and Forest Management in Priority Agro-ecosystems of Myanmar Project (GCP/MYA/017/GFF).* Nay Pyi Taw: Food and Agriculture Organization of the United Nations (FAO) and AVSI Foundation.

Colquhoun, A., Coransson, H., Sandberg, & Nyoi, M. Y. (2016). *How the people of Myanmar Live with Climate Change and What Communication can do.* Myanmar Repoert;Climate Asia.

Drummond, S. T., Sudduth, K. A., Joshi, A., Birrell, S. J., & Kitchen, N. R. (2003). Statistical and Neural Methods for Site-Specific Yield Prediction. *American Society of Agricultural Engineers (ASAE)*, Vol.46(1), 1-10.

Enovejas, A. M., Maldia, S., Komarudin, N. A., & Hilmi, Y. S. (2021). Effect of Cliimate Variables in Rice Yield in Nueva Ecijia, Philippines. *Asia Pacific Journal of Sustainable Agriculture, Food and Energy*, Vol.9(1), 29-44.

Faradiba. (2020). Analysis of Climate Factors on Paddy Production in West Java. *Advances in Social Science, education and Humanities Research, Volume 560*.

Farmers' Portal. (2021). *Ministry of Agriculture and Farmers Welfare.* Government of India.

Farook, A. j., & Kannan, K. S. (2015). Climate Change Impact on Rice Yield in India-Vector Autoregression Approach. *Applied Statistics*, Vol.3, 160-178.

Food and Agriculture Organigation, FAO. (2008). *Climate Change anf food Security: A Framework Document.* Food and Agriculture Organigation of the United Nations.

Food and Agriculture Organization, FAO. (2020). *The State of Food and Agriculture 2020. Overcoming water challenges in agriculture.* Retrieved from https://doi.org/10.4060/cb1447en

Gershenson, C. (2003). Artificial Neural Networks for Beginners. Retrieved from https://www.researchgate.net/publication/1956697

Ghodsi, R., Yani, R. M., Jalali, R., & Ruzbahman, M. (2012). Predicting Wheat Production in Iran Using an Arrificial Neural Networks Approach. *Academic Research in Business and Social Sciences*, Vol.2, 34-46.

Greeshma.U, A. (2015). ARTIFICIAL NEURAL NETWORK. *International Journal of Scientific & Engineering Research*, Vol.6, 110-115.

Gujarati, D. N., & Porter, D. C. (2009). *Bsic Econometrics* (Fifth ed.). New York: McGraw-Hill.

Hamjah, M. A. (2014). Climatic Effects on Major Pulse Crops Production in Bangladesh:An Application of Box-Jenkins ARIMAX Model. *Economics and Sustainable Development*, Vol.5, No.15.

Haque, M. E., Hossain, M. I., & Rahman, K. M. (2004). Searching for the Best Fitting Deterministic Model for Innovative Growth Analysis and Forecasting of Rice Production in Bangladesh. *Bangladesh.J.Agric.Econs*, 15-35.

Hilal, Y. Y., Ishak, W. W., Yahya, A., & Asha'ari, Z. H. (2016). AN Artificial Neural Network with Stepwise Method for Modelling and Simulation of Oil Palm Productivity Based on Various Parameters in Sarawak. *Applied Sciences, Engineering and Technology,* 13(9): 730-740.

Horton, R., Mel, M. D., Peters, D., Lesk, C., Bartlett, R., Helsingen, H., . Rosenzweig, C. (2017). *Assessing Climate Risk in Myanmar.* New York, NY, USA: Center or Climate Systems Research at Columbia University: WWF-US and WWF-Myanmar.

Hossain, M. J., Al-Amin, A. K., & Islam, A. H. (2018). Modeling and Forecasting of Climatic Parameters: Univariate SARIMA versus Multivariate Vector Autoregression Approach . *Bangladesh Agricultural University,* Vol.1, 131-143.

Htoo, T. (2021). Macro Analysis of Climate Change and Agricultural Production in Myanmar. *The Nature, Causes, Effects and Mitigation of Climate Change on the Environment*. doi:DOI: 10.5772/intechopen.98970

International Rice Research Institute (IRRI). (2020). *World Meteorological Special Report Organization.* Nay Pyi Taw.

International Rice Research Institute, IRRI. (2008). *Myanmar After Cyclone Nargis.* Science Daily.

Jeansen, H. T., Keogh-Brown, M., & Tarp, F. (2021). Climate Change and Agricultural Productivity in Myanmar. Application of a New Computable General Equilibrium (CGE) Model. *Wider.unu.edu*.

Joshi, N. P., Maharjan, K. L., & Piya, L. (2011). Effect of Climate Variables on Yield of Major Food-crops in Nepal. *Contemporary India Studies: Space and Society, Hiroshima University*, Vol.1, 19-26.

Karkalos, N. E., Efkolidis, N., Kyratsis, P., & Markopoulos, A. P. (2019). A Comparative Study between Regression and Neural Networks for Modeling Al6082-T6 Alloy Drilling. *Machines*, 1-18.

Karsoliya, S. (2012). Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture. *International Journal of Engineering Trends and Technology*, Vol.3(6), 714-717.

Khairunniza-Bejo, S., Muataffha, S., & Wan Ismail, W. I. (2014). Application of Artificial Neural Network in Predicting Crop: A Review. *Journal of Food Science and Engineering*, Vol-4, 1-9.

Kyi, T. (2016). Influence of Climate Change Impact on Agricultural Risks in Myanmar's Dry Zone. *Agricultural risk management*, 1-9.

Lamba, V., & Dhaka, V. (2014). Wheat Yield Using Artificial Neural Network and Crop Prediction Techniques. *Research in Applied Science and Engineering Technology (IJRASET)*, Vol.2, 330-341.

Langhorn, C. M. (2015). *Simulation of Climate Change Impacts on Selected Crop Yields in Southern Alberta.* Lethbridge, Canada.

Lar, N. M., Arunrat, N., Tint, S., & Pumijumnong, N. (2018). Assessment of the Potential Climate Change on Rice Yield in Lower Ayeyarwady Delta of Myanmar Using EPIC Model. *Environment and Natural Resources Journal*, 16(2):45-57.

Laxmi, R. R., & Kumar, A. (2011). Weather Based Forecasting Model for Crops Yield using Neural Network Approach. *Statistics and Applications*, Vol.9, 55-69.

Long, T. B., Block, V., & Coninx, I. (2016). Barriiers to the Adoption and Diffusion of Technological Innovations for Climate-Smart Agriculture in Europe: Evidence from the Netherlands, France, Switzerland and Italy. *Journal of Cleaner production, Vol.112(1),* 9-2.

Mall, R. K., Singh, R., Gupta, K., & Srinivasan, G. (2006). Impact of Climate Change on Indian Agriculture: A review. *Climate Change*, 445-478.

Maponya, P., & Mpandeli, S. (2012). Climate Change and Agricultural Production in South Africa: Impacts and Adaptation options. *Agricultural Science*, Vol.4, 48-60.

Mathieu, J. A., & Aires, F. (2016, November). Statistical Weather-Impact Models: An Application of Neural Networks and Mixed Effects for Corn Production over the United States. *Applied meteorology and Climatology*, 2509-2527.

Mathieu, J. A., & Aires, F. (2018). Using Neural Network Classifier Approach for Statistically Forecasting Extreme Corn Yield Losses in eastern United States. *Earth and Space Science*, Vol.5, 622-639.

Matsmura, K., gaitan, C. F., Sugimoto, K., Cannon, A. J., & Hsieh, W. W. (2014). Maize Yield Forecasting by Linear Regression and Artificial Neural Networks in Jilin, China. *Ageicultural Science*, 399-410.

Matthews, R. B., Kropff, M. J., Bachelet, D., & Vanlaar, H. H. (1995). *Modelimg the Impact of Climate Change on Rice Production in Asia.* Manila: International Rice Reserch Institute.

Ministry of Foreign Affairs. (2015). *Agriculture in Myanmar.* Yangon: Embassy of the Kingdom of the Netherlands.

Molla, T., Tesfaye, K., Mekibib, F., Tana, T., & Taddesse, T. (2020). Rainfall Variability and Its Impact on Rice Productivity In Fogera Plain, Northwest Ethlopla. *Ethiop.j.Agric.Sic., 30(2),* 67-79.

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to Time Series Analysis and Forecasting* (Second ed.). Cananda: John Wiley & Sons, Inc.

Myanmar Agriculture at a Galance. (2018). *Ministry of Agriculture, Livestock and Irrigation.* Nay Pyi Taw: Department of Planning.

Myanmar Climate Change Alliance, MCCA. (2017). *Impact of Climate Change and the Case of Myanmar.*

Myanmar Climate-Smart Agriculture (MCSA). (2019). *Sustainable Cropland and Forest Management in Priority Agro-ecosystems of Myanmar Project.* Nay Pyi Taw: the Food and Agriculture Organization of the United Nations .

Myanmar in Brief. (2018). *Ministry of Planning and Finance.* Nay Pyi Taw: Central Statistical Organization.

Myanmar National Study. (2015). *Promotion of Climate Resilience in Rice and Maize.* Deutsche Gesellschaft für,Internationale Zusammenarbeit (GIZ) GmbH.

Myanmar's National Adaptation Programme of Action (NAPA). (2012). *Myanmar's National Adaptation Programme of Action (NAPA) to Climate Change.* Department of Meteorology and Hydrology, Ministry of Transport of Union of the Republic of Myanmar.

Myint, K. M. (2014). How Will Climate Change Impact Myanmar? *Global New Light of Myanmar*.

Myint, T. (2018). Myanmar's Rice Industry and Policies toward Value Addition and Export. *FFTC Agricultural Policy Articles*.

Naing, M. M. (2005). Paddy Field Irrigation Systems in Myanmar. *Ministry of Agriculture and Irrigation Myanmar*.

Naing, T. A., Kingsbury, A. J., Buerkert, A., & Finckh, M. R. (2008). A Survey of Myanmar Rice Production and Constraints. *Agriculture and Rural Development in the Tropics and Subtopics*, Vol.109, No.2, 151-168.

Nanda, S. K., Tripathy, D. P., Nayak, S. K., & Mohapatra, S. (2013). Prediction of Rainfall in India uaing Artificial Neural Network (ANN) Models. *I.I.Intelligent Syatems and Applications, 1*, 1-22.

Oo, A. T. (2018). Community perceptions of the Impacts of Climate Change on Agriculture in Myanmar's Central Dry Zone. *Food Security Policy Project Research Hilights Myanmar*, 1-7.

Oo, A. T. (2020). Measuring the Economic Impact of Climate Change on Crop Production in the Dry Zone of Myanmar: A Ricardian Approach Climate: MDPIAG.

Panchal, F. S., & Panchal, M. (2014). Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural Network. *Journal of Computer and Mobile Computing*, Vol.3, pg 455-464.

Pandey, K., Maurya, D., Gupta, G., & Mishra, S. (2016). Yield Forecasting Models Based on Weather Parameters for Eastern U.P. *Plant Research*.

Parekh, F. P., & Suryanarayana, T. M. (2012). Impact of Climatological Parameters on Yield of Wheat Using Neural Network Fitting. *Modern Engineering Research*, Vol.2, 3534-3537.

Park, J.-K., Das, A., & Park, J.-H. (2018). Integrated Model for Predicting Rice Yield with Climate Change. *International Agrophysics*, Vol.32, 203-215.

Pheakdy, D. V., Xuan, T. D., & Khanh, T. D. (2017). Influence of Climate Factors on Rice Yields in Cambodia. *AIMS Gosciences, 3(4)*, 561-575.

Rahman, M. A., Kang, S., Nagahhatia, N., & Macnee, R. (2017). Impacts of temperature and rainfall variation on rice productivity in major ecosystems of Bangladesh. *Agriculture & Food Security*. doi::10.1186/s40066-017-0089-5

Raj, E. E., Ramesh, K., & Rajkumar, R. (2019). Modelling the Impact of Climate on Tea Yield Variability in South India: 1981-2015. *cogent food & agriculture*.

Ranjeet, T. R., & Armstrong, L. J. (2014). An Artificial Neural Network for Predicting Crop Yields in Nepal. *Asian Fedration for Information Technology in Agriculture*, (pp. 376-386).

Ray, M. (2016). Influence of Different Weather Parameters on Rice Production. A Review. *Advances in life Sciences*, 5(16), 5776-5782.

Santosh. (2020). *Assumptions which makes Artificial Neural Network.* Retrieved from https://medium.com/analytics-vidhya.

Saud, S., & Depeng Wang, S. F. (2022). Comprehensive Impacts of Climate Change on Rice production and Adaptive Strategies in China. *Frontiers in Microbiology, Vol.13*.

Sengar, R. S., & Sengar, K. (2015). *Climate Change Effect on Crop Productivity.* New York: Taylor & Francis Group.

Sharma, M. A. (2012). *Comparative Study of Forecasting Models Based on Weather Parameters.* India: A deemed-to-be university.

Shrestha, S., Thin, N. M., & Deb, P. (2014). Assessment of Climate Change Impacts on Irrigation Water Requirement and Rice Yield for Ngamoeyeik Irrigation Project in Myanmar. *Journal of Water and Climate Change*, 427-442.

Sitienei, B. J., Juma, S. G., & Opere, E. (2017). On the Use of Regression Models to Predict Ta Crop Yield Responses to Climate Change: A Case of Nandi East, Sub-County of Nandi County, Kenya. *Climate*, 1-14.

Slagle, J. T. (2014). *Climate Change in Myanmar: Impacts and Adaptation.* California USA: Naval Postgraduate School.

Somah, T. P. (2013). *Climatic Change Impacts on Subsistence Agriculture in the Sudano_Sahel of Cameroon.* Cameroon.

Song, Y., & Chunyi Wang, H. W. (2022). The negative impact of increasing temperatures on rice yields in southern China. *Science of the total Environment* .

Stastny, J., Konecny, V., & Trenz, O. (2011). Agricultural Data Prediction by Means of Neural Network. *Agricultural Economics*, Vol.7, 356-361.

Syeda, J. A. (2017). Impact of Climate Change on Wheat Production in Dinajpur Region of Bangladesh: an Econometric Analysis. *Environ. Sci. & Natural Resources*, Vol.10(2), 157-162.

Than, M. (1990). *Agriculture in Myanmar.* Yangon.

Thanda Kyi. (2016). Overview of Agriculture Policy in Myanmar. *FFTC Agricultural Policy Articles*.

Tsay, R. S. (2014). *Multivariate Time Series Analysis.* Cananda: John Wiley & Sons, Inc.

UN Environment. (2019). *New Sustainable rice project to guard Myanmar's rice sector against climate change, Myanmar.* Reliefweb.

UN-Habitat. (2020). *United Nations Human Settlement Programme. Myanmar Climate Change Alliance to start 2nd Phase of the Programme.* https://unhabitat.org.mm.

United States Agency International Development (USAID). (2023). *Rice Area and Production Estimates for The 2022 Monsoon Season.* Myanmar Agricultural Crop Yield Estimation Project.

United States Department of Agriculture, USDA. (2008). *Commodity Intelligence Report.* Foreign Agricultural Service.

Wei, W. W. (2006). *Time Series Analysis, Univariate and Multivariate Methods, Second Edition.* United States: Greg Tobin.

Wenjiao, S., Fulu, T., & Zhao, Z. (2013). A Review on Statistical Models for Identifying Climate Contributions to Crop Yields. *Journal of Geographical Sciences*, 567-575.

Wiah, E. N., & Twumasi-Ankrah, S. (2017). Impact of Climate Change on Cocoa Yield in Ghana Using Vector . *Ghana journal of Technology*, Vol. 1, No.2, pp. 32-39.

Win, M. S. (2013). Analysis of Socioeconomic Factors Affecting the Rice Production : A Case Study in Theegone Village, Shwe Bo Township. *Mandalay University Research Journal*, Vol.4, 32-43.

Win, U. K. (1991). *A Century of Rice Improvement in Burma.* International Rice Research Institute.

World Bank. (2016). *Ayeyarwady Paddy farmers Set Up Productivity with Smart Farming Methods.* Retrieved from The World Bank: https://www.worldbank.org/en/news/feature/2016/08/02

Zaw, K., Lwin, N. N., Nyein, K. T., & Thandar, M. (2011). Agricultural Transformation, Institutional Changes, and Rural Development in Ayeyerwady Delta,Myanmar. Myanmar, ERIA.

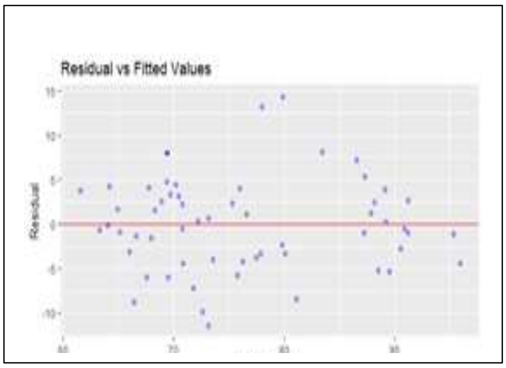
Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: Thee state of the art. *International Journal of Forecasting*, 35-62.

APPENDIX(A)

Multiple Linear Regression

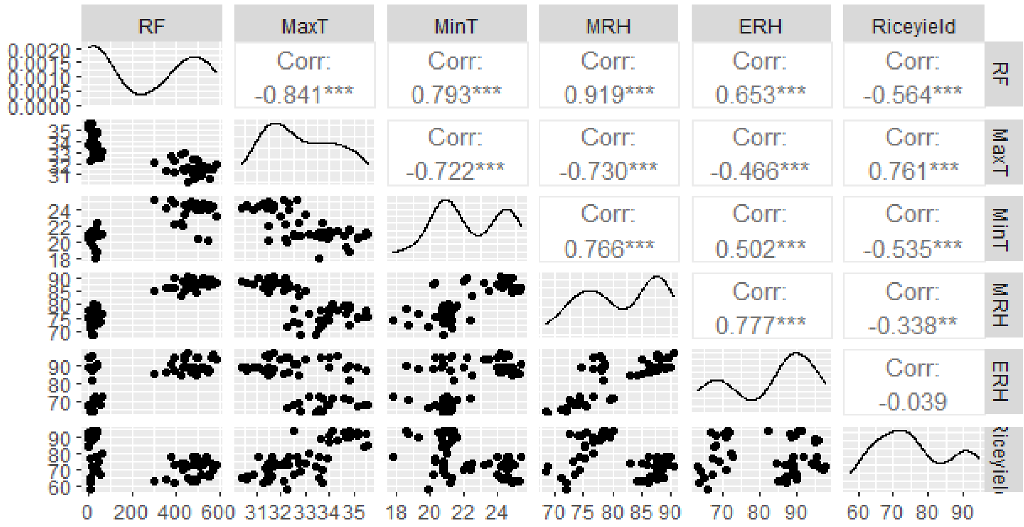
A.I Pathein

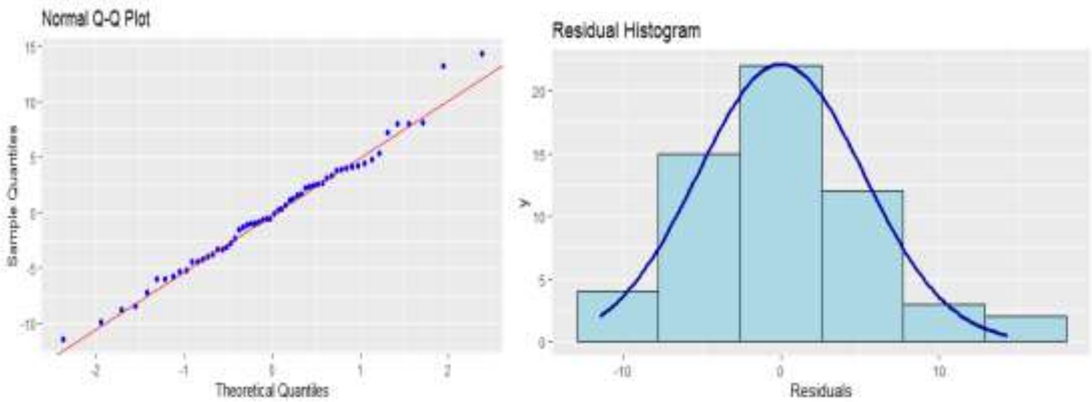
```
 Call:
lm(formula = semi_ptrice ~ semi_ptrf + semi_ptmaxt + semi_ptmint
+
    semi_ptmrh + semi_pterh, data = pathein58)

Residuals:
     Min       1Q   Median       3Q      Max
-11.4425  -3.6560  -0.3395   3.2724  14.3179

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -198.64910   38.85984  -5.112 4.65e-06 ***
semi_ptrf     -0.02690    0.01024  -2.626  0.01132 *
semi_ptmaxt    5.72106    0.96268   5.943 2.37e-07 ***
semi_ptmint   -0.98269    0.61907  -1.587  0.11849
semi_ptmrh     1.19267    0.35860   3.326  0.00162 **
semi_pterh     0.22284    0.10913   2.042  0.04624 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.477 on 52 degrees of freedom
Multiple R-squared:  0.7635, Adjusted R-squared:  0.7408
F-statistic: 33.58 on 5 and 52 DF,  p-value: 3.801e-15


 Breusch Pagan Test for Heteroskedasticity
 -----------------------------------------
 Ho: the variance is constant
 Ha: the variance is not constant
                 Data
 ---------------------------------------
 Response : semi_ptrice
 Variables: fitted values of semi_ptrice

         Test Summary
 ------------------------------
 DF            =    1
 Chi2          =    0.004293572
 Prob > Chi2   =    0.9477557
```

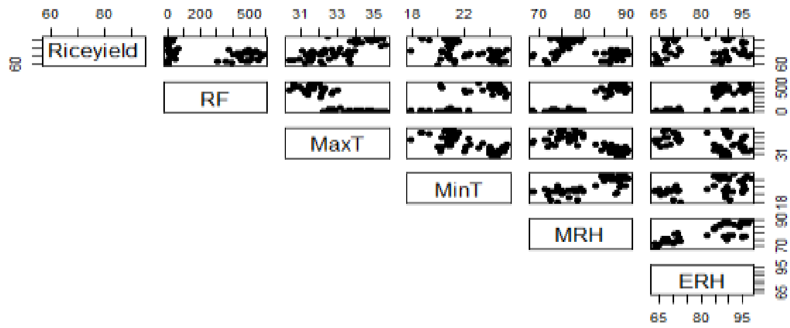



```
      Shapiro-Wilk normality test

data:  pt_MLR$residuals
W = 0.98594, p-value = 0.7377


      Rainbow test

data:  pt_MLR
Rain = 2.3823, df1 = 29, df2 = 23, p-value = 0.01803

 lag Autocorrelation D-W Statistic p-value
   1       0.2696608      1.349313   0.002
```

```
Alternative hypothesis: rho != 0
    Variables  Tolerance       VIF
1    semi_ptrf 0.08995571 11.116582
2 semi_ptmaxt 0.26828288  3.727409
3 semi_ptmint 0.33972084  2.943593
4  semi_ptmrh 0.09381153 10.659671
5  semi_pterh 0.36005116  2.777383
```







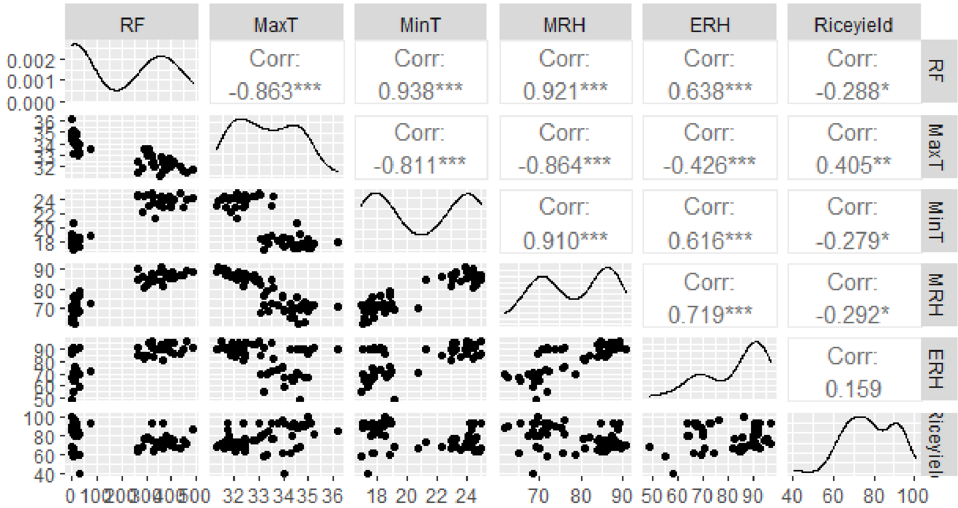
A.II Hinthada

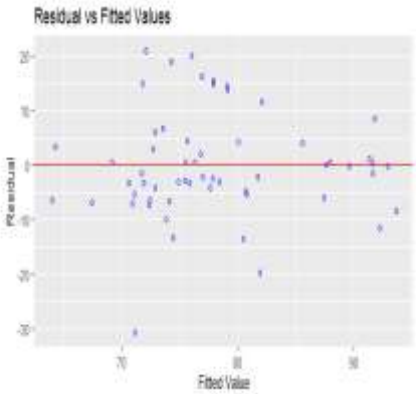
```
Residuals:
    Min      1Q  Median      3Q     Max
-30.754  -5.291  -1.577   4.254  20.979

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 100.066598 114.033762   0.878 0.384242
RF           -0.007221   0.026980  -0.268 0.790048
MaxT          0.054125   2.852096   0.019 0.984932
MinT          0.579188   1.373211   0.422 0.674927
MRH          -1.252072   0.601832  -2.080 0.042430 *
ERH           0.768858   0.197203   3.899 0.000278 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.22 on 52 degrees of freedom
Multiple R-squared:  0.3672, Adjusted R-squared:  0.3064
F-statistic: 6.036 on 5 and 52 DF,  p-value: 0.0001777


Breusch Pagan Test for Heteroskedasticity
 -----------------------------------------
 Ho: the variance is constant
 Ha: the variance is not constant

               Data
 -------------------------------------
 Response : Riceyield
 Variables: fitted values of Riceyield

        Test Summary
 ----------------------------
 DF            =    1
 Chi2          =    1.825428
 Prob > Chi2   =    0.1766686
```

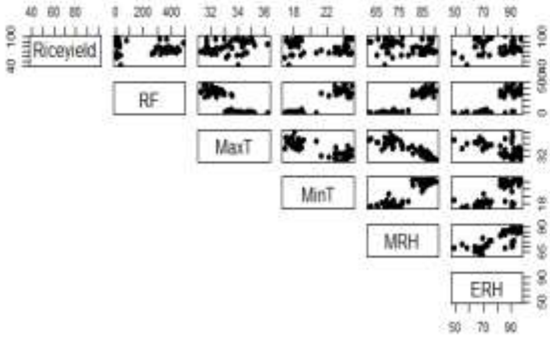
```
      Shapiro-Wilk normality test

data:  htd_MLR$residuals
W = 0.94952, p-value = 0.0173
```

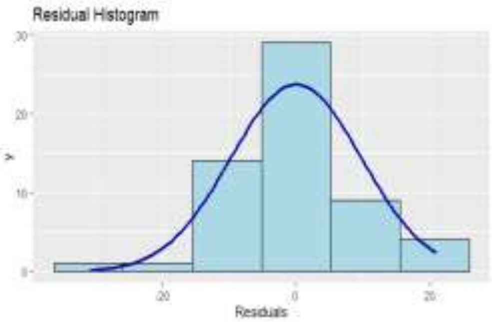
```
  Variables  Tolerance       VIF
1        RF 0.07774714 12.862210
2      MaxT 0.14680554  6.811732
3      MinT 0.09951164 10.049076
4       MRH 0.06910358 14.471031
5       ERH 0.31624952  3.162060
```

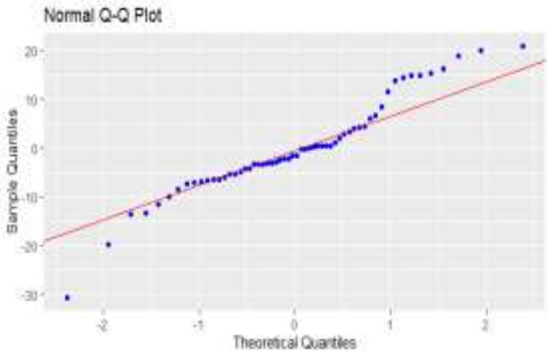
```
      Rainbow test

data:  htd_MLR
Rain = 2.4714, df1 = 29, df2 = 23, p-value = 0.01458

 lag Autocorrelation D-W Statistic p-value
   1        0.595641     0.7580186       0
 Alternative hypothesis: rho != 0
```











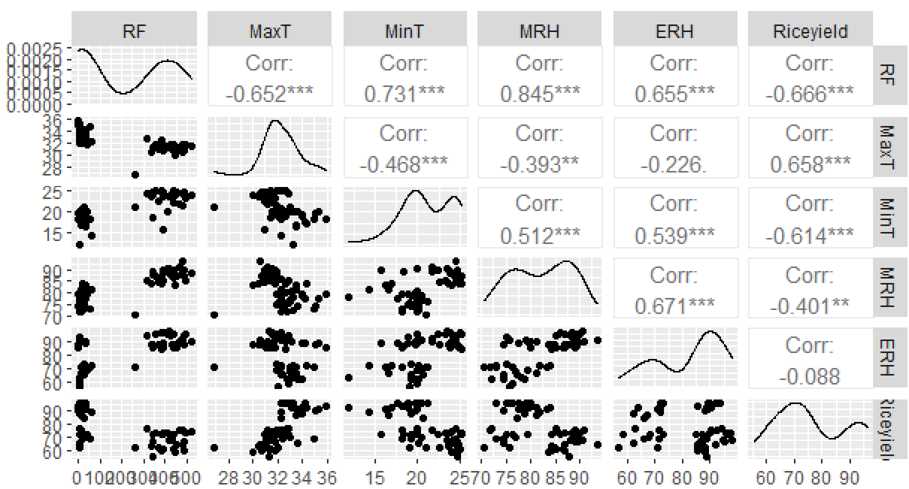
A.III Maubin

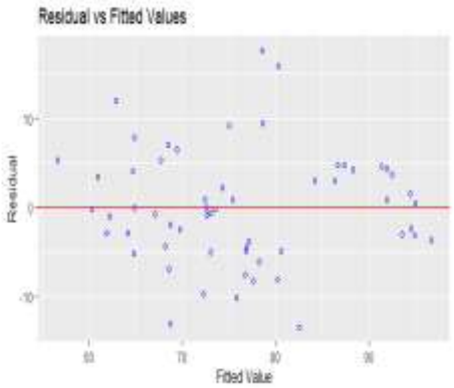
```
Call:
lm(formula = Riceyield ~ RF + MaxT + MinT + MRH + ERH, data =
mub)

Residuals:
     Min       1Q   Median       3Q      Max
-13.4981  -4.2456  -0.1788   4.1488  17.5988

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.66433   31.85339  -0.021  0.98344
RF          -0.04935    0.01362  -3.624  0.00066 ***
MaxT         1.36888    0.90172   1.518  0.13505
MinT        -1.27765    0.44996  -2.839  0.00643 **
```

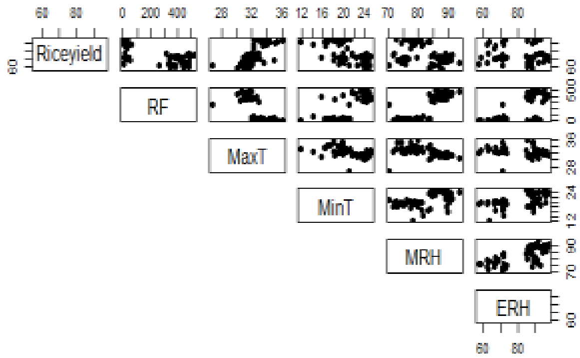
```
MRH           0.26816    0.32022   0.837  0.40619
ERH           0.58351    0.10821   5.392 1.72e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.654 on 52 degrees of freedom
Multiple R-squared:  0.7374, Adjusted R-squared:  0.7121
F-statistic:  29.2 on 5 and 52 DF,  p-value: 5.527e-14


 Breusch Pagan Test for Heteroskedasticity
 -----------------------------------------
 Ho: the variance is constant
 Ha: the variance is not constant

               Data
 -------------------------------------
 Response : Riceyield
 Variables: fitted values of Riceyield

       Test Summary
 ----------------------------
 DF            =    1
 Chi2          =    0.1193088
 Prob > Chi2   =    0.7297853
```

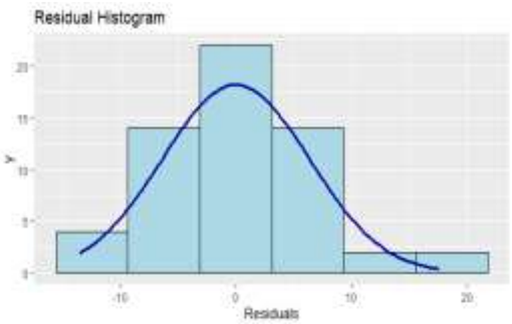
	Variables	Tolerance	VIF
1	RF	0.1018729	9.816151
2	MaxT	0.4482669	2.230814
3	MinT	0.4025437	2.484202
4	MRH	0.2053711	4.869234
5	ERH	0.4676739	2.138242

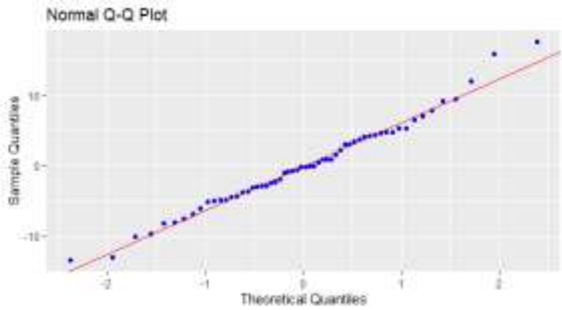
```
      Shapiro-Wilk normality test

data:  mub_MLR$residuals
W = 0.9845, p-value = 0.6655


      Rainbow test

data:  mub_MLR
Rain = 4.4279, df1 = 29, df2 = 23, p-value = 0.0002531

 lag Autocorrelation D-W Statistic p-value
   1        0.390686      1.080526       0
 Alternative hypothesis: rho != 0
```











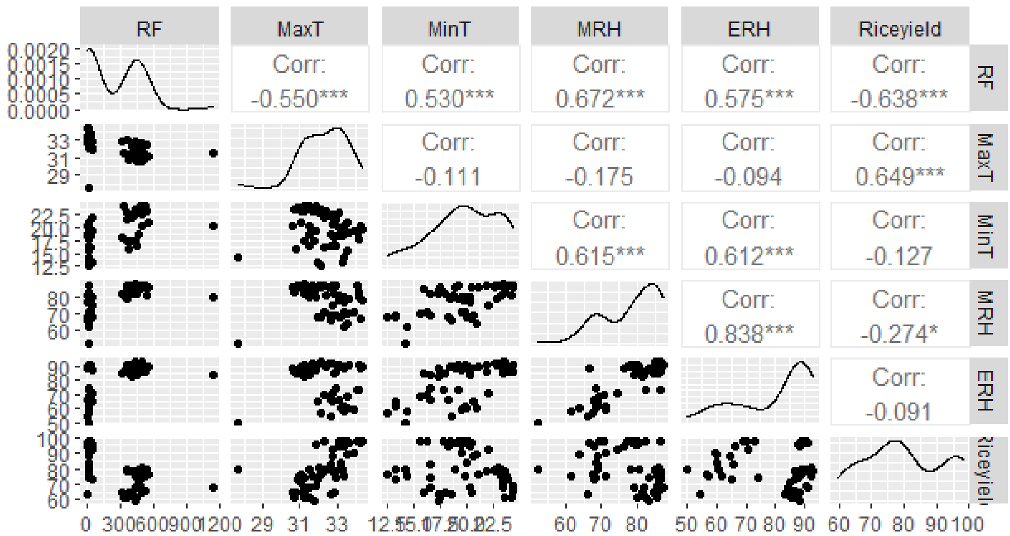
A.IV Myaungmya

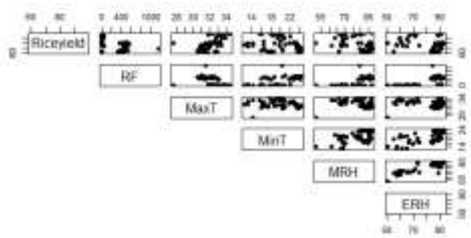
```
Call:
lm(formula = Riceyield ~ RF + MaxT + MinT + MRH + ERH, data = mm)

Residuals:
    Min      1Q  Median      3Q     Max 
-18.163  -4.465   1.029   4.669  15.410 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -34.57395   35.19481  -0.982 0.330472    
RF           -0.02930    0.00733  -3.998 0.000203 ***
MaxT          3.29643    1.11292   2.962 0.004599 ** 
MinT          0.47258    0.45254   1.044 0.301182    
MRH          -0.32145    0.26675  -1.205 0.233635    
ERH           0.38570    0.16205   2.380 0.021007 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.008 on 52 degrees of freedom
Multiple R-squared:  0.6067, Adjusted R-squared:  0.5689 
F-statistic: 16.05 on 5 and 52 DF,  p-value: 1.518e-09


 Breusch Pagan Test for Heteroskedasticity
 -----------------------------------------
 Ho: the variance is constant            
 Ha: the variance is not constant        

               Data
```

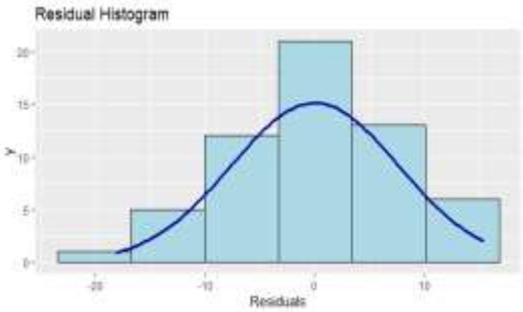
```
-------------------------------------
Response : Riceyield
Variables: fitted values of Riceyield

       Test Summary
----------------------------
DF            =    1
Chi2          =    2.153619
Prob > Chi2   =    0.1422343
```

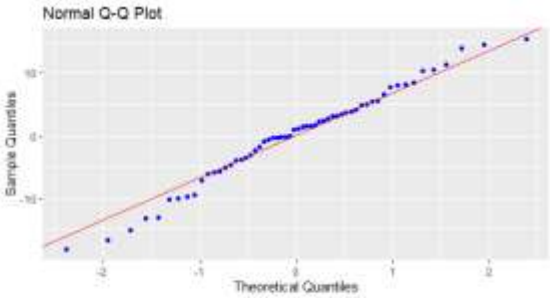
```
Variables Tolerance      VIF
1          RF 0.3334296 2.999134
2        MaxT 0.6073839 1.646405
3        MinT 0.5596875 1.786711
4         MRH 0.2365485 4.227464
5         ERH 0.2781859 3.594719
```

```
      Shapiro-Wilk normality test

data:  mm_MLR$residuals
W = 0.98257, p-value = 0.5693

      Rainbow test

data:  mm_MLR
Rain = 3.7227, df1 = 29, df2 = 23, p-value = 0.0009553

 lag Autocorrelation D-W Statistic p-value
   1       0.5088255     0.9152907       0
 Alternative hypothesis: rho != 0
```









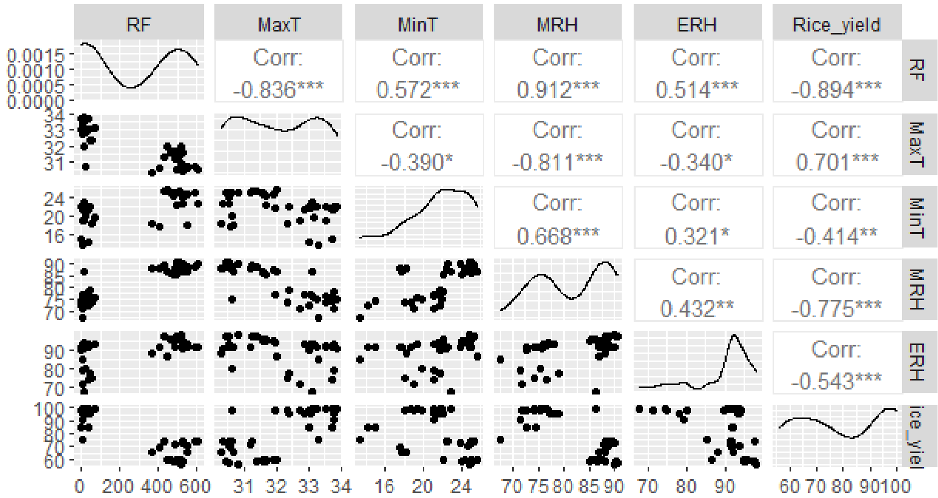
A.V Phyarpon

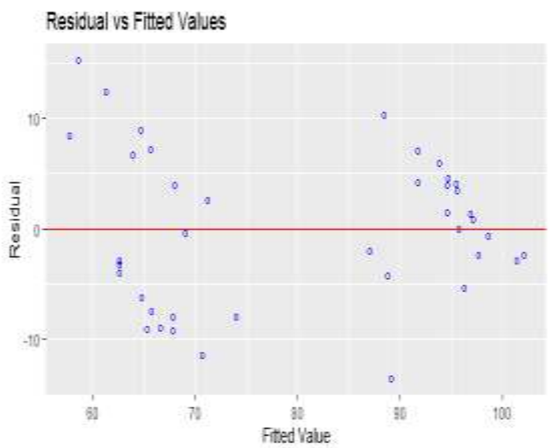
```
Call:
lm(formula = Rice_yield ~ RF + MaxT + MinT + MRH + ERH, data =
pp)

Residuals:
    Min      1Q  Median      3Q     Max
-13.604  -4.541  -0.207   4.253  15.274

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 168.27132   78.78019   2.136    0.040 *
RF           -0.07495    0.01360  -5.511 3.73e-06 ***
MaxT         -2.29748    1.91736  -1.198    0.239
MinT          0.86941    0.52964   1.642    0.110
MRH           0.06663    0.48749   0.137    0.892
ERH          -0.21550    0.18469  -1.167    0.251
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.32 on 34 degrees of freedom
Multiple R-squared:  0.8328, Adjusted R-squared:  0.8082
F-statistic: 33.88 on 5 and 34 DF,  p-value: 2.819e-12


 Breusch Pagan Test for Heteroskedasticity
 -----------------------------------------
 Ho: the variance is constant
 Ha: the variance is not constant

                Data
 --------------------------------------
 Response : Rice_yield
 Variables: fitted values of Rice_yield

       Test Summary
 -----------------------------
 DF            =    1
 Chi2          =    5.351534
 Prob > Chi2   =    0.02070404

     Shapiro-Wilk normality test

data:  pp_MLR$residuals
W = 0.98799, p-value = 0.9416
```

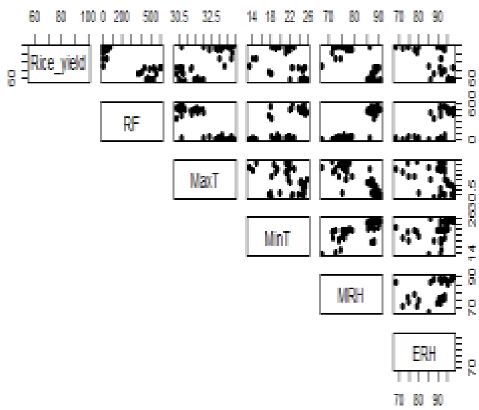
```
Variables Tolerance      VIF
1        RF 0.1224824 8.164441
2      MaxT 0.2505114 3.991834
3      MinT 0.4862094 2.056727
4       MRH 0.1196535 8.357468
5       ERH 0.7045172 1.419412
```

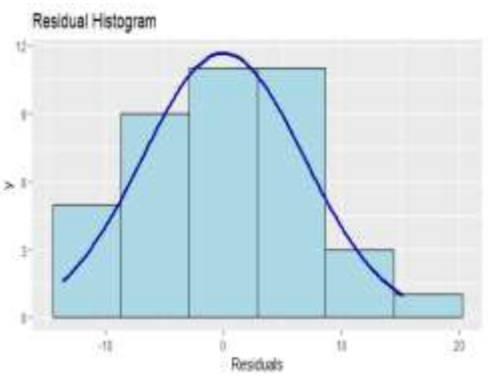
```
     Rainbow test
```

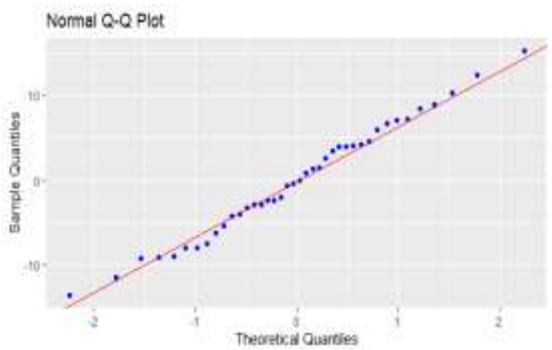
```
data:  pp_MLR
Rain = 0.97306, df1 = 20, df2 = 14, p-value = 0.5334

 lag Autocorrelation D-W Statistic p-value
   1       -0.304291      2.570899   0.072
 Alternative hypothesis: rho != 0
```











A.VI

Pathein

```
Call:
lm(formula = ptrice ~ ptrf + ptmaxt + ptmrh + pterh, data = pt)

Residuals:
     Min       1Q   Median       3Q      Max 
-12.0084  -3.3632   0.0046   2.9693  14.7775 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) -220.49636   36.85821  -5.982 1.93e-07 ***
ptrf          -0.02930    0.01028  -2.852  0.00619 ** 
ptmaxt         6.01139    0.95860   6.271 6.69e-08 ***
ptmrh          1.05714    0.35325   2.993  0.00419 ** 
pterh          0.24735    0.10957   2.258  0.02812 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.555 on 53 degrees of freedom
Multiple R-squared:  0.7521,	Adjusted R-squared:  0.7333 
F-statistic: 40.19 on 4 and 53 DF,  p-value: 1.867e-15

Rainbow test

data:  ptmodel
Rain = 2.204, df1 = 29, df2 = 26, p-value = 0.0860

 Breusch Pagan Test for Heteroskedasticity
 -----------------------------------------
 Ho: the variance is constant            
 Ha: the variance is not constant        

              Data                
 ----------------------------------
 Response : ptrice 
 Variables: fitted values of ptrice 

        Test Summary          
 -----------------------------
 DF            =    1 
 Chi2          =    0.05093126 
 Prob > Chi2   =    0.8214507 

        Shapiro-Wilk normality test

data:  ptmodel$residuals
W = 0.98378, p-value = 0.6292

  Variables  Tolerance       VIF
1      ptrf 0.09196134 10.874135
2    ptmaxt 0.27832964  3.592862
3     ptmrh 0.09944920 10.055386
4     pterh 0.36740786  2.721771

lag Autocorrelation D-W Statistic p-value
   1       0.2542836      1.394119    0.01
 Alternative hypothesis: rho != 0

RMSE [1] 5.309737

mape [1] 5.41029
```

Hinthada

```
Call:
lm(formula = htdrice ~ htdmrh + htderh, data = hin)
```

```
Residuals:
    Min      1Q  Median      3Q     Max 
-31.295  -5.077  -1.594   4.080  21.018 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 109.0713    12.1455   8.980 2.27e-12 ***
htdmrh       -1.2020     0.2216  -5.424 1.34e-06 ***
htderh        0.7648     0.1553   4.924 8.12e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 9.959 on 55 degrees of freedom
Multiple R-squared:  0.3649,	Adjusted R-squared:  0.3418 
F-statistic:  15.8 on 2 and 55 DF,  p-value: 3.786e-06

	Rainbow test

data:  htdmodel
Rain = 1.865, df1 = 29, df2 = 26, p-value = 0.05608

 Breusch Pagan Test for Heteroskedasticity
 -----------------------------------------
 Ho: the variance is constant            
 Ha: the variance is not constant        

              Data               
 -----------------------------------
 Response : htdrice 
 Variables: fitted values of htdrice 

        Test Summary         
 ----------------------------
 DF            =    1 
 Chi2          =    1.958035 
 Prob > Chi2   =    0.1617237 

	Shapiro-Wilk normality test

data:  htdmodel$residuals
W = 0.94393, p-value = 0.009667

  Variables Tolerance    VIF
1    htdmrh 0.4837227 2.0673
2    htderh 0.4837227 2.0673

lag Autocorrelation D-W Statistic p-value
   1       0.5957406     0.7551959       0
 Alternative hypothesis: rho != 0

RMSE [1] 9.697592
MAPE [1] 9.762927
```

Maubin

```
Call:
lm(formula = mbrice ~ mbrf + mbmint + mberh, data = mub)

Residuals:
     Min       1Q   Median       3Q      Max 
-12.6196  -4.6133  -0.2767   3.6085  17.7652 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 61.561817  10.399364   5.920 2.29e-07 ***
mbrf        -0.050787   0.007301  -6.956 4.85e-09 ***
mbmint      -1.431790   0.430123  -3.329  0.00158 ** 
```

```
mberh        0.674877   0.100685   6.703 1.25e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.796 on 54 degrees of freedom
Multiple R-squared:  0.7155,   Adjusted R-squared:  0.6997
F-statistic: 45.27 on 3 and 54 DF,  p-value: 9.203e-15

        Rainbow test

data:  mbmodel
Rain = 1.9484, df1 = 29, df2 = 25, p-value = 0.04707

 Breusch Pagan Test for Heteroskedasticity
 -----------------------------------------
 Ho: the variance is constant
 Ha: the variance is not constant

               Data
 ----------------------------------
 Response : mbrice
 Variables: fitted values of mbrice

        Test Summary
 ----------------------------
 DF            =    1
 Chi2          =    0.4883446
 Prob > Chi2   =    0.4846667

        Shapiro-Wilk normality test

data:  mbmodel$residuals
W = 0.98078, p-value = 0.4855

  Variables Tolerance      VIF
1      mbrf 0.3697702 2.704382
2    mbmint 0.4595028 2.176265
3     mberh 0.5635079 1.774598

 lag Autocorrelation D-W Statistic p-value
   1       0.5535663     0.7549234       0
 Alternative hypothesis: rho != 0
RMSE
[1] 6.557238
mape
[1] 6.802359
```

Myaungmya

```
Call:
lm(formula = mmrice ~ mmrf + mmmaxt + mmerh, data = mm)

Residuals:
    Min      1Q  Median      3Q     Max
-17.486  -5.026   1.046   5.364  16.912

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -42.869924  34.605331  -1.239  0.22077
mmrf         -0.030706   0.006539  -4.696 1.86e-05 ***
mmmaxt        3.295618   1.100971   2.993  0.00416 **
mmerh         0.297492   0.110795   2.685  0.00961 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8.027 on 54 degrees of freedom
Multiple R-squared:  0.5896,   Adjusted R-squared:  0.5668
F-statistic: 25.86 on 3 and 54 DF,  p-value: 1.664e-10
```

```
        Rainbow test

data:  mmmodel
Rain = 3.0886, df1 = 29, df2 = 25, p-value = 0.00273

 Breusch Pagan Test for Heteroskedasticity
 -----------------------------------------
 Ho: the variance is constant
 Ha: the variance is not constant

              Data
 ----------------------------------
 Response : mmrice
 Variables: fitted values of mmrice

        Test Summary
 -----------------------------
 DF            =    1
 Chi2          =    5.266724
 Prob > Chi2   =    0.02173688


Shapiro-Wilk normality test

data:  mmmodel$residuals
W = 0.97541, p-value = 0.2859

  Variables Tolerance      VIF
1      mmrf 0.4209208 2.375744
2    mmmaxt 0.6236827 1.603379
3     mmerh 0.5980149 1.672199

 lag Autocorrelation D-W Statistic p-value
   1       0.5067291     0.9499648       0
 Alternative hypothesis: rho != 0
RMSE
[1] 7.745396
MAPE
[1] 8.225762
```

Phyarpon

```
Call:
lm(formula = pprice ~ pprf, data = pp)

Residuals:
    Min      1Q  Median      3Q     Max 
-19.988  -6.656   2.315   4.769  14.758 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 95.901613   1.758269   54.54  < 2e-16 ***
pprf        -0.060709   0.004929  -12.32 7.73e-15 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.58 on 38 degrees of freedom
Multiple R-squared:  0.7997,	Adjusted R-squared:  0.7944 
F-statistic: 151.7 on 1 and 38 DF,  p-value: 7.732e-15


        Rainbow test

data:  ppmodel
Rain = 1.2301, df1 = 20, df2 = 18, p-value = 0.3316
```

```
 Breusch Pagan Test for Heteroskedasticity
 -----------------------------------------
 Ho: the variance is constant
 Ha: the variance is not constant

              Data
 ----------------------------------
 Response : pprice
 Variables: fitted values of pprice

        Test Summary
 ----------------------------
 DF            =    1
 Chi2          =    1.221379
 Prob > Chi2   =    0.2690902

        Shapiro-Wilk normality test

data:  ppmodel$residuals
W = 0.95337, p-value = 0.09904

lag Autocorrelation D-W Statistic p-value
   1     -0.09819864      2.159171   0.422
 Alternative hypothesis: rho != 0
RMSE
[1] 7.387737
MAPE
[1] 8.64007
```

APPENDIX(B)

Autoregressive Integrated Moving Average with Predictors (ARIMAX) and Seasonal Integrated Moving Average with Predictors (SARIMAX)

B.I. Pathein

```
Regression with ARIMA(1,1,1)(1,0,1)[2] errors : 310.115
 Regression with ARIMA(0,1,0)           errors : 364.3904
 Regression with ARIMA(1,1,0)(1,0,0)[2] errors : 310.9435
 Regression with ARIMA(0,1,1)(0,0,1)[2] errors : 342.8256
 Regression with ARIMA(0,1,0)           errors : 361.9616
 Regression with ARIMA(1,1,1)(0,0,1)[2] errors : 307.4122
 Regression with ARIMA(1,1,1)           errors : 304.4772
 ARIMA(1,1,1)(1,0,0)[2] with drift         : Inf
 Regression with ARIMA(0,1,1)           errors : 346.8483
 Regression with ARIMA(1,1,0)           errors : 308.2459
 Regression with ARIMA(1,1,1)           errors : 303.6955
 ARIMA(1,1,1)(1,0,0)[2]                    : Inf
 Regression with ARIMA(1,1,1)(0,0,1)[2] errors : 306.3011
 Regression with ARIMA(1,1,1)(1,0,1)[2] errors : 307.7855
 Regression with ARIMA(0,1,1)           errors : 346.1584
 Regression with ARIMA(1,1,0)           errors : 308.9257

 Best model: Regression with ARIMA(1,1,1)           errors
Regression with ARIMA(1,1,1) errors

Coefficients:
          ar1     ma1    ptrf  ptmaxt   ptmint    ptmrh    pterh
      -0.9899  0.3827  0.0010  0.0497  -0.2034  -0.1587  -0.0107
s.e.   0.0124  0.1192  0.0057  0.5630   0.2956   0.2153   0.1177

sigma^2 = 9.33:  log likelihood = -142.35
AIC=300.7   AICc=303.7   BIC=317.04
```

```
Training set error measures:
                    ME    RMSE      MAE       MPE     MAPE
MASE
Training set 0.5003729 2.83607 1.978273 0.5710126 2.666476
0.8938459
                     ACF1
Training set 0.0001147797

z test of coefficients:

         Estimate Std. Error  z value  Pr(>|z|)    
ar1    -0.9898684  0.0123715 -80.0122 < 2.2e-16 ***
ma1     0.3826549  0.1192145   3.2098  0.001328 ** 
ptrf    0.0010320  0.0056914   0.1813  0.856111    
ptmaxt  0.0496828  0.5629601   0.0883  0.929676    
ptmint -0.2034253  0.2955579  -0.6883  0.491279    
ptmrh  -0.1586657  0.2153389  -0.7368  0.461233    
pterh  -0.0106750  0.1177013  -0.0907  0.927735    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


      Ljung-Box test

data:  Residuals from Regression with ARIMA(1,1,1) errors
Q* = 1.5387, df = 3, p-value = 0.6734

Model df: 2.   Total lags used: 5

R squared
 [1] 0.9319871

Call:
arimax(x = ptrice, order = c(1, 0, 1), seasonal = list(order =
c(3, 1, 3), period = 2),
    xreg = reg, method = "ML", xtransf = reg)

Coefficients:
         ar1     ma1     sar1     sar2     sar3    sma1    sma2
sma3
      0.2385  0.2482  -0.8169  -0.1700  -0.0431  1.2158  0.6743
-0.2259
s.e.  0.2974  0.2969   0.5256   0.5873   0.2552  0.5168  0.7673
0.5242
         ptrf   ptmaxt  ptmint    ptmrh   pterh
      -0.0016  -0.5734  0.0427  -0.3334  0.0454
s.e.   0.0029   0.2951  0.1598   0.1684  0.1056

sigma^2 estimated as 5.181:  log likelihood = -130.44,  aic =
286.88

Training set error measures:
                    ME     RMSE      MAE       MPE     MAPE
MASE
Training set 0.3693254 2.236773 1.589887 0.4625826 2.147346
0.1224709
                    ACF1
```

```
Training set 0.007586958

z test of coefficients:

         Estimate Std. Error z value Pr(>|z|)
ar1     0.2385356  0.2973751  0.8021  0.42247
ma1     0.2481801  0.2969024  0.8359  0.40321
sar1   -0.8169293  0.5256038 -1.5543  0.12012
sar2   -0.1700194  0.5873053 -0.2895  0.77221
sar3   -0.0431450  0.2551631 -0.1691  0.86573
sma1    1.2158316  0.5168348  2.3525  0.01865 *
sma2    0.6743288  0.7672589  0.8789  0.37947
sma3   -0.2258811  0.5242092 -0.4309  0.66654
ptrf   -0.0016088  0.0029464 -0.5460  0.58505
ptmaxt -0.5734197  0.2950906 -1.9432  0.05199 .
ptmint  0.0427287  0.1598420  0.2673  0.78922
ptmrh  -0.3333572  0.1684441 -1.9790  0.04781 *
pterh   0.0454400  0.1056468  0.4301  0.66711
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


      Ljung-Box test

data:  Residuals from ARIMA(1,0,1)(3,1,3)[2]
Q* = 10.46, df = 3, p-value = 0.01503

Model df: 8.   Total lags used: 11

Call:
arimax(x = ptrice, order = c(0, 0, 1), seasonal = list(order =
c(1, 1, 3), period = 2),
    xreg = reg, method = "ML", xtransf = reg)

Coefficients:
        ma1     sar1    sma1    sma2     sma3     ptrf   ptmaxt
ptmint    ptmrh   pterh
      0.451  -0.6414  1.1389  0.5523  -0.3092  -0.0015  -0.5703
0.0358  -0.3351  0.0494
s.e.  0.123   0.1539  0.2248  0.2793   0.1857   0.0029   0.2975
0.1614   0.1638  0.1021

sigma^2 estimated as 5.249:  log likelihood = -130.78,  aic =
281.55

Training set error measures:
                    ME     RMSE      MAE       MPE     MAPE
MASE       ACF1
Training set 0.3780531 2.251402 1.600317 0.4744839 2.156512
0.1232743 0.03953285
[1] 0.9575974

z test of coefficients:

         Estimate Std. Error z value  Pr(>|z|)
ma1     0.4510072  0.1229756  3.6675  0.000245 ***
sar1   -0.6414242  0.1539408 -4.1667 3.090e-05 ***
sma1    1.1388566  0.2247644  5.0669 4.044e-07 ***
```

```
sma2    0.5523134  0.2793387  1.9772  0.048017 *
sma3   -0.3091698  0.1856940 -1.6649  0.095924 .
intercept9.2265452  8.4798216 1.0881  0.2765691
ptrf   -0.0014768  0.0029418 -0.5020  0.615661
ptmaxt -0.5703193  0.2974654 -1.9173  0.055205 .
ptmint  0.0357703  0.1613567  0.2217  0.824559
ptmrh  -0.3350968  0.1637847 -2.0460  0.040760 *
pterh   0.0493818  0.1020814  0.4837  0.628564
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


      Ljung-Box test
data:  Residuals from ARIMA(0,0,1)(1,1,3)[2]
Q* = 11.371, df = 3, p-value = 0.019878

arimax(x = ptrice, order = c(0, 0, 1), seasonal = list(order =
c(1, 1, 3), period = 2),
    xreg = sreg, method = "ML", xtransf = sreg)

Coefficients:
         ma1     sar1    sma1    sma2     sma3   ptmaxt    ptmrh
      0.4377  -0.6123  1.0936  0.5162  -0.3378  -0.5624  -0.2940
s.e.  0.1235   0.1601  0.2359  0.2781   0.1808   0.2483   0.1363

sigma^2 estimated as 5.323:  log likelihood = -131.11,  aic =
276.23

Training set error measures:
                    ME     RMSE      MAE       MPE     MAPE
MASE       ACF1
Training set 0.3946509 2.267093 1.617656 0.4947128 2.172091
0.12461 0.03201438

z test of coefficients:

       Estimate Std. Error z value  Pr(>|z|)
ma1     0.43766    0.12348  3.5444 0.0003935 ***
sar1   -0.61230    0.16009 -3.8248 0.0001309 ***
sma1    1.09361    0.23594  4.6351 3.568e-06 ***
sma2    0.51625    0.27813  1.8561 0.0634350 .
sma3   -0.33784    0.18076 -1.8690 0.0616224 .
ptmaxt -0.56242    0.24834 -2.2647 0.0235314 *
ptmrh  -0.29398    0.13632 -2.1565 0.0310466 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R squared
[1] 0.957059

      Ljung-Box test

data:  Residuals from ARIMA(0,0,1)(1,1,3)[2]
Q* = 7.1626, df = 3, p-value = 0.06689

Model df: 5.   Total lags used: 8
```

B.II Hinthada

```
Regression with ARIMA(1,1,1)(1,0,1)[2] errors : 387.9455
 Regression with ARIMA(0,1,0)           errors : 419.9694
 Regression with ARIMA(1,1,0)(1,0,0)[2] errors : 385.716
 Regression with ARIMA(0,1,1)(0,0,1)[2] errors : 393.1644
 Regression with ARIMA(0,1,0)           errors : 417.4555
 Regression with ARIMA(1,1,0)           errors : 383.5001
 Regression with ARIMA(1,1,0)(0,0,1)[2] errors : 385.7612
 Regression with ARIMA(1,1,0)(1,0,1)[2] errors : 388.6535
 Regression with ARIMA(1,1,1)           errors : 383.6851
 Regression with ARIMA(0,1,1)           errors : 402.9395
 Regression with ARIMA(1,1,0)           errors : 381.4578
 Regression with ARIMA(1,1,0)(1,0,0)[2] errors : 383.9025
 Regression with ARIMA(1,1,0)(0,0,1)[2] errors : 383.9348
 Regression with ARIMA(1,1,0)(1,0,1)[2] errors : 386.6858
 Regression with ARIMA(1,1,1)           errors : 381.3171
 Regression with ARIMA(1,1,1)(1,0,0)[2] errors : 382.7714
 Regression with ARIMA(1,1,1)(0,0,1)[2] errors : 382.8774
 Regression with ARIMA(1,1,1)(1,0,1)[2] errors : 385.7027
 Regression with ARIMA(0,1,1)           errors : 400.8757

 Best model: Regression with ARIMA(1,1,1)           errors

Series: htdrice
Regression with ARIMA(1,1,1) errors

Coefficients:
          ar1     ma1    htdrf  htdmaxt  htdmint   htdmrh
htderh
      -0.9582  0.3600  -0.0091   1.1503   1.1137  -0.0338  -
0.1590
s.e.   0.0545  0.1859   0.0165   1.4291   0.8745   0.3209
0.1605

sigma^2 = 37.27:  log likelihood = -181.16
AIC=378.32   AICc=381.32   BIC=394.66

Series: htdrice
Regression with ARIMA(1,1,1) errors

Coefficients:
          ar1     ma1    htdrf  htdmaxt  htdmint   htdmrh
htderh
      -0.9582  0.3600  -0.0091   1.1503   1.1137  -0.0338  -
0.1590
s.e.   0.0545  0.1859   0.0165   1.4291   0.8745   0.3209
0.1605

sigma^2 = 37.27:  log likelihood = -181.16
AIC=378.32   AICc=381.32   BIC=394.66

Training set error measures:
                   ME     RMSE      MAE         MPE     MAPE
MASE
Training set 0.408517 5.668053 3.666344 0.004792548 5.173931
1.046352
                    ACF1
```

```
Training set -0.04737495

z test of coefficients:

          Estimate Std. Error  z value Pr(>|z|)
ar1     -0.9581542  0.0545452 -17.5663   <2e-16 ***
ma1      0.3600446  0.1859200   1.9366   0.0528 .
htdrf   -0.0090797  0.0165220  -0.5496   0.5826
htdmaxt  1.1502982  1.4291366   0.8049   0.4209
htdmint  1.1136760  0.8744872   1.2735   0.2028
htdmrh  -0.0338209  0.3208626  -0.1054   0.9161
htderh  -0.1590192  0.1605353  -0.9906   0.3219
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R squared
[1] 0.7895511

      Ljung-Box test

data:  Residuals from Regression with ARIMA(1,1,1) errors
Q* = 1.1395, df = 3, p-value = 0.7675

Model df: 2.   Total lags used: 5

arimax(x = hinsed, order = c(2, 0, 0), seasonal = list(order =
c(0, 1, 0), period = 2),
    xreg = reg, method = "ML", xtransf = reg)

Coefficients:
         ar1      ar2  intercept      rf     maxt     mint
mrh      erh
      0.2758  -0.4306    29.0431  0.0106  -0.2623  -0.6864  -
0.0032  -0.0814
s.e.  0.1249   0.1358    60.4338  0.0144   1.5283   0.7306
0.2716   0.0954

sigma^2 estimated as 32.87:  log likelihood = -177.48,  aic =
370.96

Training set error measures:
                    ME   RMSE      MAE      MPE     MAPE
MASE       ACF1
Training set -0.1299564 5.7335 3.587045 1.300904 4.295366
0.7495156 0.04129289

z test of coefficients:

            Estimate Std. Error z value Pr(>|z|)
ar1        0.2757750  0.1249048  2.2079 0.027252 *
ar2       -0.4306464  0.1358367 -3.1703 0.001523 **
intercept 29.0430965 60.4338208  0.4806 0.630817
rf         0.0105537  0.0143675  0.7346 0.462609
maxt      -0.2623094  1.5283317 -0.1716 0.863727
mint      -0.6863959  0.7306317 -0.9395 0.347497
mrh       -0.0031728  0.2716075 -0.0117 0.990680
erh       -0.0813526  0.0953901 -0.8528 0.393747
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

[1] 0.2199048

      Ljung-Box test

data:  Residuals from ARIMA(2,0,0) with non-zero mean
Q* = 4.7311, df = 3, p-value = 0.1926

Series: htdrice
ARIMA(2,0,0)(0,1,0)[2]

Coefficients:
         ar1      ar2
      0.3072  -0.3377
s.e.  0.1260   0.1322

sigma^2 = 37.4:  log likelihood = -180
AIC=366   AICc=366.46   BIC=372.07

Training set error measures:
                   ME     RMSE     MAE       MPE     MAPE
MASE       ACF1
Training set 0.813484 5.900971 3.48407 0.6112431 4.856636
0.9943325 0.02403286

z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ar1  0.30720    0.12604  2.4373  0.01480 *
ar2 -0.33768    0.13217 -2.5549  0.01062 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R squared
[1] 0.7807626

      Ljung-Box test

data:  Residuals from ARIMA(2,0,0)(0,1,0)[2]
Q* = 1.8979, df = 3, p-value = 0.5939

Model df: 2.   Total lags used: 5
```

B.III. Maubin

```
ARIMA(1,1,1)(1,0,1)[2] with drift         : Inf
 Regression with ARIMA(0,1,0)           errors : 363.8766
 Regression with ARIMA(1,1,0)(1,0,0)[2] errors : 294.0561
 Regression with ARIMA(0,1,1)(0,0,1)[2] errors : 326.8265
 Regression with ARIMA(0,1,0)           errors : 361.5882
 ARIMA(1,1,0) with drift         : Inf
 Regression with ARIMA(1,1,0)(2,0,0)[2] errors : 296.947
 Regression with ARIMA(1,1,0)(1,0,1)[2] errors : 296.9866
 Regression with ARIMA(1,1,0)(0,0,1)[2] errors : 294.0675
 Regression with ARIMA(1,1,0)(2,0,1)[2] errors : Inf
 ARIMA(0,1,0)(1,0,0)[2] with drift         : Inf
 Regression with ARIMA(1,1,1)(1,0,0)[2] errors : Inf
```

```
 Regression with ARIMA(0,1,1)(1,0,0)[2] errors : Inf
 Regression with ARIMA(1,1,0)(1,0,0)[2] errors : 294.8869

 Best model: Regression with ARIMA(1,1,0)(1,0,0)[2] errors

Series: mbrice
Regression with ARIMA(1,1,0)(1,0,0)[2] errors

Coefficients:
          ar1    sar1   drift     mbrf  mbmaxt   mbmint    mbmrh
mberh
      -0.9773  0.1801  0.4273  -0.0067  0.1342  -0.0551  -0.0694
0.0461
s.e.   0.0233  0.1523  0.2095   0.0055  0.2610   0.1655   0.1113
0.0695

sigma^2 = 7.592:  log likelihood = -136.11
AIC=290.23   AICc=294.06   BIC=308.61

Training set error measures:
                      ME     RMSE      MAE        MPE     MAPE
MASE
Training set -0.02013238 2.532537 1.747725 -0.1123474 2.428512
0.9245475
                   ACF1
Training set 0.08119264

z test of coefficients:

         Estimate Std. Error  z value Pr(>|z|)
ar1    -0.9772589  0.0233446 -41.8623  < 2e-16 ***
sar1    0.1801016  0.1522909   1.1826  0.23696
drift   0.4272936  0.2094560   2.0400  0.04135 *
mbrf   -0.0066647  0.0054891  -1.2142  0.22469
mbmaxt  0.1342481  0.2610019   0.5144  0.60700
mbmint -0.0551007  0.1654864  -0.3330  0.73916
mbmrh  -0.0694464  0.1112689  -0.6241  0.53254
mberh   0.0460909  0.0694635   0.6635  0.50699
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R squared
[1] 0.9577671

      Ljung-Box test

data:  Residuals from Regression with ARIMA(1,1,0)(1,0,0)[2]
errors
Q* = 3.5494, df = 3, p-value = 0.3144

Model df: 2.   Total lags used: 5

Call:
arimax(x = mbrice, order = c(2, 0, 1), seasonal = list(order =
c(0, 1, 0), period = 2),
    xreg = reg, method = "ML", xtransf = reg)

Coefficients:
```

```
         ar1     ar2      ma1     mbrf  mbmaxt   mbmint    mbmrh
mberh
      0.5701  0.1181  -0.3635  -0.0056  0.0634  -0.0149  -0.1269
0.0689
s.e.  0.3356  0.1839   0.3148   0.0054  0.2637   0.1642   0.1091
0.0704

sigma^2 estimated as 6.582:  log likelihood = -132.33,  aic =
280.65

Training set error measures:
                    ME    RMSE      MAE       MPE     MAPE
MASE        ACF1
Training set 0.3897657 2.52095 1.767157 0.4593043 2.460984
0.1060383 -0.03348187

z test of coefficients:

         Estimate Std. Error z value Pr(>|z|)
ar1     0.5700891  0.3355862  1.6988  0.08936 .
ar2     0.1180683  0.1839393  0.6419  0.52095
ma1    -0.3635391  0.3147661 -1.1549  0.24811
mbrf   -0.0056411  0.0054026 -1.0441  0.29642
mbmaxt  0.0633732  0.2637344  0.2403  0.81010
mbmint -0.0148617  0.1641781 -0.0905  0.92787
mbmrh  -0.1268638  0.1090773 -1.1631  0.24480
mberh   0.0688655  0.0703536  0.9788  0.32766
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R squared
[1] 0.9594292

      Ljung-Box test

data:  Residuals from ARIMA(2,0,1)(0,1,0)[2]
Q* = 4.6155, df = 3, p-value = 0.2022

Model df: 3.   Total lags used: 6

Call:
arimax(x = sddmb, order = c(0, 1, 1), seasonal = list(order =
c(0, 1, 0), period = 2),
    xreg = reg, method = "ML", xtransf = reg)

Coefficients:
          ma1  intercept       rf    maxt    mint      mrh
erh
      -0.9999    -0.9763  -0.0022  0.0712  0.0691  -0.0214  -
0.0266
s.e.   0.0524     1.7168   0.0042  0.2990  0.0534   0.0975
0.0531

sigma^2 estimated as 6.232:  log likelihood = -130.37,  aic =
274.74

Training set error measures:
```

```
                     ME     RMSE     MAE MPE MAPE      MASE
ACF1
Training set -0.2339107 2.496404 1.67486 -0.3124133 2.483727
0.4097795 0.08380196

z test of coefficients:

            Estimate  Std. Error  z value Pr(>|z|)
ma1       -0.9999204  0.0524315 -19.0710   <2e-16 ***
intercept -0.9762667  1.7168028  -0.5687   0.5696
rf        -0.0021592  0.0041972  -0.5144   0.6069
maxt       0.0711936  0.2989598   0.2381   0.8118
mint       0.0691159  0.0533749   1.2949   0.1954
mrh       -0.0213710  0.0974605  -0.2193   0.8264
erh       -0.0266351  0.0530958  -0.5016   0.6159
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


      Ljung-Box test

data:  Residuals from ARIMA(0,0,1) with non-zero mean
Q* = 5.1645, df = 3, p-value = 0.1601

Model df: 7.   Total lags used: 10

[1] 0.4784285

ARIMA(0,1,1)(0,1,0)[2]

Coefficients:
          ma1
      -0.9276
s.e.   0.1074

sigma^2 = 7.418:  log likelihood = -133.63
AIC=271.26   AICc=271.49   BIC=275.27

Training set error measures:
                     ME     RMSE    MAE        MPE     MAPE
MASE      ACF1
Training set -0.3251768 2.628011 1.7549 -0.5269496 2.460186
0.9283433 0.1222492

z test of coefficients:
    Estimate Std. Error z value  Pr(>|z|)
ma1 -0.92756    0.10737 -8.6393 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R squared
[1] 0.9552801

      Ljung-Box test
data:  Residuals from ARIMA(0,1,1)(0,1,0)[2]
Q* = 2.4993, df = 3, p-value = 0.4754

Model df: 1.   Total lags used: 4
```

B.IV Myaungmya

```
ARIMA(1,1,1)(1,0,1)[2] with drift         : Inf
 Regression with ARIMA(0,1,0)           errors : 358.6095
 Regression with ARIMA(1,1,0)(1,0,0)[2] errors : 312.3746
 Regression with ARIMA(0,1,1)(0,0,1)[2] errors : 346.2754
 Regression with ARIMA(0,1,0)           errors : 356.4565
 Regression with ARIMA(1,1,0)           errors : 309.5532
 Regression with ARIMA(1,1,0)(0,0,1)[2] errors : 312.3656
 Regression with ARIMA(1,1,0)(1,0,1)[2] errors : Inf
 ARIMA(1,1,1) with drift         : Inf
 Regression with ARIMA(0,1,1)           errors : 348.2178
 Regression with ARIMA(1,1,0)           errors : 311.867

 Best model: Regression with ARIMA(1,1,0)           errors

Series: mmrice
Regression with ARIMA(1,1,0) errors

Coefficients:
          ar1   drift    mmrf  mmmaxt   mmmint    mmmrh   mmerh
      -0.9781  0.4688  0.0007  0.3169  -0.3227  -0.1055  0.0357
s.e.   0.0205  0.2044  0.0031  0.3522   0.2213   0.0977  0.0860

sigma^2 = 10.33:  log likelihood = -145.28
AIC=306.55   AICc=309.55   BIC=322.9

Training set error measures:
                       ME     RMSE      MAE         MPE     MAPE
MASE
Training set -0.001932151 2.984834 1.772922 -0.09721582 2.358447
0.8539057
                  ACF1
Training set 0.1772359

z test of coefficients:

          Estimate  Std. Error  z value Pr(>|z|)
ar1    -0.97810382  0.02045488 -47.8176  < 2e-16 ***
drift   0.46883018  0.20441293   2.2935  0.02182 *
mmrf    0.00071244  0.00306678   0.2323  0.81630
mmmaxt  0.31692305  0.35224903   0.8997  0.36827
mmmint -0.32266550  0.22126380  -1.4583  0.14476
mmmrh  -0.10554252  0.09772541  -1.0800  0.28015
mmerh   0.03569567  0.08604425   0.4149  0.67825
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R squared
[1] 0.9398214

      Ljung-Box test

data:  Residuals from Regression with ARIMA(1,1,0) errors
Q* = 7.2952, df = 3, p-value = 0.06306

Model df: 1.   Total lags used: 4
```

```
arimax(x = mmrice, order = c(1, 1, 1), seasonal = list(order =
c(2, 1, 0), period = 2),
    xreg = reg, method = "ML", xtransf = reg)

Coefficients:
         ar1     ma1    sar1     sar2    mmrf  mmmaxt   mmmint
mmmrh   mmerh
      0.2588  -1.000  0.0201  -0.2943  0.0014  0.1872  -0.3853  -
0.0463  0.0461
s.e.  0.1363   0.463  0.1577   0.1532  0.0028  0.3599   0.2396
0.0908  0.0833

sigma^2 estimated as 8.288:  log likelihood = -138.36,  aic =
294.72

Training set error measures:
                    ME     RMSE      MAE        MPE     MAPE
MASE
Training set -0.231984 2.803413 1.815288 -0.3638835 2.417426
0.1081556
                    ACF1
Training set 0.004250585

z test of coefficients:

         Estimate Std. Error z value Pr(>|z|)
ar1     0.2587856  0.1362592  1.8992  0.05754 .
ma1    -0.9999844  0.4630220 -2.1597  0.03080 *
sar1    0.0200886  0.1577286  0.1274  0.89865
sar2   -0.2943475  0.1532386 -1.9208  0.05475 .
mmrf    0.0014244  0.0028252  0.5042  0.61413
mmmaxt  0.1872338  0.3599099  0.5202  0.60291
mmmint -0.3853391  0.2395663 -1.6085  0.10773
mmmrh  -0.0462530  0.0908286 -0.5092  0.61059
mmerh   0.0460837  0.0832665  0.5534  0.57996
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R squared
[1] 0.9476837

      Ljung-Box test

data:  Residuals from ARIMA(1,1,1)(2,1,0)[2]
Q* = 6.1748, df = 3, p-value = 0.1034

Model df: 4.   Total lags used: 7

Call:
arimax(x = mmrice, order = c(1, 1, 1), seasonal = list(order =
c(0, 1, 0), period = 2),
    xreg = reg, method = "ML", xtransf = reg)

Coefficients:
         ar1      ma1   mmrf  mmmaxt   mmmint    mmmrh   mmerh
      0.2299  -1.0000  7e-04  0.2717  -0.3278  -0.0900  0.0391
s.e.  0.1334   0.0986  3e-03  0.3534   0.2446   0.0967  0.0877
```

```
sigma^2 estimated as 8.936:  log likelihood = -140.05,  aic =
294.11

Training set error measures:
                    ME     RMSE      MAE        MPE     MAPE
MASE       ACF1
Training set -0.3150299 2.911037 1.820575 -0.4708899 2.420361
0.1084706 0.01071571

z test of coefficients:

          Estimate  Std. Error  z value Pr(>|z|)
ar1     0.22987153  0.13339428   1.7232  0.08484 .
ma1    -0.99999387  0.09863033 -10.1388  < 2e-16 ***
intercept0.2041370 12.4529162    0.0164  0.98692
mmrf    0.0034193   0.0018818    1.8170  0.06922 .
mmmaxt  0.27171439  0.35340502   0.7688  0.44198
mmmint -0.32777595  0.24459580  -1.3401  0.18022
mmmrh  -0.09004735  0.09665060  -0.9317  0.35150
mmerh   0.03912710  0.08773343   0.4460  0.65561
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


      Ljung-Box test

data:  Residuals from ARIMA(1,1,1)(0,1,0)[2]
Q* = 11.321, df = 3, p-value = 0.01011

Model df: 7.   Total lags used: 10

[1] 0.9439395

Call:
arimax(x = mmrice, order = c(1, 1, 1), seasonal = list(order =
c(0, 1, 0), period = 2),
    xreg = reg, method = "ML", xtransf = reg)

Coefficients:
         ar1      ma1    mmrf
      0.2458  -1.0000  0.0004
s.e.  0.1325   0.1258  0.0029

sigma^2 estimated as 9.414:  log likelihood = -141.47,  aic =
288.94

Training set error measures:
                    ME     RMSE      MAE        MPE     MAPE
MASE
Training set -0.3980664 2.987863 1.863448 -0.5897515 2.476858
0.111025
                     ACF1
Training set -0.0005257895
> coeftest(model)

z test of coefficients:

        Estimate  Std. Error z value  Pr(>|z|)
```

```
ar1   0.24581183  0.13254803  1.8545   0.06367 .
ma1  -0.99998009  0.12576909 -7.9509 1.851e-15 ***
mmrf  0.00040896  0.00294390  0.1389   0.88951 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> cor(fitted(model),mmrice)^2
[1] 0.9410893

      Ljung-Box test

data:  Residuals from ARIMA(1,1,1)(0,1,0)[2]
Q* = 6.8588, df = 3, p-value = 0.07654

Model df: 2.   Total lags used: 5
```

B.V Phyarpon

```
Regression with ARIMA(1,0,1)(1,1,1)[2] errors : 224.709
 Regression with ARIMA(0,0,0)(0,1,0)[2] errors : 214.9387
 Regression with ARIMA(1,0,0)(1,1,0)[2] errors : 217.4208
 Regression with ARIMA(0,0,1)(0,1,1)[2] errors : 217.5157
 Regression with ARIMA(0,0,0)(0,1,0)[2] errors : 212.1702
 Regression with ARIMA(0,0,0)(1,1,0)[2] errors : 216.4162
 Regression with ARIMA(0,0,0)(0,1,1)[2] errors : 216.646
 Regression with ARIMA(0,0,0)(1,1,1)[2] errors : 219.7542
 Regression with ARIMA(1,0,0)(0,1,0)[2] errors : 214.5486
 Regression with ARIMA(0,0,1)(0,1,0)[2] errors : 215.5144
 Regression with ARIMA(1,0,1)(0,1,0)[2] errors : 217.4571

 Best model: Regression with ARIMA(0,0,0)(0,1,0)[2] errors

Series: pprice
Regression with ARIMA(0,0,0)(0,1,0)[2] errors

Coefficients:
         pprf  ppmaxt  ppmint   ppmrh    pperh
      -0.0051  0.2225  0.3365  0.0763  -0.0409
s.e.   0.0082  0.8792  0.2818  0.2151   0.1153

sigma^2 = 12.18:  log likelihood = -98.73
AIC=209.46   AICc=212.17   BIC=219.29

Training set error measures:
                    ME     RMSE      MAE        MPE     MAPE
MASE      ACF1
Training set 0.2423059 3.169544 1.750137 0.08624623 2.368374
1.050137 0.2817892

z test of coefficients:

         Estimate Std. Error z value Pr(>|z|)
pprf   -0.0051321  0.0082448 -0.6225   0.5336
ppmaxt  0.2225249  0.8791575  0.2531   0.8002
ppmint  0.3365162  0.2817658  1.1943   0.2324
ppmrh   0.0762554  0.2150772  0.3545   0.7229
pperh  -0.0409296  0.1153497 -0.3548   0.7227
```

```
R squared
[1] 0.9633835

      Ljung-Box test

data:  Residuals from Regression with ARIMA(0,0,0)(0,1,0)[2]
errors
Q* = 5.2274, df = 4, p-value = 0.2648

Model df: 0.   Total lags used: 4

Call:
arimax(x = sdpp, order = c(0, 1, 1), seasonal = list(order = c(0,
1, 0), period = 2),
    xreg = reg, method = "ML", xtransf = reg)

Coefficients:
          ma1  intercept       rf    maxt     mint     mrh
erh
      -1.0000   -54.9142  -0.0090  0.2681  -0.2055  0.6130
0.0336
s.e.   0.0708    20.0300   0.0035  0.3681   0.0785  0.1834
0.0376

sigma^2 estimated as 5.48:  log likelihood = -85.79,  aic =
185.58

Training set error measures:
                    ME     RMSE       MAE      MPE     MAPE
MASE         ACF1
Training set 0.0553274 2.340991 1.549626 32.09708 1.947326
0.3913747 -0.02598377

z test of coefficients:

             Estimate  Std. Error  z value  Pr(>|z|)
ma1        -0.9999672   0.0707918 -14.1255 < 2.2e-16 ***
intercept -54.9141964  20.0299907  -2.7416 0.0061141 **
rf         -0.0090114   0.0034649  -2.6008 0.0093017 **
maxt        0.2680868   0.3680607   0.7284 0.4663831
mint       -0.2055279   0.0784684  -2.6192 0.0088125 **
mrh         0.6130050   0.1833970   3.3425 0.0008303 ***
erh         0.0336365   0.0376122   0.8943 0.3711618
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


      Ljung-Box test

data:  Residuals from ARIMA(0,1,1)(0,1,0)[2] with non-zero mean
Q* = 5.6096, df = 3, p-value = 0.1322

Model df: 7.   Total lags used: 10

[1] 0.6929227

arimax(x = sdpp, order = c(0, 1, 1), seasonal = list(order = c(0,
1, 0), period = 2),
```

```
    xreg = reg, method = "ML", xtransf = reg)

Coefficients:
          ma1  intercept       rf     mint     mrh
      -0.8570   -45.6821  -0.0121  -0.1600  0.6399
s.e.   1.3274    50.5338   0.0168   0.6257  0.7454

sigma^2 estimated as 6.932:  log likelihood = -88.98,  aic = 
187.97

Training set error measures:
                     ME    RMSE      MAE       MPE     MAPE     
MASE       ACF1
Training set 0.09942646 2.63294 1.802753 -5.797297 2.370599 
0.4553046 0.07374637
> coeftest(fit1)

z test of coefficients:

            Estimate Std. Error z value Pr(>|z|)
ma1        -0.856967   1.327406 -0.6456   0.5185
intercept -45.682089  50.533771 -0.9040   0.3660
rf         -0.012074   0.016846 -0.7167   0.4735
mint       -0.159999   0.625686 -0.2557   0.7982
mrh         0.639905   0.745352  0.8585   0.3906

> cor(fitted(fit1),sdpp)^2
[1] 0.6004276
> checkresiduals(fit1)

      Ljung-Box test

data:  Residuals from ARIMA(0,0,1) with non-zero mean
Q* = 0.97488, df = 3, p-value = 0.8073

Model df: 1.   Total lags used: 4
```

APPEMDIX (C)

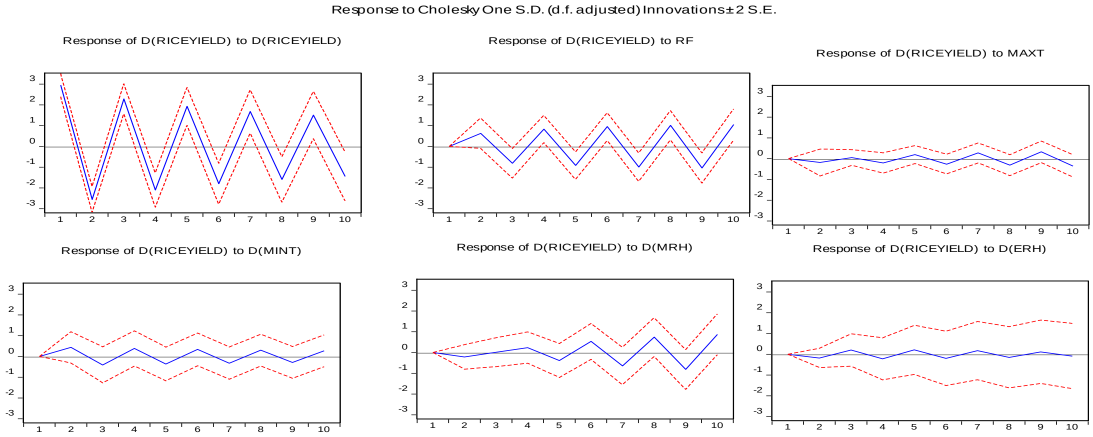
Vector Autoregressive (VAR) Model

C.I (Pathein District)

Equation: D(RICEYIELD) = C(1)*D(RICEYIELD(-1)) + C(2)*RF(-1) + C(3)*MAXT(-1) + C(4)*D(MINT(-1)) + C(5)*D(MRH(-1)) + C(6)*D(ERH(-1)) + C(7)

Observations: 56

R-squared	0.961888	Mean dependent var	0.597679
Adjusted R-squared	0.957221	S.D. dependent var	14.28780
S.E. of regression	2.955148	Sum squared resid	427.9120
Durbin-Watson stat	1.299806		



VAR Residual Heteroskedasticity Tests (Levels and Squares)

Joint test:

Chi-sq	df	Prob.
305.8296	252	0.0115

VAR Residual Normality Tests

Orthogonalization: Cholesky (Lutkepohl)

Null Hypothesis: Residuals are multivariate normal

Component	Jarque-Bera	df	Prob.
1	12.63192	2	0.0018
2	0.334720	2	0.8459
3	2.862786	2	0.2390
4	5.039789	2	0.0805
5	1.423564	2	0.4908
6	44.86212	2	0.0000
Joint	67.15490	12	0.0000

C.II (Hinthada District)

Equation: D(RICEYIELD) = C(1)*D(RICEYIELD(-1)) + C(2)*RF(-1) + C(3)
*D(MAXT(-1)) + C(4)*MINT(-1) + C(5)*MRH(-1) + C(6)*D(ERH(-1)) +
C(7)

Observations: 56

R-squared	0.769630	Mean dependent var	0.635536
Adjusted R-squared	0.741422	S.D. dependent var	12.61214
S.E. of regression	6.413353	Sum squared resid	2015.424
Durbin-Watson stat	1.717768		

VAR Residual Serial Correlation LM Tests
Date: 06/25/23 Time: 00:48
Sample: 1992S1 2020S2
Included observations: 56

Null hypothesis: No serial correlation at lag h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	68.19585	36	0.0009	2.078438	(36, 169.6)	0.0010
2	59.05514	36	0.0091	1.753874	(36, 169.6)	0.0095

VAR Residual Heteroskedasticity Tests (Levels and Squares)

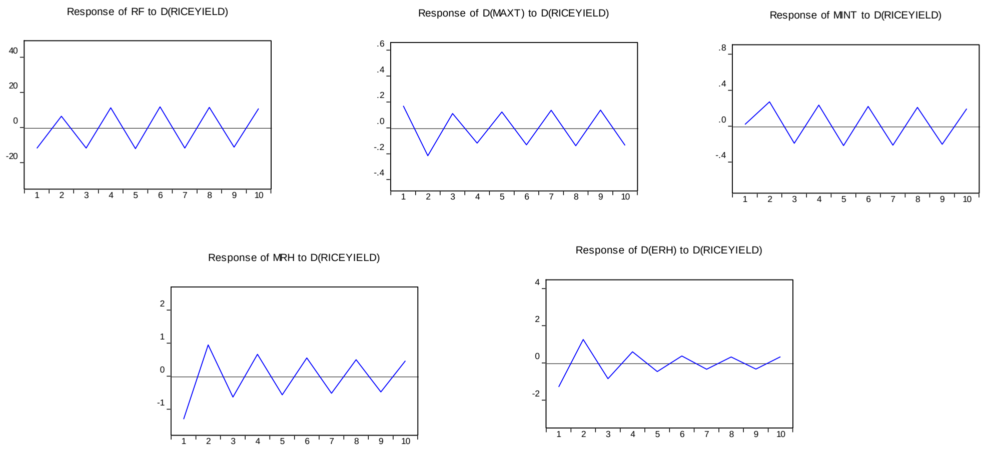
Joint test:

Chi-sq	df	Prob.
1001.669	1008	0.5503

VAR Residual Normality Tests
Orthogonalization: Cholesky (Lutkepohl)
Null Hypothesis: Residuals are multivariate normal

Component	Jarque-Bera	df	Prob.
1	59.76419	2	0.0000
2	0.089252	2	0.9564
3	0.830603	2	0.6601
4	0.774123	2	0.6790
5	0.532114	2	0.7664
6	69.27505	2	0.0000
Joint	131.2653	12	0.0000

Response to Cholesky One S.D. (d.f. adjusted) Innovations



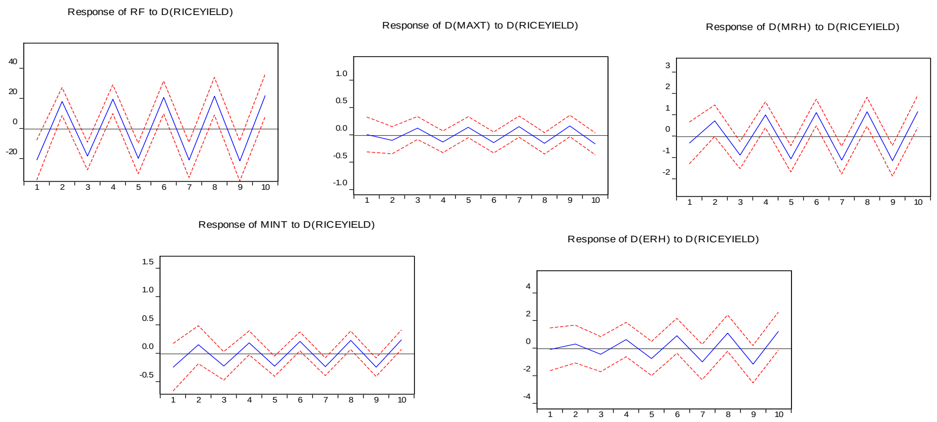
C.II (Maubin District)

Equation: D(RICEYIELD) = C(1)*D(RICEYIELD(-1)) + C(2)*RF(-1) + C(3)*D(MAXT(-1)) + C(4)*MINT(-1) + C(5)*D(MRH(-1)) + C(6)*D(ERH(-1)) + C(7)

Observations: 56

R-squared	0.980047	Mean dependent var	0.599643
Adjusted R-squared	0.977603	S.D. dependent var	18.14674
S.E. of regression	2.715756	Sum squared resid	361.3913
Durbin-Watson stat	1.616836		

Response to Cholesky One S.D. (d.f. adjusted) Innovations ± 2 S.E.



VAR Residual Serial Correlation LM Tests

Null hypothesis: No serial correlation at lag h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	43.80447	36	0.1741	1.246524	(36, 169.6)	0.1780
2	36.20246	36	0.4592	1.008694	(36, 169.6)	0.4643

VAR Residual Heteroskedasticity Tests

Joint test:

Chi-sq	df	Prob.
297.1872	252	0.0266

VAR Residual Normality Tests

Orthogonalization: Cholesky (Lutkepohl)

Null Hypothesis: Residuals are multivariate normal

Component	Jarque-Bera	df	Prob.
1	51.30130	2	0.0000
2	0.462715	2	0.7935
3	59.62378	2	0.0000
4	6.173676	2	0.0456
5	0.147260	2	0.9290
6	10.09472	2	0.0064
Joint	127.8035	12	0.0000

C.IV (Myaungmya District)

Equation: D(RICEYIELD) = C(1)*D(RICEYIELD(-1)) + C(2)*RF(-1) + C(3)
*MAXT(-1) + C(4)*D(MINT(-1)) + C(5)*D(MRH(-1)) + C(6)*D(ERH(-1)) + C(7)

Observations: 56

R-squared	0.973301	Mean dependent var	0.617679
Adjusted R-squared	0.970031	S.D. dependent var	17.71550
S.E. of regression	3.066810	Sum squared resid	460.8610
Durbin-Watson stat	1.707303		

VAR Residual Serial Correlation LM Tests

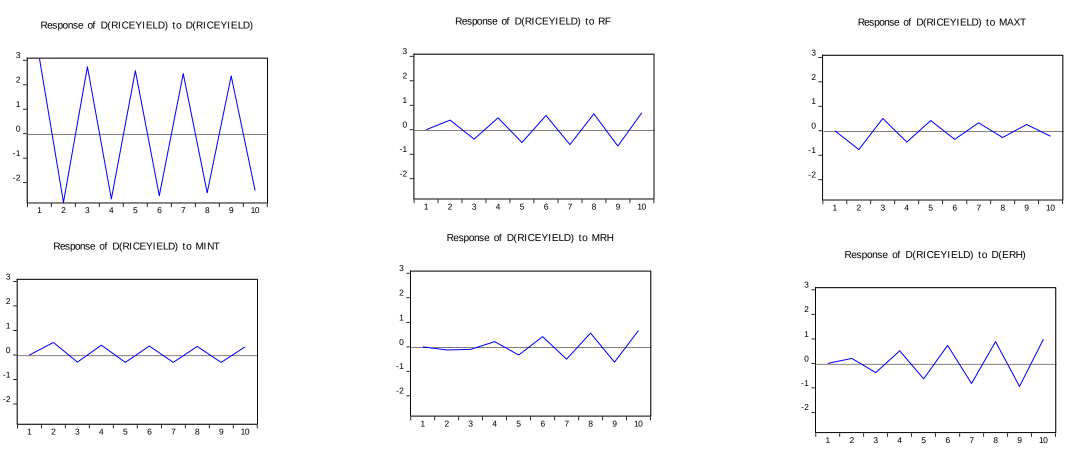
Null hypothesis: No serial correlation at lag h

Sample: 1992S1 2020S2

Included observations: 56

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	60.42800	36	0.0066	1.801612	(36, 169.6)	0.0070
2	43.47355	36	0.1831	1.235969	(36, 169.6)	0.1871

Response to Cholesky One S.D. (d.f. adjusted) Innovations



VAR Residual Heteroskedasticity Tests
(Levels and Squares)

Joint test:

Chi-sq	df	Prob.
275.9557	252	0.1436

VAR Residual Normality Tests
Orthogonalization: Cholesky (Lutkepohl)
Null Hypothesis: Residuals are multivariate normal

Component	Jarque-Bera	df	Prob.
1	175.3103	2	0.0000
2	2677.406	2	0.0000
3	155.0459	2	0.0000
4	1.598327	2	0.4497
5	1.323596	2	0.5159
6	11.66173	2	0.0029
Joint	3022.345	12	0.0000

C.V (Phyarpon District)

Equation: D(RICE_YIELD) = C(1)*D(RICE_YIELD(-1)) + C(2)*D(RF(-1)) + C(3)*D(MAXT(-1)) + C(4)*D(MINT(-1)) + C(5)*D(MRH(-1)) + C(6) *D(ERH(-1)) + C(7)

Observations: 38

R-squared	0.991466	Mean dependent var	0.635526
Adjusted R-squared	0.989814	S.D. dependent var	32.76937
S.E. of regression	3.307204	Sum squared resid	339.0655
Durbin-Watson stat	1.410420		

VAR Residual Heteroskedasticity Tests
(Levels and Squares)

Joint test:

Chi-sq	df	Prob.
289.4904	252	0.0523

VAR Residual Normality Tests
Orthogonalization: Cholesky (Lutkepohl)
Null Hypothesis: Residuals are multivariate normal

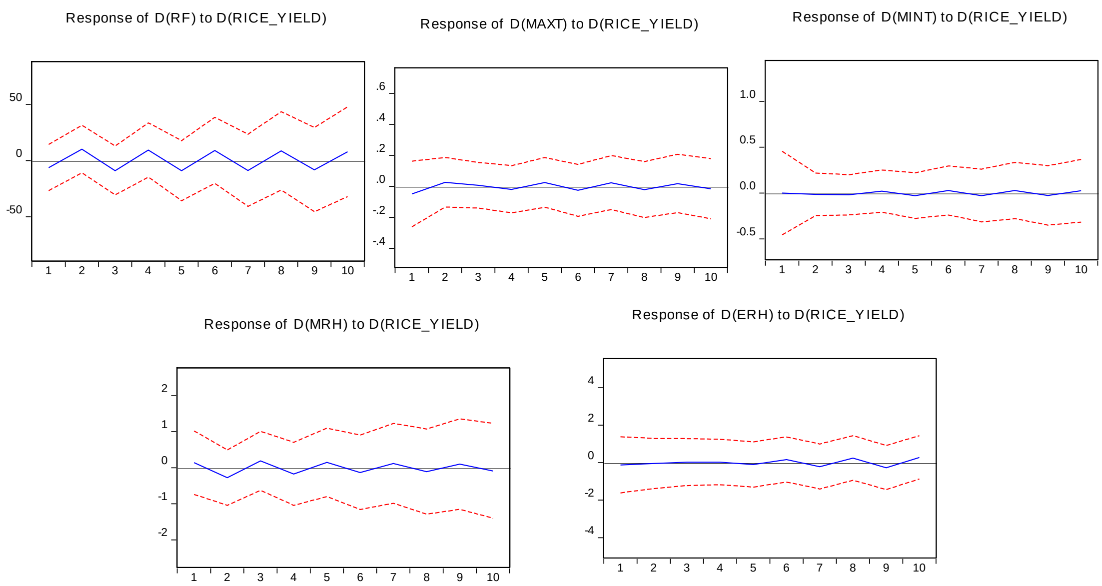
Component	Jarque-Bera	df	Prob.
1	168.0839	2	0.0000
2	0.786519	2	0.6749
3	0.570174	2	0.7519
4	12.25209	2	0.0022
5	6.273341	2	0.0434
6	18.21117	2	0.0001
Joint	206.1772	12	0.0000

VAR Residual Serial Correlation LM Tests

Null hypothesis: No serial correlation at lag h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	39.88062	36	0.3016	1.129124	(36, 90.6)	0.3164
2	53.04780	36	0.0333	1.603725	(36, 90.6)	0.0375

Response to Cholesky One S.D. (d.f. adjusted) Innovations ± 2 S.E.



APPENDIX(D)

D.I Pathein

Augmented Dickey-Fuller test

	t-Statistic	Prob.*
Null Hypothesis: RICEYIELD has a unit root	-1.982881	0.2934
Null Hypothesis: D(RICEYIELD) has a unit root	-5.143958	0.0001
Null Hypothesis: RF has a unit root	-4.002648	0.0029
Null Hypothesis: MAXT has a unit root	-2.973932	0.0436
Null Hypothesis: MINT has a unit root	-2.756801	0.0715
Null Hypothesis: D(MINT) has a unit root	-8.547572	0.0000
Null Hypothesis: MRH has a unit root	-1.702786	0.4240
Null Hypothesis: D(MRH) has a unit root	-4.890310	0.0002
Null Hypothesis: ERH has a unit root	-1.557336	0.4974
Null Hypothesis: D(ERH) has a unit root	-4.810144	0.0002

*MacKinnon (1996) one-sided p-values.

D.II Hinthada

Augmented Dickey-Fuller test

	t-Statistic	Prob.*
Null Hypothesis: RICEYIELD has a unit root	-1.982881	0.2934
Null Hypothesis: D(RICEYIELD) has a unit root	-5.908750	0.0000
Null Hypothesis: RF has a unit root	-3.460762	0.0130
Null Hypothesis: MAXT has a unit root	-0.729048	0.8301
Null Hypothesis: D(MAXT) has a unit root	-7.764704	0.0000
Null Hypothesis: MINT has a unit root	-4.779535	0.0002
Null Hypothesis: MRH has a unit root	-3.166552	0.0274
Null Hypothesis: ERH has a unit root	-1.887426	0.3357
Null Hypothesis: D(ERH) has a unit root	-44.91251	0.0001

*MacKinnon (1996) one-sided p-values.

D.III Maubin

Augmented Dickey-Fuller test

	t-Statistic	Prob.*
Null Hypothesis: RICEYIELD has a unit root	-1.704917	0.4233
Null Hypothesis: D(RICEYIELD) has a unit root	-99.06358	0.0001
Null Hypothesis: RF has a unit root	-2.994922	0.0417
Null Hypothesis: MAXT has a unit root	-1.714221	0.4185

Null Hypothesis: D(MAXT) has a unit root	-9.462166	0.0000
Null Hypothesis: MINT has a unit root	-3.023043	0.0389
Null Hypothesis: MRH has a unit root	-2.631906	0.0927
Null Hypothesis: D(MRH) has a unit root	-7.588361	0.0000
Null Hypothesis: ERH has a unit root	-1.652849	0.4493
Null Hypothesis: D(ERH) has a unit root	-41.86634	0.0001

*MacKinnon (1996) one-sided p-values.

D.IV Myaungmya

Augmented Dickey-Fuller test

	t-Statistic	Prob.*
Null Hypothesis: RICEYIELD has a unit root	-2.044925	0.2674
Null Hypothesis: D(RICEYIELD) has a unit root	-81.56483	0.0001
Null Hypothesis: RF has a unit root	-3.905243	0.0037
Null Hypothesis: MAXT has a unit root	-3.960942	0.0032
Null Hypothesis: MINT has a unit root	-1.925222	0.3186
Null Hypothesis: D(MINT) has a unit root	-5.525860	0.0000
Null Hypothesis: MRH has a unit root	-0.677727	0.8434
Null Hypothesis: D(MRH) has a unit root	-8.969934	0.0000
Null Hypothesis: ERH has a unit root	-1.214744	0.6618
Null Hypothesis: D(ERH) has a unit root	-8.233185	0.0000

*MacKinnon (1996) one-sided p-values.

D.V Phyarpon

Augmented Dickey-Fuller test

	t-Statistic	Prob.*
Null Hypothesis: RICEYIELD has a unit root	-2.332903	0.1673
Null Hypothesis: D(RICEYIELD) has a unit root	-116.8550	0.0001
Null Hypothesis: RF has a unit root	-2.376257	0.1555
Null Hypothesis: D(RF) has a unit root	-11.50188	0.0000
Null Hypothesis: MAXT has a unit root	-1.075761	0.7147
Null Hypothesis: D(MAXT) has a unit root	-6.586138	0.0000
Null Hypothesis: MINT has a unit root	-2.905257	0.0546
Null Hypothesis: D(MINT) has a unit root	-4.389679	0.0014
Null Hypothesis: MRH has a unit root	-2.704165	0.0826
Null Hypothesis: D(MRH) has a unit root	-54.62787	0.0001
Null Hypothesis: ERH has a unit root	-1.560451	0.4920
Null Hypothesis: D(ERH) has a unit root	-3.848257	0.0056

*MacKinnon (1996) one-sided p-values.