Project Report

On

FORECASTING RICE PRODUCTION IN ALAPPUZHA DISTRICT USING ADVANCED TIME SERIES MODELS

Submitted in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

In

APPLIED STATISTICS AND DATA ANALYTICS

By

MALAVIKA JAIN

(Reg No. SM23AS008)

(2023-2025)

Under the Supervision of

JESNA BABU

ST. TERESA'S COLLEGE
(AUTONOMOUS)
ESTD 1925

DEPARTMENT OF MATHEMATICS AND STATISTICS

ST TERESA'S COLLEGE(AUTONOMOUS)

ERNAKULAM, KOCHI – 682011

APRIL 2025

ST TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

CERTIFICATE

This is to certify that the dissertation entitled, **FORECASTING RICE PRODUCTION IN ALAPPUZHA DISTRICT USING ADVANCED TIME SERIES MODELS** is a bonafide record of the work done by **MALAVIKA JAIN** under my guidance as partial fulfilment of the award of the degree of **Master of Science in Applied Statistics and Data Analytics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date: 30/04/2025

Place: Ernakulam

Jesna

JESNA BABU
Assistant Professor,
Department of Mathematics and Statistics
St. Teresa's College (Autonomous)
Ernakulam.

Nisha Oommen
Assistant Professor & HOD
Department of Mathematics and Statistics
St. Teresa's College (Autonomous)
Ernakulam.

ST. TERESA'S COLLEGE (AUTONOMOUS) DEPARTMENT OF MATHEMATICS AND STATISTICS ERNAKULAM

External Examiners

1. Sangeetha Chandran 30.04.25

2. Anju N.B 30.04.25

DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of **JESNA BABU**, Assistant Professor, Department of Mathematics and Statistics, St. Teresa's College (Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam

Date: 30|04|2025

MALAVIKA JAIN

SM23AS008

ACKNOWLEDGEMENT

I take this opportunity to thank everyone who has encouraged and supported me to carry out this project.

I am very grateful to my project guide Jesna Babu for her immense help during the period of work.

In addition, I acknowledge with thanks to the Department for all the valuable support and guidance that has significantly contributed to the successful completion of this project.

I would also like to thank the HOD for her valuable suggestions and critical examinations of the project.

Ernakulam:

Date: 30/04/2025

MALAVIKA JAIN

SM23AS008

ABSTRACT

This research analyses rice production data over seasons for paddy in the Kuttanad region, Alappuzha, covering the period between June 1996 and February 2024, to estimate future production estimates. The prediction of future rice is also found to October 2026, by was using the SARIMA model, Holt-Winters Exponential Smoothing, and a Convolutional Neural Network (CNN) model for forecasting. A comparison of performances was made for the three models using RMSE and MAE, with CNN showing the most accurate results as the best candidate for forecasting.

ST.TERESA'S COLLEGE (AUTONOMOUS) ERNAKULAM

Certificate of Plagiarism Check for Dissertation

Author Name	MALAVIKA JAIN
Course of Study	M.Sc. Applied Statistics & Data Analytics
Name of Guide	Ms. Jesna Babu
Department	P.G. Dept. of Mathematics & Statistics
Acceptable Maximum Limit	20
Submitted By	library@teresas.ac.in
Paper Title	FORECASTING RICE PRODUCTION IN ALAPPUZHA DISTRICT USING ADVANCED TIME SERIES MODELS
Similarity	7% AI - 6%
Paper ID	3409024
Total Pages	46
Submission Date	2025-03-17 15:51:36

Signature of Student

Signature of Guide

Checked By
College Librarian

CONTENTS

CHAPTER 1
INTRODUCTION

Rice farming is a central part of the agricultural landscape of Kerala, India. Kerala has centuries-old paddy farming traditions, and historical documents point to the significance of rice in the food culture of the region. Kerala's geography and climate, its high monsoon rainfall, fertile alluvial soil, and many backwaters, are all conducive to paddy farming. Early farming techniques, such as the practice of using conventional organic farming techniques and native rice varieties, have largely influenced the state's agricultural past. Paddy farming in Kerala has seen major technological interventions. Studies on better seed varieties, water management strategies, and integrated pest management systems have helped ensure higher productivity. Research has investigated adoption levels of new agricultural technologies among farmers, stressing the role of extension services and farmer training. Ecological sustainability of rice production has been of interest in literature. Rice farming covers 7.46% of the entire cropped area in the state. Palakkad, Alappuzha, Thrissur and Kottayam constitute the top rice producers in Kerala.

This project addresses analysis of rice cultivation in Kuttanad, a distinct low-lying area covering the Alappuzha.

Kuttanad, also known as the "Rice Bowl of Kerala," is a scenic area in the southern Indian state of Kerala. Famous for its large area of paddy fields, linked together by a network of peaceful backwaters and canals, Kuttanad is an important centre for farming activities. The low-lying topography of the region, with its closeness to sea level, requires innovative farming methods like paddy fields below sea level. It is the only place in India where rice is grown below sea level, facilitated by a complex system of bunds, canals, and dikes that shield the fields from inundation In the annual years some studies were made regarding the problems and possibilities of paddy cultivation in Ramankari village's Kuttanad area.

By 1980-81 the share of the state in paddy area and rice production reduced to 2 percent and 2.37 percent respectively. According to the study, farmers are experiencing a range of issues such as the unavailability of needed number of labourers during the crop season, the falling profitability of the crop, militant trade unionism, tardy pace mechanization, non-availability of easy credit and adequate marketing facilities, frequent failures of crops, uneconomic size of holdings and so on (Thomas,2002). Over the last decade one of the major factors which affected the production of rice was the flood in 2018. The people of Kerala have suffered the worst natural calamity, deluge from 15th august 2018, and this continued for a week.

The effect of the flood had curled the life of farmers; the sudden wave had swept away their crops The primary issues encountered by farmers are the lack of support, decrease in production and absence of pure water, delay in receiving support, repayment of the loan, etc. It provides that production prior to a wave is not like production subsequent to a flood (Santhi and Veerakumaran, 2019) The region, production and productivity of paddy crop in India during 2009-10 to 2018-19 reveal rising trend year by year. Years 2010-11 and 2011-12 reveal a sudden boost in the production and productivity of paddy in India, which then reveals a declining trend during the coming years and the trend regained from 2016-17 (Thomas,2020)

The agronomy industry had been ruined by the Kerala flood of 2018, particularly rice farming involving a huge economic loss.

Climate-resilient agriculture system is required to compete with the disasters and their effects. The current study intends to unveil the significance of increasing Pokkali rice farming across the state's coastal wetland regions to develop a flood-resilient agriculture system (Baidya et.al, 2023). According to the study, lack of pre and post-flood measures impacts the adaptive capacity of the residents. Due to the severity of flood risk, channelised policy actions that are archetypal in nature need to be recommended to enhance their resilience and adaptive capacity (Varughese and Mathew, 2023). Agriculturists of the Kuttanad region follow the puncha cultivation system, where a single or two crops are grown every year based on the water supply and monsoon trends. The area generates approximately 37% of the state's rice production by estimate (Kerala state economic and statistic department).

1.1 Objective of the study

1. To model and forecast the production of rice production using SARIMA model
2. To model and forecast of rice production by applying the Holt-Winter's method.
3. To forecast the future production of rice using the CNN Model.
4. To compare the forecast by SARIMA, Holt Winter's and CNN and to find the best fitted model.

CHAPTER 2
REVIEW OF LITERATURE

This chapter illustrates the outcome from the studies associated with this analysis that compared the different Rice production dataset and the prediction and interpretation with the aid of different statistical analysis, machine learning algorithms etc. A literature review is the process of writing up a summary, synthesis and critique of the literature that was found as a result of a literature search. It can be utilized as background or context for an original research project.

Thomas (2002) conducted a case study related to problems and prospects of paddy cultivation in Kuttanad region A case study of Ramankari village in Kuttanad taluk its objective is a development-oriented village level study. The major thrust of the study is on the revival and development of paddy farming sector in Kuttanad region. With this focus the study aims a) To examine the salient features in the transformation of the rice economy of Kuttanad region, b) To examine the socio-economic profile of paddy farmers in the study area, c)To analyse the changes in paddy farming operations in the study area, d) To assess the economic viability of paddy cultivation in its present level, e) To identify the current problems in paddy farming and f) To suggest appropriate measures for the development of paddy farming sector in the study area

Hamjah (2014) carried out research on forecasting rice production in Bangladesh using the Box-Jenkins ARIMA model. The research was focused on identifying the optimal ARIMA models in rice yield forecasting in different seasons (Aus, Boro, and Aman) using data from the Bangladesh Agricultural Ministry. ARIMA (2,1,2) was optimal for total production and Aman and ARIMA (1,1,3) was optimal for Boro. The difference between actual and estimated values confirmed the models' validity as the optimal models to use in short-term forecasting.

Santhi and Veerakumaran (2019) examined the effect of the 2018 Kerala flood on agriculture in Edathua Panchayat, Kuttanad Taluk, Alappuzha district. Heavy rainfall (164% above normal) and dam spill-overs led to the flood, which had adverse effects on farmers through the destruction of crops, houses, livestock, and farm equipment. Through interviews with 45 affected farmers, the study established important issues such as lower yields, lack of support, shortage of water, problems in repaying loans, and delayed assistance. Despite the government's provision of immediate relief, long-term assistance was considered necessary to prevent farmers from taking loans and leaving agriculture. The study was to examine the effect of the flood on agriculture, livestock, and financial and technical assistance required for rehabilitation.

Raj et al. (2020) carried out a study on the issues of paddy farmers in Upper Kuttanad, Alappuzha district, Kerala. Alappuzha is the second largest state in rice cultivation area, production, and productivity. The Kuttanad area, the "Rice Bowl of Kerala," is the only place in the world to cultivate rice below mean sea level. In order to evaluate the constraints of the farmers, a survey was carried out in Upper Kuttanad, and primary data were gathered from 25 paddy farmers in saltwater-affected fields and 25 in unaffected fields in the Haripad block. Based on Henry Garrett's ranking method, the study identified that the most critical problem for farmers in unaffected fields was weed infestation, followed by the non-availability of hired labour. For saltwater-affected farmers, however, the biggest problem was saltwater intrusion because of which paddy quality decreased due to salinity.

Thomas (2020) examined the trends in area, production, and productivity of paddy cultivation in Kerala's Kuttanad, also referred to as the "Rice Bowl of Kerala" and the largest wetland ecosystem on India's west coast. The analysis spanned 2009-10 to 2018-19 data, which indicated a general rising trend in paddy cultivation in India, with a sharp rise in 2010-12, followed by a decline and recovery from 2016-17. Although Kerala experienced a decline in absolute area and production, its average decadal growth trend experienced a marginal rise. In 2018-19, Kerala's paddy productivity (2,920 kg/hectare) was higher than the all-India average (2,659 kg/hectare), with Kuttanad recording record production. There has been renewed interest in paddy cultivation in recent years, particularly in Kuttanad, with the help of policy interventions by central, state, and local governments. The study aimed to assess trends in paddy cultivation in India, Kerala, and the Kuttanad region.

John (2020) compared Holt-Winters and Seasonal ARIMA (SARIMA) models in predicting quarterly rice and corn production in the Philippines. Rice and corn are the lifeline of the economy and food security of the country, yet the agriculture sector is faced with climate variability, land limitation, and reliance on imports to meet demand. Good decision-making and resource allocation require accurate production forecasting. Philippine Statistics Authority quarterly production figures from 1987 to 2023 were used in the study. Results showed that the Holt-Winters additive seasonality model outperformed the SARIMA model, with smaller Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) values. Results can be utilized by planners, agriculture stakeholders, and commodity traders in forecasting import requirements. Under volatility in global food prices and exchange rates, accurate forecasting can assist in managing import costs and food security. Applying the Holt-Winters model and enhancing forecasting techniques will allow the Philippines to achieve greater food sufficiency and propel rural economic growth. The study identified the necessity for credible forecasting models in the delivery of stable rice and corn supplies to support sustainable agricultural growth. Agricultural forecasting studies continue to play a critical role in managing evolving challenges and predictive accuracy.

Baidya et al. (2023) investigated the flood resilience of Pokkali rice with specific reference to the 2018 Kerala flood. The research brought to the limelight the rising frequency of climate change-related natural disasters, with Kerala witnessing its worst monsoon in a century. Heavy rainfall, runoff excess, low reservoir storage, inadequate drainage, constricted sea outlets, and high spring tides were some of the contributing factors that worsened the flood calamity. The extreme meteorological and hydrological conditions caused extensive loss to the state, especially in the agricultural sector, which is a major component of Kerala's economy.Rice cultivation was economically ruined, emphasizing the significance of climate-resilient agriculture to withstand such catastrophes. Resilience refers to the ability to survive and recover from adverse situations without losing stability. Pokkali, a specific rice crop cultivated in Kerala's coastal wetlands, has demonstrated an exceptional capacity to withstand floods—unlike any other rice crop. Thus, Pokkali can potentially play a vital role in ensuring food security during catastrophes, particularly floods, and serve as a Nature-Based Solution for flood-resilient agriculture. Further, large-scale cultivation of Pokkali in other suitable coastal and wetland areas can enhance flood resilience and develop agricultural sustainability.

Varughese and Mathew (2023) analysed the impact of climate change, repeated flooding, and consequent mass migration in Kerala's wetland ecosystem, Kuttanad. Despite the immense ecological and economic value of wetlands, its inhabitants are confronted with mounting climatic extremes. Unpredictable rains in Kuttanad are particularly felt through repeated flooding. This study analyses the vulnerability of the inhabitants to mounting threats of flooding, climate change migration triggers, and coping capacity. Drawing from a primary survey-based research in a vulnerability and adaptive capacity framework, the research captures major challenges such as loss of residences, sinking houses, heightened health risks, and loss of livelihood, all of which have resulted in heightened flood vulnerability. Heightened vulnerability has led to large-scale migration, and Probit regression analysis reveals the influence of socio-economic determinants in shaping the decision of households to migrate. Marginalized social groups and environmentally dependent households are most at risk. The research also suggests that the lack of appropriate pre-flood and post-flood measures compromises the capacity of residents to adapt. With the intensity of flood hazards, the research suggests the need for well-designed policy interventions in strengthening resilience and adaptive capacity in the area.

CHAPTER 3
MATERIALS AND METHODS

3.1 Data Collection

The dataset used for this study consist of the season wise data of rice production collected from the official website of Kerala State Economics and Statistics Department and also from Directorate of Economics and statistics Department of Agriculture and farmers Welfare Ministry Government of India. The data contains of seasonal production of rice in Kerala, from June 1995 to February 2023

3.2 Methodology

The first important step in the analysis was to study the data in detail. It outlines the data collection methods, processing techniques, and analytical models applied to achieve the study's objectives. The objective of the study is to understand the patterns and trends of rice production in Kuttanad region and to predict the future production by using various statistical techniques and machine learning algorithms. This was conducted as a step-by-step procedure by analysing the data and understands the characteristics of the dataset. Then the model was forecasted using SARIMA. Then the model was forecasted using the Holt-Winter's model. Then it was also forecasted using convolution neural network. Then the comparison of the three models is done by calculating the MAE and RMSE values of both models.

3.3 Tools for Analysing and Forecasting

1. Seasonal ARIMA (Seasonal Autoregressive integrated moving average) model.
2. Holt-Winters method
3. Convolution Neural Network (CNN)

3.4 Tools for Comparison

1. Mean Absolute Error (MAE)

The formula for MAE is:

$$MAE = \frac{1}{n} \sum_{i=0}^{n} |y_i - \hat{y_i}|$$

Where,

n is total number of observations

y_i is the actual value

$\hat{y_i}$ is the predicted value

| y_i - $\hat{y_i}$| is the absolute error for each observation

2. Root Mean Square Error (RMSE)

The formula for RMSE is:

$$RMSE = \sqrt{\frac{\Sigma(y_{i-}\widehat{y_{i)}}}{N-P}}$$

Where,

y_i is the actual value for the i th observation

$\hat{y_i}$ is the predicted value for the i th observation

N is the number of observations

P is the number of parameter estimates, including the constant

3.5 Python

In this research we take programming language as python. Python is an interpreted high-level programming language that is most famous for simplicity and readability. Python is frequently used in web development, artificial intelligence, data science, automation, etc. It can even be utilized for performing the EDA to observe the concealed pattern of the dataset and to make predictions using SARIMA, Holt-Winter's model and CNN

CHAPTER 4

EXPLORATORY DATA ANALYSIS

Exploratory Data analysis (EDA) is the first step of data analysis process, it helps to understand the underlying structure and pattern of data. It also helps to understand the relationships within dataset. EDA can be also used to understand the quality of data, check null values, missing values etc.

Using EDA after checking for null and missing values, next step is to summarize and visualize the data to understand it more. This may include calculating summary statistics such as mean, median, mode and standard deviation etc. Visualizing the data, decomposing the data to trend, seasonal and residual also include in this.

4.1 Descriptive Statistics:

Descriptive statistics describe the variability, distribution and central tendency of the dataset by summarizing the mean, median, mode, standard deviation and quartiles. Measures like standard deviation and quartiles are used to identify outliers and extreme values of the dataset. Overall descriptive statistics is a complete summary of data's main characteristics, which will help to understand data more and do further analysis accurately.

Table 1

count	84
mean	32025.372012
std	19905.579864
min	1374.000000
25%	18709.500000
50%	27636.500000
75%	40995.250000
max	89983.768000

4.2 Time Series Plot

A Time Series Plot is a graphical representation of data points collected over time, helping to analyze trends, patterns, and fluctuations. It consists of a time variable on the x-axis and the measured data on the y-axis. This type of plot is widely used in various fields such as economics, finance, weather forecasting, and agriculture to understand how a variable changes over time. By visualizing time series data, one can identify trends, seasonal variations, and anomalies that might not be apparent in raw numerical data. A time series plot is essential for making predictions and analyzing historical patterns, making it a crucial tool for statistical and machine learning models like ARIMA.

4.3 Seasonal Decomposition:

Using seasonal decomposition technique time series data is decomposed in to its components trend, seasonal and residual. Using this we can understand whether the dataset shows seasonality and also understand the trend of the data. Understanding the trend of the data and seasonality will helps to choose correct model and increase accuracy of the forecasting

Trend: A trend refers to the long-term increase or decrease in data over time.

Seasonality: In time series data, seasonality indicates a recurring or predictable pattern that repeats over a short period, such as days, weeks, or months. An example of this is rainfall patterns.

Residual: Residuals represent the difference between the observed and predicted values.

CHAPTER 5
TIME SERIES

Time series analysis is the analysis statistical data collected according to the time of occurrence. It allows for the identification of trends, patterns, and fluctuations. The data in a time series is typically collected at regular intervals, such as hourly, daily, monthly, or yearly, making it essential for forecasting and predictive analysis. Time series data may contain noise or irregular variations caused by unpredictable events. Time series analysis is widely used in various fields such as finance, economics, climate science, healthcare, and agriculture. Visualization techniques such as time series plots are commonly used to understand data trends, detect anomalies, and make data-driven decisions.

5.1 Components of Time series

The components of a time series describe the different patterns and variations observed in time-dependent data. They are as follows:

1. The trend indicates the overall direction of the data to rise or fall over a long span of time. A trend is a smooth, overall, long-term, average direction. It is not always the case that the rise or fall is in the same direction over the entire period of time. It can be linear and non-linear trend.
2. Seasonality is periodic and recurring patterns that happen on a regular basis, like every day, month, or year
3. Cyclic variations refer to the long-run fluctuations which are not of fixed duration like seasonality but result from economic, environmental, or social cycles. It is a four-stage cycle that includes the stages of prosperity, recession, depression, and recovery. The cyclic variation can be regular are not periodic.
4. Irregular or random variations are unforeseen fluctuations that occur because of unexpected events. They are not regular variations and are entirely random or irregular. These fluctuations are unexpected, uncontrollable, unforeseen, and are random. These forces are earthquake, wars, flood, famine, and any other calamities

5.2 Mathematical Models for Time Series Analysis

In time series analysis, it is assumed that there exist two models commonly for the decompositions of a time series into its components.

Additive Model: According to the additive model the decomposition of time series is done on the assumption that the effects of various components are additive or in other words,

$$Y_t = T_t + S_t + C_t + I_t$$

Where Y_t is the time series value and T_t , S_t, C_t and I_t stands for trend, seasonal variations, cyclical variations and irregular variations respectively. In this model S_t, C_t and I_t are absolute quantities and can have positive or negative values. The model assumes that four components of the time series are independent of each other and none has any effect on the remaining three components.

Multiplicative Model: According to the multiplicative model the decomposition of a time series is done on the assumption that the effects of the four components of the time series are not independent of each other. According to the multiplicative model,

$$Y_t = T_t \text{ x } S_t \text{ x } C_t \text{ x } I_t$$

In this model S_t, C_t and I_t are not absolute amounts as in the case of the additive model. There are relative variations and are expressed as rates or indices fluctuating above or below unity. The multiplicative model can be expressed in terms of the logarithm.

$$\log Y_t = \log T_t + \log S_t + \log C_t + \log I_t$$

5.3 Time Series Modelling

5.3.1 Basic Definitions

(i) Autocorrelation Function (ACF)

The autocorrelation function of a stationary time series $\{Z_t\}$, ρ(k) at lag k is defined as the correlation at lag k between Z_t and Z_{t+k}. Thus, the autocorrelation function at lag k is given by

$$\rho(k) = \gamma(k)/\gamma(0)$$

where $\gamma(k) = \mathrm{Cov}(Z_t, Z_{t+k})$

(ii) Partial Autocorrelation Function (PACF)

The partial autocorrelation function, like the autocorrelation function, conveys vital information regarding the dependence structure of a stationary of lag k can be thought of as the partial regression coefficients kk in the representation.

$$X_t = \varphi_{k1} X_{t-1} + \varphi_{k2} X_{t-2} + \ldots + \varphi_{kk} X_{t-k} + \varepsilon_t$$

Thus, the partial autocorrelation at lag k, φ_{kk}, measures the correlation between X_t and X_{t-k} after adjusting for the effects of ($X_{t-1}, X_{t-2} \ldots X_{t-k+1}$).

(iii) Autoregressive (AR) process

A time series $\{Z_t\}$ is said to be an autoregressive process of order p, abbreviated as AR(p) if it is a weighted linear sum of the past p values plus a random shock so that,

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \ldots + \varphi_p Z_{t-p} + \varepsilon_t$$

where $\{\varepsilon_t\}$ denotes a purely random process with 0 mean and constant variance σ^2. Using the backward shift operator B, such that $BZ_t = Z_{t-1}$, the AR (p) model may be written more succinctly in the form $\varphi(B) Z_t = \varepsilon_t$ Where $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 + \ldots - \varphi_p B^p$ is a polynomial in B of order p.

(iv) Moving Average (MA)processes

A time series $\{Z_t\}$ is said to be a moving average process of order q, abbreviated as MA(q) if it is a weighted linear sum of the last q random shocks so that $Z_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_p \varepsilon_{t-p}$, where $\{\varepsilon_t\}$ denotes a purely random process with 0 mean and constant variance σ^2 . Using the backward shift operator B the MA (q) model may be written in the form,

$$Z_t = \theta(B)\, \varepsilon_t ,$$

where $\theta(B) = 1 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_p \varepsilon_{t-p}$ is a polynomial in B of order q.

(v) Stationarity

Stationarity is a crucial concept in time series analysis, especially when working with ARIMA models. A time series is stationary if its statistical properties remain constant over time. This means the series does not exhibit trends or seasonal patterns. There is a commonsense reason for requiring stationarity: we could not get useful estimates of the parameters of a process otherwise. If p=0, we have either a pure MA model or a white noise series. All pure MA models and white noise are stationary, so there are no stationarity conditions to check.

For an AR (1) or ARMA (1, q) process, the stationary requirement is that the absolute value of φ_1 must be less than one: $|\varphi_1| < 1$.
For an AR (2) or ARMA (2, q) process, the stationary requirement is a set of three conditions: $|\varphi_2| < 1$, and $\varphi_1 - \varphi_2 < 1$.

(vi) Invertibility

Invertibility is a property of the Moving Average (MA) component in SARIMA models. A model is invertible if its MA process can be rewritten as an equivalent Autoregressive (AR) process with stable coefficients. This ensures that past values can be reconstructed without exploding errors. There is a commonsense reason for invertibility: a non-invertible ARIMA model implies that the weights placed on past Z observations do not decline as we move further into the past, but commonsense says that larger weights should be attached to more recent observations. Invertibility ensures that these results hold.

If q=0, we have either a pure AR process or a white noise series. All pure AR processes and white noise are invertible and no further checks are required. For an MA (1) or ARMA (p, 1) process, invertibility requires that the absolute value of Θ_1 must be less than one, $|\Theta_1| < 1$, $\Theta_1 + \Theta_2 < 1$, and $\Theta_1 - \Theta_2 < 1$.

(vii) ARMA Process

An Autoregressive Moving Average (ARMA) process is a time series model that combines autoregressive (AR) and moving average (MA) components to describe the behaviour of a variable over time. Mathematically, an ARMA (p, q) process is defined as

$$X_t = c + \sum_{i=1}^{p} \varphi_1 \, X_{t-i} + \sum_{j=1}^{q} \Theta_j \varepsilon_{t-j} + \varepsilon_t$$

where X_t represents the time series at time t, c is a constant, φ_i are the autoregressive coefficients that capture dependencies on past values, and Θ_j are the moving average coefficients that account for past errors. The white noise error term ε_t has a mean of zero and constant variance. The parameter p determines the number of lagged values included in the AR component, while q determines the number of past error terms in the MA component. If q=0, the model simplifies to an AR(p) process, and if p=0, it becomes an MA(q) process. By incorporating both AR and MA terms, the ARMA model effectively captures patterns of persistence and random fluctuations in time series data, making it a powerful tool for forecasting.

(viii) ARIMA and SARIMA

In practice, many time series exhibit non-stationary behaviour, making it necessary to transform them before applying standard AR, MA, or ARMA models. One common approach to achieve stationarity is differencing, where the first difference is given by, $Z_t - Z_{t-1} = (1 - B)Z_t$. If required, this process can be repeated, leading to second difference and beyond. In general, differenced d differences to a series can be expressed as $((1 - B)^d Z_t$. When a time series is differenced d times before fitting an ARMA (p, q) model, the resulting model is known as an Autoregressive Integrated Moving Average (ARIMA) process, denoted as ARIMA (p, d, q). Here, the letter "I" stands for Integrated, referring to the differencing process, and d indicates the number of differences applied. Mathematically, an ARIMA model is represented as:

$$\varphi(\mathrm{B})((1-B)^d Z_t = \Theta(B)\,\varepsilon_t$$

where $\{\varepsilon_t\}$ represents a white noise process with mean zero and constant variance σ^2 . In addition to autoregressive and moving average components, ARIMA models may also include a constant term

When a time series exhibits seasonality, it follows a repeating pattern every S periods. For example, in monthly data, a seasonal cycle often spans 12 months (S=12). A Seasonal ARIMA (SARIMA) model extends the ARIMA framework by incorporating seasonal autoregressive (SAR) and seasonal moving average (SMA) terms. These seasonal terms capture dependencies at lags that are multiples of S. The seasonal ARIMA model, denoted as ARIMA (p, d, q) × (P, D, Q) S, combines non-seasonal and seasonal components in a multiplicative form:

$$\phi(B^s)\,\varphi(\mathrm{B})(1-)^D(1-B^s)^d\,X_t = \mathit{\Theta}(B^s)\,\Theta(B)\,Z_t$$

where,

$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p$ represents the non-seasonal AR component.

$\Theta(B) = 1 - \Theta_1 B - \Theta_2 B^2 - \cdots - \Theta_q B^q$represents the non-seasonal MA component

$\phi(B^s) = 1 - \phi_1 B^s - \phi_2 B^{2s} - \cdots - \phi_p B^{2p}$ represents the seasonal AR component.

$\mathit{\Theta}(B^s) = 1 - \mathit{\Theta}_1 B^s - \mathit{\Theta}_2 B^{2s} - \cdots - \mathit{\Theta}_Q B^{Qs}$ represents the seasonal AR component.

$\{Z_t\}$ is a white noise process with zero mean and constant variance σ^2

p, d, q are the non-seasonal AR order, differencing order, and MA order, respectively.

P, D, Qare the seasonal AR order, differencing order, and MA order, respectively.

S represents the number of periods in a full seasonal cycle.

(ix) Diagnostic Checking

An equally important step in the time series modelling process is diagnostic checking that evaluates how well a fitted model accounts for the observed data prior to using it for prediction. After estimation of an ARIMA or SARIMA model, diagnostic checks are completed so that the residuals are found to be white noise, or it means that should be independently and having the same distribution with mean zero and constant variance. This usually entails inspecting residual plots, performing nonetheless statistical assessments. Patterns, autocorrelation or heteroscedasticity in the residuals imply that the model may not have been appropriately specified, and adjustments including changing the model order, or adding new terms may be helpful.

(x) Akaike Information Criterion (AIC)

Akaike information criterion (AIC) indicates the goodness of fit of time series models with respect to model fit. It is defined as:

$$AIC = -2 \ln (L) + 2k$$

where represents the likelihood of the model given the data, k is the number of estimated parameters in the model. A lower AIC value indicates a better balance between model fit and complexity, helping to prevent overfitting. When selecting the best ARIMA or SARIMA model, different candidate models are compared, and the one with the lowest AIC is preferred. However, AIC should not be used in isolation; other diagnostic checks are also necessary to confirm model adequacy

5.4 Statistical Test

5.4.1 Augmented Dickey-Fuller (ADF) test

One of the methods used to test the stationarity of a time series is the Unit Root Test, an example of which is the Augmented Dickey-Fuller (ADF) test. The Augmented Dickey-Fuller test is used to test the null hypothesis whether the given time series is non-stationary or stationary. This is one of the frequently used statistical tests for doing stationarity analysis on time series data.

Null Hypothesis (H_0): The time series is non-stationary

Alternative Hypothesis (H_1): The time series is stationary

If the p-value calculated from the test is less than a selected significance level (e.g., α= 0.05), thus rejecting the null hypothesis.

5.5 Seasonal Auto Regressive Integrated Moving Average (SARIMA)

A Seasonal Autoregressive Integrated Moving Average (SARIMA) model is an extension of the ARIMA (Autoregressive Integrated Moving Average) model that includes seasonal components. SARIMA models are used when time series data exhibits seasonal patterns or variations. SARIMA models are widely used for forecasting The SARIMA model is denoted by SARIMA (p, d, q) (P, D, Q, s), where:

p: Autoregressive order for non-seasonal components.

d: Degree of differencing for non-seasonal components.

q: Moving average order for non-seasonal components.

P: Seasonal autoregressive order.

D: Degree of differencing for seasonal components.

Q: Seasonal moving average order.

s: Seasonal period length

5.6 Holt-Winter's Model

Holt Winter's method can be used to forecast the future values by analysing the trend and patterns that occur in the dataset. The Holt method has two version, additive and multiplicative, which depend on the characteristic of the dataset. Let $L_{t,} T_t, I_t$ be the local level, trend and seasonal index respectively at time t. The dataset considered is additive in nature thus we have,

Its interpretation will be based on whether seasonality is assumed to be additive or multiplicative. In the additive scenario i, y_t-I_t is the depersonalized value, whereas in the multiplicative category it is y_t / I_t. The levels of the 3 quantities L_t, T_t, and I_t, all have to be estimated and so we require 3 updating equations with three smoothing parameters, say ∝, γ and δ. As with the smoothing parameters typically being in the range (0, 1) beforehand. The shape of the updating equations is again intuitively plausible.

Suppose the seasonal variation is multiplicative. Then the (recurrence form) equations for updating L_t, T_t, I_t, when a new observation y_t becomes available are

$$L_t = \alpha\left(\frac{y_t}{I_{t-s}}\right) + (1-\alpha)(L_t + T_{t-1})$$

$$T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)T_{t-1}$$

$$I_t = \delta(y_t/L_t) + (1-\delta)I_{t-s}$$

and the forecasts from time (t) are then,

$$y_{t+h} = (L_t + hT_t)I_{t-s+h}$$

for h= 1,2,…,s. If the seasonal variation is additive, the equations for updating L_t, T_t, I_t, when a new observation y_t becomes available are

$$L_t = \alpha(y_t - I_{t-s}) + (1-\alpha)(L_{t-1} + T_{t-1})$$

$$T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)T_{t-1}$$

$$I_t = \delta(y_t - L_t) + (1-\delta)I_{t-s}$$

and the forecasts from time (t) are then

$$y_{t+h} = L_t + hT_t + I_{t-s+h}$$

For starting values, it seems sensible to set the level component L0, equal to the average observation in the first year. i.e.,

$$L_0 = \sum\nolimits_{t=1}^{s} y_t$$

where s is the number of seasons. The starting values for the slope component can be taken from the average difference per period between the first- and second-year averages. i.e.,

$$T_0 = \left(\frac{\sum_{t=s+1}^{2s} y_t}{s} - \frac{\sum_{t=1}^{s} y_t}{s}\right)/s$$

The seasonal index can be calculated after allowing for trend adjustment as:

$$I_0 = \frac{y_k - (k-1)T_0}{2}$$

$$I_0 = \frac{y_k - (L_0 + (k-1))T_0}{2}$$

Where k= 1,…,s. This will lead to (s) separate values for I_0, which is what is required to gain the initial seasonal pattern.

5.7 Convolution Neural Network

A Convolutional Neural Network (CNN) is a deep learning model that is usually applied to image processing but is also applicable to time-series forecasting, such as the forecasting of rice production. While classical forecasting relies on simple patterns, CNNs are capable of extracting intricate relationships between sequential data. When applied to rice production data within a quarter, CNNs recognize seasonality and relationships through processing the dataset as a structured matrix. The forecast process starts by preparing the data, where the dataset is arranged in a two-dimensional format, with years in rows and quarters in columns. Normalization is done in order to increase model efficiency. Convolution layers are then used, with small filters that move across the data to identify trends and patterns over various time frames. In order to further reduce the extracted features, pooling layers are used to decrease dimensionality while maintaining key information. The data is flattened into a one-dimensional array from the processed data before being input into fully connected layers, which project the learned patterns onto future production values. The model goes through a training process, in which it tunes network weights from historical data in order to minimize prediction errors. After training, the CNN can be employed to feed in current quarterly data and output predictions for future rice yields, utilizing the intricate patterns it has acquired.

CHAPTER 6

RESULTS AND DISCUSSIONS

This chapter deals with a comparative study of time series modelling and forecast of quarterly production of rice in Kuttanad a place in Alappuzha district of Kerala using the SARIMA model, Holt-Winter forecasting procedure and CNN model. The dataset consists of 84 observations starting from June 1996 to February 2024.

6.1Modelling of rainfall using SARIMA model

6.1.1 Descriptive Statistics

Table 1

count	84
mean	32025.372012
std	19905.579864
min	1374.000000
25%	18709.500000
50%	27636.500000
75%	40995.250000
max	89983.768000

6.1.2 Time series plot

Figure 1 represents the time series plot of quarterly rice production data. It is visible that there is a seasonality in the data.

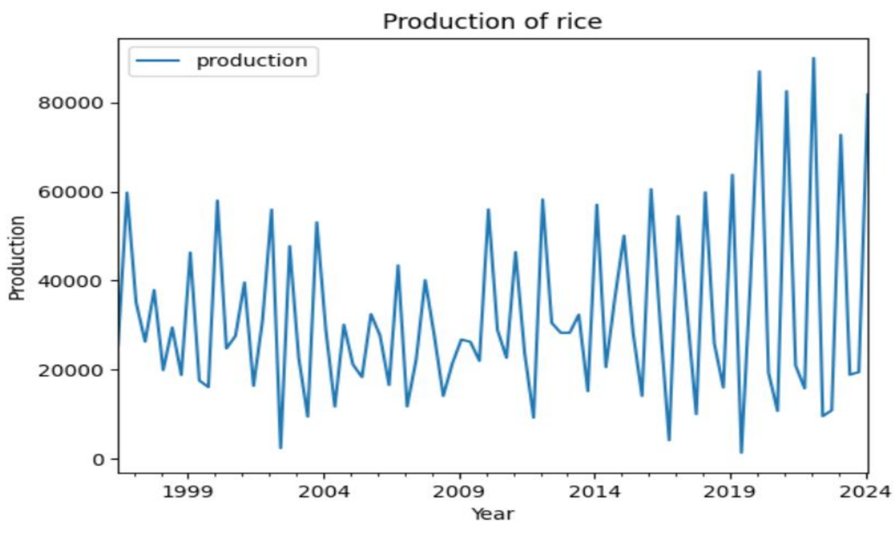

Figure 1 Time series plot

6.1.3 Decomposition of Time

The second step is to perform the decomposition to capture the trend,seasonal and random components of time series . Figure 2 represent the plot,

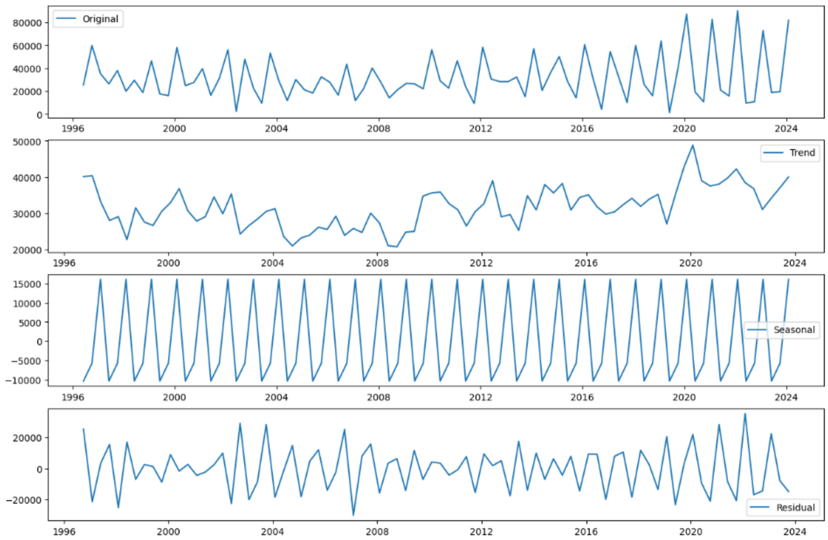

Figure 2: Decomposition plot

From the Figure 2, it can be concluded that the data has seasonality.

6.1.4 Augmented Dickey-Fuller for checking stationarity

This step is to examine the stationary condition of the dataset by using Augmented Dickey-Fuller test also known as ADF test. This test follows a hypothesis testing approach.

The null hypothesis Ho is given by

Ho: The data is stationary

The alternative hypothesis is given by

H_1: The data is not stationary

The outcome achieved,

ADF test statistics = -1.9527362729599802

p-value = 0.30767305630819997, which is greater than 0.05, so accept the null hypothesis. i.e.; the data is not stationary.

Then apply seasonal differencing to the dataset, the seasonal patterns have been removed, making the data more stationary and suitable for time series modelling. After that it can perform the n order differencing until the time series is stationary. Then perform differencing with n = 1, now again check the stationarity using ADF test.

Here, tests the hypothesis

Ho: The data is stationary

Against

H_1: The data is not stationary

ADF test statistics = -4.380504451510238, p-value = 0.00032132326608105014. The ADF test gives the p-value which is smaller than 0.05, so fail to accept Ho and Thus, conclude that the data is stationary. Fig 3 gives the differenced seasonal production

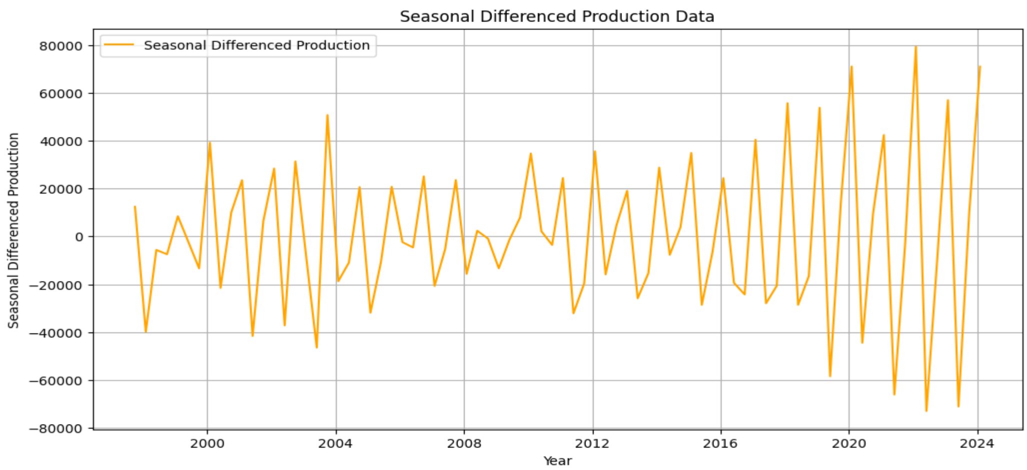

Figure 3: Seasonal Differenced Production data

6.1.4.1 ACF Plot of Seasonally Differenced Data

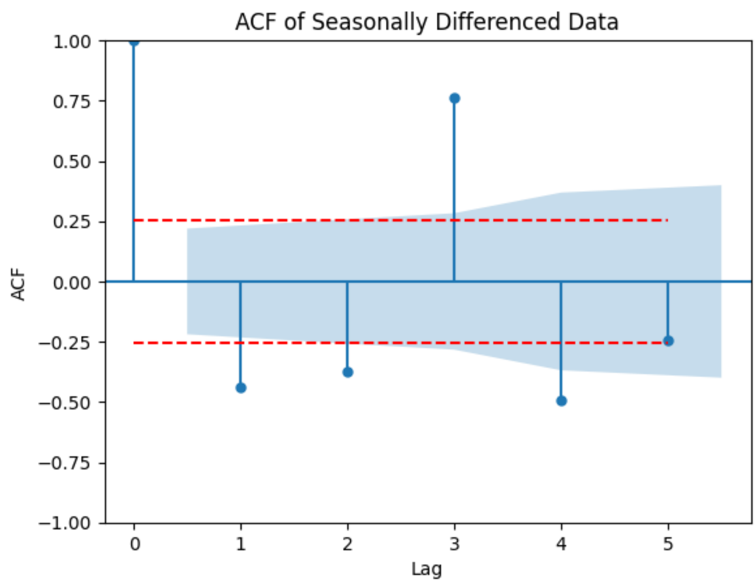

Figure 4: ACF Of Seasonally Differenced Data

6.1.4.2 PACF of Seasonally Differenced Data

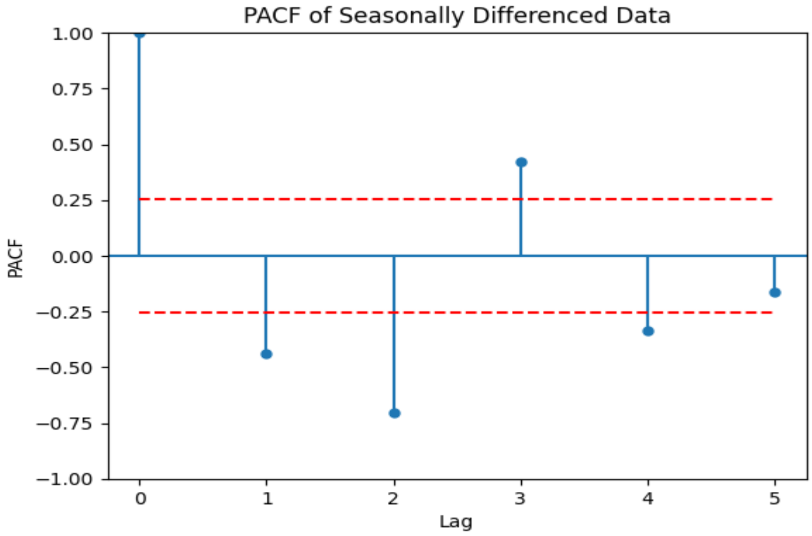

Figure 5:PACF of Seasonally Differenced Data

Since the p-value is less than 0.05 , its clear that no more seasonal differencing is needed. Now D = 1 and d = 0 , to find the seasonal AR order to (P) and the seasonal MA order(Q) have to plot the ACF and PACF of stationary data at seasonal lags.

6.1.4.3 ACF of Stationary Data at seasonal lags

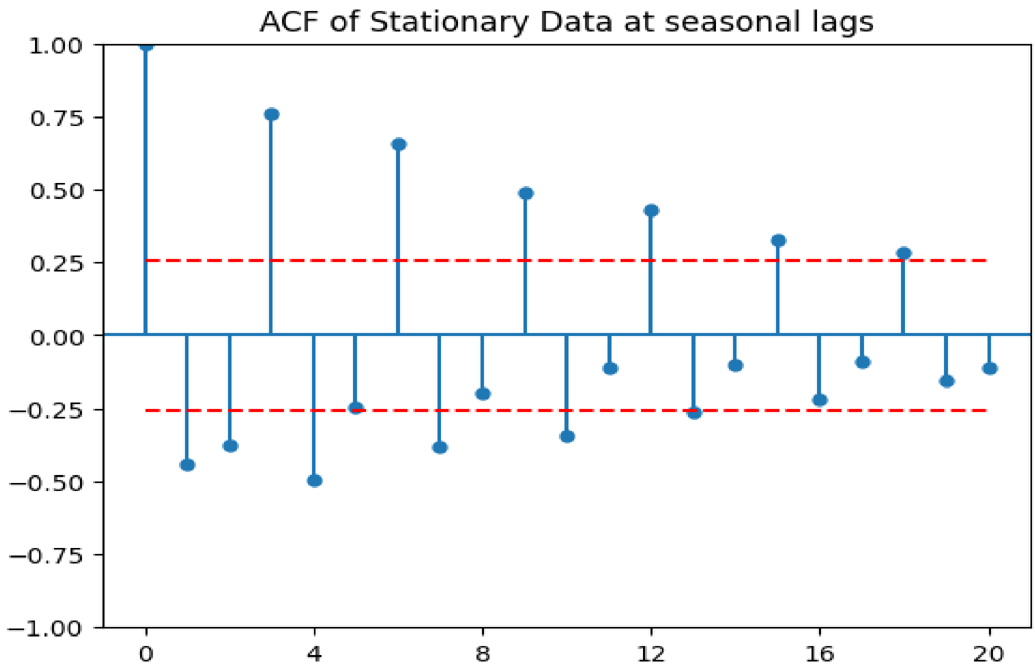

Figure 6: ACF of Statonary Data at seasonal lags

6.1.4.4 PACF of Seasonal Differenced Data

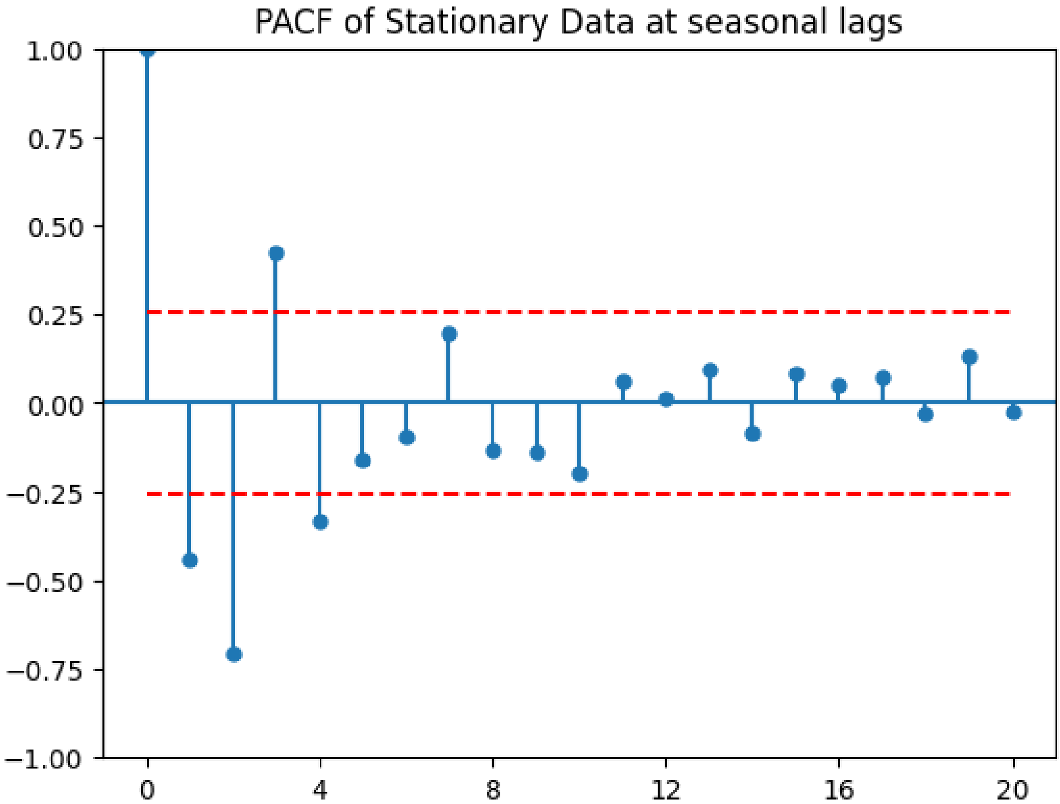

Figure 7: PACF of statonary Data at seasonal lags

From Figure 6 and Figure 7 shows the ACF and PACF of seasonally differenced data at seasonal. From the above the ACF and PACF of sesonally differenced data non seasonal AR order and non- seasonal MA order is maximum p = 4 and maximum q = 1 , and also know that d = 0. From the ACF and PACF of seasonally differenced data at seasonal lags maximum P = 4 and maximum Q = 9 and D = 1

In the next step we choose the best model for forscasting the values .Its done by choosing one model from all possible models according Akaike Information Criterion(AIC).

Here we obtain SARIMAX(0, 1, 0)x(3, 1, 0, 4) as the best model with AIC value 1849.323

Table 2

Parameters	coefficient	Std Error	z	P>|z|
ar . S. L4	-1.3565	0.126	-10.777	0.000
ar . S. L8	-1.3364	0.174	-7.662	0.000
ar . S. L12	-0.3833	0.154	-2.486	0.013
sigma2	7.23e+08	2.58e-11	2.54e+19	0.000

6.1.5 Diagnostic Checking

Diagnostic checking is performed for confirming the validity, effectiveness and reliability of statistical models. The main objective of it is to choose the right and best model. It is given in Figure 8

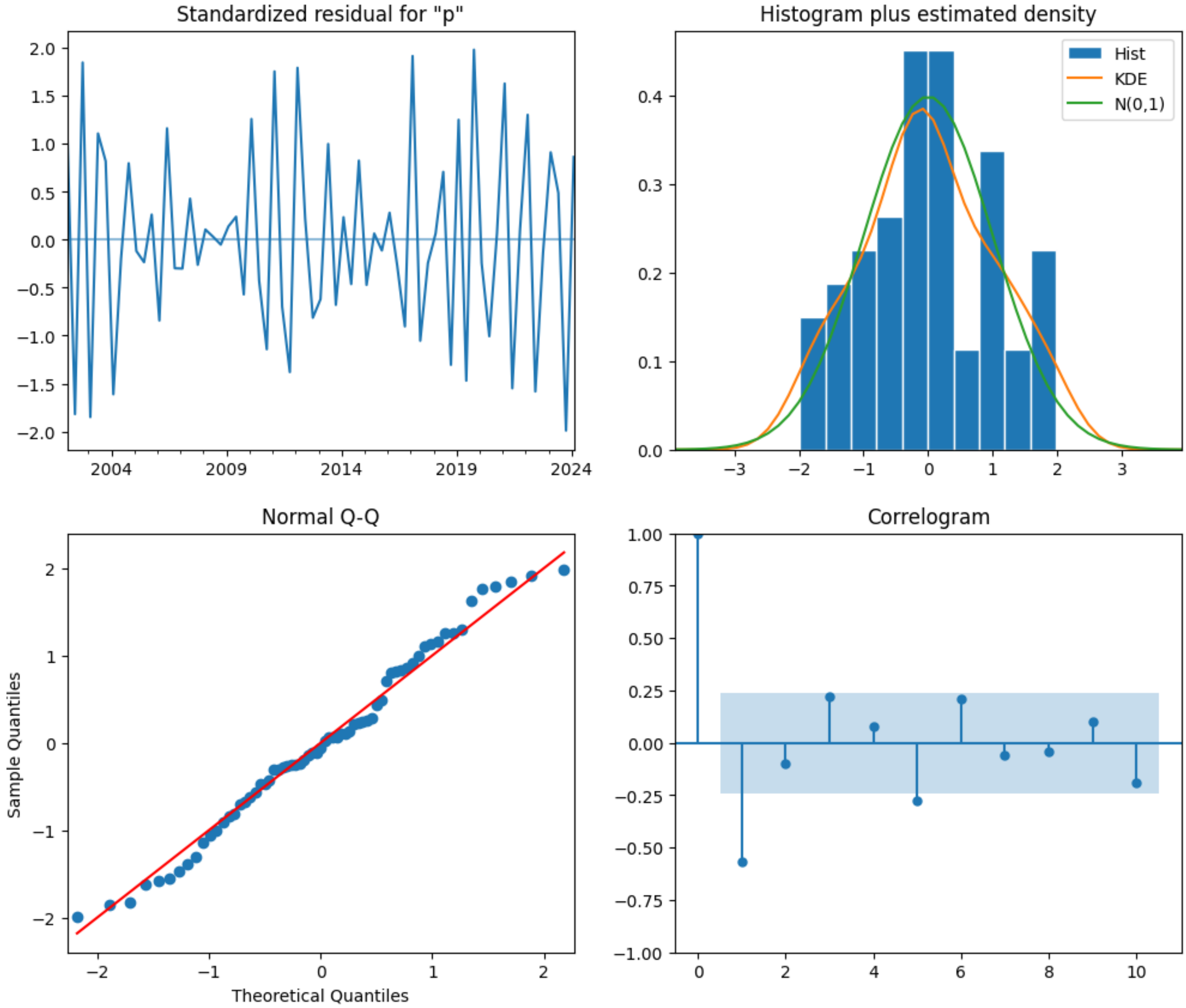

Figure 8: Diagnostic Plot

Figure 8 depicts the diagnostic values of the dataset which include the Normal Q-Q plot, which indicate that the residuals are approximately normally distributed. The correlogram give the idea that most of the lags die to zero means that there is less significant autocorrelation present in the data, indicating a good model fit. The histogram of standardized residuals overlaid with the kernel density estimate (KDE) and the standard normal distribution N (0,1). Since the histogram matches the normal distribution curve (green line), the residuals are approximately normal.

6.1.6 In-Sample forecast

Forecasting the sample means to forecast the actual data points or the training datapoints. Here we can evaluate model performance on training dataset. The table 3 is the actual and in sample forecasted values and the figure 9 is the plot of actual and predicted values

Table 3

MONTH	ACTUAL PRODUCTION	PREDICTED PRODUCTION
2015-02-01	50082.000	62738.687711
2015-06-01	28432.000	26665.938185
2015-10-01	14114.000	17111.812580
2016-02-01	60549.000	52942.713802
2016-06-01	30662.000	37930.716161
2016-10-01	4172.000	28468.372092
2017-02-01	54501.000	2939.510085
2017-06-01	32592.000	60910.477421
2017-10-01	10025.000	16536.335379
2018-02-01	59822.000	58075.442070
2018-06-01	25903.000	6845.901015
2018-10-01	16003.000	51079.759423
2019-02-01	63770.000	30100.175108
2019-06-01	1374.000	40890.071035
2019-10-01	40184.000	-13139.391889
2020-02-01	87003.000	93512.843997
2020-06-01	19304.000	46355.443705
2020-10-01	10742.000	9004.326905
2021-02-01	82546.000	38706.384108
2021-06-01	20913.149	62569.940595
2021-10-01	15812.553	13045.442040
2022-02-01	89983.768	54920.392997
2022-06-01	9569.407	52125.593874
2022-10-01	10874.000	14650.958521
2023-02-01	72749.512	48209.805885
2023-06-01	18880.800	5748.189566
2023-10-01	19462.750	72984.852139

6.1.6.1 Fitted Versus Actual values

The actual and fitted are plotted in the figure 9. The red line shows the predicted(fitted) values and the blue line shows the actual values

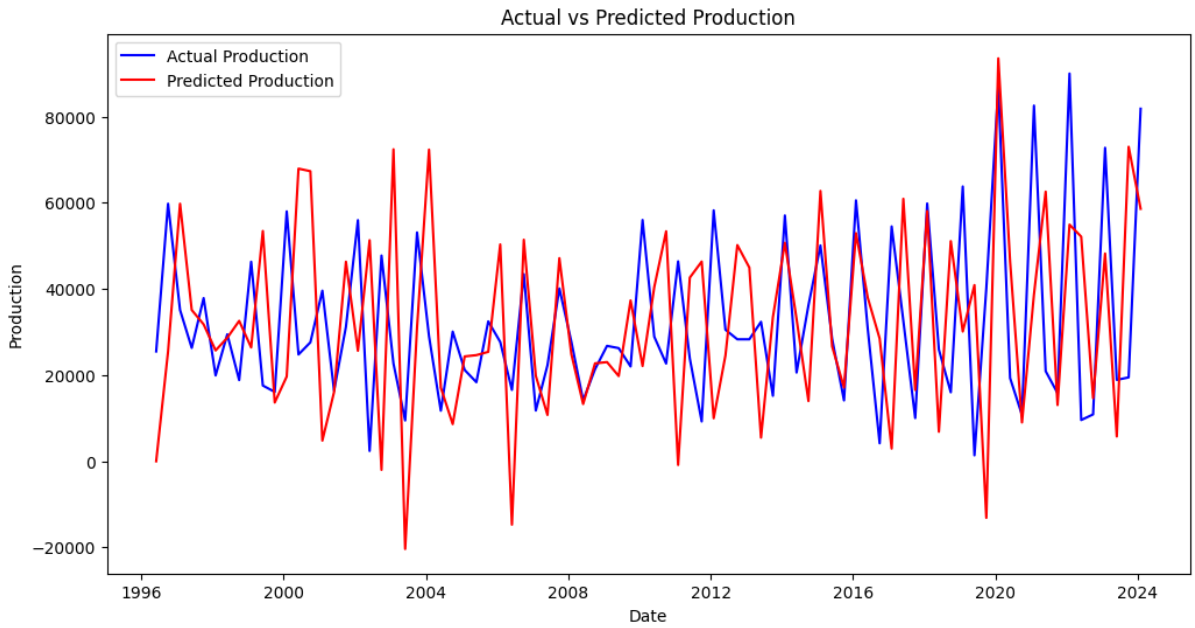

Figure 9: Actual Vs Predicted Production in SARIMA Model

6.1.6 Forecast of the Rice Production using the SARIMA Model

The model can be used for forecasting. The production from June 2024 to October 2030 is forecasted using the SARIMA model fitted. The forecasted values are computed and plotted (Table 4 and Figure 10). The dotted green line indicates the forecasted production.

Table 4

Months	Forecasted values	LCL	UCL
01-06-2024	13436.639333	0	66137.251545
01-10-2024	2058.857273	0	76588.777808
01-02-2025	83546.585207	0	174826.723148
01-06-2025	20542.499326	0	125943.723749
01-10-2025	15357.728148	0	126080.256557
01-02-2026	85743.828075	0	201543.389752
01-06-2026	7578.522118	0	128241.684150
01-10-2026	8916.487717	0	134254.665695

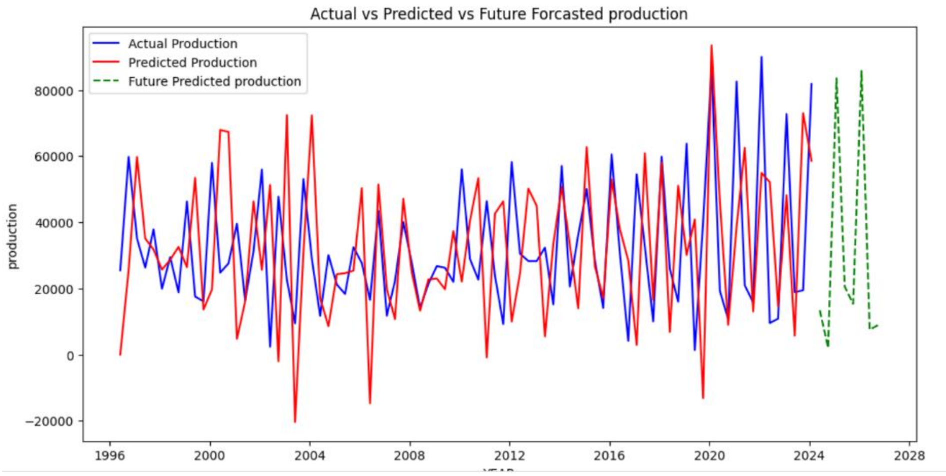

Figure 10: Actual vs Predicted vs Future Forecast Production

6.2 Modelling and Forecasting Using Holt-Winters Method

Holt-Winters Forecasting is used to find the forecast values of rice production. Table 5 shows the parameters estimate pf the mode

Table 5

Parameters	Parameter Estimation
Alpha (level)	0.0401474
Gamma (Trend)	0.0393733
Delta (season)	0.0677946

6.2.1 Diagnostic Checking

It is a crucial step to ensure the reliability, effectiveness and validity of statistical model. It helps to understand how precise the model is to improve the prediction accuracy.

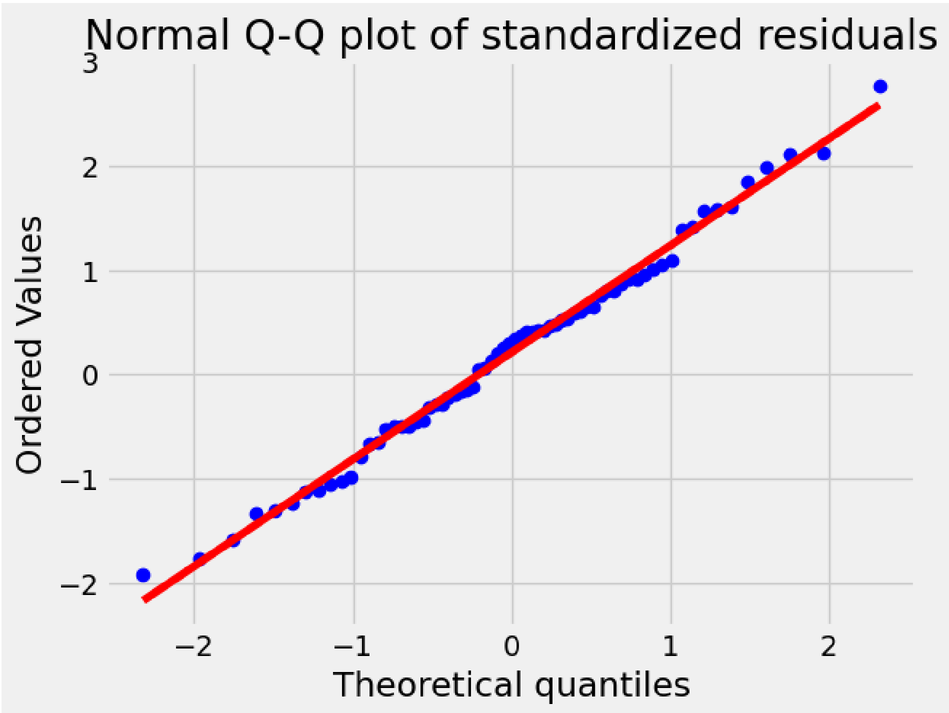

Figure 11: Normal Q-Q plot

Figure 11, shows the quantile-quantile (Q-Q) plot comparing the distribution of residuals with the normal distribution. It is clear that the most of the residual values lie on the straight line, which indicate that the residuals are approximately normally distributed.

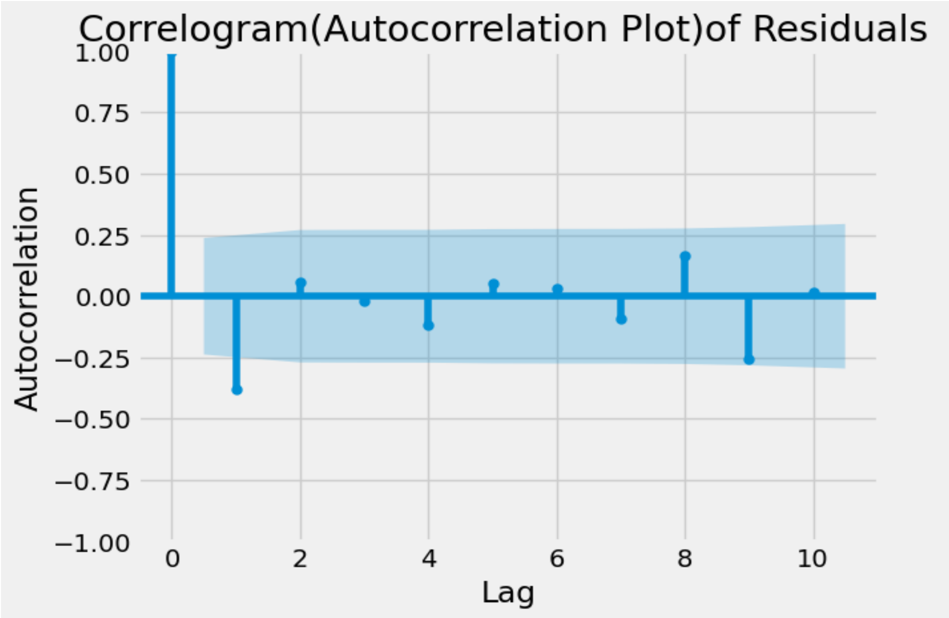

Figure 12: Correlogram

Figure 12, represents the correlogram. From the inspection, it is obvious that most of the lags fall to zero, which means that there is not much significant autocorrelation in the data.

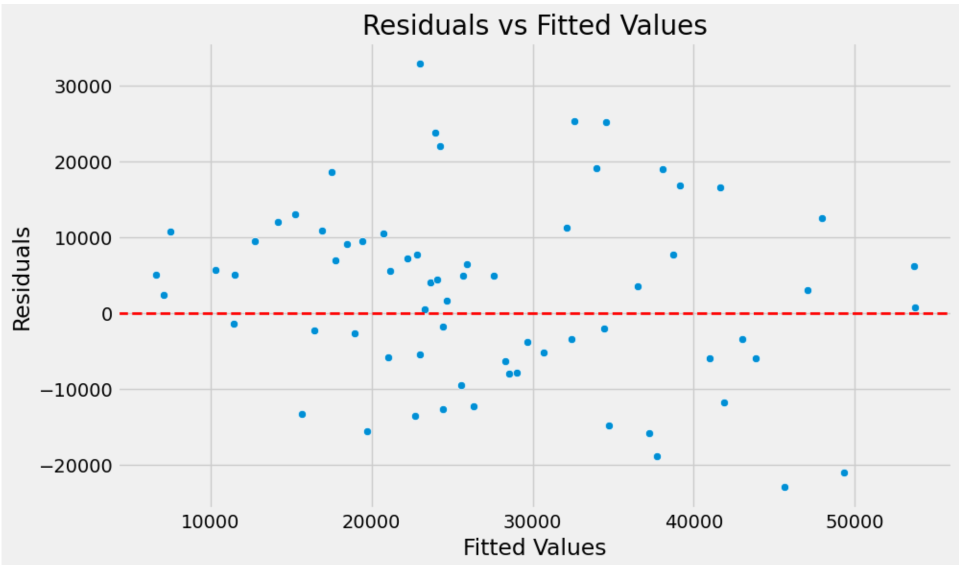

Figure 13: Residuals vs Fitted Values

The figure 13, represents scatter plot of residual indicates the variance of the residual is constant.

6.2.2 In-Sample Forecast

Now the fitted time series model is used to do the in-sample forecasting. The in-sample forecasting is done from February 2019 to February 2024 is given in table 6.

Table 6

Months	Actual Value	Predicted Value
01-02-2019	63770.000	56335.071047
01-06-2019	1374.000	27397.853975
01-10-2019	40184.000	9938.363415
01-02-2020	87003.000	55946.522288
01-06-2020	19304.000	27021.954175
01-10-2020	10742.000	9574.700796
01-02-2021	82546.000	55594.698476
01-06-2021	20913.149	26681.583764
01-10-2021	15812.553	9245.410930
01-02-2022	89983.768	55276.128433
01-06-2022	9569.407	26373.384568
01-10-2022	10874.000	8947.244964
01-02-2023	72749.512	54987.669072
01-06-2023	18880.800	26094.315819
01-10-2023	19462.750	8677.261122
01-02-2024	81817.310	54726.474383

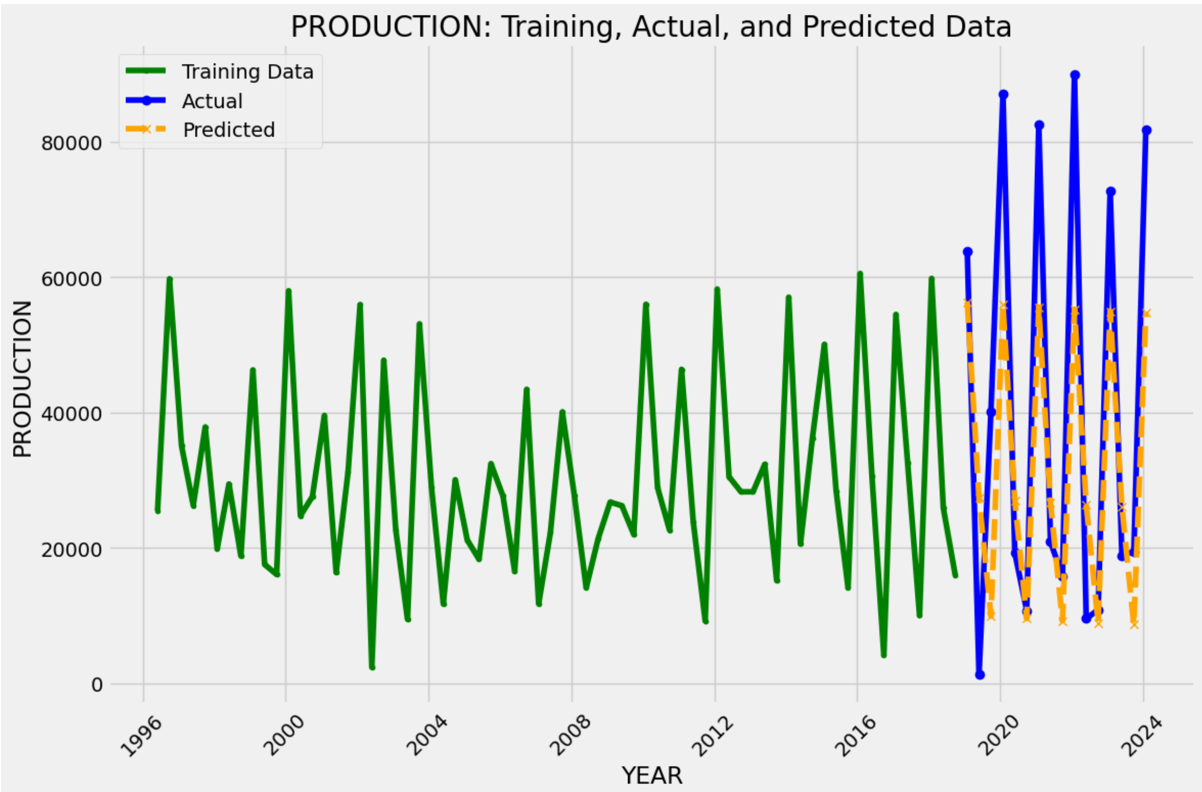

Figure 14: Production Training, Actual and Predicted Data for Holt- Winter's Model

From figure 14, it is given that the green line shows the training data, blue line shows the testing data and the orange line show the predicted data.

6.2.3 Forecast of Rice Production using Holt-Winters Model

The rice production from June 2024 to October 2030 is Forecasted. Table 7 and figure 15 gives these insights.

Table 7

Months	Forecasted Values	LCL	UCL
01-06-2024	25841.624159	20222.020260	92448.121835
01-10-2024	8432.795679	0	63510.904762
01-02-2025	54489.967359	0	56051.414203
01-06-2025	25612.816471	19833.471501	92059.573075
01-10-2025	8211.436681	0	63135.004962
01-02-2026	54275.814563	0	55687.751583
01-06-2026	25405.635285	19481.647688	91707.749263
01-10-2026	8011.000147	0	62794.634552

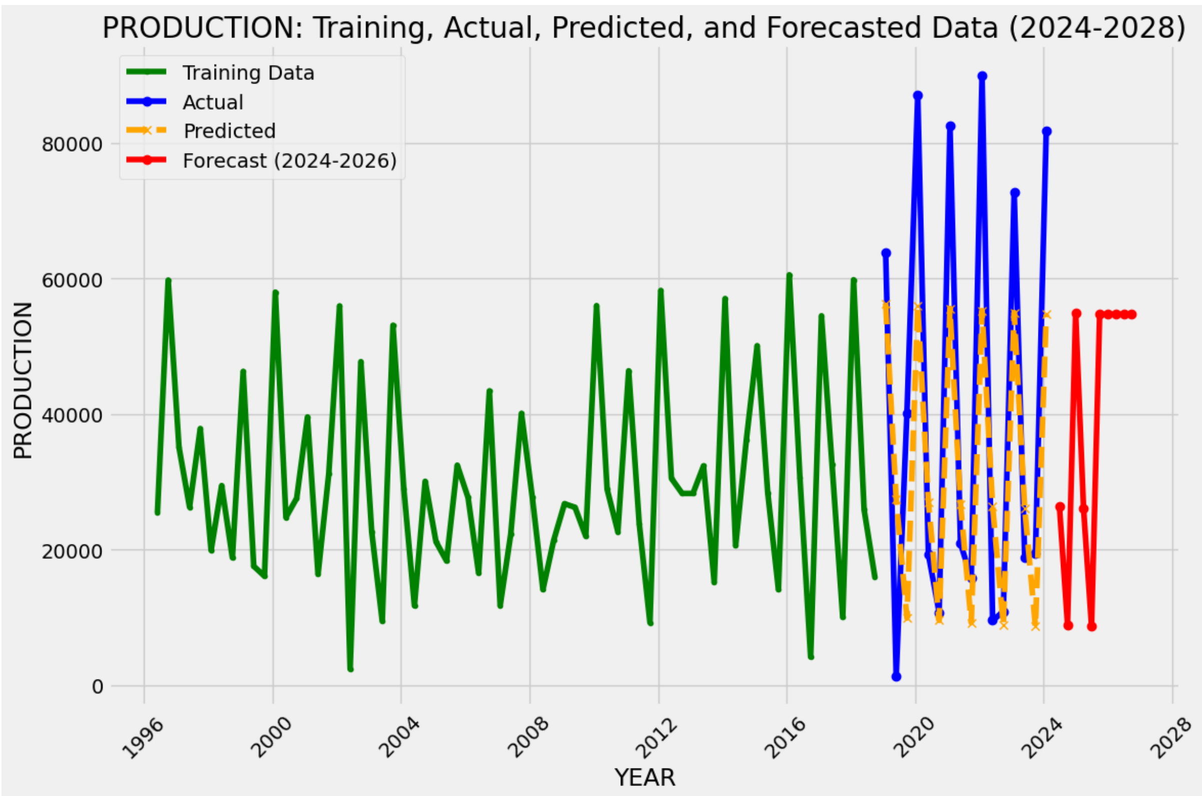

Figure 15: Production Training, Actual, Predicted and Forecast data for Holt- Winter's Model

6.3 Modelling and forecasting using CNN

CNN refers to the Convolution Neural Networks is used for the time series forecasting. Even though they are mainly used for image processing tasks, they can also extract patterns and temporal features by applying convolution filters over sequences of data

6.3.1 Diagnostic Checking

Here data is represented in a graphical model as in figure 16

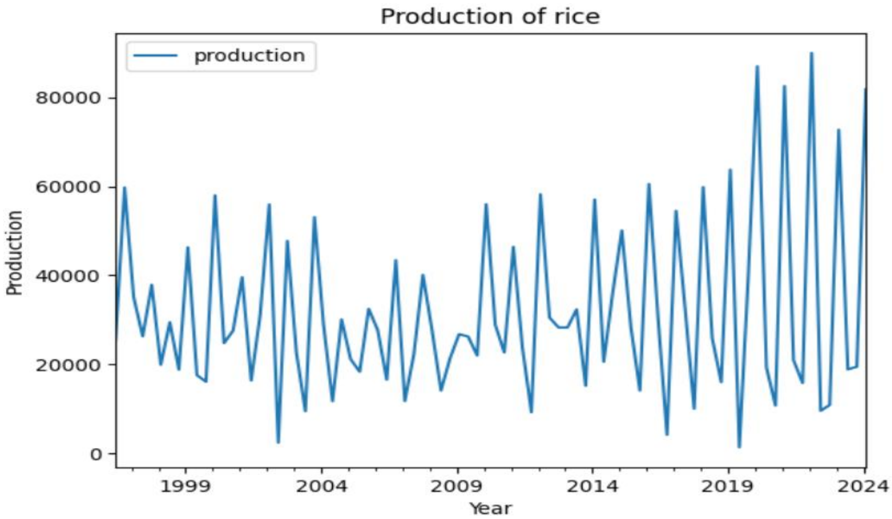

Figure 16: Production of Rice

After the data is plotted it is converted into a range of [0,1] and is processed for multiple iterations. A 1D Convolutional layer used to extract patterns and features from sequential or time series data. The histogram of standardized residuals overlaid with the kernel density estimate (KDE) and the standard normal distribution N (0,1) is also checked in figure 17.

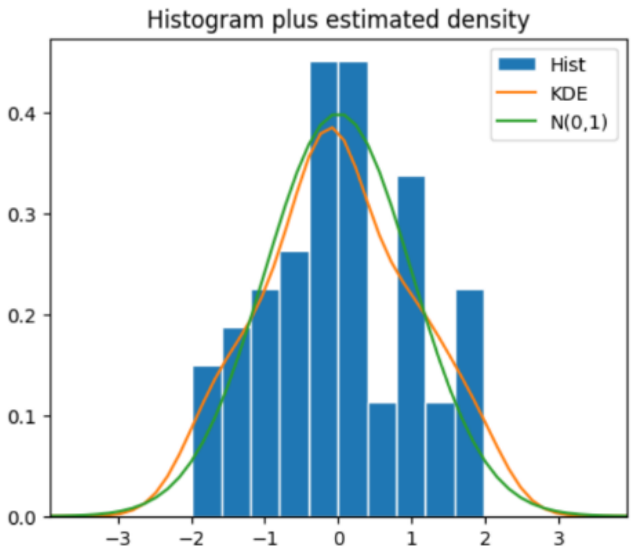

Figure 17: Histogram for Normal distribution of data

In this model, it has considered 64 number of filters, kernel-size of 2 and the activation function as relu, which

6.3.2 Forecast Values using CNN Model

It introduces non linearity. The data is then used to train the model and future values are forecasted and plotted in Table 8 and figure 18, respectively.

Table 8

Months	Forecasted Values	LCL	UCL
01-06-2024	14776.804688	0	32144.319370
01-10-2024	28261.480469	10893.967162	45628.996218
01-02-2025	67960.359375	50592.839830	85327.868886
01-06-2025	17025.845703	0	34393.360761
01-10-2025	31724.246094	14356.733370	49091.762426
01-02-2026	66986.632812	49619.116737	84354.145793
01-06-2026	13744.004883	0	31111.519458
01-10-2026	36785.136719	19417.623800	54152.652856

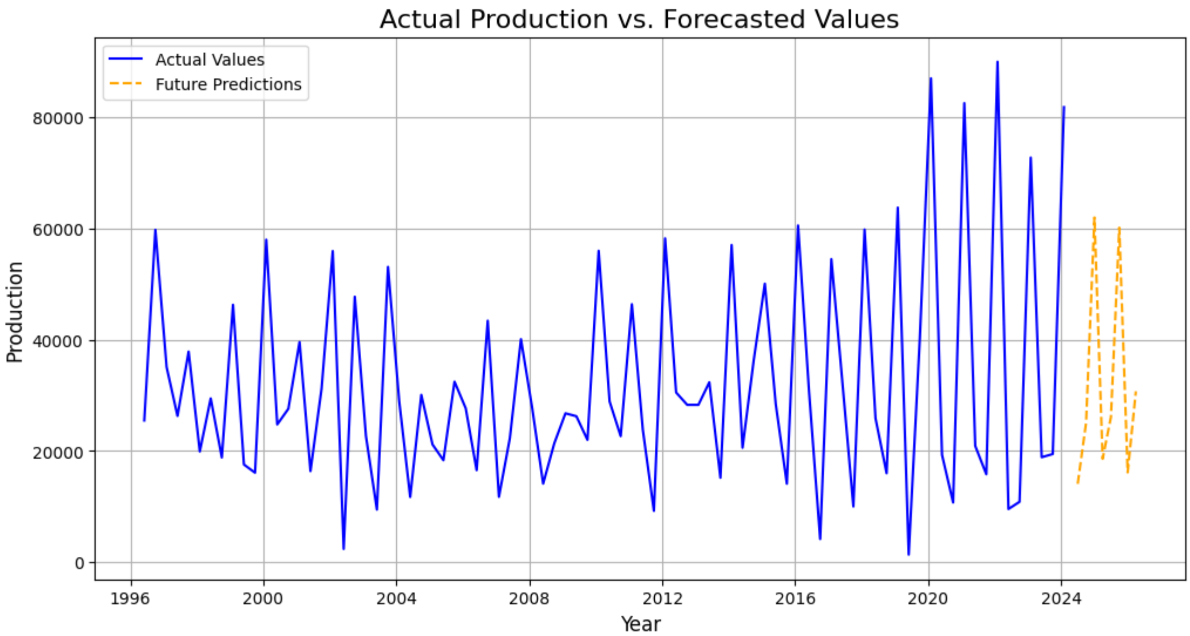

Figure 18: Actual Production vs Forecasted Values in CNN Model

6.4 Comparison Between SARIMA, Holt- Winter's and CNN Model

To determine the best model, the performance metrices should be compared and considered. Here the Mean Absolute Error (MAE) and Root Mean Square Error were provided for the three model.

Table 9

	SARIMA Model	Holt- Winter's Model	CNN Model
MAE	18590.44	16201.4427734708	15319.19
RMSE	26557.76879417141	19659.589021460324	18110.60

From the table 9, it is clear that the CNN model has the lower MAE and RMSE value compared to SARIMA and Holt-Winter's model. Thus, CNN is found to be a better model.

CONCLUSION

This study utilizes three forecasting models: the SARIMA Model, Holt's Winters Model, and the CNN Model. The performance of these models is evaluated using RMSE and MAE values to determine the most accurate method for predicting future values.

For the SARIMA Model, the best-performing configuration is ARIMA (0,1,0) (3,1,0) [4], with an AIC value of 1849.323. It yields a RMSE of 26557.76879417141 and a MAE of 18590.44.

In the Holt's Winters Model, the optimal configuration results in a RMSE of 19659.589021460324 and MAE of 16201.4427734708.

The CNN Model, which forecasts values based on pattern evaluation, achieves the future forecast values with a RMSE of 18110.60 and a MAE of 15319.19.

Therefore, the CNN Model is identified as the most effective for forecasting rice production. Furthermore, this study provides valuable insights for enhancing future production through various strategic measures.

Finding of the study:

Over the last few years Kerala has witnessed an increase in production in the months of October to January and February to May, but shows a decrease in production during the months of June to September. It can be seen that after the flooding in August 2018, a major increase in the production of rice can be witnessed, because the sediments and silt from various dams and rivers get settled in that land which increases the fertility of the soil (Thomas, 2020).

REFERANCE

Baidya, S., Radhakrishnan, P., & Gupta, A. K. (2023). Flood Resilience of Pokkali Rice: A Study of the 2018 Flood in Kerala. In The Palgrave handbook of socio-ecological resilience in the face of climate change: contexts from a developing country (pp. 197-212). Singapore: springer nature Singapore

Hamjah, M. A. (2014). Rice production forecasting in Bangladesh: An application of Box-Jenkins ARIMA model. Mathematical theory and modelling, 4(4), 1-11.

Nithin Raj, K., Lazarus, T. P., Aswathy Vijayan, D. A., Aparna, B., Joseph, B., & Stephen, R. (2020). Constraints in paddy cultivation faced by the farmers in upper Kuttanad: A study in Alappuzha district of Kerala. Constraints, 50(10).

Parreño, S. J. E. (2023). forecasting quarterly rice and corn production in the Philippines: A comparative study of seasonal ARIMA AND Holt winter model. ICTACT Journal on Soft Computing, 14(2).

Thomas K C International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349- 5162, Vol.7, Issue 12, page no.311-318, December-2020

Thomas, P. M. (2002). Problems and prospects of Paddy cultivation in Kuttanad region–A case study of Ramankari village in Kuttanad Taluk. Project report of Kerala Research Programme on Local Level Development (KRPLLD), Thiruvananthapuram.

Varughese, A. R., & Mathew, V. (2023). Climate change, flooding woes and mass exodus of inhabitants: an analysis of Kuttanad wetland ecosystem in Kerala, India. GeoJournal, 1-22

Veerakumaran, G., & Santhi, S. L. (2019). Impact assessment of Kerala flood 2018 on agriculture of farmers in Edathua Panchayat, Kuttanad Taluk of Alappuzha District.