Forecasting Production of Rice In India – Using Arima And Deep Learning Methods

Dr. S. A. Jyothi Rani[1], N. Chandan Babu[2]

Assistant Professor, Department of Statistics, O.U,
Research Scholar, Department of Statistics, O.U.
Osmania University, Hyderabad, Telangana

ABSTRACT *-- In this paper, forecasting of production of Rice (Million Tones) using Auto Regressive Integrated Moving Averages (ARIMA) method, Recurrent Neural Network, Multilayer Perceptron (MLP) and Convolution Neural Networks (CNN) are presented. The appropriate best model is evaluated by comparing mean square error (MSE), Root mean square error (RMSE), mean absolute percentage error (MAPE). The study of the results shows that CNN is performing better than the other models ARIMA, RNN and MLP.*

KEYWORD *-- ARIMA, Rice, RNN, CNN, MLP.*

I. INTRODUCTION

Paddy is one of the most important food crops of India. Paddy also called rice paddy, small, level, flooded field used to cultivate rice in Southern and Eastern Asia. The difference between rice and paddy is that, rice is (uncountable) cereal plants (oryza saliva) of the grass family whose seeds are used as food while paddy is rice, before it is milled. Rice is the most important crop of India as it is the staple food for most of the people of the country. This crop is the foundation of livelihood for millions of rural households and plays an important role in the country's food security, so the term "rice is life" is most appropriate in Indian context, especially in South India. In the area and Production of rice India occupies a predominant position. To keep pace with the increase in demand of the growing population the country has improved its yield by adopting modern production technologies such as high-yielding varieties/hybrids, expansion of irrigation potential, and use of chemical fertilizer. Demand for rice is expected to increase in future so the production of rice also needs to be increased. Taking the demand into consideration there is a dire need in the increase of rice productivity as the land area used for rice cultivation is declining day by day. In the country the inefficient use of inputs (fertilizer, water, labor, machines, new technology)and the increasing scarcity of water and labor especially for rice cultivation, new emerging challenges from climate change, rising fuel prices, increasing cost of cultivation, and socioeconomic changes such as migration of labor, urbanization, less liking for agricultural work by youths, and concerns from environmental pollution are the major conditions for productivity and sustainability of rice-based systems. The increased productivity per unit of area of rice with enhanced resources used efficiently is the only way to sustain rice production for meeting the increasing population demand. High-yielding varieties that have a resistance capacity to multiple stresses especially to the climatic changes can enhance the future productivity gain of rice in India, Exploring the efficient utilization of inputs (water, fertilizers, pesticides, machineries, labour, seeds etc.) reduce cultivation cost, enhance profit, and provide safe environment and these crop production techniques in rice could also increase factor productivity and sustainability of rice production in India.

II. REVIEW OF LITERATURE

Rahul Tripathi (2014)[1]: In this paper he Forecasted area of rice, production, and productivity of Odisha. He collected data from 1950-51 to 2008-09 by using univariate autoregressive integrated moving average (ARIMA) models and was compared with the forecasted values with actual values.

N. N. Jambhulkar (2013)[2]: In this paper he studied, ARIMA stochastic modeling is used for describing rice production in Punjab. The yearly rice production data collected from 1960-61 to 1999-2000 has been used for train the data and the data from 2000-01 to 2009-10 has been used for tested validation of the model. The best model has been chosen based on the minimum Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)

values. It has been found that ARIMA (1, 1, 2) model described the rice production data in Punjab by using this model they forecasted the production of rice for next 5 years.

E. Manjula, S. Djodiltachoumy (2017)[3]: In this paper they predicted crop yield from past data, achieved by applying association rule mining on agriculture data. They focused on creation of a prediction model which may be used to future prediction of crop yield. They analyzed crop yield prediction using data mining technique based on association rules for the selected region i.e. district of Tamil Nadu in India. The experimental results shows that the proposed work efficiently predict the crop yield production.

Narayanan Balakrishnan and Dr.Govindarajan Muthukumarasamy(2016)[4]: They studied AdaSVM and AdaNaive are the proposed ensemble model used to project the crop production over a period of time. This ensemble model is compared to SVM and Naive Baye's methods. The two parameters used separately for prediction of output are the accuracy and the classification error. The finding yields that AdaSVM and AdaNaive are agreeable than SVM and Naive Baye's for the data set analyzed.

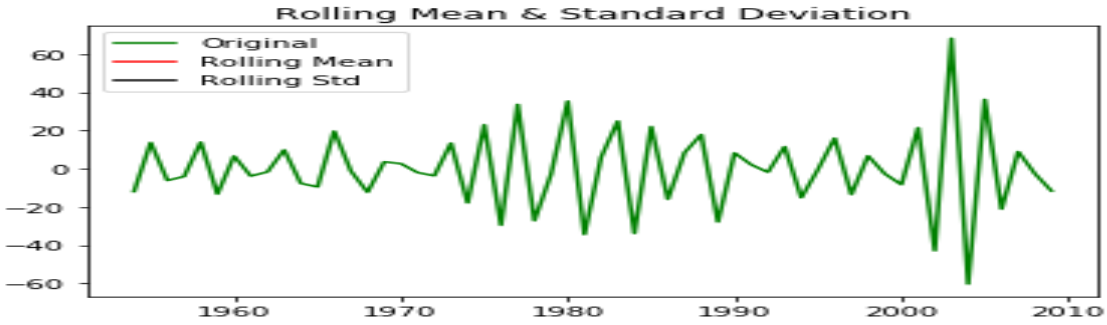
Dr. Yamin Hassan, Naranarayan Buragohain and Shahidul Islam (2018)[5]: They Forecasted rice production by using time series analysis techniques may be helpful to find out the best fitted model to forecast the production of rice in Assam. Forecasted Prediction of rice will help the higher authorities to adopt appropriate strategy to tackle the management of economy related to rice production.

III. OBJECTIVES OF STUDY

1. To fit the best model using ARIMA, RNN, MLP and CNN using python code.
2. To find MSE, RMSE, MAPE values for the models: ARIMA, RNN, MLP and CNN.
3. To forecast the production of Rice for the next 10 years using the best model among the models used in this study.

IV. RESEARCH METHODOLOGY

To fit an ARIMA model, Recurrent Neural network, Multilayer perceptron and Convolution Neural networks requires a sufficiently large data set. In this study, we collected data from RBI website and used the data for Non-commercial crop production of Rice for the year 1950-51 to 2017-18. The development of ARIMA model for a single variable involves identification, estimation and verification. Each of these steps is now explained for a non-commercial crop production.




```
Results of Dickey-Fuller Test:
Test Statistic                  -5.076572
p-value                          0.000016
#Lags Used                      11.000000
Number of Observations Used     44.000000
Critical Value (1%)             -3.588573
Critical Value (5%)             -2.929886
Critical Value (10%)            -2.603185
dtype: float64
```

First we tested whether the data is stationary by using Dickey-Fuller test:
Null Hypothesis H0: The data is not stationary
Alternative Hypothesis H1: The data is stationary
P-value less than 0.05: Reject the null hypothesis (H0), the data does not have a unit root and is stationary.

Once the data is stationary we trained years from 1951 to 2008 overall 58 years data based on this we tested the production of sugarcane for 10 years from 2009 to 2018.

Statespace Model Results

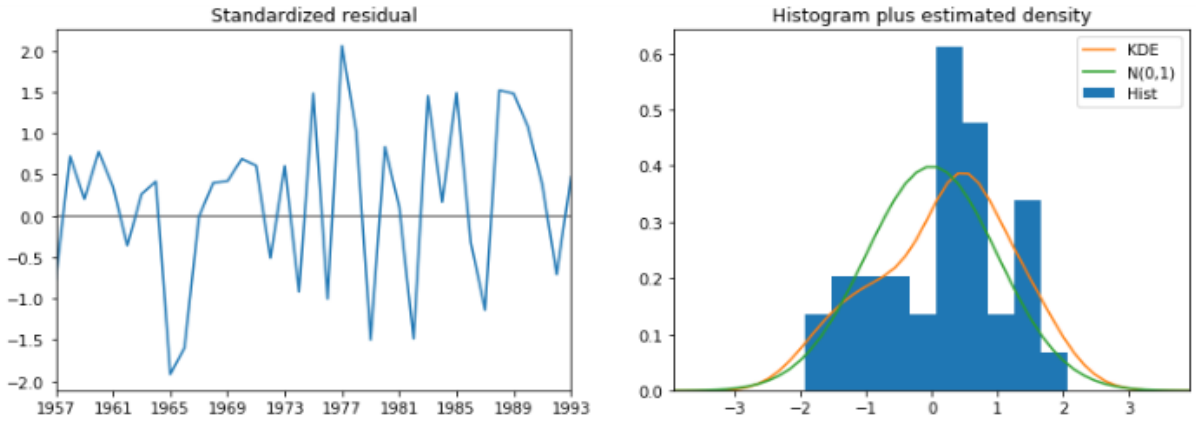
Dep. Variable:	Actuals_Rice	No. Observations:	43
Model:	SARIMAX(1, 2, 3)	Log Likelihood	-107.423
Date:	Mon, 27 Jan 2020	AIC	224.847
Time:	10:43:17	BIC	232.901
Sample:	01-01-1951	HQIC	227.686
	- 01-01-1993		
Covariance Type:	opg		

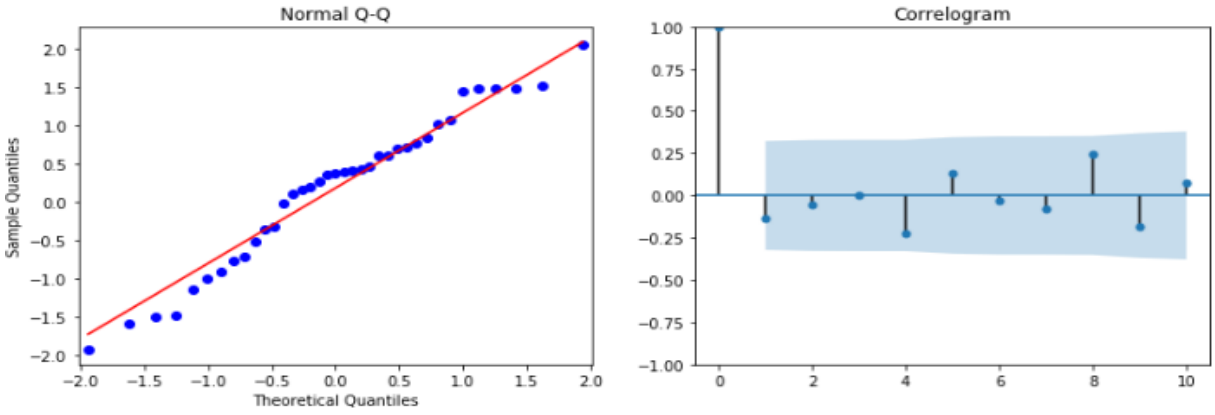
	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-0.6547	0.448	-1.460	0.144	-1.534	0.224
ma.L1	-1.1383	3.04e+04	-3.75e-05	1.000	-5.96e+04	5.96e+04
ma.L2	-0.5176	7610.895	-6.8e-05	1.000	-1.49e+04	1.49e+04
ma.L3	0.7879	2.49e+04	3.17e-05	1.000	-4.87e+04	4.87e+04
sigma2	16.9331	5.34e+05	3.17e-05	1.000	-1.05e+06	1.05e+06

Ljung-Box (Q):	22.12	Jarque-Bera (JB):	1.32
Prob(Q):	0.97	Prob(JB):	0.52
Heteroskedasticity (H):	1.67	Skew:	-0.34
Prob(H) (two-sided):	0.38	Kurtosis:	2.36

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).

Plot – Diagnostics





A recurrent neural network (RNN) is a class of ANN, whereby the network's connections form a time sequence for dynamic temporal behavior. Furthermore, RNNs benefit from extra memory to analyze the time series sequences. In ANNs, back propagation (BP) is a multi-layered neural network where weights are calculated using the propagation of the backward error gradient. In Back propagation, there are several steps to learn cycle, using a function for activation to send signals to the other nodes. Among various ANNs, the back propagation ANN (BPNN) was identified for prediction tool suitable for production of rice crop.

Multilayer Perceptron (MLP) is also used to forecast the production of sugarcane crop.

Convolution Neural Network (CNN) is a Deep Learning algorithm which can take in an input image, assign importance to various aspects/objects in the image and be able to differentiate one from the other.

V. CONCLUSION

Test Accuracy

	ARIMA	***MLP***	***RNN***	***CNN***
MSE	434.67318	1975.8261	480.76	149.71
RMSE	20.848817	44.4502	21.92	12.23
MAPE	20.399434	0.4606	0.18	**0.102**

Test of accuracy can be verified based on Mean square error (MSE), Root mean square error (RMSE) and Mean absolute percentage error (MAPE)

The performance of the above models is evaluated by comparing mean square error (MSE), Root mean square error (RMSE), mean absolute percentage error (MAPE). The results show that CNN is performing better than the other models ARIMA, RNN and MLP. The forecasted production of Rice (Million Tones) from 2018-19 to 2027-28.

Year	2018-19	2019-20	2020-21	2021-22	2022-23	2023-24	2024-25	2025-26	2026-27	2027-28
Forecasted Rice (Million tons)	91.84	91.54	91.3	91.09	90.83	90.63	90.42	90.22	90.04	89.85

REFERENCES

[1] Rahul Tripathi (2014): Research Article | Open Access, Volume 2014 |Article ID 621313 | 9 pages | https://doi.org/10.1155/2014/621313, Forecasting Rice Productivity and Production of Odisha, India, Using Autoregressive Integrated Moving Average Models.

[2] N. N. Jambhulkar (2013): 'Modeling and forecasting of rice production in West Bengal using ARIMA model' under theme VI 'Socio-economics is ues and livelihood security' in an ARRW Golden Jubilee International Symposium on "Sustainable Rice Production and Livelihood Security: Challenges and Opportunities" held at C. R. R. I. Cuttack from 2-5th March 2013.

[3] E. Manjula, S. Djodiltachoumy (2017): International Journal of Computational Intelligence and Informatics, Vol. 6: No. 4, March 2017 ISSN: 2349-6363 298

[4] Dr. Yamin Hassan, Naranarayan Buragohain and Shahidul Islam (2018): International Journal of Development Research, Volume: 08, Article ID: 14166

[5] Narayanan Balakrishnan and Dr.Govindarajan Muthukumarasamy (2016): International Journal of Computer Science and Software Engineering (IJCSSE), Volume 5, Issue 7, July 2016 ISSN (Online): 2409-4285 www.IJCSSE.org Page: 148-153.