PARIPEX - INDIAN JOURNAL OF RESEARCH | Volume - 12 | Issue - 09 | September - 2023 | PRINT ISSN No. 2250 - 1991 | DOI : 10.36106/paripex

ORIGINAL RESEARCH PAPER

Statistics

A COMPARISON OF ARIMA & NNAR MODELS FOR PRODUCTION OF GROUND NET IN THE STATE OF ANDHRA PRADESH

KEY WORDS: AR, MA, ARMA, ARIMA, NNAR, RMSE and Akaine's Information Criterion AIC

G. Vijayalakshmi

ABSTRACT

If the data is linear and non-stationary, the models viz. Auto-Regressive (AR), Moving Average (MA), and Auto-Regressive Moving Average (ARMA) models cannot be used. So, another important forecasting technique called Auto-Regressive Integrated Moving Average (ARIMA) with (p, d, q) terms can be used. The best feature of Artificial Neural Networks when it is applied to forecasting data is its inherent capability of nonlinear modeling without any presumption about the statistical distribution of the given data. Model selection criteria based on RMSE for ARIMA and Neural Network Auto-regressive (NNAR) models are computed. An appropriate model has to be framed effectively for the production of Wheat data in the state of Andhra Pradesh taken during the period from 1982 to 2022 (for 40 years).

INTRODUCTION

The Most widely used important statistical tools for traditional forecasting techniques for stationary and linear data are Auto-regressive (AR) with p terms, and Moving Average (MA) with q terms in these models. They are combined together to form Auto-regressive Moving Average (ARMA) with (p,q) terms in the model, where p is the Auto-regressive terms and q is the Moving Average terms. When the data is non-stationary, we use ARIMA (p,d,q) model which is also known as Box-Jenkins's Methodology, where d is the time lagged differencing. When d= 0, it becomes simply ARMA with p and q terms model.

A Neural Network is a simplified model of the same way that the human brain processes information. It works by stimulating a large number of inter-connected processing units that resembles abstract versions of neurons. The processing units are organized in layers. They are arranged into three parts in a neural network:

a) An input layer with unit(s) representing the input field(s),
b) One or more hidden layers, and
c) An output layer with unit(s) representing the target field(s).

The units are connected with varying connection strengths (or weights). Input data are presented in the first layer and the values are propagated from each neuron to every neuron in the next layer. Eventually, a result shall be delivered from the output layer.

The main contributors in the field of forecasting and neural networks are Yule (1926), Box and Jenkin's (1976), Young (1982), Arash Bahrammirzaee (2010), Mehdi Khashei., Mehdi Bijari (2010) [15], Prapanna Mondal, Labani Shit, and Saptarsi Goswami (2014)[16].

Objectives

The important objectives of our current paper are outlined as follows:

1. To study the forecasting techniques by applying ARIMA and Neural Network Autoregressive (NNAR) Models in our methodology.
2. To compare the above models by computing the RMSE.
3. To study the patterns in the production of Rice in the state of Andhra Pradesh during 40 (for Forty) time periods (i.e., from 1982 to 2022).
4. To forecast the production of Rice for the next 8 years.
5. To compute AIC for ARIMA model.
6. To analyze the forecasted results by applying the suitable forecasting.
7. To point out the future development in view of Indian agricultural scenario.

METHODOLOGY

a) ARIMA Model

The terms ARIMA (p , d , q) model can be represented as

$$\left[1-\beta(1+\alpha_1)+\alpha_1\beta^2\right]X_s=\lambda_1+e_s-\mu_1 e_{s-1}$$

$$X_s=(1+\alpha_1)X_{s-1}-\alpha_2 X_{s-2}+\lambda_1+e_s-\mu_1 e_{s-1} \quad \ldots\ldots \quad (1)$$

In this above form, the ARIMA models look like a conventional Regression Equation except that there is more than one error on the right-hand side.

Suppose p is the number of auto-regressive terms, q is the number of Moving Average terms and d is the degree of differencing and the model is represented as ARIMA (p,d,q) models.

Further, derivatives can also be taken into account by considering the Auto-Regressive or Moving Average trends that occur at certain points of time.

Let us have ARIMA model with p^{th} order auto-regressive terms given by

$$Y_s=\alpha_0+\alpha_1 Y_{s-1}+\alpha_2 Y_{s-2}+\ldots+\alpha_p Y_{s-p}+\varepsilon_s \quad \ldots\ldots \quad (2)$$

The ARIMA model having Moving Average model with q terms is given by

$$Y_s=\lambda+\varepsilon_s-\theta_1\varepsilon_{s-1}-\theta_2\varepsilon_{s-2}-\ldots-\theta_q\varepsilon_{s-q} \quad \ldots\ldots \quad (3)$$

ARIMA model having AR with p terms and MA with q terms is given by

$$Y_s=\alpha_0+\alpha_1 Y_{s-1}+\alpha_2 Y_{s-2}+\ldots+\alpha_p Y_{s-p}+\varepsilon_s-\theta_1\varepsilon_{s-1}-\theta_2\varepsilon_{s-2}-\ldots-\theta_q\varepsilon_{s-q} \quad \ldots\ldots \quad (4)$$

Now, ARIMA (0,1,1) model is given by

$$Y_s-Y_{s-1}=\varepsilon_s-\theta_1\varepsilon_{s-1} \quad \ldots\ldots \quad (5)$$

Now, ARIMA (0,1,1) forecasting model in exponential smoothing is given by

$$\hat{y}_{s+1}=y_s-\theta_1(y_s-\hat{y}_s)=(1-\theta_1)y_1+\theta_1\hat{y}_s \quad \ldots\ldots \quad (6)$$

B) Neural Networks

If the time series data is non-stationary, then an effective forecasting techniques are introduced, called Artificial Neural Networks. These techniques are data driven and self adaptive by nature. In the last few decades, lot of research has been carried-out in Artificial Neural Networks.

Neural networks approach has been suggested as an alternative technique to forecasting and gained huge popularity in last few years. The basic objective of neural networks is to construct a model for stimulating the intelligence of human brain into machine. Similar to the work of a human brain, artificial neural networks try to recognize regularities and patterns in the input data, learn from experience and then provide generalized results based on their known previous knowledge.

Neural Network Autoregressive (NNAR) Model: -

Simple mathematical models of the brain form the basis of artificial neural networks (ANN), which are used in forecasting. Complex linear and nonlinear relationships between the response and its predictors are possible with their help. Lagged values of the dependent variable y are used as inputs to the feed-forward neural network, which also has a single hidden layer of size nodes. The model is valid for a wide variety of fitted repetition networks, all of which have initial weights chosen at random. When making predictions, these are then averaged. The network is optimized for making predictions in a single step.

Neural Network Auto Regressive (NNAR) With One Hidden Layer: -

In this research, we focus exclusively on feed-forward networks with a single hidden layer, and we designate the number of lagged inputs and the number of hidden nodes in the network with the notation NNAR (p, kp, k). For instance, a neural network with a hidden layer consisting of five neurons and using the previous nine observations (yt1, yt2,..., yt9) to predict the current output yt yt is called a NNAR (9, 5) model. Without the constraints on the parameters to assure stationary, the NNAR (p, kp, 0) model is similar to the ARIMA (p, 0, kp, 0, 0) model.

Each layer of nodes in a multilayer feed-forward network receives inputs from the layers above it. All of a layer's nodes' outputs become inputs for the following layer. A weighted linear combination of the inputs to each node is used. A nonlinear function is then applied to the result, and the result is then output. For example, the inputs into hidden neuron jj in Figure 1 are merged linearly to provide

$$z_j = b_j + \sum w_{i,j} x_i \quad \ldots\ldots\ldots\ldots \quad (7)$$

In the hidden layer, this is then modified using a non – linear function such as a sigmoid,

$$s(z) = \frac{1}{1+e^{-z}} \quad \ldots\ldots\ldots\ldots\ldots \quad (8)$$

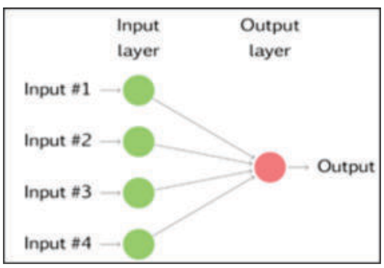

Fig 1: Simple Neural Network With Input Layer

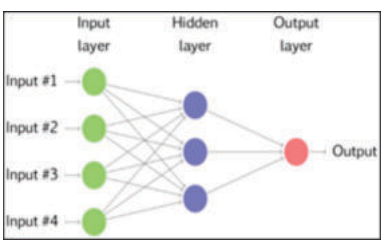

Fig 2: Neural Network With Four Inputs And One Hidden Layers To Give The Input For The Next Layer.

This tends to reduce the effect of extreme input values, thus making the network somewhat robust to outliers.

The values for b1, b2, b3, and w1, 1,..., w4, 3 are "learned" from the data, as are the values for w1, 1,..., w4, 3. Weights typically have their values capped so they don't get too large. The "decay parameter," or weight-restricting parameter, is typically equal to 0.1. The initial values for the weights are chosen at random, and they are subsequently modified based on the data that has been collected. As a result, the predictions made by a neural network contain some degree of chance. For this reason, the network is often trained multiple times with varying random seed values. There needs to be an up-front agreement on how many nodes will make up each hidden layer.

We shall apply the different forecasting methods are Auto Regressive Integrated Moving Average (ARIMA) and Neural Networks Models to forecast the production of Rice in the State of Andhra Pradesh.

Empirical Analysis Forecasting For Wheat Production Using Arima

Discriptive statistics for Ground net production. We can see that the average yield of groundnut production is 73.6853 (Kg./Hectare) and the standard deviation is 15.6761 from the average in Table No. 4.21, which displays the descriptive statistics of groundnut production. Since no outliers were found in the Sugarcane data set, the result of the Grubb's test is 2.0722. We infer that the Groundnut production data is normally distributed because the Jarque - Bera test, used to examine its normality, yielded a result of 0.7164, which was not statistically significant at the 0.05 level.

Table1: Descriptive Statistics Of The Yield Of Ground Nut (1982-2022)

GROUNDNUT PRODUCTION (IN LAKHS TONNES)	
Mean	73.6853
Median	73.7000
Maximum	102.400
Minimum	41.2000
Standard Deviation	15.6761
Skewness	-0.0432
Kurtosis	2.3582
Grubbs test	2.0722 (0.6990)
Jarque-Bera test	0.7164 (0.6989)

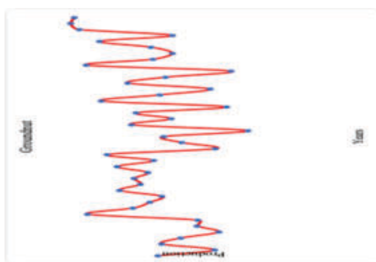

Fig 3: Time Series Graph For Ground Net Production Data.

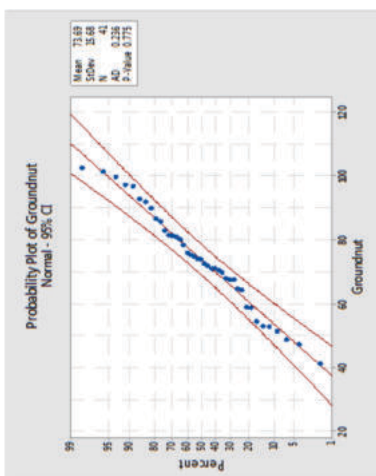

Fig4.: Trend Of Yield Of Groundnut Production

All of the results are insignificant for ACF at the 0.05 level. Both sets of univariate time series data are insignificant when the PACF is considered separately from the others. The ARIMA system is present in both series, since the other PACF points exhibit a wave with positive and negative values. Therefore, series filtering is accomplished by the process of differencing.

Model Selection For Analysis Of Errors

The following are the best forecasting models for different ARIMA (p, d, q) and NNAR (p, q) models as given below:

Table4: ARIMA (p, d, q) and NNAR (p, q) Models

Model	ARIMA (p, d, q)	RMSE	MAPE	MSE	AIC
1	ARIMA (0, 1, 1)	53.1440	4.8300	0.7380	439.66
2	ARIMA (1, 1, 0)	61.8370	6.1360	0.9718	442.04
3	ARIMA (0, 1, 2)	62.0700	6.1352	0.9651	451.60
4	ARIMA (1, 1, 1)	63.5865	6.2600	0.9938	451.13
5	ARIMA (2, 1, 0)	61.3855	5.9700	0.9414	450.77
6	NNAR(1, 1)	64.8142	6.0901	0.9063	448.69

The production of groundnuts is best described by an ARIMA (0, 1, 1) model. Parameter estimation taking into account ideas from Chatfield (1971), Box-Jenkins (1976), and Cressie (1989). (1988). Different series were employed to make the series stationary for both area and production. In this investigation, groundnut cultivation land and output were found to be stationary in the first difference series. The best-fitting ARIMA model is selected based on minimal Akaike's Information Criterion (AIC) and Mean Absolute Percentage Error (MAPE) in Table 4.24.

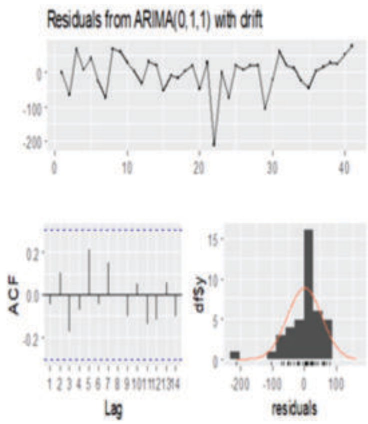

Fig5: Residuals Analysis For Groundnut ARIMA (0, 1, 1) Model

Figures 4.31 and 4.32 show residual plots of ACF and PACF, which are used for diagnostic testing of the models. Since every ACF and PACF falls within the confidence interval, the model guarantees that the residuals, or mistakes, have a white-noise character.

Forecasting Groundnut Production Using Arima And Nnar Models

The following table displays the predicted values of groundnut output over the next eight years using the ARIMA (0, 1, 1) model and the NNAR (1, 1) model of neural networks.

Table 5: Forecasted Value Of Groundnut Production Using ARIMA and NNAR Model

Year	GROUNDNUT PRODUCTION (IN LAKHS TONES)	
	Forecasted ARIMA (0, 1, 1)	Forecasted NNAR (1, 1)
2023	1266.449	1345.535
2024	1284.607	1394.18
2025	1302.766	1449.703
2026	1320.924	1512.927
2027	1339.083	1584.481
2028	1357.241	1664.559
2029	1362.55	1702.63
2030	1388.65	1795.62

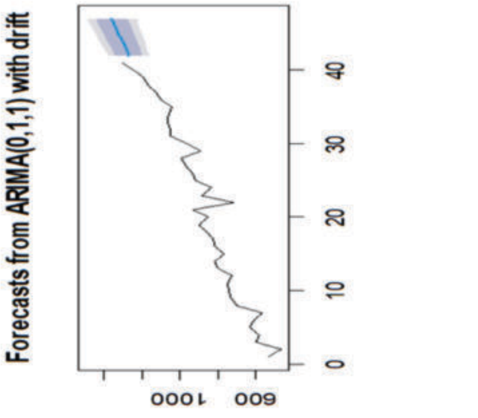

Fig6: Forecasts Values Of Groundnut Production With ARIMA (0, 1, 1) Model

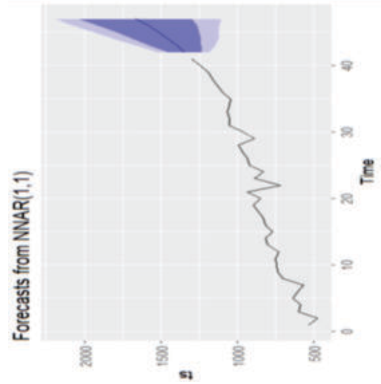

Fig7: Forecasts Values Of Groundnut Production With NNAR (1, 1) Model

Using their own model selection criteria, we compared the ARIMA and NNAR models. In this study, we apply the ARIMA and NNAR models we created to the problem of predicting India's groundnut harvest. Forecasts can be divided into two categories: those based on the ARIMA Model (with periods of zero, one, and two) and those based on the NNAR Model (with periods of zero and two). For the next eight years, from 2023 to 2030, genuine projections are generated for use in planning and other contexts. The results of these projections are shown in Table 4.25. ARIMA performed better than NNAR in terms of both predictive capacity and forecasting capacities, as demonstrated by this study. We find that the ARIMA (0, 1, 1) model is superior to the NNAR model when making predictions into the future. When compared to the NNAR (1, 1) model, the values for the RMSE (53.1440), MAPE (4.8300), and AIC (439.56) are all lower.

CONCLUSIONS

In this paper, we have studied the forecasted and future forecasted values of Ground net Production in the State of Andhra Pradesh using ARIMA model and Neural Network Autoregressive (NNAR) Models. These models are studied and applied. In our study,

We can observe that the forecasted values of Ground net Production from 2023 to 2030 for both models in the State of Andhra Pradesh (Graph No. 6 & 7) shows higher production values in ARIMA (0, 1, 1) Model while it shows lesser production in NNAR (1,1) Model.

REFERENCES

1. Arifovic J, Gencay R. Using genetic algorithms to select architecture of a feed-forward artificial neural network. Physica A. 2001;289:574–594.
2. Atiya FA, Shaheen IS. A comparison between neural-network forecasting techniques-case study: River flow forecasting. IEEE Transactions on Neural Networks, 1999, 10(2).
3. Box P, Jenkins GM. Time series analysis: Forecasting and control. San

Francisco, CA: Holden-day Inc. 1976.
4. Mitrea CA, Lee CKM, Wu Z. A comparison between Neural Networks and Traditional Forecasting Methods: A case study International Journal of engineering Business Management. 2009;1(2):19-24.
5. Carlos Gershenson., Artificial Neural Networks for Beginners.
6. Chen A, Leung MT, Hazem D. Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. Computers and Operations Research. 2003;30:901–923.
7. Zhang G, Patuwo BE, Hu MY. Forecasting with artificial neural networks: The state of the art", International Journal of Forecasting. 1998;14:35-62.
8. Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model", Neurocomputing. 2003;50:159–175.
9. Ghiassi M, Saidane H. A dynamic architecture for artificial neural networks. Neurocomputing. 2005;63:97–413.
10. Ginzburg I, Horn D. Combined neural networks for time series analysis. Advance Neural Information Processing Systems. 1994;6:224–231.
11. Ina Khandelwal, Ratnadip Adhikari, Ghanshyam Verma. Time Series Forecasting using Hybrid ARIMA and ANN Models based on DWT Decomposition, International Conference on Intelligent Computing, Communication & Convergence (ICCC-2014)
12. Zurada J. Introduction to artificial neural systems. 0-314-93391-3. West Publishing Co., St. Paul, MN, USA, 1992.
13. Kihoro JM, Otieno RO, Wafula C. Seasonal Time Series Forecasting: A Comparative Study of ARIMA and ANN Models", African Journal of Science and Technology (AJST) Science and Engineering Series 5(2), 41-49.
14. Joarder Kamruzzaman, Rezaul Begg, Ruhul Sarker. Artificial Neural Networks in Finance and Manufacturing", Idea Group Publishing, USA.
15. Mehdi Khashei, Mehdi Bijari. An artificial neural network (p,d,q) model for time series forecasting., Expert systems with applications. 2010;37:479-489.
16. Prapanna Mondal, Labani Shit, Saptarsi Goswami. Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices., International Journal of Computer Science, Engineering and Applications (IJCSEA), 2014 April, 4(2).
17. Satish Kumar, Neural Networks, A Classroom Approach, Tata McGraw-Hill Publishing Company Limited.
18. Shibata R. Selection of the order of an autoregressive model by Akaike's information criterion. Biometrika. 1976;63:117-26.
19. Taskaya T, Casey MC. A comparative study of autoregressive neural network hybrids. Neural Networks. 2005;18:781–789.