FORECASTING THE PRODUCTION OF GROUNDNUT CROP USING BOX-JENKIN'S METHODOLOGY

B. Saidulu, Research Scholar, Department of Statistics, Osmania University.
Dr. M. Raghavender Sharma, Department of Statistics, Osmania University.

Abstract

Statistical Modelling of non-stationary – non-linear statistics has become an enormous challenge in altogether field of the research project. A number of popularly used models square measure ARIMA and ANN. This text presents a comparison of Artificial Neural Network (ANN) and Box-Jenkins methodology for predicting the true production of groundnut crop worth in Telangana. The most important objective of this analysis is to develop a forecasting model to predict the production of the Groundnut crop with high accuracy. During this paper, a statistic forecasting model victimization Box-Jenkins methodology and Artificial Neural Networks was developed for forecasting the yearly production of the Groundnut crop in Telangana. The forecasting performance of the model was evaluated victimization Root mean squared error (RMSE), Mean Absolute percent Error (MAPE). The yearly forecasts counsel that, the production of Groundnut crop with a regular deviation of 15% error measure with the accuracy of 90% for the forecasted period of 10 years.

Keywords: ARIMA, Box-Jenkins Methodology, ANN and MAPE.

1. Introduction

A statistic is the basic object of the study in varied sectors of analysis. Conventionally statistic modeling includes an underlying assumption that there's a linear underlying relationship between the past and future values of the series.

Agriculture is known to be the spine of the Indian Economy for the last many decades for any kind of growth and productivity of all the crops. Groundnut is cultivated in two lakhs hectares across the Telangana region making it one of the major crops of the state. It is widely grown in Mahbubnagar, Warangal, Nalgonda, and Karimnagar Districts. Crop rotation is very important in Groundnut farming, this helps to utilize nutrients efficiently utilization and to reduce soil-borne diseases. Groundnut is more beneficial to human nutrition, and is an important product since it is used in the production of many foods and ranked fourth among oil-seed plants after cotton, soyabean, and sugarcane. Groundnut also has a particular economic value since its oil, kernels, shell, and straw can be used commercially as well as extensively throughout the state of Telangana. The production amounts of the groundnut for Telangana and based on the Indian harvested area. Increasing production would also be an important tool in the development of rural areas of the state by increasing producer revenues. Thus, it is important for Telangana to formulate policies aiming at increasing groundnut production for the future sustainability of the oil industry, export revenues, and food safety. An effort is made during this paper to assess the yearly production of the Groundnut crop in the state of Telangana and to forecast the same for a brief tenure by victimization applied mathematics strategies. The subsequent section presents the results that supported the Box-Jenkins methodology and artificial neural networks.

2. Materials and Methods

BOX-JENKINS METHODOLOGY

During this section, the modeling of groundnut production of Telangana Box-Jenkins methodology is mentioned. The Box-Jenkins procedure relates to the fitting of the associate ARIMA model of the subsequent type for the given set of information and therefore the general kind of ARIMA (p, d, q) model is given by

$$\Phi(B)\nabla^{d}Z_t = \theta(B)a_t$$

$$\text{Where } \Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2 - - - - - \Phi_p B^p$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - - - -\theta_q B^q$$

ISSN: 2278-4632
Vol-12 Issue-10 No.02 October 2022

$$\text{And } \nabla^d = (1 - B)^d$$

We have $B^K Z_t = Z_{t-k} and a_t$ is a white noise process with zero mean and variance σ^2_a.The Box-Jenkins procedure consists of the subsequent four stages. (1) Model Identification, wherever the orders d, p, and q are determined by perceptive the behavior of the corresponding Autocorrelation function (ACF) and Partial Autocorrelation Function (PACF). (2) Estimation: wherever the parameters of the model are a unit calculable
by the most probability methodology. 3) Diagnostic checking by the "Portmanteau Test", where the adequacy of the fitted model is checked by the Ljung-Box datum, applied to the residual of the model. (4) Forecasts area unit obtained from associate degree adequate model victimization minimum mean square error methodology. If the model is judged to be inadequate, stages 1-3 area unit perennial with completely different values of d, p, and q, till associate the adequate model is obtained (Box et al; 1994).

Ann Model

An Artificial Neural Network could be a mathematical model that is impressed by the structure and useful aspects of the biological neural network, a powerful predictive model. Associate degree ANN will estimate any nonlinear continuous function up to any desired degree of accuracy. It is widely used in a range of industries, business engineering, and sciences. It has the power to perfectly predict the longer term and is prime to several call processes in designing, scheduling, purchasing, strategy formulation, policy-making, and providing chain operations.

The characteristics of ANN that build it applicable for predictions are its non-linear structure, flexibility, knowledge-driven learning method, and its ability to estimate method universal functions. Neural networks area units precisely shown to possess the universal sensible approximating potential during which they will accurately approximate several varieties of advanced sensible relationships. This can be a very important and powerful characteristic, as any prediction model aims to accurately capture the useful relationship between the variable to be foreseen and different relevant factors or variables. The mixture of the abovementioned characteristics makes ANN a really general and versatile modeling tool for prediction. Finally, ANNs are unit non-linear models. The very fact that globe systems are unit typically non-linear has led to the event of many non-linear statistical models in the last decade. (Hornik,1993; Ramakrishna et al., 2011).

3. Results

FORECASTING PRODUCTION OF GROUNDNUT CROP USING BOX-JENKINS METHODOLOGY

In this paper, the building of prediction models victimization Box-Jenkins methodology for the yearly production of groundnut crops is mentioned.

The data on the yearly production of groundnut were collected from the year 1979 to 2019 (40 years) from the Bureau of Economics and Statistics. The yearly production of groundnut crops from 1979 to 2010 was used for model building and therefore the yearly groundnut crop production from 2011 to 2019was used for model validation. The prediction models for the prediction of the yearly production of groundnut crops were developed victimization Box-Jenkins methodology and Artificial Neural Networks.

The yearly production of the groundnut crop varied with an average production of thirty-two thousand. The subsequent chart shows the time trend of the yearly production of the groundnut crop from 1979 to 2019. The yearly production of the groundnut crop shows a non-stationary time series (Fig 1). The average Groundnut production was comparatively low in the year 2015 and high in 2017 due to the low and high rain fall during the above years (Fig 1).

ISSN: 2278-4632
Vol-12 Issue-10 No.02 October 2022

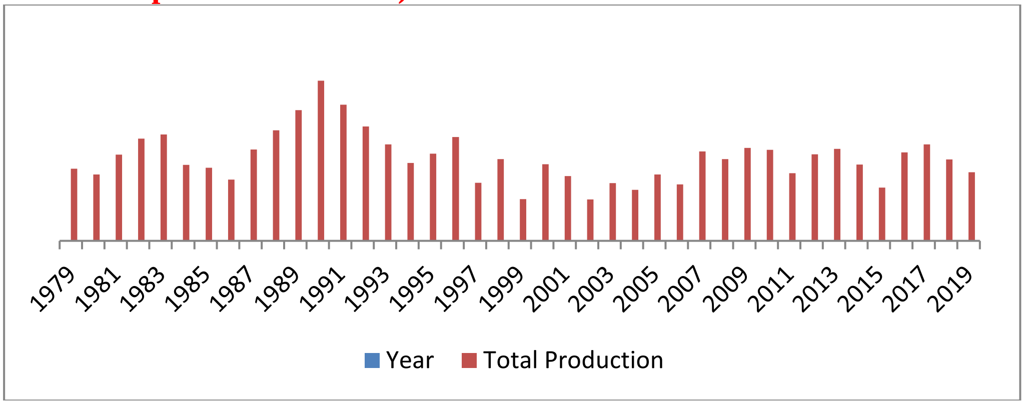

Figure1*: The yearly average for the production of the Groundnut Crop (in tones)*

The python skilled creator was accustomed to determining the most effective ARIMA model for the prediction of groundnut production, as this plan automatically determines and estimates the best-fitting ARIMA for one or additional variable series, therefore eliminating the necessity to spot an applicable model through the trial-and-error methodology. It is discovered that the ARIMA (1, 1, 0) model fits the data well and therefore the same is tested on the validation set with the original series and order-wise differencing for ACF and PACF. The model parameters square measure is given in the following Table 1 and Table 2

Original Series, Differencing, and ACF (Auto Correlation Function)

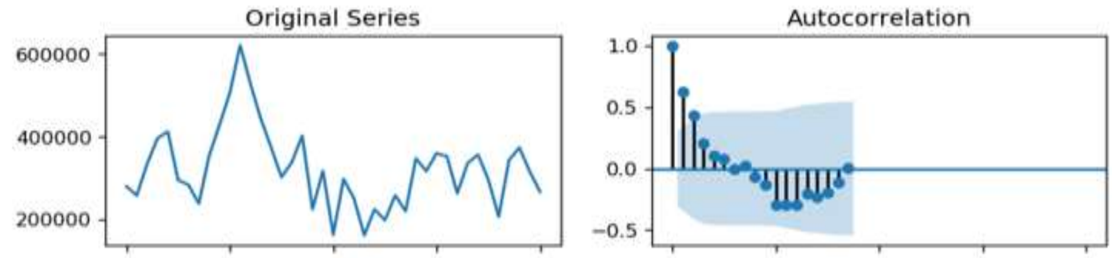

Figure2: Time series of Groundnut production (tonnes) in Telangana Original Series with ACF

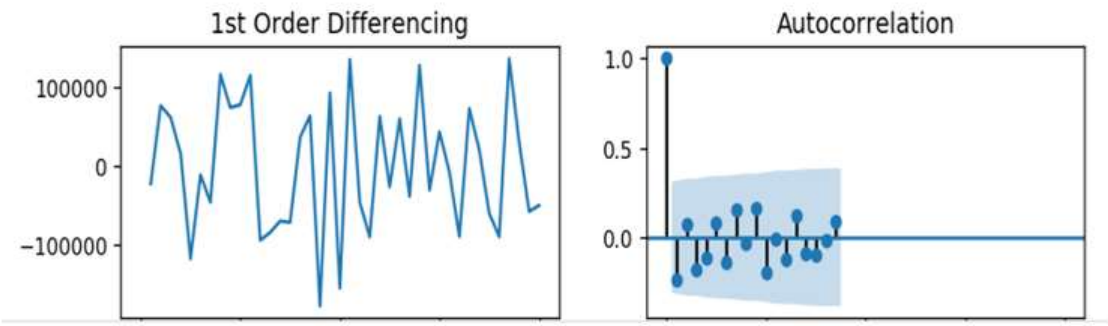

Figure3: Time series of Groundnut production (tones) in Telangana 1 st Order differencing with ACF

Original Series, Differencing, and PACF (Partial Auto Correlation Function)

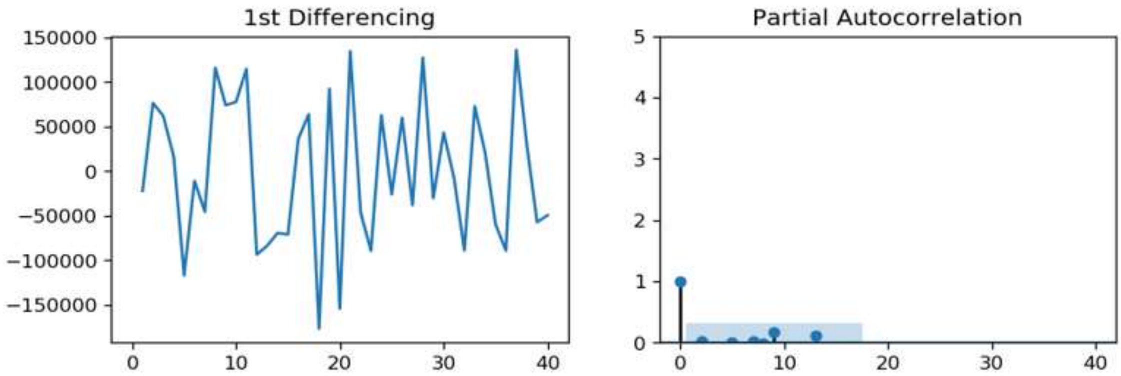

Figure4: Time series of Groundnut production (tones) in Telangana, 1 st Order differencing with PACF

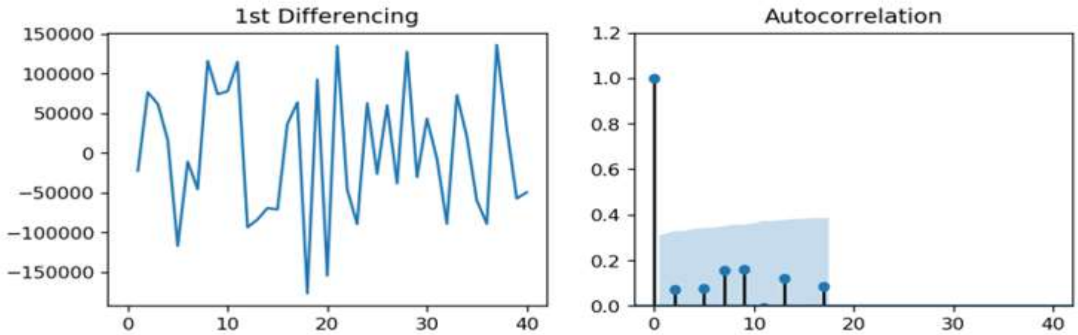

Figure5: Time series of Groundnut production (tones) in Telangana 1 st Order differencing with ACF

Table 1. ARIMA Model Parameters

```
                              ARIMA Model Results
==============================================================================
Dep. Variable:                D.Total   No. Observations:                   40
Model:                 ARIMA(1, 1, 0)   Log Likelihood                -508.023
Method:                       css-mle   S.D. of innovations          79294.432
Date:                Fri, 18 Mar 2022   AIC                           1022.046
Time:                        18:23:18   BIC                           1027.113
Sample:                             1   HQIC                          1023.878

=================================================================================
                    coef    std err          z      P>|z|      [0.025      0.975]
---------------------------------------------------------------------------------
const          -337.4250   1.02e+04     -0.033      0.974   -2.03e+04    1.96e+04
ar.L1.D.Total    -0.2361      0.152     -1.551      0.121      -0.535       0.062
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1           -4.2347           +0.0000j            4.2347            0.5000
-----------------------------------------------------------------------------
```


ISSN: 2278-4632
Vol-12 Issue-10 No.02 October 2022

```
                              ARIMA Model Results
==============================================================================
Dep. Variable:                D.Total   No. Observations:                   40
Model:                 ARIMA(0, 1, 1)   Log Likelihood                -508.004
Method:                       css-mle   S.D. of innovations          79245.753
Date:                Fri, 18 Mar 2022   AIC                           1022.008
Time:                        18:25:36   BIC                           1027.075
Sample:                             1   HQIC                          1023.840

=================================================================================
                    coef    std err          z      P>|z|      [0.025      0.975]
---------------------------------------------------------------------------------
const          -337.4250   9396.917     -0.036      0.971   -1.88e+04    1.81e+04
ma.L1.D.Total    -0.2571      0.173     -1.485      0.138      -0.596       0.082
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
MA.1            3.8898           +0.0000j            3.8898            0.0000
-----------------------------------------------------------------------------
```

Hence, the fitted model for the forecasting of Groundnut production in the state of Telangana is ARIMA (0.1.1) and ARIMA (1,1, 0)

$$\nabla^1 Z_t = (1 + 0.043B^8 - 0.112B^{13})a_t.$$

The adequacy of the model was checked by exploitation Ljung-Box q test statistic and therefore the same is discovered that Q=11.376 at 20 degrees of freedom. The corresponding p-value of the letter check datum is 0.85 and which is far larger than 0.05, hence, the null hypothesis of the adequate model was accepted and therefore the given ARIMA (0,1,1) model may be an appropriate model for the prediction of the production of groundnut. Similarly, a synthetic neural networks model was developed for the prediction of the production of groundnut using python.

Artificial Neural Networks Model

ANN model may be a feed-forward neural networks model having one input layer, one hidden layer Associate in Nursing, and an output layer. The standardized values of previous observation (Lag-1 or Zt -1) were used as Associate in nursing input for the one-step a head prediction of the silver costs during this ANN model. The hidden layer consists of two hidden neurons to capture the nonlinearity element within the statistic. The hyperbolic tangent function is employed as an activation function in the hidden layer and the identity is used as an activation function in the output layer. The subsequent Figure 6 shows a typical feed-forward neural network used for the prediction of Groundnut production in Telangana state. The ANN model was trained exploitation backpropagation rule until the error measures of the testing sample are smaller than the coaching sample on a trial-and-error basis.

Figure6: Feed forward neural network prediction model for the yearly production of groundnut crops.

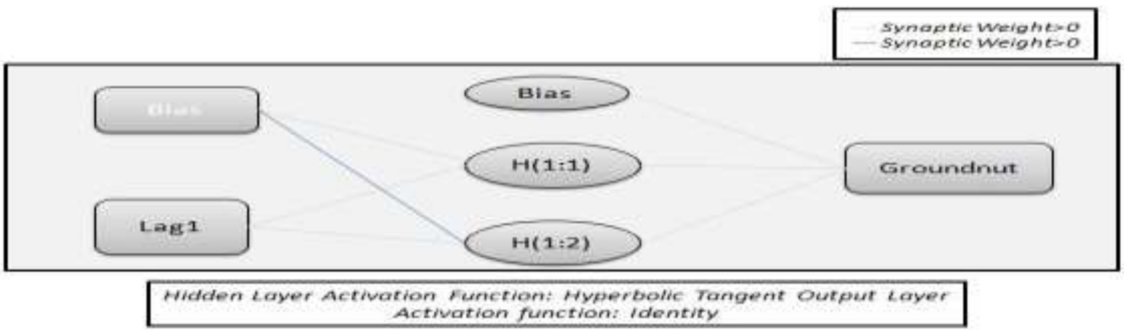

Table 2. Feed forward Neural Networks Model Parameters

Predictor		Predicted		
		Hidden Layer 1		Output Layer
		H(1:1)	H(1:2)	S_t
Input Layer	(Bias)	0.018	-0.065	
	lag1	0.142	0.284	
Hidden Layer 1	(Bias)			0.134
	H(1:1)			2.302
	H(1:2)			3.107

Hence the ANN Model is here.
$L=(Z_{t-1}\ 39717.32)/3613.655$; $h_1=\tanh(0.018+0.142*L)$; $h_2=\tanh(-0.065+0.284*L)$;
$S_t=0.134+2.302*h1+3.071*$ h2and $Z_t= 39715.91+3612.785*S_t$

Comparison of Arima and Ann Models

The predictions from the two models were compared in the training sample and testing samples supported the mean absolute error, root means square error and mean absolute percentage errors. The subsequent Table 3 presents the error measures from the ARIMA and ANN prediction models.

Table 3. Comparison of the forecasting performance of ARIMA and ANN models

Measure	Training Sample		Testing Sample	
	ARIMA	ANN	ARIMA	ANN
MAE	471.13	468.21	462.11	461.37
RMSE	546.03	539.14	545.21	531.34
MAPE	0.15	0.14	0.17	0.21

The ANN model has comparatively lower error measures in the testing sample as compared to the ARIMA Model. The ANN fits well stronger than ARIMA within the training and testing samples not show stronger performance in the testing sample. The following figure 7 shows the out-of-sample forecasts supported by ARIMA and ANN models. The ANN model forecasts show a similar trend to the original costs, whereas the ARIMA forecasts show a linear trend over time. The forecasts from the 2 models are presented in the following table.

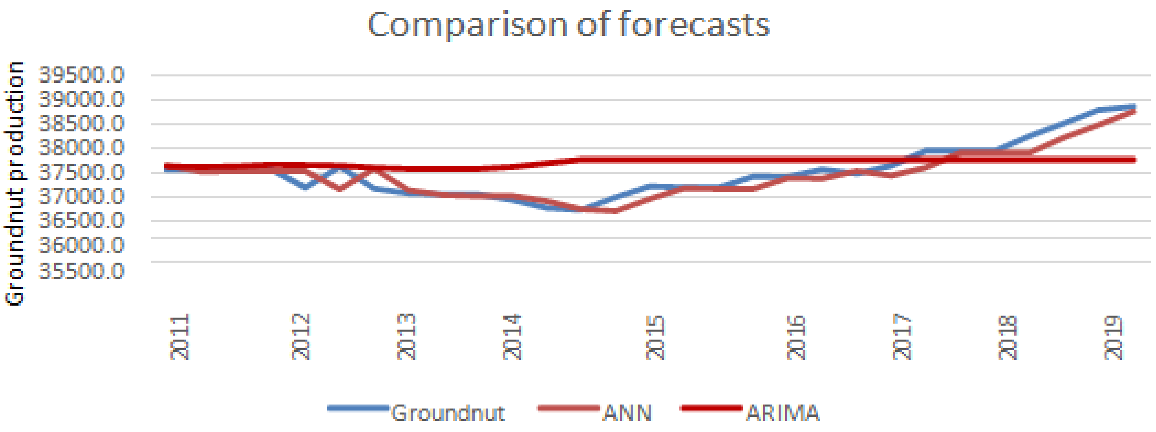

Figure 7. Comparison of forecasts for Groundnut production in Telangana state (in tonnes)

Table 4. Out-of-sample forecasts of Groundnut production in Telangana state using ARIMA and ANN Models

	Production		
Year	**(in tonnes)**	**ANN**	**ARIMA**
2020	321837.4878	321836.2	321811.2
2021	322911.3250	322910.1	322885.1
2022	324609.7692	324608.5	324583.5

2023	324384.2368	324383.5	324358.2
2024	322478.9459	322477.7	322452.7
2025	320022.1944	320020.9	319995.9
2026	320774.4857	320773.2	320748.2
2027	321897.6765	321896.4	321871.4
2028	324472.5152	324471.3	324446.3
2029	323583.4063	323582.2	323557.2
Mean	322697.20	322695.95	322670.75
SD	1590.437815	1590.438	1590.438

Discussion and Conclusion

The forecasts recommend that the ANN model predicts well the groundnut production in Telangana State as compared to the ARIMA model. ARIMA model provides only linear trends whereas the ANN model presents the nonlinear fluctuations within the forecasts.

References

[1] Box.G.E.P. Jenkins, G.M. and Reinsel, G.C. (1994). Time Series analysis forecasting and control, 3rd ed, Engelalwood cliffs, N.J. PrenticeHall,
[2] BROCKWELL P.J., DAVIS R.A (1996), An introduction to time series and Forecasting, Springer, Verlag.
[3] C.Gershenson, "Artificial Neural Networks for Beginners" Networks, 2003., Vol. cs.NE/0308, pp.8.
[4] Haykin, S.S., (1999), "Neural Networks: A Comprehensive Foundation", Upper Saddle River, N.J., PrenticeHall.
[5] Hornik, K, (1993), Some new results on neural network approximation, Neural Networks, 6, 1069-1072.
[6] K.Murali Krishna, Dr.M.Raghavender Sharma and Dr.N.Konda Reddy, Forecasting of silver prices using Artificial Neural Networks. JARDCS, Volume 10, 06 issue 2018.R. J. Vidmar. (1992, Aug).
[7] Makridakis S., Wheel Wright. S.C., Hyndman R.J., 2003, Forecasting Methods and Applications, John Wiley &Sons.
[8] Raghavender, M. (2009), Forecasting paddy yield in Andhra Pradesh using seasonal Time Series Model, Bulletin of Pure and applied sciences.
[9] Ramakrishna, R., Naveen Kumar, B. and Krishna Reddy.M. (2011), Forecasting daily electricity load using neural networks, International Journal of Mathematical Archive, Vol.2, 1-11.
[10] Tang, Z., Almeida, C.D. and Fishwick, P.A., 1991, "Time Series Forecasting using Neural Networks Vs. Box-Jenkins Methodology", Simulation, Vol.57, No.5,303-310.
[11] The Economic Times, 3rdDec,2010.