Journal
Since 1954

The Andhra Agric. J 62(1):185-189, 2015

Comparative Performance of Time Series Models on Groundnut in North Coastal Zone of Andhra Pradesh

G Srinivas, V Srinivasa Rao, G Raghunadha Reddy and V V Krishna
Department of Agricultural Statistics, Agricultural College, Bapatla -522 101

ABSTRACT

The present study was carried out to compare the performance of time series models of area, production and productivity of groundnut in North coastal zone of Andhra Pradesh. It has been undertaken to estimate the future trends and to fit the adequate model for the future projections by 2020 AD by using time series data from 1971 to 2011. Different linear, non linear and time series models were fitted to the area, production and productivity of groundnut and the best-fitted model was chosen based on least Mean Absolute Percent Error (MAPE) value and highest R^2 value for future projections. The result showed that there would be slight decrease in the future projection of groundnut area and production i.e. 49.71 thousand hectares and 64.89 thousand tonnes by 2020 AD. The linear model revealed that there would be substantial increase in productivity of the groundnut i.e. 3379 kg ha^1 by 2020 AD.

Key words :ARIMA, ARCH, GARCH, Exponential smoothing, Linear trend, MAPE, R^2.

Groundnut is the sixth most important oilseed crop in the world. It contains 48-50% oil and 26-28% protein and is a rich source of dietary fiber, minerals and vitamins. Groundnut is grown on 26.4 M ha worldwide with a total production of 37.1 million Mt and an average productivity of 1.4 MT./ha (FAO, 2003). In India it is cultivated on 7.5 M ha, with an annual production of 8Mt. More than 9 million small farmers depend on this and other small oil seed crops for their livelihoods. In this study an attempt has been made to assess the growth rates in area, production and productivity of groundnut crop in North coastal zone of Andhra Pradesh by using 41 years of data from 1971 to 2011. Besides, the projections were also estimated up to 2020 AD. The data of the study for a period of 41 years (1971 to 2011) in North coastal zone of Andhra Pradesh pertaining to area, production and productivity of groundnut were collected from the statistical abstracts of Andhra Pradesh and web resources http://www.Indiastat.com.

The future projections of area, production and productivity of groundnut crop in North coastal zone up to 2020 AD were estimated upon the best fitted time series model used for fitting the trend equations. The trend equations were fitted by using different linear, non linear and time series models. Among these models the model with least MAPE and highest R^2 was considered as the best fitted model for the projection purpose.

MATERIAL AND METHODS

Linear model:

Most naturally-occurring time series data are not at all stationary, they exhibit various kinds of trends i.e, cycles, seasonal and irregular patterns. Linear trend model is one of the simplest method used to fit the time series data. The linear trend model is:

$$Y(t) = \alpha + \beta t$$

Where 't' is the time index. The parameters alpha and beta (the "intercept" and "slope" of the trend line) are usually estimated by using least square method in which Y is the dependent variable and the time index t is the independent variable.

Exponential smoothing:

The time series data is Y_1, Y_2…... Y_t. to forecast the next value of time series Y_{t+1} that is yet to be observed with forecast for Y_t denoted by F_t. then the forecast F_{t+1} is based on weighing the most recent observation Y_t with a weight value α and weighing the most recent forecast F_t with a weight of (1-α). The weight is between 0 and 1. Thus the forecast for the period t+1 is given by

$F_{t+1} = F_t + \alpha (Y_t - F_t)$
Where α is a smoothing constant.

TIME SERIES MODELS:
Auto Regressive Process of order (p) is:
$Y_t = \mu + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \ldots + \emptyset_p Y_{t-p} + \varepsilon_t$

Moving Average process of order (q) is:
$Y_t = \mu - \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t$

And the general form of ARIMA model of order (p, d, q) is
$Y_t = \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \ldots + \emptyset_p Y_{t-p} - \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t$

Where, Y_t is the dependant variable, ε_t s are independently and normally distributed with zero mean and constant variance for t = 1, 2…..n; d is the fraction differenced, Ø's and θ's are coefficients to be estimated.

ARIMA: Four steps
Identification:

This step involves determining the values of p, d and q. When a time series data is non-stationary, it can be often made by stationary by taking first differences of the series i.e., creating a new time series of successive differences i.e., (Y_t-Y_{t-1}), this is first order differencing (d=1). If first difference does not convert the series to stationary, then go for second order differencing (d=2). Values of p and q are determined by Autocorrelation function (ACF) and Partial Autocorrelation function (PACF), when the data is stationary only.

Estimation of parameters:

In this step, the precise estimates of parameters of the model are obtained by least-squares. Here, Standard computer package SAS 9-3 was used for finding the estimates of relevant parameters.

Diagnostic checking:

The estimated model must be checked to verify whether it is adequately represents the series or not. For evaluating the adequacy of AR, MA and ARIMA processes, various reliability statistics are available, but Mean Absolute Percentage Error (MAPE) and R^2 are chosen as diagnostic checks in the present study.

Auto Regressive Conditionally Heteroscedastic (ARCH) model:

If the series is being high volatility in the data to capture the volatility ARCH model can be used. In this model, ε_t denote the error terms (return residuals, with respect to a mean process) i.e. the series terms. These ε_t are split into a stochastic piece Z_t and a time-dependent standard deviation ε_t characterizing the typical size of the terms so that $\sigma_t = \sigma_t Z_t$ The random variable Z_t is a strong process. The series σ_t^2 is modeled by Engle R.F (1982).

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \cdots + \alpha_q \epsilon_{t-q}^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2$$

Where $\alpha_0 > 0$ *and* $\alpha_i \geq 0, i > 0$

Generalized Autoregressive Conditional Heteroscedasticity (GARCH) Process:

Bollerslev (1986) developed a GARCH (p, q) model which takes both the components of autoregressive and moving average in the form of the heteroscedasticity variance, GARCH model was the extension of the ARCH model which was proposed by Engle (1982).

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \cdots + \alpha_q \epsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$$

MODEL SELECTION:

The choice of the trend equation amongst the available alternatives is very crucial. Many researchers use coefficient of multiple determination (R^2) and MAPE as the criterion of model selection are given below

R^2-Criteria

$$R^2 \equiv 1 - \frac{SS_{\text{err}}}{SS_{\text{tot}}}. \quad \text{Or} \quad R^2 = \frac{SS_{\text{reg}}}{SS_{\text{tot}}}.$$

Mean Absolute Percentage Error (MAPE)

$$\mathrm{M} = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|.$$

Table 1. Different models on area, production and productivity of groundnut in North coastal zone of Andhra Pradesh.

Groundnut						
	criteria	Linear	Exp.Smoothing	AIMA	ARCH	GARCH
Area	R^2	0.28	0.80	0.83	0.81	0.81
	MAPE	23.27	9.64	9.52	10.20	10.20
Production	R^2	0.17	0.50	0.52	0.52	0.51
	MAPE	27.64	18.14	18.11	19.45	19.21
Productivity	R^2	0.17	0.07	0.08	0.06	0.06
	MAPE	9.73	9.83	10.34	10.56	10.56

Table 2. Forecasted values of area, production and productivity of groundnut in North coastal zone of Andhra Pradesh.

Year	Predicted Area ('000 hectares	Predicted Production ('000 tonnes)	Predicted Productivity (kg ha^{-1})
2012	57.12	73.16	3272
2013	51.83	64.94	3285
2014	51.54	64.88	3298
2015	49.12	64.89	3312
2016	49.89	64.89	3325
2017	49.65	64.89	3338
2018	49.72	64.89	3352
2019	49.70	64.89	3365
2020	49.71	64.89	3379

Where, A_t is the actual value and F_t is the forecast value. The difference between A_t and F_t is divided by the actual value A_t again.

RESULTS AND DISCUSSION

To understand performance of the time series models of area, production and productivity of groundnut in North coastal zone during the period 1971 to 2011, time series data was analyzed by using linear, non linear and time series models *viz.,* linear, exponential smoothing, ARIMA, ARCH, GARCH models were fitted and the results were presented in table 1.

Groundnut area

The average area of the groundnut crop for the entire period of study was 130.77 thousand hectares. It can be seen that the area of groundnut was in declining trend from1994 onwards. The area of the groundnut crop ARIMA (1,1,4) model with least MAPE value of 9.52 and highest R^2 value of 0.83 was considered as the best fit among all the models considered.

Groundnut production

The average production of the groundnut crop for the entire period was 127.1 thousand tones, ARIMA (1,1,3) model was considered as the best fit among all the models with least MAPE value of 18.11 and highest R^2 value of 0.52.

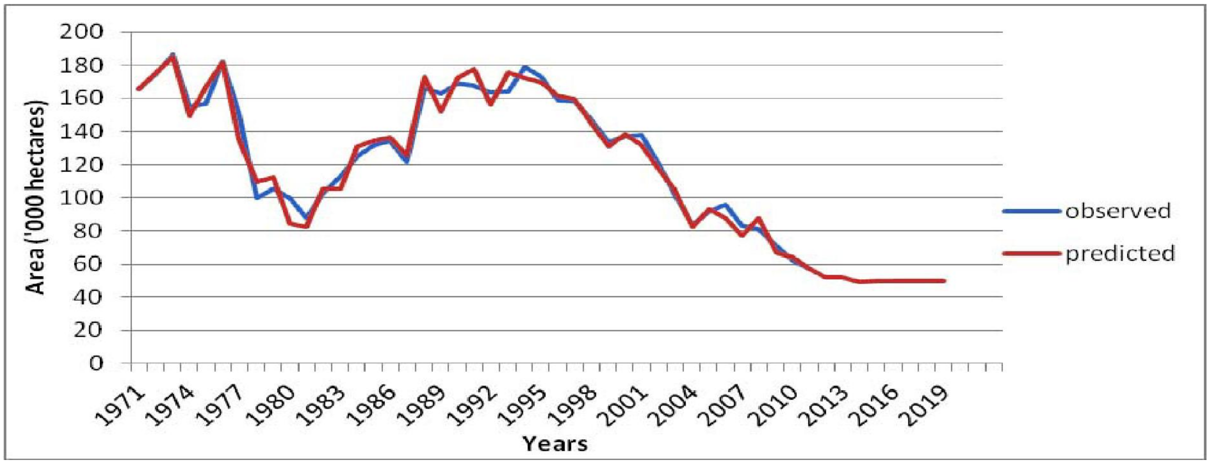
Groundnut Productivity

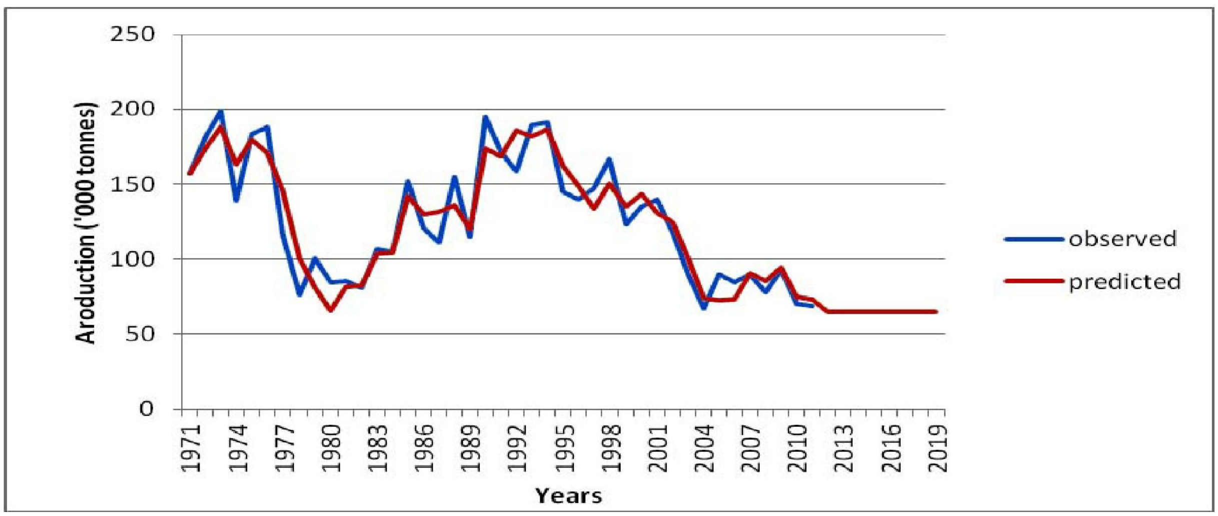
The average productivity of groundnut crop for the entire period was 2990.19 kg ha^{-1}. The linear model is having with least MAPE value of 9.73 and highest R^2 value of 0.17 was considered as the best fit among all the models.

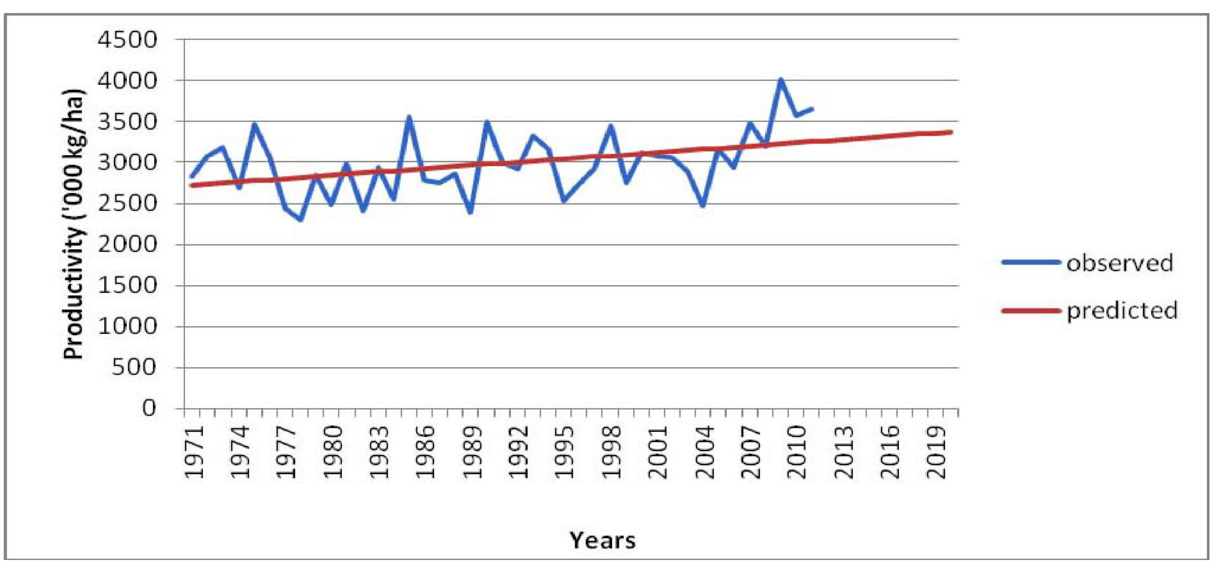
Projections

The future projections of area, production and productivity of groundnut by 2020 AD were

Projections of groundnut area, production and productivity in north coastal zone of Andhra Pradesh







calculated based on the selected models. The area of the groundnut crop, ARIMA (1,1,4) model was identified as the best fitted model with least MAPE value of and highest R^2 value. The projected area of groundnut by 2020 AD would be 49.71 thousand hectares and it was presented in Table 2.

For production ARIMA (1,1,3) model was identified as the best model for future projections as it has highest R^2 value with the least MAPE value for long term projections. The projected production of groundnut by 2020 A.D would be 64.89 thousand tones.

Linear model was found to be best fitted model for the projection of groundnut productivity as it has highest R^2 and least MAPE values. So, the productivity of groundnut projected by 2020 AD would be 3379 kg ha^{-1}.

CONCLUSIONS

The study was under taken to obtain a suitable time series models for forecasting the area, production and productivity of the groundnut crop in the North coastal zone of Andhra Pradesh were projected up to 2020 AD. The average area under groundnut was 130.77 thousand hectares with an average production of 127 thousand tones and productivity of 2990.19 kg ha^{-1}. The ARIMA model revealed that there would be decrease in the future projection of groundnut area and production i.e. 49.71 thousand hectares and 64.89 thousand tons by 2020 AD. The linear model revealed that there would be substantial increase in productivity of the groundnut i.e. 3379 kg ha^{-1} by 2020 AD.

Selection of an appropriate model is made by comparing the highest R^2 and lowest MAPE values. When there is volatility in the data linear, exponential smoothing and ARIMA models may not capture its trend. Hence ARCH and GARCH models were used to capture the volatility in the data.

LITERATURE CITED

Balangammal D, Ranganathan C R and Sundaresan R 2000 Forecasting of agricultural scenario in Tamil Nadu- A time Series analysis. *Journal of Indian Society of Agricultural Statistics*, 53(3): 273-286.

Bollerslev T 1986 Generalized Auto Regressive Conditional Heteroscedasticity. *Journal fof Econometrics,* 31: 307-327.

Devi B I, Raghuram P, Lavanya T and Suman V 2011 Forecasting of prices of sunflower and groundnut in Andhra Pradesh- An application of ARIMA model. *The Andhra Agricultural Journal*, 58(3): 368-370.

Engle R F 1982 Auto Regressive Conditional Heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica,* 50 (4): 987-1007.

Iqbal N, Bakhsh K, Maqbool A and Ahamad A S 2005 Use of the ARIMA model for forecasting wheat area and production in Pakistan. *Journal of Agriculture and Social Sciences*, 1(2): 120-122.

Irfan M, Irfan M and Tahir M 2011 Modelling of province wise yield of rice crop in Pakistan using GARCH model. *International Journal of Science and Technology,* 6(1): 224-228.

Muhammad F, Bashir M and Ahmad S 1992 Forecasting rice production in Pakistan using ARIMA models. *Journal of Animal Plant Science,* 2: 27-31.

Nochai R and Nochai T 2006 ARIMA model for forecasting oil palm price. http://www. titida@s_t. au. Ac. Th.

Pankratz A 1983 *Forecasting with Univariate Box – Jenkins Models. Concepts and Cases.* John Wiley and sons, New York.

Reddy V M 1978 Growth rates. *Economic and Political Weekly,* 132: 806-812.

Saeed N, Saeed A, Zakria M and Bajwa T M 2000 Forecasting of wheat production in Pakistan using ARIMA models. *International Journal of Agriculture and Biology*, 2(4): 352-353.

(Received on 12.09.2013 and revised on 07.08.2013)