Economic Affairs, Vol. 62, No. 1, pp. 121-127, March 2017
DOI: 10.5958/0976-4666.2017.00035.3

Forecasting of Pearl Millet Productivity in Gujarat Under Time Series Framework

Satvinder Kour[1], U.K. Pradhan[2], Ranjit Kumar Paul[3] and P.R. Vaishnav[4]

[1 & 4] Department of Agricultural statistics, B.A. College of Agriculture, Anand Agricultural University, Anand, Gujarat, India,
[2&3] ICAR-Indian Agricultural Statistics Research Institute, New Delhi-110012, India

Corresponding author: rani.fardeen@gmail.com

Received: 05-11-2016 Revised: 12-01-2017 Accepted: 22-02-2017

ABSTRACT

Pearl millet (*Pennisetumglaucum*) is one of the most extensively cultivated cereals in the world, ranking fourth after rice, wheat and sorghum. In Gujarat, for last 20 years, it has been noticed that the area under pearl millet production has been decreased gradually although productivity increase. The importance of pearl millet productivity forecasting is more relevant in semi-arid state like Gujarat where the precipitation is confirmed to short period of four months. In this paper, we have applied ARIMA model for forecasting of productivity of pearl millet of Gujarat. In the present study, time series data of pearl millet productivity (Kg./ha)of Gujarat for 52 years from 1960-61 to 2011-12 were collected from Directorate of Agriculture, Gandhinagar, Gujarat and partially from Directorate of Economics and statistics. The ARIMA model is validated on the basis of relative mean absolute prediction error (RMAPE), Mean absolute deviation (MAD) and root mean square error(RMSE) values. It may be noted that ARIMA (0, 1, 1) model performs quite satisfactorily as the RMAPE value is less than 6 percent.

Keywords: *ARIMA, Forecast, Pearl millet productivity*

Agriculture is backbone of Indian economy, among various production commodities of basic importance, agricultural production is one which is subject to wide and irregular fluctuations of output. Agriculture contributing about 40 per cent towards the gross national product and provide livelihood to about 70 per cent of the population. The share of agriculture in total GDP is 18 percent in 2013-14 (national income, Economic survey 2014-15, by the CSO). In India, it is possible to cultivate large number of crops due to diverse climatic conditions. India has made considerable progress in agriculture since independence in terms of production, productivity and area with respect to many crops. Crop failure on account of drought or flood will have a severe repercussion not only on the country's economy but also on food security. Agriculture now-a-days has become highly input and cost intensive area without judicious use of fertilizers and plant protection measures, agriculture no longer remains as profitable as before because of uncertainties of weather, production, policies, prices etc that often lead to losses to the farmers. Crop productivity forecast provided useful information to farmers, marketers, government agencies and other agencies and useful in formulation of policies regarding stock, distribution and supply of agricultural produce to different areas in the country.

Under the changed scenario today, forecasting of various aspects relating to agriculture are becoming essential. The forecast of crop production and productivity well in advance has its own importancein every region in order to planning & policy making for future food distribution, In policy decision regarding export and import, pricing,and for exercising several administrative measures for storage and marketing of agriculture commodities. Forecasting is the process of making statements about events whose actual outcomes (typically) have not yet been observed. The importance of timely and reliable forecasts of area and productivity of principal crops need not be over-emphasized for the country like India where the economy is mainly based on agricultural production. The primary advantage of forecasting is that it provides various stakeholders with valuable information that can be used to make decisions about the future. However, statistical techniques employed should be able to provide objective crop forecasts with reasonable precisions well in advance before harvests for taking timely decisions. Various approaches have been used for forecasting such agricultural systems. A lot of literatures can be found regarding applications of time series models in forecasting agricultural produce or prices. The most widely used technique in this regard has been the autoregressive integrated moving average (ARIMA) model. To see some the applications in agriculture one may refer to Paul *et al.* (2014), Paul and Sinha (2016). Paul and Das (2010, 2013) applied ARIMA model for modelling and forecasting of Inland fish production in India as well as fish landing in Ganga basin. Paul *et al.* (2013) applied Seasonal ARIMA (SARIMA) model for forecasting of total meat export from India One advantage of the ARIMA approach is that it is able to provide a good understanding of the system. This model has been dominating time series analysis for several decades.

Pearl millet (*Pennisetumglaucum*) is one of the most extensively cultivated cereals in the world, ranking fourth after rice, wheat and sorghum in terms of area planted to these crops. It is a principal cereal cultivated in drought-prone semi-arid regions of Africa and the Indian subcontinent, mostly for food uses. Pearl millet is the most widely grown types of millet. It is also known as bajra, pearl millet is a popular crop for food and fodder grown under limited moisture supply. Pearl millet crop has wide adaptability to local environments. It is a hardy crop and can be grown in areas which are very hot and dry and on soils too poor for crops like maize and sorghum. Pearl millet is considered more efficient in utilization of soil moisture and has a higher level of heat tolerance than sorghum and maize. Due to higher potentiality, it requires more nutrients and known as exhaustive crop and is primarily consumed in the states of Haryana, Rajasthan, Gujarat and Madhya Pradesh. India is the largest producer of this crop, both in terms of area (9.1 m ha) and production (7.3 m t), with an average productivity of 780 kg/ha (Anon. 2010-11). Among the states, Gujarat ranked third in area after Rajasthan and Maharashtra. Gujarat has 101 lakh hectare of net sown area and 128 lakh hectare of total cropped area with 8 agro climatic zones. The total area under pearl millet cultivation in Gujarat is (1.07m ha). In 2010-11, production of Pearl millet was (.60 m t), Productivity of Pearl millet during the period was 1720 kg/ha (Anon. 2010-11). In Gujarat for last 20 years, it has been noticed that the area under pearl millet production has been decreased gradually although productivity increase, which can be clearly shown from the figure 1.

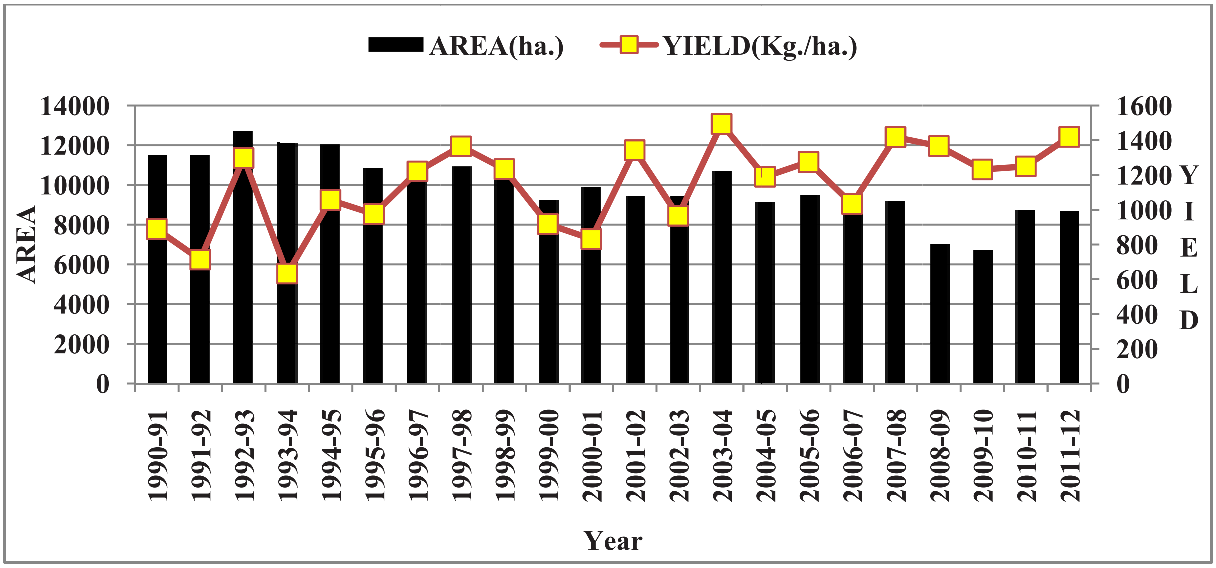


Fig. 1: Graphical presentation of 20 years pearl millet Area and Productivity data

The pre-harvest estimates of crop yield are considered mainly as an aid to conjecture the final production and therefore, sufficient attention needs to be paid towards their improvement. The forecast of crop yield in every region in order to planning & policy making future for food providing distribution, pricing and also its import and export. The importance of pearl millet productivity forecasting is more relevant in semi-arid state like Gujarat where the precipitation is confirmed to short period of four months. In this paper, we have applied an ARIMA model for forecasting of productivity of pearl millet of Gujarat and trend of production over the five decades has been studied.

Data and Methodology

In the present study, pearl millet productivity time series data(Kg./ha)of Gujarat for 52 years from 1960-61 to 2011-12 were collected from Directorate of Agriculture, Gandhinagar, Gujarat and partially from Directorate of Economics and statistics. From the year 1960-61 to 2007-08 was used for model development, and data from 2008-09 to 2011-12 were used for model validation purpose. The SAS 9.4 statistical software package has been used for data analysis.

Trend analysis

Linear trend analysis of area and productivity of pearl millet, it has been clear that the trend over year for both area and productivity, found to be significantly negative and positive respectively. But the issue is critical as area decreases at a faster rate whereas productivity increases at an average rate and overall production decreases. The parameter estimates along with standard error shown in table1.

Table 1. Trend analysis of Area and Productivity

Trend analysis for Area Parameter Estimates				
Variable	DF	Parameter Estimate	Standard Error	Pr > |t|
Intercept	**1**	17399.00	503.54	<.0001
Trend	**1**	-178.05	16.53	<.0001
Trend analysis for Productivity Parameter Estimates				
Variable	DF	Parameter Estimate	Standard Error	Pr > |t|
Intercept	**1**	441.85	55.64	<.0001
Trend	**1**	17.02	1.83	<.0001

Autoregressive integrated Moving Average (ARIMA) Model

A generalization of ARMA models which incorporates a wide class of non- stationary time-series is obtained by introducing the differencing into the model. ARIMA econometric modelling takes into account historical data and decomposes it into an autoregressive (AR) process where there is a memory of past events and integrated (I) process which accounts for stabilizing or making the data stationary, making it easier to forecast, and a moving average (MA) of forecast errors, such that the longer the historical data, the more accurate forecast will be as it learns from over time. The simplest example of a non-stationary process which reduces to a stationary one after differencing is Random Walk. A process $\{y_t\}$ is said to follow an Integrated ARMA model, denoted by ARIMA (*p*, *d*, *q*), if is ARMA (*p*, *q*). The model is written as

$$\varphi(B)(1-B)^d y_t = \theta(B)\varepsilon_t$$

Where

$$\varphi(B) = 1r\varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p$$

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

$\varepsilon_t \sim WN(0, \sigma^2)$, *WN* indicating White Noise. *B is* the backshift operator such that $By_t = y_{t-1}$,

The integration parameter d is a non-negative integer. When $d = 0$, ARIMA (*p*, *d*, *q*) model reduces to ARMA (*p*,*q*) model. The ARIMA methodology is carried out in three stages, viz. identification, estimation and diagnostic checking. Parameters of the tentatively selected ARIMA model at the identification stage are estimated at the estimation stage and adequacy of tentatively selected model is tested at the diagnostic checking stage. If the model is found to be inadequate, the three stages are repeated until satisfactory ARIMA model is selected for the time-series under consideration. An excellent discussion of various aspects of this approach is given in Box *et al.* (2007). Most of the standard software packages, like SAS, SPSS and E-Views contain programs for fitting of ARIMA models.

Estimation of parameters

At the estimation stage parameters are estimated for the ARIMA model tentatively chosen at the identification stage. Estimation of parameters for ARIMA model is generally done through iterative least squares method. The software package SAS has been used for fitting of ARIMA models. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) values for ARIMA model has been computed by:

$AIC = T\log(\sigma^2) + 2(p+q+1)$ and

$BIC = T\log(\sigma^2) + (p+q+1)\log T$

Where T denotes the number of observations used for estimation of parameters and denotes the Mean square σ^2 error.

Diagnostic-checking

At this stage testing is done to see if the estimated model is statistically adequate *i.e.* whether the error terms are white noise which means error terms are uncorrelated with a zero mean and constant variance. For this purpose Ljung-Box test is applied to the residuals after fitting a model. The null hypothesis is that the series is white noise, and the alternative hypothesis is that one or more autocorrelations up to lag m are not zero. The test statistic is given by:

$$Q^* = T(T+2)\sum_{k=1}^{m}\frac{r_k^2}{T-k}$$

where T is the number of observations used to estimate the model and m is the number of lags included for computing the test statistic. The statistic Q^* approximately follows a Chi-squared distribution with *(T-k)* degrees of freedom, where k is the number of parameters estimated in the ARIMA model and r_k is the autocorrelation function of residual at lag k.

RESULTS AND DISCUSSION

In present investigation, time series data (Kg./ha) for pearl millet productivity of Gujarat for 52 years from 1960-61 to 2011-12 are considered. For model development data from the year 1960-61 to 2007-08 was used, and data from 2008-09 to 2011-12 were used for model validation purpose. The SAS 9.4 statistical software package has been used for data analysis.

One of the important assumptions in ARIMA model fitting is that the data should be stationary. But from the fig.-2, it can easily be seen that though the trend is stationary, but the mean is not stationary.

In the present investigation we have used Augmented Dickey Fuller (ADF) testfor testing the stationarity. It is found that after first differentiating the data has become stationary at 1% level of significance. The result of ADF test is reported in table 2.

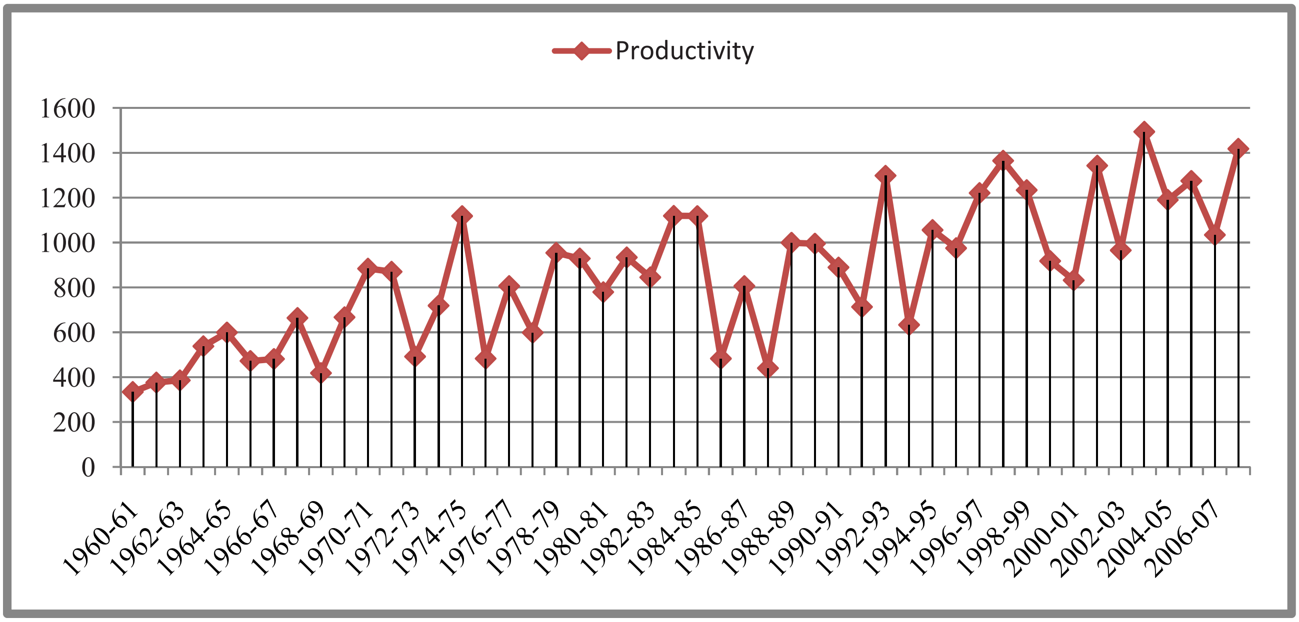


Fig.2. Graphical representation of pearl millet productivity data

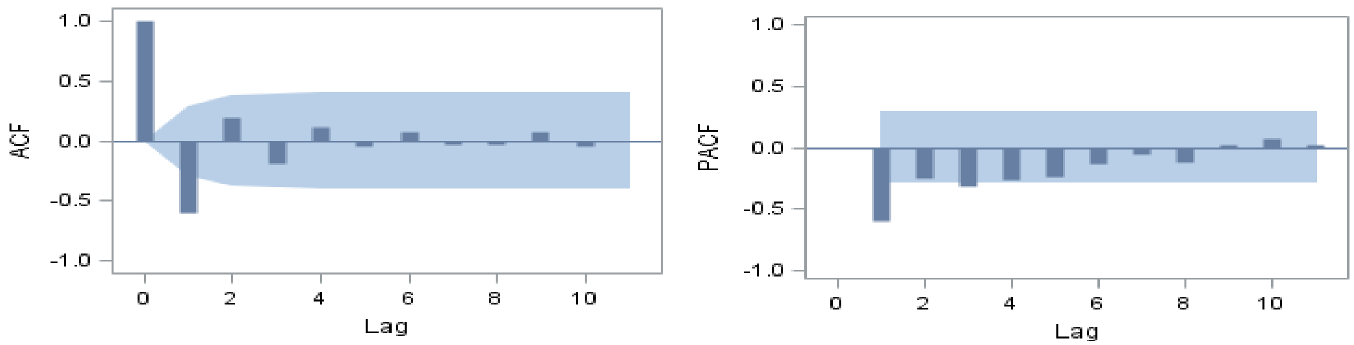


Fig. 3: ACF and PACF value for different Lag

Table. 2: Stationarity check of Pearl millet productivity data

Augmented Dickey-Fuller Unit Root Tests									
		Actual series				First differenced series			
Type	Lags	Tau	Pr< Tau	F	Pr > F	Tau	Pr< Tau	F	Pr > F
Zero Mean	**0**	-0.69	0.4117			-13.28	<.0001		
	1	0.19	0.7364			-7.40	<.0001		
	2	0.48	0.8154			-6.52	<.0001		
Single Mean	**0**	-4.02	0.0028	8.25	0.0010	-13.27	0.0001	88.01	0.0010
	1	-2.27	0.1841	2.98	0.3341	-7.48	0.0001	27.97	0.0010
	2	-1.96	0.3039	2.49	0.4537	-6.71	0.0001	22.49	0.0010
Trend	**0**	-8.04	<.0001	32.29	0.0010	-13.11	<.0001	86.04	0.0010
	1	-4.94	0.0012	12.21	0.0010	-7.39	<.0001	27.36	0.0010
	2	-4.69	0.0025	11.04	0.0010	-6.62	<.0001	21.96	0.0010

Table 3. AIC and BIC value for different combination of ARIMA model.

Model	**AIC**	**BIC**	**Parameter**
P=0,q=1	641.1552	644.8555	MU**;MA1,1**
P=0,q=2	642.3032	647.8537	MU**;MA1,1**;MA1,2
p=0,q=3	644.6765	652.0771	MU*;MA1,1*;MA1,2;MA 1,3
P=1,q=0	655.4257	659.1259	MU;AR1,1**
P=1,q=1	642.177	647.7275	MU**;MA1,1**;AR1,1
P=1,q=2	643.4471	650.8477	MU**;MA1,1;MA1,2;AR1,1
P=1,q=3	*The estimation algorithm did not converge after 50 iterations.*		
P=2,q=0	654.3806	659.9311	MU;AR1,1**;AR1,2
P=2,q=1	643.9213	651.3219	MU**;MA1,1**;AR1,1;AR1,2
P=2,q=2	645.4381	654.6888	MU**;MA1,1;MA1,2;AR1,1;AR1,2
P=2,q=3	647.473	658.5738	MU*;MA1,1;MA1,2**;MA 1,3;AR1,1**;AR1,2
P=3,q=0	651.2397	658.6402	MU;AR1,1**;AR1,2**;AR1,3*
P=3,q=1	645.0012	654.2519	MU**;AR1,1;AR1,2;AR1,3;MA 1,1**
P=3,q=2	652.282	663.3829	MU**;AR1,1;AR1,2;AR1,3;MA 1,1;MA1,2
P=3,q=3	648.9003	661.8514	MU**;MA1,1;MA1,2;MA 1,3;AR1,1;AR1,2;AR1,3

denotes significant at 5% level, denote at 1% level.*

In order to select the order of the ARIMA model, its autocorrelation functions (acf) and partial autocorrelation functions (pacf) are examined and on the basis of minimum AIC and BIC values, the model is selected. From fig.2 we have decided the maximum order of p and q i.e. 3. Again we have tried different combination of p and q and calculated the respective AIC and BIC. The values of AIC and BIC for different order are given in table 3.

From table-3, based on minimum AIC and BIC values, it is found that ARIMA (0,1,1) found to be best model for productivity forecast. The parameters are estimated for the selected model and the same are provided in the table 4.

Table 4. Parameter estimates along with standard errors (SE)

Parameter	Estimate	Standard Error	t Value	P value
MU	19.570	2.666	7.34	<.0001
MA1,1	0.963	0.052	18.29	<.0001

Validation of models for hold-out data

One-step ahead forecasts of pearl millet productivity for the year's 2007-08 to2011-12 in respect of above fitted models are reported in Table 4. The developed ARIMA model is validated on the basis of relative mean absolute prediction error (RMAPE),Mean absolute deviation(MAD) and root mean square error(RMSE) values defined as

$$\text{RMAPE} = 1/4\sum_{i=1}^{4}\{|y_{t+i} - \hat{y}_{t+i}|/y_{t+i}\}\times 100$$

$$\text{MAD} = 1/4\sum_{i=1}^{4}\{|y_{t+i} - \hat{y}_{t+i}|\}$$

$$\text{RMSE}=\sqrt{\frac{\Sigma(y_{t+i} - \hat{y}_{t+i})^2}{4}}$$

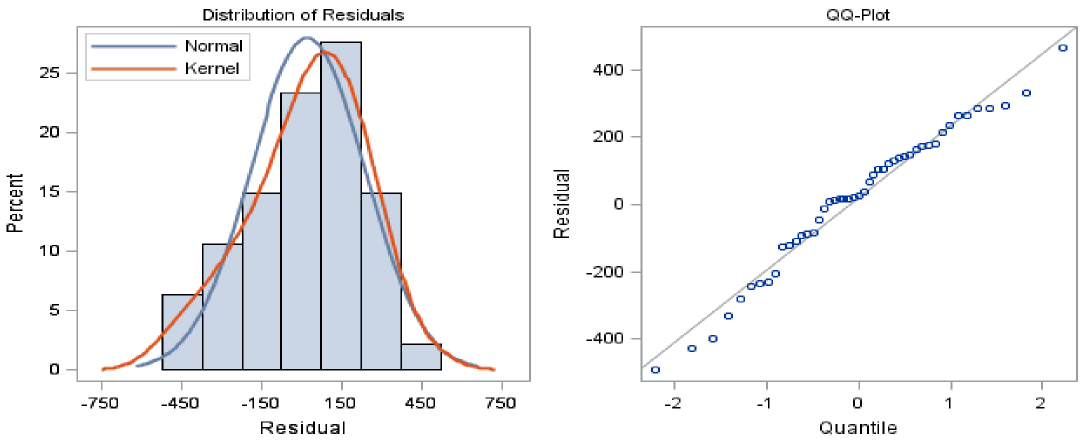
All the statistical measures are presented in Table 5. It may be noted that ARIMA (0,1,1) models perform quite satisfactorily as the RMAPE value is less than 6 percent.

Table 5. Forecasts of Pearl millet productivity (in Kg/Hectare) and their performance for hold-out data.

Year	Actual(Kg./ha)	Forecast(Kg./ha)
2008-09	1366.84	1310.73
2009-10	1232.14	1330.30
2010-11	1250.00	1349.87
2011-12	1419.03	1369.44
RMAPE (%)	5.88	
MAD	75.93	
RMSE	79.40	

The residuals from the fitted model were also investigated in order to examine presence of any autocorrelation among them. It is found that all the residuals are independent and distributed with normal distribution with zero mean and constant variance. Which indicates the proper specification of the model for forecasting of Pearl millet productivity.

Fig. 4: Residual Normality Diagnosis for fitted model



CONCLUSIONS

From the trend analysis of area and productivity of pearl millet in Gujarat state, it has been clear that the trend over year for both area and productivity, found to be significantly negative and positive respectively. But the issue is critical as cultivated area under pearlmillet decreases at a faster rate whereas productivity increases at an average rate and overall production decreases. The importance of pearl millet productivity forecasting is more relevant in semi-arid state like Gujarat where the precipitation is confirmed to short period of four months. In this paper, we have applied an ARIMA model for forecasting of productivity of pearl millet of Gujarat and trend of production over the five decades has been studied. It is found that ARIMA (0,1,1) model found to be best model for productivity forecast and it shows an increasing trend in future years. The residuals from the fitted model were also investigated in order to examine presence of any autocorrelation among them. It is found that all the residuals are independent and distributed with normal distribution with zero mean and constant variance, which indicates the proper specification of the model for forecasting of Pearl millet productivity.

REFERENCES

Anonymous. District-wise Area, Production and yield of important food and non-foodcrops in Gujarat state.2010-11.

Anonymous. Annual Report, Government of India,Ministry of Statistics and Programme Implementation.New Delhi. 2014–15.

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. *Time-SeriesAnalysis: Forecasting and Control*. 3rd edn. India: Pearson education; 2007.

Dickey, D.A. and Fuller W.A. Distribution of the estimators for the autoregressive time series with a unit root. *Journal of the American Statistical Association*.1979; 74: p427-431.

Paul, R.K.,Alam, W. and Paul,A.K.Prospects of livestock and dairy production in India under time series framework. *Indian Journal of Animal Sciences*.2014; 84(4): p130-134.

Paul, R.K. and Das, M. K. Statistical modelling of inland fish production in India.*Journal of the Inland Fisheries Society of India*.2010; 42 (2): p1–7.

Paul, R.K. and Das, M.K. Forecasting of average annual fish landing in Ganga Basin. *Fishing Chimes*.2013; 33(3): p51–54

Paul, R.K., Panwar,S., Sarkar, S.K., Kumar, A., Singh, K.N., Farooqi, S. and Chaudhary,V.K. Modelling and forecasting of meat exports from India. *Agricultural Economics Research Review*.2013;26(2): p249–56.

Paul, R. K. and Sinha,K. Forecasting crop yield: a comparative assessment of ARIMAX and NARX model. 2016; *RASHI*, 1(1): p77-85.