Forecasting of Area, Production, and Productivity of Food Grains in India : Application of ARIMA Model

** Pushpa Savadatti*

Abstract

Food grains occupy a dominant place in Indian agriculture. The demand for food grains is continuously increasing due to steady increase in the population. Food grains are an important source of energy and protein to majority of the Indians, who are vegetarians. Apart from this, the Government of India enacted the National Food Security Act (NFSA) which came into force with effect from July 5, 2013. This further put pressure on the demand for food grains in the country. Realizing the importance of food grains, the Government of India initiated various measures to boost the production and productivity of food grains since independence. As a result of this, the production of food grains has increased since 1950s, but still there is a gap between demand for and supply of food grains in the country which needs to be addressed urgently. In view of this, the projections for the area, production, and productivity of food grains for 5 years starting from 2016-17 onwards, based on the univariate time series analysis known as ARIMA analysis, was conducted in this paper. ARIMA (2,1,2), ARIMA (4,1,0), and ARIMA (3,1,3) models were fitted to the data on area, production, and productivity of food grains, respectively and these models were found to be adequate. The forecast values indicated that production and productivity will increase during the forecast period but that of area exhibited near stagnancy, calling for timely measures to enhance the supply of food grains to meet the increasing demand in the years to come.

Keywords: forecasts, autocorrelation, partial autocorrelation, residuals

JEL Classification : C22, C32, C53

Paper Submission Date : November 18, 2017 ; Paper sent back for Revision : December 6, 2017 ; Paper Acceptance Date : December 13 , 2017

Agriculture accounts for a very important position in the Indian economic development even today. More than 56% of the Indian population depends on agriculture for their livelihood. The agriculture sector provides food for more than 1.2 billion people. Food grains occupy a dominant place in the Indian agriculture as 80% of the cropped area was under food grains production during 2015-16. India is the largest producer of millets and is the second largest producer of rice, wheat, and pulses (Deshpande, 2017). The demand for food grains is continuously increasing due to steady increase in the population. Food grains are an important source of energy and protein to the majority of Indians, who are vegetarians. Apart from this, the Government of India enacted National Food Security Act (NFSA) which came into force with effect from July 5, 2013. The purpose of the act is to provide subsidized food grains to nearly 75% of the rural population and up to 50% of the urban population under Targeted Public Distribution System (TPDS), the act has already covered 80.54 crore persons as against the targeted of 81.35 crore people (Ministry of Finance, Dept of Economic Affairs, 2017). This further put pressure on the demand for food grains in the country. Supply of food grains is influenced by multiplicity of constraints like limited availability of land, water resources, competition from competing crops for the resources, market imperfections, weather risk, price risk, etc. Realizing the importance of food grains, the

* *Dean,* School of Business Studies, **Professor and Head*, Department of Economic Studies and Planning, Central University of Karnataka, Kadaganchi, Aland Road, Gulbarga - 585 367, Karnataka. E-mail : pmsavadatti@gmail.com

Government of India initiated various measures to boost the production and productivity of food grains since independence. As a result of this, food grains production increased from 74231 thousand tonnes in 1965 - 66 to 275680 thousand tonnes in 2015 -16. However, still there is a gap between demand for and supply of food grains in the country which needs to be addressed urgently. In view of this, the projections for the area, production, and productivity of food grains based on the sound technical analysis will help the policymakers and the government to take timely measures to enhance the supply of food grains to meet the increasing demand in the years to come. Hence, the present time series analysis of forecasting the food grains' area, production, and yield with the help of auto regressive integrated moving average (ARIMA) process becomes important, which is the main objective of this research paper.

Literature Review

Many researchers have adopted the ARIMA forecasting model, the popular and very widely used forecasting models for univariate time-series data for the field of agriculture. ARIMA technique was adopted to forecast the productivity of 34 different agricultural products in India (Padhan, 2012) and the study forecasted 5 years ahead from 2011 onwards. The author concluded that though the forecasted values for productivity of few selected agricultural products had been done based upon various criteria like mean absolute percentage error (MAPE), Akaike information criteria (AIC), etc., there were many other factors that could be influencing the productivity of the selected crops. Forecasting rice area, production, and productivity of Odisha was made based upon historical data by using the ARIMA model (Tripathi, Nayak, Raja, Shahid, Kumar, Mohanty, Panda, Lal, & Gautam, 2014). Based on validation results, it was concluded that ARIMA model could be successfully used for forecasting rice area, production, and yield. Similarly, a number of studies had used extensively univariate ARIMA technique to model various aspects of the agriculture sector in India (Biswas, Dhaliwal, Singh, & Sandhu, 2014 ; Darekar & Reddy, 2017 ; Gurung, Panwar, Singh, Banerjee, Gurung, & Rathore, 2017; Mishra, Sahu, Padmanaban, Vishwajith & Dhekale, 2015 ; Prabakaran & Sivapragasam, 2014).

Numerous studies used this technique to forecast area, production, and productivity of agricultural products in other countries also. To quote few such studies, short term forecasting for production of different varieties of rice employing ARIMA model had been done for Bangladesh (Awal & Siddique, 2011) and concluded that the forecasts could be used by policy makers, researchers, as well producers in their decision making. Another study modelled and forecasted area, production, and yield of total seeds of rice and wheat in SAARC countries (Sahu, Mishra, Dhekale, Vishwajith, & Padmanaban, 2015) and based upon the forecasts, the study emphasized the need for quantum jump in the per hectare yield of these two crops for the region. A study had been conducted in Nigeria to forecast the cultivation area and production of maize (Badmus & Ariyo, 2011) using univariate ARIMA process and concluded that total cropped area could be increased in the future, if land reclamation and conservation measures were adopted. An empirical study on agricultural products price forecasting based on ARIMA model in China showed that ARIMA model provided high accuracy of short term prediction for cucumber prices in Shandong Shouguang wholesale market (Xin & Can 2016).

It is evident from the literature review that time series analysis with the help of ARIMA model is widely used in forecasting of different variables pertaining to agricultural crops. In the present study, this univariate technique is adopted to estimate area, production, and yield of food grains grown in India, which are playing a very important role in the Indian economy in terms of food and nutrition to the population, employment creation, etc.

Methodology

(1) Data : The present analysis is based on the secondary data collected from Centre for Monitoring Indian Economy (CMIE). The annual time series data were collected for a period from 1966 - 67 to 2015 - 16 on area, production, and productivity of food grains, respectively at the all India level. The area, production, and yield of food grains were measured in thousand hectares, thousand tonnes, and kgs/hectare, respectively.

(2) The Box -Jenkins (BJ) Methodology : The present study used the methodology popularly known as BJ methodology for forecasting area, production, and productivity of food grains in India. But technically, this methodology is known as the ARIMA methodology (Guajarati & Sangeeta, 2007). Here, our emphasis is on univariate ARIMA model. The BJ methodology involves the following steps :

(i) Identification : This step is to identify suitable values for p,d, and q in the ARIMA (p,d,q) model. p indicates the number of autoregressive (AR) terms, d indicates the order of integration, that is, the number of times the time series data has to be differenced so that series are stationary, and q denotes the number of moving average terms. The BJ procedure is applicable to stationary time series data. So, it is necessary to ensure that series are stationary ; if not, they have to be made stationary through appropriate transformation. Examining the plots of the auto correlation functions (ACF) and partial auto correlation function (PACF) of the series, appropriate values for the p,d,q may be identified.

The general ARIMA (p,d,q) model may be written as :

$$\phi(B)\Delta^d Y_t = \theta(\mathrm{B})\varepsilon_t \quad \text{-------------------------------------} \quad (1)$$

where,

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ ------------------- (2), where p = autoregressive parameters.

$\theta(B) = 1 - \theta_1 \mathrm{B} - \theta_2 B^2 - \cdots - \theta_q B^q$ ------------------ (3), where q = moving average parameters.

$\phi(B)$ is autoregressive operator and $\theta(B)$ is the moving average operator (Pindyck & Rubinfeld, 1998).

(ii) Estimation and Checking : After selecting the suitable values for p,d,q in the first step, the next step is to estimate the parameters of AR and MA terms with the suitable estimation method. Once the model is estimated, then there is need for the model adequacy check by considering the properties of the estimated residuals. The estimated residuals from the selected ARIMA model are normal and randomly distributed and will be tested based on the skewness, kurtosis, Jarque - Bera (JB) test and residual plots of ACF and PACF along with the Ljung-Box Q statistics. The Q statistics can be expressed as :

$$Q_m = n(n+2)\Sigma_{k=1}^{l} \frac{r_k^2}{n-k} \sim \chi^2_{l-m} \quad \text{---------------------------} (4)$$

where, n is the number of observations or residuals, k is the order of residual correlation, l is the number of autocorrelations included in the test, m is the number of parameters estimated, and $l - m$ equals degrees of freedom and it follows the chi-square distribution (Nazeem, 1998).

If the estimated residuals are white noise, then the estimated model is adequate ; otherwise, another ARIMA model has to be selected starting from identification stage once again. Since, BJ methodology is an iterative procedure, identification, estimation, and checking stages are repeated until we get a satisfactory model.

(iii) Forecasting : The main purpose of the ARIMA modelling is to forecast and they are known for forecasting accuracy. Hence, once the selected model satisfies the model adequacy tests, then it will be used for forecasting.

The annual time series data collected on the area, production, and productivity of food grains are used for the present analysis with the help of e-views software to identify the appropriate ARIMA models for all the three variables.

Analysis and Results

The BJ methodology involves the process of identification, estimation, and forecasting. The same steps are followed for the analysis. The analysis has been done with the help of E - views 9 software and the results of the data analysis are presented in figures and tables.

(1) Model Identification : Before we begin the modelling of the data, it is necessary to check whether the series under consideration are stationary or not. This can be done by observing the plots of ACF and PACF for the series under consideration and also with the help of Augmented Dicky Fuller (ADF) test. The ACF and PACF plots for the area under food grains (at levels) are presented in the Figure 1. The ACF and PACF of area time series data presented in Figure 1 indicates that series are non-stationary as PACF dies down slowly. The ADF test results for area series are presented in the Table 1.

The ADF test for area (at levels) presented in the Table 1 substantiates the conclusion of non-stationarity as p - values is > 0.05. So, the area series are non-stationary at levels. There is need for appropriate data transformation to make the series stationary. The area series were transformed by taking first difference. The ACF and PACF of transformed series are shown in the Figure 2.

Both ACF and PACF cut off after first lag for the transformed area series indicate that the series are stationary.

Table 1. Stationarity Test for Area Under Total Food Grains in India

Area		At levels		First difference	
		t-statistics	Probability*	*t*-statistics	Probability*
ADF test statistics		-1.773913	0.3887	-12.76053	0.0000
Test Critical Value	1% level	-3.574446		-3.574446	
	5% level	-2.923780		-2.923780	
	10% level	-2.599925		-2.599925	

*MacKinnon (1996) one-sided p - values ; ADF = Augmented Dicky Fuller

Figure 1. ACF and PACF of the Area (at Levels) under Food Grains in India

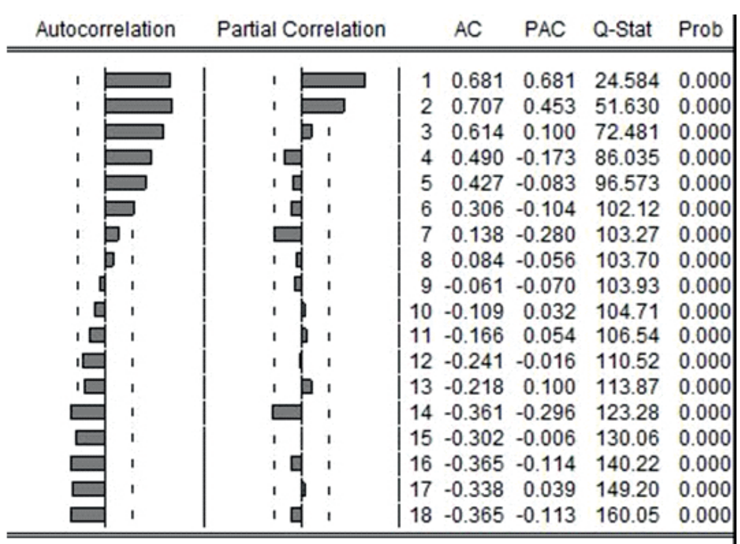
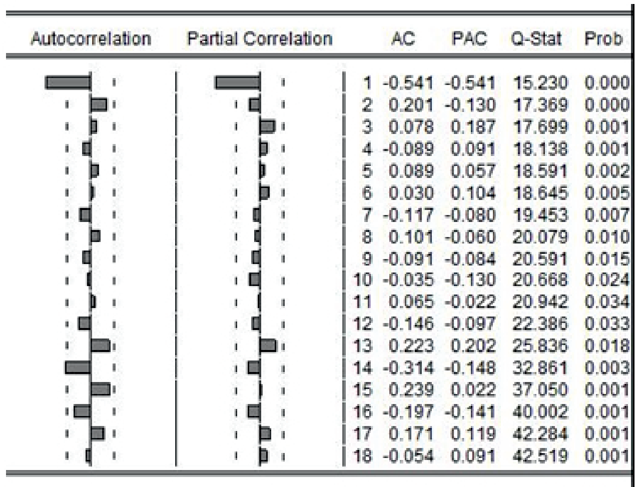


Figure 2. ACF and PACF of the Area (First Difference) Under Food Grains in India



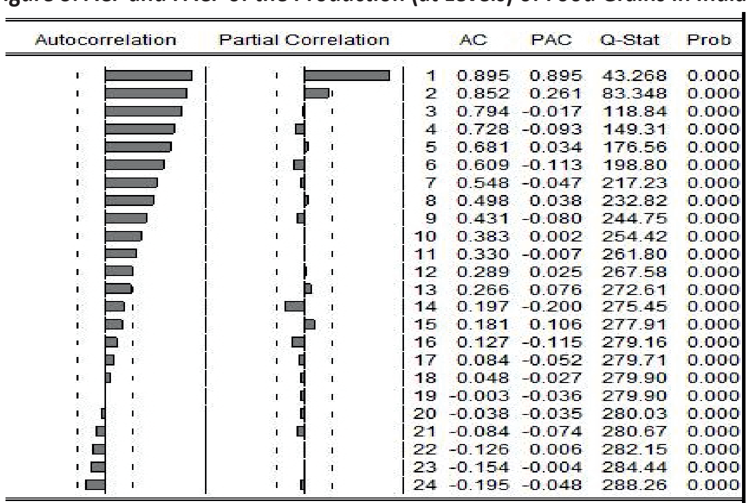
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.541	-0.541	15.230	0.000
		2	0.201	-0.130	17.369	0.000
		3	0.078	0.187	17.699	0.001
		4	-0.089	0.091	18.138	0.001
		5	0.089	0.057	18.591	0.002
		6	0.030	0.104	18.645	0.005
		7	-0.117	-0.080	19.453	0.007
		8	0.101	-0.060	20.079	0.010
		9	-0.091	-0.084	20.591	0.015
		10	-0.035	-0.130	20.668	0.024
		11	0.065	-0.022	20.942	0.034
		12	-0.146	-0.097	22.386	0.033
		13	0.223	0.202	25.836	0.018
		14	-0.314	-0.148	32.861	0.003
		15	0.239	0.022	37.050	0.001
		16	-0.197	-0.141	40.002	0.001
		17	0.171	0.119	42.284	0.001
		18	-0.054	0.091	42.519	0.001

Table 2. Stationarity Test for Production of Total Food Grains in India

Production		At levels		First difference	
		t-statistics	Probability*	*t*-statistics	Probability*
ADF test statistics		-0.198722	0.9315	-13.10399	0.0000
Test Critical Value	1% level	-3.571310		-3.574446	
	5% level	-2.922449		-2.923780	
	10% level	-2.599224		-2.599925	

*MacKinnon (1996) one-sided *p*-values ; ADF = Augmented Dicky Fuller

Figure 3. ACF and PACF of the Production (at Levels) of Food Grains in India



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.895	0.895	43.268	0.000
		2	0.852	0.261	83.348	0.000
		3	0.794	-0.017	118.84	0.000
		4	0.728	-0.093	149.31	0.000
		5	0.681	0.034	176.56	0.000
		6	0.609	-0.113	198.80	0.000
		7	0.548	-0.047	217.23	0.000
		8	0.498	0.038	232.82	0.000
		9	0.431	-0.080	244.75	0.000
		10	0.383	0.002	254.42	0.000
		11	0.330	-0.007	261.80	0.000
		12	0.289	0.025	267.58	0.000
		13	0.266	0.076	272.61	0.000
		14	0.197	-0.200	275.45	0.000
		15	0.181	0.106	277.91	0.000
		16	0.127	-0.115	279.16	0.000
		17	0.084	-0.052	279.71	0.000
		18	0.048	-0.027	279.90	0.000
		19	-0.003	-0.036	279.90	0.000
		20	-0.038	-0.035	280.03	0.000
		21	-0.084	-0.074	280.67	0.000
		22	-0.126	0.006	282.15	0.000
		23	-0.154	-0.004	284.44	0.000
		24	-0.195	-0.048	288.26	0.000

The results of the ADF test for differenced series presented in Table 1 confirm that the data series are stationary after first difference as the p - value is < 0.05. The ACF and PACF plots of production series at levels are presented in the Figure 3. It is observed from the Figure 3 that ACF dies down very slowly, signalling that the production series are non - stationary. To validate this decision, the ADF test is also done for production series at level and

Figure 4. ACF and PACF of the Production (First Difference) of Food Grains in India

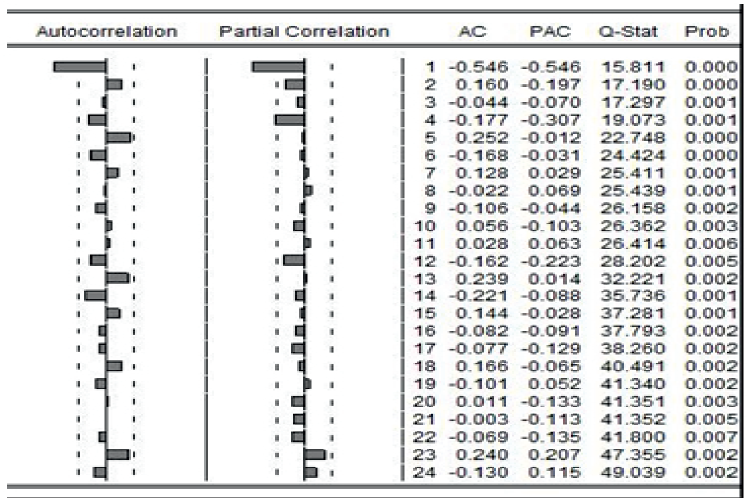


Figure 5. ACF and PACF of the Yield (at Levels) of Food Grains in India

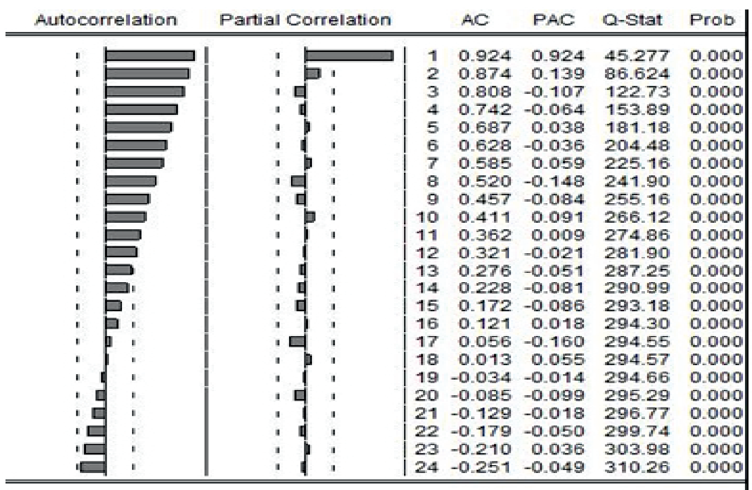


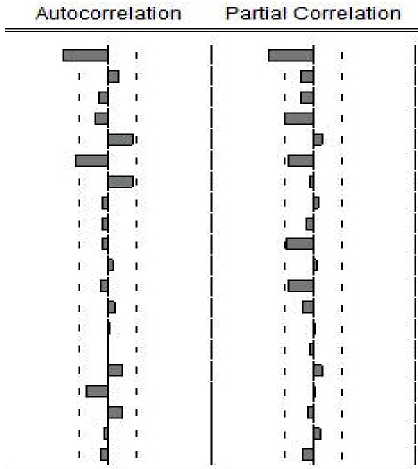
Table 3. Stationarity Test for Productivity of Total Food Grains in India

Productivity		**At levels**		**First difference**	
		***t*-statistics**	**Probability***	***t*-statistics**	**Probability***
ADF test statistics		**-0.156105**	**0.9368**	**-11.34520**	**0.0000**
Test Critical Value	1% level	-3.574446		-3.574446	
	5% level	-2.923780		2.923780	
	10% level	-2.599925		-2.599925	

*MacKinnon (1996) one-sided *p*-values ; ADF = Augmented Dicky Fuller

first difference. The results of the same are presented in the Table 2. The ADF test results (Table 2) for production series at levels display that the series are non-stationary at levels as the probability is 0.9315 > 0.05. The production series are differenced once to make them stationary. The ACF and PACF plots of the differenced series are presented in the Figure 4.

Figure 6. ACF and PACF of the Yield (First Difference) of Food Grains in India



	AC	PAC	Q-Stat	Prob
1	-0.456	-0.456	10.837	0.001
2	0.103	-0.132	11.406	0.003
3	-0.089	-0.124	11.840	0.008
4	-0.130	-0.279	12.774	0.012
5	0.257	0.088	16.529	0.005
6	-0.326	-0.249	22.699	0.001
7	0.249	-0.033	26.403	0.000
8	-0.055	0.054	26.589	0.001
9	-0.057	-0.081	26.793	0.002
10	-0.061	-0.267	27.032	0.003
11	0.051	0.030	27.206	0.004
12	-0.070	-0.254	27.538	0.006
13	0.078	-0.105	27.965	0.009
14	0.017	0.017	27.987	0.014
15	0.003	-0.041	27.987	0.022
16	0.133	0.086	29.325	0.022
17	-0.221	0.015	33.126	0.011
18	0.137	-0.060	34.643	0.010
19	-0.034	0.081	34.741	0.015
20	-0.080	-0.102	35.298	0.019

The Figure 4 clearly shows that ACF and PACF cut off after first lag depicting first differenced series are stationary. This is further authenticated by the ADF test results for the differenced production series presented in the Table 2. The next task is to test whether the yield series are stationary or not. For this, we examine first ACF and PACF of yield series at levels which are presented in the Figure 5.

The ACF plots clearly show that the series are non-stationary. This is confirmed by the ADF test results presented in the Table 3 for yield series at levels and for the first differenced series. The yield series are transformed to first differenced series. The differenced series are tested for stationarity. ACF and PACF plots of the differenced yield series are presented in the Figure 6.

The Figure 6 indicates that the series are stationary at first difference as ACF and PACF cut off after first lag and none of the other lagged autocorrelation and partial correlations appear to be significant except 6th lag of autocorrelation and 4th lag of partial correlations. The ADF test results for differenced series are also in conformity of the conclusion that yield series are stationary at first difference (Table 3). All the three-series - area, production, and yield of food grains are made stationary after first difference. Once the series are stationary, the next task is to it identify the suitable ARIMA model for these series based on the ACF and PACF of the stationary series. We need to try various specifications of the model and that takes us to the next step, that is, estimation and diagnostic checking.

(2) Estimation and Checking : As a part of the identification process, we have examined the ACF and PACF of time series pertaining to area, production, and productivity of food grains and all the three series were made stationary after first difference and the ACF and PACF of differenced series were also closely examined which is necessary for identification of the model. Based on the study of the plots of autocorrelations and their partials, various ARIMA specifications have been tried for all the three variables. The optimal ARIMA models were selected considering the various criteria like Akaike information criterion (AIC), Schwarz information criterion (SIC), significance of the parameters, R^2, F-statistics, residual series examination, etc. The results of the fitted ARIMA model of the order (2,1,2) for area is presented in the Table 4.

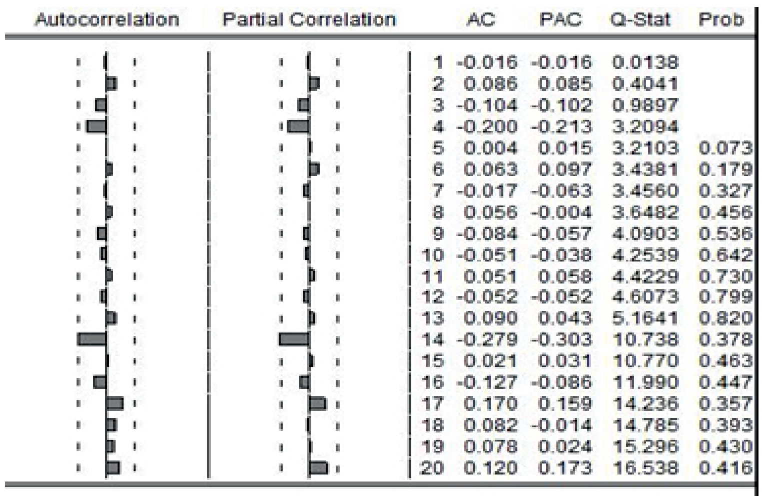
The results indicate that coefficients of AR(1) is significant and R^2 is also significant. Next, it needs to be checked whether the estimated residuals from the fitted model are white noise and for that, we need to study the correlogram of the residuals, which is presented in the Figure 7. It can be seen from the Figure 7 that none of the residual autocorrelations are statistically significant and this can also be confirmed by looking at the Box-Pierce

Table 4.Details of ARIMA (2,1,2) Model Fitted for Area Under Food Grains

Variable	Coefficient	Standard Error	*t* - Statistics	Probability
C	242.7500	616.2080	0.393942	0.6956
AR(1)	0.615258	0.188365	3.266313	0.0021***
AR(2)	-0.209240	0.257031	-0.814065	0.4201
MA(1)	-1.373985	407.2704	-0.003374	0.9973
MA(2)	0.999995	592.8114	0.001687	0.9987
R^2	0.445966***		Akaike info criterion	19.42029
Adjusted R^2	0.381543		Schwarz criterion	19.65194
F- statistics	6.922510		Hannan-Quinn criteria	19.50818
Prob(*F*-statistic)	0.000081		Durbin-Watson stat	1.995968

Note: *** indicates significance at 1%

Figure 7. Correlogram of Residuals - Area Model : ARMA (2,1,2)



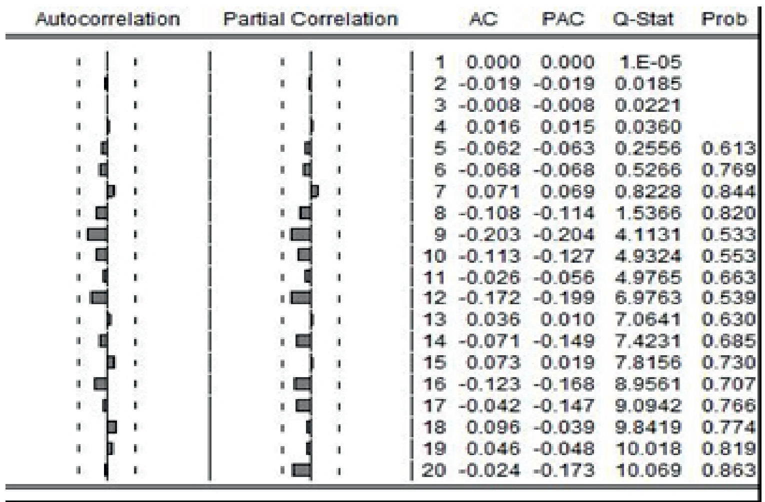
	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1			-0.016	-0.016	0.0138	
2			0.086	0.085	0.4041	
3			-0.104	-0.102	0.9897	
4			-0.200	-0.213	3.2094	
5			0.004	0.015	3.2103	0.073
6			0.063	0.097	3.4381	0.179
7			-0.017	-0.063	3.4560	0.327
8			0.056	-0.004	3.6482	0.456
9			-0.084	-0.057	4.0903	0.536
10			-0.051	-0.038	4.2539	0.642
11			0.051	0.058	4.4229	0.730
12			-0.052	-0.052	4.6073	0.799
13			0.090	0.043	5.1641	0.820
14			-0.279	-0.303	10.738	0.378
15			0.021	0.031	10.770	0.463
16			-0.127	-0.086	11.990	0.447
17			0.170	0.159	14.236	0.357
18			0.082	-0.014	14.785	0.393
19			0.078	0.024	15.296	0.430
20			0.120	0.173	16.538	0.416

Table 5. Details of ARIMA (4,1,0) Model Fitted for Production of Food Grains

Variable	Coefficient	Standard Error	*t* - Statistics	Probability
C	3574.196	735.5284	4.859359	0.0000
AR(1)	-0.715859	0.184713	-3.875519	0.0004***
AR(2)	-0.311934	0.216430	-1.441266	0.1568
AR(3)	-0.279771	0.267304	-1.046638	0.3011
AR(4)	-0.319389	0.165101	-1.934509	0.0596*
R^2	0.420848***		Akaike info criterion	21.63764
Adjusted R^2	0.353504		Schwarz criterion	21.86929
F-statistic	6.249284		Hannan-Quinn criteria	21.72553
Prob(*F*-statistic)	0.000194		Durbin-Watson stat	1.937778

Note: ***, * indicate significance at 1% and 10%, respectively

Figure 8 : Correlogram of Residuals - Production Model: ARIMA (4,1,0)



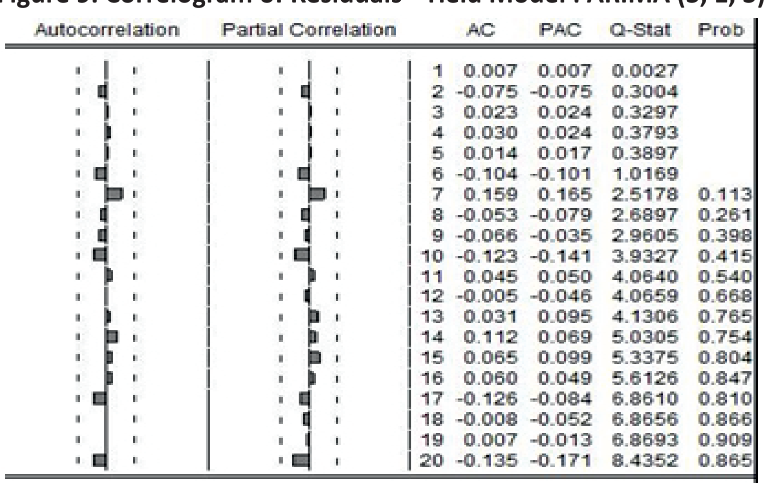
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.000	0.000	1.E-05	
		2	-0.019	-0.019	0.0185	
		3	-0.008	-0.008	0.0221	
		4	0.016	0.015	0.0360	
		5	-0.062	-0.063	0.2556	0.613
		6	-0.068	-0.068	0.5266	0.769
		7	0.071	0.069	0.8228	0.844
		8	-0.108	-0.114	1.5366	0.820
		9	-0.203	-0.204	4.1131	0.533
		10	-0.113	-0.127	4.9324	0.553
		11	-0.026	-0.056	4.9765	0.663
		12	-0.172	-0.199	6.9763	0.539
		13	0.036	0.010	7.0641	0.630
		14	-0.071	-0.149	7.4231	0.685
		15	0.073	0.019	7.8156	0.730
		16	-0.123	-0.168	8.9561	0.707
		17	-0.042	-0.147	9.0942	0.766
		18	0.096	-0.039	9.8419	0.774
		19	0.046	-0.048	10.018	0.819
		20	-0.024	-0.173	10.069	0.863

Table 6. Details of ARIMA (3, 1, 3) Model Fitted for Productivity of Food Grains

Variable	Coefficient	Standard Error	*t*-Statistics	Probability
C	28.80512	1.627084	17.70352	0.0000
AR(1)	-0.799759	0.310260	-2.577708	0.0136**
AR(2)	-0.176910	0.380225	-0.465277	0.6442
AR(3)	0.406331	0.235845	1.722871	0.0924*
MA(1)	0.230986	7.145491	0.032326	0.9744
MA(2)	-0.233580	8.374210	-0.027893	0.9779
MA(3)	-0.993993	6.660986	-0.149226	0.8821
R^2	0.491373***		Akaike info criterion	11.53346
Adjusted R^2	0.404535		Schwarz criterion	11.84233
F-statistic	5.658458		Hannan-Quinn criteria	11.65064
Prob (*F*-statistic)	0.000128		Durbin-Watson stat	1.926914

Note: ***, **, * indicate significance at 1%, 5%, and 10%, respectively.

Figure 9. Correlogram of Residuals - Yield Model : ARIMA (3, 1, 3)



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.007	0.007	0.0027	
		2	-0.075	-0.075	0.3004	
		3	0.023	0.024	0.3297	
		4	0.030	0.024	0.3793	
		5	0.014	0.017	0.3897	
		6	-0.104	-0.101	1.0169	
		7	0.159	0.165	2.5178	0.113
		8	-0.053	-0.079	2.6897	0.261
		9	-0.066	-0.035	2.9605	0.398
		10	-0.123	-0.141	3.9327	0.415
		11	0.045	0.050	4.0640	0.540
		12	-0.005	-0.046	4.0659	0.668
		13	0.031	0.095	4.1306	0.765
		14	0.112	0.069	5.0305	0.754
		15	0.065	0.099	5.3375	0.804
		16	0.060	0.049	5.6126	0.847
		17	-0.126	-0.084	6.8610	0.810
		18	-0.008	-0.052	6.8656	0.866
		19	0.007	-0.013	6.8693	0.909
		20	-0.135	-0.171	8.4352	0.865

Q - statistics presented in the Figure 7, which are high. Therefore, we can say that the estimated area model may be used to forecast the area series.

The results of the estimated ARIMA (4,1,0) model for production are presented in the Table 5. It is clear from the Table 5 that AR(1) and AR(4) are statistically significant at 1% and 10% , respectively. The R^2 is also significant at the 1% level. The adequacy of the model is also checked on the basis of estimated residual series whose ACF and PACF plots along with Q statistics are presented in the Figure 8. It is clear from the residual correlogram plot (Figure 8) that all the AC and PAC are statistically insignificant, indicating that there is no pattern left in the residuals for the production model. This is further substantiated by the higher probability values of the Q statistics.

The results of the estimated ARIMA (3,1,3) model for data on yield are presented in the Table 6. The results indicate that AR(1) and AR(3) terms are statistically significant at the 5% and 10% levels, respectively. R^2 is also significant at the 1% level. The ACF and PACF of the estimated residuals of the fitted model are presented in the Figure 9. It is amply clear from the Figure 9 that there is no pattern left in the residuals as all the residuals are statistically insignificant substantiated by the value of the Q statistics and high value of the corresponding probability. Further, the test for presence of heteroskedasticity in residuals is also done for all the three selected ARIMA models based on the correlogram of the squared residuals respectively for area, production, and yield, which indicates the absence of the problem. The normality test of the residuals for all the fitted models were also examined based on the skewness, kurtosis, and Jarque - Bera (JB) tests and the results of the same are presented in the Table 7.

The normality assumption of the residuals is satisfied for all the models and substantiated by the statistics presented in the Table 7. Skewness and kurtosis are near to 0 and 3, respectively and JB probability is > 0.05 in all the cases, indicating that the residuals are normally distributed. The graphs of the actuals and fitted values for area, production, and yield are presented in Figures 10, 11, and 12, respectively. The important pattern present in the area, production, and yield series are captured by the fitted values of area (Figure 10) , production (Figure 11), and yield (Figure 12), respectively. Based on the above diagnostic checking, it is concluded that the all the

Table 7. Normality Test of the Residuals for the Fitted ARIMA Models

Variable/Model	Area/ARIMA (2,1,2)	Production/ARIMA (4,1,0)	Productivity/ARIMA (3,1,3)
Skewness	-0.0270	-0.4752	-0.1855
Kurtosis	2.8370	4.2200	3.0161
Jarque-Bera/	0.0602	4.8834	0.2818
Probability	0.9704	0.0870	0.8686

Figure 10. Actual and Fitted Values -Area Model : ARMA (2,1,2)

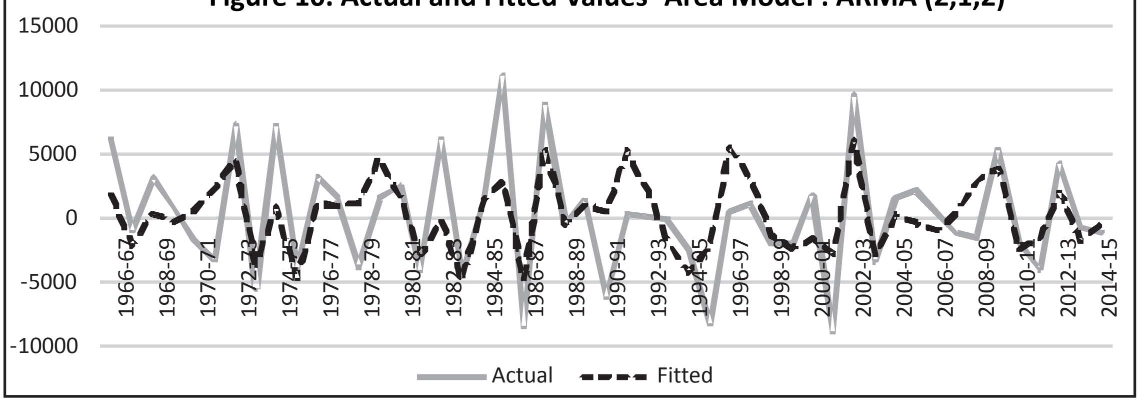


Figure 11. Actual and Fitted Values - Production Model : ARIMA (4,1,0)

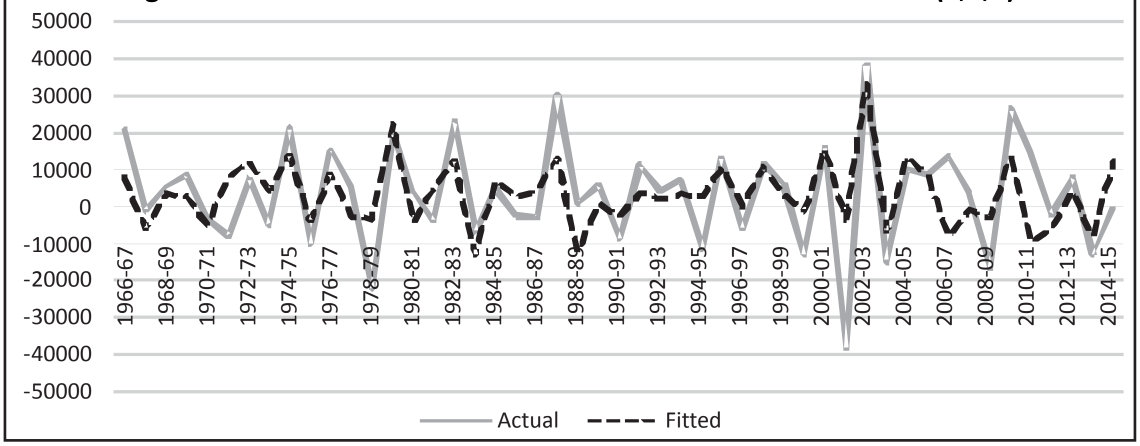
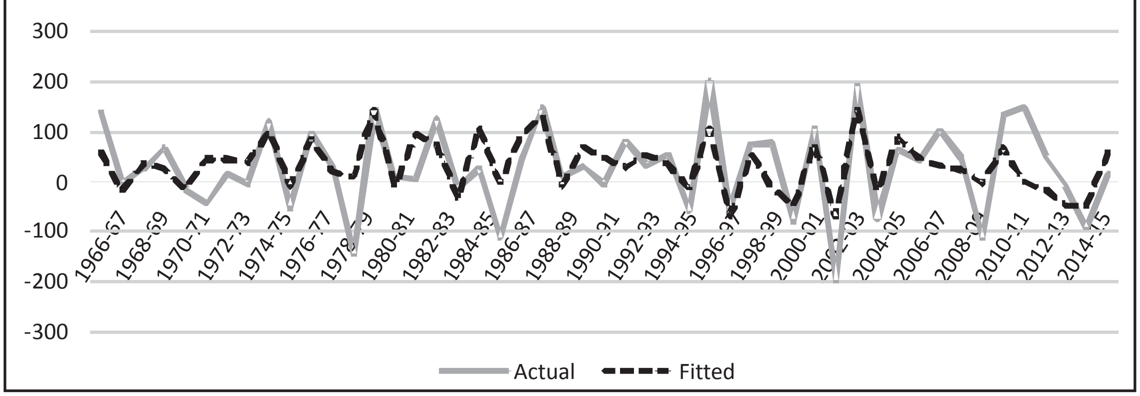


Figure 12. Actual and Fitted Values - Yield Model : ARIMA (3,1,3)



estimated ARIMA models are considered as reasonably satisfactory and may be used for forecasting. That takes us to the next step of forecasting values for the area, production, and yield of total food grains in the country.

(3) Forecasting : The main purpose of the fitted models is to forecast the values. The fitted models are used to forecast the values for the area, production, and yield for the next 5 years from 2015-16 onwards. The forecast values for area, production, and yield are presented in the Table 8. The forecasting accuracy tests for the models are presented in the Table 9. The accuracy of the forecasts is evaluated with the help of root mean square error (RMSE), mean absolute percentage error (MAPE), and Theil's inequality coefficient. The results of the tests are presented in the Table 9. The Theil's equality coefficient is near to zero in all the three cases, indicating that forecasting accuracy is reasonably good (Pindyck & Rubifeld, 1998). The forecasting precision of the estimated ARIMA models for area, production, and yield have been done. The forecasting precision results for the area model are presented in the Table 10.

The results for area under food grains displays that the forecast model can control 100% of the predicted value relative error in 10 % . The actual, estimated, and relative percentage error for the production model is presented in the Table 11.

In case of production, the predicted value relative error is within 5% in all the cases except for the years 2006-07 and 2009-10 (Table 11). This means 80% of the predicted value relative error is within 5%. Therefore, it may be concluded that the production model is found to be satisfactory as a whole. The Table 12 presents the forecasting precision for yield model.

Table 8. Forecasts for Area, Production, and Yield of Food Grains in India

Year	Area ('000' hectares)	Production ('000' tonnes)	Yield (Kgs/hectare)
2016-17	133442.42	267811.84	2141.55
2017-18	133685.17	271385.93	2170.63
2018-19	133927.92	274960.11	2199.25
2019-20	134170.67	278534.39	2228.07
2020-21	134413.42	282108.48	2257.01

Table 9. Forecast Evaluation of the Fitted ARIMA Models

Variable/Model	Area/ARIMA (2,1,2)	Production/ARIMA (4,1,0)	Productivity/ARIMA (3,1,3)
RMSE	7200.49	12844.57	107.9404
MAPE	4.8709	6.0895	7.3897
Theil's Inequality Coefficient	0.0284	0.0345	0.0365

Table 10. Forecasting Precision for Area Model - ARIMA (2,1,2)

Area under Food grains ('000' hectares)			
Year	Actuals	Estimated	Error %
2006-07	123,708.00	131,014.90	-5.91
2007-08	124,067.50	131,257.70	-5.80
2008-09	122,833.50	131,500.40	-7.06
2009-10	121,333.60	131,743.20	-8.58
2010-11	126,671.30	131,985.90	-4.20
2011-12	124,754.90	132,228.70	-5.99
2012-13	120,770.70	132,471.40	-9.69
2013-14	125,046.80	132,714.20	-6.13
2014-15	124,298.70	132,956.90	-6.97
2015-16	123,217.40	133,199.70	-8.10

Table 11. Forecasting Precision for Production Model - ARIMA (4,1,0)

Production of Food grains ('000 tones)			
Year	Actuals	Estimated	Error %
2006-07	217,282.10	232,070.10	-6.81
2007-08	230,775.00	235,644.30	-2.11
2008-09	234,466.20	239,217.60	-2.03
2009-10	218,107.40	242,793.00	-11.32
2010-11	244,482.00	246,366.10	-0.77
2011-12	259,286.00	249,941.00	3.60
2012-13	257,134.60	253,515.00	1.41
2013-14	265,045.20	257,089.00	3.00
2014-15	252,022.90	260,663.50	-3.43
2015-16	251,566.30	264,237.40	-5.04

In case of yield, all the predicted values are within the relative error of 8% and 90% of the predicted value, and the relative errors are within 5%, confirming the forecasting accuracy of the selected model (Table 12).

The forecasting accuracy tests amply substantiate that the forecasts done for the area, production, and yield of food grains are fairly accurate. The plots of the actual and forecasted values for area, production, and yield are

Table 12. Forecasting Precision for Yield Model - ARIMA (3,1,3)

Yield of Food grains (Kgs/Hectare)			
Year	Actuals	Estimated	Error %
2006-07	1,756.41	1,853.75	-5.54
2007-08	1,860.00	1,882.64	-1.22
2008-09	1,908.81	1,910.92	-0.11
2009-10	1,797.59	1,940.33	-7.94
2010-11	1,930.05	1,968.78	-2.01
2011-12	2,078.36	1,997.55	3.89
2012-13	2,129.11	2,026.69	4.81
2013-14	2,119.57	2,055.09	3.04
2014-15	2,027.56	2,084.15	-2.79
2015-16	2,041.65	2,112.96	-3.49

Figure 13. Actual & Forecasted Values - Area under Food Grains

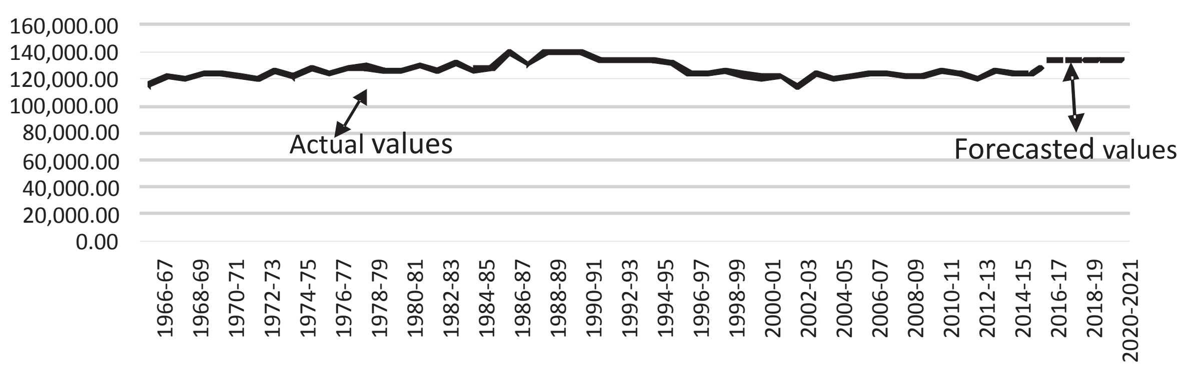
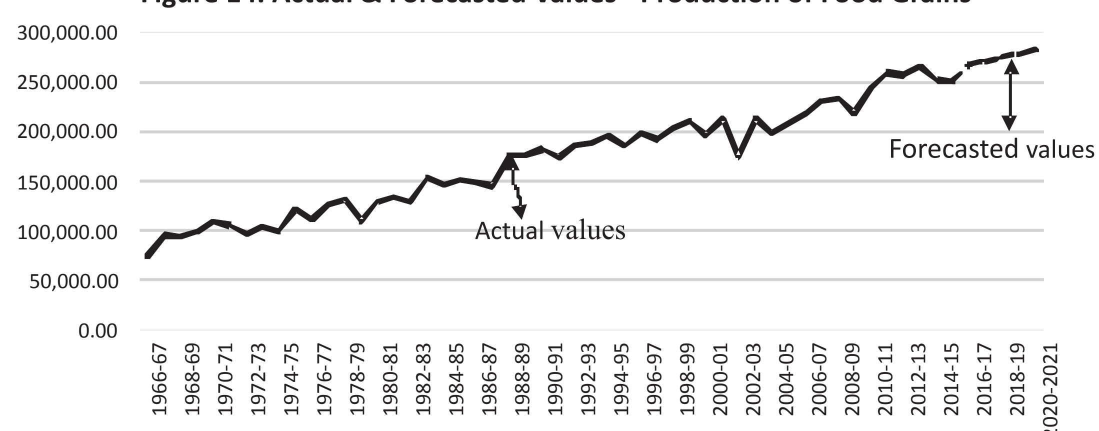
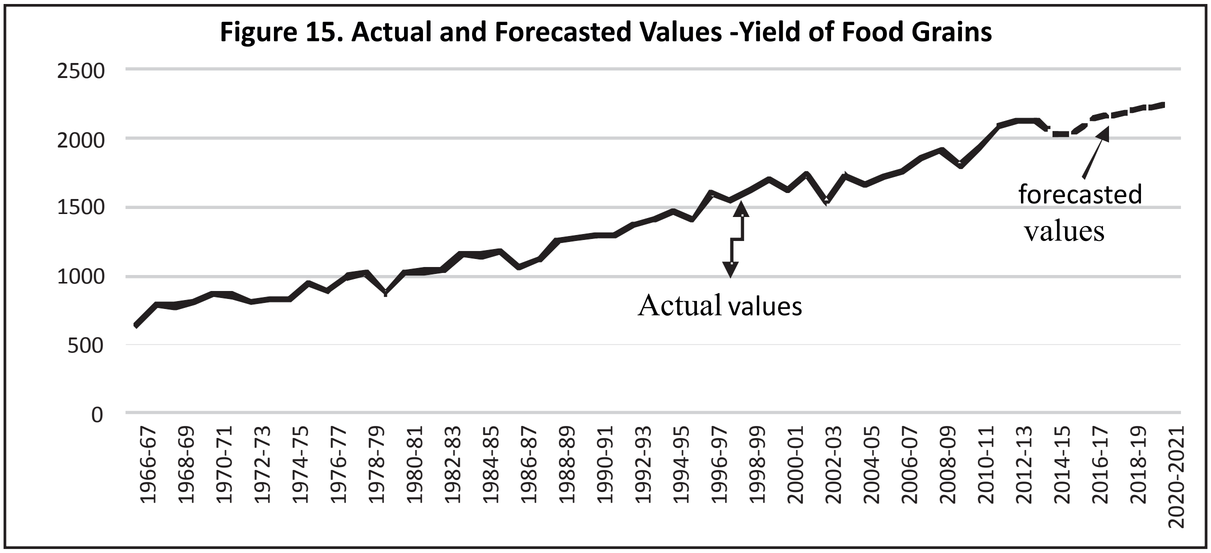


Figure 14. Actual & Forecasted Values - Production of Food Grains





presented in the Figures 13 to 15, respectively. The Figure 13 shows that the area under food grains will be almost stagnant during forecast period, that is, 2016-17 to 2020-21 but that of production and yield of food grains shows increasing trend during the same period (Figure 14 and Figure 15, respectively). It may be inferred that the increase in production will be brought about by increased yield during the forecast period.

Research and Policy Implications

The present study makes an effort to forecast the values for the yield, production, and productivity of food grains in India based on the past historical data using sound econometric technique known as ARIMA technique. The ARIMA technique is popular for its forecasting accuracy. Hence, in the present study, univariate ARIMA analysis is adopted to forecast the values for food grain's production, area, and yield in India for 5 years starting from 2016-17 to 2020-21. The selected ARIMA models fulfil all the technical requirements of a good model ; hence, the forecasts are fairly accurate.

These forecasts would be of great help to policy makers in their planning and future policy decisions as these forecasts indicate the direction of movements of the food grains' area, production, and yield in the country. Food grains are an important source of food to millions of people in India who are poor and living below the poverty line. It is the responsibility of the state to ensure food security to all these poor people. Secondly, the Indian population is growing continuously. This makes it necessary to take measures to ensure the food safety for the future by filling the gap between expected demand and supply of food grains. At this juncture, these forecast values will be of great help to the policy makers to take appropriate investment and policy decisions to ensure food security for all the citizens of the country in the future. It is evident from the forecasts that though production and productivity show an increasing trend, but area under food grains is not showing much increase. So, there is need for more concerted efforts by the government in enhancing the area under food grains production on one hand and on the other, there is need to motivate farmers to grow more food grains as more and more farmers are shifting their cultivable land towards commercial crops. Incentives in terms of better prices for the products, enough warehousing facilities, marketing facilities, supply of inputs at subsidized rates, assured markets, etc., are essential to encourage the farmers to bring more land under cultivation of food grains. Hence, this calls for concerted efforts by the government to increase the area under cultivation of food grains in order to increase the production of food grains on a sustainable basis.

Limitations of the Study and Scope for Further Research

The present analysis is based on the univariate time-series analysis in forecasting the area, production, and yield of food grains, which is the limitation of the study as these variables are influenced by a multiplicity of factors. For example, area under food grains is influenced by cost of cultivation, farm gate prices of the products, timely availability of inputs at ease, rainfall and irrigation facilities, etc. Similarly, production and yield are also influenced by various other factors. So, while forecasting the future values for the area, production, and yield, if we take important exogenous variables in the analysis, that enhances the accuracy of the forecasts. This could be considered for any further research on forecasting.

References

Awal, M.A., & Siddique, M. A. B. (2011). Rice production in Bangladesh employing by ARIMA model. *Bangladesh Journal of Agricultural Research, 36*(1), 51-62. doi:https://www.banglajol.info/index.php/BJAR/article/view/9229

Badmus, M.A., & Ariyo, O. (2011). Forecasting cultivation area and production of Maize in Nigeria using ARIMA model. *Asian Journal of Agricultural Sciences, 3* (3), 171 - 176.

Biswas, B., Dhaliwal, L. K., Singh, S. P., & Sandhu, S. K. (2014). Forecasting wheat production using ARIMA model in Punjab. *International Journal of Agricultural Sciences, 10* (1), 158 - 161.

Darekar, A., & Reddy, A. A. (2017). Forecasting of common paddy prices in India. *Journal of Rice Research, 10* (1), 71-75.

Deshpande, T.(2017). *State of agriculture in India : Report.* PRS legislative Research. Retrieved from www.prsindia.org/parliamenttrack/analytical-reports/state-of-agriculture-in-india-4669/

Gujarati, D.N., & Sangeeta. (2007). *Basic econometrics* (4th ed.) New Delhi : Tata McGraw-Hill .

Gurung, B., Panwar, S., Singh, K. N., Banerjee, R., Gurung, S. R., & Rathore, A.(2017). Wheat yield forecasting using detrended yield over a sub-humid climatic environment in five districts of Uttar Pradesh, India. *Indian Journal of Agricultural Sciences, 87*(1), 87-91.

Ministry of Finance, Dept of Economic Affairs. (2017). *Economic Survey 2016-17*. New Delhi : Government of India.

Mishra, P., Sahu, P.K., Padmanaban, K., Vishwajith, K.P., & Dhekale, B.S.(2015). Study of instability and forecasting of food grain production in India. *International Journal of Agriculture Sciences, 7* (3), 474 - 481.

Nazeem, S.M. (1998). *Applied time series analysis for business and economic forecasting.* New York : Marcel Dekker, Inc.

Padhan, P.C. (2012). Application of ARIMA model for forecasting agricultural productivity in India. *Journal of Agriculture and Social Sciences, 8,* 50 - 56. doi: 11-017/AWB/2012/8-2-50-56

Pindyck, R. S., & Rubinfeld, D.L. (1998). *Econometric models and economic forecasts* (4th ed.) Boston : McGraw-Hill.

Prabakaran, K., & Sivapragasam, C. (2014). Forecasting areas and production of rice in India using ARIMA model. *International Journal of Farm Sciences, 4*(1), 99-106. doi: https://www.inflibnet.ac.in/ojs/index.php/IJFS/article/viewFile/2341/1905

Sahu, P.K., Mishra, P., Dhekale, B.S., Vishwajith, K.P., & Padmanaban, K. (2015). Modelling and forecasting of area, production, yield and total seeds of rice and wheat in SAARC countries and the world towards food security. *American Journal of Applied Mathematics and Statistics, 3* (1), 34 - 48. doi: 10.12691/ajams-3-1-7.

Tripathi, R., Nayak, A.K., Raja, R., Shahid, M., Kumar, A., Mohanty, S., Panda, B.B., Lal, B., & Gautam, P. (2014). Forecasting rice productivity and production of Odisha, India, using autoregressive integrated moving average models. *Advances in Agriculture.* doi:10.1155/2014/621313

Xin, W., & Can, W. (2016). Empirical study on agricultural products price forecasting based on internet-based timely price information. *International Journal of Advanced Sciences and Technology, 87,* 31 - 36. doi: 10.14257/ijast.2016.87.04

About the Author

Dr. Pushpa Savadatti, trained in the London School of Economics, is a Chairperson, Department of Economics and Dean, School of Business Studies at Central University of Karnataka, Kalburgi, Karnataka. She has three decades of research and teaching experience in Economics. Her areas of research interest are Agriculture Economics and Applied Econometrics.