



## State of the art in total pulse production in major states of India using ARIMA techniques

Pradeep Mishra<sup>a</sup>, Aynur Yonar<sup>b</sup>, Harun Yonar<sup>c</sup>, Binita Kumari<sup>d</sup>, Mostafa Abotaleb<sup>e,\*</sup>, Soumitra Sankar Das<sup>f</sup>, S.G. Patil<sup>g</sup>

<sup>a</sup> College of Agriculture, Powarkheda, Jawaharlal Nehru Krishi Vishwa Vidyalaya, Hoshangabad, Madhya Pradesh, India

<sup>b</sup> Department of Statistics, Faculty of Science, Selçuk University, Konya, Turkey

<sup>c</sup> Department of Biostatistics, Faculty of Veterinary Medicine, Selçuk University, Konya, Turkey

<sup>d</sup> Department of Agricultural Economics, Rashtriya Kisan (PG) College, Shamli (affiliated to Chaudhary Charan Singh University, Meerut), India

<sup>e</sup> Department of System Programming, South Ural State University, Chelyabinsk, Russia

<sup>f</sup> Birsa Agricultural University, Kanke, Ranchi, Jharkhand, India

<sup>g</sup> Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

### ARTICLE INFO

#### Keywords:

Autoregressive integrated moving average

Autocorrelation function

Forecasting

Modeling

Production

### ABSTRACT

Pulses are staple protein-rich food for Indian vegetarians, and India is one of the largest producers in the world. The present investigation is an attempt to study the trend in the production of total pulses in India using the autoregressive integrated moving average (ARIMA) method. For stochastic trend estimation, yearly data were used for the period from 1961 to 2019. On the basis of the performance of several goodness of model fit criteria, the most suitable ARIMA model is chosen to capture the trend of pulse production. Forecasting for the 10 years from 2020 to 2029 is done, and it is observed that India has the highest forecast value (31.03302 million tonnes) in 2029. This study will play an important role in determining the gap between production of and demand for pulses in the future.

### 1. Introduction

Pulses are leguminous edible dry seeds rich in protein, minerals, and fibers. They play a diverse role in agriculture as a food crop, provender, cash crop, and rotation crop or intercrop. Many sustainable development goals aiming to achieve sustainable development by 2030 cannot be met without inclusion of pulses in our consumption and production basket (Rawal and Navarro, 2019). Pulses as a commodity group fit in all the Feed the Future Initiative themes aiming toward sustainable poverty and hunger reduction and enhancement of nutrition and health conditions alongside protection of the environment (Maredia, 2012).

Total pulse production worldwide was recorded as 92.28 million tonnes in 2018 (FAO, 2018), of which the major pulses were dry beans (32.98%), chickpeas (18.63%), peas (13.53%), cowpeas (7.83%), lentils (6.86%), and pigeon peas (6.45%). India is the main producer of pulses for most nations, and accounts for 25% of worldwide production. It is also the leading consumer of pulses, with 27% of global consumption (Srivastava et al., 2010). Although India is the largest producer of pulses (23020 tonnes in 2019), domestic production is not sufficient to meet

internal demand, and the country has to import 3 million to 5 million tonnes (15% of global imports) of pulses every year, making it the top pulse importer worldwide (Suresh and Reddy, 2016). Despite the imports, in 2019, the consumption of pulses in India amounted to 48 g per capita per day, slightly less than the 50 g per capita per day recommendation of the Indian Medical Research Council. One of the major hurdles in meeting self-sufficiency in pulses is policies that promote staple crop production, such as subsidies for fertilizers and credit and irrigation facilities that discourage the production of pulses and other legumes (FAO et al., 2020).

In India, pulses are grown mostly under rain-fed conditions. Besides other external factors, erratic rainfall has a serious impact on the production of pulses (Reddy, 2009). There is already a demand and supply gap for pulses in the country, and the uncertainty caused by vagaries in rainfall further widens the gap. Therefore, forecasting production, productivity, and prices is important for effective planning and decision-making related to the production of pulses. The time-series approach of forecasting is the most reliable one. On the basis of the past pattern in data, a very common method applied for forecasting a

\* Corresponding author.

E-mail address: [abotalebmostafa@bk.ru](mailto:abotalebmostafa@bk.ru) (M. Abotaleb).

<https://doi.org/10.1016/j.crfs.2021.10.009>

Received 12 February 2021; Received in revised form 18 October 2021; Accepted 22 October 2021

Available online 28 October 2021

2665-9271/© 2021 The Authors.

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Table 1**  
Descriptive statistics of total pulse production data.

State or country	Minimum	Maximum	Mean	Standard deviation	Skewness	Kurtosis
Karnataka	239.00	1951.21	697.94	421.19	1.54	1.78
Madhya Pradesh	873.00	8111.58	2803.88	1509.61	1.82	3.72
Maharashtra	451.00	3768.06	1470.84	723.27	1.13	0.84
Rajasthan	227.00	3700.00	1509.11	773.86	0.97	0.87
Uttar Pradesh	1212.00	3835.00	2535.55	576.71	-0.01	0.21
India	8347.00	25416.22	13110.88	3531.55	1.69	3.10

time series (Ray and Bhattacharyya, 2020) is the autoregressive integrated moving average (ARIMA) method. In a study by Vishwajith et al. (2018) on forecasting mung production, ARIMA(4,1,4) was the best-fitting model over ARIMAX and generalized autoregressive conditional heteroscedasticity (GARCH) models. Mishra et al. (2021) considered ARIMA models for forecasting of sugarcane production by major states for 2025.

In contrast, Ray and Bhattacharyya (2020) found the ARIMAX(1,1,1) model for pulse production better suited than the ARIMA model. Price prediction is an important tool to forecast the market price, which is necessary for framing policies for sustained production and remunerative prices (Darekar and Reddy, 2017). Savadatti (2017) applied the ARIMA model for projection purposes and observed stagnancy in the area of pulse production but a rise in pulse production and productivity. Many other studies have used the ARIMA model for forecasting; for example, for forecasting sugarcane production (Muhammad et al., 1992) and sugarcane and cotton crop production and yield (Ali et al., 2015) in Pakistan. In Tamil Nadu, ARIMA models were used for area, production, and productivity forecasting for various crops (Balanagammel et al., 2000) and sugarcane yield (Suresh and Krishna Priya, 2011). Conditional variances are taken into account by use of the GARCH model. Yaziz et al. (2011) used both the GARCH model and the ARIMA model to predict crude oil prices, and they concluded that the GARCH model was superior to the ARIMA model. However, Vishwajith et al. (2014) could not establish the superiority of either the GARCH model or the ARIMA model in modeling data for pulses in India.

**2. Material and methods**

In the present investigation, the data relate to total pulse production for five major producing states and India from 1950 to 2019. To set model structure, 80% of the total data is selected for training and to approve the model, the remaining 20% is chosen for the test. The statistical software package R was used for model building.

**2.1. ARIMA models**

Box and Jenkins (1976) introduced the ARIMA model, and thus it is also known as the Box-Jenkins method in the literature. ARMA model includes an autoregressive (AR) and a moving average (MA) model. While these models are suitable for stationary series, the ARIMA model is performed in nonstationary series (Tekindal et al., 2020).

The easiest way to make a time series stationary is by taking the difference. The process of subtracting the values of a certain period from the last values of the time series is called the difference operation.

In non-stationary data, ARMA(p, q) model is known as the ARIMA(p, d, q) models if the d-order difference operation is performed to make the data stationary. In the kind of equations in the ARIMA (p, d, q) models, p represents the degree of the AR model, q represents the degree of the MA model, and d represents the number of differences needed to stabilize the data (Yonar et al., 2020; Ray et al., 2021; Mishra et al., 2021).

The equation for the ARIMA(p, d, q) model is as follows:

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + \alpha_1 - \theta_1 \alpha_{t-1} - \alpha_2 - \theta_2 \alpha_{t-2} - \dots - \alpha_q - \theta_q \alpha_{t-q}, \tag{1}$$

where  $\varphi_p$  represents the parameter values relating to the AR operator,  $\alpha_q$  is the error term coefficient,  $\theta_q$  represents the parameter values relating to the MA operator, and  $Y_t$  represents the data with  $d$ th differences of the original data (Brockwell et al., 2016; Gujarati and Porter, 2012).

The following steps can be applied for fitting time-series data to an ARIMA model (Hyndman and Khandakar, 2007).

**Step 1.** Plot the data, detect any unusual observations, and transform the data to stabilize the variance if necessary.

**Step 2.** Determine the values of p and q by analyzing the autocorrelation function (ACF) and the partial ACF (PACF), and from the selected model try to identify the best ARIMA model by using Akaike’s information criterion (AIC) with correction (AICc).

**Step 3.** From the best model, check the residuals by plotting the ACF and the PACF. Try to modify the model if the plotted ACF and PACF do not look like white noise.

**Step 4.** If the residuals look like white noise, calculate the forecasts.

For model selection, the mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), AIC, AICc, and Bayesian information criterion (BIC) are statistical measures that evaluate the performance of the fit of the forecasting. The goodness-of-fit approach for a time series is based on the residuals. Therefore, the measure of forecast accuracy should be checked using many metrics together. The model with the best forecasting ability has the smallest error criterion value.

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2,$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2},$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |e_t|,$$

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \frac{|e_t|}{\hat{y}_t},$$

Where  $e_t$  is the error term,  $y_t$  is the observation, and  $\hat{y}_t$  is the forecast; also  $e_t = y_t - \hat{y}_t$ .

$$AIC = -2LL + 2k$$

$$AIC_c = AIC + \frac{2(k^2 + 1)}{n - k - 1}$$

$$BIC = k \ln(n) - 2 \ln \hat{L}$$

Here, k indicates the number of estimated parameters in the model and L stands the maximum value of the likelihood function for the model, n denotes sample sizes.

In this study, statistical analysis was performed with the statistical

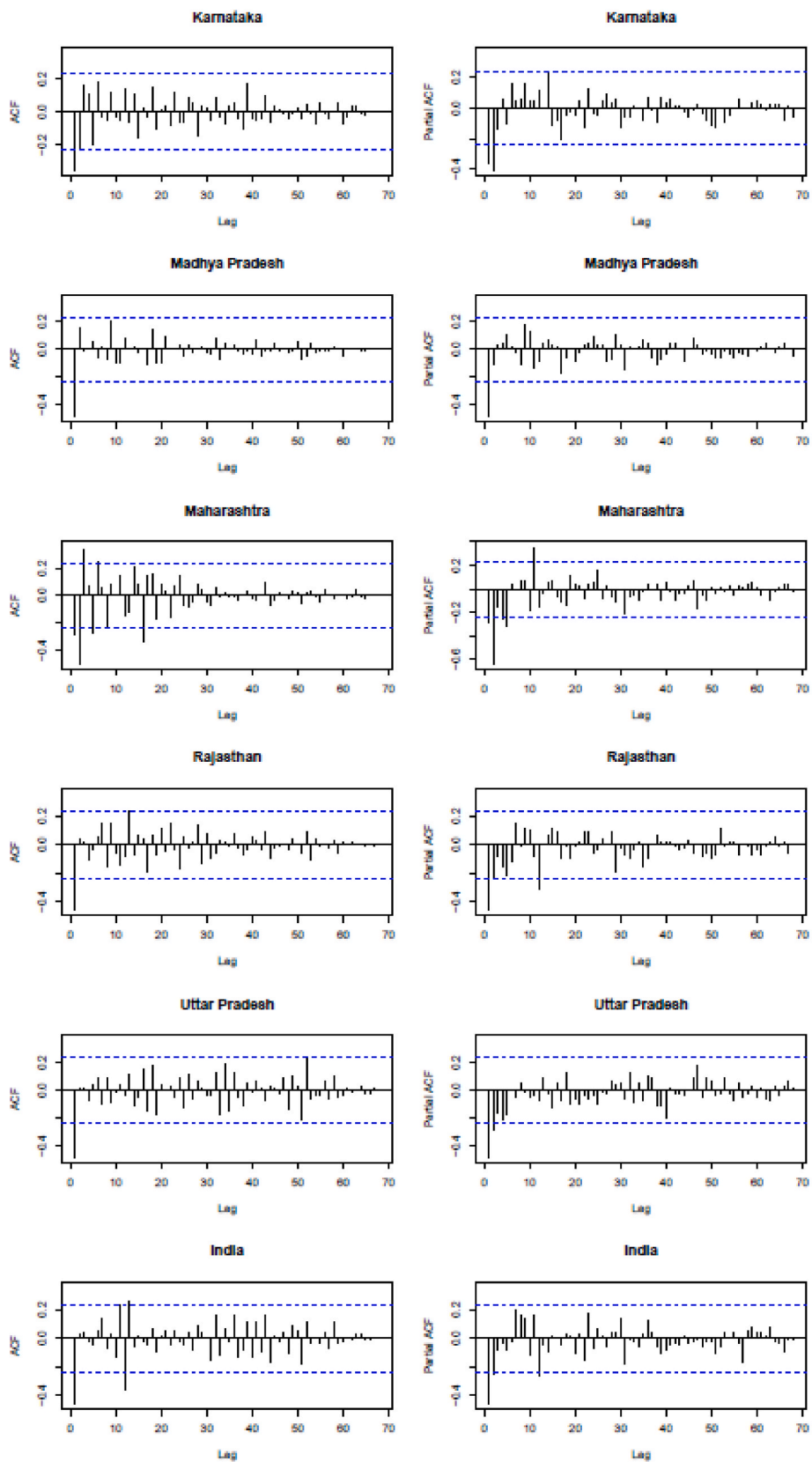


Fig. 1. Autocorrelation function (ACF) and partial ACF graphs of the first differences of the data in pulses production in Major states in India using ARIMA models.

**Table 2**  
Model fitting for total pulse production data.

State or country	Model	ME	RMSE	MAE	MPE	MAPE	MASE	LL	AIC	AICc	BIC
Karnataka	ARIMA(0,1,1)	-0.77	153.22	111.28	-7.04	17.94	0.87	-445.80	897.59	897.69	904.29
Madhya Pradesh	ARIMA(1,1,0)	-0.18	590.21	378.94	-6.85	16.95	0.85	-538.79	1083.58	1083.95	1090.28
Maharashtra	ARIMA(2,1,1)	-0.50	340.47	254.81	-6.55	19.28	0.77	-501.66	1013.32	1014.27	1024.49
Rajasthan	ARIMA(0,1,1)	2.54	562.45	425.73	-16.06	37.15	0.85	-535.61	1077.21	1077.58	1083.91
Uttar Pradesh	ARIMA(0,1,1)	10.12	388.00	285.48	-2.71	12.93	0.79	-510.25	1026.50	1026.87	1033.20
India	ARIMA(1,1,2)	-42.17	1509.31	1148.00	-1.73	9.17	0.79	-604.27	1218.55	1219.50	1299.72

AIC, Akaike’s information criterion; AICc, Akaike’s information criterion with correction; ARIMA, autoregressive integrated moving average; BIC, Bayesian information criterion; LL, likelihood; MAE, mean absolute error; ME, mean error; MAPE, mean absolute percentage error; MASE, mean absolute scaled error; MPE, mean percentage error; RMSE, root mean squared error.

software package R. The auto. Arima() function in R was used to determine the best ARIMA model. This function uses the Hyndman-Khandakar algorithm (Hyndman and Khandakar, 2007) to obtain an ARIMA model by combining the unit root test and minimization of the AICc. The Hyndman-Khandakar algorithm involves only Steps 2 and 3 given above; therefore, we conducted the other steps ourselves. Furthermore, the Arima() function in R (Version 4.03, <https://stat.ethz.ch/pipermail/r-announce/2020/000662.html>) was used to test other models we thought may be suitable on the basis of ACF and PACF graphs.

### 3. Results and discussion

Descriptive statistics for total pulse production between 1950 and 2019 are presented in Table 1 for Karnataka, Madhya Pradesh, Maharashtra, Rajasthan, Uttar Pradesh, and India. While the minimum value for total pulse production in India is 8347, the maximum value is 25416.22 between 1950 and 2019. Thus, it can be said that there has been an increase of approximately 205% in total pulse production in India since 1950. Moreover, in Karnataka, Madhya Pradesh, Maharashtra, Rajasthan, and Uttar Pradesh, production increased by approximately 716%, 830%, 735%, 1530%, and 230%, respectively. As a result, the mean total pulse production is 13110.88 for India. The state with the highest average pulse production is Madhya Pradesh, with 2803.88, and the lowest is Karnataka, with 697.94. When the standard deviations of pulse production are analyzed, the highest standard deviation is seen in India, with 3531.55 (Table 1).

To test the normality of the data set, the descriptive statistics are calculated and then by dividing the skewness coefficient by the standard errors, the normality of the data is tested (Das et al., 2017). Skewness and kurtosis should be within the range from +2 to -2 (a few authors also use the more lenient +3 to -3). With use of this rule of thumb and as the data size is large, it can be concluded that the data sets are normally distributed. The positively skewed and platykurtic nature of the data for the states of Karnataka, Maharashtra, and Rajasthan indicate that there was a marginal change of the area in favor of pulse production during the early period and it remained almost the same in the study period (Vishwajit et al., 2018). The leptokurtic and positively skewed nature of the data for Madhya Pradesh and India indicates a very marginal change of the area during the early period. In Uttar Pradesh, the platykurtic and negatively skewed nature of the data indicates a marginal change of the area during the late period, and it remained almost the same during the of the study.

ACF and PACF graphs of the first differences of the series are presented in Fig. 1. According to the goodness-of-fit parameter, especially the AICc, the best models given in Table 2 were selected for five major states and India, and the best model was used to determine the forecast value for the 10 years from 2020 to 2029. It is clearly seen from Fig. 1 and Table 2 that the data became stationary at first differences. As seen in Table 2, the goodness of fit of ARIMA models was also assessed by

various information criteria, such as the mean error, RMSE, MAE, mean percentage error, MAPE, mean absolute scaled error, likelihood, AIC, BIC, and the AICc used in the Hyndman-Khandakar algorithm.

The residuals from the fitted models were checked by means of graphs of the ACF and PACF concerning the residuals given in Fig. 2A corrected result is observed if all autocorrelations are within the threshold limits and the residuals look like white noise.

To assess future quantity on the basis of recent information, the application mainly uses time series in forecasting models (Das et al., 2019). The present investigation aimed to establish the importance of ARIMA models and attempted to make short-term predictions for pulses in India and Indian states. The forecast values were generated between the years 2020 and 2029. Tables 3–8 show point forecasts and 80% and 95% prediction intervals obtained through the ARIMA models in Table 2 for five major states and India, respectively. Lo80 and Hi80 are the lower and upper bounds of the prediction interval for significance level  $\alpha = 0.20$ , and Lo95 and Hi95 are the lower and higher bounds of the prediction interval for significance level  $\alpha = 0.05$ . From these tables, it can be said that total pulse production will show a continuous increase in Karnataka, Madhya Pradesh, Rajasthan, and India and will decrease continuously in Uttar Pradesh (depicted in Fig. 1).

Pulse production in Karnataka, Madhya Pradesh, Rajasthan, and India is expected to follow a rising trend in the next few years on the basis of the forecast. It was estimated that in 2029, pulse production will reach 2090.95 in Karnataka (Table 3), 8705.09 in Madhya Pradesh (Table 4), 3846.63 in Rajasthan (Table 6), and 31033.02 in India (Table 8). However in Maharashtra, the pulse production trend is highly fluctuating (Table 5), and this is also clearly visible from Fig. 3. In Uttar Pradesh (Table 7), the production trend is for a decrease, and this is also supported by Fig. 3. Agricultural funding, price support programs, better management practice, research workers, etc., for long-term production will be the major factors to sustain this trend.

### 4. Conclusion

Pulses are an important part of a healthy, well-balanced diet, and they are particularly prevalent in the Indian diet. Pulses are classified as both a mixed crop and an intercrop in the agricultural industry. Making predictions about pulse production would aid in determining whether or not demand would be met in the foreseeable future. It is indisputably established by the outcomes of this study that India would have the highest predicted value in the year 2029. Uttar Pradesh is experiencing a decline in overall pulse output, whilst Karnataka, Madhya Pradesh, and Rajasthan are experiencing an increase in production. Agriculture funding, price support programs, improved management practices, research employees, and other variables that will contribute to long-term output will be the most important factors in maintaining this trend. This type of project aids in the implementation of policy and the long-term planning for a particular crop.

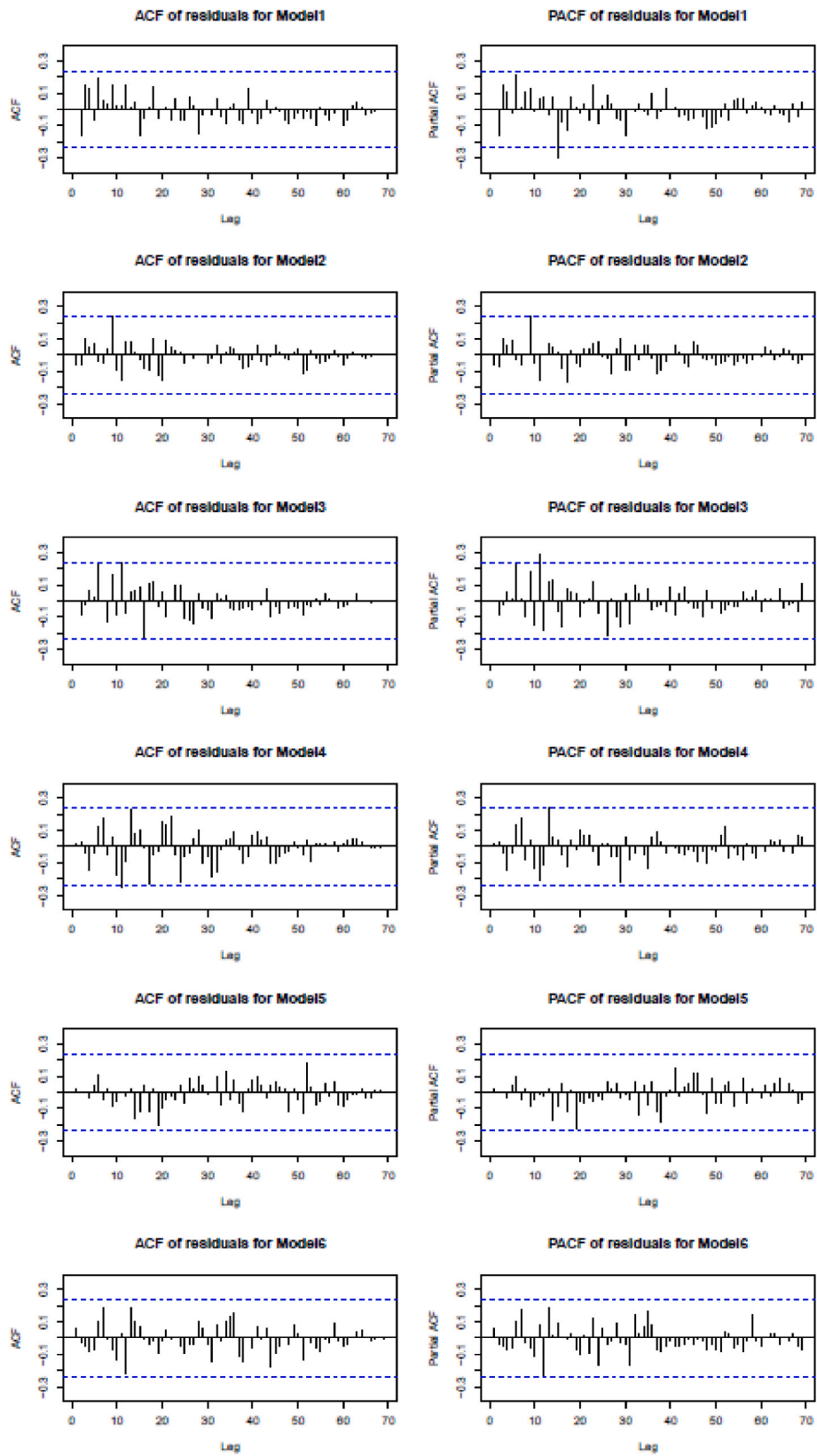


Fig. 2. Autocorrelation function (ACF) and partial ACF(PACF) of the residuals of the fitted models on pulses production in Major states in India using ARIMA models.

**Table 3**  
Total pulse production forecasting for Karnataka(millions of kilograms).

Year	Forecast	Lo80	Hi80	Lo95	Hi95
2020	1886.47	1685.75	2087.18	1579.50	2193.43
2021	1909.19	1690.25	2128.13	1574.35	2244.03
2022	1931.91	1696.15	2167.67	1571.34	2292.47
2023	1954.63	1703.17	2206.09	1570.05	2339.20
2024	1977.35	1711.11	2243.58	1570.18	2384.52
2025	2000.07	1719.84	2280.30	1571.49	2428.65
2026	2022.79	1729.23	2316.35	1573.83	2471.75
2027	2045.51	1739.20	2351.82	1577.05	2513.97
2028	2068.23	1749.68	2386.78	1581.05	2555.41
2029	2090.95	1760.61	2421.29	1585.74	2596.16

Lo80, lower bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Hi80, upper bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Lo95, lower bound of the predictive interval for significance level  $\alpha = 0.05$ ;  
 Hi95, higher bound of the predictive interval for significance level  $\alpha = 0.05$ .

**Table 4**  
Total pulse production forecasting for Madhya Pradesh(millions of kilograms).

Year	Forecast	Lo80	Hi80	Lo95	Hi95
2020	7888.21	7115.08	8661.35	6705.80	9070.63
2021	7931.14	7061.58	8800.70	6601.26	9261.02
2022	8050.87	7005.62	9096.13	6452.29	9649.45
2023	8133.33	6978.21	9288.46	6366.72	9899.95
2024	8233.88	6961.06	9506.70	6287.27	10180.49
2025	8325.65	6953.06	9698.25	6226.45	10424.86
2026	8421.68	6952.52	9890.84	6174.80	10668.57
2027	8515.65	6957.53	10073.77	6132.71	10898.59
2028	8610.61	6967.59	10253.64	6097.83	11123.40
2029	8705.09	6981.69	10428.49	6069.38	11340.81

Lo80, lower bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Hi80, upper bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Lo95, lower bound of the predictive interval for significance level  $\alpha = 0.05$ ;  
 Hi95, higher bound of the predictive interval for significance level  $\alpha = 0.05$ .

**Table 5**  
Total pulse production forecasting for Maharashtra(millions of kilograms).

Year	Forecast	Lo80	Hi80	Lo95	Hi95
2020	3160.45	2707.65	3613.25	2467.95	3852.95
2021	2991.43	2529.40	3453.45	2284.82	3698.04
2022	2497.88	2023.12	2972.63	1771.81	3223.95
2023	2738.69	2242.58	3234.80	1979.96	3497.42
2024	3033.16	2492.23	3574.09	2205.88	3860.45
2025	2880.08	2336.73	3423.43	2049.10	3711.07
2026	2780.13	2235.32	3324.94	1946.91	3613.34
2027	2936.62	2374.14	3499.10	2076.38	3796.86
2028	3012.71	2435.93	3589.50	2130.60	3894.83
2029	2951.37	2370.31	3532.43	2062.71	3840.02

Lo80, lower bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Hi80, upper bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Lo95, lower bound of the predictive interval for significance level  $\alpha = 0.05$ ;  
 Hi95, higher bound of the predictive interval for significance level  $\alpha = 0.05$ .

**Author statement**

All authors read and approved the final manuscript. Preparation of original manuscript: PM, AY, and HY. Data compilation: BK, MA, SSD, and SGP. Coding and analysis: BK, SSD, and SGP. Results and discussion: PM, AY, and HY. Finalization of the manuscript: PM, AY, HY, and MA.

**Credit authors statement**

All authors have read and approved the final manuscript.

**Table 6**  
Total pulse production forecasting for Rajasthan(millions of kilograms).

Year	Forecast	Lo80	Hi80	Lo95	Hi95
2020	3454.25	2717.47	4191.02	2327.45	4581.04
2021	3497.84	2717.11	4278.58	2303.81	4691.88
2022	3541.44	2719.09	4363.79	2283.77	4799.12
2023	3585.04	2723.08	4447.00	2266.79	4903.30
2024	3628.64	2728.82	4528.47	2252.48	5004.81
2025	3672.24	2736.08	4608.40	2240.50	5103.98
2026	3715.84	2744.70	4686.98	2230.61	5201.07
2027	3759.44	2754.54	4764.34	2222.57	5296.30
2028	3803.04	2765.47	4840.60	2216.22	5389.85
2029	3846.63	2777.41	4915.86	2211.39	5481.88

Lo80, lower bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Hi80, upper bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Lo95, lower bound of the predictive interval for significance level  $\alpha = 0.05$ ;  
 Hi95, higher bound of the predictive interval for significance level  $\alpha = 0.05$ .

**Table 7**  
Total pulse production forecasting for Uttar Pradesh(millions of kilograms).

Year	Forecast	Lo80	Hi80	Lo95	Hi95
2020	1621.39	1113.15	2129.64	844.10	2398.69
2021	1599.40	1082.28	2116.51	808.54	2390.26
2022	1577.41	1051.57	2103.24	773.21	2381.60
2023	1555.41	1021.00	2089.82	738.10	2372.72
2024	1533.42	990.57	2076.27	703.20	2363.64
2025	1511.43	960.26	2062.59	668.50	2354.36
2026	1489.43	930.08	2048.78	633.98	2344.89
2027	1467.44	900.02	2034.86	599.64	2335.24
2028	1445.45	870.07	2020.83	565.48	2325.41
2029	1423.45	840.23	2006.68	531.49	2315.42

Lo80, lower bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Hi80, upper bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Lo95, lower bound of the predictive interval for significance level  $\alpha = 0.05$ ;  
 Hi95, higher bound of the predictive interval for significance level  $\alpha = 0.05$ .

**Table 8**  
Total pulse production forecasting for India(millions of kilograms).

Year	Forecast	Lo80	Hi80	Lo95	Hi95
2020	24989.44	22982.17	26996.71	21919.58	28059.29
2021	25895.95	23837.77	27954.12	22748.24	29043.65
2022	26722.13	24545.24	28899.02	23392.86	30051.39
2023	27479.32	25113.14	29845.50	23860.56	31098.08
2024	28177.26	25561.62	30792.91	24176.98	32177.55
2025	28824.32	25914.12	31734.53	24373.55	33275.10
2026	29427.68	26192.12	32663.25	24479.31	34376.06
2027	29993.51	26413.22	33573.80	24517.93	35469.09
2028	30527.11	26591.24	34462.98	24507.71	36546.50
2029	31033.02	26736.77	35329.27	24462.48	37603.57

Lo80, lower bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Hi80, upper bound of the predictive interval for significance level  $\alpha = 0.20$ ;  
 Lo95, lower bound of the predictive interval for significance level  $\alpha = 0.05$ ;  
 Hi95, higher bound of the predictive interval for significance level  $\alpha = 0.05$ .

Preparation of original manuscript: PM, AY, and HY. Data compilation: BK, MA, SSD and SGP. Coding and analysis: BK SSD & SGP. Results and discussion: PM, AY, HY,. Finalization of the manuscript: PM, AY, HY, and MA.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

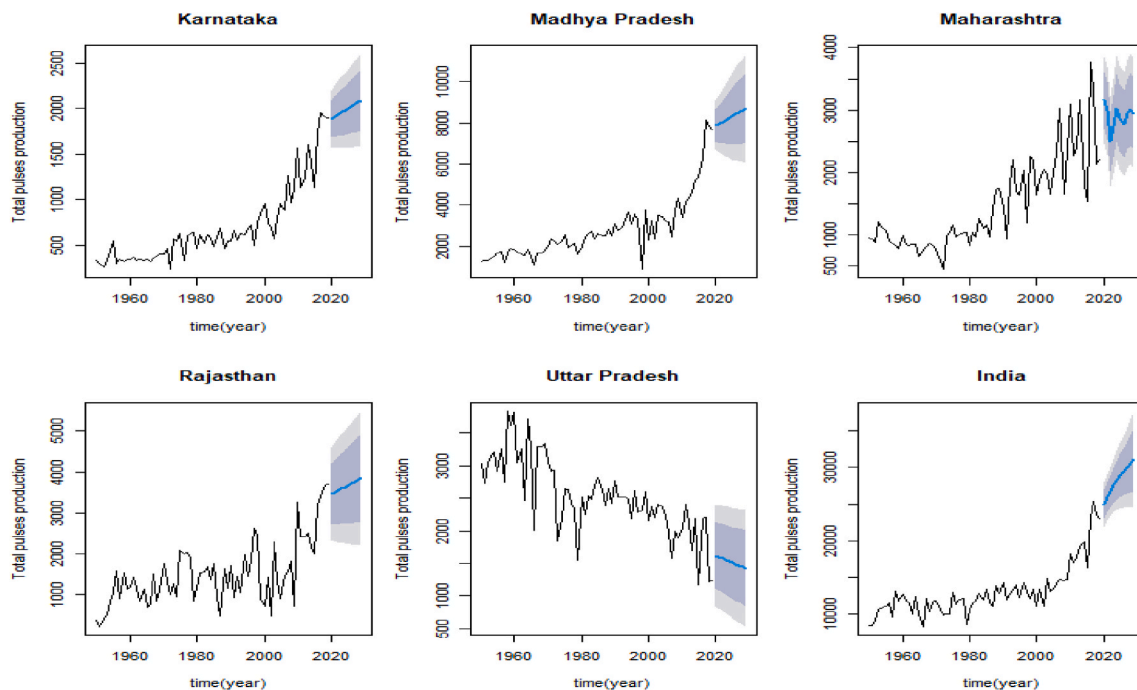


Fig. 3. Point forecasts and 80% and 95% prediction intervals obtained with autoregressive integrated moving average models on pulses production for major states in India. (From 2020–2029).

## Acknowledgment

The work was supported by Act 211 Government of the Russian Federation, contract No. 02.A03.21.0011. The work was supported by the Ministry of Science and Higher Education of the Russian Federation (government order FENU-2020-0022).

## References

- Ali, S., Badar, N., Fatima, H., 2015. Forecasting production and yield of sugarcane and cotton crops of Pakistan for 2013-2030. *Sarhad J. Agric.* 31 (1), 1–10.
- Balanagammal, D., Ranganathan, C.R., Sundaresan, K., 2000. Forecasting of agricultural scenario in Tamil Nadu: a time series analysis. *J. Indian Soc. Agric. Stat.* 53 (3), 273–286.
- Box, G.E., Jenkins, G.M., 1976. *Time series analysis, control, and forecasting*. San Francisco: Holden Day 3226 (3228), 10.
- Brockwell, P.J., Davis, R.A., 2016. *Introduction to Time Series and Forecasting*. Springer.
- Darekar, A., Reddy, A.A., 2017. Price forecasting of pulses: the case of pigeon pea. *Journal of Food Legumes* 30 (3), 212–216.
- Das, S.S., Das, K., Debnath, M.M., 2017. Different methods for judging the normality assumption for univariate and bivariate data and its remedial measure. *Journal of Crop and Weed* 13 (1), 7–16.
- Das, S.S., Ray, S., Sen, A., Siva, G.S., Das, S., 2019. Statistical study on modeling and forecasting of jute production in West Bengal, India. *International Journal of Current Microbiology and Applied Sciences* 8 (7), 1719–1730.
- FAO, 2018. *Food Outlook*. FAO, Rome.
- FAO, IFAD, UNICEF, WFP, WHO, 2020. *The State of Food Security and Nutrition in the World 2020. Transforming Food Systems for Affordable Healthy Diets*. FAO, Rome. <https://doi.org/10.4060/ca9692en>.
- Gujarati, D.N., Porter, D.C., Gunasekar, S., 2012. *Basic Econometrics*. Tata McGraw-Hill Education.
- Hyndman, R.J., Khandakar, Y., 2007. *Automatic Time Series for Forecasting: the Forecast Package for R* (no. 6/07). Monash University, Department of Econometrics and Business Statistics, Clayton.
- Maredia, M., 2012. Global pulse production and consumption trends: the potential of pulses to achieve 'Feed the Future' food and nutritional security. In: Presented at Global Pulse Researchers Meeting, Rwanda, vols. 13–19. February.
- Muhammad, F., Javed, M.S., Bashir, M., 1992. Forecasting sugarcane production in Pakistan using ARIMA models. *Pakistan J. Agric. Sci.* 9 (1), 31–36.
- Mishra, P., Al Khatib, M.G., Sardar, I., et al., 2021a. Modeling and forecasting of sugarcane production in India. *Sugar Tech.* <https://doi.org/10.1007/s12355-021-01004-3>.
- Mishra, P., Matuka, A., Abotaleb, M.S.A., Weerasinghe, W.P.M.C.N., Karakaya, K., Das, S.S., 2021b. Modeling and forecasting of milk production in the SAARC countries and China. *Modeling Earth Systems and Environment*. <https://doi.org/10.1007/s40808-021-01138-z>.
- Rawal, V., Navarro, D.K., 2019. *The Global Economy of Pulses*. FAO, Rome.
- Ray, S., Bhattacharyya, B., 2020. Time series modelling and forecasting of pulses production behaviour of India. *Indian J. Ecol.* 47 (4), 1140–1149.
- Ray, S., Das, S.S., Mishra, P., Al Khatib, A.M.G., 2021. Time series SARIMA Modelling and forecasting of monthly rainfall and temperature in the south Asian countries. *Earth Systems and Environment* 5 (3), 531–546.
- Reddy, A.A., 2009. Pulses production technology: status and way forward. *Econ. Polit. Wkly.* 44, 73–80.
- Savadatti, P.M., 2017. Trend and forecasting analysis of area, production and productivity of total pulses in India. *Indian Journal of Economics and Development* 5 (12), 1–10.
- Srivastava, S.K., Sivaramne, N., Mathur, V.C., 2010. Diagnosis of pulses performance of India. *Agric. Econ. Res. Rev.* 23, 137–148.
- Suresh, K.K., Krishna Priya, S.R., 2011. Forecasting sugarcane yield of Tamilnadu using ARIMA models. *Sugar Tech* 13 (1), 23–26.
- Suresh, A., Reddy, A.A., 2016. Total factor productivity of major pulse crops in India: implications for technology policy and nutritional security. *Agric. Econ. Res. Rev.* 29, 87–98.
- Tekindal, M.A., Yonar, H., Yonar, A., Tekindal, M., Çevrimli, M.B., Alkan, H., İnanç, Z.S., Mat, B., 2020. Analyzing COVID-19 outbreak for Turkey and eight country with curve estimation models, Box-Jenkins (ARIMA), Brown linear exponential smoothing method, autoregressive distributed lag (ARDL) and SEIR models. *Eurasian Journal of Veterinary Science, Covid-19 Special Issue* 142–155.
- Vishwajith, K.P., Dhekale, B.S., Sahu, P.K., Mishra, P., Noman, M.D., 2014. Time series modelling and forecasting of pulses production in India. *Journal of Crop and Weed* 10 (2), 147–154.
- Vishwajith, K.P., Sahu, P.K., Mishra, P., Dhekale, B.S., Singh, R.B., 2018. Modelling and forecasting of Arhar production in India. *International Journal of Agricultural and Statistical Sciences* 14 (1), 73–86.
- Yaziz, S.R., Ahmad, M.H., Nian, L.C., Muhammad, N., 2011. A comparative study on Box-Jenkins and GARCH models in forecasting crude oil prices. *J. Appl. Sci.* 11 (7), 1129–1135.
- Yonar, H., Yonar, A., Tekindal, M.A., Tekindal, M., 2020. Modeling and forecasting for the number of cases of the COVID-19 pandemic with the curve estimation models, the Box-Jenkins and exponential smoothing methods. *EJMO* 4 (2), 160–165.