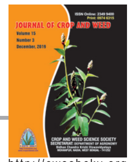
Journal of Crop and Weed, 17(1): 48-55 (2021)

ISSN- O : 2349 9400 ; P : 0974 6315



http://cwssbckv.org
www.cropandweed.com

Paddy price forecasting in India using ARIMA model

*B. KATHAYAT AND A.K DIXIT

DESM Division, ICAR-NDRI, Karnal (Haryana)

Received : 07.07.2020 ; Revised : 22.12.2020 ; Accepted : 31.12.2020

DOI : https://doi.org/10.22271/09746315.2021.v17.i1.1405

ABSTRACT

The year 2020 being an extraordinary year of pandemic is likely to have far reaching effects on the prices of essential food commodities. The study was undertaken with an aim to forecast wholesale paddy prices for the next agricultural year i.e., 2020-21 in five major states namely Punjab, UP, Tamil Nadu, West Bengal and Delhi. ARIMA model has been used to forecast the prices. Analysis of the seasonal price variations reflects that price variations have been more prominent in northern states than others. Results indicate that wholesale prices for the agricultural year 2020-21 would be in the range of Rs1810.23–Rs.2239.59 qt^{-1} for Punjab, Rs.1662.91-Rs.1674.98 qt^{-1} for Tamil Nadu, Rs.3010.00-Rs.3133.36 qt^{-1} for Delhi, Rs.1835.05-Rs.1902.22 $qt^{-1}$for West Bengal and Rs.1080.90-Rs.1495.35 qt^{-1} for Uttar Pradesh. ARIMA models were evaluated using RMSE and MAPE. Best fit ARIMA model for different markets obtained were ARIMA(4,0,12) ARIMA(0,1,6) ARIMA(0,1,12) ,ARIMA(0,1,3) and ARIMA(3,1,12) for Punjab, Tamil Nadu, Delhi, West Bengal and Uttar Pradesh, respectively. Since crucial decisions with respect to production, marketing and consumption are dependent on certain price expectation, findings of the study might prove helpful in guiding the stakeholders involved in making timely decisions.

Keywords : Agricultural marketing, ARIMA model, paddy, price forecast

Paddy which becomes rice after removal of the husk,is an important staple food crop for more than 50 per cent people all around the world especially in South East Asia. Asia alone accounts for more than 90.7% of world rice production share (FAO,2018), China, India, and Indonesia being the leading producers in the region. Apart from being major consumers,India also accounts for more than 20% share in global production levels with an estimated area of 43.19 mha under paddy cultivation in 2016-17(DAC,MOA&FW,2018). Due to increasing population growth, demand for paddy is consistently rising from 1950s onward which evidently explains the upward trend observed in its area under cultivation over time. In terms of volume of output produced India ranks high but productivity levels are still below the world's average and even lower than Egypt, Japan, China, Vietnam, USA and Indonesia (Darekar and Reddy, 2017). Low yield could be attributed to cultivation under less fertile and degraded lands, pest and disease infestation, faulty cropping pattern and lower technological adoption among farmers. In India paddy is mainly cultivated in three different seasons as indicated in Table 1 varying according to region. In north India it is mainly cultivated as *kharif* crop while in south and eastern region it could be grown in all seasons.

Wheat and rice among others are two most vital food items under public distribution management system which are heavily procured by government at Minimum Support Price (MSP) to enable the access to poor at subsidized prices while maintaining adequate support to growers at the same time. Despite reaping bumper harvest in last few seasons, outbreak of COVID-19 has put an additional strain on the food grain reserves. In the light of recent migrant crisis government introduced several relief packages like PM Garib KalyanYojna to provide sufficient food grains mainly wheat, rice and pulses to millions of poor households who have become the worst victims of the pandemic. Consumption dependence of a huge population makes paddy an essential food crop, reduced availability might endanger the food security of entire nation. The year 2020 being an extraordinary year of pandemic inducing a severe crisis like situation on every front is likely to have far reaching effects on prices of such essential food commodities in the upcoming months however nothing concrete on the kind of effect can be inferred at the present situation.

Even though price forecasting is a risky venture due to several unseen factors which may render the forecast invalid (Jadav *et al*., 2017) but accuracy of price forecast for cereal has been reported somewhat higher than the perishable commodities. Forecasting helps in making timely decisions when faced with an uncertain future as witnessed globally in current times.ARIMA model popularly known as the Box-Jenkins model has been widely used in forecasting of time series data. Utilization of the same model has been done in this paper to forecast the paddy prices across five major spatially separated markets namely Punjab, Delhi, Uttar Pradesh, Tamil Nadu and West Bengal.

Email: babitakathayat1@gmail.com

A number of decisions with respect to production, marketing and consumption are dependent on certain price expectation. Therefore it is essential to monitor the market price movement over time. The present study is undertaken with an aim of forecasting paddy prices for the agricultural year 2020-21.

Several other studies have used ARIMA model for the purpose of price forecast. Punitha (2007) applied this model to forecast the arrivals and prices of maize and ground nut in Hubli and Devangere markets of Karnataka. Verma *et al.*(2016) utilized ARIMA modeling for price forecast of coriander in Rajasthan and found ARIMA (0, 1, 1) model to be the best fit. Darekar and Reddy (2018) applied ARIMA modeling in wheat price forecasting and ARIMA (0,1,1) (0,1,1) was found to be suitable fit. Ohyver and Padjihastuti (2018) made medium quality rice price forecast in Indonesia and found ARIMA (1,1,2) to be the suitable fit.

MATERIALS AND METHODS

Table 1: Rice growing seasons in India

Particulars	Crop season	Harvest season	Lean season
Autumn Rice	May to August	Sept to Oct	Nov to April
Winter Rice	June to July	Nov to Dec	Jan to May
Summer Rice	Nov to February	March to June	July to Oct

Table 2: Selection criteria

States	Basis for selection production (% share to total)	Region
Punjab	11.85	North
West Bengal	13.26	East
Uttar Pradesh	11.75	North
Tamil Nadu	6.45	South
Delhi	Nil production but major market	North

Data Collection: The study was undertaken mainly with an objective to forecast paddy prices in major markets. Monthly wholesale prices of paddy spanning from July 2002 to June 2020 were collected for 5 major states (Table 2) in India from the Agmarknet price portal. Analytical tools such as descriptive statistics, instability indices, seasonal indices and ARIMA model were employed to analyze the data and make appropriate forecast.

Seasonal Indices: Seasonal indices of monthly prices were computed using 12 month moving averages to deseasonalize and smoothen the data. By deseasonalizing, we remove the effect of any pattern or seasonal fluctuations in the data.

Instability index: Instability can be measured by different methods such as Coefficient of variation (CV), dispersion, Cuddy Della Valle Index (CDVI) etc. Present study applies CDVI to show the instability in the time-series data. It is a better measure of capturing instability over CV since it first de-trends the series and gives clear indication about the instability. $\bar{R}^2$ Indicates the adjusted R square (Coefficient of determination).

$$CDVI = CVX\sqrt{1-\bar{R}^2}$$

Seasonal Variation: To measure the extent of variation in seasonal indices, the coefficient of average seasonal price variation (ASPV), intra year price rise (IPR) and coefficient of variation (CV) have been computed.

$$\textbf{ASPV} = \left[\frac{HSPI - LSPI}{\frac{HSPI + LSPI}{2}}\right]\times 100$$

$$\textbf{IPR} = \left[\frac{HSPI - LSPI}{LSPI}\right]\times 100$$

Where, LSPI = lowest seasonal price index
HSPI= Highest seasonal price index

$$\textbf{CV} = \left[\frac{\sigma}{\bar{X}}\right]\times 100$$

Where, σ = Standard Deviation
$\bar{X}$ = Mean value

Price forecasting by ARIMA model

ARIMA approach was initially popularized by Box and Jenkins (1976), also known as Box-Jenkins models. ARIMA model is most commonly used to analyze and forecast univariate time series data. It predicts the value in response to the linear combination of its own past time series values. Time series models are generally assume to be stationary *i.e.* the mean and variance for the series are constant and its covariant is time invariant. If the time series is stationary then it can be inferred that ARMA (p, q) is applied. If the series is not stationary, it can be differenced "d" times to make it stationary then the model is referred as ARIMA (p,d,q). The term 'p' indicates order of partial autocorrelation,'d' reflects the order of difference *i.e.* degree of differencing of time series and 'q' indicates order of moving average. When the model is characterized by seasonality element, then it should be extended with seasonal parameters known as the seasonal ARMA (in case of stationary time-series) or SARIMA (in case of non-stationary time series).They are represented as SARIMA(p,d,q)(P,D,Q) where P stands for seasonal autoregressive order, D for seasonal differencing number and S for seasonal moving average order.

The Box-Jenkins methodology is as follows consisting of four steps.

Identification: This is the first step to check the stationarity of the data series since estimation can be done only on stationary series.The structure of autocorrelation and partial correlation coefficients provides clues for the presence of stationarity. If the model is found to be non-stationary, differencing can be done to achieve stationarity in the data series. Seasonality may result from seasonal or trend fluctuations. Next step is to find initial values for the orders of seasonal and non-seasonal parameters, p, q. These values could be obtained by looking for significant autocorrelation and partial autocorrelation coefficients.

Estimation: After identification stage, autoregressive and moving average parameters are estimated using OLS, Maximum likelihood method, backcasting etc. techniques. One or more models are tentatively chosen for best fit and accuracy of models is tested using diagnostic statistics.

ARMA (Autoregressive moving average) model is defined as follows:

$$Y_t = \phi_1 Y_{t-1} \dots \phi_P Y_{t-P} + e_t - \theta_t e_{t-1} \dots \theta_q e_{t-q}$$

Y_t = Original time series

e_t = series of unknown random errors

ϕ_p = autoregressive parameters

ϕ_1 = Moving average parameter

Box-Jenkins uses backshift operator (B) to show the effect of changing time period from t to t-1. The purpose is to find the best estimate of ϕ_1 to fit the time series that is being modeled.

$$\mathbf{(1-B)\ Y_t = (1-\phi_1\ B)\ e_t}$$

B= Backshift operator

Diagnostic Checking: After estimating parameters of identified ARIMA model diagnostic checking is essential to verify that the model is best fit. In this ACF and PCF of residuals are checked to see if they are white noise.

Following diagnostic tools were applied in this study.

1) Root of Mean Squared Error

$$\text{(RMSE)} = \sqrt{\frac{\sum (A_t - F_t)^2}{n}}$$

2) Mean Absolute Percent Error

$$\text{(MAPE)} = \frac{1}{n}\sum_{t=1}^{n}\left\|\frac{A_t - F_t}{A_t}\right\|$$

F_t =Forecasted value

At = Observed value

n= No. of observation

Forecasting : The forecasts obtained by ARIMA are more reliable than those obtained from the traditional econometric modeling, particularly for short-term (Makridakis *et al.*, 1998).

RESULTS AND DISCUSSION

The monthly wholesale prices of paddy from 5 leading spatially separated states namely Punjab, West Bengal,Uttar Pradesh, Tamil Nadu and Delhi in India have been analysed and forecasted for appropriate decision making. Price behaviour of the paddy across different states as revealed by seasonal index show that higher prices generally prevailed during crop season *i.e.*, July to September and harvest season in some cases. Paddy wholesale prices have also been rising consistently from July 2002 to June 2020 irrespective of markets and production share.

1. Summary statistics

Analysis of collected data reflect that among the selected states, average prices in Delhi have been highest and lowest in West Bengal as exhibited (Fig. 1) in the past 18 years.

Instability in the paddy wholesale prices

The instability in prices as depicted by Cuddy Della Index was highest in Punjab and lowest instability was recorded in West Bengal (10.93) followed by Tamil Nadu (13.81). Other than being a major producer West Bengal also has huge consumption base for rice while in most north Indian states like Punjab, Uttar Pradesh consumption preference is more towards wheat than rice

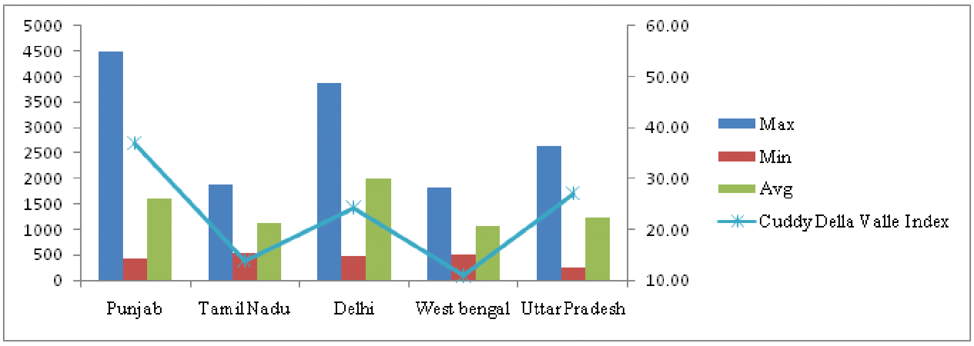


Fig. 1.Monthly price movement in wholesale selected paddy markets (AY 2002-03 to 2019-20)

Source: Author's computation based on data collected from AGMARKNET portal

which could possibly made the variation in northern state.

Seasonal variations: Since most agricultural commodities are seasonal in nature their prices are subjected to seasonal fluctuations therefore continuous monitoring is essential. Rice is mainly grown in three seasons in different parts of India as a result different states have varying harvest and peak crop season. As per Fig. 2, state wise seasonal price variations were highest in Punjab followed by Uttar Pradesh while in other states prices have been more or less stable. The variation in seasonal indices exhibited a wider range. The average seasonal price variation (ASPV) and intra-year price rise (IPR) were found to be highest in Punjab followed by Uttar Pradesh which could be attributed to

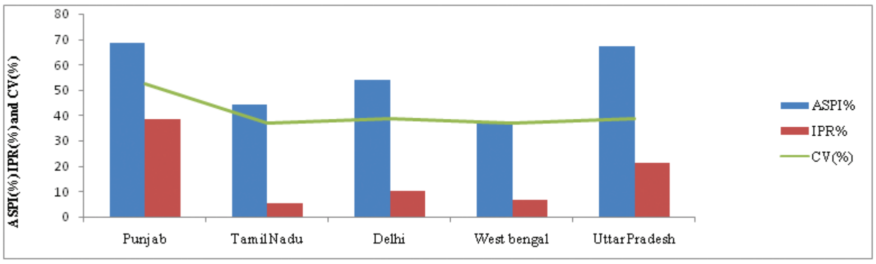


Fig. 2: Seasonal variations in prices

Source: Author's computation based on data collected from AGMARKNET portal

Table 3: Average deseasonalised monthly wholesale paddy prices (Rs Qt^{-1}) in AY 2002-03 to 2019-20

Months	Punjab	Tamil Nadu	Delhi	West Bengal	Uttar Pradesh
July	1704.11	1133.30	2016.54	1065.37	1222.38
August	1691.25	1142.79	2059.11	1062.90	1236.48
September	1601.34	1140.40	2009.16	1062.76	1255.75
October	1562.88	1136.28	1990.94	1054.58	1256.74
November	1559.12	1124.93	1982.33	1059.06	1256.95
December	1526.00	1134.93	1989.27	1057.10	1286.44
January	1683.92	1143.88	1985.13	1058.18	1205.43
February	1634.06	1143.95	1984.23	1060.14	1222.30
March	1547.57	1144.44	2001.49	1063.72	1286.07
April	1489.38	1140.11	2010.04	1061.76	1274.25
May	1544.34	1141.62	1998.03	1060.94	1252.99
June	1639.61	1134.95	1965.96	1068.50	1068.50

AY=Agricultural year

Table 4: ARIMA model fit statistics for the monthly wholesale paddy prices

Market	Fitted model	RMSE	MAPE
Punjab	ARIMA(4,0,12)	481.15	243.80
Tamil Nadu	ARIMA(0,1,6)	90.42	46.71
Delhi	ARIMA(0,1,12)	188.15	51.04
West Bengal	ARIMA(0,1,3)	38.95	27.93
Uttar Pradesh	ARIMA(3,1,12)	268.43	217.46

Table 5: Diagnostic testing (Ljung-Box Q test)

Market	Fitted model	Ljung-Box Q		
		Statistics	DF	Sig
Punjab	ARIMA(4,0,12)	33.45	14	0.002
Tamil Nadu	ARIMA(0,1,6)	37.40	17	0.003
Delhi	ARIMA(0,1,12)	20.16	15	0.171
West Bengal	ARIMA(0,1,3)	21.01	16	0.178
Uttar Pradesh	ARIMA(3,1,12)	24.96	14	0.035

Table 6: Price forecasts of northern states (in Rs. Q^{-1})

Month	Punjab ARIMA(4,0,12)			Uttar Pradesh ARIMA(3,1,12)			Delhi ARIMA(0,1,12)		
	Forecast	Lower Limit	Upper Limit	Forecast	Lower Limit	Upper Limit	Forecast	Lower Limit	Upper Limit
July-20	2097.05	1157.40	3036.71	1682.73	1154.37	2211.08	3080.05	2709.53	3450.64
Aug-20	2239.59	1154.59	3324.59	1495.35	869.17	2121.53	3040.66	2458.90	3622.43
Sep-20	2074.80	945.51	3204.08	1384.54	716.35	2052.73	3023.53	2322.48	3724.59
Oct-20	1810.23	666.57	2953.90	1487.11	806.27	2167.95	3044.75	2241.94	3847.55
Nov-20	1836.85	683.67	2990.02	1526.84	816.78	2236.90	3010.00	2116.96	3903.04
Dec -20	1986.12	806.45	3165.78	1507.94	755.53	2260.34	3019.26	2044.30	3994.22
Jan-21	2171.47	968.08	3374.85	1438.64	641.86	2235.42	3024.87	1974.36	4075.38
Feb-21	2149.95	931.39	3368.50	1364.09	533.61	2194.58	3012.11	1991.14	4133.09
Mar-21	1873.04	644.65	3101.43	1362.45	503.56	2194.58	3057.96	1870.69	4245.23
Apr -21	1812.35	573.01	3051.69	1129.71	242.84	2016.58	3086.12	1836.06	4336.18
May -21	1898.34	647.87	3148.80	1080.90	164.58	1997.45	3123.08	1813.25	4432.91
Jun-21	1981.13	721.24	3241.02	1179.49	2125.68	233.30	3133.36	1766.37	4500.36

Table 7: Price forecasts of southern and eastern states (in Rs. Q^{-1})

Month	West Bengal ARIMA(0,1,3)			Tamil Nadu ARIMA(0,1,6)		
	Forecast	Lower Limit	Upper Limit	Forecast	Lower Limit	Upper Limit
July-20	1835.05	1758.32	1911.78	1674.98	1496.76	1853.20
Aug-20	1841.14	1721.11	1961.16	1662.91	1410.87	1914.95
Sep-20	1848.19	1696.78	1999.59	1670.70	1362.01	1979.38
Oct-20	1854.19	1682.64	2025.75	1674.05	1317.61	2030.49
Nov-20	1860.19	1670.63	2049.77	1668.39	1269.86	2066.90
Dec -20	1866.20	1660.17	2072.23	1674.86	1238.31	2111.41
Jan-21	1872.20	1650.95	2093.45	1674.86	1211.97	2137.74
Feb-21	1878.20	1642.70	2113.71	1674.86	1187.05	2162.66
Mar-21	1884.21	1635.27	2133.15	1674.86	1163.33	2186.37
Apr -21	1890.21	1628.53	2151.89	1674.86	1140.69	2209.02
May -21	1896.21	1622.38	2170.04	1674.86	1118.96	2230.76
Jun-21	1902.22	1616.75	2187.69	1674.86	1098.22	2251.67

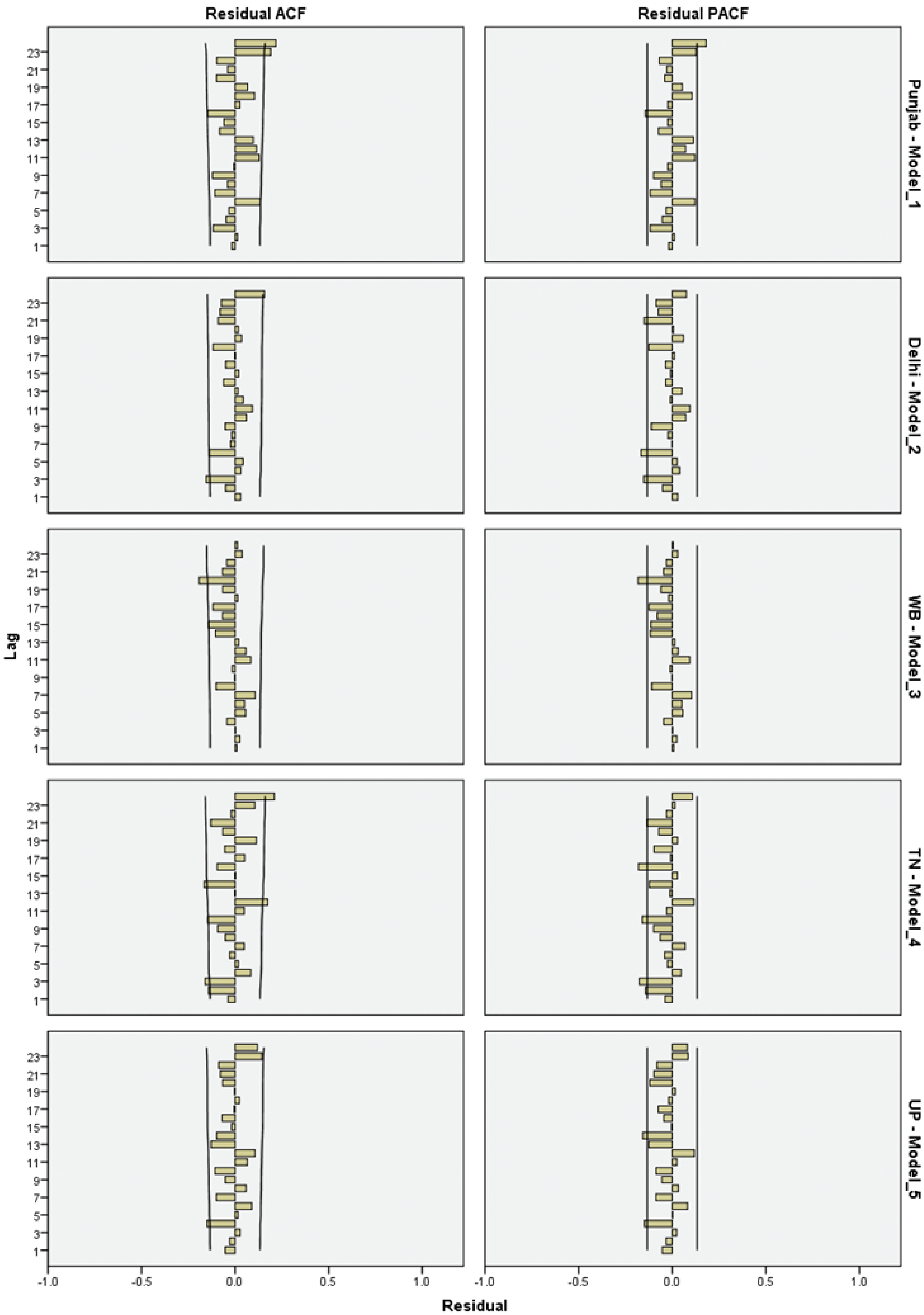


Fig 3. Residual ACF and PCF plot

the fact that these regions being major procurement centers where prevailing MSP has a strong influence on wholesale prices.

2. Wholesale paddy price forecast

1. Checking for stationarity : Examination of ACF and PACF revealed the presence of non-stationarity in the respective data series as depicted.Therefore differencing has been carried out to make the data set stationary. Residuals ACF and PCF plots for respective markets are given in Fig. 3.

2. Residuals were checked based on the diagnostic measures such as Root Mean Square Error (RMSE)and Mean Absolute Percentage Error (MAPE). Best fit ARIMA model were then estimated.ARIMA (4,0,12) ARIMA (0,1,6), ARIMA (0,1,12), ARIMA (0,1,3) and ARIMA (3,1,12) for Punjab, Tamil Nadu, Delhi, West Bengal and Uttar Pradesh respectively were found best fit.

3. Forecasting: After fitting the models, prices have been forecasted for next agricultural year. Results of the paddy price forecast along with the lower and upper limit for five different markets have been shown in Table 5 and 6. Forecast is based on past data of 18 years and the forecasted prices obtained as per the analysis for the agricultural year 2020-21 would be in the range of Rs1810.23– Rs.2239.59 qt^{-1} for Punjab, Rs.1662.91-Rs.1674.98 qt^{-1} for Tamil Nadu, Rs.3010.00-Rs.3133.36 qt^{-1} for Delhi, Rs.1835.05-Rs.1902.22 qt^{-1} for West Bengal and Rs.1080.90-Rs.1495.35 qt^{-1} for Uttar Pradesh. Forecasted wholesale prices are expected to be running highest in Delhi although it is not a producing region but major consuming centre and a national market makes the prices very volatile in the region. For northern states like Punjab and Uttar Pradesh where mainly Kharif rice is grown and harvested in Nov-Dec, prices are found to be on higher side both during crop season as well as lean season. Especially for Punjab prices are expected to be very high for the month of January to February 2021 in the range of Rs.2171qt^{-1} to Rs.2149 qt^{-1}. However in UP, prices are on lower side than those prevailing in corresponding markets of Delhi and UP where the price rise during harvest season October 2020 to January 2021 could be observed. Prices across northern states registered extreme volatility. While in West Bengal prices have been more or less stable fluctuating within a narrow range (Rs.1835.05 qt^{-1}-Rs.1902.02 qt^{-1})which could be attributed mainly to round the year availability and more than one growing season. Similar is the trend in case of Tamil Nadu where rice is grown in two seasons, forecasted prices are reported to be the lowest with minor variations depending on season.

CONCLUSION

The year 2020 being an extraordinary year of pandemic is likely to have far reaching effects on the prices in the upcoming time. Therefore, the study has been undertaken to analyse the seasonal variations in paddy prices as well as forecast the wholesale paddy prices for the agriculture year 2020-21. Analysis of the seasonal variations in wholesale prices reflected through various indices such as ASPI,IPR, Coefficient of variation observed to much more prominent in Northern states like Punjab, Uttar Pradesh, Delhi, these regions being major procurement centers could be the possible reason. Price forecast was made by fitting ARIMA model for five major states using wholesale data of last 18 years which revealed variability in forecasted prices region to region. Results further indicated that prices for the agricultural year 2020-21 would be in the range of Rs1810.23– Rs.2239.59 qt^{-1} for Punjab, Rs.1662.91-Rs.1674.98 qt^{-1} for Tamil Nadu, Rs.3010.00-Rs.3133.36 qt^{-1} for Delhi, Rs.1835.05-Rs.1902.22 qt^{-1} for West Bengal and Rs.1080.90-Rs.1495.35qt^{-1} for Uttar Pradesh. Among all the selected markets, prices have been reported highest for Delhi market and price expectation are likely to be on higher side during harvest season for Northern states the i.e., in month of December 2020 to February 2021. Best fit ARIMA model for different markets were ARIMA(4,0,12) ARIMA(0,1,6), ARIMA(0,1,12), ARIMA(0,1,3) and ARIMA(3,1,12) for Punjab, Tamil Nadu, Delhi, West Bengal and Uttar Pradesh, respectively. Wider price variation has been observed for northern states than Tamil Nadu and West Bengal where forecasted prices are likely to be more stable and varying within a narrow range. . Since crucial decisions with respect to production, marketing and consumption are dependent on certain price expectation, results obtained by the study might prove helpful in guiding the stakeholders involved in making timely decisions.

REFERENCES

AGMARKNET portal, Directorate of Marketing and Inspection,Department of Agriculture and Farmers Welfare, GOI 2020

Abadan, Sarah and Shabri, Ani. 2014.Hybrid Empirical Mode Decomposition-ARIMA for Forecasting Price of Rice, Applied Mathematical Sciences. **8**(63): 3133 - 3143

Box, G.E.P. and Jenkin, G.M. 1976. Time Series of Analysis, Forecasting and Control, Sam Franscico, HoldenDay, California. USA.

Darekar, A. and Reddy, A. A. 2018. Forecasting wheat prices in India. *Wheat and Barley Research*.**10** (1): 54-60.

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

Food and Agriculture Organization, FAOSTAT 2018, Production share of paddy by region (Average 1994 – 2018),FAO,Rome

Jadhav, V., Reddy, C. and Gaddi, G M. 2017. Application of ARIMA model for forecasting agricultural prices. *Journal of Agri. Science & Technology*.**19**: 981-992

Makridakis, S., Wheelwright, S C and Hyndman, R J.1998. Forecasting - methods and applications, New York: John Wiley and Sons, Inc.

MOA&FW, 2018.Agricultural Statistics Division, DES, Dept of Agriculture, Cooperation and Farmer's Welfare. New Delhi.

Ohyver, M. and Pudjihastuti, H. 2018. Arima Model for forecasting the price of medium quality rice to anticipate price fluctuations. *In Procedia Computer Science*, **135**:707–711

Punitha, S.B. 2007. A Comparative analysis of market performance of agricultural commodities - An economic approach. M. Sc. (Ag.) Thesis, University of Agriculture Science, Dharwad, Karnataka, India.

Sendhil, R., Arti, Lal. P., Gururaj, B.M., Jamaludheen, A., Chaudhary, U. and Rathore, R. 2019. Price dynamics and extent of integration in Indian wholesale and retail wheat markets, *Journal of Agricultural Science and Technology* **21**(3):517-530.

Verma, V., Kumar, P., Singh, S. and Singh, H. 2016. Use of ARIMA modeling in forecasting coriander prices for Rajasthan. *International Journal of Seeds and Spices,* **6** (2):40-45.