The Andhra Agric. J 68 (1): 126-132, 2021

Price Forecasting of Ragi in Hassan Market of Karnataka – A Timeseries Approach

Govindaraju, V Srinivasa Rao, D Ramesh and V Sitarambabu

Department of Statistics and Computer Applications, Agricultural College, Bapatla, A. P.

ABSTRACT

The present study is an attempt to forecast the prices of ragi at Hassan market of Karnataka till May 2022 using various sources of secondary data. The ARIMA methodology was employed to forecast and for diagnostics of the model MAPE, MAE, AIC, and R^2 were used. It is found that the model ARIMA(2,0,2)(1,0,1)[12] is appropriate over the other possible models and the prices are estimated as 2327 Rs./qtl in August, 2021, which is the highest during the forecast period. Hence, farmers need to plan the production process in such a way that a good price for the product can be expected.

Keywords: *ARIMA, Forecasting, Price and Ragi,*

Millets are the most important cereals of the semi-arid zones of the world. For millions of people in Africa and Asia, they are staple food crops. Among millet crops, finger millet figures prominently; it ranks fourth in importance after sorghum, pearl millet and foxtail millet. In India, it is cultivated over an area of 1.30 million hectares with a total production of about 2.04 million tonnes. It is an important crop of Karnataka, with more than 60% area of the country followed by Uttarakhand (10%) and Maharashtra (9.6%). Forecasting techniques are used to assist in decision-making and planning the future more effectively and economically. The time series approach to forecasting is one such approach that relies on past patterns. A data set containing observations on a single phenomenon observed over multiple periods is called a time series. In time series data, both the values and ordering of the data points have meaning. For many agricultural commodities, data is usually collected over time. Analysis of time series has been a part of statistics for a long. Various statistical approaches are available for time series analysis, modeling, and forecasting. One of the most important and widely used time series models is the Auto-Regressive Integrated Moving Average (ARIMA). The popularity of the ARIMA model is due to its statistical properties as well as the use of the well-known Box-Jenkins's methodology in the modeling process. The present study was undertaken with the objective of analyze, model and forecast the prices of ragi in the Hassan markets of Karnataka.

MATERIAL AND METHODS

The Hassan market was purposefully chosen for the current study based on the continuous market arrivals of ragi throughout the year and the availability of data. Secondary monthly time series data on modal prices of ragi was collected from the Krishi Marata Vahini portal for a period of 19 years from January-2002 to May-2021. The analysis was conducted using R software version 4.0.4 with the available packages of outliers, tseries, aTSA and forecast model and forecast ragi prices over a 12-months from June-2021 to May-2022. Before analysis, as the study was dealing with time series, the present data set had been verified initially for the existence of outliers.

Test for Outlier

For detecting the outlier in the time series, Grubb's test was used in the current scenario as the test is particularly useful in case of a large sample and is easy to follow. Using R software and the function *grubbs.test()*, outliers are attempted to be identified and, if found, replaced by the median of the respective series (Sahu 2010).

Trend analysis

The long-term variation in the series is known as a trend. The standard moving average method eliminates the seasonal and irregular components. Using R, we employ *decompose()* function to obtain the moving average trend value as *decompose(time.series.object)$trend*. These moving average trend values are refined as a linear line by fitting a simple regression equation.

Seasonal component

When a series is influenced by seasonal factors and recurs on a regular, periodic basis, it is known as seasonality. The method of simple averages is used to find the monthly seasonal index. Using R, we apply *decompose()* function to obtain the seasonal component value as *decompose (time.series.object) $seasonal*.

The ARIMA Model

In an Auto-Regressive Integrated Moving Average (ARIMA) model, the time series variable is assumed to be a linear function of previous actual values and random shocks. In general, an ARIMA model is characterized by the notation ARIMA (P,D,Q), where p, d and q denote orders of Auto-Regression (AR), Integration (differencing) and Moving Average (MA), respectively. ARIMA is a parsimonious approach which can represent both stationary and non-stationary processes.

An ARMA (p, q) process is defined by:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ... \theta_q \varepsilon_{t-q}$$

where, y_t and ε_t are the actual value and random error at time period t, respectively, ϕ_i (i=1, 2…,p) and θ_j (j=1, 2…,q) are model parameters. The random errors, are assumed to be independently and identically distributed with a mean of zero and a constant variance of σ^2. Seasonal ARIMA model is denoted by (p, d, q) (P, D, Q), where p denotes the order of autoregressive process, q: order of moving average process and d: number of times to be a series must be differenced to induce stationary. P: number of seasonal autoregressive components, Q: number of seasonal moving average terms and D denotes the number of seasonal differences required to induce stationary.

The first step in the process of ARIMA modelling is to check for the stationarity of the series as the estimation procedure is available only for a stationary series. A series is regarded as stationary if its statistical characteristics, such as the mean and the autocorrelation structures, are constant over time. Stochastic trend of the series is removed by differencing. Multiple ARMA models are chosen on the basis of an Auto-Correlation Function (ACF) and a Partial Auto-Correlation Function (PACF) that closely fits the data. Then, the parameters of the tentative models are estimated through any non-linear optimization procedure such that the overall measure of errors is minimized or the likelihood function is maximized. Lastly, diagnostic checking for model adequacy is performed for all the estimated models through the plot of residual ACF and using Ljung-box test. The most suitable ARIMA model is selected using the highest coefficient of determination (R^2) value and the lowest root mean square error (RMSE), mean absolute percentage error (MAPE) and Akaike Information Criterion (AIC) value. In this study, all estimations and forecasting of ARIMA model were performed using R software.

Forecast Evaluation Methods

The forecasting ability of different models is assessed with respect to common performance measures, *viz*. the root mean squared error (RMSE), mean average percentage error (MAPE), and percentage of forecast error.

- Root mean squared error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}(\hat{y}_t - y_t)^2}{T}}$$

- Mean absolute error (MAE)

$$MAE = \frac{\sum_{t=1}^{n}|y_t - \hat{y}_t|}{N}$$

- Mean absolute percentage error

$$MAPE = \left[\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100\right]/n$$

- Akaike Information Criterion (AIC)

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

- Deviations of predicted prices from the actual prices

$$forecast\ error(\%) = \frac{y_t - \hat{y}_t}{y_t} \times 100$$

Where, y_t= actual prices, $\hat{y}_t$= predicted prices, ε_t=forecast error, T= sample size, $\hat{L}$= maximum value of the likelihood function for the model, k= number of estimated parameters, N = Number of observations

RESULTS AND DISCUSSION

At first, ragi price series of Hassan market from January-2002 to May-2021 was tested for outliers by Grubb's method. It was observed that the number of extreme observations in the present data was zero, as confirmed by results in Table 1.

Table 1. Grubbs test for detecting Outliers

Statistic	Value
Grubbs test statistic	2.317
p-value	1
Outliers detected	0

The test statistic is G = 2.317 and the corresponding p-value is 1. Since this value is not less than 0.05, we fail to reject the null hypothesis. We do not have sufficient evidence to say that the minimum value of series is an outlier. This is also confirmed by Box- whisker plot as shown in Figure 1.

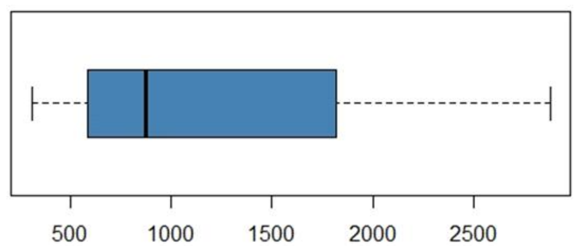

Figure 1. Box-plot of ragi prices in Hassan market

The ragi price series was then subjected to trend analysis. It was showed that ragi prices in the Hassan market have been rising over time, as represented by the linear trend line (red colour) in Figure 2. The maximum price was recorded to be 2878 Rs./qtl in the month of March 2017 and the minimum price was 312 Rs./qtl in the month of March 2002.

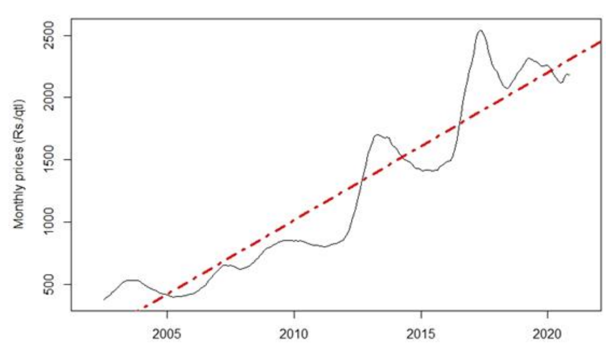

Figure 2. Trend of ragi prices in Hassan market

On the other hand, Seasonal indices prepared to know about the presence of seasonality in monthly ragi prices, as shown in Figure 3. The highest seasonal index was observed in the month of August and the lowest was in the month of June.

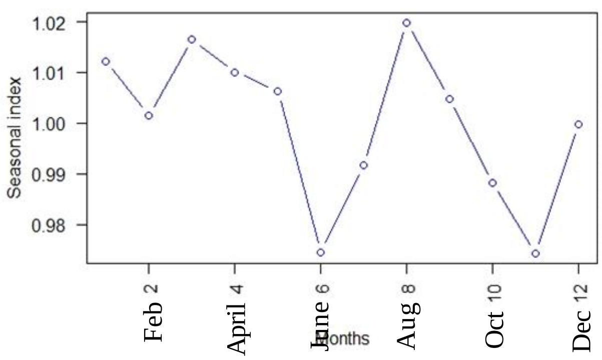

Figure 3. Seasonal indices of ragi prices in Hassan market

Table 2. Augmented Dickey Fuller test for prices of ragi at Hassan market

with drift and trend				
Lags	at levels		at first difference	
	ADF	p-value (0.05)	ADF	p-value (0.05)
0	-3.28	0.075	-14.59	0.01
1	-3.45	0.047	-11.1	0.01
2	-3.27	0.076	-9.08	0.01
3	-3.27	0.076	-8.02	0.01
4	-3.21	0.087	-7.54	0.01

Later, the data was subjected to the ARIMA methodology. The evidence of trend and seasonality, which clearly shows the non-stationarity of price series, is further confirmed by the ADF test and the results have been presented in Table 2. As a result, we used

Table 3. The tentative models for ragi prices in Hassan market

ARIMA MODELS	Model selection criterion				
	R^2	AIC	RMSE	MAE	MAPE
ARIMA(1, 1, 1)(1, 0, 1)$_{[12]}$	0.976	2852.14	111.366	74.99	5.993
ARIMA(1, 1, 0)(1, 0, 0)$_{[12]}$	0.976	2850.53	111.46	75.027	5.994
ARIMA(2, 1, 0)(1, 0, 0)$_{[12]}$	0.976	2851.64	111.244	74.967	5.991
ARIMA(0, 1, 1)(0, 0, 1)$_{[12]}$	0.976	2850.49	111.451	75.029	5.996
ARIMA(2, 1, 2)(1, 0, 1)$_{[12]}$	0.977	2850.46	108.802	73.492	5.953
ARIMA(1, 1, 0)(1, 0, 0)$_{[12]}$	0.976	2851.32	111.152	74.614	5.981

Table 4. ARIMA(2, 1, 2)(1, 0, 1)$_{[12]}$ model parameter table

Model	Coefficients		Estimate	SE
ARIMA(2, 1, 2)(1, 0, 1)$_{12}$	AR	Lag 1	0.37	0.756
		Lag 2	0.202	0.801
	MA	Lag 1	-0.348	0.739
		Lag 2	-0.293	0.786
	AR, Seasonal	Lag 1	-0.87	0.102
	MA, Seasonal	Lag 1	0.762	0.129
	Estimated σ^2		11889	
	Log likelihood		-1418.33	

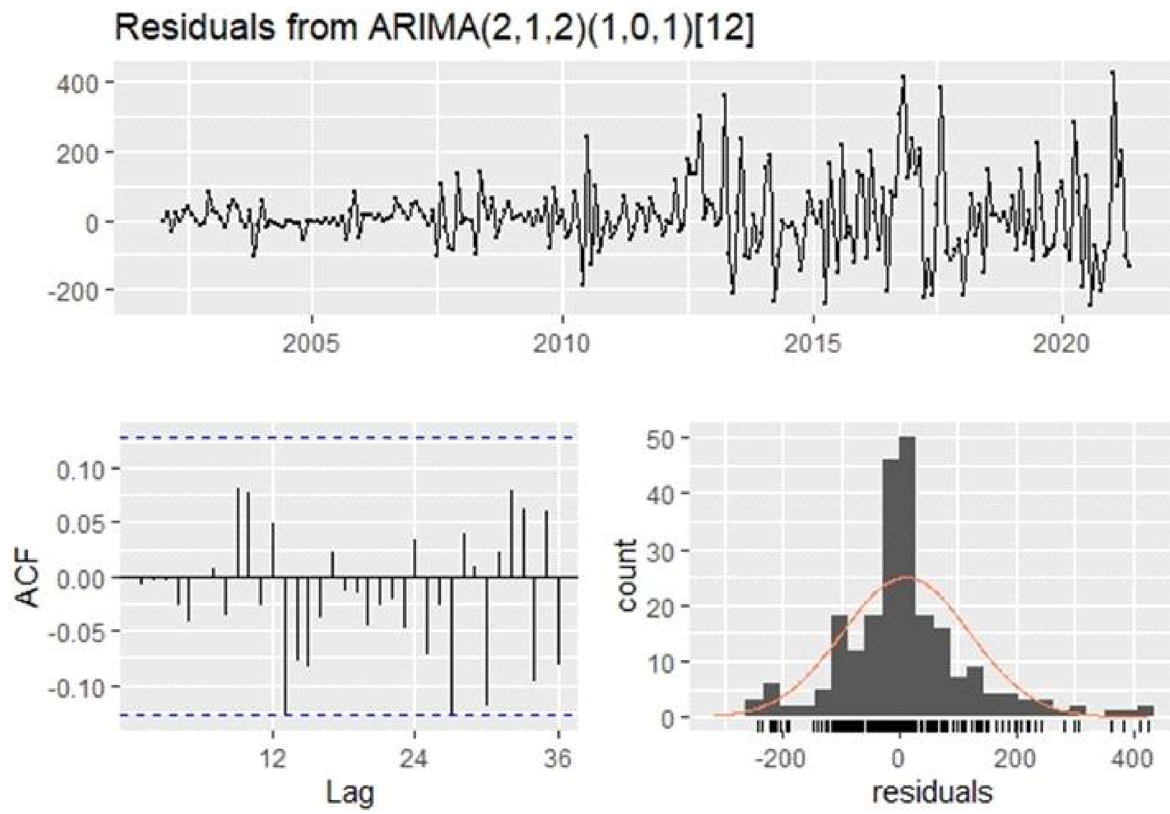

Figure 4. Residual plot of ARIMA(2,1,2)(1,0,1)$_{[12]}$ model

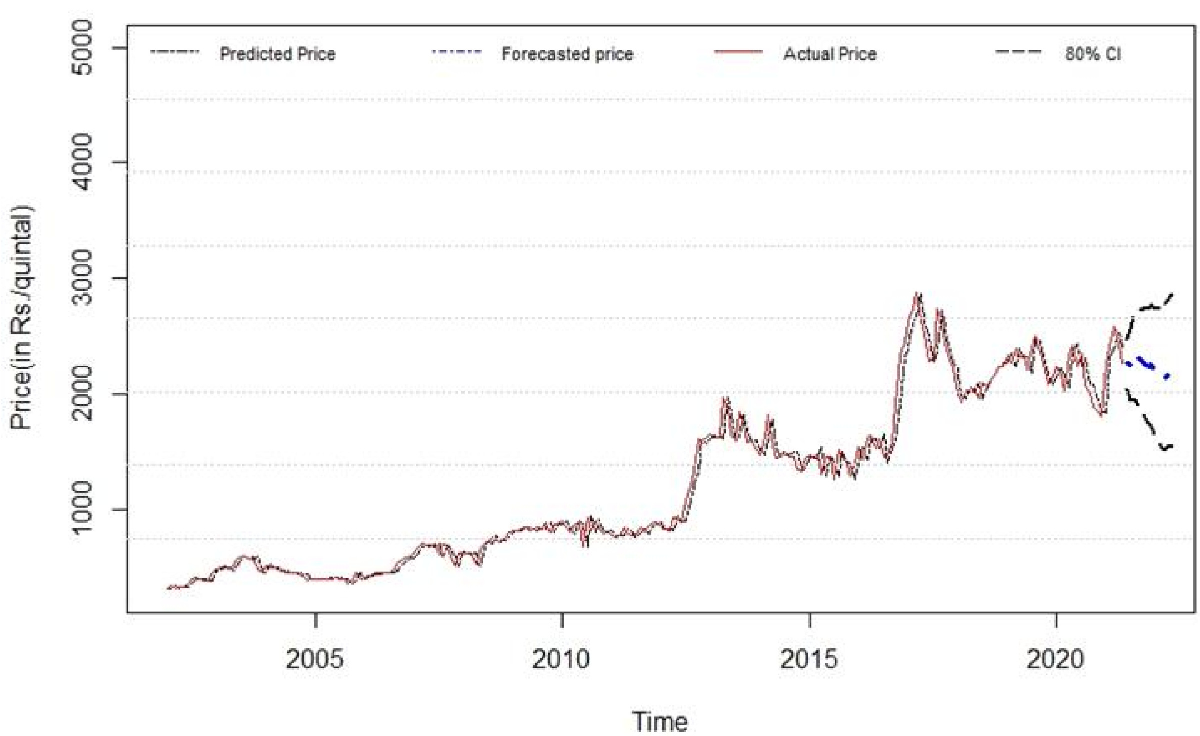

Figure 5. Actual vs fitted graph of ARIMA for ragi prices in Hassan market

Table 5. Forecasting of ragi prices in Hassan market

Year	Month	Actual prices (Rs./qtl)	Predicted prices (Rs./qtl)	Forecast error (%)
2021	January	2263	1836	18.86
2021	February	2400	2301	4.11
2021	March	2590	2392	7.66
2021	April	2437	2539	-4.18
2021	May	2266	2399	-5.89
2021	June	*	2259	*
2021	July	*	2254	*
2021	August	*	2327	*
2021	September	*	2303	*
2021	October	*	2276	*
2021	November	*	2243	*
2021	December	*	2246	*
2022	January	*	2181	*
2022	February	*	2157	*
2022	March	*	*2125*	*
2022	April	*	2174	*
2022	May	*	2200	*

the first differences for the price series. The first differenced series were found to be stationary as indicated in Table 2 and, hence, further differences were not tried.

Now, by fixing d=1, different ARIMA models are tried and the better six models are represented in Table-3. Based on the lowest values of diagnostic criterion (RMSE, MAPE and MAE) ARIMA(2,1,2)(1,0,1)$_{[12]}$ was chosen for prices and the parameters of the model were estimated and presented in Table 4. Later, independence of residuals was confirmed by Ljung-box test, as p-value of the test is 0.854 (> 0.05). The same also further revealed the residual ACF chart as per Figure 4.

Finally, forecasting was done for ragi prices of Hassan market from June-2021 till May-2022 by using ARIMA(2,1,2)(1,0,1)$_{[12]}$ while keeping the first five months of 2021 data for validation. The predicted values obtained through the ARIMA model were compared to actual prices, and forecast errors were calculated, as shown in Table 5 along with the forecasted prices. It was observed that the actual and predicted values were closely related and the predicted values were within control limits as captured in Figure 5. As per model forecasts, ragi prices have been studly decreasing . The highest price, 2327 Rs./qtl, was forecasted in August-2021 and the lowest was in the month of March, 2022 with 2125 Rs./qtl.

CONCLUSION

In this study, prices of ragi in the Hassan market were analyzed using ARIMA and also forecasted up to May-2022 by using the best-fitted ARIMA model ARIMA(2,0,2)(1,0,1)[12]. Based on the forecasted values, it can be concluded that ragi prices will continue to fall in the coming months, and it has been estimated at 2327 Rs./qtl, which was the highest forecasted for August-2021. The studies on forecasting market prices would help the farming community and policymakers to make appropriate decisions. These predicted values may help to maintain the pace of arrivals and prices in the markets.

LITERATURE CITED

A Ian McLeod, Hao Yu and Esam Mahdi 2012 Time Series Analysis with R. Time Series Analysis: Methods and Applications, 30: 661-712.

Akhilesh K G 2018 Forecasting of Arrivals and Prices of Pulses in Chhattisgarh- A Statistical Approach. M Sc Thesis, Indira Gandhi Krishi Vishwavidyalaya, Raipur, C.G.

Box G E P and Jenkins G M 1976 Time series analysis : Forecasting and control. Holden Day, San Francisco.

Jadhav V, Reddy B V C and Gaddi G M 2017 Application of ARIMA Model for Forecasting Agricultural Prices. Journal Agricultural Science and Technology, 19: 981-992.

Nireesha V 2014 Time Series Analysis of Area, Production and Productivity of Major Coarse Cereals in Andhra Pradesh. M Sc Thesis, Acharya N. G. Ranga Agricultural University, Guntur, A.P.

Ramesh D, Rao V S, D V S Rao and G R Reddy 2013 Forecasting of Guntur District Rainfall by using ARIMA models. The Andhra Agricultural Journal, 60 (2): 405-408.

Ramesh D, Pal S, Rao V S and Bhattacharyya B 2015 Time Series Modeling for Trend Analysis and Forecasting Wheat Production of India. International Journal of Agriculture, Environment and Biotechnology, 8 (2): 303-308.

Robert H, Shumway and David S 2019 *Time series: a data analysis approach using R*, Chapman & Hall/CRC Press, 259 pp.

Sahu CR, Lakhera ML and Naidu GM 2019 Forecasting Arrivals and Prices of Paddy in Kurud Market of Chhattisgarh State using SARIMA Approach. The Andhra Agricultural Journal, 66 (spl): 68-70.

Sahu PK 2010 Forecasting production of major food crops in four major SAARC countries. International Journal of Agricultural and Statistics Science, 10: 71-92.

Venkataviswateja B, Rao V S, Umar S N and C S Reddy 2018 A study on arrivals and prices of red Chillies in Guntur market yard– A Time Series approach. International Journal of Current Research in Multidisciplinary, 3 (5): 06-10.

Vinayak N J, Chourad R, Ahmad N D G, Shreya A and Sarfraz S 2014 Forecasting the prices of onion in Belgaum market of Northern Karnataka using ARIMA technique. Internat. Res. J. Agric. Eco. & Stat., 5 (2) : 153-159.

Yogisha G M, Vijaya Kumar H S and Basavaraja H 2007 Trends and seasonal variation in arrivals and prices of potato in Kolar district. Indian Journal of Agricultural Marketing, 69 (4): 26-28.

Received on 24.12.2020 and revised on 09.03.2021