Original Research Paper

Volume -10 | Issue - 3 | March - 2020 | PRINT ISSN No. 2249 - 555X | DOI : 10.36106/ijar

Statistics

FORECASTING OF EGG PRICES IN TELANGANA USING R-SOFTWARE

M. Udayshankar	Department Of Statistics, Osmania University, Hyderabad
Dr. M. Raghavender Sharma*	Department Of Statistics, Osmania University, Hyderabad *Corresponding Author

ABSTRACT This paper presents the Egg is a common man's nutritious food and widely used in different food preparations and also in governments programs like mid-day meals and ICDS. The prices of egg affect not only a common man but also the above programs. An attempt is made in this paper to develop a forecasting model for prediction of monthly average prices of egg. Analysis of seasonal variations in the prices and forecasting using SARIMA model was discussed. The data reveals the prices raise in November and relatively down in April of the year. A SARIMA(1,0,0)(2,1,0)12 model is selected as a an appropriate and adequate models for forecasting of egg prices in Telangana state.

KEYWORDS : Time Series, Sarima, Mape, Egg, Necc.

1. INTRODUCTION

Egg is a common man food and it is a main source of quality protein at an affordable price. In One Egg (boiled) contains 84 calories, 8.3g protein, 5.7g fat, and 1.6g sat fat. A study published in paediatrics magazine has suggested that for 6 months in a family young children's eats egg along with their diet with reduced sugar-sweetened foods, they may achieve a healthy height (Ref). Eating one egg (Per a day) may reduce the risk of many diseases like heart disease and malnutrition problems. The government has declared national egg day as June 3rd. In this paper, monthly data on average prices of 100 eggs during January 2009 to November 2019 collected from National Egg co-ordination committee (NECC.org) was used to model and forecast the future prices.

In the recent days, the egg eaters and usage of egg in industries increasing at face, the forecasts of the egg prices are beneficial. The objective of this paper is to analyze the seasonal price variations of egg and to develop a seasonal autoregressive integrated moving average (SARIMA) and attempt is made to forecast the average price of egg in Hyderabad in Telangana state.

2. MATERIAL AND METHODOLOGY

A historical data of monthly egg prices in 100 units during January 2009 - November 2019(containing 131 observations). The times series was analysed using Box-Jenkins methodology. This methodology is an algorithm to select an adequate model by fitting a SARIMA model to historical time series. This method having several advantages, mainly to obtain the minimum numbers of parameters (AR and MA and Differentiation) and handling with stationary and non-stationary time series in seasonal as well as non-seasonal components (Permanasari et al., 2009). This procedure includes in the four steps. They are Identification, Estimation, Diagnostic testing and Forecasting (Sivapathasundaram and Bogahawatte, 2012). By using auto correlation and partial auto correlation functions the stationary and non-stationary was identified. Once, the model is finalized, the parameters to be estimated. Diagnostic check is to test the model adequacy and this is primarily to check the assumption about the errors are random. The final selected model can be used for forecast the future prices. The model performance can be derived from the error measures like mean absolute percentage error (MAPE), mean absolute error (MAE) and root mean square error (RMSE). The model is implemented using R software and MS-Excel is for preparation of charts and tables.

Several models were tested for the given data and identified one model for estimating the accurate parameters and they are statistically adequate presentation of the accessible data. These estimators are obtained by using minimum mean square error (MMSE) method. After determine the auto-regressive and moving averages and stationary Represent p, q and d the seasonal ARIMA would be SARIMA (p,d,q)(P,D,Q)s.

$$Y_t = \nabla^d X_t \quad \text{------------------------------- (1)}$$

Where $\nabla = (1-B)$ and

d is the differentiation to achieve stationarity in non-seasonal com p onents

D is the differentiation to achieve the stationarity in seasonal components

then SARIMA(p,d,q)(P,D,Q) is or ARIMA(p,d,q)(P,D,Q)s

$$\Phi_P(B^s)\,\Phi_p(B)(\nabla_s)^D(\nabla)^d Y_t = \Theta_Q(B^s)\,\Theta_q(B)\,X_t \quad \text{--------------------------------(2)}$$

In the above equation (2), non-seasonal components (AR and MA) and seasonal components (SAR and SMA)

3. LITERATURE REVIEW:

CHENGAPPA (1980):

Applied the Box-Jenkins model to forecast pool sale and export auction prices of coffee. Monthly data used and due to the distinct seasonal variation in prices, the ARIMA seasonal model was applied. The pool sale prices forecasts were found to be accurate when compared to forecasts of export prices. This attributed to possible lack of stationary to the data. Hence, the adoption of differencing procedure or transformations to make the data stationary found necessary for a better estimate of export prices.

SRIVASTAVA AND BRAHMAPRAKASH (1994):

Used ARIMA models for forecasting sugarcane productivity based on the time series data of fifty years (1940-41 to 1989-90) in Bihar. Their findings, therefore ascertained that time series data on sugarcane productivity for the state of Bihar was described by an ARIMA (0, 1, 1) model

MASTNY (2001):

Used ARIMA models, also called Box- Jenkins model for the analysis of time series for agricultural commodities. The study contains a basic mathematical explanation of ARIMA models together with the practical illustration of prices development forecast for a selected agricultural commodity.

MOMBENIETAL., (2013):

Employed the seasonal autoregressive integrated moving average (SARIMA) model for residential water demands. Most residential water consumption series are seasonal and non- stationary. The model was fit to monthly residential water consumption in Iran from May 2001 to March 2010. The results found that three-parameter log-logistic distribution fits the model residuals adequately. Therefore the forecast values for 1 year ahead using the fitted SARIMA model

VINAYAK N.JALIKATTI AND B.L.PATIL 2013:

Proposed Onion price forecasting in Hubli market of Northern Karnataka using ARIMA technique

V. N., RAGHVENDRA CHOURAD., MAHANTESH GUM G OLMATH., SAFARAZ SHAEIKH AND SHREYA AMAR AP URKAR., 2014:

Price Forecasting of Onion in Bijapur Market of Northern Karnataka using ARIMA Technique. International Journal of Commerce and Business Management.

DR.SANGITA VISHNU WARADE., 2016:

Forecasting Of Onion Prices in Maharashtra: An Approach to Support

Volume -10 | Issue - 3 | March - 2020 | PRINT ISSN No. 2249 - 555X | DOI : 10.36106/ijar

Vector Regression and ARIMA Model, International Journal of Social Science and Humanities Research.

4. RESULTS AND DISCUSSION

The first step is to identify the stationary and non-stationary by plotting sequence charts and autocorrelation function plots.

Chart 1: Egg prices in Telangana state and ACF and PACF plots

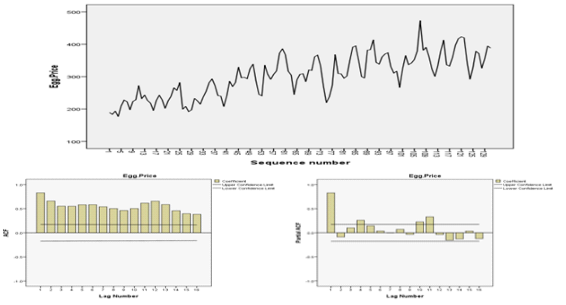


Figure1. Egg average prices in Telangana State and ACF and PACF plots.

The table (1) is represented as data for years (2009-2017) average prices. The monthly egg prices compiled and collected from the National Egg co-ordination committee (NECC.org). The data consists of years 2009-2019 from monthly average egg prices in Telanaga state, India. To identify and fit the model in 2009 – 2017 years are used and 23 months are used for forecasting and comparing predicted egg prices.

Table 1. Monthly average prices in Telangana State

Months\Years	2009-2011	2012-2014	2015-2017
January	238.24	321.98	353.07
February	203.06	298.91	310.36
March	206.73	276.92	278.83
April	188.05	232.39	266.46
May	211.08	257.23	325.94
June	234.41	309.47	372.22
July	225.39	295.10	353.32
August	204.85	286.00	331.38
September	228.49	319.09	328.84
October	240.89	311.20	347.23
November	273.18	343.91	397.65
December	260.72	348.75	382.14

From the table 1, it is clear that, more than 62% of the average prices are allocated in three months and the prices increasing almost (48-62) %. The maximum prices are observed during November, December and June like (397.65, 382.14 and 372.2). A significant average price increases are in November, whereas a very minimum average price in the month is observed from April. The prices are increased in January, June and November in the state for the past nine years with monthly averages, however the price are increased for the last 5 years. Average prices are high during November-January and also in June, whereas the same prices are relatively low during March-April in a year (Figure 2).

Chart 2: seasonal indices for monthly average prices

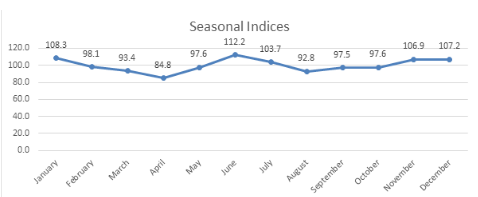


Figure2. Seasonal indices for monthly average prices in Telangana State

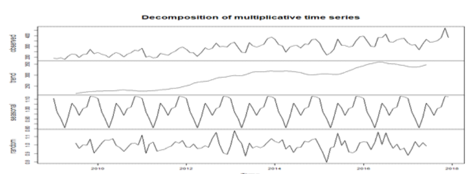


Figure3. Decomposition of multiplicative time series for Monthly average prices in Telangana State

From figure 3, it was observed that the data plot from January 2009 to November 2019 having the increasing order of components of time series. We need to make them as consistent mean and variance. In this some of the Analysis for variation in variance level and tested the different models and identified the best model. The SARIMA model for the data having number of possible models are occurred. In R programming provides an automatic algorithm, Running auto.arima() default values used and parapets for AR and MA will be associated along with the best model.

The best model is SARIMA (1,0,0)(2,1,0)12 and the following are the parameters of the model.

Table 2. SARIMA (1,0,0)(2,1,0)$_{12}$ parameters are

Ar1	SAR1	SAR2	drift
0.5241	-0.737	-0.345	1.5347

The above parameters are substitute in the SARIMA (1,0,0)(2,1,0)$_{12}$, then

$[(1-\varphi B)(1-\Phi_1 B^{12}-\Phi_2 B^{24})(1-B^{12})]Y_t = \theta a_t$

$Y_t = 0.5241Y_{t-1} + 0.2634Y_{t-12} - 0.1380Y_{t-13} + 0.3867\ Y_{t-24} - 0.2026Y_{t-25} - 0.3499Y_{t-36} - 0.18338Y_{t-37} + 1.5347$ ....(*)

The SARIMA (1,0,0)(2,1,0)12 model is where denote the average price and t a represents the residual at time t. For testing the model adequacy, we use the Box-Ljung test and it is adequate model. The table 4 shows that, the all possible model and SARIMA (1,0,0)(2,1,0)12 occurs minimum mean absolute percent error (MAPE) as 5.82, minimum mean absolute error is 17.68 and with a low normalized BIC value of 934.85.

Table 4: We are fitting the several models using data from Jan 2009 to Dec 2017. The results are summarized in the following table.

Models	AIC	BIC	RMSE	MAE	MAPE	P-value	t-test
SARIMA(1,0,0)(1,1,0)$_{12}$	937.06	944.75	28.45	20.55	6.70	0.15	26.40
SARIMA(1,0,0)(0,1,0)$_{12}$	958.18	963.31	32.78	24.26	7.86	0.04	11.55
SARIMA(1,0,0)(2,1,0)$_{12}$	922.03	934.85	25.34	17.68	5.82	0.35	21.85
SARIMA(1,0,1)(2,1,0)$_{12}$	932.29	945.12	26.86	19.26	6.30	0.02	12.95
SARIMA(0,0,1)(2,1,0)$_{12}$	964.75	975.01	32.86	23.89	7.80	0.01	15.86
SARIMA(1,0,0)(2,1,0)$_{12}$	935.35	945.61	27.58	19.82	6.47	0.07	30.24

The Ljung-Box test is used to test the lack of fit in time series models and it is used to test for the residuals of time series after fitting model and also it is applied to test for residuals for auto correlation functions. Below table gives the complete information about the residuals about our model. Compare to all possible models based on the below criteria that means which is having low residuals, we conclude that SARIMA (1,0,0)(2,1,0)12 is the best model for forecast the average egg prices.

The hypothesis for the model is the p- value is more than 0.05 now we can conclude that the model is fairly accurate. The following figure 3 has training data having January 2009 to December 2017 and table shows the forecasts of monthly average prices in Telangana state.

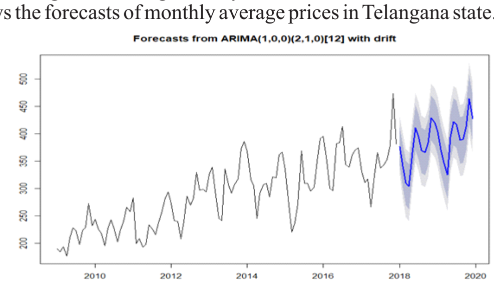


Figure3. Training data set years (2009-2017) Monthly average prices in Telangana State

Table 5: Ljung-Box test

Model	SARIMA(1,0,0)(2,1,0)12
AIC	922

BIC	934.9
RMSE	25.3434
MAE	17.6795
MAPE	5.82472
P-value	0.349
t-test	21.85
df	20

Table 6: train and test data error comparison table

	MAE	RMSE	MAPE
train data	17.6795	25.3434	5.8247
test data	17.6795	25.3434	5.8247

The train and test data the RMS and MAPE are same. Seasonal ARIMA(1,0,0)(2,1,0)12 is the better model for forecast the future values. The test data set years January 2018 to November 209 gives the complete idea about the pattern for the next two years (Figure 5)

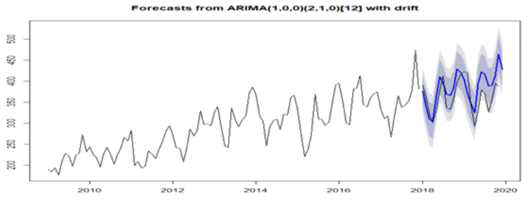
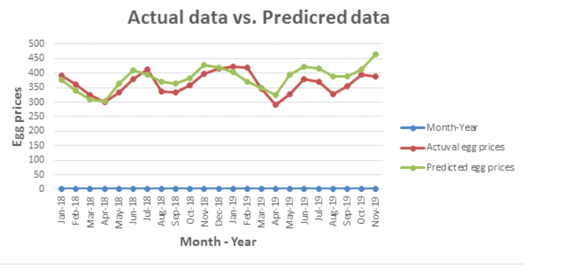


Figure4. Test data set (2018-2019 November) Monthly average prices in Telangana State

Forecast the predicted egg price and actual test data to the comparison for next two years with 95% (LCL and UCL) range values are present in the below table 7 and actual prices and predicted prices are shown in the figure 5.

Table 7: Forecast of egg prices in Telangana state

Year-Month	Actual egg Prices	Predicted egg prices	LCL	UCL
2018-Jan	390.61	376.69	322.87	430.51
2018-Feb	361.57	340.78	280.02	401.55
2018-Mar	323.71	309.96	247.43	372.50
2018-Apr	300.57	303.48	240.47	366.49
2018-May	333.00	365.64	302.49	428.78
2018-Jun	378.93	410.91	347.73	474.09
2018-Jul	412.68	395.41	332.22	458.60
2018-Aug	336.45	369.22	306.03	432.42
2018-Sep	332.57	365.57	302.38	428.77
2018-Oct	359.10	383.56	320.36	446.75
2018-Nov	396.60	428.51	365.31	491.70
2018-Dec	417.23	420.15	356.95	483.34
2019-Jan	422.84	403.14	338.37	467.91
2019-Feb	420.00	370.70	305.50	435.90
2019-Mar	344.71	347.55	282.24	412.86
2019-Apr	291.67	324.57	259.23	389.92
2019-May	328.10	393.34	327.99	458.70
2019-Jun	378.43	421.79	356.44	487.15
2019-Jul	371.52	417.11	351.75	482.47
2019-Aug	326.13	387.95	322.59	453.31
2019-Sep	354.83	389.58	324.22	454.93
2019-Oct	394.71	412.30	346.95	477.66
2019-Nov	388.50	464.10	398.75	529.46
2019-Dec		427.23	361.87	492.58
2020-Jan		405.98	336.07	475.89
2020-Feb		376.74	305.63	447.85
2020-Mar		360.57	289.13	432.01
2020-Apr		334.60	263.08	406.13
2020-May		396.97	325.42	468.52
2020-Jun		436.34	364.78	507.90
2020-Jul		419.32	347.76	490.88
2020-Aug		403.04	331.48	474.60
2020-Sep		405.82	334.26	477.38
2020-Oct		427.90	356.34	499.46
2020-Nov		491.61	420.05	563.17
2020-Dec		446.90	375.34	518.46



Figuare5: The comparison of Actual prices versus predicted prices

4. CONCLUSION:

In the present study, the fitted model is SARIMA (1,0,0)(2,1,0)12 and it was verified with the test data. The test data contains nearly twenty months of data (January 2018 - November 2019).The comparison of the actual prices and predicted prices was given in figure 5. The Predicted pattern is very close to actual pattern for the test data. It was observed that, the forecast using seasonal ARIMA (1,0,0)(2,1,0)12 are about with 6% of error variation. Hence, SARIMA (1,0,0)(2,1,0)12 is a better option for forecasting egg prices as RMSE and MAPE are very minimum. The short-term forecast shows relatively lower variation around the predicted values as compared to the long term forecasts. As the MAPE was 6%, the policy makers and the people related to egg industry may consider this variation for better planning of programs and the above result would show a light on the cost effective and benefit program for the public and government.

5. R OUTPUT:

ARIMA(1,0,0)(2,1,0)[12] with drift

COEFFICIENTS:

ar1 sar1 sar2 drift
0.5241 -0.7366 -0.3449 1.5347
s.e. 0.0889 0.1075 0.1123 0.2503
sigma^2 estimated as 754: log likelihood=-456.01
AIC=922.03 AICc=922.69 BIC=934.85

ERROR MEASURES:

ME RMSE MAE MPE MAPE MASE
-0.3013183 25.34338 17.67945 -0.6162822 5.824721
0.5397259
-ACF1
-0.02038747

Box-Ljung test
data: fit$residuals
X-squared = 21.848, df = 20, p-value = 0.3488

REFERENCES:

1. Adanacioglu,H.And Yercan, M¬¬., 2012,An Analysis of Tomato Prices at Wholesale level in Turkey: An Application of SARIMA model. Custose agronegocio, 8 (4): 52-75.
2. Anonymous. 2012a. Report on Fresh Fruits and Vegetables, Republic of Turkey, Ministry of Economy, 8p. [http://www.tcp.gov.tr/english/sectors/index.cfm]
3. Anonymous. 2012c. Seasonal ARIMA models, Online Course on Applied Time Series Analysis, Lesson4, The Pennsylvania State University [https://onlinecou rse s.s ci ence.psu.edu/stat510/?q=node/67]
4. Anonymous. 2011b. Turkish farmers from the EU of double standards, Antalya Fresh Fruit and Vegetable Exporters' Association-Chairman Mustafa Satıcı, [http://haberpan.com/news/turkish-farmers-from-the-eu-of-double-standards]
5. Albright, S.C., Winston, W., and Zappe, C. 2011. Data Analysis and Decision Making, 4th Edition, Time Series Analysis and Forecasting, ISBN-10: 0538476125 ISBN-13: 9780538476126, 669-743.
6. Bhar, L.M. and Sharma, V.K. 2007. Time Series Analysis, an Electronic-book on Statistical Methods for Agricultural Research, ed. by Jaggi, S., Varghese, C.,
7. Batra, P.K., and Sharma, V.K., New Delphi. [http://w ww.iasr i.res.in /ebook /EB_SMA R/eb ook_ pdf%20files/Manual%20II/8-Time%20series%20analysis.pdf].
8. Sivapathasundaram, V. and Bogahawatte, C.(2012), Forecasting of Paddy Production in Sri Lanka: A Time Series Analysis using ARIMA Model, Tropical Agricultural Research Vol. 24 (1), pp 21 - 30.
9. Halim, S. And Bisono, I.N. 2008. Automatic seasonal auto regressive moving average models and unit root test detection, International Journal of Management Science and Engineering Management, Vol. 3, No. 4, pp. 266-274.
10. Kumar, R., K. Srinivas and N. Sivaramane (2013), Assessment of the maize situation, outlook and investment opportunities in India. MAIZE-CRP Project Report, National Academy of Agricultural Research Management, Hyderabad, India.
11. Ljung, G. M. and Box, G. E. P. (1978), On a Measure of Lack of Fit in Time Series Models, Biometrika, 65, 297–303.
12. Raghavender, M. (2010), Forecasting paddy production in Andhra Pradesh with ARIMA model, Int. J. Agricult. Stat. Sci., Vol 6 (1), pp 251-258.
13. Ranjit Kumar, K. Srinivas, Naveen Kumar Boiroju and Pravin C. Gedam, (2014), Production performance of maize in India: Approaching an inflection point, Int. J. Agricult. Stat. Sci., Vol. 10 (1), pp241-248.