Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, September : 2023

Predictive Analysis on Tobacco Prices in India using Box-Jenkins methodology

P. Naresh Kumar[1], M. Uday Shankar [2] and M. Raghavender Sharma [3]

1. *P. Naresh Kumar, Research Scholar, Department of statistics, University College of Science, Osmania University, Hyderabad, Telangana. Email:naresh.stats@gmail.com:,Cell:8897374504*
2. *Dr. M. Udayshankar, University College of Science, Osmania University, Hyderabad, Telangana. Email:udayshankar.malyala@gmail.com, , Cell: 91-95380 02202.*
3. *Dr. M. Raghavender Sharma, Assistant Professor, Department of Statistics, Osmania University, Hyderabad, Telangana. Email:drmrsstatou@gmail.com , Cell: 91-94410 90239.*

Abstract:

In this paper studies the Box-Jenkins methodology is used to analyse and develop the forecasted models on prices of tobacco and public impact on uses of tobacco in different categories in India. In India alone one million deaths are create annually by smoking. The auto regressive integrated moving average model has been used for the forecast the next one month's monthly average prices of tobacco in India. The model performance is used for validation data set based on the error measures.

Keywords: ARIMA, MAPE, MAE, RMSE, Tobacco.

1. Introduction:

The plant leaves have more level of addictive chemical nicotine and after harvesting, the leaves are cured, agred, and continued in multiple approaches. The final outcome products are type of smoke categories such as cigarettes, cigars and pipes and claim to the gums. The cured tobacco leaves and the products prepared from them exists several serious cancers reasons chemical compounds and uses of tobacco and improve the second hand tobacco smoke have been actively connected to several forms of cancer and other sickness. The uses of tobaccos are drastically increasing and that impact showed in the prices. The prices of tobacco (Rs./Quintal) in India presented in the following figure 1.1, it is observed that the 2010 April to 2011 Dec the prices are decreased due to lack of tobacco leaves and production also less. From 2012 January again the prices are increased as 36% more than the actual one like 7381.25 (Rs./Quintal) .

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, September : 2023

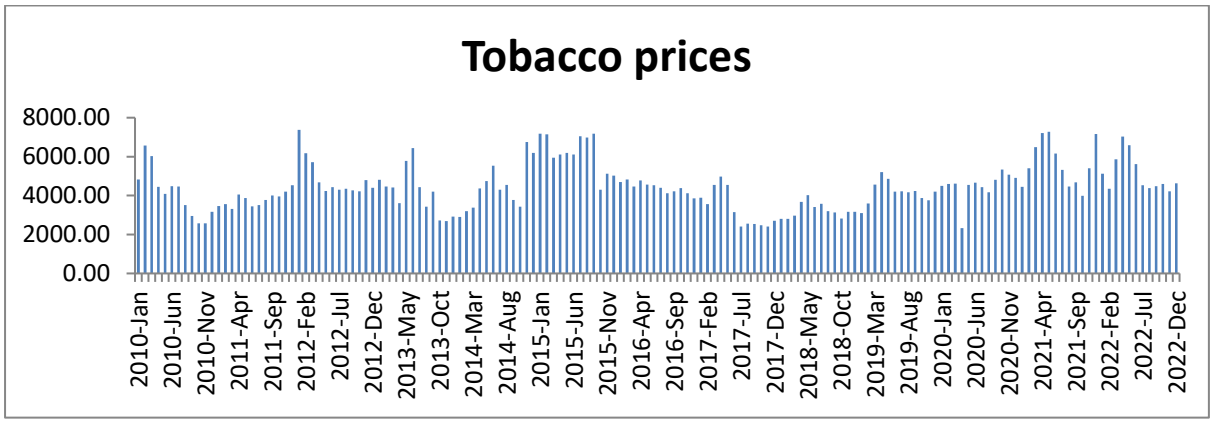

Figure 1.1: Time Series plot on monthly average prices of tobacco in India (Rs./Quintal)

2. Material and Methods:

A historical data on monthly average prices of tobacco in India collected from the agmarknet from January 2010 to December 2022 in India. The whole data set is separated into development and validation data sets. The Box-Jenkins methodology is used for forecasting the future monthly average prices of tobacco prices in India. In R software is used for the model development and validation and for the preparation of tables and charts used in Ms-Excel.

2.1 Box-Jenkins methodology:

The Box-Jenkins methodology is used for the identifying the best model and develop the auto regressive integrated moving average model for the given data sets. The Box-Jenkins methodology consists several benefits which are to check the less number of parameters and also test the stationary of the data. This methodology includes four major steps for building and validating the better model. Identification, Estimation, Diagnostic checking and Forecasting. The first step model identification is to get the model parameters by using auto correlation and partial auto correlation functions for the stationary data. The ACF dies out for the several lags and q spikes in the ACF plots, then parameter q and PACF dies out for several lags and p spikes in the PACF plots, then this is p parameter. The Diagnostic check is to check the model by using LJung Box – Q Statistics and verify the underline assumptions are satisfied based on errors are random. The

selected parameters must be significant, if not repeat the process until to get the significant parameter with adequate model. Test the several models and identify the best model for forecasting the future monthly average prices of tobacco in India. The model performance is tested by using the error measures such as root mean square error (RMSE). Mean absolute error (MAE) and Mean absolute percentage error (MAPE). Therefore the ARIMA model is

$$\Phi(B)\,\nabla^d = \theta(B)\,a_t$$

Where $\Phi(B) = (1 - \Phi_1 B - \Phi_2 B^2 \text{----} \Phi_p B^p)$

$$\theta(B) = (1 - \theta_1 B \text{------} \theta_q B^q) a_t$$

and $\nabla = (1-B)$

3. Results and Discussion:

The time series plot of monthly average prices of tobacco in India from January 2010 to December 2022. The data is separated in to training data for model development (contain 146 observation) and test data set for model validation (contain 12 observations. The following figure 3.1 and 3.2, gives the data patter and ACF and PACF plots.

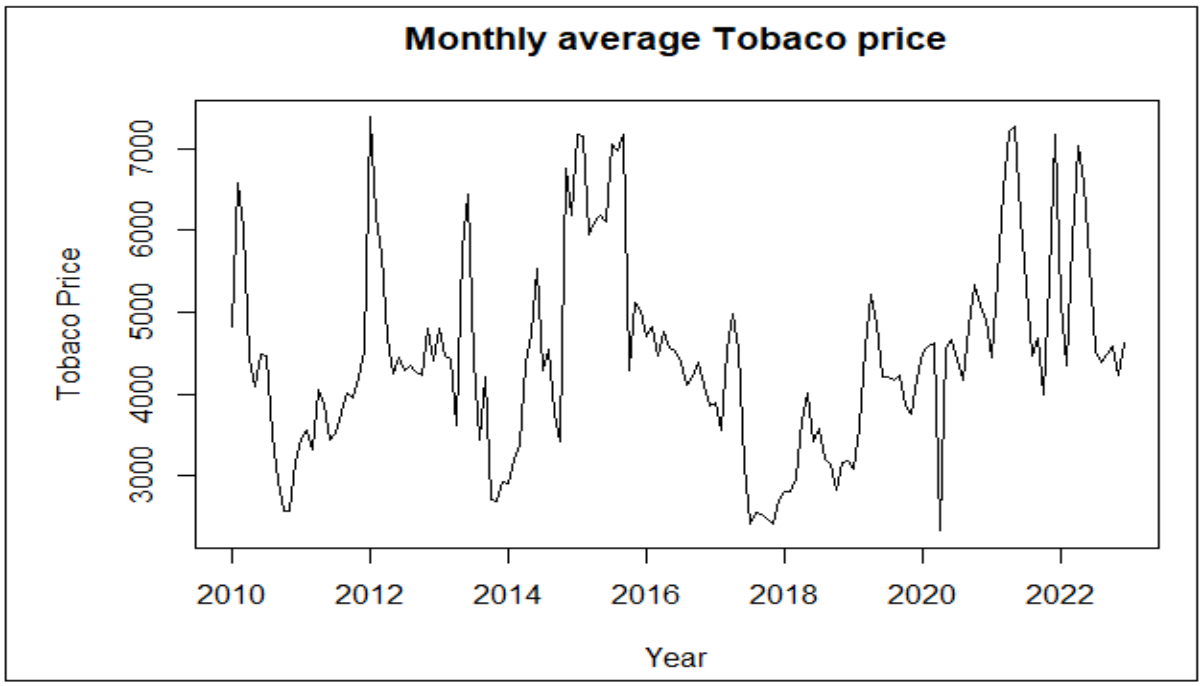

Figure 3.1: Data patter of monthly average prices of tobacco in India (Rs./Quintal)

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 9, September : 2023

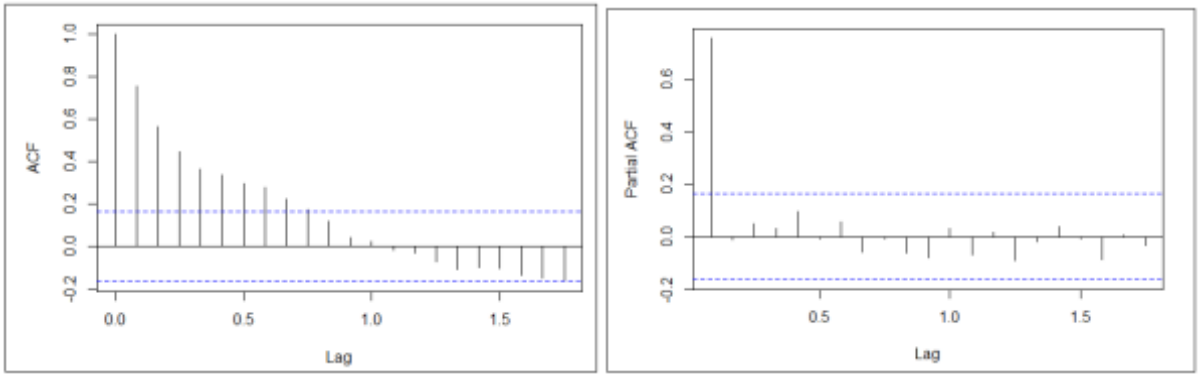

Figure 3.2: ACF and PACF plots for monthly average prices of tobacco in India (Rs./Quintal).

From the above figure 3.1 and 3.2, it was observed that the data is not a constant with mean and variance. Therefore data is not stationary and also more fluctuations are presented. The lowest monthly average prices of tobacco was observed in April 2020 and maximum monthly average prices of tobacco was observed January 2012. For making the stationary of the data applied transformation with one difference (d=1) then the data makes the stationary.

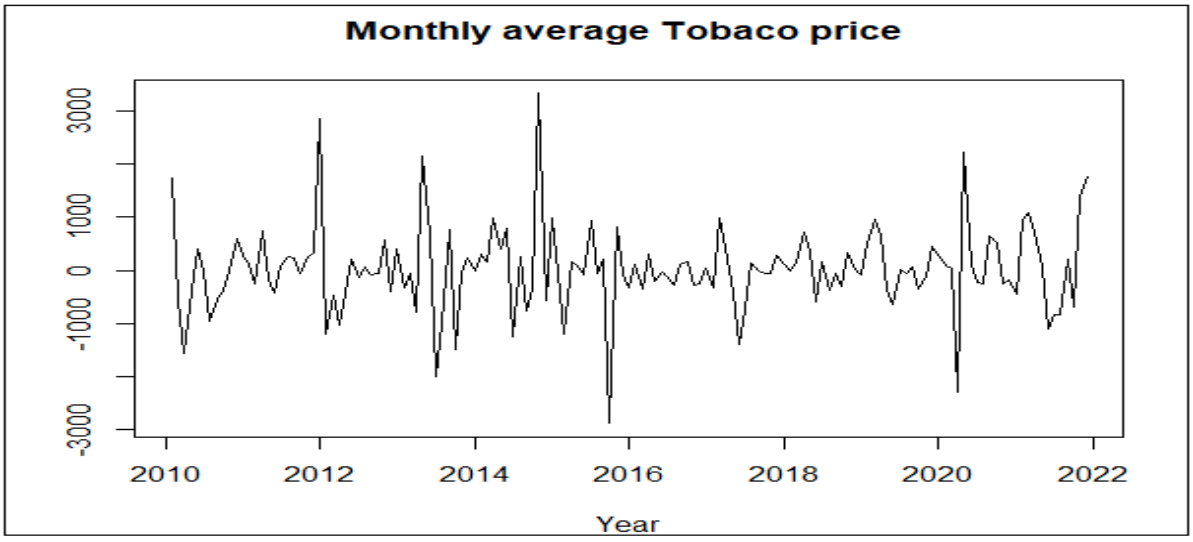

Figure 3.3: Transformed data patter of monthly average prices of tobacco in India

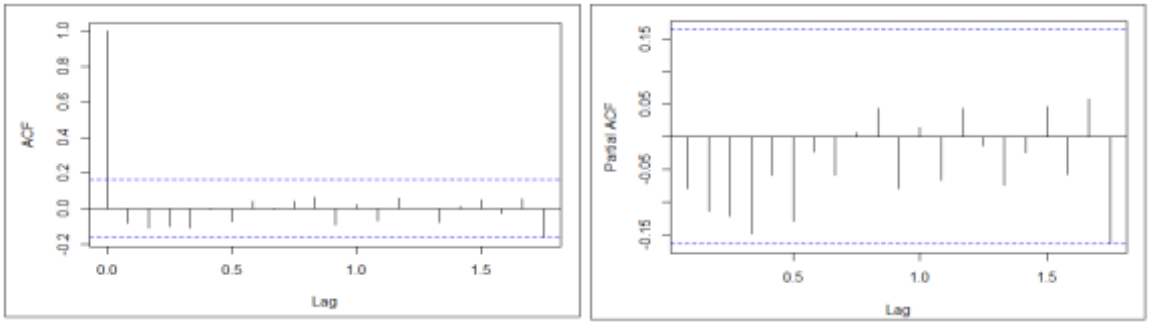

Figure 3.4: ACF and PACF plots for monthly average prices of tobacco in India

From the above figure 3.3 and 3.4, it was observed that the data is a constant with mean and variance. Therefore it is stationary. The suitable and best model is also identified based on the ACF and PACF plots and tested the several models. The following table 3.1.1, shows the some possible models on monthly averages prices of tobacco in India.

Table 3.1.1: Possible models on ARIMA

ARIMA(p,d,q)			AIC	BIC	Significance of the parameter	LJung Box Q statistics	P-Value	Adequacy
1	0	0	2327.95	2336.86	Significant	12.066	0.9694	Adequate
1	1	0	2323.82	2329.75	Insignificant	17.382	0.7901	Adequate
0	1	1	2323.47	2329.4	Insignificant	17.598	0.7789	Adequate
1	1	1	2314.76	2323.65	Significant	12.279	0.9514	Adequate
2	1	0	2323.71	2332.6	Insignificant	17.47	0.7369	Adequate
0	1	2	2320.67	2329.55	Significant	15.96	0.8179	Adequate
2	1	1	2317.64	2329.49	Insignificant	11.984	0.9401	Adequate
1	1	2	2316.67	2328.52	Insignificant	12.369	0.9291	Adequate

From the above table 3.1.1, it was observed that the AROIMA (1, 1, 1) model is selected based on the minimum AIC and BIC with parameters are significant and model is adequate as compared with the other models. The ARIMA (1, 1, 1) model is best model for forecast the future monthly average prices of tobacco. The following table 3.1.2 presents the model parameters on AR1 and MA 1.

Table 3.1.2: Model Parameters:

Parameters	Estimated	Std.Error	z value	Pr(>|z|)
ar1	0.79	0.06	14.39	<0.001
ma1	-1.00	0.02	-40.10	<0.001

The above table 3.1.2, shows that the parameters values are submitted in the model and optimum equations is

$(1-0.79B)\ (1-B)\ Y_t = 2a_t$

Now, the forecasted monthly average prices of tobacco in India are listed in the following table from January 2022 to December 2022.

Table 3.1.4: Forecasting of monthly average prices of tobacco using ARIMA(1, 1, 1)

Year-Month	Actual tobacco prices	Forecasted tobacco prices
2022-Jan	5126.85	6612.73
2022-Feb	4353.85	6173.43
2022-Mar	5866.02	5825.04
2022-Apr	7026.27	5548.75
2022-May	6587.08	5329.64
2022-Jun	5609.15	5155.88
2022-Jul	4529.73	5018.07
2022-Aug	4377.50	4908.79
2022-Sep	4479.13	4822.12
2022-Oct	4589.94	4753.39
2022-Nov	4218.94	4698.88
2022-Dec	4625.15	4655.66

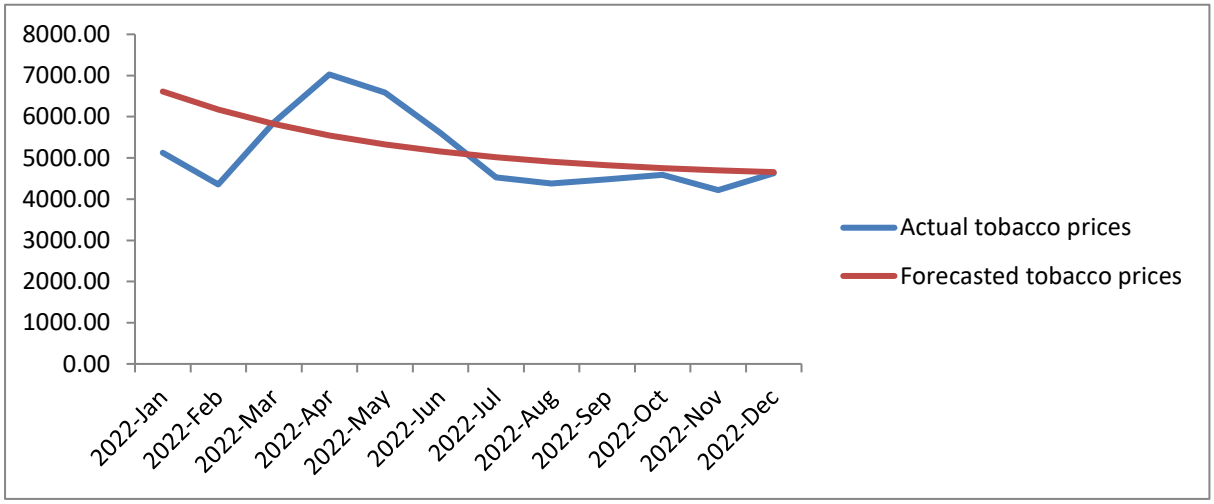

Figure 3.1.2: Forecasting of monthly average prices of tobacco using ARIMA(1, 1, 1)

The model is developed on training and testing data sets is utilized for validation of the selected model performance. The following table shows the error measure of the model ARIMA (1, 1, 1).

Table 3.1.5: Error measures of model ARIMA (1, 1, 1)

Data	RMSE	MAE	MAPE

Training set	765.22	523.99	12.09
Test set	930.34	714.27	13.82

4. Conclusion:

From the above analysis on auto regressive integrated moving average model is shows the better results for forecasting the future monthly average prices of tobacco in India. The performance of the model from the above table 3.1.5, it was observed that the error measures are very near with training and test data set such as mean absolute percentage error was 12.09 (Rs./Quintal) and 13.82 (Rs./Quintal) respectively. Therefore the MAPE was increased by 2% as compared with both the data set. We concluded that the tobacco prices are goes up because of usage of tobacco is getting high. The government has to take the some more additional activity in public and colleges to get more awareness to stop the usage of tobacco and this will help the people in their health.

5. References:

1. M.UdayShankar (2020) et al, FORECASTING OF EGG PRICES IN TELANGANA USING R-SOFTWARE, INDIAN JOURNAL OF APPLIED RESEARCH, Volume -10 | Issue - 3 | March - 2020 | PRINT ISSN No. 2249 - 555X | DOI : 10.36106/ijar.
2. D Srinivas (2022) et al, Forecasting of SBI Stock Prices in India using ARIMA models, JUNI KHYAT JOURNAL, ISSN:2278-4632, Volume -12| | Issue - 08 | August – 2022.
3. Anderson, B. and Ledolder, J.: Statistical Methods for Forecasting, John Wiley & Sons, New York, 1983
4. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31(3) (1986) 309-317.
5. Srinivas Dande, Dr.M.Udayshankar, Dr.M. Raghavender Sharma[2022].: Forecast of SBI stock prices in India Using ARIMA models, JUNI KHYAT JOURNAL ISSN: 2278-4632
6. Banhi Guha and Gautam Bandyopadhyay,Gold Price Forecasting using Arima model, Journal of Advanced Management Science, Vol 4, No.2, March 2016.
7. Raghavender, M. (2009), Forecasting paddy yield in Andhra Pradesh using seasonal Time Series Model, Bulletin of Pure and applied sciences.
8. R.Rama Krishna & Naveen Kumar.B, 2013, Forecasting yield per Hectare of Rice in AP; IJMCAR, ISSN 2249-6955, Volume 3, Issue 1, March 2013, 9-14.
9. Hornik, 1993, Some New result on neural network approximation, Neural Network, 6(1993), pp. 1069-1072.
10. Box, G. E. P. and Pierce, D.: Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models, Journal of the American Statistical Association 65 (1970) 1509-1526.
11. M.A.Khan, 2013, “Forecasting of Gold Prices (Box Jenkins Approach)”. ‘International Journal of Emerging Technology and Advanced Engineering, Vol.3, PP.662-67.
12. Raghavender, M. (2010), Forecasting paddy production in Andhra Pradesh with ARIMA model, Int. J. Agricult. Stat. Sci., Vol 6 (1), pp 251-258.