Acta Scientific COMPUTER SCIENCES

Volume 4 Issue 5 May 2022

Research Article

Use of Auto Regressive Moving Average and Neural Network Method for Predicting Tea Prices in Assam

Mahuya Deb*

Department of Advanced Computing, St. Joseph's College, India

***Corresponding Author:** Mahuya Deb, Department of Advanced Computing, St. Joseph's College, India.

Received: November 25, 2021

Published: April 26, 2022

Abstract

Tea constitutes a major cash crop and earns huge revenue for the state of Assam. Assam being a big exporter of tea can benefit if an early warning mechanism for the price of tea could be obtained in advance. This could benefit the buyers to calculate the price and profit before they plan to engage in the buying activity so as to retain their position both nationally as well as internationally. Such calculative decisions are important in respect to quality, quantity, time and venue since auctions are held in different parts of the country. But many a times some random fluctuation may creep in which might pose a threat for maintaining a uniformity in the tea prices over the years. As per reports due to pandemic, the local consumption of tea has increased by upto 5% therefore skewing past data trends. The packet tea enterprise have shifted their focus from their own particular blends and this has resulted in a sudden crash in prices. These factors present a serious challenge in predicting the tea prices of a particular category. The price fluctuations are dependent on the availability of specific type of tea, credit periods and the total demand at the time of bidding. Hence these fluctuations have been analyzed in detail from historical data using various statistical techniques to get an understanding of the process behaviour Amongst the various techniques used for forecasting, a time series forecasting tool called ARIMA (Auto Regressive Integrated Moving average) method has been used from time to time. In addition to this, a neural network method has also been applied and further a valid comparison has been drawn between ARIMA and Neural Network. The data of tea prices from the year July 2005 to May2020 has been collected and trained using ARIMA and Neural Network model with different parameters The test criterion like Akaike Information Criterion ~~is~~ has been applied to analyse the accuracy of the model. The one with the least AIC is chosen. The model is then used to forecast the price of tea for the next ten data points where the prediction error appears small. The maximum variability in tea price has been observed in the latter months of 2021 specifically May, June, July and August. A reduced variability in the tea price has been observed in the months of June, July and August when the pandemic was taking a peak. For this, R programming software is used which is an effective tool for visualization, statistical computing, scientific inference, and graphical interface.

Keywords: Arima; Neural Network Method; Forecasting

Introduction

Traditionally, tea constitutes an important agricultural export item and is one of the oldest organized industries in the State. The sector has seen many ups and downs in recent years perhaps due to factors like decline in cultivation area, production and export of tea, increase in domestic consumption etc. These fluctuations arise an interest to understand the pattern in tea prices so that appropriate measures could be adopted by the respective stakeholders.

Against this perspective time series analysis with some margin of error would give an insight to estimate the future values of a series from its past values. Time series has wide application in different field ranging from economy, finance, business etc to name a few. Among the various methods of forecasting ARIMA model is the most widely used. However, in recent times neural network methods has occupied an important place in predictive analytics. Neural networks are a comprehensive family of machine learning and of late their application in economics and finance has increased considerably. The aim of this paper is to see the application of both ARIMA and neural networks in predicting the tea prices in Assam. With this objective an exhaustive literature review is undertaken to get an insight into the forecasting techniques adopted from time to time by researchers in predicting a certain phenomenon.

Literature review

[1] studied the pattern of tea growth in Assam, which was the first tea growingstate in India. She gave a historical perspective of the growth of tea and the measure that the British Govt took for the tea growers' paper [2]. In their study considered two important regions which grew tea i.e., north and south India for studying the volatility in tea prices. Autoregressive integrated moving average (ARIMA) model was used for this analysis. The possible factors identified were weather conditions, and variousproduction related factors [3]. considered India's weekly tea auction price have been considered in the year 2013 to 2014 followed by application of the ARMA model to analyse the data and tested it for stability, autocorrelation and partial correlation test. Finally, the ARIMA (1,1) was established [4]. in their article reviewed and examined the approaches for modelling tea price. They found the VAR techniques more useful as it could model the non-structural relationship of tea price alongside other time series variables which are endogenous and exogenous in nature. They contributed to the existing literature regarding the use of various forecasting tools in tea prices [5]. used seasonal autoregressive integrated moving average modelling techniques in an analysis of monthly cocoa prices [6].

Considered the Auto-Regressive Integrated Moving Average (ARIMA) model to predict the amount of tea production in Bangladesh. The predicted model behaved good enough statistically. Similar work was carried out by [7] in Sri Lanka. Box and Jenkins ARIMA model approach was applied to fit the time sris model [8]. In his paper studied the use of artificial neural network as an advanced time series forecasting models through a systematic literature review.

Objective

The primary objective in this study is to use forecasting tools like ARIMA and neural networkmethods for predicting the tea prices of Assam. Data for the same is collected from July 2005 to March 2020(www. indexmundi.com)

Method and Methodology

Autoregressive Integrated Moving Average (ARIMA) model of Time Series Analysis Technique which is a generalization of Auto Regressive (AR) and Moving Average (MA) including the notion of stationary (integration) is used in this paper. Considered to be one of the most popular models it is used to predict a linear time series data, Auto Regression Integrated Moving Average is specified by three ordered parameters (p, d, q) wherep denotes the order of the autoregressive model (time lags), d is the degree of differencing, q is the order of moving average model.

Before developing an ARIMA model stationarity of the data needs to be ascertained i.e., whether mean and variance of the data is constant over time. For this it is essential to check the correlogram or the autocorrelation function (ACF) and partial ACF (PACF) of the time series data. In ARIMA model the future value of the variable is expressed as a linear combination of past values and past errors.

ARIMA model in words

Predicted Yt is given by the following equation

Y_t = constant +Lagged values of Y (upto p lags) + Lagged forecast errors (upto q lags)

Predicted Yt = Constant + Linear combination Lags of Y (upto p lags) + Linear Combination of Lagged forecast errors (upto q lags)

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + .. + \beta_p Y_{t-p} \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + .. + \phi_q \epsilon_{t-q} \quad \text{........ Eq (1)}$$

i.e. $$Y_t = c_0 + \sum_{j=1}^{p} a_j Y_{t-j} + \sum_{i=1}^{q} c_i Y_{t-i}$$

The data is analysed using R programming. The various steps that are followed here are stated below:

- Test for stationarity - adf. test () from library t series
- Testing for autocorrelations -acf () and pacf () plots from R base or tsdisplay () from library(forecast)
- Use of dfw () for differencing
- Setting up an ARIMA Model
- Checking the AIC information criterion for model accuracy
- Forecast the future based on the selected model
- Checking the residuals -check residuals () from library 'forecast'

The data related to the tea prices is obtained from the year June 2005 -May 2020. The graph for the data is obtained as follows. The graph for the same is plotted with tea price along the y-axis and time points along the x-axis as viewed in figure 1. The graph clearly indicates fluctuations in the price of tea with a steep rise around the year 2015-17 and again a sharp decline thereafter little ups and downs throughout the time periods. A closer look at the data is revealed through decomposing the data which is presented in figure 2. Clearly there is no evidence of stationarity in the given data. This is tested using the Augmented Dicky Fuller test.

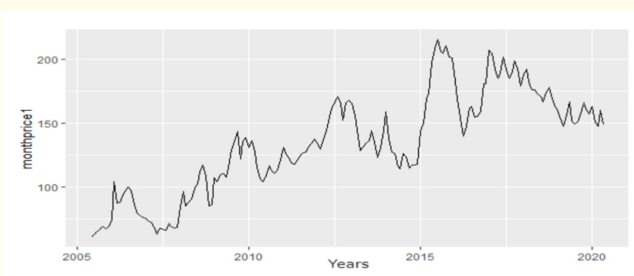

Figure 1: The graph of the time series.

Now we use the following test for ARIMAadf. test (monthprice1)

The software generated a set of indicators which suggest the presence/absence of stationarity in the data.

Augmented Dickey-Fuller Test

data: monthprice1

Dickey-Fuller = -3.3012, Lag order = 5, p-value = 0.07304alternative hypothesis: stationary

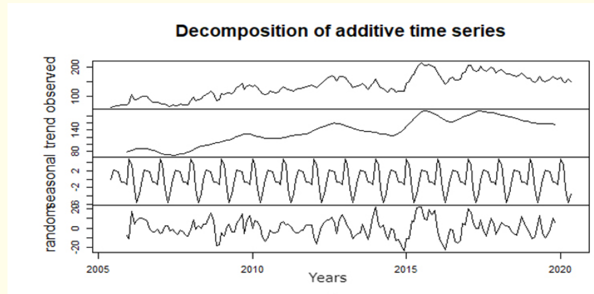

Figure 2: Decomposition of time series.

Acf (monthprice1)

As the p value is greater than 0.05 this suggests that we accept the null hypothesis of no stationarity in the data. This isalso confirmed using the ACF and PACF plots.

The plots to suggest that the data is not stationary as the values cross the threshold values at every point

Therefore, to ensure that the data is stationary we take the difference of 1st order using Durbin-Watson test. Thisis ensured using the following function:

Dw test (monthprice1 [-180] ~ monthprice1 [-1])

Durbin-Watson test

data: (monthprice1[-180] ~ monthprice1[-1]) DW = 1.7102, p-value = 0.08151

Alternative hypothesis: true autocorrelation is greater than 0

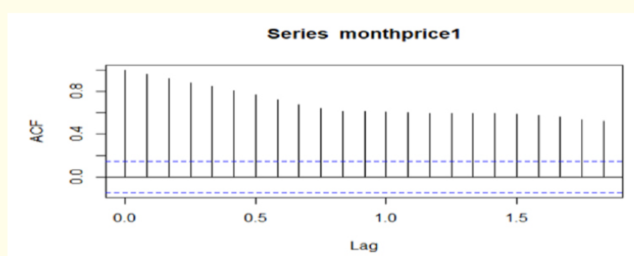

Figure 3

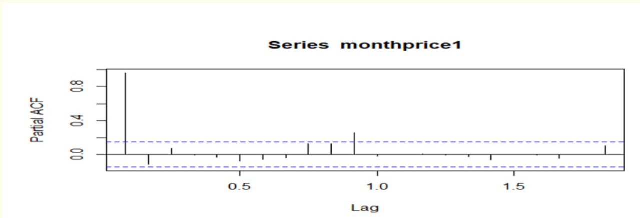

Figure 4

As p > .05 it confirms the absence of autocorrelation is the data. Hence it is desired to fit an ARIMA model. The choice of the best model depends on the AIC criterion. Among the various models tested the one with the lowestAIC is fitted.

auto. arima (teaprice1, trace = T, stepwise = F, approximation = F)Best model: ARIMA (2, 1, 3)

Series: month price1ARIMA (2, 1, 3)

Coefficients:

	ar1	ar2	ma1	ma2	ma3
	1.1882	00.7885	-1.1062	0.5400	0.2445
S. E	0.0885	0.0840	0.1057	0.1321	0.0764

Table a

sigma ^2 estimated as 78.22: log likelihood = -642AIC = 1296 AICc = 1296.49 BIC = 1315.12

Once the ARIMA model is fitted next step that follows is to forecast the future values of tea prices in Assam.

Arimafore < -forecast (monthprice1, h = 15) arimafore

Now the model is checked for residuals.Check residuals (arimafore)

Ljung-Box test

data: Residuals from ETS (A, N, N)

Q* = 65.608, df = 22, p-value = 3.187e-06Model df: 2. Total lags used: 24

Point Forecast	Point Forecast	Lo95	Hi 95
Jun-20	149.0511	130.91	167.18
Jul-20	149.0511	123.4	174.69
Aug-20	149.0511	117.64	180.45
Sep-20	149.0511	112.78	185.31
Oct-20	149.0511	108.5	189.59
Nov-20	149.0511	104.63	193.46
Dec-20	149.0511	101.07	197.02
Jan-21	149.0511	97.76	200.33
Feb-21	149.0511	94.65	203.44
Mar-21	149.0511	91.71	206.39
Apr-21	149.0511	88.91	209.18
May-21	149.0511	86.23	211.86
Jun-21	149.0511	83.67	214.4
Jul-21	149.0511	81.2	216.89
Aug-21	149.0511	78.82	219.27

Table b

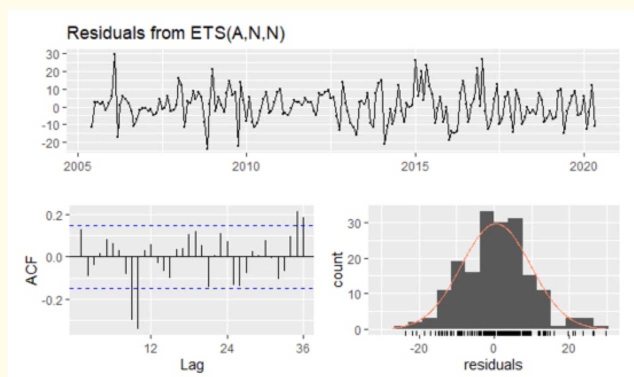

Figure 5

The residual test ensures normality of the model.

Now we use the neural network modelfit = nnetar (monthprice1)

nnet forecast<-forecast (fit, h = 15, PI = F) nanoforests

Year	Point forecast
Jun 2020	146.8527
July	146.5185
Aug	146.7491
Sept	147.5081
Oct	148.9843
Nov	149.9267
Dec	150.3161
Jan 2021	150.9743
Feb	150.7753
Mar	150.2871
April	150.5262
May	150.2109
June	149.7146
July	149.2214
August	148.7960

Table c

plot(nnetforecast)

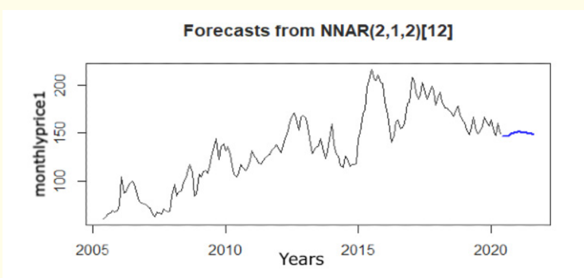

Figure 6

A comparison between ARIMA and Neural Network can be obtained from the graph below. While the ARIMA plot shows a horizontal straight line, neural network plot shows a small descend and then again, a rise in the prices giving a more real time fluctuation.

Autoplot (monthprice1) + forecast: autolayer (nnetforecast $ mean, series = 'nnetar model') + forecast: autolayer (arimafore $ mean, series = 'ARIMA model')

Figure 7

Conclusion

One of the most important elements of decision making is predicting the occurrence ofmacroeconomic variables. Exponential smoothing tools and ARIMA have been used for quite sometimes and of late neural networks have occupied a much better position with regard to forecasting.While the ARIMA forecasted values were mostly similar, the neural network model reflected a pattern. This calls for introspection on the part of those people engaged in this trade to understand the factors responsible for such ups and downs in tea prices. It also makes us believe that the better the tool used the better is the prediction or in other words as we go on increasing the complexities in the model betterare the predictions. However, it deserves special mention that this modelling is obtained ignoring the COVID influence prevalent in recent times.

Bibliography

1. Roy S. “Historical Review of Growth of Tea Industries in India: A Study of Assam Tea”. InternationalConference on Social Science and Humanity, IPEDR. IACSIT Press, Singapore 5 (2011).
2. Chaudhry S., *et al*. “A time series analysis of auction prices of Indian tea”. *International Journal of Research in Economics and Social Sciences* 7.6 (2017): 100-111.
3. Gunathilaka RD and Tularam GA. “The tea industry and a review of its price modelling in major teaproducing countries”. *Journal of Management and Strategy* 7 (2016): 21-36.
4. Hossain M and Abdulla F. “Forecasting the Tea Production of Bangladesh: Application of ARIMA Model (2015).

5. Kumarasinghe HPASS and Peiris BL. "Forecasting annual tea production in Sri Lanka (2018).

6. Liu H and Shao S. "India's tea price analysis based on ARMA model". *Modern Economy* 7.2 (2016): 118-123.

7. Sapsford D and Varoufakis Y. "An ARIMA analysis of tea prices". *Journal of AgriculturalEconomics* 38.2 (1987): 329-334.

8. Tealab A. "Time series forecasting using artificial neural networks methodologies: A systematic review". *Future Computing and Informatics Journal* 3.2 (2018): 334-340.