Jordan Journal of Mathematics and Statistics (JJMS) 8(3), 2015, pp. 257- 270

FORECASTING THE TEA PRODUCTION OF BANGLADESH:APPLICATION OF ARIMA MODEL

MD. MOYAZZEM HOSSAIN [(1)] AND FARUQ ABDULLA [(2)]

ABSTRACT: Bangladesh is the world's 10^{th} largest tea producer and fifteen number exporters and sixteen number consumers in the world. The consumption is increasing day by day mainly due to the rapid increase in population. The main purpose of this research is to identify the Auto-Regressive Integrated Moving Average (ARIMA) model that could be used to forecast the production of tea in Bangladesh. This study considered the published secondary data of yearly tea production in Bangladesh over the period 1972 to 2013. According to AIC, AIC_C and BIC, the most suitable model to forecast the tea productions in Bangladesh is ARIMA (0,2,1). Adequacy of the fitted model has been tested using Run test and Jarque and Bera test criteria followed by residual analysis. The comparison between the original series and forecasted series shows the same manner indicating the fitted model behaved statistically well and suitable to forecast the Tea productions in Bangladesh i.e., the models forecast well during and beyond the estimation period.

1. INTRODUCTION

Tea Camellia sinensis L. [28] is a unique crop relative to any others typical crop due to its cultivation and harvesting system. It is a type of crop which shows wide adaptability and grows in a range of climates and soils in various parts of the world Hamid [12]. Mondal et al. [32] mentioned in his research that Tea is the oldest non alcoholic caffeine containing beverage in the world. Khisa and Iqbal [29] said that now-a-days tea is the most popular drink all over the world. Modern man cannot think of starting a day without having a cup of

2000 Mathematics Subject Classification. 62E05, 62E99.

Keywords and phrases. Tea, ARIMA Model, Forecasting, Bangladesh

Received: April 20, 2015 Accepted: Oct. 6, 2015

tea. It refreshes the mind and gives energy. Almost all people take tea once or twice a day. Bekhit [28] said that the Chinese were the first to use tea as medicinal drink, later as beverage and have been doing so for the past 3000 years. Redowan and Kanan [27] mentioned that the art of tea cultivation in Bangladesh began over a century and a half ago in the 1840s near the Chittagong Club and first tea garden for commercial purpose was established at Malnicherra in Sylhet in 1854.

Its commercial production began shortly thereafter in 1857 and this same year Bangladesh Tea Board was established in Dhaka and Bangladesh Tea Research Institute (BTRI) was founded in Srimangal see Nasir and Shamsuddhoa [33]. The tea cultivation in Bangladesh has been expanding since then. At present there are 163 tea gardens in the country see BTD [5]. In Bangladesh, tea grows well at only 80300 ft. above from sea level mainly in the hilly regions northeast of Sylhet (Sylhet, Moulvibazar and Habibgonj districts) and southeast of Chittagong see BTRI [6]. But few tea gardens also present in Brahmanbaria and Panchagar districts. From total annual production, 94% comes from Sylhet (63% Moulvibazar district) from and rest from others part of the country. Bangladesh is the world’s 10th largest tea producer and fifteen number exporters and sixteen number consumers in the world BTD [5]. Rahman [25] has investigated the impact of price and other factors on the supply of tea in Bangladesh by the modified form of the Cobweb supply model.

The relationship between tea drinking and human health has become a subject of intense study by scientists throughout the world in the past two decades. These studies suggest that tea may be beneficial to human health for details see Weisburger [20]; Chen [37]; Yang et al. [9]; Blumberg [18]. Researcher like Weisburger [20][21]; Yang et al. [8]; McKay and Blumberg [10]; Blumberg [18]; Chen [38]; Huang and Xu [16] conclude that Tea may have great potential for the development of medication and in the prevention of a number of human diseases including cancer. Yong-xing Zhu et al. [36] showed that tea and tea constituents have various biological activities and suggested that tea drinking might be beneficial to human health. Tea has potential in the prevention or adjuvant treatment of several diseases including cancer, cardiovascular diseases and obesity. Kamath et al. [2]

mentioned that Researchers from Brigham and Women's Hospital and Harvard University published novel data indicating that tea contains a component that can help the body ward of infection and disease and that drinking tea may strengthen the immune system. ADA [3] said that Type 2 diabetes is considered a global epidemic. Stote and Baer [24]; Anderson and Polansky [30] said that Catechins in tea have been shown to help reduce blood sugar and provide insulinboosting activity, which may be benefiial for people with both type 1 and type 2 diabetes. Linke and LeGeros [15] said that a recent study conducted at the New York University Dental Center examined the effects of Black Tea extract on dental caries formation in hamsters. Compared to those who were fed water with their food, hamsters that were fed water with Black Tea extract developed up to 63.7 percent fewer dental caries.

Hazarika [19] has been developed an Autoregressive Integrated Moving Average (ARIMA) model with the help of Box- Jenkins approach for the tea production of Assam. He has shown that the selected ARIMA model is quite appropriate for the tea production of Assam, India. Gijo, E.V. [11] was modelled by Box-Jenkins seasonal auto regressive integrated moving average (ARIMA) model for a tea packaging company in India and shown that his model has helped the organisation to plan the production activities more efficiently. Dhekale, B. S., et al. [4] analyze the growth and trend behavior of tea production scenario in West Bengal. They used ARIMA model to forecast tea production. Ahammed [22] analyzed the trends in Bangladesh tea using polynomial models. Borodoloi [23] analyzed global tea production and export trend of India using linear regression analysis. Dissanayake [26] was forecasted Tea production in Sri Lanka using ARIMA models. Dutta et al. [31] studied the linear relationship between rainfall and fertilizer with north east tea production. Nasir and Shamsuddoha [33] mentioned that In Bangladesh the production has increased by 1.84 % and contributes 1.37 in export in the word tea trade and earns near about 1775 million Taka (Taka 69 = USD 1.00) every year. Kamal and Bhuiyan [35] said that the tea industry of Bangladesh is not only provides a huge amount of foreign currency, but also provides a lot of employment.

Taking tea is an integral part of social life in Bangladesh. The consumption is increasing day by day mainly due to the rapid increase in population. The tea producing industry has been traditionally regarded as one of the major agro-based labor intensive industry and occupies an important role in the national economy of Bangladesh. Tea is the most important agriculture crop which plays a great role to earn foreign money. Moreover, a large number of people were involved in the production and marketing of tea. Thus, it is necessary to estimate the tea production in Bangladesh which leads us to do this research. The main purpose of this research is to identify the Auto-Regressive Integrated Moving Average (ARIMA) model that could be used to forecast the tea production in Bangladesh. The R programming language of version 3.1.3 is used to analyze this data set.

2. MATERIALS AND METHODS

2.1 Data Source.

This study considered the published secondary data of yearly tea production in Bangladesh which was collected over the period 1972 to 2013 from the Food and Agricultural Organization (FAO) website (http://faostat3.fao.org).

2.2 ARIMA Model.

Suppose that $\{\zeta_t\}$ is a white noise with mean zero variance σ^2, then $\{Y_t\}$ is defined by $Y_t = \zeta_t + \beta_1\zeta_{t-1} + \beta_2\zeta_{t-2} + \ ... + \beta_q\zeta_{t-q}$ is called a moving average process of order q and is denoted by $MA(q)$. If the process $\{Y_t\}$ is given by $Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ ... \ + \alpha_p Y_{t-p} + \zeta_t$ is called an auto-regressive process of order p and is denoted by $AR(p)$. Models that are combination of AR and MA models are known as $ARMA$ models. An $ARMA(p,q)$ model is defined as $Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ ... \ + \alpha_p Y_{t-p} + \zeta_t + \beta_1\zeta_{t-1} + \beta_2\zeta_{t-2} + \ ... \ + \beta_q\zeta_{t-q}$, where, Y_t is the original series, for every t, we assume that ζ_t is independent of $Y_{t-1}, Y_{t-2}, ..., Y_{t-p}$. A time series $\{Y_t\}$ is said to follow an integrated autoregressive moving average (ARIMA) model if the d^{th} difference $W_t = \nabla^d Y_t$ is a stationary ARMA process. If $\{W_t\}$ follows an

$ARMA(p,q)$ model, we say that $\{Y_t\}$ is an $ARIMA(p,d,q)$ process. Fortunately, for practical purposes, we can usually take $d=1$ or at most 2. An $ARIMA(p,1,q)$ process is defined as, $W_t = \alpha_1 W_{t-1} + \ldots + \alpha_p W_{t-p} + \zeta_t + \beta_1 \zeta_{t-1} + \ldots + \beta_q \zeta_{t-q}$, where, $W_t = Y_t - Y_{t-1}$.

2.3 Box-Jenkins Method.

The influential work of Box and Jenkins [14] shifted professional attention away from the stationary serially correlated deviations from deterministic trend paradigm toward the $ARIMA(p,d,q)$ paradigm. It is popular because it can handle any series, stationary or not with or without seasonal elements. The basic steps in the Box-Jenkins methodology consist of the following five steps:

Preliminary Analysis: Create conditions such that the data at hand can be considered as the realization of a stationary stochastic process.

Identification of a Tentative Model: Specify the orders *p*,*d*,*q* of the ARIMA model so that it is clear the number of parameters to estimate. Empirical autocorrelation functions play an extremely important role to recognize the model.

Estimation of the Model: The next step is the estimation of the tentative ARIMA model identified in step-2. By maximum likelihood method we estimate the parameters of the model.

Diagnostic Checking: Check if the model is a good one using tests on the parameters and residuals of the model.

Forecasting: If the model passes the diagnostics step, then it can be used to interpret a phenomenon, forecast.

2.4 Jarque-Bera Test.

The normality assumption is checked by using Jarque and Bera [7] test, which is a goodness of fit measure of departure from normality, based on the sample kurtosis (k) and skewness (s). The test statistics Jarque-Bera (JB) is defined as $JB = \frac{n}{6}\left(s^2 + \frac{(k-3)^2}{4}\right) \sim \chi^2_{(2)}$, where n is the number of observations and k is the sample kurtosis an s is the sample skewness. The statistic JB has an asymptotic chi-square distribution with 2 degrees of freedom, and can be used to test the hypothesis of skewness being zero and excess kurtosis being zero, since sample from a normal distribution have expected skewness of zero and expected excess kurtosis of zero.

2.5 Evaluation of Forecast Error.

Before performing growth analysis it is necessary to estimate the growth model that best fits the time series. There are many summary statistics available in the literature for evaluating the forecast errors of any model, time series or econometric. Here, an attempt is made to identify the best models for tea production in Bangladesh using the following contemporary model selection criteria, such as RMSPE, MPFE and TIC.

Root Mean Square Error Percentage (RMSPE): Root Mean Square Error Percentage (RMSPE) is defined as, $RMSPE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{Y_t^f - Y_t^a}{Y_t^a}\right)^2}$, where Y_t^f is the forecast value in time t and Y_t^a is the actual value in time t.

Mean Percent Forecast Error (MPFE): Mean Percent Forecast Error (MPFE) is defined as, $MPFE = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{Y_t^a - Y_t^f}{Y_t^a}\right)$, where Y_t^a is the actual value in time t and Y_t^f is the forecast value in time t.

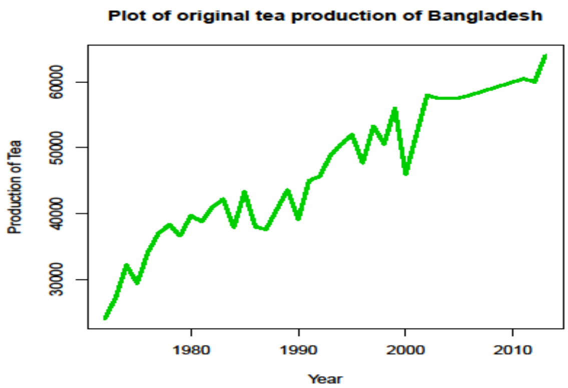
Theil Inequality Coefficient (TIC): Theil (1966) Inequality Coefficient (TIC) is defined as

$$TIC = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t^f - Y_t^a\right)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t^a\right)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t^f\right)^2}}$$

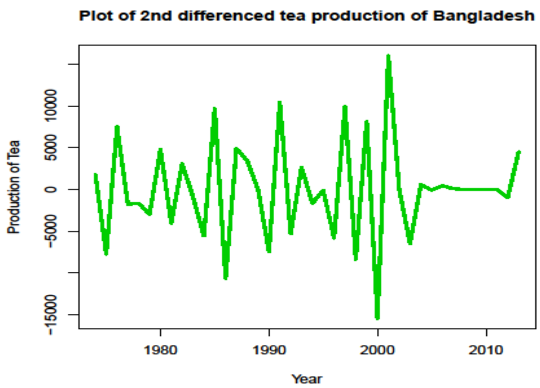
, where Y_t^f is the forecast value in time t and Y_t^a is the actual value in time t.

3. RESULTS AND DISCUSSION

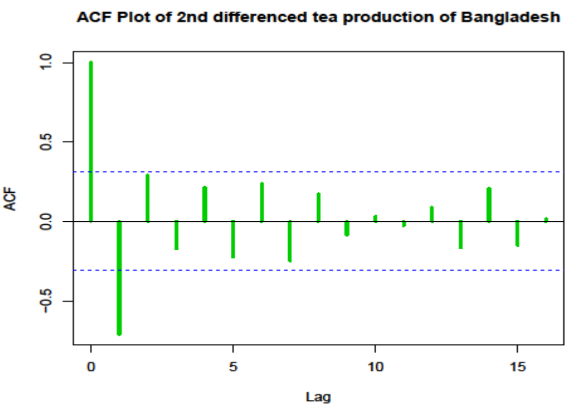
In order to make forecasting a time series it is necessary to check the time series is stationary or not first. In this study Augmented-Dickey-Fuller (ADF) unit root test, Phillips-Perron (PP) unit root test and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) unit root test are used to check whether the data series is stationary or not. After second differencing the Augmented-Dickey-Fuller (ADF) test with $\Pr\left(|\tau| \geq \text{-}5.837\right) < 0.01$, Phillips-Perron (PP) test with $\Pr\left(|\tau| \geq \text{-}59.7534\right) < 0.01$ and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) unit root test with $\Pr\left(|\tau| \geq 0.0451\right) > 0.1$ at 5% level of significance adequately declared that the data series is stationary and suggest that there is no unit root. The graphical representations of the original and second differenced series are presented in Figure 1(a), (b).



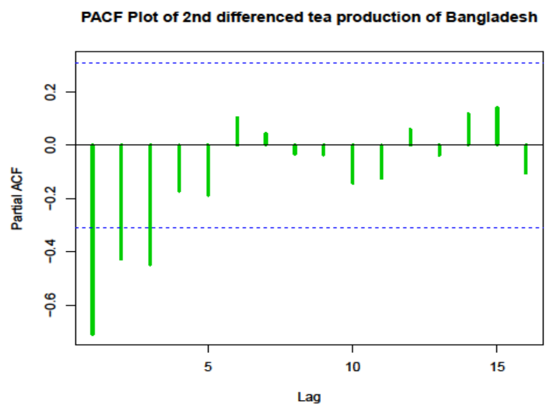
(a)



(b)



(c)



(d)

Figure 1. (a) Time series (original series) plot, (b) Time series (2nd differenced) plot (c) ACF and (d) PACF of 2nd differenced tea production in Bangladesh.

It is clear that the yearly tea production in Bangladesh has gradually an increasing trend with some fluctuation over the study period 1972-2013 i.e., the variance is unstable which leads the tea production data series is not stationary (Figure 1(a)). However, it is clear that the second differenced tea production data series shows stable variance which leads the data becomes stationary. To stabilize the variance and to make the data stationary second difference is enough that is difference order is 2 and it is said that integrated of order 2 (Figure 1(b)). The alternative positive and negative ACF (Figure 1(c)) and exponentially decay PACF (Figure 1(d)) indicates an autoregressive moving average process. The PACF with significant spike at lag 3 and ACF with significant spike at lag 1 suggest that third order autoregressive and first order moving average are effective on tea production in Bangladesh. However, using the tentative procedure, it is clear that ARIMA(0,2,1) model with $\text{AIC} = 774.68$, $\text{AIC}_\text{C} = 775.01$ and $\text{BIC} = 778.06$ is the best selected model for forecasting the tea production in Bangladesh. The estimates of the parameters of the fitted ARIMA(0,2,1) model are shown in Table 1. The value of the most useful “forecasting criteria” of the fitted ARIMA(0,2,1) model are $RMSPE = 0.08049284$, $MPFE = -0.01332399$, and $TIC = 0.0358971$.

Table 1: Summary statistics and forecasting criteria of the fitted ARIMA (0,2,1) model

Coefficients	Estimates	Std.Error	t-value	p-value
ma1	-1.000	0.0674	-14.8368	0.02142169

Several plots of the residuals for the fitted ARIMA(0,2,1) model are presented in Figure 2, suggest that there is no significant pattern and hence there is no autocorrelation among the residuals. Also, the lower and upper values of the runs are 15 and 28 respectively which are obtained from Swed and Eisenhart [13] Tables at 5% level of significance, and here the observed number of runs is 25. So the "Run" test with $\Pr(15 \le R \le 28) = 0.95$ at 5% level of significance strongly suggest that there is no autocorrelation among the residuals of the fitted ARIMA(0,2,1) model. Here "Jarque-Bera" test is used to check the normality assumption of the residuals of the fitted model. We observe that the "Jarque-Bera" test with $\Pr\left(\left|\chi_2^2\right| \ge 4.041\right) = 0.1326$ at 5% level of significance moderately suggests to accept the normality assumption that is the residuals of the fitted ARIMA(0,2,1) model are normally distributed. Therefore, it is clear that our fitted ARIMA(0,2,1) model is the best fitted model and adequately used to forecast the tea production in Bangladesh.

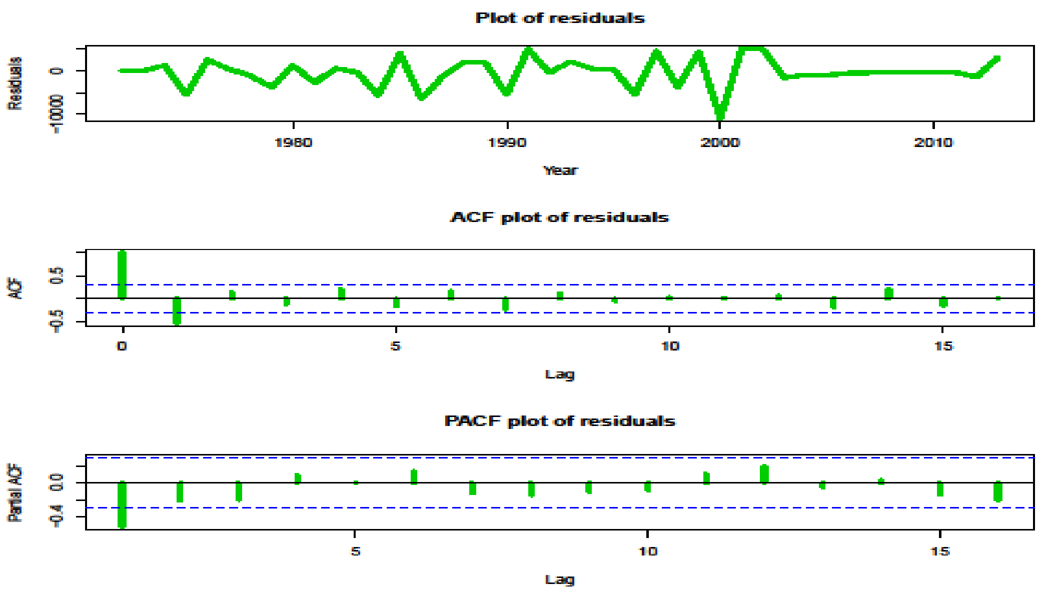


Figure 2. Several plots of residual.

By using the best fitted model ARIMA(0,2,1), the forecasted tea production and 95% confidence level for twelve years are shown in Table 2.

Table 2: Forecasted tea production (tonnes) in Bangladesh

Year	Forecasted	LCL	UCL
2014	64973.95	57977.68	71970.23
2015	65947.91	55936.58	75959.23
2016	66921.86	54518.78	79324.93
2017	67895.81	53412.14	82379.48
2018	68869.76	52497.59	85241.94
2019	69843.72	51715.01	87972.43
2020	70817.67	51029.21	90606.13
2021	71791.62	50417.64	93165.60
2022	72765.58	49864.88	95666.27
2023	73739.53	49359.88	98119.18
2024	74713.48	48894.42	100532.54
2025	75687.43	48462.24	102912.63

Note: LCL= Lower Confidence Limit and UCL=Upper Confidence Limit

The graphical comparison of the original series and the forecast series is shown in Figure 3. It is observed that the forecast series (blue-color) fluctuated from the original series (dark-green-color) with a very small amount that is it shows the production in same manner of the original series (Figure 3). Therefore, the forecasted series is really better representation of the original tea production series in Bangladesh.

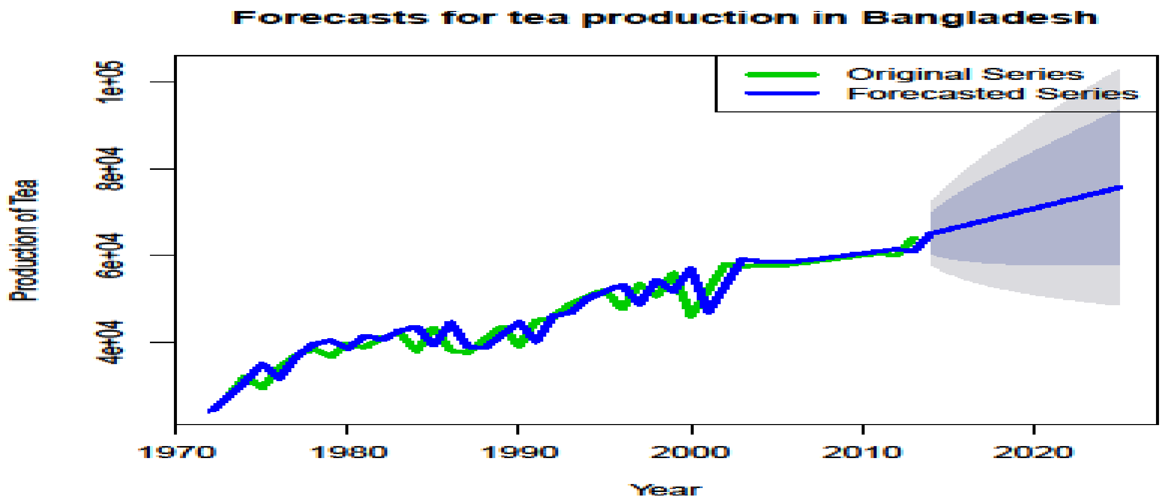


Figure 3: Comparison between the original and forecasted tea production in Bangladesh.

4. Conclusion

The best selected Box-Jenkins ARIMA model for forecasting the tea productions in Bangladesh is ARIMA (0,2,1). The comparison between the original series and forecasted series shows the same manner indicating the fitted model behaved statistically well and suitable to forecast the Tea productions in Bangladesh i.e., the models forecast well during and beyond the estimation period. Thus, this model can be used for policy purposes as far as forecasts the tea production in Bangladesh.

REFERENCES

[1]. Consul, P.C., A simple urn model dependent on predetermined strategy, *Sankhya, Ser. B,* 36(3) (1974) 391-399.

[2]. A. B. Kamath, L. Wang, H. Das, L. Li, V. N. Reinhold, and J. F. Bukowski, Antigens in tea-beverage prime human Vgamma 2Vdelta 2T cells in vitro and in vivo for memory and nonmemory antibacterial cytokine responses, Proc Natl Acad Sci USA 100(2003), 6009-6014.

[3]. American Diabetes Association (ADA), Diagnosis and Classifiation of Diabetes Mellitus, Diabetes Care 35(2012), S64– S71.

[4]. B. S. Dhekale, P. K. Sahu, K. P. Vishwajith, P. Mishra and M. Noman, Modeling and forecasting of tea production in West Bengal, Journal of Crop and Weed 10(2014), 94-103.

[5]. BTD (Bangladesh Tea Board) (2012). http://www. Teaboard.gov.bd/

[6]. BTRI (Bangladesh Tea Research Institute), Brief note on tea culture for the BTRI, Annual Report, Shromongal, Bangladesh, (2012).

[7]. C. M. Jarque and A. K. Bera, A test for normality of observations and regression residuals, International Statistical Review 55(1987), 163–172. JSTOR 1403192.

[8]. C. S. Yang, J. Y. Chung, G. Yang, S. K. Chhabra and M. J. Lee, Tea and tea polyphenols in cancer prevention, The Journal of Nutrition 130(2000), 472S–478S.

[9]. C. S. Yang, P. Maliakal and X. Meng, Inhibition of carcinogenesis by tea, Annual Review of Pharmacology and Toxicology 42(2002), 25–54.

[10]. D. L. McKay and J. B. Blumberg, The role of tea in human health: an update, Journal of the American College of Nutrition 21(2002), 1–13.

[11]. E. V. Gijo, Demand forecasting of tea by seasonal ARIMA model, Int. J. of Business Excellence 4(2011), 111 – 124.

[12]. F. S. Hamid, Yield and Quality of Tea under Varying Conditions of Soils and Nitrogen Availability, Pakistan Research Repository, Higher Education Commision Pakistan (2006). http:// eprints.hec.gov.pk/ 2348/1/2203.htm.

[13]. F. S. Swed and C. Eisenhart, Tables for Testing Randomness of Grouping in a Sequence of Alternatives, The Annals of Mathematical Statistics 14(1943), 66-87.

[14]. G. Box and G. Jenkins, Time Series Analysis: Forecasting and Control, San Francisco: Holden-Day, (1970).

[15]. H. A. Linke and R. Z. LeGeros, Black tea extract and dental caries formation in hamsters, Int. J. Food Sci. Nutr. 54(2003), 89-95.

[16]. H. Huang and X. Xu, Anticancer activity of tea: evidence from recent animal experiments and human studies, Journal of Tea Science 24(2004), 1–11.

[17]. H. Theil, Applied Economic Forecasting. North-Holland Publishing Company, Amsterdam (1966).

[18]. J. Blumberg, Introduction to the proceedings of the third International scientific symposium on tea and human health: role of flavonoids in the diet, The Journal of Nutrition 133(2003), 3244s–3246s.

[19]. J. Hazarika, Development of ARIMA Model for Tea Production of Assam, India: A Case Study, Int. J. Agricult. Stat. Sci. 6(2010), 11-18.

[20]. J. H. Weisburger, Tea and health: a historical perspective, Cancer Letters 114(1997), 315–317.

[21]. J. H. Weisburger, Tea and health: the underlying mechanisms, Proceedings of the Society for Experimental Biology and Medicine 220(1999), 271–275.

[22]. K. M. Ahammed, Investment for Sustainable Development of Bangladesh Tea Industry – An Empirical Study, BEA XVIII Biennial Conference Papers 2012, 1-20.

[23]. K. P. Borodoloi, Global tea production and export trend with special reference to India, Two and a Bud 59(2013), 152-156.

[24]. K. S. Stote, and D. J. Baer, Tea consumption may improve biomarkers of insulin sensitivity and risk factors for diabetes, J. Nutr. 138(2008), 1584S–1588S.

[25]. M. L. Rahman, Impact of Price and Other Factors on Tea in Bangladesh: Sources of Variation and Disparity over Division, BRAC University Journal 4(2007), 9-11.

[26]. M. R. Dissanayake, Time Fluctuation Models to Forecast Tea Production and Prices in Sri Lanka, 10th International Conference on Sri Lankan Studies., 16-18 December 2005.

[27]. M. Redowan and A. H. Kanan, A Study on Maximization of Land Use with Associated Crops Other than Tea and Management, Int. J. of Ecol. and Devt. 25(2013), 57-70.

[28]. M. Y. Bekhit, Levels of Essential and Non- Essential Metals in Leaves of the Tea Plant (Camellia sinensis L.) and Soils of Wushwush Farms, Ethiopia, (2006). http://etd.aau.edu.et/dspace/bitstream/123456789/307/1/Michael%20Yemane.pdf.

[29]. P. Khisa and M. Iqbal, Tea Manufacturing in Bangladesh: Problems and Prospects, 4th International Conference on Mechanical Engineering, December 26-28, Dhaka, Bangladesh, 6(2001), 85-91.

[30]. R. A. Anderson and M. M. Polansky, Tea enhances insulin activity, J. Agric. Food Chem. 50(2002), 7182-7186.

[31]. R. Dutta, T. A. Stein, M. A. Smaling, R. M. Bhagat and M. Hazarika, Modelling and quantifying tea productivity in Northeast India, Two and a Bud 59(2012), 56-63.

[32]. T. K. Mondal, A. Bhattacharya, M. Laxmikumaran and P. S. Ahuja, Recent advances of tea (Camellia sinensis) biotechnology, Plant Cell, Tissue and Organ Culture 76(2004), 195–254.

[33]. T. Nasir and M. Shamsuddhoa, Tea production, Consumption and Exports: Bangladesh Perspective, International Journal of Education Research and Technology 2(2011), 68-73.

[34]. X. X. Zheng, Y. L.Xu, S. H. Li, R. Hui, Y. J. Wu, and X. H. Huang, Effects of green tea catechins with or without caffeine on glycemic control in adults: a meta-analysis of randomized controlled trials, Am. J. Clin. Nutr. 97(2013), 750-762.

[35]. Y. Kamal and N. U. Bhuiyan, The Tea Industry of Bangladesh – A Synopsis of its Problems, Prospects and Export Potentialities, IBS Business Review 2(2004), 45-58.

[36]. Yong-xing Zhu, Herve´ Huang and You-ying Tu, A review of recent studies in China on the possible beneficial health effects of tea, International Journal of Food Science and Technology 41(2006), 333–340.

[37]. Z. Chen, Major progress in the investigation on tea and human health in 20th century, China Tea 4(2001), 8–10.

[38]. Z. Chen, Biochemical and molecular biological basis of the anticarcinogenic activity of tea polyphenolic compounds, Journal of Tea Science 23(2003), 83–93.

(1) Department of Statistics, Jahangirnagar University, Savar, Dhaka-1342, Bangladesh.
E-mail address: mmhmm.justat@gmail.com

(2) Department of Statistics, Islamic University, Kushtia-7003, Bangladesh.
E-mail address: faruqiustat09mnil@gmail.com