Asian Journal of Agricultural Extension, Economics & Sociology

38(4): 21-29, 2020; Article no.AJAEES.55859
ISSN: 2320-7027

Detection of Tea Export Potentiality from Bangladesh in the World Market: An Autoregressive Integrated Moving Average (ARIMA) Approach

A. U. Noman[1], S. Majumder[1], M. F. Imam[1], M. J. Hossain[1], F. Elahi[1*] and K. Akter[2]

[1]Department of Agricultural Statistics, Bangladesh Agricultural University, Mymensingh-2202, Bangladesh.
[2]Department of Agricultural Statistics, Sylhet Agricultural University, Sylhet, Bangladesh.

Authors' contributions

This work was carried out in collaboration among all authors. Author AUN designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Author SM managed the analyses, supervised and edited the work. Authors MFI and MJH reviewed the analysis and all drafts of the manuscript. Authors FE and KA managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJAEES/2020/v38i430333
Editor(s):
(1) Dr. Jurislav Babić, University of Osijek, Croatia.
Reviewers:
(1) Peter Yacob, Universiti Tunku Abdul Rahman, Malaysia.
(2) Venkata Sanyasi Seshendra Kumar Karri, GITAM University, India.
Complete Peer review History: http://www.sdiarticle4.com/review-history/55859

Original Research Article

Received 16 February 2020
Accepted 22 April 2020
Published 02 May 2020

ABSTRACT

Export plays an important role in promoting economic growth and development. The study is conducted to make an efficient forecasting of tea export from Bangladesh for mitigating the risk of export in the world market. Forecasting has been done by fitting Box-Jenkins type autoregressive integrated moving average (ARIMA) model. The best ARIMA model is selected by comparing the criteria- coefficient of determination (R^2), root mean square error (RMSE), mean absolute percentage error (MAPE), mean absolute error (MAE) and Bayesian information criteria (BIC). Among the Box-Jenkins ARIMA type models for tea export the ARIMA (1,1,3) model is the most appropriate one for forecasting and the forecast values in thousand kilogram for the year 2017-18, 2018-19, 2019-20, 2020-21 and 2021-22, are 1096.48, 812.83, 1122.02, 776.25 and 794.33 with

**Corresponding author: E-mail: fazle.elahi1111@gmail.com;*

upper limit 1819.70, 1348.96, 1862.09, 1288.25, 1318.26 and lower limit 660.69, 489.78, 676.08, 467.74, 478.63, respectively. So, the result of this model may be helpful for the policymaker to make an export development plan for the country.

Keywords: Time series; ARIMA; stationary; forecasting.

1. INTRODUCTION

Bangladesh is a small country surrounded by hilly areas on it's three sides. Tea is an intensively managed perennial monoculture crop cultivated on large and small-scale plantation in the hilly areas. It is made from the twigs containing tender leaf and bud of the tea plant, a member of Camellia genus, is characterized as *Camellia sinensis* and *Camellia assamica* in botanical terminology [1].

Tea is the second largest export oriented cash crop of Bangladesh, following jute [2]. It has been one of the major exportable items of Bangladesh since 1971 [3]. It had a captive export market in some countries. Leading importers of Bangladeshi tea are: Pakistan, Kazakistan, Afghanistan, United Arab Emirates and Saudi Arabia. In 2016, China, Japan and K.S.A. also import tea from Bangladesh. It plays a very vital role in the national economy through export earning, trade balancing as well as in employment generation. It produces 2% of world production and exports 0.07% of world export [4]. The industry offers rural employment opportunities especially to women and contributes to poverty alleviation in the rural areas. About 0.35 million people including 0.09 million permanent workers and 0.21 million dependents live on the Bangladesh tea industry. The industry accounts for 1% of national GDP in the economy of Bangladesh [5].

Bangladesh exported tea to eight countries over the years among which Pakistan was the top listed country during 2016-17. UAE, K.S.A, Kuwait, USA, China and Japan possessed the subsequent six positions respectively of the importers of tea from Bangladesh in a descending order in 2016-17. Pakistan was the main importer of tea and Bangladesh earned 3424.52 thousand US $ which was 81.55% of total export earnings from tea in 2016-17. The earnings of Bangladesh in thousand US $ from the subsequent six leading importers in the same year were 575.58, 137.56, 38.01, 21.72, 1.81 and 0.78, constituting 13.71%, 3.27%, 0.91%, 0.52%, 0.04% and 0.02% respectively of the total earnings from tea [6].

Nine exporters perform export performance of Bangladeshi tea. The exporter is AbulKhair, Consol, Monir Shah, Ispahani, Shaw Wallace, Kazi & Kazi, Meghna, Haji Ahmed and Halda Valley. Among nine exporters Consolwas the main exporter of tea and earned 1898.69 thousand US $ which was 45.22% of total export earnings from tea in 2016-17. Meghna, AbulKhair, Haji Ahmed, Ispahani, Monir Shah, Shaw Wallace, Halda Valley and Kazi&Kazi possessed the subsequent eight positions respectively of the exporters of Bangladeshi tea in a descending order in 2016-17. The earnings in thousand US $ of the exporters from the tea export of Bangladesh in the same year were 785.54, 597.3, 416.3, 367.43, 115.84, 16.29, 1.81 and 0.78, constituting 18.71%, 14.22%, 9.91%, 8.75%, 2.76%, 0.39%, 0.04% and 0.02%, respectively of the total earnings from tea [6].

2. MATERIALS AND METHODS

The yearly data of the export (Thousand Kilogram) quantity of tea for the year 1990-91 to 2016-17 have been collected from the Bangladesh Tea Board (BTB). These Secondary data are used to analyze and achieve the objective of the study.

A very popular process in econometric time series is the autoregressive integrated moving average (ARIMA) process. Time series models are based on the assumption that it is stationary. But many of the econometric time series are non-stationary, that is, integrated. If a time series is integrated, ARIMA models a non-stationary time series is made stationary by applying finite differencing of the data points. The most popular type ARIMA process of order p, d and q denoted as ARIMA (p, d, q), may be defined as follows [7]:

$$\emptyset(B)(\nabla^d Y_t - \mu) = \theta_q(B)\, \varepsilon_t$$

where,

Y_t = Tea export at time t
μ = The mean of $\nabla^d Y_t$
$\emptyset_p(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p$
$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$
$\emptyset_i$ = The i^{th} autoregressive parameter; i =1, 2, …, p

θ_i = The i^{th} moving average parameter; i =1, 2, …, q
p = The autoregressive order
q = The moving average order
d = The times that the series is differenced
∇ = The difference operator and
B = The back shift operator

The estimation methodology for ARIMA, Box-Jenkins uses a three-step iterative approach of model identification, parameter estimation and diagnostic checking to determine the best model from a general class of ARIMA models. The very first step of Box-Jenkins methodology for selecting ARIMA model is to test the stationary of a series and to determine the true degree of differencing d in order to make the series stationary. Having found the value of the second parameter d of the ARIMA model, the next step is to identify the appropriate autoregressive order p and moving average order q. This step can be performed by studying the autocorrelation function (ACF) & partial autocorrelation function (PACF) and Augmented Dickey Fuller (ADF) & Phillips-Perron (P-P) unit root test. In second stage, the parameters of the model are estimated using the software SPSS 20. In order to select the ARIMA model of the best fit for forecasting the latest available model selection criteria R^2, RMSE, MAPE, MAE and BIC are used. The third step examined the residual of ACF and PACF with Ljung-Box statistics to check whether the chosen model fits the data reasonably well.

2.1 Coefficient of Determination (R^2)

The coefficient of determination (R^2) proposed by Theil [8] is the ratio of the regression sum of squares to the total sum of squares i. e.

$$R^2 = \frac{\text{Regression sum of squares}}{\text{Total sum of squares}}$$

It is generally considered that the more the value of R^2, the better is the fit.

2.2 Root Mean Square Error (RMSE)

The root mean square error is defined as

$$\text{RMSE} = \sqrt{\left(\frac{1}{n-k}\sum_{t=1}^{n}\varepsilon_t^2\right)}$$

where, n is the sample size and k is the total number of estimable parameters. The model with minimum RMSE is assumed to describe the data series more adequately.

2.3 Mean Absolute Error (MAE)

Mean absolute error is the average of the absolute deviations. An absolute deviation is the difference between the observed and estimated values. Symbolically,

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|\varepsilon_t|$$

where, ε_t is the difference between the observed and estimated values. The model with minimum MAE is assumed to describe the data series more adequately.

2.4 Mean Absolute Percentage Error (MAPE)

Mean absolute percentage error is a measure of accuracy in a fitted time series values in statistics, specifically trending. It usually expresses accuracy as a percentage, and is defined by the formula:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|\varepsilon_t|}{y_t}\times 100$$

where, ε_t is the difference between the observed and estimated values, Y_t y_t and n stand for observed value and fitted points, respectively. The model with minimum MAPE is assumed to describe the data series adequately.

2.5 Bayesian Information Criterion (BIC)

The Bayesian information criterion (BIC) is a criterion for model selection among a finite set of models. The BIC was developed by Gideon E. Schwarz in [9]. The formula for the BIC is-

$$\text{BIC} = n\log(\text{MSE}) + k\log n$$

where, n is the sample size, MSE is the mean square error and k is the total number of observation. The model with minimum BIC is assumed to describe the data series more adequately.

3. RESULTS AND DISCUSSION

3.1 Logarithmic Transformation of the Tea Export Series

In Fig. 1(a) for the original series of tea export time series data posted on the study. It can be clearly seen from the Fig. 1(a) that the fluctuation increases as time increases. This variation in the magnitude of the fluctuation with time is termed as non-stationary in variance of the data series. The main approach for achieving stationary in variance is through a logarithmic transformation of the data [10]. Fig. 1(b) shows the logarithmic transformation of the original series of tea export. It is clear that the magnitude of the fluctuations in the log transformed series does not vary so much with time. Thus the logarithmic transformation makes the series stationary in variance.

3.2 Test of Stationary Using ACF and PACF

The correlogram of the tea export time series plotted using the tea export time series log transformed data posted on the study. The correlogram up to 16 lags is shown in Table 1(a) and Fig. 2(a) & (b). The Table 2(a) shows that autocorrelation and partial autocorrelation coefficients starts with a very high value at lag 1(0.864) and 1(0.864) and decline very slowly. The Fig. 2(a) shows that maximum spikes of the autocorrelation is out of the boundary and Fig. 2(b) shows that the first spike in the partial autocorrelation is out of the boundary. Thus it seems that the log transformed tea export time series data is non-stationary in mean.

For making stationary we take the first difference of the tea export log transformed data. The first difference correlogram up to 16 lags is shown in Table 1 and Fig. 2(c) & (d). The Table 1 shows

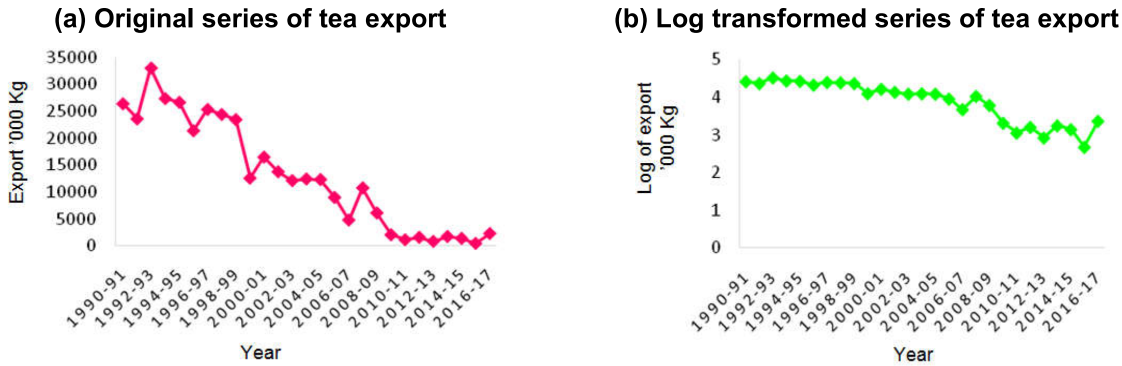

Fig. 1. Original series of tea export from Bangladesh (a) and log transformed series of tea export (b) over 1990-91 to 2016-17

Table 1. Autocorrelation and partial autocorrelation of log transformed value of tea export (a) first difference of log transformed value of tea export from Bangladesh (b) over 1990-91 to 2016-17

(a) Log transformed value of tea export from Bangladesh					(b) First difference of log transformed value of tea export from Bangladesh				
Leg	**ACF**	**PACF**	**Box-Ljung Statistic**	**Prob.**	**Leg**	**ACF**	**PACF**	**Box-Ljung Statistic**	**Prob.**
1	.864	.864	22.481	.000	1	-.326	-.326	3.097	.078
2	.742	-.019	39.713	.000	2	-.166	-.305	3.932	.140
3	.665	.114	54.162	.000	3	.255	.068	5.533	.137
4	.550	-.188	64.466	.000	4	-.076	-.011	5.725	.221
5	.421	-.110	70.789	.000	5	.001	.048	5.725	.334
6	.311	-.066	74.386	.000	6	.118	.119	6.229	.398
7	.154	-.271	75.317	.000	7	-.146	-.060	7.049	.424
8	.037	.050	75.374	.000	8	-.265	-.391	9.890	.273
9	-.001	.181	75.374	.000	9	.230	-.123	12.151	.205
10	-.052	.001	75.496	.000	10	-.044	-.089	12.239	.269
11	-.129	-.110	76.314	.000	11	.026	.188	12.271	.344
12	-.192	-.147	78.248	.000	12	-.092	-.045	12.711	.390
13	-.227	.001	81.139	.000	13	.077	.142	13.041	.445
14	-.281	-.175	85.880	.000	14	-.097	-.166	13.606	.479
15	-.318	-.068	92.484	.000	15	-.015	-.205	13.620	.554
16	-.348	-.010	101.101	.000	16	.099	-.195	14.335	.574

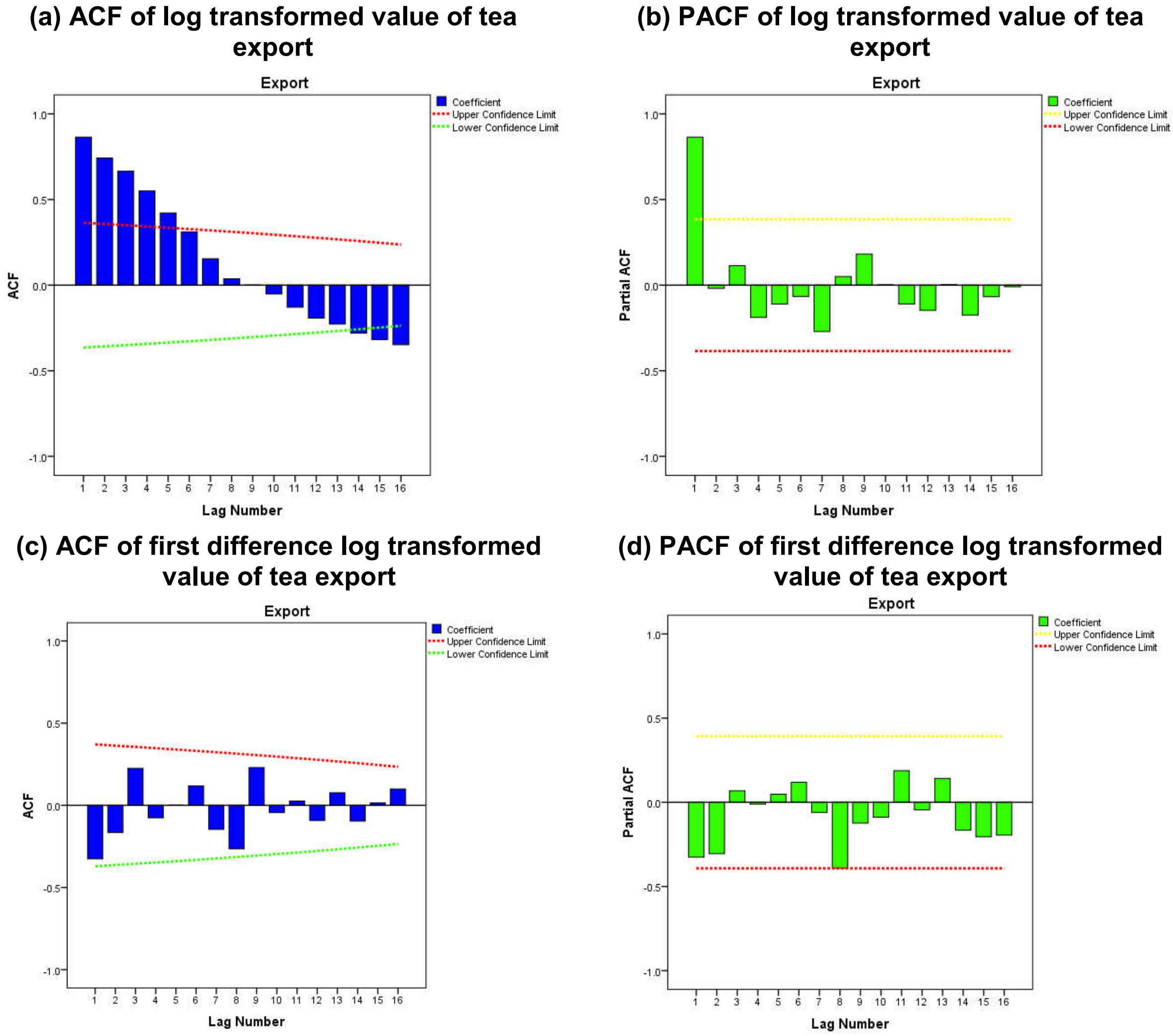

Fig. 2. Autocorrelation and partial autocorrelation of log transformed value of tea export (a & b) and first difference of log transformed value of tea export (c & d) over 1990-91to 2016-17

that autocorrelation and partial autocorrelation coefficients start with a very low value at lag 1(-0.326) and 1(-0.326). The Fig. 2(c) shows that all the spikes of the autocorrelation lie in the boundary and Fig. 2(d) shows that all the spikes in the partial autocorrelation also lie in the boundary. Thus it seems that the log transformed tea export time series data is stationary in mean.

3.3 Test of Stationary Using Augmented Dickey Fuller (ADF) & Phillips-Perron (P-P) Unit Root Test

To apply standard statistical inference problem in a dynamic time series model, variables included should have to be stationary because economic theory so far developed is based on stationary assumption. Thus unit root test in a priori is important to avoid spuriosity in the result of the ultimate analysis. Apart from the graphical method of using ACF and PACF for determining stationary of the time series, a very popular method of determining stationary is the Augmented Dickey Fuller [11] and Phillips-Perron [12] test. In the present study, the Augmented Dickey Fuller (ADF) and Phillips-Perron (P-P) test are done for the test of stationary.

The computed ADF test-statistic is -1.160 with p-value 0.691 > 0.05 indicating the acceptance of the null hypothesis. That means log transformed value of tea export series has a unit root problem and the series is a non-stationary series.

The computed P-Ptest statistic is -0.923 with p-value 0.780 > 0.05 indicating the acceptance of the null hypothesis. That means log transformed value of tea export series has a unit root problem and the series is a non-stationary series.

Table 2. ADF and P-P tests for log transformed value of tea export series

Null Hypothesis: Log transformed value of tea export has a unit root Lag Length: 16		
Test	**Test statistic**	**p-value**
ADF-test	-1.160	0.691
P-P test	-0.923	0.780

Table 3. ADF and P-P test for first difference log transformed value of tea export series

Null Hypothesis: Log transformed value of tea export has a unit root Lag Length: 16		
Test	**Test statistic**	**p-value**
ADF-test	-3.970	0.001
P-P test	-3.883	0.002

The computed ADF test-statistic is -3.970 with p-value 0.001< 0.01 indicating the rejection of the null hypothesis. That means log transformed value of tea export series has no unit root problem and the series is a stationary series.

The computed P-P test statistic is -3.883 with p-value 0.002< 0.01 indicating the rejection of the null hypothesis. That means log transformed value of tea export series has no unit root problem and the series is a stationary series.

3.4 Determination of the Degree of Differencing

For this first difference of the log transformed data series and their ACF and PACF plots plotted against the lag length are considered. Fig. 2(c) & (d) represent the ACF and PACF plots of the data series at log transformed first difference. No trend in this series is observed, perhaps suggesting that the first difference of the time series is stationary. A formal application of the Augmented Dickey Fuller (ADF) and Phillips-Perron (P-P) unit root tests (Table 3) show the same result. So, degree of differencing d = 1 is selected for the final model.

3.5 Determination of the Autoregressive Order

After determining the degree of differencing, the next step remains is to determine the appropriate ARIMA level of stationary series. Examining both ACF and PACF the order of auto-regression is determined. Fig. 2(a) & (b) represent the ACF and PACF plots of the log transformed data series. It is evident that the ACF plots show exponentially declining values when the PACF has only one significant spike. This suggest that the maximum order of auto-regression should be considered to be 1. So, ARIMA models can be tested with the order of p = 0 and 1.

3.6 Determination of the Moving Average Order

In case of MA process, the order of q exhibits reverse property to that of AR process of order p. This means MA (q) models have precisely q spikes in the first q values of the ACF and exponentially declining values of PACF. Fig. 2(a) & (b) indicates that the ACF has 5 significant spikes in first 5 values and the PACF shows sharply declining values. Thus, the ARIMA models can be tested with the order of q = 0, 1, 2, 3, 4 and 5, respectively.

3.7 Tentative ARIMA Specification

Twelve ARIMA models with tentatively selected values of p, d and q are tried. These are ARIMA (0,1,0), ARIMA (0,1,1), ARIMA (0,1,2), ARIMA (0,1,3), ARIMA (0,1,4), ARIMA(0,1,5), ARIMA (1,1,0), ARIMA (1,1,1), ARIMA (1,1,2), ARIMA (1,1,3), ARIMA (1,1,4) and ARIMA (1,1,5).

3.8 Final Model Selection

Table 4 shows the estimated values of the criteria for computing ARIMA models. It is evident from the Table 4 that the values of RMSE, MAPE and MAE for ARIMA (1,1,3) are minimum than that of other models. ARIMA (1,1,5) is better only for the criterion R^2 and ARIMA (0,1,1) is better only for the criterion BIC. So, ARIMA (1,1,3) can be considered as the best model for forecasting exported quantity of tea from Bangladesh.

3.9 Best Estimated Model for Log Transformed Tea Export

The above discussion about the fitness of various models to the time series of log transformed value of tea export in Bangladesh reveals that ARIMA (1,1,3) is finally chosen for forecast. The best model with estimated parameters are given bellow-

$$(1 - .548B)(\Delta Y_t - .050) = (1 - .114B + .476B^2 - .404B^3)\ \varepsilon_t$$

3.10 Diagnostic Checking

For diagnostic checking ACF and PACF of residuals and Ljung-Box Q statistic are widely used in practice. In Table 5, the Ljung-Box Q statistic is given for the best selected stochastic model with p-value. The Ljung-Box Q value is insignificant. It implies that the residuals of ACF and PACF of ARIMA (1,1,3), for the respected log transformed time series of tea export are white noise implying all the spikes lie in the boundary (Fig. 3). So, the model's fitness is acceptable.

3.11 Forecasting Time Series Log Transformed Value of Tea Export

An important limitation of making forecasts is that forecasting error increases as the period of forecasts increases. For this reason, short-term forecasts are more reliable than long term forecasts. So, five-year forecasts are used to make with 95% confidence interval of log transformed value of tea export estimated by using the best selected model (Table 6). The prediction period is extended from 2017-18 to 2021-22.

Table 4. Estimated values of model selection criteria of log transformed value of tea export series

Model	Values of selection criteria					Ljung –Box Q test	
	R^2	RMSE	MAPE	MAE	BIC	Statistics	Sig.
ARIMA(0,1,0)	.794	.251	5.087	.177	-2.639	15.567	.623
ARIMA(0,1,1)	.841	.225	4.608	.160	**-2.730**	9.895	.908
ARIMA(0,1,2)	.842	.229	4.556	.158	-2.572	9.330	.899
ARIMA(0,1,3)	.853	.226	4.424	.155	-2.475	9.184	.868
ARIMA(0,1,4)	.863	.225	4.487	.160	-2.369	10.039	.759
ARIMA(0,1,5)	.863	.229	4.490	.161	-2.195	10.103	.685
ARIMA(1,1,0)	.828	.234	4.656	.160	-2.655	11.024	.855
ARIMA(1,1,1)	.841	.230	4.579	.159	-2.567	9.569	.888
ARIMA(1,1,2)	.837	.238	4.502	.155	-2.371	10.400	.794
ARIMA(1,1,3)	.863	**.224**	**4.320**	**.153**	-2.370	**8.446**	**.865**
ARIMA(1,1,4)	.862	.229	4.329	.154	-2.193	8.558	.805
ARIMA(1,1,5)	**.875**	.225	4.445	.156	-2.112	8.988	.704

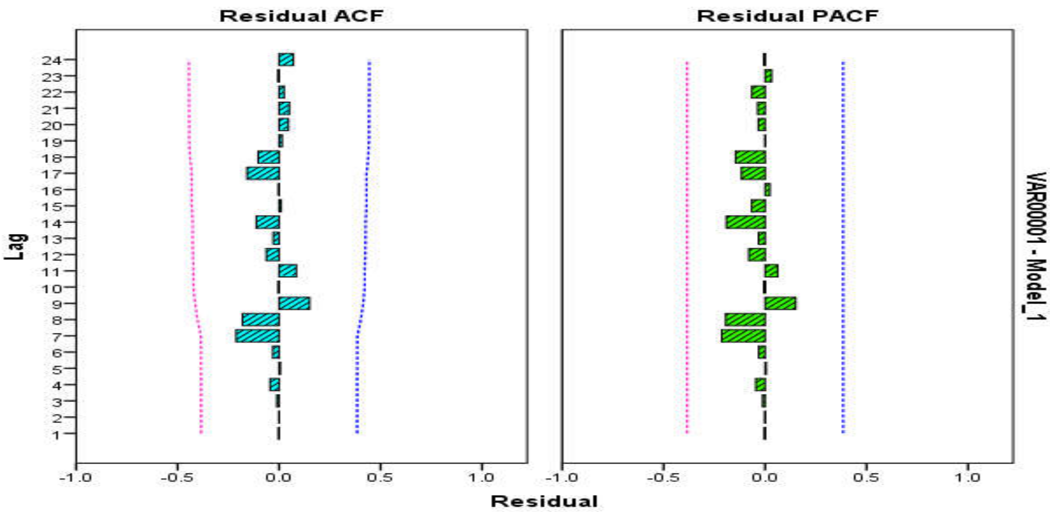

Fig. 3. Residuals of ACF and PACF for ARIMA (1,1,3) model

Table 5. Diagnostic tools and model selection criteria for the best fitted model of tea export

Model	Values of selection criteria					Ljung –Box Q test	
	R^2	RMSE	MAPE	MAE	BIC	Statistics	Sig.
ARIMA(1,1,3)	Not satisfied	.224	4.320	.153	Not satisfied	8.446	.865

Table 6. Forecasting log transformed value of tea export series from Bangladesh for the period 2017-18 to 2021-22

Forecast year	ARIMA (1,1,3)					
	Forecast value		UPL		LPL	
	Log value	Actual value ('000 Kg)	Log value	Actual value ('000 Kg)	Log value	Actual value ('000 Kg)
2017-18	3.04	1096.48	3.26	1819.70	2.82	660.69
2018-19	2.91	812.83	3.13	1348.96	2.69	489.78
2019-20	3.05	1122.02	3.27	1862.09	2.83	676.08
2020-21	2.89	776.25	3.11	1288.25	2.67	467.74
2021-22	2.9	794.33	3.12	1318.26	2.68	478.63

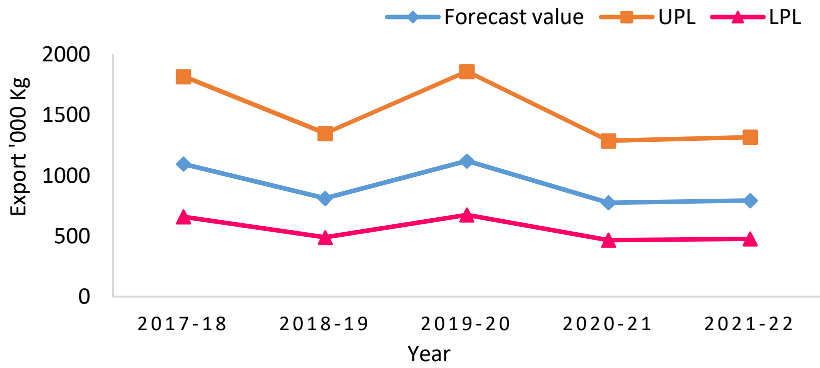

Fig. 4. Forecasting the tea export from Bangladesh by ARIMA (1,1,3) model

4. CONCLUSION

A time series model accounts for patterns in the past movement of a variable and uses that information for predicting its future movements. In a sense a time series model is just a sophisticated model for extrapolation. Time series data and models have become very popular to be intensively used in empirical research and econometrics. The ARIMA model offers a good technique for predicting magnitude of these variables. The forecasting results show fluctuate pattern of tea export over the forecasted period. The forecasting of tea export can help the policy makers for future planning to increase export.

ACKNOWLEDGEMENTS

The authors would like to express their heartiest gratitude to Mohammad Abdul Wazed, Director, Bangladesh Bureau of Statistics (BBS), Manuja Peiris, Chief Executive, International Tea Committee, (ITC), United Kingdom, Kaison Chang, Secretary, FAO Intergovernmental Group on Tea, Major General Jahangir Al Mustahidur Rahman, Chairman Bangladesh Tea Board (BTB), Baizid Bostami Road, Nasirabad, Chattogram for their kind cooperation during the period of study. They are also very much grateful to the Ministry of Science and Technology for providing financial support to collect data for conducting the research work through the award of "National Science and Technology Fellowship 2017-2018".

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Alam AFMB. Profile of tea industry in Bangladesh. Global Advances in Tea Science, New Delhi: Aravali Books. 1999; 882.

2. Hossain MM, Abdulla F. Jute production in Bangladesh: A time series analysis. Journal of Mathematics and Statistics. 2015;11(3):93-98.
3. Sabur SA, Molla MMU, Azad MAK. Tea industry in Bangladesh: Marketing system and price behaviour. Bangladesh Journal of Agricultural Economics. 2000;23(454-2016-36391):77-90.
4. Golder PC, Sastry RK, Srinivas K. Research priorities in Bangladesh: Analysis of crop production trends. SAARC Journal of Agriculture. 2013;11(1):53-70.
5. BTB. Strategic plan for Bangladesh tea industry. Bangladesh Tea Board, Baizid Bostami Road, Chittagong. 2016;2-5.
6. BTB. Statistical bulletin of Bangladesh tea board for the month of October. Baizid Bostami Road, Chittagong. 2017;9-11.
7. Box G, Jenkins G. Time series analysis: Forecasting and control. San Francisco: Holden-Day; 1976.
8. Theil H. Economic forecasts and policy. Second Edition, North-Holland, Amsterdam. 1961;216-221.
9. Schwarz G. Estimating dimension of a model. Annals of Statistics. 1978;6:461-464.
10. Makridakis S, Wheelwright SC, Hyndman RJ. Forecasting methods and applications. 3rd Edition, John Willey and Sons. Inc, USA; 1998.
11. Dickey DA, Fuller WA. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica. 1981; 49(4):1057-1072.
12. Phillips PCB, Perron P. Testing for a unit root in time series regression. Biometrika. 1988;65:335-346.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/55859