International Journal of Mathematics and Computation,
ISSN 0974-570X (Online), ISSN 0974-5718 (Print),
Vol. 30; Issue No. 2; Year 2019,

Forecasting Rice Production in Jessore, Dinajpur and Kushtia Districts of Bangladesh by Time Series Model

M. A. Rahman[1] and **M. M. Hossain**[2]

[1]Department of Statistics, Jahangirnagar University,
Savar, Dhaka-1342, Bangladesh;
Email :

[2]Department of Statistics, Jahangirnagar University,
Savra, Dhaka-1342, Bangladesh);
Email : mmhmm.justat@gmail.com

ABSTRACT

The economy of Bangladesh mainly depends on agriculture in which rice is the leading crop. Rice is the principal food, imitated in the high per capita rice consumption in this country. Most of the people in Bangladesh fulfill the nutritional demand by rice. Over its long history, rice production of Bangladesh has gradually changed in terms of yield potentials, cultivations techniques, and cropping patterns. Despite the pressure from over-population, the country has reached self-sufficiency in rice production. Thus, this paper attempts to identify the appropriate ARIMA (Autoregressive Integrated Moving Average) model that is used to forecast the production of rice in Jessore, Dinajpur and Kushtia districts. In this paper, $ARIMA(1,2,1)$ *model for both Jessore and Dinajpur and* $ARIMA(1,2,2)$ *model for Kushtia is found to be suitable for forecasting the rice production also the test result indicating that the errors of the selected models are not auto-correlated as well as follows the normal distribution. Finally, these models are used to forecast the rice production of the selected districts for the upcoming 20 years which help the decision makers to establish the rice production management.*

Keywords: Rice production, Box- Jenkins methodology, Ljung- Box, Jarque-Berra Test .

Mathematics Subject Classification: 62P10

Journal of Economic Literature (JEL) Classification : Q12

1. INTRODUCTION

Bangladesh has a long history of rice cultivation and stands fourth rice production in the world. It is grown all over the country except the hilly areas. The agro-climate conditions of the country are suitable for growing rice year-round. The production of rice in Bangladesh depends on the seasonal changes and water supply. Among the verities of paddy, Aman is harvested more widely in November and December and accumulating more than half of the annual production. According to the amount of harvested Aus has the second position, including traditional strains but more often with high-yielding varieties. Rice for the Aus harvest is sown March or April, benefits from April and May rains, matures during in the summer rain, and its harvested during.

By familiarizing and using the high yielding seed varieties, fertilizers, pesticides, irrigation, and mechanized cultivation system in Bangladesh, rice production has been increased at a significant rate.

Rice production has increased by 23.7% in fiscal year 2012-2013compared to the fiscal year 2006-2007 (BBS, 2010). However, the growth rate of rice production is lower than the demand rate for rice in the country. Thus, to satisfy the additional demand the country has to import rice almost every year in the previous decades (Nargis and Lee, 2013). Moreover, an increase in total rice production is required to feed this ever-increasing population. On the other hand, simultaneously the total cultivable land is decreasing at a rate of more than one percent annually due to the construction of industries, factories, houses, roads, and highways. In addition, due to rapid urbanization, food habits of the people tend to change which demand the cultivation of new crops that must share land used for rice cultivation. Therefore, it is necessary to initiate the attempts for increasing the yield of rice per unit area (BBS, 2012).

About two-third of the people lives in rural areas, and near about half of the total manpower is involved in agricultural activities in Bangladesh. Increasing food production and attaining food security in Bangladesh require sustainable growth of the agricultural sector. In Bangladesh, agricultural contributes 22.7% of Gross Domestic Product (GDP) of the country (Bangladesh Finance Bureau, 2016). Therefore, it is necessary to identify the amount of rice production which assist to make the food policy in Bangladesh. Thus, the main objective of this paper is to forecast the rice production in three districts namely Jessore, Dinajpur and Kushtia whose production has a significant effect in total rice production of Bangladesh.

2. METHODS AND MATERIALS

2.1. Sources of data

For this study, the data of the rice production for Jessore, Dinajpur and Kushtia districts are collected from the different Statistical Year Books Bangladesh from 1971 to 2015.

2.2. ARIMA model

If $\{Z_t\}$ is a white noise with mean zero and variance σ^2 and $\{Y_t\}$ is defined as $Y_t = Z_t + \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + \ldots + \beta_q Z_{t-q}$ is said to be a moving average process having order q and is symbolized by $\mathrm{MA}(q)$. Also, the process $\{Y_t\}$ is given by $Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ldots + \alpha_p Y_{t-p} + Z_t$ is called an auto-regressive process of order p and is represented by $\mathrm{AR}(p)$. The combined model of AR and MA model is known as ARMA models. An $\mathrm{ARMA}(p,q)$ model can be written as $Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ldots + \alpha_p Y_{t-p} + Z_t + \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + \ldots + \beta_q Z_{t-q}$, where, Y_t is the original series, for every t, and assume that Z_t is independent of $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$. 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.

2.3. Ljung-Box test

Ljung-Box (Ljung and Box, 1978) test is used to detect presence of auto-correlation among the residuals. In order to check the auto-correlation, the null hypothesis $H_0 : \rho_1(e) = \rho_2(e) = \ldots = \rho_k(e) = 0$

is tested by the Ljung-Box statistic, $Q^{*} = N(N+1)\sum_{i=1}^{k}(N-k)\rho_k^2(e)$, which is approximately chi-square distribution with $(k-q)$ degrees of freedom, where, N is the number of observations, q is the number of estimated parameters of the model.

.

2.4. Jarque-Bera test

The normality assumption is checked by using Jarque and Bera (1987) test and is mainly based on the sample kurtosis (k) and skewness (s). The Jarque-Bera (JB) test statistics can be written as $JB = \frac{n}{6}\left(s^2 + \frac{(k-2)^2}{4}\right)$, where n is the number of observations and k is the sample kurtosis an s is the sample skewness. The JB statistic has an asymptotic chi-square distribution with 2 degrees of freedom.

2.5. Ljung-Box test

In order to evaluate the accuracy of the fitted model researchers use different model selection criteria. Here, we consider the followings:

Akaike (1974) suggested measuring the goodness of fit of a model. He defined the criterion as $AIC = \log\hat{\sigma}_k^2 + \frac{n+2k}{n}$, where, $\hat{\sigma}_k^2 = \frac{SSE_k}{n}$, SSE_k denotes the residual sum of squares under the model with k number of parameters and n is the number of observations. However, a corrected form of AIC is suggested by Sugiura (1978) and expanded by Hurvich and Tasi (1989) is defined as follows: $AIC_C = \log\hat{\sigma}_k^2 + \frac{n+k}{n-k-2}$. The minimum AICC specifies the best model.

3. RESULTS

Before model fitting, we first need to check the stationarity of the data as well as have to check what type of variations is occurred in the data set. We get the following Time Series plots of original dataset and differenced data set of rice production for Jessore, Dinajpur and Kushtia districts respectively and presented in Figure 1.

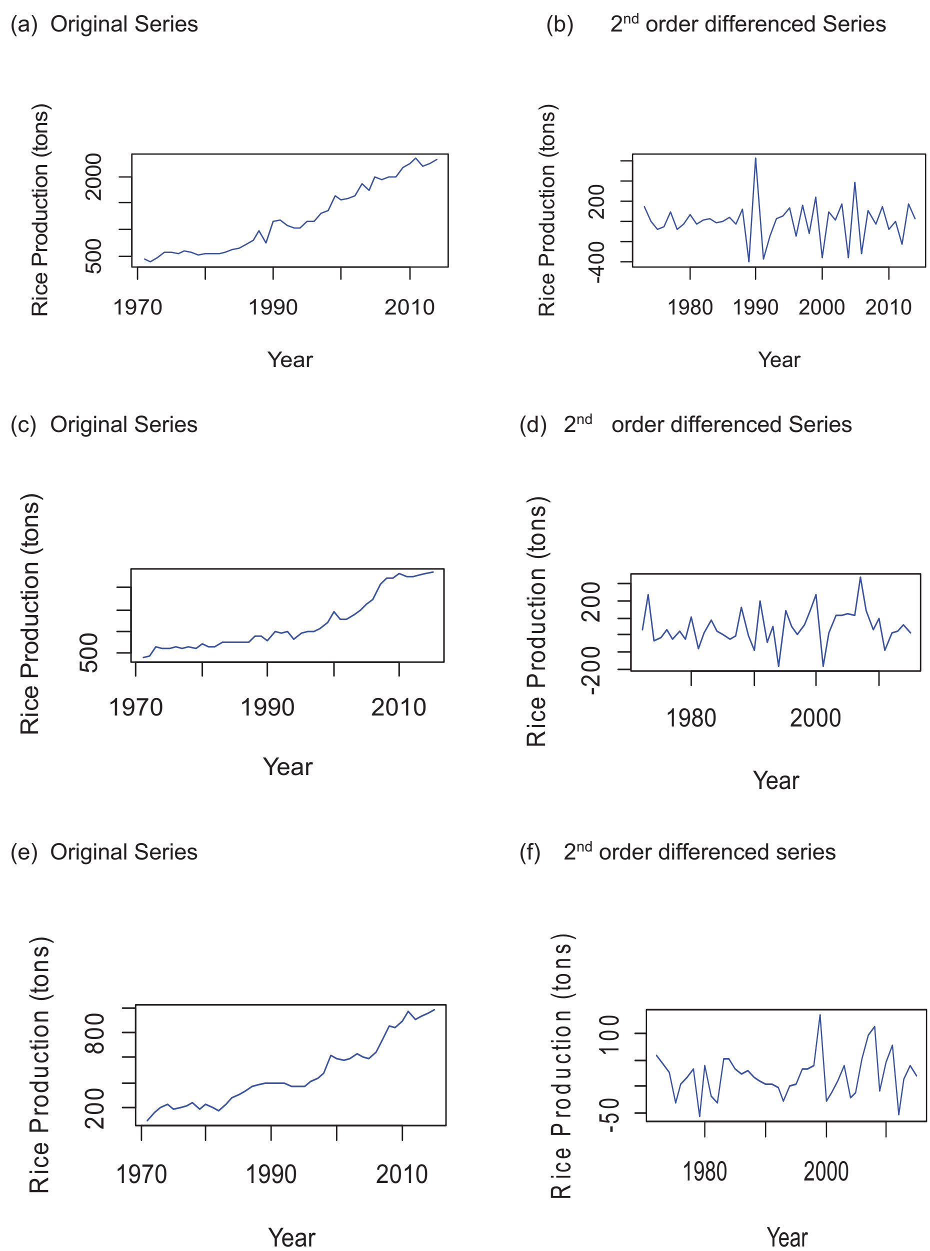


Figure 1. (a) Time series (original series) plot, (b) Time series (2nd differenced) plot, (c) Time series (original series) plot, (d) Time series (2nd differenced) plot (e) Time series (original series) plot and (f) Time series (2nd differenced) plot of rice production for Jessore, Dinajpur and Kushtia districts respectively.

From the Time Series plot given in (Figure 1 (a), (c) and (e)), it can be seen that there is an increasing trend in the rice production in Jessore, Dinajpur, and Kushtia districts. However, the random fluctuations in the rice production here are not constant in the amount over time. So, all the Time

Series data i.e., the rice production for Jessore, Dinajpur, and Kushtia districts are non-stationary. But the second difference of rice production in Jessore, Dinajpur, and Kushtia districts are stationary (Figure (b), (d) and (f)). Besides the graphical test, we also use ADF and KPSS test to check the stationarity problem. The test results are given in Table 1.

Table 1: Results of unit root test

Test	Jessore		Dinajpur		Kushtia	
	Original data	2nd differenced data	Original data	2nd differenced data	Original data	2nd differenced data
P value of ADF	0.4743	0.01	0.7891	0.01	0.4874	0.01501
P value of KPSS	0.01	0.10	0.01	0.10	0.01	0.10

The p-value of Augmented Dickey-Fuller (ADF) test for the original data set of Jessore, Dinajpur and Kushtia are 0.4743, 0.7891 and 0.4874 respectively, which are greater than 0.05. So, the null hypothesis i.e., the rice production time series data is non-stationary, may not be rejected at 5% level of significance. Also, all the p-value of this test for the 2nd difference data of the three districts is smaller than 0.05. Thus, the null hypothesis is rejected. Hence, we may conclude that the 2nd difference rice production time series data of Jessore, Dinajpur, and Kushtia are stationary. A similar decision is made if we consider the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for all districts considered in this study (Table 1). Now, we estimate the parameters of different models and the most appropriate model is selected with the help of AIC_C. The results are presented in the following table (Table 2).

Table 2: Model Estimation and Selection

Jessore		Dinajpur		Kushtia	
Model	AICc	Model	AICc	Model	AICc
ARIMA (0,2,0)	714.2557	ARIMA (0,2,0)	714.2557	ARIMA (0,2,0)	660.9741
ARIMA (0,2,1)	672.7382	ARIMA (0,2,1)	672.7382	ARIMA (0,2,1)	625.7238
ARIMA (0,2,2)	638.3642	ARIMA (0,2,2)	638.3642	ARIMA (0,2,2)	611.8534
ARIMA (1,2,1)	**569.1909**	**ARIMA (1,2,1)**	**569.1909**	ARIMA (1,2,0)	606.3256
ARIMA (1,2,0)	572.4249	ARIMA (1,2,0)	572.4249	ARIMA (1,2,1)	603.6418
ARIMA (1,2,2)	570.1332	ARIMA (1,2,2)	570.1332	**ARIMA (1,2,2)**	**601.2548**

According to the value of AIC_C, we observed that the ARIMA (1,2,1), ARIMA (1,2,1) and ARIMA (1,2,2) models are more appropriate for forecasting the rice production in Jessore, Dinajpur and

Kushtia districts respectively. The following table (Table 3) gives the maximum likelihood estimates and their standard errors for the selected models.

Table 3: Parameter of the Selected ARIMA Model

	Coefficient	Estimate	Standard Error
Jessore	ar1	-0.4074	0.1417
	ma1	-0.9795	0.1887
	Intercept	2173.0621	983.3706
Dinajpur	Coefficient	Estimate	Standard Error
	ar1	0.9922	0.0104
	ma1	0.4034	0.2604
	Intercept	1328.4527	826.0853
Kushtia	Coefficient	Estimate	Standard Error
	ar1	1.4527	0.5203
	ma1	-0.2292	0.4642
	ma2	-0.7707	0.4566
	Intercept	541.7423	186.3585

Here, we check the normality assumption and autocorrelation problem of the residuals of fitted models. The normality is checked by the well-known Jarque-Bera (JB) test. Moreover, the autocorrelation is examined by the Ljung-Box test. The test results are given in Table 4.

Table 4: Results of Jarque-Bera and Ljung-Box test

Test Statistic	Jessore		Dinajpur		Kushtia	
	Statistic value	P value	Statistic value	P value	Statistic value	P value
Jarque-Bera	1.0135	0.0	1.0666	0.02653	3.7611	0.02039
Ljung-Box	5.585	0.01811	14.913	0.02462	10.251	0.0

At 5% level of significance, all the p-value of Jarque-Berra test is less than 0.05. So, we may reject the null hypothesis i.e., it can be concluded that the residuals of the fitted ARIMA models are normally distributed. Also, the p-value for the Ljung-Box test of all models is smaller than 0.05 indicating that there is no significant evidence of autocorrelations among the residuals of the fitted models at 5% level of significance. The following table presents the forecasted rice production (in Thousand Metric tons) of Jessore, Dinajpur and Kushtia districts.

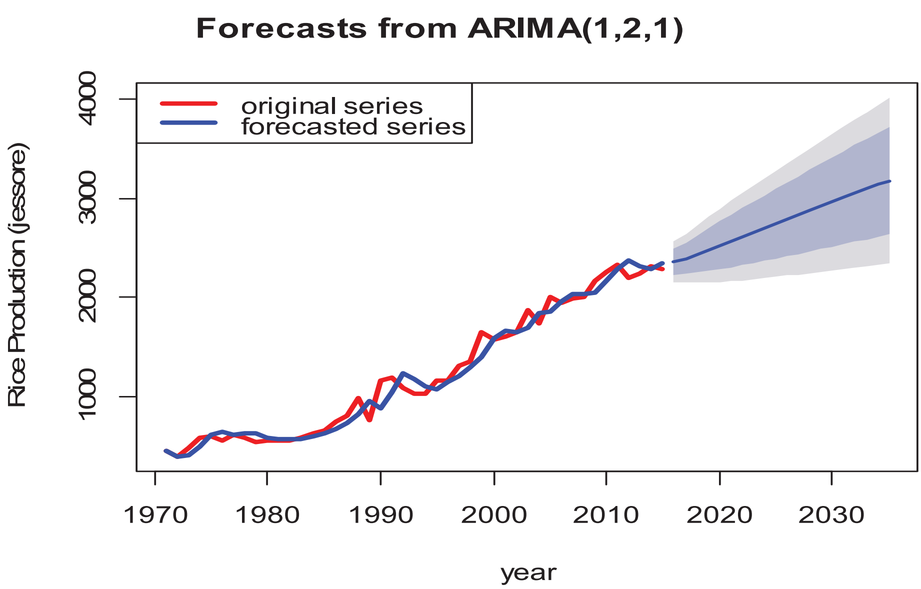
Table 5: Forecasted values of rice production in Jessore, Dinajpur and Kushtia districts

Year	Jessore	Dinajpur	Kushtia
2016	2304.223	2327.125	184.573
2017	2294.544	2319.312	210.288
2018	2284.963	2311.560	232.648
2019	2275.479	2303.869	252.091
2020	2266.091	2296.238	268.998
2021	2256.798	2288.666	283.699
2022	2247.599	2281.154	296.382

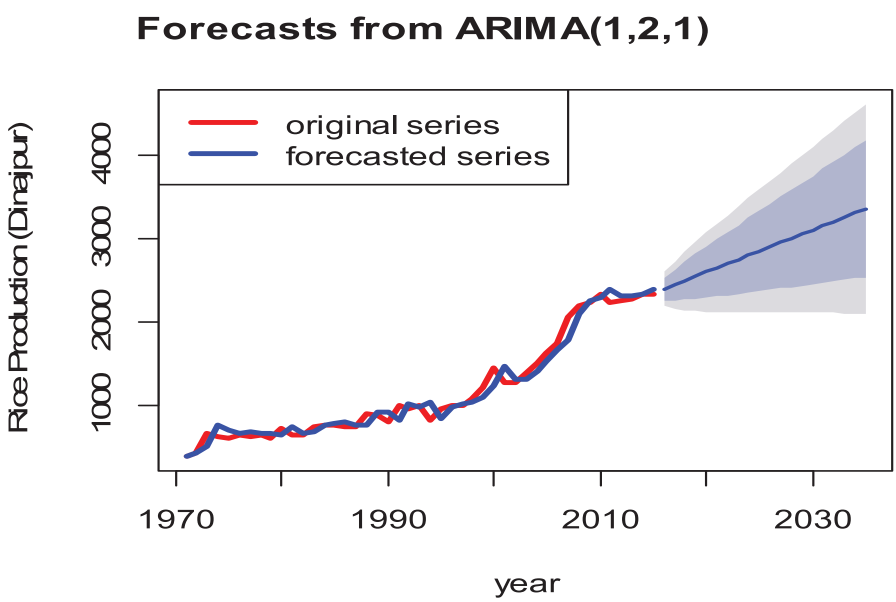
Year	Jessore	Dinajpur	Kushtia
2023	2238.492	2273.701	307.598
2024	2219.478	2266.306	317.263
2025	2220.554	2258.968	325.668
2026	2211.721	2251.689	332.976
2027	2202.977	2244.466	339.331
2028	2194.321	2237.299	344.856
2029	2185.753	2230.189	349.661
2030	2177.272	2223.134	353.839
2031	2168.876	2216.135	357.326
2032	2160.564	2209.190	360.631
2033	2152.337	2202.300	363.378
2034	2144.193	2195.463	365.767
2035	2136.131	2188.680	367.844

The graphical presentation of original and forecasted rice productions of three districts considered in this study are displayed by the following Figure 2.

(a) Jessore



(b) Dinajpur



(c) Kushtia

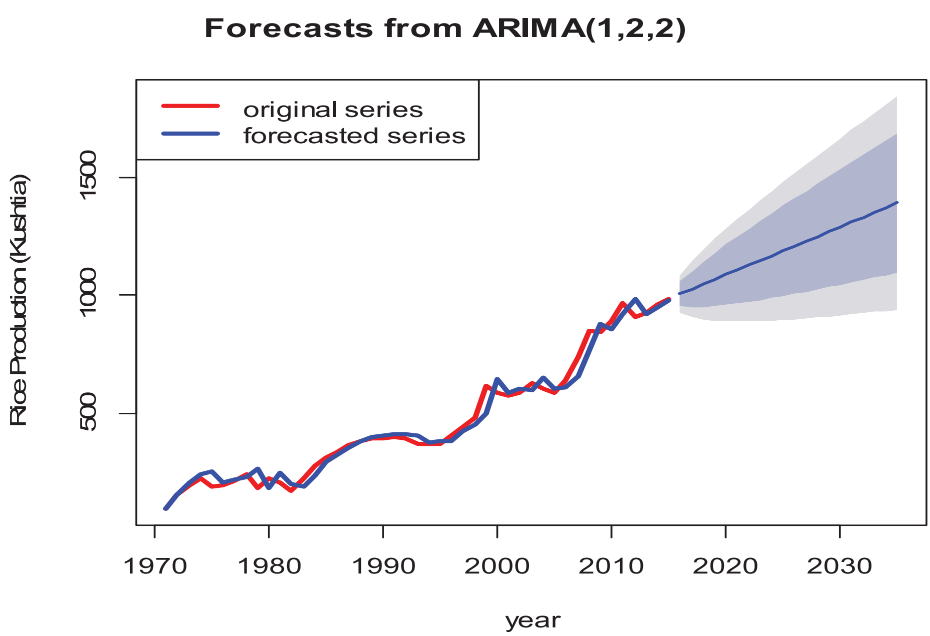


Figure 2. Visual inspection of Forecasted with Original rice productions

From the visual inspection of plotted values of original and forecasted rice productions, it is quite evident that the chosen model is good enough as the forecasted values are very close enough to the original data series.

4. CONCLUSION

At first, it is very essential to find the stationary condition of the data series, which is checked by either graphically or analytically. Here, we consider the Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) to check the stationarity of the rice production of Jessore, Dinajpur and Kushtia districts. The p-value of ADF test for the original series of rice production is 0.4743, 0.7891 and 0.4874 for the districts of Jessore, Dinajpur and Kushtia respectively and the p-value of KPSS the original series of rice production is 0.01 for all districts considered in this study. These results indicate that the rice production for all districts considers in this study are non-stationary. Now, we make them the stationary by using difference. Also, we conduct the ADF and KPSS for difference series. The results of ADF and KPSS indicate that the 2nd difference series of Jessore, Dinajpur, and Kushtia are stationary. From the result of ACF and PACF, we identify the tentative order of ARIMA model and final model is selected by AICc. Here, ARIMA (1,2,1), ARIMA (1,2,1) and ARIMA (1,2,2) models are found to be suitable for forecasting the rice production in Jessore, Dinajpur and Kushtia districts respectively. Then we made a diagnostic check to examine that the model is either appropriate for forecasting or not. Hence, the Ljung-Box test is used to test the autocorrelation problem. The results of Ljung- Box test with p-value 0.018110, 0.02462 and 2.2e-16 indicates that there is no autocorrelation in the errors of the selected model. Moreover, Jarque-Berra test is used to test the normality of the errors of the selected models. From the results, it is evident that the errors follow the normal distribution. Finally, we compare the graphically of the accuracy of the fitted model to the original series for all the districts considered in this study. Additionally, the forecasted value also shows in the graphs. Finally, with the help of selected models, we forecast the rice production of the selected districts for the upcoming 20 years which help the decision makers to establish the rice production management.

5. REFERENCES

Akaike, H., 1974, A new look at the statistical model identification. *IEEE Transactions on Automatic Control* ***AC-19***, 716–723.

BBS, 2010, *Statistical Year Book of Bangladesh*. Bureau of Statistics Divisions. Ministry of Planning. Government of the people Republic of Bangladesh. Dhaka. Bangladesh.

BBS, 2012, *Statistical Year Book Bangladesh*. Bangladesh Bureau of Statistics. Statistics & Informatics Division (SID). Ministry of Planning. Government of the People's Republic of Bangladesh. Dhaka. Bangladesh.

BBS, 2012, *Year Book of Agricultural Statistics*. Bangladesh Bureau of Statistics. Statistics & Informatics Division (SID). Ministry of Planning. Government of the People's Republic of Bangladesh. Dhaka. Bangladesh.

Hurvich, C. M., Tsai, C. L., 1989, Regression and time series model selection in small samples. *Biometrika* **76**, 297-307.

Jarque, C. M., Bera, A.K., 1987, A test for normality of observations and regression residuals. *International Statistical Review* **55(2)**, 163–172.

Ljung, G. M., Box, G. E. P., 1978, On a Measure of a Lack of Fit in Time Series Models. *Biometrika* **65(2)**, 297–303. Doi: 10.193/biomet/65.2.297.

Nargis, F., Lee, S. H., 2013, Efficiency Analysis of Boro Rice Production in North-Central Region of Bangladesh. *The Journal of Animal and Plant Sciences* **23(2)**, 527-533.

Sugiura, N., 1978, Further analysis of the data by Akaike's information criterion and the finite corrections. *Comm. Statist. A Theory Methods* **7**, 13-26.