IUBAT Review—A Multidisciplinary Academic Journal, 5 (1): 1-14
https://doi.org/10.3329/iubatr.v5i1.64587

Analysis and Forecasting of Rice Production Behavior in Bangladesh

Rehana Parvin*
Department of Statistics, International University of Business Agriculture and Technology

Received: November, 2021 Revised: April, 2022 Accepted+Published: December, 2022

Abstract

This research looks at how numerous factors, including the price of rice, area, improved seed distribution, areas irrigated by different improved methods, and fertilizer, influence rice production in Bangladesh. This study employed the autoregressive distributed lag (ARDL) model to investigate the short-run and long-run effects of each exogenous variable on rice production. In this research work, auto regressive integrated moving average (ARIMA) and hybrid ARIMA-GRACH (general autoregressive conditional heteroskedastic) were used to forecast rice production in Bangladesh from 1983 to 2030. The parameters of price of rice, area irrigated by different improved methods, area, and fertilizer were found to have a significant and positive impact on rice production in this study, both in the short run and the long run, according to the ARDL model. According to the lowest Root Mean Squared Error, Mean Absolute Error and Theil Inequality Coefficient criteria, the linear hybrid ARIMA (11, 1, 11)–EGARCH (1, 1)–M model performed better in predicting the rice production in Bangladesh.

Keywords: *Bangladesh; Mean Absolute Error; Rice production; Root Mean Squared Error; Theil Inequality; Autoregressive distributed lag (ARDL).*

1. Introduction

Bangladesh is a predominantly agrarian nation. Agriculture is the most important sector of the Bangladeshi economy, accounting for 19.6% of gross domestic product (GDP) and employing 63% of the workforce (Bangladesh Economic Review, 2020). About 85 percent of the population lives in rural areas, while agriculture employs 75 percent of the rural labor force. As such, Bangladesh's economic freedom is mostly dependent on the growth of agricultural productivity. The agricultural industry plays a critical role in most agricultural countries and has a significant impact on overall economic performance. Both commercially and culturally, rice is Asia's most important food crop, and its production is considered the world's most important economic activity. Therefore, it is crucial to figure out how rice production in Bangladesh is affected by different factors and how to predict it.

To investigate the influence of various factors on rice production, the autoregressive distributed lag (ARDL) model was utilized. Rice output was used as a predictor in the study, with independent variables including production prices, distribution of improved seeds, area irrigated by various enhanced technologies, area, and fertilizer. Although the auto-regressive integrated moving average (ARIMA) model is an effective tool for forecasting time series data, it is not able to manage all the volatility and nonlinearity. To solve these drawbacks, the general autoregressive conditional heteroskedastic (GARCH) model or a hybrid ARIMA-GARCH model is used to capture the volatility and nonlinearity of time series data. Furthermore, the hybrid ARIMA-GARCH model might be a useful tool for overcoming constraints and improving forecast validity. So, the presentation of likely univariate time hybridization was investigated in this work.

The most significant component of GDP is agricultural production. The agricultural output of a country has a significant impact on its GDP, which reflects the country's economic growth. Agriculture generates nearly half of the country's value added. Due to changes in agricultural output, the overall pace of economic growth has been unstable. Agriculture has also helped the country generate foreign currency. All these considerations point to the importance of this sector's expansion in the development of the national economy. Bangladesh's agricultural sector is underdeveloped due to our farmers' lack of access to new technologies because of poverty. Agricultural imports account for a significant percentage of our foreign currency each year. It would boost our foreign currency reserves if we could expand crop production to fulfill our demands while reducing agricultural imports. Agriculture employs roughly three-fifths of the population. It is the backbone of Bangladesh's economy, accounting for 45 percent of the country's GDP (Rehman *et al*., 2019). Therefore, agricultural output analysis and forecasting are critically important in national policymaking.

*Corresponding Author's Email: rehana.parvin@iubat.edu

Review of the literature shows, no research into the elements that influence rice production has yet been conducted in Bangladesh. Therefore, it is crucial to figure out how rice production in Bangladesh is affected by production prices, distribution of improved seeds, area irrigated by various enhanced technologies, area, and fertilizer. Production forecasting is an integral component of the decision-making process and the appraisal of investment scenarios, both of which are critical for an upstream firm.

1.1 Objectives of this study

- To investigate the short-run and the long-run effects of production prices, distribution of improved seeds, area irrigated by various enhanced technologies, area, and fertilizer on rice production.
- To gather the ARIMA and hybrid ARIMA-GARCH models' expected presentations.
- To predict the rice production behavior in Bangladesh using a hybrid ARIMA-GARCH model.

2. Literature review

The ARDL model has been used in most agricultural studies in recent years (Yurtkuran, 2021; Aziz *et al*., 2020; Adebayo *et al*., 2021; Sarkodie and Owusu, 2016). Warsame et al. (2021) used the ARDL model to investigate how climate change affects crop productivity, considering recorded rainfall, temperature, and CO2. Rainfall increases crop production in the long run but hinders it in the short run, according to the authors' results, whereas temperature has a negative influence on crop output in both the long and short runs. However, agricultural output is unaffected by carbon dioxide emissions.

Koondhar *et al*. (2020) used the ARDL model to observed the relationship between air pollution, energy consumption, and agricultural GDP contribution. The F-statistic values exceeded the upper bound value in the ARDL bounds testing model's results. According to the authors' findings, the entire model was changed at a rate of 2.364 in the direction of long-run equilibrium. Rehman *et al*. (2019) investigated the association between agricultural gross domestic product (AGDP) and variables like cultivated area, fertilizer usage, credit distribution, and water availability in Pakistan using econometric analysis. The study revealed that fertilizer use, improved seed distribution, and credit distribution & had a positive and significant impact on AGDP, but water availability had a negative but negligible impact. Greater use of improved maize varieties and mineral fertilizers, along with increased extension services and the absence of disastrous droughts, according to Abate *et al*. (2015), are the primary factors boosting Ethiopia's maize productivity development. In the research and development of maize and other commodities, Ethiopia used a homegrown solution method. Ethiopia's maize experience has shown us that continued investment in agricultural research and development, as well as national government policy support, is critical for agriculture's sustainable growth.

In recent years, several forecasting models have been integrated with the ARIMA model in the agriculture sector using time series data to improve forecasting results (Mahto *et al*., 2019; Iqbal *et al*., 2005; Wang *et al*., 2018; Wen *et al*., 2019). Compared to other forecasting approaches like simulation and intelligence techniques, ARMA–GARCH models offer the advantages of accuracy and explicitness, as well as the capacity to accommodate heteroskedasticity. It is important to note that GARCH models come in a variety of flavors, with several of them capable of representing asymmetric time-varying volatility. The goal of this article is to thoroughly evaluate and compare numerous models to determine which one is the best. Many forecasting models have been combined with the ARIMA/ARCH/ANN models in recent years (Bhardwaj *et al*., 2014; Chuan *et al*., 2013; Jannah *et al*., 2021).

Alshammari *et al*. (2020) used wavelet approaches to model and forecast the Saudi Arabia stock market, including the ARIMA and GARCH models. The authors suggest integrating the best MODWT function (the BL14 function) with the fitted GARCH model to create a new forecasting model. Asadullah *et al*. (2021) used ARIMA, Naive, Exponential smoothing, and one multivariate model to forecast the Chinese Yuan/USD, and the authors claim that Naive beats all other time series models individually and in combination. Nguyen and Nguyen (2020) examined the factors that influence Vietnamese farmers' readiness to accept organic farming using the theory of planned behavior. Based on the degree of impact of these elements on persuading Vietnam to accept organic farming, the authors offered a few recommendations to state management authorities. Wulandari *et al*. (2020) look at the nexus between production inputs and pricing commodities in East Java, Indonesia.

Mohammadi and Su (2010) used ARIMA-GARCH models to forecast oil prices in eleven worldwide marketplaces. The authors explored the four-variance models of GARCH, APARCH, EGARCH, and FIGARCH. The APARCH model outperforms all four variance models. The invariability of oil returns is captured better by restrictive standard deviation than by restrictive variance. By using an ARIMA-GARCH Model to Improve Subway Short-Term Ridership Prediction by Accounting for Dynamic Variance, Ding *et al*. (2017) improved subway short-term ridership prediction. The findings of the model demonstrate that the recommended model outperformed the standard model at all three locations. By reducing the length of the average gap in the prediction of the ridership and increasing the prediction gap coverage probability, the hybrid model considerably improved the reliability of the anticipated point value. The short-term traffic was also modeled, considering the various traffic designs between weekdays and weekends.

The ARIMA-GARCH combination model has enhanced evaluation and prediction precision when compared to the standard forecasting method, according to the findings. The results indicate that while the ARIMA model is powerful and versatile, it is unable to handle the data set's invariability. On the other hand, hybrid ARIMA-GARCH can deal with both non-linearity and invariability in a data set. Therefore, the authors propose that the ARIMA-GARCH combination is the best and most successful concept for forecasting and modeling gold prices. Uwilingiyimana *et al*. (2015) constructed ARIMA-GARCH models to anticipate inflation rates in Kenya. In comparison to typical forecasting models, the linear hybrid ARIMA (1, 1, 12)-GARCH (1, 2) model gives the appropriate outputs for modeling and forecasting inflation rates. Zhou *et al*. (2006) introduced a non-linear GARCH model and a linear ARIMA model for forecasting a new traffic network. The results revealed that the hybrid ARIMA-GARCH model performs better in both the short and long runs and can handle both non-linearity and invariability. In terms of prediction accuracy, the authors claimed that the ARMA/GARCH model surpasses the conventional FARIMA model.

Using hybrid ARFIMA–GARCH models such as ARFIMA–IGARCH, ARFIMA–FIGARCH, ARIMA–GARCH, and ARFIMA–GARCH, Kang and Yoon (2013) investigated the tenacity and invariability of market returns. The ARFIMA–FIGARCH model outperforms the others in terms of market returns and volatility for three different types of petroleum. Different ARMA models paired with GARCH processes, specifically ARMA–GARCH models, as well as their adjusted variants, ARMA–GARCH-in-mean, were used to predict and forecast power prices by Liu and Shi (2013). To model and forecast electricity prices, it integrates GARCH processes, especially ARMA–GARCH models and their adjusted variants, ARMA–GARCH-in-mean. The empirical results reveal that the ARMA–GARCH-M models are the best instruments for estimating and predicting the average and variance of electricity prices, while the ARMA–SGARCH-M models are simple and the ARMA–GJRGARCH-M model is ruthless. They also discovered that power prices fluctuate over time and exhibit an inverse nonlinear leverage effect.

In most of the above-mentioned literature, the ordinary least square (OLS) method was employed to investigate how different factors can explain worldwide agricultural production. The OLS approach, on the other hand, was unable to detect the independent variable's lag effects on the dependent variable. To fill this gap, we used the ARDL model in this work to examine how a variety of parameters, such as rice price, area, improved seed distribution, areas irrigated using various improved methods, and fertilizer, affect rice output in the short and long run. The ARIMA model is a useful tool for forecasting time series data, but it cannot manage all the nonlinearity and volatility. The GARCH model or a hybrid ARIMA-GARCH model was employed in this study to capture the volatility and nonlinearity of time series data to overcome these disadvantages. The hybrid ARIMA-GARCH model may also be effective for overcoming restrictions and enhancing forecast validity.

3. Methodology

The methodology involved analysis of data obtained from secondary sources in terms of existing model of ARDL-GARCH model.

3.1 Data information

The data came from the Bangladesh Bureau of Statistics Agriculture Yearbook, 45 years of Agriculture Statistics of Major Crops published by the Bangladesh Bureau of Statistics (BBS), and other yearbooks produced by BBS. The yearly secondary time series data for empirical analysis were collected from 1983 to 2020. Rice production (in lac Metric ton) is a dependent variable in this study, whereas production price (in Taka/Metric ton), distribution of improved seeds (in kg), area irrigated by various improved methods (in acres), fertilizer (in kg), and area (in lac Hectare) are independent variables.

3.2 Methods

The overall study is divided into two sections: analysis and forecasting. The analysis section looks at how a variety of factors influence rice production, including the price of rice, area, improved seed distribution, area irrigated using various improved methods, and fertilizer. We applied the ARDL model to check the lag effects of the price of rice, area, improved seed distribution, area irrigated by improved methods, and fertilizer on rice production in the short and long runs. To forecast rice production in Bangladesh from 1983 to 2030, the forecasting division employed ARIMA and several forms of hybrid ARIMA-GRACH models. This paper applied three different measurements for the forecasting performance measures of rice production. The very common error is the Root Mean Squared deviation (RMSE), which is a very common one. The two other common criteria are Mean Absolute Error and Theil Inequality Coefficient, abbreviated as MAE and TIC, respectively.

3.2.1 ARDL model

This paper employed the augmented Phillips curve approach to explore the relationship between rice, area, improved seed distribution, area irrigated by improved methods and fertilizer in the rice production of Bangladesh from 1983-2020. The ARDL equation is formulated as follows when the price of rice, area, improved seed distribution, area irrigated using various improved methods and fertilizer are considered as independent variables, and rice production is considered as a dependent variable:

$$\Delta rice_t = \alpha_0 + \sum_{i=1}^{p} a_{1i}\,\Delta rice_{t-i} + \sum_{i=1}^{q} a_{2i}\,\Delta price_{t-i} + \sum_{i=1}^{r} a_{3i}\,\Delta area_{t-i} + \sum_{i=1}^{s} a_{4i}\,\Delta seed_{t-i} + \sum_{i=1}^{t} a_{5i}\,\Delta method_{t-i} + \sum_{i=1}^{u} a_{6i}\,\Delta fer_{t-i} + \mu_t \quad (1)$$

Where, rice stands for rice production, price stands for price of rice, seed stands for improved seed distribution, methods stand for area irrigated using various improved methods and fer is stand for fertilizer.

This study requires some pre-tests to estimate the equation (1). We apply the well-known Augmented Dickey-Fuller (ADF) test to examine the order of integration of the selected variables. The ARDL approach, on the other hand, is applicable whether the variables are purely I (0) or I (1), even if fractionally integrated. However, this method is not applicable to the I (2) series because, the estimators become invalid in the I (2) series. We also apply a general-to-specification approach to get the ARDL model's final specification, and the Schwarz Information Criterion (SIC) are chosen as the best lag lengths. To evaluate the existence of a long-run link between the variables, the limits testing approach proposed by Shin *et al*. (2014) was used to conduct a co-integration test. Using F-test, we check the null hypothesis of $a_{1i} = a_{2i} = a_{3i} = a_{4i} = a_{5i} = a_{6i} = 0$ jointly.

3.2.2 The Box-Jenkins Method of ARIMA Model

Box and Jenkins introduced the Box-Jenkins ARIMA model in 1976. The ARMA models, as well as the autoregressive (AR) and moving average (MA) models, are combined in the Box-Jenkins ARIMA model. The ARIMA model is a simple and successful technique for forecasting a range of data sets. The Autoregressive (AR) model predicts how a variable will change in relation to its previous or previous values. The difference between the original data values and the prior values is denoted by the term integrated (I). The Moving Average (MA) model uses prior data to infer the link between a residual error and an observation from an MA structure. ARIMA writes an ARIMA model, which has three parameters (p, d, q). The letters q and p stand for an autocorrelation function (ACF) element and a partial correlation function (PCF) element, respectively (PACF). The number of times the real time series observations are different is known as d, and the lag observation order is known as P. The MA window size is known as q, and the number of times the actual time series observations are different is known as q. The prediction equation for a stationary time series is:

Predicted value of Z = a fixed value and one or more current values of Z or errors.

The qualities of stationary data are not affected by the time of collection. If the mean and variance of a data series do not change over time, it is called stationary. Various transformations are employed to convert non-stationary data to stationary data. The following prediction equation was constructed to apply the relevant notions to the problem. At first, consider, z is the dth difference of Z, that is indicates that,

$$\text{If } d = 0\,;\ z_t = Z_t \tag{2}$$

Where, $d = 0$ indicates series is at level.

$$\text{If } d = 1\,;\ z_t = Z_t - Z_{t-1} \tag{3}$$

Where, d = 1 indicates series is at first difference.

$$\text{If } d = 2\,;\ z_t = (Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2}) = Z_t - 2Z_{t-1} + Z_{t-2} \tag{4}$$

Where, d = 2 indicates series is second difference. With reference to z, the general predicting equation:

$$\hat{z}_t = \alpha + \phi_1 z_{t-1} + \dots\dots\dots\dots + \phi_p z_{t-p} - \theta_1 e_{t-1} - \cdots \dots\dots .. - \theta_q e_{t-q} \tag{5}$$

Here is the final predicting equation, which is introduced by Box and Jenkins. Where, θ'sare the MA parameters and ϕ's are AR parameters. The coefficients are known by AR (1), AR (2) … and MA (1), MA (2) … etc.

3.2.3 The ARCH Family Models

The Autoregressive Conditional Heteroscedasticity Model, or ARCH model, defines the volatility of time series data in econometrics. The ARCH model is appropriate if time series data follows an AR model. ARCH models are used to describe financial time series data. This time-varying volatility can be captured using the ARCH model. GARCH models, on the other hand, are typically used when the ARMA model fails to capture the error variance.

3.2.4 The ARCH Model

For ARCH model, this study $\varepsilon_t$considered as an error term. To forecast the conditional volatility, the AR(q) is an easy way. The AR(q) process equation is:

$$\hat{\varepsilon}_t^2 = \partial_0 + \partial_1 \hat{\varepsilon}_{t-1}^2 + \partial_2 \hat{\varepsilon}_{t-2}^2 + \cdots + \partial_q \hat{\varepsilon}_{t-q}^2 + \pi_t \tag{6}$$

Where, π_t is known white noise term. At the time t+1 the forecast equation will be:

$$E_t \hat{\varepsilon}_{t+1}^2 = \partial_0 + \partial_1 \hat{\varepsilon}_t^2 + \partial_2 \hat{\varepsilon}_{t-1}^2 + \cdots + \partial_q \hat{\varepsilon}_{t+1-q}^2 \tag{7}$$

In 1982, Engel introduced equation (6) as an ARCH model.

3.2.4 The GARCH –M Model

EGARCH (Nelson, 1991), nonlinear GARCH (NGARCH) (Engle and Ng, 1993), quadratic GARCH (QGARCH) (Sentana, 1995), and Glosten–Jagannathan–Runkle GARCH (GJRGARCH) models are some of the most common GARCH variants (Glosten *et al*., 1993). These models employ a nonlinear function to explain time-varying volatility as well as leverage effects that are not captured by traditional GARCH models.

A heteroskedasticity factor is included in the mean equation of a GARCH-M model to highlight the impact of volatility on mean prediction. The GARCH-M model can be expressed in a variety of ways, including NGARCH-M and EGARCH-M. When the variance is not fixed, the GARCH -M model can successfully capture invariability. The GARCH-M model is frequently referred to as the invariability model. The GARCH -M mode lenhanced by Bollerslev in 1986.

The following equation can be an error equation:

$$\varepsilon_t = \pi_t \sqrt{\gamma_t} \tag{8}$$

Consider that $\sigma_\pi^2 = 1$ and

$$\gamma_t = \partial_0 + \sum_{i=1}^{q} \partial_t \, \hat{\varepsilon}_{t-i}^2 + \sum_{j=1}^{p} \partial_j \, \gamma_{t-j} \,. \tag{9}$$

Equation (9) in recognized as a GARCH-M model with two parameter p and q. where p represents MA elements and q represents AR elements and written by GARCH (p, q)-M.

3.2.5 The IGARCH –M Model

For non-linear time series, the data GARCH model with respect to meaning is suitable. The IGARCH -M model is the new edition of the GARCH-M model. The IGARCH-M model is written by IGARCH (p, q)-M and has two parameters, p and q, where p represents MA elements and q represents AR elements. The following general restriction has been applied to the GARCH –M model to develop the IGARCH-M model:

$$\sum_{i=1}^{q} \partial_i + \sum_{j=1}^{p} \mu_i = 0 \text{ and } 0<\mu<1 \tag{10}$$

3.2.6 Exponential GARCH - M Model

In 1991, Nelson introduced another model of the ARCH family, named the exponential GARCH model. The GARCH model has a restriction in that all parameters must be non-negative. However, in real life, sometimes, the calculated parameter value is negative. In that case, the EGARCH –M model is successfully applied. On the other hand, the EGARCH model can capture the asymmetric effects in the data series. The following equation is known as the Exponential GARCH –M Model:

$$ln(\gamma_t) = \partial_0 + \partial_1(\varepsilon_{t-1}/\gamma_{t-1}^{0.5}) + \rho_1|\varepsilon_{t-1}/\gamma_{t-1}^{0.5}| + \alpha_1 ln(\gamma_{t-1}) \tag{11}$$

3.2.7 The Threshold GARCH –M (TGARCH) Model / GJR Model

In 1993, Glosten et al., suggested the TGARCH (p, q)–M model. This is the invariability measurement model and is quadratic in nature. To capture the leverage effect, the TGARCH (p, q)-M model is the perfect one. The GJR-GARCH –M model is written by:

$$\gamma_t = \partial_0 + \sum_{i=1}^{q} \partial_t \, \hat{\varepsilon}_{t-i}^2 + \sum_{j=1}^{p} \alpha_j \, \gamma_{t-j} + \sum_{i=1}^{q} \tau_i \, \pi_{t-i} \hat{\varepsilon}_{t-i}^2 \tag{12}$$

3.2.8 Forecasting Performance Measures

This article used three independent measurements to forecast rice production performance measures. The Root Mean Squared Error (RMSE) is a relatively prevalent type of error. Mean Absolute Error and Theil Inequality Coefficient, abbreviated as MAE and TIC, respectively, are two criteria that are more popular. This study also makes a comparison of the value of the above-mentioned criteria. An appropriate model will be that one has the lowest value of above-mentioned criteria.

RMSE: The RMSE is calculated by taking the differences between the observed values and the predicted values. The RMSE result is always positive; if the RMSE value is zero, the model fits well. The lowest value of the RMSE provides the better forecasting capacity of the model. The RMSE formula is as follows:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(\widehat{z_t} - z_t)^2}{n}}$$

Where, n is the number of forecasts at time t and z_t is actual values, $\hat{z}_t$ is the forecasted values.

MAE: Itconsiders the whole meaning of the error's total value. It calculates the absolute value of the residual's mean value. It is as follows:

$$MAE = \frac{\sum_{t=1}^{n} |\widehat{z_t} - z_t|}{n}$$

Where, n is the number of forecasts at time t and z_t is actual values, $\hat{z}_t$ is the forecasted values.

TIC: Theil inequality coefficient is a different kind of forecasting criterion. The TIC range is always zero to one, with zero indicating the best match.

$$TIC = \frac{\sqrt{\sum_{t=1}^{n} (\widehat{z_t} - z_t)^2}}{\sqrt{\sum_{t=1}^{n} \widehat{z_t^2}} + \sqrt{\sum_{t=1}^{n} z_t^2}}$$

Where, n is the number of forecasts at time t and z_t is actual values, $\hat{z}_t$ is the forecasted values.

4. Results and Discussion

4.1 Analysis section

The unit root test, also known as the stationary test, is the most important condition for time series data when looking at the order of integration of variables. The Augmented Dickey-Fuller (ADF) test was used to conduct our empirical research. Table 1 summarizes the results of the ADF test. For the best lag structure, involving intercept and linear time trend at level but eliminating time trend from the first difference, the Schwarz information criteria (SIC) was used. All of our study variables are stationary at first difference, showing that they are I (1) according to the ADF test. In the absence of I (2) variables but two variables like rice production and fertilizer also stationary at level at 5% level of significance. So, we can utilize the bound testing methodology for ARDL model to approximate the equation (1).

Table 1**:** Unit root test results

Variables	Level/Difference	ADF t-statistic
Rice	Level	0.899068**
	First difference	-7.288810**
Price	Level	0.104313
	First difference	-9.207845**
Area	Level	0.205396
	First difference	-7.302215**
Seed	Level	-1.72151
	First difference	-6.661854**
Method	Level	-0.280750
	First difference	-6.608569**
Fertilizer	Level	0.172578**
	First difference	-8.722619**

Note: ** refer significant at 5% levels of significance. For optimal lag order AIC criterion is used and constant and time trend are included in level, but only constant is included in 1st difference.

The The long-run connection is determined by the model's optimal lag section, according to Bahmani-Oskooee and Bohl (2000). For optimal lag order selection, the Akaike information criterion (AIC) is used. Model estimation results for symmetric long-run relationships (co-integration) are shown in Table 2. Because the F-bound statistic result of 7.103855 is larger than the necessary upper limit of 3.79 at 5 percent, the bound test confirms that there is a long-run relationship or co-integration in a linear fashion. Some diagnostic tests were also undertaken to support the ARDL model's dependability. The Jarque-Bera (J-B) test, the Ramsey RESET test, the Autocorrelation Conditional Heteroskedasticity (ARCH) up to order 2 for heteroskedasticity, and the serial autocorrelation LM test up to level 2 for serial autocorrelation were all used to assess error normality. The results of all the tests are shown on the bottom panel of Table 2.

ARDL model passes all diagnostic tests, implying that it is reliable. The speed of adjustment (SOA) is a metric that measures how quickly companies close the difference between their prior year's leverage and their current period's desired leverage. The adjustment speed, according to our data, is -4.47, indicating a 47 percent increase in significance in the previous period to reach equilibrium. We also used the cumulative sum (CUSUM) and cumulative sum square (CUSUM Square) tests to ensure that the model was stable. Fig. 1 shows the results of these tests and shows that the model is quite stable.

Table 2 presents the results of the ARDL long run parameters. It shows that rice prices, area, distribution of improved seeds, and fertilizer have positive, significant lag effects on rice production at the 5% level of significance. Rice has a coefficient of 0.006527, meaning that if the price of rice rises by 1%, rice production climbs by 0.0065%. These findings could be explained by the notion that the higher the price, the better the quality. Some consumers may believe that if they choose the middle-priced option, they will get an average-quality product, and if they choose the lowest-priced option, they will get the lowest-quality product. The coefficient of improved seed distribution is 0.3325, implying that if improved seed distribution for rice production rises by 1%, rice production will increase by 0.3325 percent. These results could be explained by the assumption that seeds are a crucial component in enhancing crop production and yield. Seed quality can significantly boost a crop's yield potential, making it one of the most cost-effective and efficient agricultural inputs.

Table 2**:** ARDL Model estimates result

Variables	Coefficient	Std. Error	t-statistic
Price	0.006527**	0.003359	1.943137
Method	7.756709	3.357331	2.310380
Fertilizer	0.142832**	0.248591	0.574569
Area	4.779619**	0.548012	8.721737
Seed	0.332590**	0.280440	1.185958
cointEq(-1)*		-0.472130**	
Constant		-210.2643	
Adjusted R-square		0.990053	
F- bound test		7.103855 [2.62, 3.79]	
t- test		-5.749423 [-2.86, -4.19]	
J-B [prob]		0.4171	
R-R [prob]		0.1310	
LM (1) [prob]		0.5522	
LM (2) [prob]		0.8194	
ARCH (1) [prob]		0.8063	
ARCH (2) [prob]		0.3067	

Note: ** refer significant at 5% levels of significance. For optimal lag order, SIC criterion is used. The critical values for F-bound test and t-test are from the Narayan (2005).

The coefficient of area is 4.779, which means that expanding the cultivated area for rice production by 1% will result in an increase of 4.779 percent in rice production. The concept that areas are a major component in boosting agricultural productivity could explain these findings. If the land area is large enough, it can considerably increase the output potential of a crop, making it one of the most cost-effective and efficient agricultural inputs. Fertilizer has a coefficient of 0.1428, which means that if fertilizer use increases by 1%, rice production will increase by 0.1428 percent. Fertilizers, which boost output and promote healthy production by delivering the right balance of nutrients to the soil, are one of the likely causes of these findings. If fertilizers are not applied, the soil will become exhausted, making it difficult for plants to thrive.

The value of adjusted is reasonably high. This means that rice production is strongly influenced by rice price, area, the distribution of better seeds, the amount of land irrigated in various ways, and fertilizer. The fitted model can explain 99 percent of the variation in rice production based on the value of adjusted (coefficient of multiple determinations).The P-value and predicted parameters have been considered in the comments. The significance level was set at 5%.

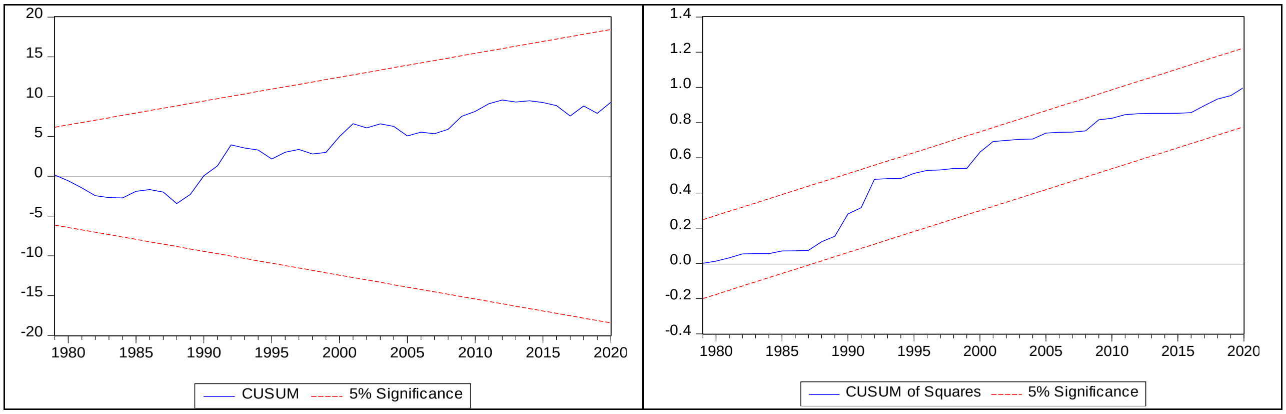


Fig. 1: Model stability check using the Cumulative Sum (CUSUM) and Cumulative Sum square (CUSUMSQ) test.

4.2 Forecasting

4.2.1 Model Estimation and Forecasting Using the ARIMA Box-Jenkins Method

The initial stage in this procedure was to look at the time series data's stationary analysis. The findings of the unit root test have already been reviewed in detail (see Table 1).Calculating the value of the AR components q and the MA elements p is the next stage in the identification procedure. The parametric values for these can be obtained by casting them into the autocorrelation and partial autocorrelation coefficients. It is assumed that the value of parameter ‘d’ is one (1). The ACF and PACF correlogram series have defined some guidance that appears to prepare a statistical demarcation of the available data. As shown in Appendix A1(a), the Auto Correlation Function (ACF) significantly cut-across the confidence line only for lag 11, and the Partial Autocorrelation Function (PACF) significantly cut-across the confidence line only for lag 11. The value of "q" and "p" is reported by two separate series of the correlogram, ACF and PACF. Therefore, ARIMA (11, 1, 11) is recommended as the most frugal model in this investigation.

4.2.2 Diagnostic Check of the Model ARIMA (11, 1, 11)

After evaluating the coefficients of a suggested parsimonious model, an examination is required to confirm whether the model is good enough for prediction. The model's residual must be white noise (Ljung-Box Q statistic) for a solid univariate process to exist; the null hypothesis is that the residual is white noise. The process should be stationary in terms of covariance, and it should also be invertible.To estimate all of the information contained by this model, the ACF and PACF of residuals were tested for ARIMA (11, 1, 11). All lags are flattened within a 95 percent confidence level, according to the residual correlogram of the ARIMA (11, 1, 11) model. The null hypothesis was found to be acceptable (i.e., residuals are white noise). All autoregressive roots are inside the circle, showing that the ARMA process is covariance stationary, and all moving average roots are also inside the circle, suggesting that the ARMA process is invertible, as shown in Appendix A1(b). So, the chosen model is considered the best fit and is utilized for predicting.

4.2.3 Estimating and Forecasting Based on the Various GARCH Models

Because of the ARIMA model's poor performance, R-Square was relatively low, while RMSE, MAE, and TIC were all quite high. Therefore, a new model for predicting rice production must be developed. Table 3 summarizes the joint estimation for several hybrid ARIMA-GARCH-M models, including ARIMA (11, 1, 11) – GARCH (1, 1)-M, ARIMA (11, 1, 11) – EGARCH (1,1)-M, ARIMA (11, 1, 11) – GJR-GARCH (1,1)-M, and ARIMA (11, 1, 11) – PARCH (1,1)-M. In the appendices for the individual models, the results of forecasting measures (RMSE, MAE, and TIC) and residual analysis are also shown. In terms of performance, all models performed admirably. The ARCH order was set to 1, the GARCH order was set to 1, the asymmetric order was set to 1, and the student's error distribution was set to 1.

4.2.4 Relative Analysis

Some forecasting accuracy metrics, including MAE, RMSE, MAPE, and TIC, have been incorporated and contrasted in the proposed models in this study. The Akaike information criterion (AIC), the Schwarz information criterion (SIC), and the R-square are used to compare models. The estimated results of RMSE, MAE, TIC, AIC, SIC and R-square are shown in Table 3. Table 3 shows that the RMSE, MAE, MAPE, and TIC criteria for the ARIMA (11, 1, 11) – EGARCH (1, 1)-M model are all relatively low, with values of 15.31, 12.93, 5.631, and 0.030, respectively. Compared to other models, the AIC (6.79) and SC (7.07) are both low, whereas the R-square (0.987) is high. Therefore, ARIMA (11, 1, 11) – EGARCH (1,1)-M is the best model for predicting performance. The ARIMA (11, 1, 11) – EGARCH-M model clearly reflects the nonlinearity and invariability to forecast the rice production in Bangladesh.

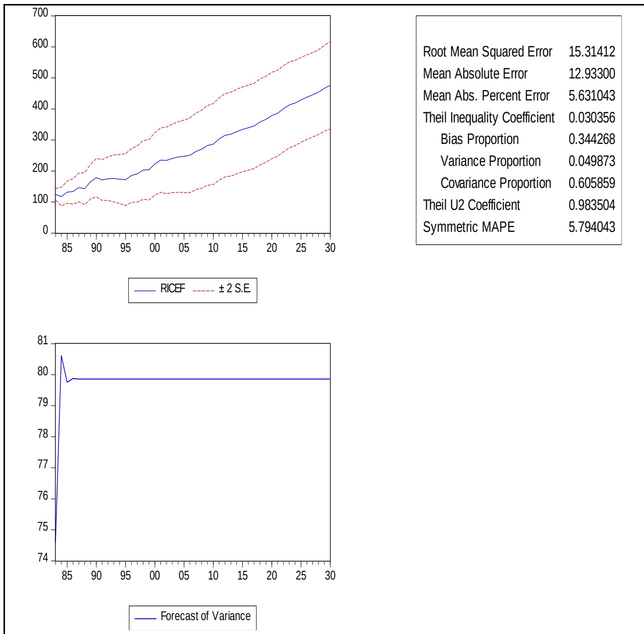
Table 3**:** Comparison of test statistics for ARIMA and family of hybrid ARIMA-GARCH –M models

Model	RMSE	MAE	MAPE	TIC	AIC	SBIC	R-square
ARIMA (11, 1, 11)	29.36	23.40	8.67	0.0594	7.796806	7.951240	0.110370
ARIMA (11,1,11)-GARCH-M (1,1)	47.30	41.69	20.13	0.0857	6.875068	7.219823	0.980
ARIMA (11,1,11)-EGARCH-M (1,1)	**15.31**	**12.93**	**5.631**	**0.030**	**6.79**	**7.07**	**0.987**
ARIMA (11, 1, 11) -GJR-GARCH-M (1,1)	24.71	20.03	10.40	0.046	7.33	7.71	0.977
ARIMA (11, 1, 11)-PARCH –M (1,1)	32.33	26.46	10.89	0.064	7.75	8.18	0.9825

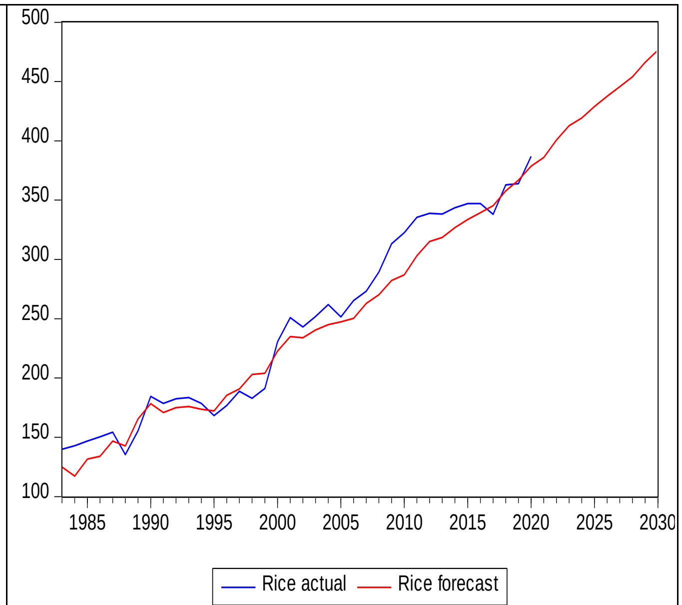
Note: RMSE presents Root Mean Square Error, MAE presents Mean Absolute Error, MAPE presents Mean Absolute Percentage Error, TIC presents Theil Inequality Coefficient, AIC presents Akaike Information Criterion, and SBIC presents Schwarz Criterion respectively.

4.2.4 ARIMA (11, 1, 11) – EGARCH (1, 1) -M Model for Forecasting of Rice Production in Bangladesh

If the observed and anticipated values are near, the forecast error is low. As a result, researchers always favor the lowest forecasting accuracy criterion. All forecast errors, such as RMSE, MAE, MAPE, and TIC, were found to be lower for the ARIMA (11, 1, 11) – EGARCH (1, 1)-M model than for other models. The residual distribution of the histogram demonstrates that it is generally normal. As a result, from 1983 to 2030, the ARIMA (11, 1, 11) – EGARCH (1, 1)-M model performed better in forecasting rice output in Bangladesh. Fig. 2(a) shows the ARIMA (11, 1, 11) – EGARCH (1,1)–M forecasting graph for rice production in Bangladesh, with predicted outcomes within the 95 percent confidence range (+/-2SE). Fig. 2 depicts the accuracy of the forecasted outcome by presenting the same graph for both the actual and forecasted values. The actual and anticipated values of rice production are fairly close on this graph, and several points cross. As a result, the proposed model is found to be effective in predicting Bangladesh's rice production. Fig. 2(b) also displays the projected value for the following ten years (from 2021 to 2030), indicating that rice output in Bangladesh is increasing.This research used time series data from 1983 to 2020 and was forecasted from 1983 to 2030. Rice production in Bangladesh is expected to reach 385.9236 million tons in 2021, 400.727 million tons in 2022, and 476.4641 million tons in 2030. The expected results are shown in detail in Table 4.



a) Forecasting graph by using ARIMA (11, 1, 11) – EGARCH (1, 1) -M model



b) Comparing forecast graph of actual and forecasted rice production by using ARIMA (11, 1, 11) – EGARCH (1, 1) -M model

Fig. 2: Forecast graph

Table 4**:** Forecasting result of rice production (in Lac Metric tons)

Year	Actual rice production	Forecasted rice production	Year	Actual rice production	Forecasted rice production
1983	139.9	124.9932	2007	273.18	262.8427
1984	142.8	117.1904	2008	289.31	270.2693
1985	146.85	131.57	2009	313.17	282.2212
1986	150.37	133.9925	2010	322.57	286.8905
1987	154.25	146.8856	2011	335.41	303.1102
1988	135.37	142.7613	2012	338.89	315.1696
1989	155.44	165.4607	2013	338.33	318.4515
1990	184.56	178.4105	2014	343.56	326.8279
1991	178.52	170.8623	2015	347.1	333.776
1992	182.52	175.108	2016	347.1	339.3964
1993	183.41	175.9633	2017	338.04	345.4333
1994	178.51	173.662	2018	362.79	357.751
1995	168.39	172.2639	2019	363.91	366.7379
1996	176.87	185.4206	2020	387	378.7084
1997	188.83	190.7564	2021	NA	385.9236
1998	182.91	202.971	2022	NA	400.727
1999	191.09	204.0483	2023	NA	412.8216
2000	230.67	222.7226	2024	NA	419.1813
2001	250.86	235.0232	2025	NA	428.8981
2002	243	233.9053	2026	NA	437.697
2003	251.88	240.534	2027	NA	445.644
2004	261.9	244.957	2028	NA	453.8822
2005	251.57	247.3277	2029	NA	466.2555
2006	265.3	250.3071	2030	NA	476.4641

5. Conclusion

The primary objective of this study was to examine and forecast rice production in Bangladesh. The ARDL model was used to look at the effects of various factors on production. According to this study, rice prices, the distribution of improved seeds, and fertilizer all have a positive and significant lag effect on rice production in both the short and long runs. In this study, it was discovered that the hybrid ARIMA (11, 1, 11) – EGARCH (1,1) – M model outperforms the ARIMA model in terms of RMSE, MAE, and MAPE values. A forecast has been constructed for the years 1983 to 2030 using the ARIMA (11, 1, 11) – EGARCH (1, 1) – M model, which is a linear hybrid ARIMA-GARCH model with respect to mean.

The findings of the study can be used by farm administration to improve post-harvest management, as well as local industries to plan for their raw material needs. For a lead year, hybrid time series forecasting can be used to its full potential in agri-business management.

Acknowledgement

The author is thankful to Miyan Research Institute, International University of Business Agriculture and Technology (IUBAT) to give the resources (software facility, books, journal etc) from the institute.

References

Abate, T., Shiferaw, B., Menkir, A., Wegary, D., Kebede, Y., Tesfaye, K., ... & Keno, T. (2015). Factors that transformed maize productivity in Ethiopia. *Food security*, *7*(5), 965-981.

Adebayo, T. S., Akinsola, G. D., Kirikkaleli, D., Bekun, F. V., Umarbeyli, S., &Osemeahon, O. S. (2021). Economic performance of Indonesia amidst CO2 emissions and agriculture: a time series analysis. *Environmental Science and Pollution Research*, *28*(35), 47942-47956.

Alshammari, T. S., Ismail, M. T., AL-WADI, S., SALEH, M. H., & JABER, J. J. (2020). Modeling and forecasting saudi stock market volatility using wavelet methods. *The Journal of Asian Finance, Economics, and Business*, *7*(11), 83-93.

Asadullah, M., Uddin, I., Qayyum, A., Ayubi, S., &Sabri, R. (2021). Forecasting Chinese Yuan/USD Via Combination Techniques During COVID-19. *The Journal of Asian Finance, Economics and Business*, *8*(5), 221-229.

Asumadu-Sarkodie, S., & Owusu, P. A. (2016). The relationship between carbon dioxide and agriculture in Ghana: A comparison of VECM and ARDL model. *Environmental Science and Pollution Research*, *23*(11), 10968-10982.

Aziz, N., Sharif, A., Raza, A., & Rong, K. (2020). Revisiting the role of forestry, agriculture, and renewable energy in testing environment Kuznets curve in Pakistan: evidence from Quantile ARDL approach. *Environmental Science and Pollution Research*, *27*(9), 10115-10128.

Bahmani-Oskooee, M., &Bohl, M. T. (2000). German monetary unification and the stability of the German M3 money demand function. *Economics letters*, *66*(2), 203-208.

Bhardwaj, D., Ansari, M. W., Sahoo, R. K., &Tuteja, N. (2014). Biofertilizers function as key player in sustainable agriculture by improving soil fertility, plant tolerance and crop productivity. *Microbial cell factories*, *13*(1), 1-10.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307-327.

Ding, C., Duan, J., Zhang, Y., Wu, X., & Yu, G. (2017). Using an ARIMA-GARCH modeling approach to improve subway short-term ridership forecasting accounting for dynamic volatility. *IEEE Transactions on Intelligent Transportation Systems*, *19*(4), 1054-1064.

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The journal of finance*, *48*(5), 1749-1778.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, *48*(5), 1779-1801.

Iqbal, N., Bakhsh, K. H. U. D. A., Maqbool, A. S. I. F., & Ahmad, A. S. (2005). Use of the ARIMA model for forecasting wheat area and production in Pakistan. *Journal of Agriculture and Social Sciences*, *1*(2), 120-122.

Jannah, M., Sadik, K., &Afendi, F. M. (2021, March). Study of Forecasting Method for Agricultural Products Using Hybrid ANN-GARCH Approach. In *Journal of Physics: Conference Series, 1863(1),* 20-52

Kang, S. H., & Yoon, S. M. (2013). Modeling and forecasting the volatility of petroleum futures prices. *Energy Economics, 36 (1),* 354-362.

Koondhar, M. A., Li, H., Wang, H., Bold, S., & Kong, R. (2020). Looking back over the past two decades on the nexus between air pollution, energy consumption, and agricultural productivity in China: a qualitative analysis based on the ARDL bounds testing model. *Environmental Science and Pollution Research*, *27*(12), 13575-13589.

Liu, H., & Shi, J. (2013). Applying ARMA–GARCH approaches to forecasting short-term electricity prices. *Energy Economics*, *37 (5)*, 152-166.

Mohammadi, H., &Su, L. (2010). International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models. *Energy Economics*, *32*(5), 1001-1008.

Narayan, P. K. (2005). The saving and investment nexus for China: evidence from cointegration tests. *Applied economics*, *37*(17), 1979-1990.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society,* 59(2), 347-370.

Nguyen, V. H., &Nguyen, T. P. L. (2020). Intention to accept organic agricultural production of Vietnamese farmers: An investigation using the theory of planned behavior. *The Journal of Asian Finance, Economics, and Business*, *7*(10), 949-957.

Rehman, A., Chandio, A. A., Hussain, I., &Jingdong, L. (2019). Fertilizer consumption, water availability and credit distribution: major factors affecting agricultural productivity in Pakistan. *Journal of the Saudi Society of Agricultural Sciences*, *18*(3), 269-274.

Sentana, E. (1995). Quadratic ARCH models. *The Review of Economic Studies*, *62*(4), 639-661.

Shin, Y.; Yu, B.; and Greenwood-Nimmo, M. (2014). modeling asymmetric cointegration and dynamic multipliers in a nonlinear ARDL framework. In: Sickles, R.; Horrace, W. (eds.) Festschrift in Honor of Peter Schmidt. Springer, New York, NY. https://doi.org/10.1007/978-1-4899-8008-39

Uwilingiyimana, C., Munga'tu, J., &Harerimana, J. D. D. (2015). Forecasting Inflation in Kenya Using Arima-Garch Models. *International Journal of Management and Commerce Innovations*, *3*(2), 15-27.

Wang, B., Liu, P., Chao, Z., Junmei, W., Chen, W., Cao, N., ... & Wen, F. (2018). Research on hybrid model of garlic short-term price forecasting based on big data. *CMC: Comput. Mater. Continua*, *57*(2), 283-296.

Warsame, A. A., Sheik-Ali, I. A., Ali, A. O., & Sarkodie, S. A. (2021). Climate change and crop production nexus in Somalia: empirical evidence from ARDL technique. *Environmental Science and Pollution Research*, *28*(1), 19838-19850.

Wei, Y., Wang, Y., & Huang, D. (2010). Forecasting crude oil market volatility: Further evidence using GARCH-class models. *Energy Economics*, *32*(6), 1477-1484.

Wen, Q., Wang, Y., Zhang, H., & Li, Z. (2019). Application of ARIMA and SVM mixed model in agricultural management under the background of intellectual agriculture. *Cluster Computing*, *22*(6), 14349-14358.

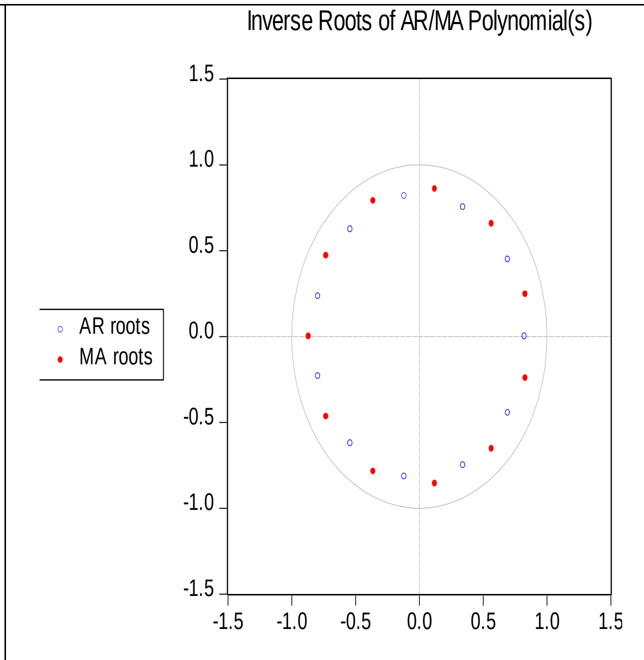
Wulandari, D., Narmaditya, B. S., Prayitno, P. H., Ishak, S., Sahid, S., &Qodri, L. A. (2020). Nexus between production input and price commodity: An integration analysis of rice barns in East Java of Indonesia. *The Journal of Asian Finance*, *Economics*, *and Business*, *7*(10), 451-459.

Yurtkuran, S. (2021). The effect of agriculture, renewable energy production, and globalization on CO2 emissions in Turkey: A bootstrap ARDL approach. *Renewable Energy*, *171*, 1236-1245.Zhu, F. (2011). A negative binomial integer valued GARCH model. *Journal of Time Series Analysis*, *32*(1), 54-67.

Appendix A1

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.072	-0.072	0.2666	0.606
		2	-0.114	-0.119	0.9544	0.621
		3	0.085	0.069	1.3473	0.718
		4	-0.036	-0.039	1.4189	0.841
		5	-0.200	-0.193	3.6934	0.594
		6	-0.124	-0.177	4.5860	0.598
		7	0.139	0.081	5.7368	0.571
		8	-0.033	-0.022	5.8029	0.669
		9	0.072	0.098	6.1233	0.728
		10	0.204	0.161	8.7858	0.553
		11	0.278	0.331	13.881	0.240
		12	-0.096	0.031	14.498	0.270
		13	-0.197	-0.164	17.179	0.191
		14	0.081	0.014	17.643	0.224
		15	-0.149	-0.066	19.265	0.202
		16	-0.036	0.105	19.366	0.250
		17	-0.047	-0.082	19.538	0.299
		18	-0.009	-0.168	19.544	0.359
		19	0.149	0.081	21.387	0.316
		20	0.045	-0.005	21.557	0.365

A1(a): Auto Correlation Function



A1(b): Autoregressive roots