European Journal of Business and Management
ISSN 2222-1905 (Paper) ISSN 2222-2839 (Online)
Vol 4, No.5, 2012

Time Series Modelling of the Contribution of Agriculture to GDP of Bangladesh

Sayedul Anam [Corresponding Author]
Department of Business Administration, Faculty of Business and Economics
Daffodil International University, Dhaka-1207, Bangladesh.
Email: anam_ju@yahoo.com

Mina Mahbub Hossain
Department of Agricultural Statistics, Faculty of Agribusiness Management
Sher-e-Bangla Agricultural University, Dhaka-1207, Bangladesh.
Email: mahbubfhisrt@gmail.com

Abstract

The Gross Domestic Product (GDP) is a basic measure of a country's overall economic performance. Bangladesh is basically an agriculture country and therefore agriculture plays important role on GDP. To assist in the decision making process, this article proposes a time series model based on the contribution of Agriculture to GDP from fiscal year 1972 to 2010. In this paper, we have identified ARIMA (1, 2, 1) model as a reasonable model to forecast the yearly growth rate of GDP of Bangladesh. We also found that the GARCH (1, 1) model with a specified set of parameters is the best fit for our concerned data set. The proposed models help to identify the influence of the agricultural sector on GDP.

Key words: GDP, Stationarity, Stochastic Model, Arima Model, Garch Model, AIC, AICc and BIC.

Introduction

The gross domestic product (GDP) is a basic measure of a country's overall economic performance. It is the market value of all final goods and services made within the borders of a country in a year. It is often positively correlated with the standard of living.

Agriculture sector includes crops, forestry, livestock and fisheries. The main agriculture food products are cereals, sugar, milk, meat, fish, fruits, vegetables, oil, etc. In Bangladesh the major industrial crops are tea, jute, tobacco, etc.

In the survey of 1983-84 the total cultivable area was 2.26 core areas, which decreased to 1.64 core areas in 1995-96. In the agriculture sector, there has been a decreasing growth rate from 1941 to 2009. For instance, in 1950 the agriculture sector contributed to 70.0% of the total GDP but in 2009 the contribution had decreased to 20.6%. This is dreadful for Bangladesh because most of the village people depend on agriculture.

Rahman (2010) fitted an ARIMA model for forecasting Boro rice production in Bangladesh. Palit (2006) had presented a comprehensive review of the evolution of strategic policies, particular analysis of agricultural & industrial reforms, policy implementation stages and efficient uses of human resources for the economic development of Bangladesh.

It was found that impressive progress on gender disparity issues particularly in primary and secondary education, and female economic participation have been achieved despite of economic difficulties and at all levels. Gallo (2004) studied the evolution of GDP disparities among 138 European regions over the period from 1980 to 1995. The results of the analysis indicate the persistence of regional disparities, a progressive bias toward poverty trap and the importance of geography in the convergence process. Schumacher and Breitung (2008) discussed a factor model for short-term forecasting of GDP growth using a large number of monthly and quarterly time series in real-time. The factors were estimated by applying an EM algorithm, combined with a principal components estimator and they proposed alternative methods for forecasting quarterly GDP with monthly factor.

Alam, *et al*. (2009) showed the contribution of agriculture on GDP in different years and compared with other South Asian countries. Karim *et al*. (2008) conducted a study to estimate the contribution made by Bangladesh Agricultural Research Institute (BARI) cereal crops to GDP in the national economy of Bangladesh during the fiscal year 2005-06. In Bangladesh rice is grown in three distinct seasons namely, Boro (January to June), Aus (April to August) and Amon (August to December). Modern rice varieties were introduced for the boro and aus seasons in 1967 and for the amon seasons in 1970. Boro rice is cultivated in nearly 35% of the 10.80 million ha of rice harvested area and it contributed to 50% of the 38.7 million tons of rice produced in 2001/2002.

Due to the immense importance of agriculture in Bangladesh, we undertake time series modelling to determine the contribution of agriculture to GDP.

Methodology

The time series models used in this paper are briefly described. An important parametric family of stationary time series are the Autoregressive Moving Average (ARMA) processes and they play a key role in the modeling of time series data. When a time series is not stationary usually differencing operations are applied at the appropriate lag in order to achieve stationarity. The mean is usually subtracted and a ARMA model is fitted to the data set. A stationary zero mean ARMA (*p*,*q*) model is defined as (see Brockwell and Davis, 2002) a sequence of random variables $\{X_t\}$ which satisfy, $X_t - \phi_1 X_{t-1} - ...... - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q}$ for every t and where $\{Z_t\}$ is a sequence of uncorrelated random variables with zero mean and constant variance σ^2. A process is said to be a ARMA process with mean μ, if $\{X_t - \mu\}$ is an ARMA (*p*, *q*) process. A process is called an ARMA (*p*, *d*, *q*) process if *d* is a nonnegative integer such that $(1-B)^d X_t$ is an ARMA (*p*, *q*) process and where *B* is the usual backward shift operator.

In this study, we also apply the Autoregressive Conditionally Heteroscedastic (ARCH) models which were introduced by Engle (1982) and its extension to GARCH (generalized ARCH) model Bollerslev(1986). In these models, the key concept is the *conditional variance*, that is, the variance conditional on the past. In the classical GARCH models, the conditional variance is expressed as a linear function of the squared past values of the series. Due to the conditional property of GARCH, the mechanism depends on the observations of the immediate past, thus including past variances into explanation of future variances. A process $\{\varepsilon_t\}$ is called a GARCH (*p*,*q*) process (see Francq and Zakoian, 2010) if its first two conditional moments exists and satisfy,

$$E(\varepsilon_t / \varepsilon_u , u < t) = 0, \qquad t \in \mathrm{Z} \quad \sigma_t^2 = \omega + \alpha(B)\varepsilon_t^2 + \beta(B)\sigma_t^2 , t \in \mathrm{Z}$$

Where, *B* is the standard Backshift Operator, and α and β are polynomials of degrees *p* and *q* given as

$$\alpha(B) = \sum_{i=1}^{q} \alpha_i B^i \quad \text{and} \quad \beta(B) = \sum_{j=1}^{p} \beta_j B^j .$$

Our model selection also includes the Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (AICC) and Bayesian Information criterion (BIC). The AIC statistic is defined as, $AIC = -2\ln L + 2(p+q+1)$, where *L* is the Gaussian Likelihood for an ARMA(*p*, *q*) process. On the other hand, the AICC statistic is defined as,

$$AICC = -2\ln L + \frac{2(p+q+1)n}{(n-p-q-2)}.$$

Since, the AICC criterion has a more extreme penalty than the AIC statistics it would counteract fitting very large models. The Bayes Information Criterion (BIC) is given by,

$$BIC = -2\,(\text{Log likelihood}) + p\log(n).$$

In general, BIC penalizes models with more parameters more strongly than AIC.
In this research we used the “forecast”, “fArma” and “fGarch” packages in R programming to analyze the data set. The “forecast”, “fArma” and “fGarch” packages are included in the book

by Cryer and Chan (2008).

Results and Discussion

The data used in this study was obtained from the "Statistical Year Book of Bangladesh (2008)" published by Bangladesh Bureau of Statistics (BBS). By summing over all the districts we find the GDP of Bangladesh. Yearly data on contribution of Agriculture to Gross Domestic Product of Bangladesh from the fiscal year 1972 to 2010 was obtained from the Monthly Bulletin published by the Bangladesh Bureau of Statistics and Economic Trend Central bank of Bangladesh (Bangladesh Bank). The data consisted of 39 yearly observations from 1972 to 2010 and plot of it is shown in Figure 1. Clearly, the time series is not stationary with no apparent seasonality.

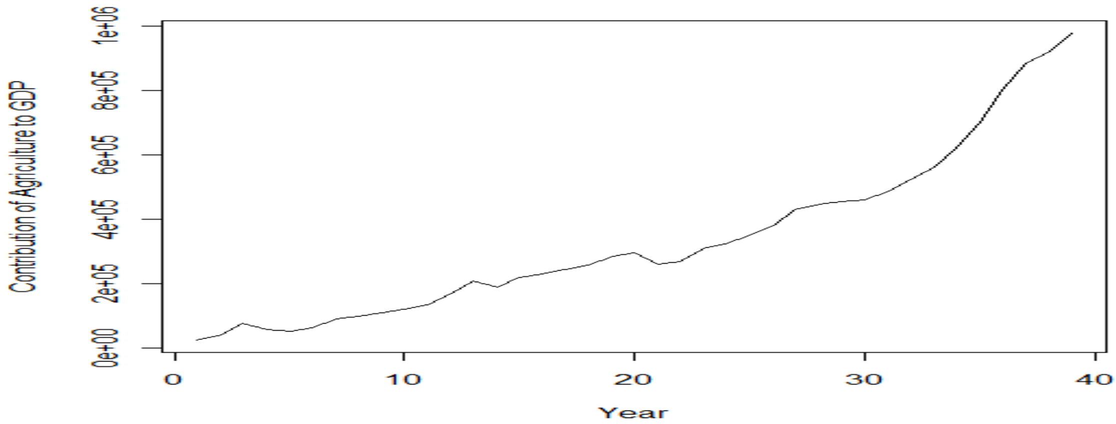

Figure 1: Time Series Plot of Yearly data on contribution of Agriculture to GDP of Bangladesh

We therefore differenced the data once at lag 1 and the plot is shown in Figure 2.

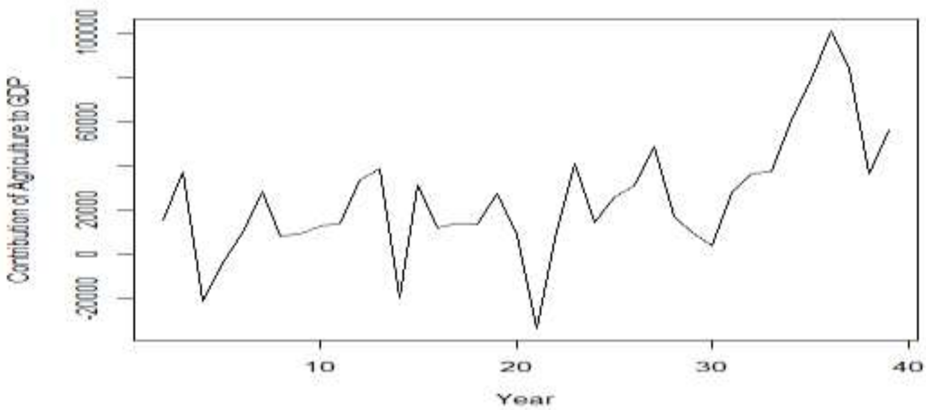

Figure 2: Time Series Plot of Yearly data on contribution of Agriculture to GDP after differencing at lag 1.

This differenced series appeared to be still not stationary and hence we conducted several tests for unit roots. The Augmented Dickey-Fuller (ADF) test statistic value was found to be −2.8784. Since

−2.8784 > −3.18 (at the 0.10 critical level), we cannot reject the unit root hypothesis at the 0.10 level. The Phillips-Peron (PP) (Phillips and Perron, 1988) test statistic value of −3.2031 (> −3.4850, at the 0.10 critical level) also failed to reject the unit root hypothesis. KPSS Test (Kwiatkowski, Phillips, Schmidt and Shin, 1992) test statistic value of 0.1513 (>0.1131, at the 0.10 critical level) also give the same result. As such we differenced the series once again at lag 1. A plot of the differenced series after differencing at lags 2 and is shown in Figure 3.

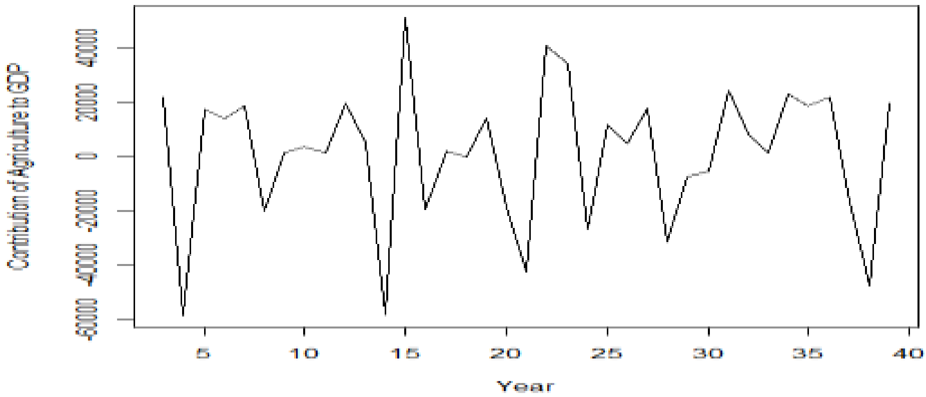

Figure 3: Time Series Plot of Yearly data on contribution of Agriculture to GDP after differencing at lag 2.

This differenced series appeared to be stationary and hence we conducted several tests for unit roots to be certain. The Augmented Dickey-Fuller (ADF) test statistic value was found to be −6.4456. Since −6.4456< −3.18 (at the 0.10 critical level), we can reject the unit root hypothesis at the 0.10 level. The Phillips-Peron (PP) test statistic value of −12.5887(< −3.205575, at the 0.10 critical level) also suggest to reject the unit root hypothesis. The KPSS Test test statistic value of 0.1165 (<.0.1965, at the 0.10 critical level) also rejected the unit root hypothesis. Hence, we can conclude that the series in Figure 3 (after differencing at lags 2) is stationary.

The sample ACF and PACF plots of differenced series at lag 2 are shown in Figure 4.

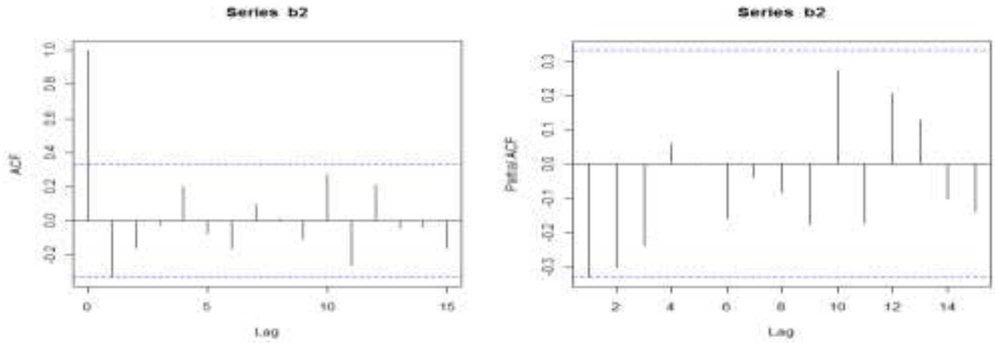

Figure 4: ACF and PACF plots of yearly data on contribution of Agriculture to GDP after differencing at lag 2.

The AIC, AICC and BIC values among the possible models considered in this study are tabulated in Tables 1, 2 and 3, respectively.

Table 1: AIC value of possible ARIMA models (*d*=2)

	0	1	2	3	4
0	809.76	807.12	806.86	808.51	810.22
1	807.89	**806.01**	808.57	809.63	810.34
2	807.01	808.24	809.53	807.28	808.36

3	807.57	809.33	810.79	808.45	810.30
4	809.11	811.08	807.95	809.72	812.44

Table 2: AICC value of possible ARIMA models (d=2)

	0	1	2	3	4
0	809.88	807.49	807.37	809.84	812.29
1	808.26	**805.52**	809.90	811.70	813.34
2	807.78	809.58	811.60	810.28	812.51
3	808.90	811.39	813.79	812.59	815.84
4	811.17	814.08	812.10	815.26	819.64

Table 3: BIC value of possible ARIMA models (d=2)

	0	1	2	3	4
0	811.32	811.23	811.27	814.73	818.00
1	811.41	**808.00**	814.79	817.41	819.67
2	811.67	814.47	817.31	816.61	819.25
3	813.79	817.10	820.12	819.33	822.75
4	816.88	820.41	818.83	822.16	826.44

Hence our fitted ARIMA (1, 2, 1) model is given as,

$$(1-0.1993B)\ (\ (1-B)^2Y_t - 1100) = (1 + 0.6622\mathrm{B})\ Z_t, \qquad (1)$$

where $\{Y_t\}$ is the original time series (i.e. Yearly data on contribution of Agriculture to Gross Domestic Product of Bangladesh a) and $\{Z_t\}$ is white noise with mean 0 and variance 5004. More specifically in equation(1), the estimated value of ϕ = 0.1993 and its standard error was found to be 0.3197, while the estimated value of $\theta = -0.6622$ and its standard error was found to be 0.2622.

To validate our fitted model, several randomness tests on the residuals were performed. The null hypothesis is that the residuals are white noise and hence if the p-value is greater than 0.05, we will not reject the null hypothesis. Since the Ljung-Box randomness test in figure 5 had p-values greater than 0.05, it indicates that the residuals behave like white noise and as such our fitted model seems appropriate.

European Journal of Business and Management
ISSN 2222-1905 (Paper) ISSN 2222-2839 (Online)
Vol 4, No.5, 2012

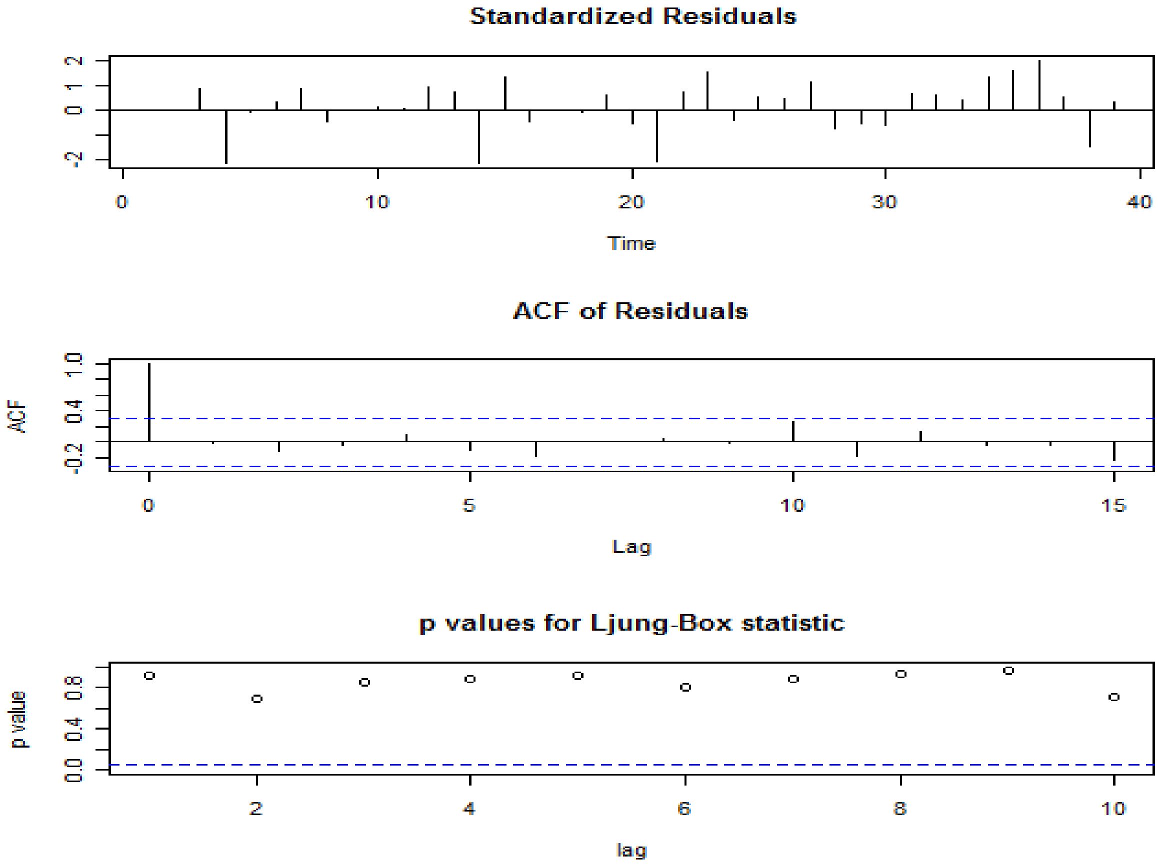

Figure 5: Model Diagnostics Plot

Again using a unique function named "auto.arima" of forecast package in R, we will have another model that may fit our data set. The ARIMA (0, 2, 1) emerged to be another plausible model with an AICC value 805.49 and BIC value 808.23 which is much more than the AIC value 805.12. Hence our fitted ARIMA (0, 2, 1) model is given as,

$$(1-B^2)Y_t - 1100.081 = (1 + 0.5095B)Z_t\,, \qquad \{Z_t\} \sim WN(0, 506294760) \;\ldots\ldots\ldots\ldots\ldots(2)$$

where Y_t is the original time series (i.e. Yearly data on contribution of Agriculture to Gross Domestic Product of Bangladesh). The standard error of the MA coefficient was found to be 0.1625.

To validate our fitted model, several randomness tests on the residuals were performed and the plots are shown in Figure 6. The null hypothesis is that the residuals are white noise and hence if the *p*-value is greater than 0.05, we will not reject the null hypothesis. However, since the Ljung-Box randomness test in Figure 6 had *p*-values smaller than 0.05, it indicates that the residuals of this model may not behave like white noise and as such our fitted model seems inappropriate.

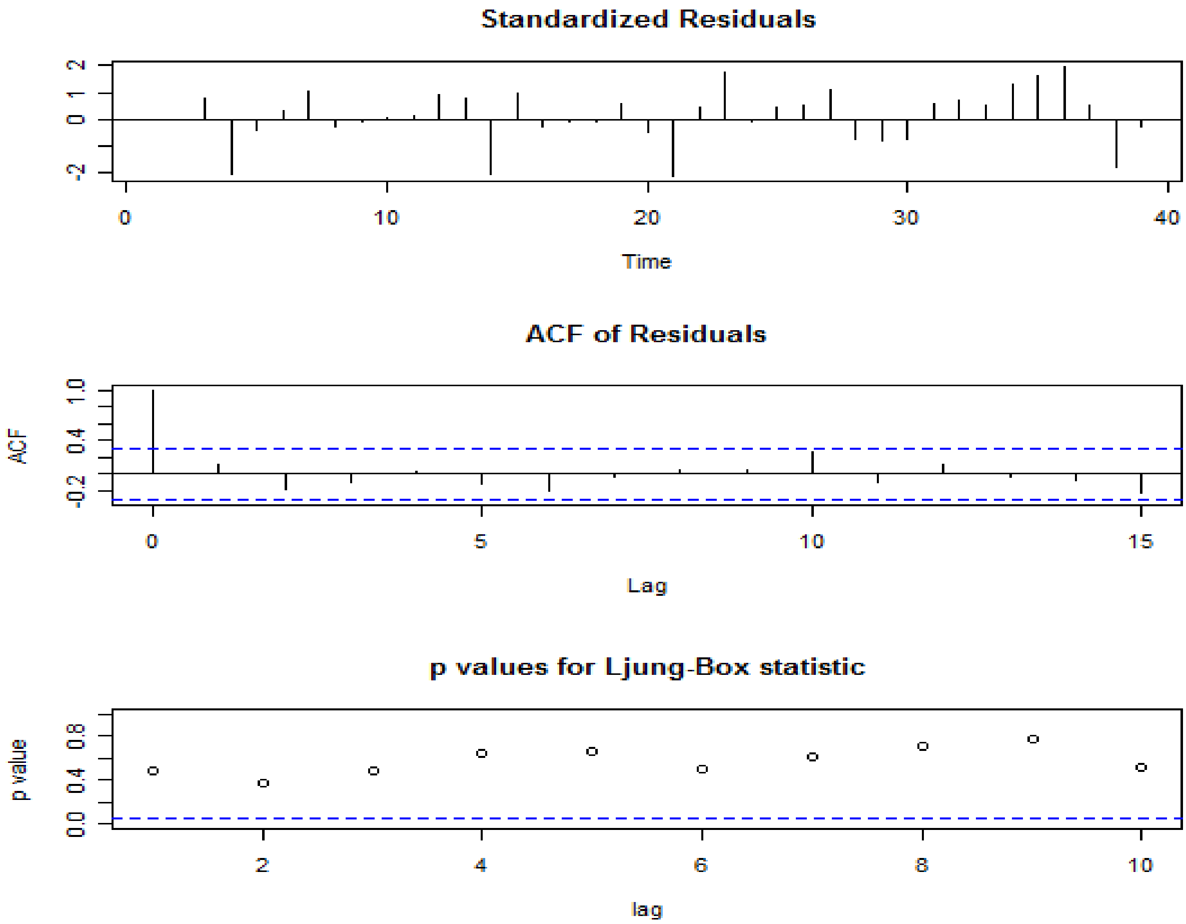

Figure 6: Model Diagnostics Plot

To evaluate the forecasting performance of this model we used the, Mean Error (*ME*), Root Mean Square Error (*RMSE*), Mean Absolute Error (*MAE*), Mean Percentage Error (*MPE*) and the Mean Absolute Percentage Error (*MAPE*), and Mean Absolute Square Error (*MASE)* values are shown in Table 4. Based on all three criteria, clearly we can see that the ARIMA (1, 2, 1) model has smaller values and hence outperforms the ARIMA (0,2,1) model.

Table 4: The computed *ME*, *RMSE*, *MAE*, *MPE*, *MAPE*, and *MASE* values.

Model	ME	RMSE	MAE	MPE	MAPE	MASE
ARIMA(1,2,1)	3622.16	21756.53	16731.51	.0290	8.74	0.59
ARIMA(0,2,1)	3818.24	21884.41	16764.77	0.403	8.98	0.95

Using the model given in (1), we produced forecast from 2010 to 2034 (24 periods) and the forecasted values are given in Table 5.

Table 5: The forecasted values.

Observations	Point Forecast	Lo 80	Hi80	Lo 95	Hi 95
40	1034639	1005776	1063502	990497.3	1078780

European Journal of Business and Management
ISSN 2222-1905 (Paper) ISSN 2222-2839 (Online)
Vol 4, No.5, 2012

41	1090711	1038777	1142644	1011284.6	1170137
42	1146780	1072325	1221236	1032910.1	1260651
43	1202849	1105731	1299968	1054319.9	1351379
44	1258918	1138600	1379237	1074906.8	1442930
45	1314987	1170721	1459254	1094350.8	1535624
46	1371056	1201991	1540121	1112494	1629618
47	1427125	1232365	1621885	1129265.4	1724985
48	1483194	1261826	1704562	1144640.3	1821748
49	1539263	1290373	1788153	1158619.1	1919907
50	1595332	1318017	1872646	1171215.7	2019448
51	1651401	1344771	1958030	1182451	2120350
52	1707470	1370651	2044289	1192349.5	2222590
53	1763539	1395673	2131404	1200937.3	2326140
54	1819607	1419856	2219358	1208241	2430974
55	1875676	1443217	2308136	1214286.9	2537066
56	1931745	1465772	2397718	1219100.6	2644390
57	1987814	1487538	2488091	1222707.1	2752921
58	2043883	1508529	2579237	1225130.1	2862636
59	2099952	1528762	2671142	1226392.5	2973511
60	2156021	1548250	2763791	1226516.3	3085525
61	2212090	1567008	2857171	1225522.3	3198657
62	2268159	1585048	2951270	1223430.6	3312887
63	2324227	1602382	3046073	1220260.2	3428195

A plot of the forecasted values is shown in Figure 7.

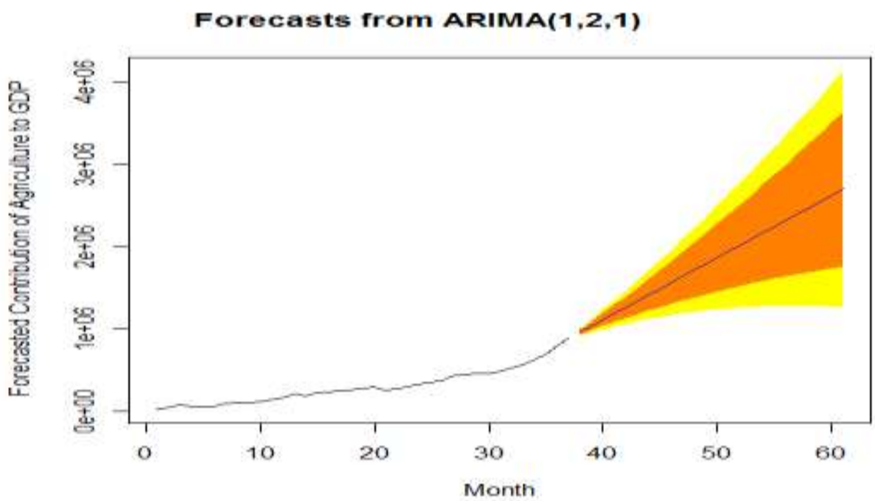

Figure 7: Time Series Plot of Forecasted Yearly data on contribution of Agriculture to Gross Domestic Product of Bangladesh

We now shift gears in an attempt to find some sort of pattern in the data that would suggest the use of a different model. Recall from the previous section the ACF showed that the returns appeared to be random. It even appeared that they were equally likely to be positive or negative from one observation to the next. This implies that the sign of the value of observations is independent of the past; however, the magnitude of change in observations may show correlation. To see this, we squared the observations and produced new ACF/PACF plot on this new data set of squared returns. The ACF/PACF plots are shown in Figure 6. Notice the slow decay of the lag plots which indicates there is correlation between the magnitude of change in the returns. In other words, there is serial dependence in the variance of the data. Figure 6 also show the ACF of the squared returns. These results indicate that a Generalized Autoregressive Conditional Heteroskedasticity(GARCH) model would be a good model to try fitting the data to.

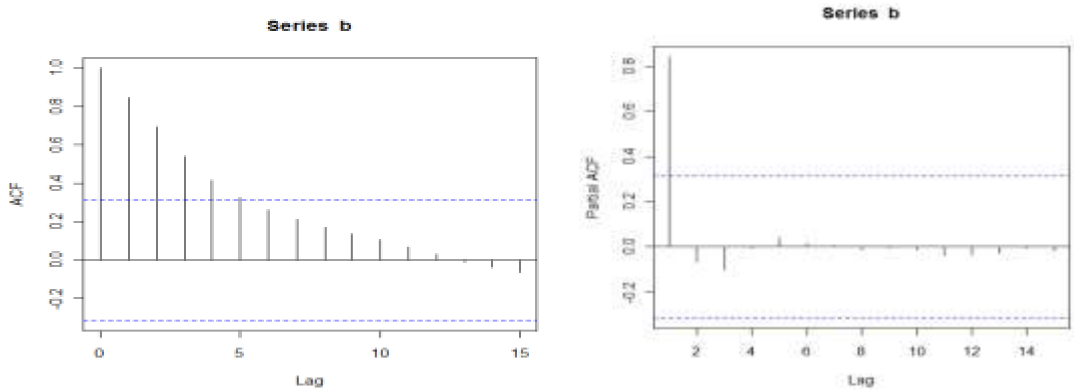

Figure 8: ACF and PACF Plot of Squared Yearly data on contribution of Agriculture to Gross Domestic Product of Bangladesh.

Using computer simulation, we obtain estimates for the parameters of the GARCH (1, 1) process. Here we tried have tried ARMA (0, 1) for our daily data and GRACH (1, 1) process for our model errors. After the modeling we have got two fitted GARCH (1, 1) model for our data set with the following results:

Model -01: ($\omega = 3.334, \quad \alpha = 1, \; \beta = 0.973$)

Mean and Variance Equation: data ~ arma(0, 1) + garch(1, 1)

Coefficient(s):

constant	ma1	omega	alpha1	beta1	shape
5.3330e+00	2.6558e-01	1.2422e-05	4.1481e-02	8.5739e-01	4.4333e+00

Standardised Residuals Tests:

			Statistic	*p*-Value
Jarque-Bera Test	R	Chi^2	45.58016	1.265882e-10
Shapiro-Wilk Test	R	W	0.8724873	0.0002813002
Ljung-Box Test	R	Q(10)	4.06473	0.9443796
Ljung-Box Test	R	Q(15)	4.799463	0.9936924
Ljung-Box Test	R	Q(20)	6.992603	0.9967092
Ljung-Box Test	R^2	Q(10)	6.533969	0.7685861
Ljung-Box Test	R^2	Q(15)	6.911123	0.9600692
Ljung-Box Test	R^2	Q(20)	7.047139	0.9965276
LM Arch Test	R	TR^2	9.686346	0.6434558

Information Criterion Statistics:

AIC	**BIC**	**SIC**	**HQIC**
8.052768	8.152678	8.050667	8.093356

Model-02: ($\omega = 1.2422e-05$, $\alpha = 9.1071e-02$, $\beta = 8.0558e-01$)

Mean and Variance Equation: data ~ arma(1, 0) + garch(1, 1)

Coefficient(s):

constant	Ar1	Omega	Alpha1	Beta1	Shape
3.0259e+00	4.4445e-01	1.2422e-05	9.1071e-02	8.0558e-01	4.7181e+00

Standardised Residuals Tests:

			Statistic	*p*-Value
Jarque-Bera Test	R	Chi^2	20.00259	4.534114e-05
Shapiro-Wilk Test	R	W	0.9067863	0.002646117
Ljung-Box Test	R	Q(10)	1.844062	0.9974016
Ljung-Box Test	R	Q(15)	2.662852	0.9998097
Ljung-Box Test	R	Q(20)	6.083134	0.9987805
Ljung-Box Test	R^2	Q(10)	8.53956	0.5762852
Ljung-Box Test	R^2	Q(15)	9.139002	0.870138
Ljung-Box Test	R^2	Q(20)	9.944842	0.9691611
LM Arch Test	R	TR^2	9.139002	0.6319168

Information Criterion Statistics:

AIC	**BIC**	**SIC**	**HQIC**
4.047526	4.298293	4.011563	4.138842

Now, considering the information criteria, since AIC, BIC, SIC, and HQIC are all show the smallest value, we can say that GARCH (1, 1) with **Mean and Variance Equation:** data ~ arma(1, 0) + garch(1, 1) best fits for our data set with the following parameter values, $\omega = 1.2422e-05$, $\alpha = 9.1071e-02$, $\beta = 8.0558e-01$. Both Figure 7 and the *p*-values concerning the fit of our coefficients indicate a good fit. We also performed diagnostics to rigorously prove our conclusions concerning goodness-of-fit. The residual plots given below also show validity of the model.

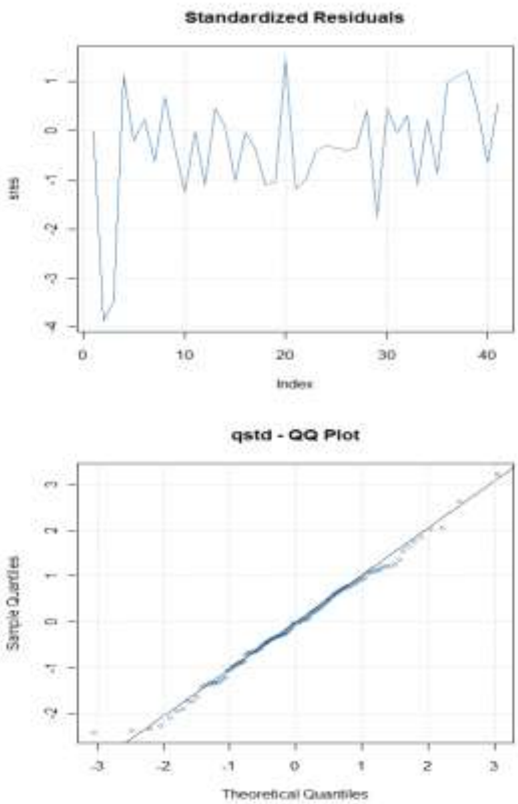

Figure-9: Residual plot of fitted GARCH (1, 1) model.

The simulation using these parameters is shown in Figure 8. We note that the graph contain 500 observations, and are plotted in order to best compare the results. Figure 8 provides further evidence that our data can be modeled by a GARCH (1, 1) model with parameters: ($\omega = 1.2422e-05$, $\alpha = 9.1071e-02$, $\beta = 8.0558e-01$).

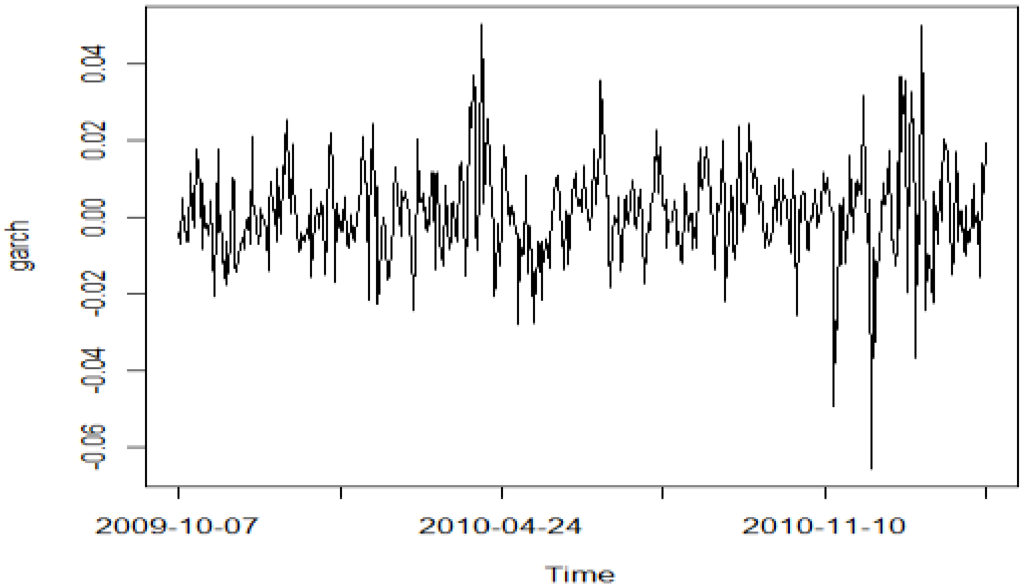

Figure 10: Time plot of Simulated GARCH (1, 1) model with the best fitted model

Conclusion

We closely examined the attributes of our chosen data set concerning the contribution of agriculture to GDP of Bangladesh. Through analysis of the visual properties, we were able to address stationarity and differencing in both the continuous and discrete cases. Particular attention was placed on the mathematical properties of various time series processes in order to determine which would provide the best fit for our data. Various combinations of autoregressive, moving average, and integrated models were considered. We carefully demonstrated the effects of slightly altering the parameters of the more common time series models through numerous simulations and discussed the autocorrelation and partial autocorrelation function and the importance thereof. We meticulously test our claim that the GARCH model with a specified set of parameters is the best fit. We do not merely accept a *p*-value as sole confirmation, but rather gain further confirmation through various statistical tests. These include test for goodness-of-fit and Ljung-Box test for independence. As a further means by which to accurately assess the strength of our conclusions, we simulated a GARCH (1, 1) process with the same specified parameters and five hundreds observations and compare our findings. Furthermore, we clearly show the way to model any time series data by ARIMA and GARCH model. A close look of the plots of our actual and simulated time series data also confirms the steps of a proper time series modeling. The aforementioned rigorous analysis and statistical testing support our conclusions concerning the results.

References

[1] Alam, G. M, Hoque, K. E., Khalifa, M. T. B, Siraj, S and Ghani, M. F., (2009), The Role of Agriculture Education and Training on Agriculture Economics and National Development of Bangladesh., *African Journal of Agricultural Research*, Vol. 12, 1334-1350.

[2] Bollerslev, T., (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307-327.

[3] Brockwell, P. J. and Davis, R. A., (2002), *Introduction to Time Series and Forecasting*, Second Edition, Springer, New York.

[4] Cryer, J. D. and Chan, K-S., (2008), *Time Series Analysis with Applications in R*, 2nd Edition, Springer, New York.

[5] Engle, R. F., (1982), Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50(4), 987-1007.

[6] Francq, C. and Zakoian, J-M, Garch Models: Structure, Statistical Inference and Finance Applications, John Wiley, United Kingdom.

[7] Gallo, J. L., (2004), Space-Time Analysis of GDP Disparities among European Regions: A Markov Chains Approach, *International Regional Science Review*, Vol. 27, No. 2, 138-163.

[8] Karim, M. R., Baset, M. A., Alam, Q. M., Hossain, S., (2008), Contribution of BARI Cereal Crops to the National GDP of Bangladesh. Bangladesh, *J. Agri. Res.*, 33(3): 565-570, December 2008.

[9] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992), Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root, *Journal of Econometrics*, 54, 159-178.

[10] Palit, S. K., (2006), The impact of socio-economic reforms on governance: The Bangladesh experience. *JOAAG*, Vol. 1, No. 1.

[11] Rahman, N. M. F., (2010), Forecasting of Boro Rice Production in Bangladesh: An ARIMA Approach, *J. Bangladesh Agril. Univ*. 8(1): 103-112, 2010.

[12] Phillips, P. C. B and Perron, P. (1988), Testing for Unit Roots in Time Series Regression, *Biometrika*, 75, 335-346.

[13] Schumacher C. and Breitung J., (2008), Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data, *International Journal of Forecasting*, Vol. 24, Issue 3, 386-398.

[14] *Statistical Year Book of Bangladesh*, (2008), Bangladesh Bureau of Statistics (BBS).