Journal of Science and Technology, 6(1 & 2): 137-151, December 2016

MODELING AND FORECASTING PER CAPITA GDP OF BANGLADESH: AN ARIMA MODELING APPROACH

MASHFIQUL HUQ CHOWDHURY* and KH. ASMAUL HOSNA

Department of Statistics, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh

Abstract

In a specified phase of time, the gross domestic product (GDP) is the crucial evaluation of an economy's economic sketch, and refers to the market value of all final services and goods created in the country. This study concerns an analysis on a yearly data of Bangladesh's per capita GDP during 1970-2014 from World Bank data base. The main objective is to find a suitable model for forecasting per capita GDP of Bangladesh; GDP data is checked for whether it is stationary or not through ACF, PACF and using Augmented Dickey-Fuller test. The data was non-stationary but it was transformed to stationary after taking second differences. Box-Jenkins approach for ARIMA model is used. Finally, using the value of widely used model selection criteria, the AIC, ARIMA (2, 2, 2) model has been selected to forecast the per capita GDP for the next sixteen years. Model diagnosis has also been performed to check the adequacy of the model. We hope that the forecast would be hopeful for researchers and policy makers.

Keywords: Gross domestic product, forecasting, ARIMA, Box-Jenkins methodology, Run test.

Introduction

Bangladesh is the most densely populated country in the world, excluding city-states such as Singapore, Bahrain, and the Vatican. According to the results of the 2011 PHC, the population of the country stood at about 149.8 million, with a population density of 1,015 persons per square kilometer (BBS, 2012b). During the past century, the population of Bangladesh has increased exponentially. Between 2001 and 2011, about 19.8 million people were added to the population, which represents a 15 percent increase and a 1.37 percent annual growth rate.

Bangladesh is struggling to emerge from poverty. Bangladesh ranks 146th among nations on the Human Development Index (HDI) as presented in the 2011 Human Development Report (UNDP, 2011). The HPI is a multidimensional measure of poverty for developing countries; it takes into account social exclusion, lack of economic opportunities, and deprivations in survival, livelihood and knowledge. The country's HDI value of 0.500 is above the average of 0.456 for countries in the low human development group and below

* correspondence author: mashfiq3026@gmail.com

the average of 0.548 for countries in South Asia. Countries in South Asia that are close to Bangladesh in its 2011 HDI rank and population size are Pakistan and Nepal, which rank 145th and 157th on the HDI, respectively.

Industry has emerged as the largest sector of the Bangladesh economy, contributing about 30 percent to the gross domestic product (GDP). The GDP exhibited a robust growth rate of 6.7 percent in fiscal year (FY) 2010-2011 compared with 6.1 percent recorded in FY 2009-2010. The overall growth was led by the manufacturing and construction sub-sectors, which recorded impressive expansions of 10 and 6 percent, respectively, in FY 2010-2011. The accelerated growth in these sectors was mainly due to huge investments in large- and medium-scale industry. Agriculture is the second largest sector of the economy, contributing 20 percent to the total GDP in FY 2010-2011. The largest contributor in the agricultural sector is crops and horticulture (11 percent) followed by the fishery sector (4 percent). The average per capita income in Bangladesh has increased from US$599 during FY 2007-2008 to US$848 during FY 2011-2012 (BBS, 2008; MOF, 2012).

Bangladesh aspires to be a middle-income country by 2021. This will require increasing GDP growth to 7.5 to 8 percent per year based on accelerated export and remittance growth. Both public and private investment will need to increase as well. Growth will also need to be more inclusive through creation of productive employment opportunities in the domestic economy. To sustain accelerated and inclusive growth, Bangladesh will need to manage the urbanization process more effectively, as well as prepare for adaptation to climate change impacts.

Background of the Study

The gross domestic product (GDP), is the crucial evaluation of an economy's economic sketch, and refers to the market value of all final services and goods created in the country in a specified phase of time. Gross Domestic Product or (GDP) represents the economic health of a country. It presents a sum of a country's production which consists of all purchases of goods and services produced by a country and services used by individuals, firms, foreigners and the governing bodies. It is the measure of final goods and services of the domestic economy evaluated at market price (Froyen, 1993). The value of the GDP is the value of goods and services determined by the common measuring rod of market prices and are sensitive to changes in the average price level occurring in the economy.

Since the emergence of Bangladesh in 1971 as an independent country, its economy has experiencing modest and reasonably steady annual Gross Domestic Product (GDP) growth rate per over 4.0 per cent. The rate is higher than that the pre-independence era, but somewhat low according to the standard set by contemporary South Asian countries.

Figure 1 shows per capita GDP of Bangladesh, Myanmar, India, and Pakistan. In 1971 (on the eve of Bangladesh's Independence), GDP per capita in Bangladesh (107 US dollars) was more, than GDP per capita in Myanmar (99 US dollars), but was less, than GDP per capita in India (111 US dollars) and Pakistan (222 US dollars). India's per capita income was higher by only 4.0 per cent in current dollar terms; but it increased to over 50 per cent by the year 2000. The gap with Pakistan was initially 64 per cent, but it increased to over 80 per cent by 2000. It would therefore appear that Bangladesh's performance in the area of economic growth has been worse than that of both India and Pakistan.

The accompanying table 1 compares the per capita GDP of Bangladesh, United States, Germany, France, Japan and China. Per capita GDP of Bangladesh and China was almost same at a time 1970-1990. But the GDP of China is suddenly chance at 20th century in 6 times (approximately). On the other hand, in 2013 GDP per capita in Bangladesh was less, than GDP per capita in United States (52392 US dollars) in 53.5 times, Germany (45091 US dollars) in 46 times, Japan (38528 US dollars) in 39.3 times and France (43339 US dollars) in 45 times.

Table 1. Per Capita GDP of Bangladesh and some developed countries

Year	Bangladesh	United States	Germany	France	Japan	China
1970	107	5126	2712	2872	2016	112
1980	233	12436	11958	12752	9378	311
1990	301	23495	21928	21868	25388	347
2000	395	36138	23316	22461	37634	932
2010	758	47925	41100	40617	43151	4375
2013	980	52392	45091	43339	38528	6626

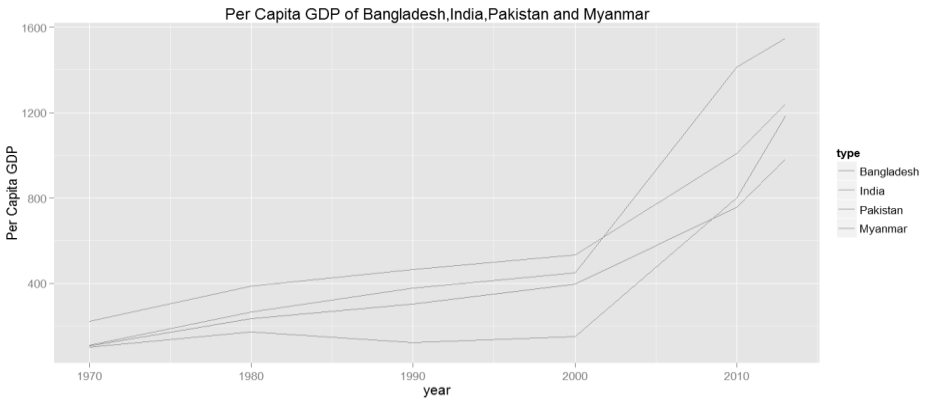

Fig. 1. Per Capita GDP of Bangladesh, Pakistan, India and Myanmar

It is noticeable that during 1970-2014 Bangladesh per capita GDP continuously showing upward trend while during 1990-2000 per capita GDP of World and Asia fluctuate two or three times. In 2014, per capita GDP of World and Asia crosses 5000 US dollars which is a matter of concern because our per capita GDP is still struggling to reach 1500 US dollars. Now the comparison graph among the World, Asia and Bangladesh according to per capita GDP is given below:

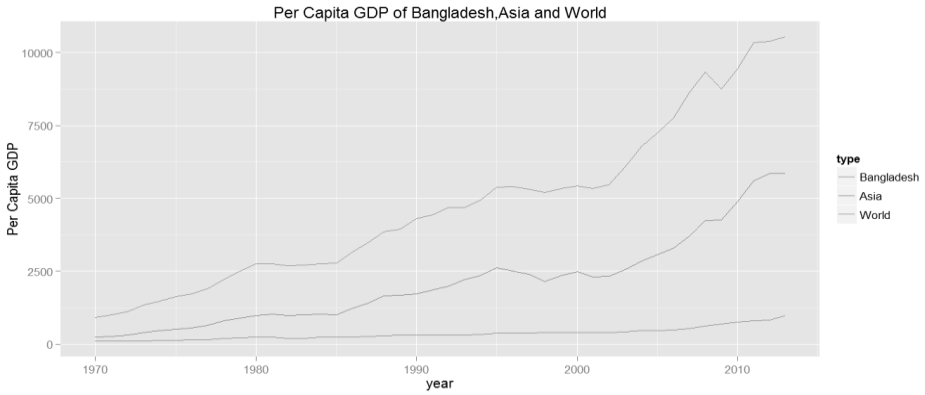

Fig. 2. Per Capita GDP of Bangladesh, Asia and World

In summary, all the studies reported in the Asia, neighbors and World, which in turn motivated this researcher to carry out this research which deals with the GDP issues of Bangladesh. Further, there is not so many studies attempted to forecast the per capita of Bangladesh GDP as well as predicts the growth rates in various forms in Bangladesh. Although some studies attempted to forecast this macro variable only as point estimates which has very little help for the policy makers and managers.

This study is focusing on the time series data of Bangladesh's GDP from 1970 to 2014. It measures the class of ARIMA model. It is aimed to reveal the regularity of Bangladesh's GDP Growth, via the fitted ARIMA model to carry out an out-fit-sample forecasting using Box-Jenkins methodology.

Objectives of the Study

The aim of this study is to apply ARIMA model to analyze the series and use forecast the per capita GDP of Bangladeshi. So the objectives are as follows:

i) to assess the trends of per capita GDP of Bangladesh.
ii) to identify the type of series by employing the correlogram and partial correlogram.
iii) to estimate the parameters of the ARIMA model.
iv) to find a suitable model for predicting per capita GDP of Bangladesh.
v) to compare the actual and predicted value using the fitted model.
vi) to forecast (2015-2030) the per capita GDP of Bangladesh in the near future.

Methodology

Data source

This study uses secondary data from World Bank database, an internationally representative sample survey. The data used in this study is reliable and publicly available. Other published reports from World Bank are also used in order to support the analysis. World Development Indicators (WDI) publication (Published by World Bank) is a collection of time-series data for 214 economies, with many indicators going back more than 50 years. WDI provides cross-country comparable statistics about development and people's lives around the globe. It is divided into six sections: World View, People, Environment, Economy, States and Markets, and Global Links.

Data source: http://data.worldbank.org/indicator/NY.GDP.PCAP.CD

The per capita GDP data from 1970 to 2014 is used for this analysis. The methodology we use in this are given below:

ARIMA model

Autoregressive integrated moving average (ARIMA) model forecasts future values of a time as a linear combination of its own past value and series of errors. In statistics and econometrics, and in particular in time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. These models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting). They are applied in some cases where data show evidence of non-stationarity, where an initial differencing step (corresponding to the "integrated" part of the model) can be applied to reduce the non-stationarity.

A non-seasonal ARIMA model is classified as an "ARIMA (p, d, q))" model, where:

- p is the of autoregressive terms,
- d is the number of nonseasonal differences needed to stationarity, and
- q is the of lagged forecast error in the prediction equation

Modeling procedure

To identity a perfect ARIMA model for a particular data series, Box and Jenkins developed a method known as Box-Jenkins methodology or in short, BJ methodology. Box-Jenkins forecasting models are based on statistical concepts and principles and are able to model a wide spectrum of time series behavior. The series also needs to be at least weakly stationary.

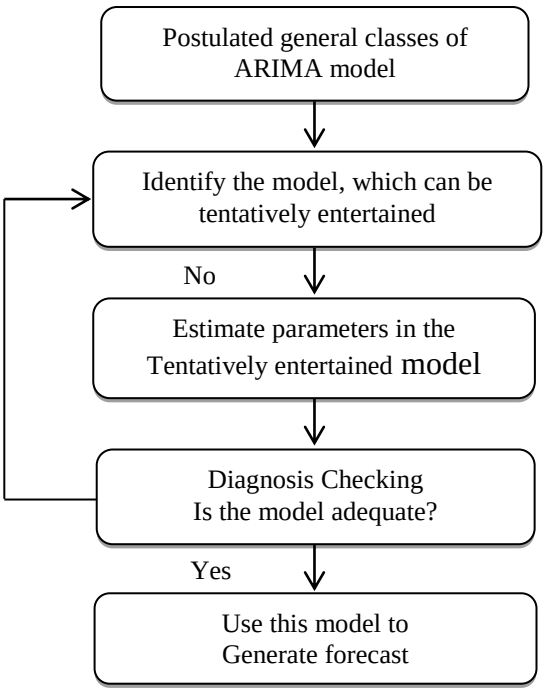

Fig. 3. The Box-Jenkins methodology for optimal model selection

The total process of selecting a model is an iterative procedure consisting of the following three phases: i) identification, ii) estimation and diagnostic checking, and iii) application.

Analysis

Checking stationarity

1970-2014 per capita GDP is considered in figure 4. The time plot presents an increasing trend in our data. It seems that the data are non-stationary. It is not possible to deal with non-stationary data because ARIMA model is defined for stationary data.

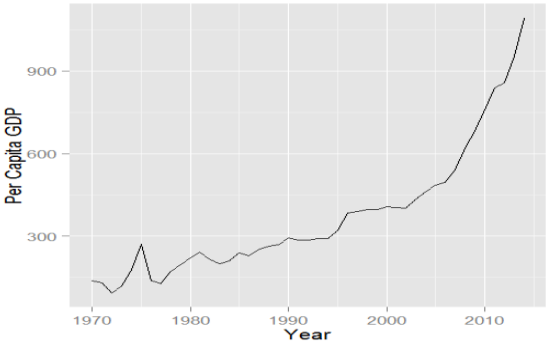

Fig. 4. Per capita GDP during 1970-2014

ACF and PACF Plots

By taking main data set, 1st difference, log transformed and 2nd differences, it is convenient to examine stationarity by using autocorrelation function (ACF) and partial auto correlation function (PACF) graphs (called correlogram) or by using different tests like Dickey-Fuller or Ljung-Box test. Figure 5 - Figure 7 illustrates ACF and PACF plots (no difference, 1st difference and log transformed) show decreasing and for several lags, the spikes fall outside the confidence limit. Thus the reasult shown by the correlogram is that the data are non stationary.

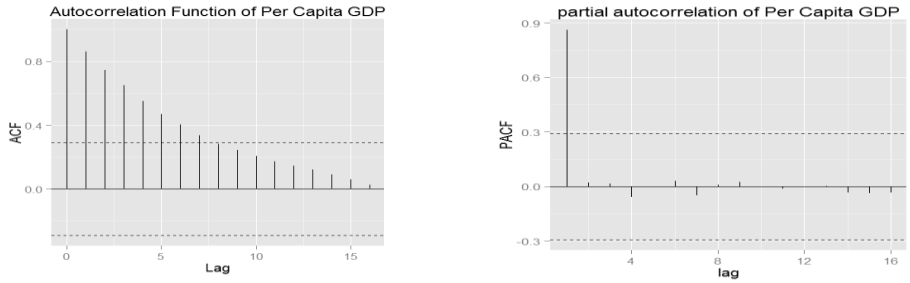

Fig. 5: per capita GDP (no difference)

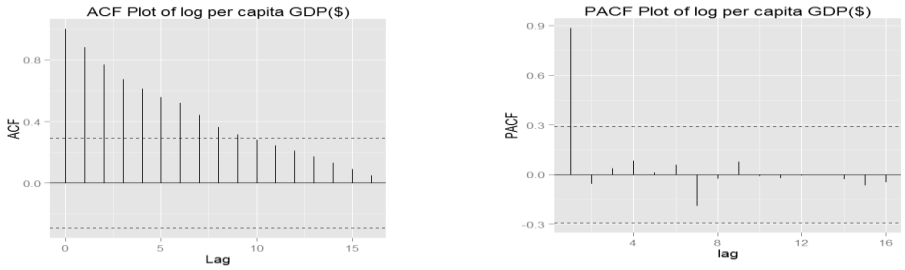

Fig. 6: per capita GDP (log transformed)

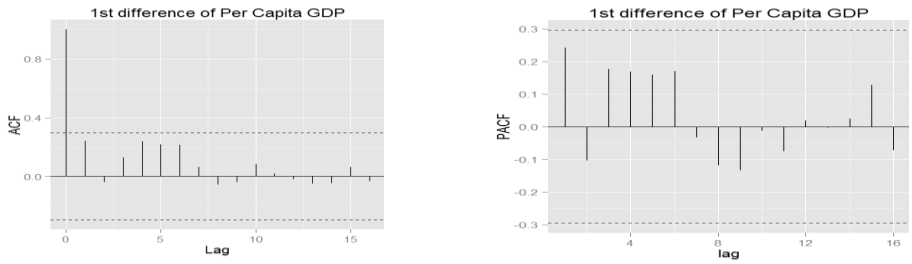

Fig. 7. per capita GDP (first difference)

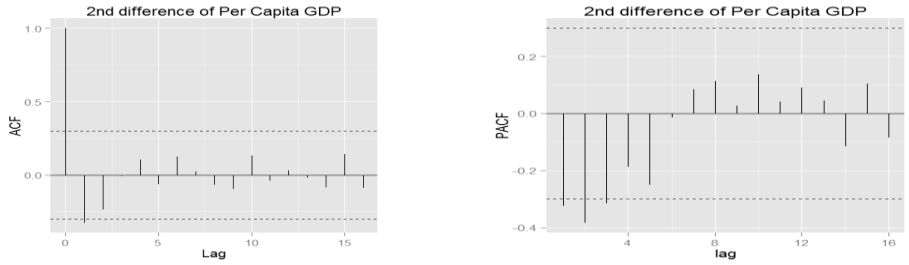

Fig. 8. per capita GDP (2nd difference)

According to Figure 8, the ACF and PACF plots (2nd difference) show that there is no lags and the spikes fall outside the confidence limit. Thus the reasult shown by the correlogram is that the data are stationary when we take 2nd difference of time series data set.

Augmented Dickey-Fuller Test about Stationarity

A statistical test to check stationarity of a time series data set is ADF test. Here, the null and alternative hypotheses are:

H_0 : Data is non-stationary

H_a : Data is stationary

The following table shows the calculated Dickey-Fuller test statistic, Lag order and *p*-value for different differences.

Table 2. Augmented Dickey-Fuller test results about stationarity

Variable	**Lag Order**	**Difference**	**ADF Statistic**	*p*-value	**Decision**
GDP	3	0	3.1472	0.99	Insignificant
Log GDP	3	0	-1.3076	0.8486	Insignificant
GDP	3	1	-2.3125	0.4501	Insignificant
GDP	3	2	-7.3174	0.01	Significant
GDP	3	3	-8.2371	0.04	Significant

For lag order 3, the calculated Dickey-Fuller test statistic for second difference is -7.3174 and the *p*-value is 0.01. For this *p*-value, the null hypothesis can be rejected almost at 1% level of significance .Thus, the test conclude the series as stationary. Since the data are concluded as stationary from the test and plots after second difference so we choose our difference parameter $d = 2$.

Model selection

Among the different plausible models, the best model can be selected by using the minimum Akaike Information Criterion (AIC) and Bayesian information criterion (BIC).

Following table shows the AIC and BIC values for different combinations of p and q, that is, for different ARIMA (p, 2, q) models.

Table 3: Values for different ARIMA (p, 2, q) model

Model	AIC	AICc	BIC
ARIMA(0,2,0)	459.9837	460.0813	461.7449
ARIMA(0,2,1)	445.0953	445.3953	448.6177
ARIMA(0,2,2)	446.7063	447.3217	451.9899
ARIMA(0,2,3)	445.4821	446.5347	452.5269
ARIMA(0,2,4)	447.4648	449.0864	456.2708
ARIMA(1,2,0)	457.3133	457.6133	460.8357
ARIMA(1,2,1)	446.9424	447.5578	452.2260
ARIMA(1,2,2)	448.0222	449.0748	455.0670
ARIMA(1,2,3)	447.4781	449.0998	456.2841
ARIMA(1,2,4)	448.2336	450.5669	458.8008
ARIMA(2,2,0)	451.5883	452.2036	456.8719
ARIMA(2,2,1)	444.6665	445.7192	451.7113
ARIMA(2,2,2)	**442.4125**	**444.0341**	**451.2185**
ARIMA(2,2,3)	444.4113	446.7446	454.9785
ARIMA(2,2,4)	446.4095	449.6095	458.7379

From the above table, we have found that the AIC and BIC value for the model ARIMA (2, 2, 2) is minimum. Hence, our required appropriate order of ARIMA model is ARIMA (2, 2, 2).

Estimation and diagnostic of the residuals

The objective of this study is to develop a general model that will help to predict per capita GDP in near future. Finally, we found our best model which is ARIMA (2, 2, 2). Here, the model includes two (2) auto-regressive, two (2) moving average coefficients and the degree of difference is 2.Thus, the model can be written as

$$\hat{y}_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$
$$\Rightarrow \hat{Y}_t - 2Y_{t-2} + Y_{t-2} = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$
$$\Rightarrow \hat{Y} = 2Y_{t-1} - Y_{t-2} + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$
$$= (2 + \phi_1) Y_{t-1} + (\phi_2 - 1) Y_{t-2} - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

Where,

$Y_t =$ Per capita GDP of Bangladesh

$\theta =$ Moving average coefficients

$e_t =$ Error term

ϕ 's are autoregressive coefficients.

Now, the parameters have to be estimated using the maximum likelihood method. Following table gives the maximum likelihood estimates of the parameters and significance of the parameters.

Table 4. The estimates and significance levels of the parameters of ARIMA (2, 2, 2) model

Parameters	Estimates	Stand Errors	*p*-value	Decision
ϕ_1	0.6209	0.2265	0.006	Highly Significant
ϕ_2	-0.4539	0.1665	0.006	Highly Significant
θ_1	-1.4637	0.2111	0.000	Highly Significant
θ_2	0.7254	0.1764	0.000	Highly Significant

From above table it is seen that, *p*-value of all coefficients are less than 0.05. Therefore, it can be said that all the parameters are highly significant. Thus the model takes the form

$$\hat{Y} = (2 + 0.6209)Y_{t-1} - (1 + 0.4539)Y_{t-2} + 1.4637\, e_{t-1} - 0.7254\, e_{t-2}$$
$$\Rightarrow \hat{Y} = 2.6209\, Y_{t-1} - 1.04539\, Y_{t-2} + 1.4637\, e_{t-1} - 0.7254\, e_{t-2}$$

To ensure that the chosen ARIMA model fits the data well, we have to check the residuals estimated from this model whether or not they are white noise.

Ljung –Box Test for Checking Independence of Residuals

Let,

H_o **:** The residuals are not correlated

H_a **:** The residuals are correlated

For 10 degrees of freedom, the test shows the calculated Chi-squared statistic is -1.6561 and the $p-$ value is 0.0977.So, at 5% level of significance, the null hypothesis cannot be rejected. Hence, it can be said that the residuals are independent of each other.

Following plot (Figure -9) shows the standardized residuals, ACF of residuals and $p-$ values for Ljung-Box statistics taking different lag values that is the diagnostic plot for the model.

From the figure the independence assumption of residuals are also satisfied since none of the ACF values of lag 1 or more is significant. The plot also shows the large $p-$ values for Ljung–Box statistics for different lag values which is also the evidence of independence of residuals.

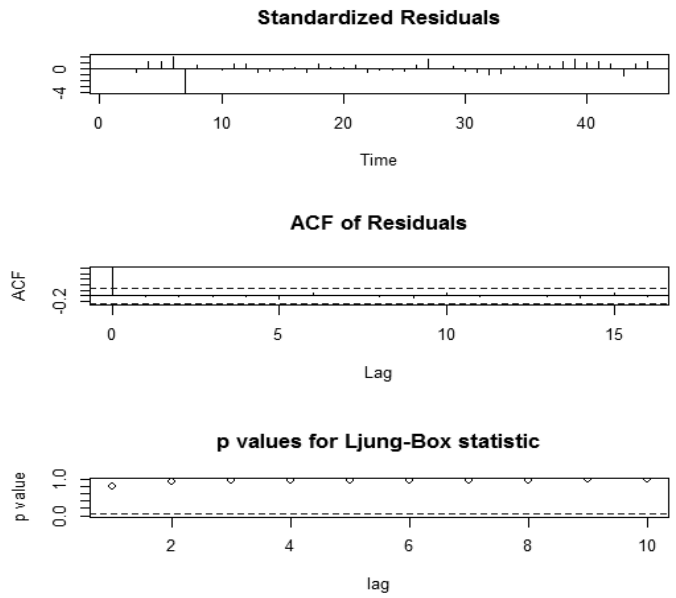

Fig. 9. Diagnostic plot for ARIMA model

Shapiro-Wilk test for Checking Normality Assumption 0f Residuals

In order to check the residuals are normally distributed or not, Shapiro-Wilk test is conducted.

Let,

H_0 : The residuals are normally distributed

H_a : The residuals are not normally distributed

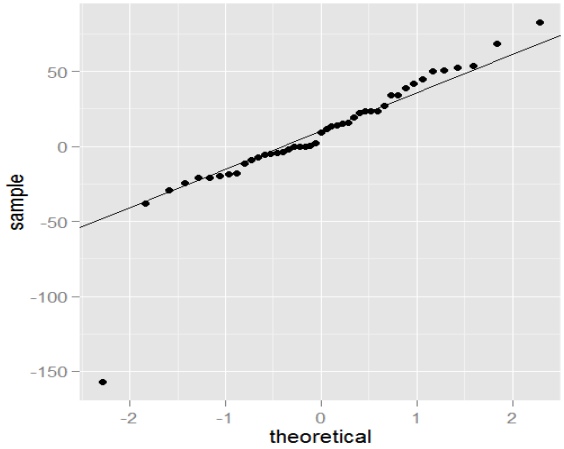

Fig. 10. The q – q plot for the residuals of ARIMA (2, 2, 2) model

Here, after conducting the test, the *p*-value founded is 0.80. Thus, the null hypothesis cannot be rejected. Hence, the residuals can be concluded as normally distributed. The normal $q - q$ plot shows that the residuals are almost normal.

Run Test for checking Randomness of the Residuals

Let

H_o : The residuals are random

H_a : The residuals are not random.

Here, after conducting the test the $p-$value founded is 0.9947.Thus, the null hypothesis cannot be rejected. Hence the residuals are random. Thus, all the diagnostic checks support that the selected model has not only the smallest AIC value but also has the better behaved residuals.

Before forecasting, it is convenient to plot the actual and fitted values for the sample points using the model to check how the model fit for the actual values. Figure 11 illustrates plot of the actual values and the fitted values using the model. From plot, it is seen that the model has a much close fit. We can conclude that the actual and fitted values are very close to each other.

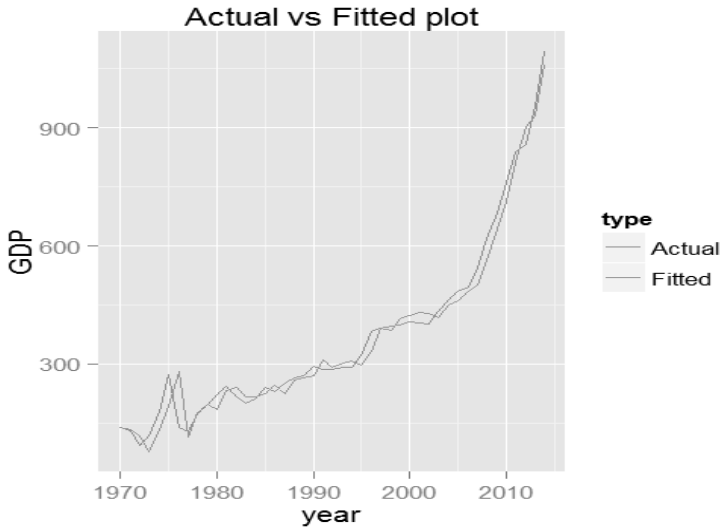

Fig. 11. Actual and Fitted values for ARIMA (2, 2, 2) model

Forecasting

The point forecast with 95% confidence interval of yearly total per capita GDP of Bangladesh for the years 2015 to 2030 by using the selected model is given in Table 5. The forecast values with the 95% confidence intervals are shown in figure 12.

Table 5: Forecasts of per capita GDP for next 16 years (2015-2030)

Time Period	Points Forecasts	95% Confidence Interval
2015	1183.233	(1112.522,1253.944)
2016	1254.086	(1145.935,1362.237)
2017	1334.363	(1202.663,1466.064)
2018	1429.433	(1272.965,1585.901)
2019	1529.402	(1337.087,1721.716)
2020	1625.695	(1386.010,1865.380)
2021	1717.852	(1424.475,2010.496)
2022	1808.151	(1459.641,2156.661)
2023	1900.162	(1494.230,2306.094)
2024	1993.518	(1527.166,2459.870)
2025	2087.098	(1556.749,2617.448)
2026	2180.207	(1582.552,2777.893)
2027	2272.921	(1605.070,2940.772)
2028	2365.605	(1625.080,3106.129)
2029	2458.448	(1642.809,3274.088)
2030	2551.405	(1658.197,3444.614)

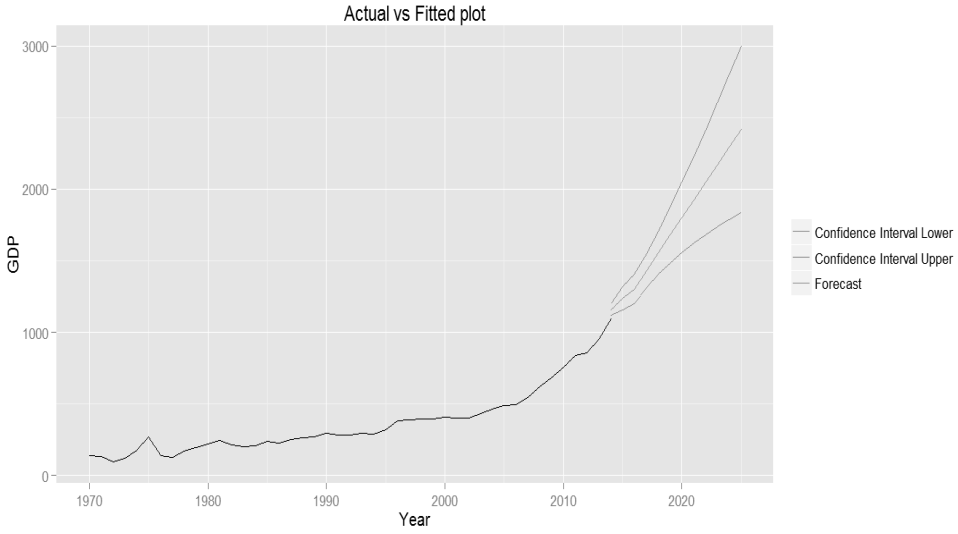

Fig. 12. Forecasts with 95% confidence intervals

Conclusion

The main objectives of this study are to find an efficient forecasting model for per capita GDP in Bangladesh and to forecast it for next few years using that model. To obtain the

objectives Box-Jenkins approach for ARIMA model has been used. The data of per capita GDP for the years 1970-2014 in US dollar has been used.

Initially, non-stationarity has been found in data. This non-stationarity has been removed first by taking second differences. Using the model selection criterion AIC and BIC, ARIMA (2, 2, 2) has been selected. The parameters of this model have been estimated using the maximum likelihood method and all the parameters of this model have been found to be significant assuming normal distributions of the estimators. The normality and independence of the residuals assumption has been checked using different plots and tests. All the result has been found satisfactory. The plot comparing actual values and fitted values using the model shows much close fit. Then the model was used for forecasting purposes. At last the point forecasts for per capita GDP of Bangladesh with the 95% confidence interval for sixteen years along with the data has been shown in a plot and it is found that the upward trend in per capita GDP will still continue for this decade.

References

Adams, Akano S.O. and J.O. (2011), Time series analysis of per capita GDP in Nigeria. *European Journal of Scientific Research*, Vol-58, pp. 30.

Adrian M. Tamayo, Forecasting of the Gross Domestic Product: *An application of Box-Jenkins Methodology.*

Amadi, Uchenna and Aboko, Simon Igboye (2013), Fitting of ARIMA Model in Forecasting Nigeria Gross Domestic Product (GDP), *Journal of Physical Science and Innovation.*

Bipasha Maity and Bani Chatterjee (2012), Forecasting GDP Growth Rates of India: An Empirical Study, *International Journal of Economics and Management Research*, Vol. 1, No. 9, 2012, pp. 52-58.

Box, George E.P. and Jenkins, Gwilym M.(1976), *Time series Analysis Forecasting and Control*, Holden-Day.

Dr. Hasan Yasien Touama, Application of The Statistical Analysis for Prediction of The Jordanian GDP by using ARIMA Time Series and Holt's Linear Trend Models for The Period (2003-2013), *Mathematical Theory and Modeling*, Vol. 4, No. 14, 2014.

Elena-Adriana Andrei and Elena Bugudui, Econometric Modeling of GDP Time Series, *Theoretical and Applied Economics*, Vol. 18(2011), No. 10(563), pp. 91-98.

Haonan Zhang and Niklas Rudholm (2013), Modeling and Forecasting Regional GDP in Sweden Using Autoregressive Models, Business Intelligence Program, School for Technology and Business Studies, Dalarna University.

Ismail Z., Jamaluddin F. (2008), Time Series Regression model for Forecasting Malaysian Electricity Load Demand, *Asian Journal of Mathematics and Statistics,*Vol-1, pp.139-149.

Kishwer Sultana, Adila Rahim, Nighat Moin, Sajida Aman and Saghir Pervaiz Ghauri, (2013) Forecasting Inflation and Economic Growth of Pakistan by Using Two Time Series Methods, *International Journal of Business and Economics Research*, Vol.2, No. 6, 2013, pp. 174-178.

Maniha Zakai (2014), A Time Series Modeling on GDP of Pakistan, *Journal of Contemporary Issues in Business Research,* Vol. 3, No. 4, pp. 200-210.

Mary V.M.Arul Flower and Senthamaraikannan K., Time Series Analysis of Wind Energy Generation, *International Journal of Statistics and Systems,* Vol-5, pp. 465-471.

M. Amin, M. Amanullah and A. Akbar(2014), Time Series Modeling for Forecasting Wheat Production of Pakistan, *The Journal of Animal & Plant Sciences*, pp. 1444-1451.

M. A. Awal and M.A.B. Siddique (2011), Rice Production in Bangladesh Employing by ARIMA Model, *Bangladesh J. Agril.* Res. 36(1): 51-62, March 2011.

V.M.Arul Flower, K. (2010), Time Series of Wind Energy Generation, *Journal of Statistics and Systems.*

Yang Lu (2009), Modeling and Forecasting China's GDP Data with Time Series Models, *D-level Essay in Statistics,* Department of Economics and Society, Högskolan Dalarna, Sweden.

Yaziz, S.R., Ahmad M. H., Nian L. C. and Muhammad N.,(2011) , A Comparative Study on Box-Jenkins and Garch Models in Forecasting Crude Oil Prices, *Journal of Applied Sciences*, Vol-11, pp.1129-1135.