East African Scholars Journal of Engineering and Computer Sciences
Abbreviated Key Title: East African Scholars J Eng Comput Sci
ISSN 2617-4480 (Print) | ISSN 2663-0346 (Online) |
Published By East African Scholars Publisher, Kenya

OPEN ACCESS

Volume-2 | Issue-4 | Apr-May -2019 |

Research Article

Assessing Students' Enrolment in Bolgatanga Polytechnic Using Time Series Analysis

Ibrahim D. Yakubu[1] and John Ayuekanbey Awaab[2]

[1]Bolgatanga Polytechnic, P. O. Box 767, Ghana -West Africa
[2]Department of Secretaryship and Management Studies Bolgatanga Polytechnic, P. O. Box 767, Ghana -West Africa

*Corresponding Author
Ibrahim D. Yakubu

Abstract: This study was conducted with the aim of determining trend of students' enrolment Bolgatanga Polytechnic. Time series analysis was applied on a historic data of students' enrolment for fifteen consecutive years. The study revealed, there is no significant increase in students' enrolment over the past fifteen year due factors such inadequate infrastructure, inadequate lectures halls, inadequate qualified lecturers and not elevating the school to a Technical University. it is observed that students' enrolment are skewed to the right, indicating that most of the enrolment values are concentrated at the left of the mean and this means that majority of the students' enrolment are below the average indicating weak enrolment in the Polytechnic. Again, since the peakness demonstrated a platykurtic flattened with a coefficient of kurtosis = 0.68384 it shows that most of the students' enrolment are at either extreme of the distribution hence flattened than normal peak. The best model that fit students' enrolment in the Polytechnic was quadratic model and ARIMA(1, 0, 0) was used to make five year forecast .These suggested that weak and inadequate strategies by the institution to increase students' enrolment. Government should endeavour and ensure that qualified and adequate lecturers are recruited to meet the future demand of increases in students' enrolment as indicated on forecasted curve. Government and stakeholder should come to the aid of the Polytechnic put-up more lectures halls to ease the pressure during lectures periods. The Polytechnic should have come out with long-term strategic policies to increase students' enrolment as indicated in the findings that no significant increase in students' enrolment over the past fifteen years.
Keywords: Time Series Analysis, Linear trend model, Quadratic trend model, Autocorrelation Function, Partial Autocorrelation Function, Stationarity, Parameter Estimation, Parsimonious model and Differencing.

INTRODUCTION

The development is critical to the progress and industrial development of every nation. The set of skills that empowers a population to contribute meaningfully to the development of a country through their various chosen fields or occupations remains the mission link in Ghana's government for since economic development.it provide advanced technical and vocational education and training the government of Ghana. Redesigned technical institutes in Accra, Kumasi and Takoradi polytechnic in 1963, cape coast polytechnic was established in 1984, tamale polytechnic which was a trade Centre was elevated to a polytechnic in 1986. Similarly Ho polytechnic technical institute was upgraded to a polytechnic in 1986 while that of sunyani and Koforidua were elevated in 1987.subsequently, Bolgatanga and Wa polytechnics were established in 1999 and 2000 to fulfill government policy of established a polytechnic in each of administrative regions of the country at the time.

Polytechnic education has become more critical today than ever as successive governments have seen the need to address the critical staffs gap in the rest for economic development.

Since 1992, when Government directed Polytechnics in Ghana to run Tertiary programmes, significant gains have been made in the output of the Polytechnic graduate. The Polytechnics provide the bulk of our people with technical education that is relevant, up-to-date in technology, and forward looking in approach (Owusu-Agyeman, 2006).

Quick Response Code

Journal homepage:
http://www.easpublisher.com/easjecs/

Article History
Received: 15.04.2019
Accepted: 28.04.2019
Published: 18.05.2019

Countries all over the world are redefining the policies that govern Tertiary education to ensure that all citizens get equal access to Tertiary education. This has been necessitated by the increasing numbers in secondary school enrolment and the need for individual development and survival. The demand for higher education is continually increasing and is triggered off by the increasing number of secondary school leavers, mobility and the presence of other groups looking for second opportunities.

Writing on trend analysis study, Bell and Best (1986) stated that "it is based upon longitudinal consideration of recorded data indicating what has happened in the past, what the present situation reveals and on the basis of this data what is likely to happen in future". Trend studies according to Koul (1995) are "undertaken through documentary analysis or survey at repeated intervals". Some researchers have used the push and pull theory of migration to explain the differences between students' objectives for enrolling in educational programmes and factors that push them to drop out of programmes. Ndudzo and Nyatanga (2013) used the push and pull theory to explore the factors which attract students enrolment as well as outlining the factors which discourage learners from pursuing studies through tertiary institutions. Other drivers include convenience, flexibility fee payment plan and personality and professional development. The major factors which discourage (push) learners to pursue polytechnic education include limited contact time with tutors, nature of induction given to learners, inaccessibility of computers and the internet, lack of financial resources to fund education and unfavorable fees payment schemes.

Concerns about decreasing enrolment has been documented by researchers (Waldrop, 2013). Daily guide (19 march ,2018) has learnt that, the Bolgatanga Polytechnic at Sumbrungu, in the Upper East Region, is currently faced with low enrollment for the 2017/2018 Academic Year.

Problem Statement

The Bolgatanga Polytechnic, established in 1999 is among two Polytechnics in the country that were not elevated to the status of Technical University. Tamale, Sunyani, Kumasi, Koforidua, Ho, Cape-Coast, Accra and Takoradi Polytechnics were elevated, leaving Wa and Bolgatanga behind. Checks revealed that before the conversion some of the polytechnics to technical universities, the Bolgatanga Polytechnic admitted higher numbers of fresh students to read various Higher National Diploma (HND) Courses in the Polytechnic. Unfortunately, the school cannot boast of attracting same number or even higher, fresh students subsequently. These situation has been attributed the several to issues among conversion of polytechnics to technical university, infrastructure, human resource and others issues which would be discovered by this research.

General Objective

The main objective of this study is to determine the trend analysis of students' enrolment in Bolgatanga Polytechnic since inception.

Specific objectives

- To determine the trend of students enrolment in Bolgatanga Polytechnic.
- To develop appropriate model that best fit students' enrolment in Bolgatanga Polytechnic.
- To make a five year forecast of the students enrolment in Bolgatanga Polytechnic to aid planning.

Research Questions

- Is there any trend of students' enrolment in Bolgatanga Polytechnic?
- What is the appropriate model that best fit students' enrolment in Bolgatanga Polytechnic?
- What is the five year forecast of the students' enrolment in Bolgatanga Polytechnic?

Literature Review

Tertiary education is a key factor in a nation's effort to develop a highly skilled workforce for competing in the global economy. There are important private and public benefits to participating in tertiary education. Higher salaries, better employment opportunities, increased savings, and upward mobility are some of the private economic benefits obtained by taking part in tertiary education. A tertiary education graduate also obtains non-economic benefits including, a better quality of life, improved health, and greater opportunities for the future. Given the extensive social and private benefits that result from tertiary education, access and inclusion are essential for achieving social justice and ensuring the realization of the full potential of all young people. First, in the interest of fairness, every individual must be given an equal chance to partake in tertiary education and its benefits irrespective of income and other social characteristics including gender, ethnicity, and language. Second, there is a strong efficiency argument in favor of equity promotion. A talented but low-income student who is denied entry into tertiary education represents a loss of human capital for society. The lack of opportunities for access and success in tertiary education will lead to under-or un-developed human resources.

School choice is an educational issue that has fueled countless, often polarized, debates in North America (Feinberg & Lubienski, 2008). The heated nature of the issue is tied to underlying concepts of democracy and educational equity. Scholars such as Musset (2012) state that the main objective, of making school choice options available for every student is to "level the playing field", allowing more disadvantaged children to access high quality schools they would otherwise not be able to attend (p. 8). Choice options are less available to parents who are socio-economically disadvantaged, and greater awareness among policy makers is needed to consider the needs of diverse families of all income levels (Bosetti, 2004).Matters of school choice are tied closely to issues of declining enrolment -although, as the literature suggests, the correlation is not always direct. With many Toronto schools operating at low capacity and concerns raised about school closures, declining enrolment poses significant challenges for school communities. According to People for Education (2014), 23 of Ontario's 72 boards are under 65% the capacity benchmark set by the province. Moreover, current utilization rates in 10Ontario school boards is less than 50%, compared to others in the Greater Toronto Area (GTA)that operate with over 100% capacity(People for Education, 2012).

The literature reveals inconclusive and often contradictory findings about the effects of school choice on student achievement and educational outcomes. On one hand, scholars (e.g.,Forster, 2013; Card, Dooley & Payne, 2008; Bosetti, 2004) strongly correlate greater choice with better student achievement. On the other hand, scholars weigh in on complex issues spanning school choice around the world (e.g.,Musset, 2012)and share inconclusive findings about the impact of school choice on student achievement outcomes. There are widely held perceptions among educators, parents and advocates that greater school choice levels the playing field and improves academic achievement. However, based on increased school choice options, the literature does not find a direct impact on increased school choice on student achievement and school enrolment.

Despite contentious debates, school choice in canada is stable (Holmes, 2008). Bosetti (2004) aptly sums up the different perspectives on school choice: the pressure for more diversity in schools, more efficiency in schools, greater parental freedom in choosing, and an interest in providing equal opportunity to students of all socio-economic backgrounds. When it comes to school choice in Ghana, however, we actually have a lot more school choice than often perceived by the public. In general, families tend to value their neighbourhood community and school (Bosetti, 2004) and feel they deserve the right to a variety of school choice options (Betts, *et al.,* 2006; Forster, 2013).

From the literature reviewed, we can conclude that there have been many periods of declining enrolment in Ghana especially Bolgatanga Polytechnic, with no "magic wand" solutions to address low enrolment in Bolgatanga Polytechnic. There is no consensus about strategies that address declining enrolment and no obvious solutions. This may be largely due to the nature government policies to elevate all the Polytechnics to Technical Universities.

METHODS

Data and Source

The data for the study was mainly secondary historical annual data of students' enrolment in Bolgatanga Polytechnic since inception, the data spans from 2004 to 2018.

Statistical Technique

Time series is an ordered sequence of values of a variable at equally spaced time interval. It can also be described as a collection of observation made sequentially in time, a set of observations (Y_t) each one being recorded at a specific time (t). Time series occur in a variety of field ranging agriculture to engineering. Many sets of data appear as time series, examples include; hourly observations made on the yield of chemical processes, a monthly sequence of goods sold in a supermarket and so on. Due to the frequent encounter of data of this form methods of analysing time series constitute a great importance in the area of statistics (Anaba and Awaab, 2018).

Objectives of Time Series Analysis

The main objectives of time series analysis are:

Description of Data

When presented with a time series data, the first step in the analysis is usually to plot the data and obtain simple descriptive measures of the main properties of the series such as seasonal effect, trend etc. The description of the data is usually done using summary statistics and or graphical methods like a time plot of the data.

Interpretation of Data

When observations are taken on two or more variables, it may be possible to use the variation in one time series to explain the variation in another series. This may lead to a deeper understanding of the mechanism which generated a given time series. For example, sales are affected by price and economic conditions.

Forecasting of Data

Given an observed time series, one may want to predict the future values of the time series. This is an important task in sales forecasting and in the analysis of economic and industrial time series. Prediction is closely related to control problems in many situations, for examples if we can predict that manufacturing process is going to move off target, the appropriate corrective action can be taken.

Control

A good forecast enables the analyst to take action so as to control a given process, whether it is an industrial process, or an economy or whatever. When a time series is generated which measures the quality of a manufacturing process, the aim of the analysis may be to control the process. In statistical quality control, the observations are plotted on control charts and the controller takes action as a result of studying the charts.

Components of a Time Series

Characteristic movement of time series may be classified into four main types often called components of time series. These four different components are trend, seasonal, cyclical and irregular or random variations.

- Trend (T_t) - The Trend of a time series also known as a secular trend is a relatively smooth pattern, regular and long term movement exhibiting the tendency of growth or decline over a period of time. The trend is that part which the series would have exhibited, has there been no other factors affecting the values.

- Seasonal effects (I_t) - Seasonal variation represents a type of periodic movement where the period is no longer than one year. The factors which cause this type of variation are the climatic changes of the different seasons, such as changes in rainfall, temperature, humidity, etc. and the customs and habits which people follow at different parts of the year. For short, a seasonal component is a pattern that is repeated throughout a time series and has a recurrence period of at most one year.

- Cycles (C_t) - Cyclical fluctuation is another type of periodic movement where the period is more than a year. Such movements are fairly regular and oscillatory in nature. One complex period is called a cycle.

- Residuals (E_t) – Irregular or Random variation movements are such variations which are caused by factors of an erratic nature. These are completely unpredictable or caused by such unforeseen events as flood, earthquake, strike, lockout, etc. It may sometimes be the result of many small forces, each of which has a negligible effect, but their combination effect is not negligible.

The idea is to create separate models for these four elements and then combine them. Generally, time series models are either additive in the form

$$Y_t = \text{Tt} + \text{It} + \text{Ct} + \text{Et} \quad (3.1)$$

Or multiplicative

$$Y_t = \text{Tt} * \text{It} * \text{Ct} * \text{Et} \quad (3.2)$$

Normality Test

The normality test is to investigate the extent to which the data collected approximate a normal distribution. The extent of normality of the data will be determined using skewness and kurtosis.

- Skewness; is the degree to which a data set is not symmetrical. Skewness can be evaluated via the skewness statistic. As data becomes more symmetrical, its skewness value approaches zero. Positive skewed or right skewed data has a value greater than 0 and the tail of such distribution points to the right. The reverse case applies for negatively skewed data.

- Kurtosis; is the degree to which a data set is peaked. Normally distributed data establishes the baseline for kurtosis not too flat or sharply peaked with a statistic of 0. A distribution with a sharper than normal peak will have a positive kurtosis value and is termed leptokurtic distribution, whereas, a platykurtic distribution has a flatter than normal peak and a negative kurtosis value.

Trend Analysis

Trend analysis fits a general trend model, thus, the linear, quadratic or exponential growth models to the time series data. This procedure is often used to fit trend when there is no seasonal component to the series. The trend most accurate to describe the series will be determined using the measures of accuracy, MAPE, MAD and MSD. The model with the minimum measure of accuracy is what best describes the series.

Trend Models

- Linear Trend Model; is estimated using the Ordinary Least Square estimation with a general model of

$$y_t = \beta_0 + \beta_1 t + e_t \quad (3.3)$$

Where y_t is the projected value of the y variable for a selected value of t, β_o is the constant intercept; β_1 represents the average change from one period to the next.

- Quadratic Trend Model; which accounts for a simple curve is of the form

$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + e_t \quad (3.4)$$

- Exponential Growth Trend Model; accounts for exponential growth or decay. Mathematically,

$$y_t = \beta_0 * \beta_1^t * e_t \quad (3.5)$$

Measures of Accuracy

Three measures of accuracy of the fitted model are computed, MAPE, MAD, and MSD for each of the simple forecasting and smoothing methods. For all three measures, the smaller the value, the better the fit of the model. We use these statistics to compare the fits of the different methods.

- Mean Absolute Percentage Error (MAPE); measures the accuracy of fitted time series values, specifically in trend estimation. It usually expresses accuracy as a percentage and is defined by,

$$MAPE = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{At - Ft}{At}\right| \qquad (3.6)$$

Where A_t is the actual value, F_t equals the fitted value, and n equals the number of observations.

- Mean Absolute Deviation (MAD); expresses accuracy in the same units as the data, which helps conceptualize the amount of error. The mean deviation is a measure of how much the fitted value of the data is likely to differ from the actual value. The absolute value is used to avoid deviation with opposite sides cancelling each other out. Its mathematical form is,

$$MAD = \frac{1}{n}\sum_{t=1}^{n}|A_t - F_t| \qquad (3.7).$$

- Mean Squared Deviation (MSD); measures the square forecast error, error variance and also recognize that longest errors are disproportionately more expensive than small errors. It is expressed mathematically as,

$$MSD = \frac{1}{n}\sum_{t=1}^{n}|A_t - F_t|^2 \qquad (3.8).$$

Autocorrelation Function

Autocorrelation is the correlation (statistical relation) between observations of a time series separated by k time units such that systematic changes in the value of one variable are accompanied by systematic changes in the other. The plot of autocorrelations is called the autocorrelation function or correlogram. The ACF is extremely useful in helping to obtain a partial description of the process for developing a forecasting model. ACF is mathematically the proportion of the autocovariance of y_t and y_{t-k} to the variance of a dependant variable y_t.

$$ACF(k) = \frac{cov(y_t, y_{t-k})}{var(y_t)} \qquad (3.9)$$

Partial Autocorrelation Function

Partial autocorrelation function measures the degree of association between Y_t and Y_{t+k} when the effect of other time lags on Y are held constant. In other words, PACF is the simple correlation betweeny_t and y_{t-k} minus the part explained by the intervening lags.

$$PACF = corr(y_t, y_{t-k}|y_{t-1}, y_{t-2}, \ldots, y_{t-k+1}) \qquad (3.10)$$

Stationarity

A stationary process has a mean and variance that do not change over time and the process does not have trends. To proceed with the estimation of an ARIMA model, the series is required to be stationary and to test for stationarity under this study we consider the Augmented Dickey-Fuller (ADF) test and the Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test.

Augmented Dickey-Fuller Test

For the ADF test, we test the hypothesis that;

- H_0: the series is not stationary.
- H_1: the series is stationary.

At 95% significance level, a p-value less than 0.05 means we reject H_0 meaning the series is stationary, else it is not stationary

Kwiatkowski, Phillips, Schmidt and Shin Test

The KPSS test has a reverse hypothesis to the ADF test hence;

- H_0: the series is stationary.
- H_1: the series is not stationary.

This means that at 95% significance level, a p-value less than 0.05 means we reject H_0 and say the series is not stationary, otherwise it is stationary.

Achieving Stationarity

To achieve stationarity, the series has to be differenced till it is stationary. However when the variability of the data increases over time, and has an exponential growth trend, the series can be log transformed before differencing to stabilise the variance. After each differencing, the tests for stationarity (ADF and KPSS tests) have to be performed again to ensure that the series is stationary.

Model Identification

After the series has been made stationary, the next step is to identify which model best describes the series. At this stage we decide how many autoregressive (p) and moving average (q) parameters are necessary to yield an effective but still parsimonious model of the process (parsimonious means that it has the fewest parameters and greatest number of degrees of freedom among all models that fit the data). In practice, the numbers of the p or q parameters very rarely need to be greater than 2 and the primary tools for doing this are the ACF and the PACF. The sample autocorrelation plot and the sample partial autocorrelation plot are compared to the theoretical behaviour of these plots shown below.

Table 1: Theoretical Behaviour of the ACF and PACF for Model Identification

Process	ACF	PACF
AR(p)	Tails off	Cut off after the order p of the process
MA(q)	Cut off after the order q of the process	Tails off
ARMA(p, q)	Tails off	Tails off

Autoregressive Process (Ar)

In statistics, an autoregressive (AR) model is a type of random process which is often used to model and predict various types of natural and social phenomena. Autoregressive models are based on the idea that the current value of the series, y_t, can be explained as a function of p past values, y_{t-1}, and y_{t-2}. . . y_{t-p}, where p determines the number of steps into the past needed to forecast the current value. Mathematically, a time series autoregressive model is given by:

$$Yt = \phi 1\, yt - 1 + \phi 2\, yt - 2 + \ldots + \phi p\, yt - p + et \qquad (3.11)$$

Where e_t is assumed to be a white noise process, ϕ_1, ϕ_2, …, ϕ_p are the autoregressive model parameters and $1<\phi<1$ for all p. Each observation is made up of a random error component (e_t) and a linear combination of prior observations. (Anaba and Awaab, 2018)

Moving Average Models (MA)

The moving average process expresses the current value of the observation in terms of the past shocks or residuals. This means that condition on the past values of the residuals, the future values of the series can be predicted. The notation MA (q) refers to the moving average model of order q. *A* moving average model of order q, abbreviated MA (q), is defined mathematically as:

$$Y_t = \mu + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \qquad (3.12)$$

Where μ is the mean of the series e_t, e_{t-1}… are white noise error terms and θ_1… θ_q are the parameters of the model.

That is, a moving average model is conceptually a linear regression of the current value of the series against previous (unobserved) white noise error terms or random shocks. The random shocks at each point are assumed to come from the same distribution. (Anaba and Awaab, 2018)

Autoregressive Moving Average Model (ARMA)

Autoregressive Moving Average (ARMA) models are typically applied to auto correlated time series data. Given a time series of data Y_t, the ARMA model is a tool for understanding and, perhaps, predicting future values in this series. The model consists of two parts, an autoregressive (AR) part and a moving average (MA) part. The model is usually then referred to as the ARMA (p, q) model where p is the order of the autoregressive part and q is the order of the moving average part. The notation ARMA (p, q) refers to the model with p autoregressive terms and q moving average terms. This model contains the AR (p) and MA (q) models. (Anaba and Awaab, 2018)

An Auto-Regressive Moving Average model of order (p, q) abbreviated as ARMA (p, q), is defined mathematically as

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q \theta_{t-q} \qquad (3.13)$$

Where Y_t is a mixed autoregressive moving average process of orders p and q abbreviated ARMA (p, q).

Autoregressive Integrated Moving Average Models (ARIMA)

If a non-stationary time series which has variation in the mean is differenced to remove the variation the resulting time series is known as integrated time series. It is called integrated because the stationary model which is fitted to the differenced data as to be summed or integrated to provide a model for the non-stationary data. All AR (p) models can be represented as ARIMA (p, 0, 0) that is no differencing and no MA (q) part, also MA (q) models can be represented as ARIMA (0, 0, q) meaning no differencing and no AR (p) component.

The general model is ARIMA (p, d, q) where p is the order of the AR part, d is the degree of differencing and q is the order of the MA part. The general ARIMA (p, d, q) model can be expressed in terms of the backward shift operator as

$$(1-\varnothing_1 B-\varnothing_2 B^2-\ldots-\varnothing_P B^p)(1-B)^d Y_t=(1+\theta_1 B+\theta_2 B^2+\ldots+\theta_q B^p)e_t \qquad (3.14)$$

Where $(1\text{-}B)^d$ is the non-seasonal differencing filter. (Anaba and Awaab, 2018)

Information Criteria

The following tools (criteria) are used in the selection of best fit model out of suggested models. The model with the minimum of these statistics is selected as the best fit.

- Akaike's Information Criteria (AIC); uses the maximum likelihood method. In the implementation of the approach, a range of potential ARMA models is estimated by maximum likelihood methods, and for each, the AIC is calculated, given by:

$$AIC(p,q) = ln(\sigma_e^2) + r\frac{2}{n} + constant \qquad (3.15)$$

Where, n is the number of observations in the historical time series data, σ_e^2 is the maximum likelihood estimate of σ_e^2, and $r = p + q + 1$ denotes the number of parameters estimated in the model. Given two or more competing models the one with the smaller AIC value will be selected.

- Schwarz's Bayesian Criterion (BIC); like AIC uses the maximum likelihood method. The BIC imposes a greater penalty for the number of estimated model parameters than does the AIC. The use of minimum BIC for model selection results in a chosen model whose number of parameters is less than that chosen under AIC. It is given by,

$$BIC(p,q) = \ln(\sigma_e^2) + r\frac{\ln(n)}{n} \qquad (3.16)$$

- Corrected Akaike's Information Criteria (AIC_C); the AIC is biased estimator and the bias can be appreciable for large parameter per data ratios. Hurvich and Tsai (1989) showed that the bias can be approximately eliminated by adding another non – static penalty term to the AIC, resulting in the corrected AIC, denoted by AIC_C and defined by the formula:

$$AIC_C = AIC + \frac{2(r+1)(r+2)}{n-r-2} \qquad (3.17)$$

Parameter Estimation

Once a model is identified the next stage of the ARIMA model building process is to estimate the parameters. Estimating the parameters for the ARIMA (Box- Jenkins) models is a quite complicated non-linear estimation problem. For this reason, the parameter estimation was done using a statistical package called gretl. (Anaba and Awaab, 2018)

Model Diagnosis

To ensure that the selected model is the best model that suits the data the following diagnostics are performed.

- Time Plot of the Residuals; is a plot of the standardized residuals against time. For a fit model, it should not show any fixed pattern, trend in the residuals, no outliers and in general case no changing variance across time. (Anaba and Awaab,2018)

- Plot of Residual ACF; allows one to examine the goodness of fit by means of plotting the ACF of residuals of the fitted model. If most of the sample autocorrelation coefficients of the residuals are within the 5% significance limits in a random pattern, then the model is a good fit.

- The Normal Q-Q Plot; is another diagnostic check on the residuals to determine whether it follows the normal distribution. This is done by using the normal probability plot Q–Q plot. It is a plot of the quantiles of two distributions against each other, or a plot based on estimates of the quantiles. The normal Q-Q plots is used to compare the distribution of a sample to a theoretical distribution. If most of the points are in line and closer to the normal line, then the model is a good fit.

- Ljung-Box Q Statistics; is a check of the overall model adequacy. The error terms are examined and for the model to be adequate the errors should be random. If the error terms are statistically different from zero, the model is not adequate. The test statistic used is the Ljung-Box statistic, a function of the accumulated sample autocorrelations, r_j, up to any specified time lag *m*. As a function of m, it is determined as:

$$Q(m) = n(n+2)\sum_{j=1}^{m}\frac{r_j^2}{n-j} \qquad (3.18)$$

Which is approximately chi-square distributed with n-p-q degree of freedom. Here p and q are orders of AR and MA respectively and n is the number of usable data points after any differencing operations. This statistic can be used to examine residuals from a time series model in order to see if all underlying population autocorrelations for the errors may be 0 (up to a specified point). If the corresponding p-value is greater than 0.05, then the model is considered adequate (Anaba and Awaab, 2018).

Forecasting

Once an appropriate time-series model is selected and established fit, we can now forecast future values of the series. Once a forecast is made for y_{T+1} it is added to the series and used to forecast for y_{T+2}. The process continues into the desired future for which a forecast is desired which for this study is the next five years.

The Box-Jenkins Method of Modelling Time Series

The Box-Jenkins methodology (Box & Jenkins, 1976) is a step-wise statistical method used in analysing and building forecasting models which best represents a time series. This method of forecasting implements knowledge of autocorrelation analysis based on autoregressive integrated moving average models.

The methodology has the following advantages;

- It is logically and statistically accurate
- It makes great use of historical time series data
- Forecasting accuracy is increase

The procedure is of four distinct stages namely; Identification, Estimation, Diagnostic checking, Forecasting (Anaba and Awaab, 2018)

- Identification: Identification methods are procedures applied to a set of data to indicate the kind of representational model that will be further investigated. The aim here is to obtain some idea of the values *p, d* and *q* needed in the general linear ARIMA model and to obtain initial estimates for the parameters. The task here is to identify an appropriate subclass of models from the general ARIMA family which may be used to represent a given time series. This requires examining the autocorrelation and partial autocorrelation coefficients calculated for the data.

- Estimation: Once the preliminary model is chosen, the estimation stage begins. The purpose of the estimation is to find the parameter estimates that minimize the mean square error. An iterative non-linear least squares procedure is applied to the parameter estimates of an ARMA (p, q) model. The method minimizes the sum of squares of error given to form the model and data. The estimates usually converge on an optimal value for the parameters with a small number of iterations.

- Diagnosis Checking: Residuals from the fitted model are examined to ensure that the model is adequate (random). Autocorrelation of the error term are estimated and plotted to determine whether they are statistically zero. Thus the observed value is test as a result of sampling error. This is the first test for adequacy. The second test for adequacy is the Q-test as discussed earlier. Under circumstances of unsatisfying results, other ARMA model may be tried until a satisfactory model is obtained (Anaba and Awaab,2018).

- Forecasting: When a model is identified and validated, forecast for one period can be made and there on several periods. As the forecast period becomes further ahead, the chance of forecast error becomes larger. As new observations for a time series are obtained, the model should be re-examined and checked for adequacy. If the time series seem to be changing over time, the parameters of the model should be recalculated or a new model may have to be developed. When small differences in forecast error are observed, only a recalculation of the model parameters is required. However, if larger differences are observed in the size of the forecast error, then a new model is required, thus, returning to the first step of the Box-Jenkins process (Anaba and Awaab, 2018).

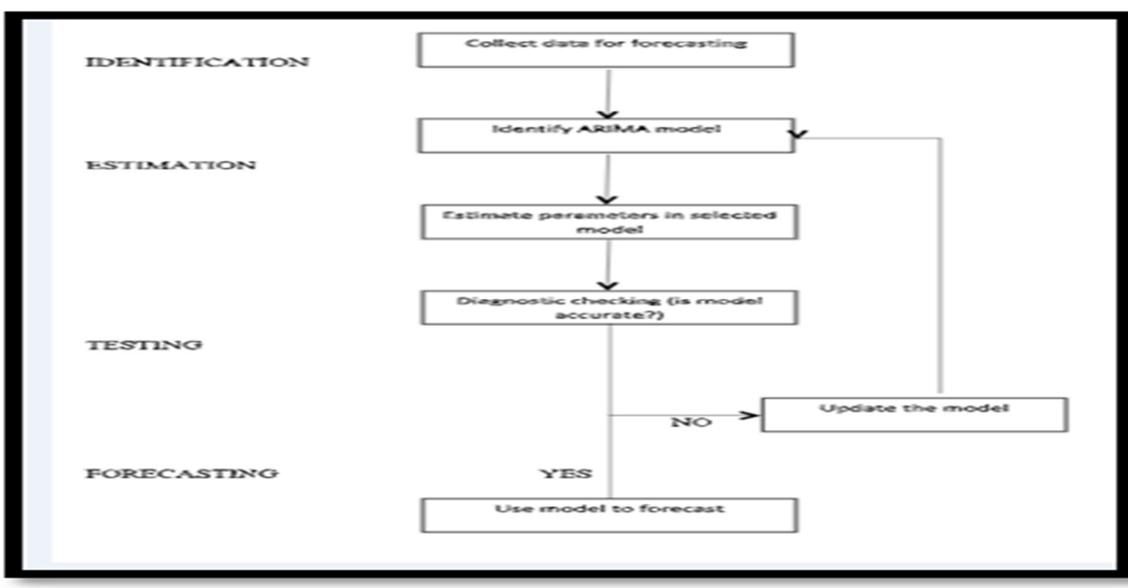


Figure 1: The Box-Jenkins Process

RESULTS AND ANALYSIS

3.1 Exploratory data analysis

An exploratory data analysis on students' enrolment for the fifteen consecutive year period using mainly Minitab software, R programme and Gretl software, and the Box-Jenkins methodology of time series analysis was also employed. Some computations were made to first obtain the descriptive statistics in relation to the students' enrolment, followed by time series plots and a trend analysis.

3.2 Descriptive Statistics of Students' Enrolment

Table 4.1: Descriptive Statistics of students enrolment in Bolgatanga Polytechnic

yearly students enrolment	
Mean	413.3076923
Standard Error	27.84535698
Median	376
Standard Deviation	100.3978624
Sample Variance	10079.73077
CV	6.74%
Kurtosis	-0.683841756
Skewness	-0.251246826
Range	337
Minimum	220
Maximum	557
Sum	5373
Count	13
Confidence Level(95.0%)	60.66982103

The minimum students' enrolment was found to be 220 and maximum 557 whilst the average enrolment was 413.3076923 with accompanying standard deviation of 27.84535698, indicating that the data is widely dispersed across the mean. The coefficient of variation of 6.74% also shows that the data has a very high variance. The students' enrolment distribution also exhibits positive skewness of 0.2512 indicating that most of the enrolment are concentrated to the left of the mean and has a positive kurtosis value of 0.68384 also indicating that the data is platykurtic, thus, has a flattened and there are many of the students enrolment at either extreme of the distribution hence flattened than normal peak. The polytechnic currently run ten programmes with maximum students' enrolment of 557 which is unacceptable and so factors that militate against the students enrolment should be address immediate to safe the school from collapsed.

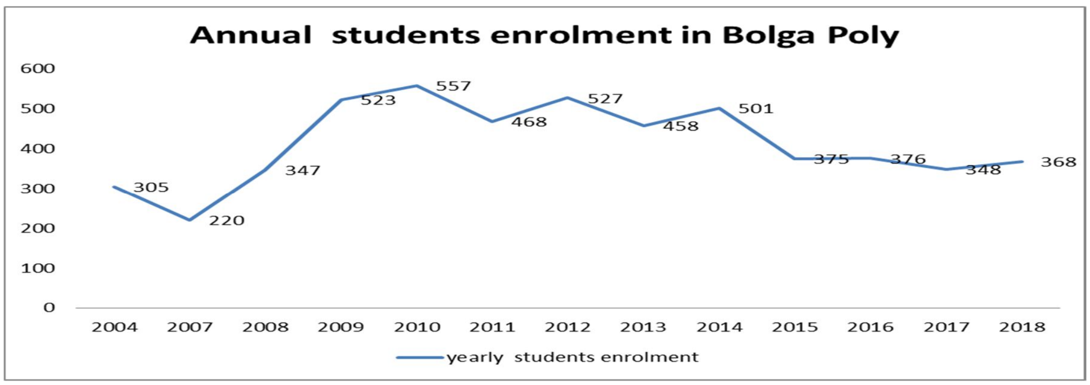


Figure 3.1 Students' enrolment in Bolga Polytechnic

Figure 3.2.1 below shows the data obtained by student's enrolment in Bolgatanga Polytechnic for fifteen-year period. The graph depicts an overall weak enrolment in the polytechnic. It can be seen that students' enrolment depicts an increased from 2007 to 2014 and decreased drastically from 2015 to 2018. In particular, 2017 and 2018 recorded the worse students' enrolment.

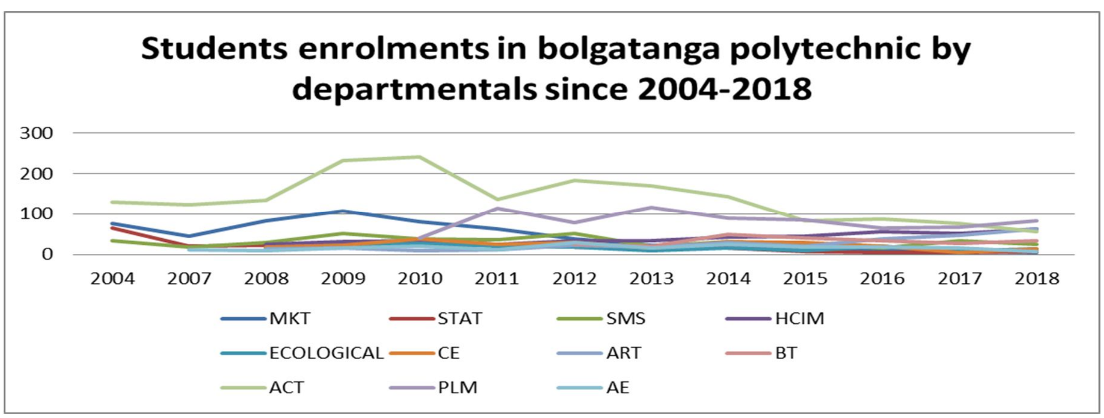


Figure 3.2 students' enrolment in Bolga Poly

From the figure 3.2 it is clear that enrolment based on departmental also experienced similarly fluctuations where Marketing, Statistics and Agricultural Engineering programmes recorded the lowest enrolment as compared other programmes in the Polytechnic

.

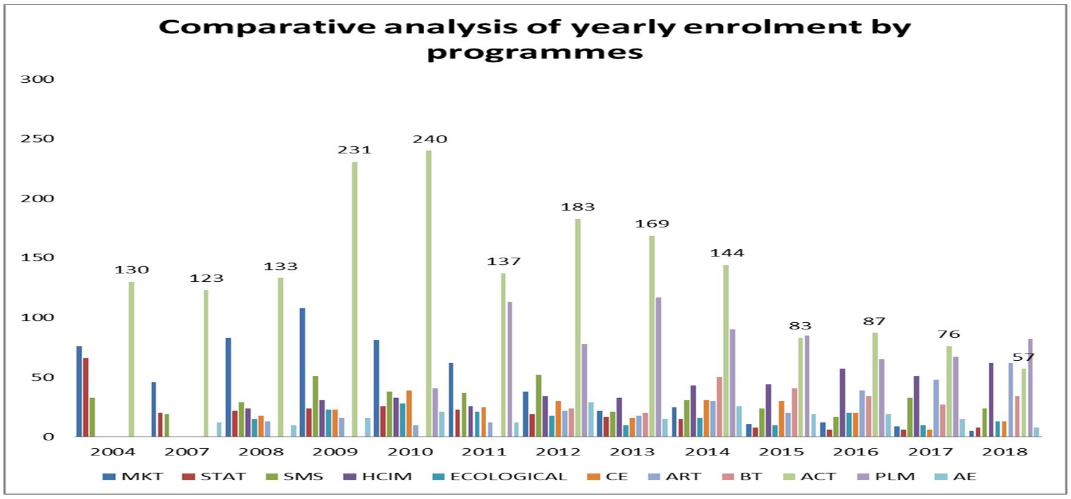
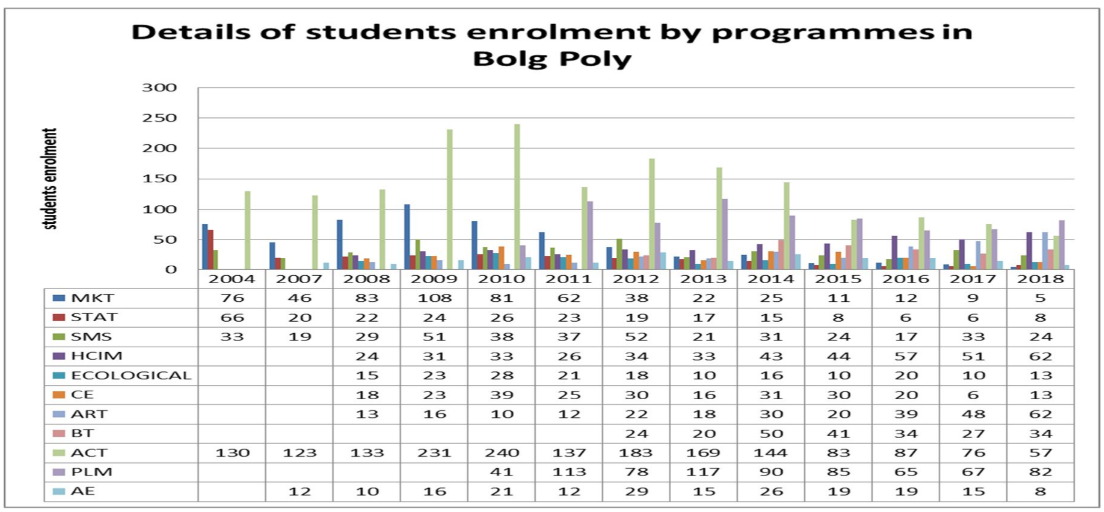


Figure 3.2: Students' enrolment in Bolgatanga Polytechnic

From the figure above Accountancy programme seen to record the highest enrolment throughout the study period as compared to the counterpart Marketing, Statistics and Agricultural Engineering programmes recorded the lowest enrolment in 2018. Below are the details of the individual's programmes and their enrolment in figure 3.3.



	2004	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
MKT	76	46	83	108	81	62	38	22	25	11	12	9	5
STAT	66	20	22	24	26	23	19	17	15	8	6	6	8
SMS	33	19	29	51	38	37	52	21	31	24	17	33	24
HCIM			24	31	33	26	34	33	43	44	57	51	62
ECOLOGICAL			15	23	28	21	18	10	16	10	20	10	13
CE			18	23	39	25	30	16	31	30	20	6	13
ART			13	16	10	12	22	18	30	20	39	48	62
BT							24	20	50	41	34	27	34
ACT	130	123	133	231	240	137	183	169	144	83	87	76	57
PLM					41	113	78	117	90	85	65	67	82
AE		12	10	16	21	12	29	15	26	19	19	15	8

Figure 3.3 yearly students enrolment in Bolga Poly

Table 3.2 means and standard deviation of the students' enrolment

SUMMARY				
Groups	***Count***	***Sum***	***Average***	***Variance***
2004'	11	205	18.63636	830.0545
2007'	11	220	20	1375
2008'	11	347	31.54545	1635.073
2009'	11	523	47.54545	4586.673
2010'	11	557	50.63636	4369.255
2011'	11	468	42.54545	1939.873
2012'	11	527	47.90909	2311.491
2013'	11	458	41.63636	2678.855
2014'	11	501	45.54545	1497.073
2015'	11	312	28.36364	491.4545
Within Groups	237674.9	130	1828.269	
Total	252163.1	142		

Table 3.3: ANOVA

Source of Variation		
Between Groups		
	Within Groups	
Total	252163.1	142

It is cleared From table 3.3 that students enrolment is not statistically significant over the study period .This suggest that the polytechnic has not been able increased students' enrolment over the past fifteen year and these pattern has associated with many factors such infrastructure, pull and push factors by parents, inadequate teaching and non-teaching staffs and government policy of elevating some of the polytechnics to Technical Universities but others were left-out due to the factors identified by the research findings which Bolgatanga Polytechnic was part.

Trend Analysis of Students' Enrolment

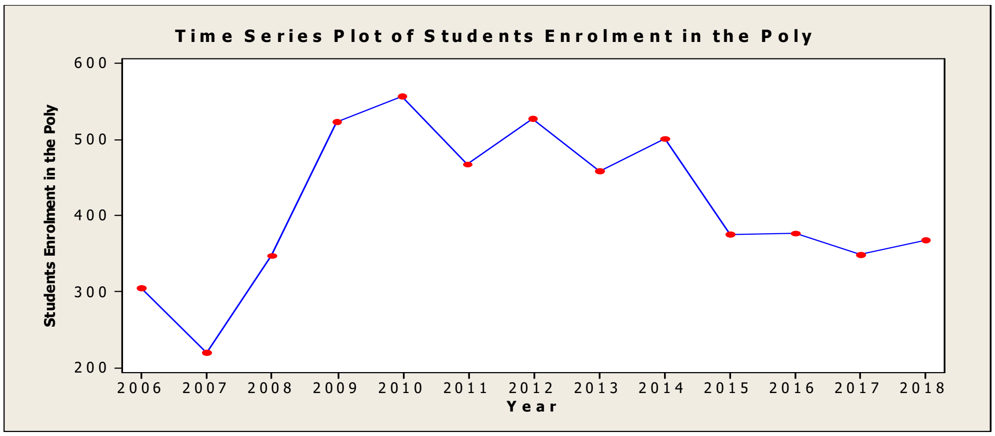


Figure 3.4 students' enrolment in Bolgatanga Polytechnic between 2006-2018

The plot in Figure 3.4 shows the fluctuation pattern of students' enrolment with respect to time. It can be observed, generally, from the figure above that decreasing trend in the plot is significantly sharp between 2015 and 2018. Students' enrolment however, took a significant downward turn at 2015, 2016, 2017 and 2018 respectively. The generally quadratic pattern in the time graph shows a no change of the mean whilst the sharper fluctuations over time shows an stable variance suggesting the series is stationary.

Tests for Stationarity

A stationary process has a mean and variance that do not change over time and the process does not have trends. To proceed with the estimation of an ARIMA model, the series is required to be stationary, as such this study employed the **Augmented Dickey-Fuller (ADF) test** and the Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test for evidence of stationarity in students' enrolment in quantitative methods.

Augmented Dickey-Fuller Test

- For the ADF test, we test the hypothesis that;
- H0: the series is not stationary.
- H1: the series is stationary.
- At 95% significance level, a p-value less than 0.05 means a rejection of H0, meaning the series is stationary, otherwise the H0 is upheld.

Kwiatkowski, Phillips, Schmidt and Shin Test

- The KPSS test has a reverse hypothesis to the ADF test hence;
- H0: the series is stationary.
- H1: the series is not stationary.

This means that at 95% significance level, a p-value less than 0.05 means we reject H0 and say the series is not stationary, otherwise it is stationary.

Table 3.4 STATIONARY TEST

TEST	**TEST STATISTIC**	**P-VALUE**
KPSS	**0.142742**	**0.067**
ADF	**-1.73045**	**0.049**

From the KPSS test values on table 3.4 above**,** at 5% significance level, the conclusion is that the series is *stationary* since the p-value (0.067) is greater than 0.05. However, the ADF test with a reverse null hypothesis indicates that the data is stationary with p-value 0.049. In all, the data is concluded to be stationary based on the evidence of the time plot, correlogram, KPSS and ADF test, hence data is stationary.

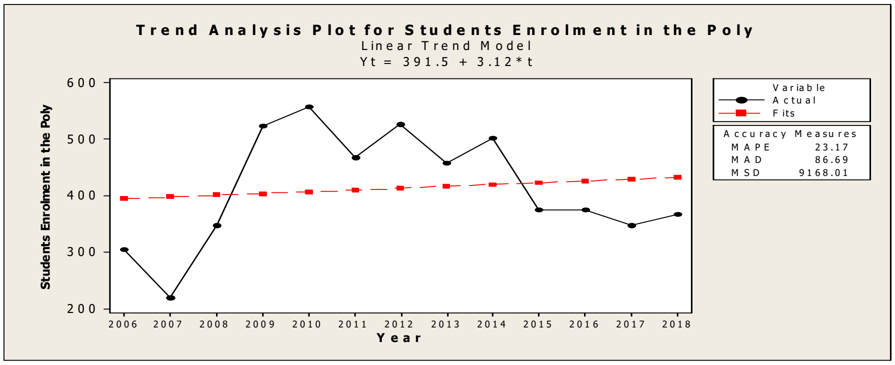


Figure 3.5: Trend plot of students' enrolment in Bolga Polytechnic

Figures 3.5 above and 3.6, 3.7 below show the linear, quadratic and exponential models respectively. In each of the figures, round dotted lines represent the actual values of students' enrolment whereas the square dotted lines represent the fitted values based on the various models.

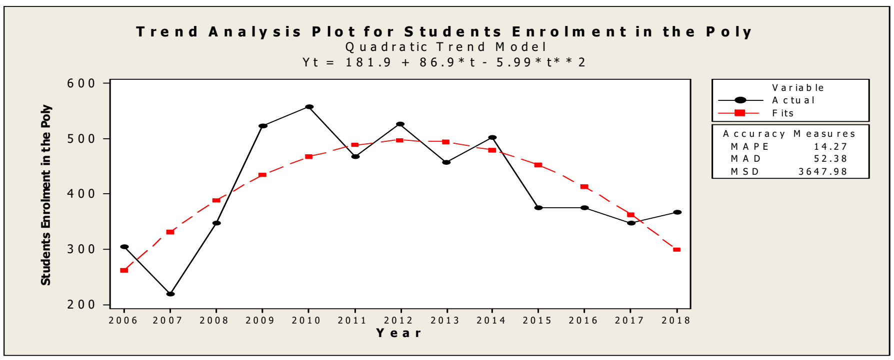


Figure 3.6: Quadratic trend plot of students' enrolment in Bolgatanga Polytechnic

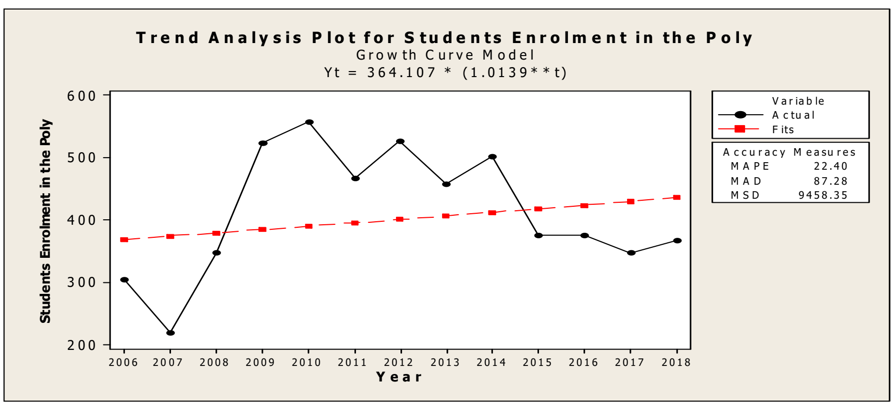


Figure 3.7: exponential trend plot of students' enrolment in Bolgatanga Polytechnic

Table 3.5: Measures of accuracy

Model	MAPE	MAD	MSD
Linear	23.17	86.69	9168.01
Quadratic	14.27	52.38	3647.98
Exponential	22.40	87.28	9458.35

From **Table 3.5** the most appropriate model to describe the trend in students' enrolment is the one with minimal errors. A closed observation of the errors produced by three models, the quadratic model has the minimum MAPE, MAD and MSD thus, is considered to be the best model in describing the trend in students' enrolment in Bolgatanga Polytechnic.

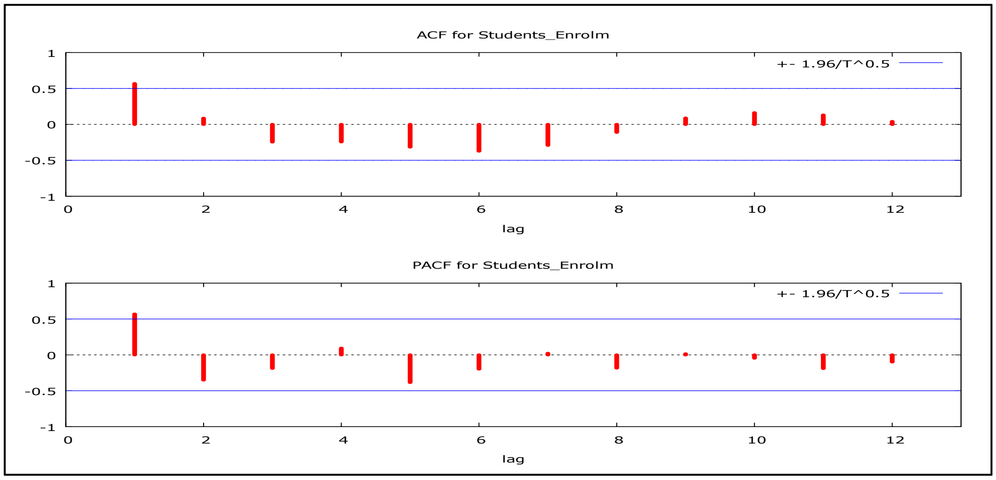


Figure 3.8 shows the ACF and PACF of students' enrolment in Bolga Polytechnic

Further analysis was conducted and checks made on the Autocorrelation Function (ACF) plots and those of the Partial Autocorrelation Function (PACF). It can be observed that with 95% confidence interval the data appears to be stationary. The ACF is significant spikes at lags 1 and 1 of the PACF as illustrated in Figures 3.8. These suggest that there is no increased or non-decreased in both ACF and PACF re-affirmed that the series is stationary.

4.3 The Box-Jenkins Method of Modeling Time Series

The Box-Jenkins methodology (Box & Jenkins, 1976) is a stepwise statistical method used in analyzing and building forecasting models which best represents a time series. This method of forecasting implements knowledge of autocorrelation analysis based on autoregressive integrated moving average models.

The methodology makes great use of historical time series data, is logically and statistically accurate and increase forecasting accuracy. The procedure is of four distinct stages namely; Identification, Estimation, Diagnostic checking, Forecasting.

4.3.2 Model identification

Table 3.4:

Model	AIC	BIC	HQ
Arima(1,0,0)	156.1739	157.8688	155.8256
Arima(1,0,1)	157.7579	160.0177	157.2934

The most appropriate model for the series is the one with the minimum Akaike Information Criteria (AIC), Bayesian information criterion (BIC) and Hannan-Quinn (HQ) . Thus, by an inspection of all the competing models in table 3.4 the ARIMA (1, 0, 0) model has the minimum values and therefore the best model for forecasting.

Parameter Estimation

Table 4.6 below displays estimates of the parameters of the ARIMA (1, 0, 0) model. The parameters of AR(1) is significant at 5% levels with coefficients and p-values of ***(***0.597190,**0.0057*)*** respectively. P-value less than 0.05 indicate the significance of the parameters.

Table 3.5: shows Parameter Estimation

Type	Coefficient	Standard error	Z value	P-value
AR 1	0.597190	0.215879	2.766	0.0057
Constant	399.043	48.7872	8.179	2.86e-016

Model Diagnosis

To ensure that the selected model is the best model that suits the data the following diagnostics are performed

Residuals Plots

The patterns of the residuals over time around the zero mean as seen in figure 3.9 below indicates that the residuals are random and independent of each other, thus, indicating that the model is fit.

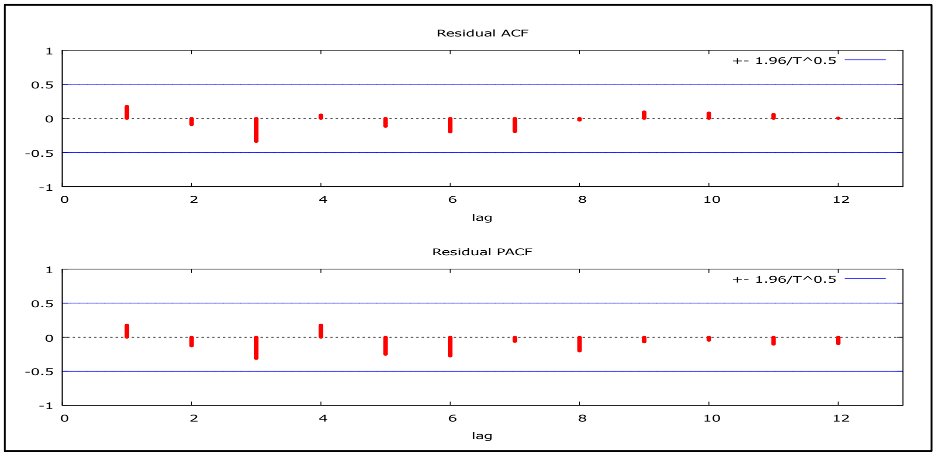


Figure 3.9: residual plot of ACF and PACF

The Normal Q-Q Plot

The Normal Q-Q Plot is another diagnostic check on the residuals to determine whether it follows the normal distribution. This is done by using the normal probability plot (Q–Q plot). It is a plot a plot based on estimates of the quantiles. The normal Q-Q plots is used to compare the distribution of a sample to a theoretical distribution. If most of the points are in line and closer to the normal line, then the model is a good fit.

The Q-Q plot in Figure 3.10 below shows all points along the normality line except for one outlier hence the model is deemed fit.

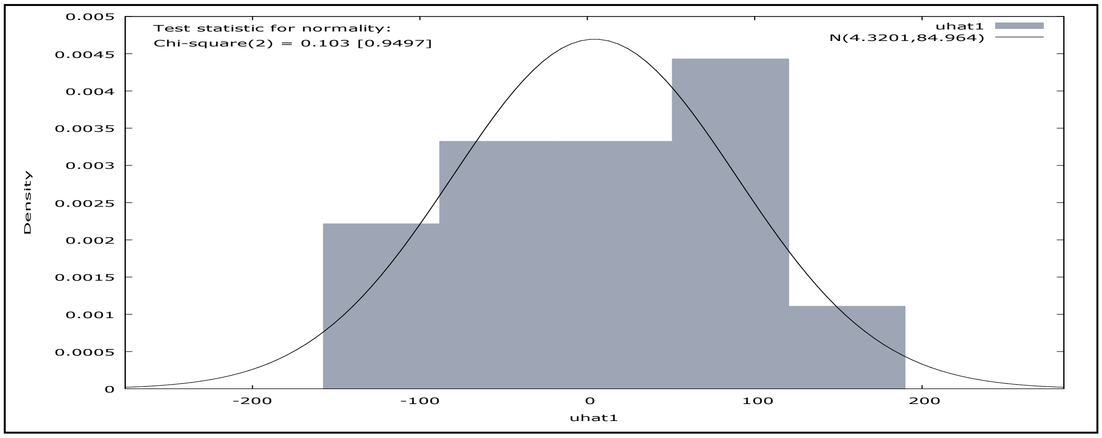


Figure 3.10: normal plot of residual

Ljung-Box Q Statistics

A check of the overall model adequacy is made using the Ljung-Box Q statistics. With a p-value of 0.834 which is way greater than 0.05 indicates that the model is generally adequate.

Table 4.3.5: Ljung-Box Q Statistics

MODEL	Statistics	DF	Sig.
Arima(1,0,0)	5.8	10	0.834

FORECAST

Since the model checks out to be of good fit, we can now forecast for future values in this instance, the next 5 observations**.**

Students'Enrolment	prediction	error	95%	interval
2019	380.50	76.709	230.16 -	530.85
2020	387.97	89.346	212.86 -	563.09
2021	392.43	93.441	209.29 -	575.57
2022	395.09	94.858	209.18 -	581.01
2023	396.68	95.359	209.78 -	583.58

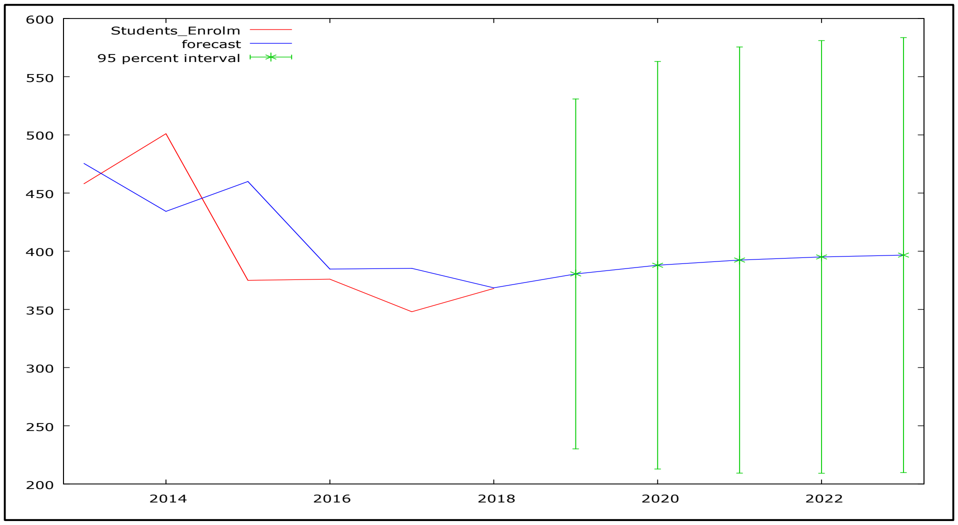


Figure 3.11: The graph for the five-year forecast of students' enrolment in Bolgatanga Polytechnic

From figure 4.11, it is believed that when the factors perceived to militate against increase in enrolment the polytechnic are address immediately the students' enrolment expected to experienced gradual increase from 2019-2023 and beyond as can be seen from the figure 3.11.

DISCUSSIONS OF RESULTS/CONCLUSION

From the descriptive statistics, in Table 3.2 it is observed that students' enrolment are positive skewed to the right, indicating that most of the enrolment values are concentrated at the left of the mean and this means that majority of the students' enrolment are below the average indicating weak enrolment in the Bolgatanga Polytechnic. Again, since the peakness demonstrated a platykurtic flattened with a coefficient of kurtosis = 0.68384 it shows that most of the students' enrolment are at either extreme of the distribution hence sharper than normal peak. These suggested that weak and inadequate strategies by the institution to increase students' enrolment.

Figure 3.4 shows an upward and flattened trend indicating a quadratic trend with the series showing a generally no increasing trend.

Figure 3.5- Figure 3.7 describe various trend models of the series and the best trend descriptor per the measures of accuracy in ***Table 3.5*** is the quadratic l model.

Secondly, even though the data was not transformed by way of differencing to achieve stationary and the tests of best fit also confirmed that the final model was adequate for the forecast, the five years predicted outcomes showed very little increase in students' enrolment over time. Thus, further demonstrating that improvement of students' enrolment in the Polytechnic is likely stagnate in the future if the Polytechnic and government are not able address the immediate challenges of the Polytechnic especially elevating it to Technical University to enable it compete among others schools.

Recommendations/Suggestions

Based on the findings of this study, couples with the fact that many students' considered factors such as infrastructure, lecturers, modern lectures hall and whether the school is Technical University before gaining admissions into the school to study a programmes: the study

Commends/suggests the following:

- Government should endeavour and ensure that qualified and adequate lecturers are recruited to meet the future demand of increases in students' enrolment as indicated on forecasted curve.
- Government and stokeholds should come to aid of the Polytechnic put-up more lectures halls to ease the pressure during lectures periods.
- The Polytechnic should have come out with long-term strategic policies to increase students' enrolment as indicated in the findings that no significant increase in students' enrolment over the past fifteen year period.

REFERENCES:

1. Bell, P., & Best, R. (1986). Supportive Education. Oxford, Blackwell.
2. Best, J. W. & Kahn, J. V. (1988) Research in education (8th ed) Boston, Allyn and Bacon
3. Betts, M. G., Diamond, A. W., Forbes, G. J., Villard, M. A., & Gunn, J. S. (2006). The importance of spatial autocorrelation, extent and resolution in predicting forest bird occurrence. *Ecological Modelling*, *191*(2), 197-224.

4. Boakye-Agyeman (2006). Polytechnic Education in Ghana: The case of the HND
5. Bosetti, L. (2004). “Determinants of school choice: understanding how parents choose elementary schools in Alberta.” Journal of Education Policy, 19, 387-405.
6. Box, G.E.P., & JENKINS, G.M. (1976). Time Series Analysis: Forecasting and Control (rev. ed.). San Francisco: Holden-Day.
7. Carr, S. (2000). Many professors are optimistic on distance learning, survey finds. *The Chronicle of Higher Education*, *46*(44), A35 Estate Management Programme. *5th FIG Regional Conference,* Accra
8. Feinberg, W., & Lubienski, C. (2008). School Choice Policies and Outcomes: Empirical and Philosophical Perspectives, State University of New York Press,New York.
9. Foster, L., & Walker, A. (2013). Gender and active ageing in Europe. European Journal of Ageing.10, 3–10. Ghana: The Relevance for Cape-Coast Polytechnic.
10. Holmes, G., DeSimone, J., & Rupp, N. (2008). “Does School Choice Increase School Quality?”, Working Paper 9683, NBER Working Paper Series, National Bureau of Economic Research.
11. KNUST-IDL. (2013). Towards a decade of distance education in KNUST (2005-2013). KNUST, Kumasi, Ghana.
12. Koul, B. N. (1995). Trends, directions and needs: A view from developing countries. In *Open and distance learning*
13. Ministry of Education. (1993). White paper on the Reforms to the Tertiary Education
14. Musset, P. (2012). “School Choice and Equity: CurrentPolicies in OECD Countries and a Literature Review”,OECD Education Working Papers, No. 66, OECDPublishing.
15. Ndudzo, D., & Nyatanga, E. K. (2013). Push and Pull Factors in ODL Service Delivery: A Case Study of Learners at
16. Owusu-Agyeman, (2006). Government Access-Policies on Polytechnic Education in *today,* ed. F. Lockwood, 23-31, London and New York: Routledge.
17. Sylvester, A. A., & John, A. A. (2018). Modeling students’ performance in quantitative methods using time series analysis. International Journal of Applied Science and Research, 1(3),1-20.
18. Waldrop, M. (2013). Campus 2.0. *Nature, 495,* (pp. 160-163).