International Journal of Mathematics and Statistics Invention (IJMSI)
E-ISSN: 2321- 4767 P-ISSN: 2321-4759
www.ijmsi.org Volume 8 Issue 6 || June, 2020 || PP 01-13

Modeling Patient Attendance in Gwarimpa General Hospital Abuja: An ARIMA Model Approach

[1]Megbatighon, John Oke and [2]Olanrewaju, Samuel Olayemi
[1]APUDI Institute for Peace Studies and Social Rehabilitation
[2]Department of Statistics, University of Abuja, Abuja Nigeria
Correspondence Author: Megbatighon, John Oke

ABSTRACTS: *This research fit a univariate time series ARIMA model to the quarterly patient attendance to Gwarinpa General Hospital from first quarter 2008 to fourth quarter 2016. The Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) model was estimated and the best fitted ARIMA model is used to obtain the post-sample forecasts for twenty (20) quarter (First Quarter 2018 to Fourth Quarter 2022). The fitted model is ARIMA (1,1,4) with Akaike Information Criteria (AIC) of 19.17464, Normalized Bayesian Information Criteria (BIC) of 19.3453. This model was further validated by Ljung-Box test with no significant Autocorrelation between the residuals at different lag times and subsequently by white noise of residuals from the diagnostic check performed which clearly portray randomness of the standard error of the residuals, no significant spike in the residual plots of ACF and PACF. The forecast values indicate clearly that patient attendance at Gwarinpa General Hospital would be on an increase rate. Therefore, Government needs to put in place more and better facilities.*

KEYWORDS: *Autoregressive Integrated Moving Average (ARIMA), Forecast, Patient,*

Date of Submission: 22-05-2020 Date of Acceptance: 09-06-2020

I. INTRODUCTION

Health status of any group of people has come to be seen as crucial not only to their well-being but also represent a strong influence on the productivity capacity of the people. The Government of Nigeria continues to looks for ways to restructure the welfare state to meet the changing needs, demands and expectations of a changing population. The need for the establishment of health insurance scheme which was informed by the general poor state of the nation's healthcare services, the excessive dependence and pressure on government provided health facilities, dwindling funding of healthcare in the face of rising costs, poor integration of private health facilities in the nation's healthcare delivery system and overwhelming dependence on out-of-pocket expenses to purchase health.

This trend according to [1] prompted the Federal Government of Nigeria to initiate the search for other means of funding health care that had been neglected in the past. Health insurance is an alternative source of health care financing that has become important in the developing world. It has been implemented as part of health reform programs and strategies aimed towards providing effective and efficient health care for citizens, most especially for the poor and vulnerable. The National Health Insurance Scheme (NHIS) was established as a social security system based on social health insurance to ensure that enrollees have access to quality and effective healthcare. The Federal government of Nigeria through the National Health Insurance Scheme (NHIS), has implemented the Tertiary Institutions Social Health Insurance Programme (TISHIP) with the hopes to achieve a more flexible, more innovative and more competitive response to the health need of people in the country, in order to ensure that every tertiary institution student has access to quality healthcare while schooling, that parents and guardian are protected from the financial hardship of huge medical bills, ensure equitable distribution of healthcare costs among different students, to ensure equitable distribution of healthcare facilities within the nation tertiary institution of learning, ensure availability of funds to the health sector for improved services ([1]).

Gwarimpa District Hospital Abuja, since its inception has received considerable amount of people, for treatment medical advice, family planning and a host of other reasons. Different categories of people have patronized the hospital for its efficiency.

It is therefore in my best interest to use my knowledge of statistics application in the attendance of ill health (patients) attending the hospital. The research will look forward to classify, arrange and record the monthly, quarterly, annual, bi-annual attendance of patients in the hospital for a period of ten years. In an attempt to introduce efficient methods and routine towards comparing the total attendance of, in and out patient this in general comprises of male, female and children patient attending the hospital. To crown it all, it shall be in form of data (secondary, primary data) depending on the set of people wishing to use it and purpose or criterion behind using the research, the data collected will be analyze, organized, summarized and compiled. Since hospital patronage is consistent and continuous process, it will be an efficient data collection, centres and will promote statistical application and voluminous data i.e. moving average and time series analysis. It is against this background that the researcher intends to statistically analyze patient attendance in Gwarimpa General Hospital using time series Analysis from 2008 to 2017.

According to the [2], patient attendance continues to increase nationally with a current per capita of 0.98 in the year 2015 to 1.09 in the year 2016. This increased cannot, however, be entirely attributed to full satisfaction of clients about quality services provided at the facilities. However, the proportion of patient attendance insured clients increased from 58.90% in 2015 to 95.56% in 2016, patient per capita increased from 0.98 in 2015 to 1.07 in 2016. This alarming statistics continues to be a major concern for key stakeholders like the Government, hospital administrators, doctors, nurses and pharmacists among others in the health sector of the country in the face of scary population-to-doctor and population to-nurse ratios of 10,032: 1 and 1240: 1 respectively. Not many studies have been conducted on patient attendance in the country. Moreover, no statistical model is available for forecasting patient cases for proper hospital planning and management. It is against this problem that researcher intends to analyze patient attendance in Gwarinpa General Hospital using time series analysis from 2008 to 2017.

The main aim of this study is to statistically analyze the rate of patient attendance in Gwarimpa General Hospital using time series Analysis from 2008 to 2017.

The ever increasing patronage of the Gwarinpa General Hospital requires that adequate preparations in terms of personnel and logistics are made in advance. The findings of the study would therefore help the Hospital management to adequately prepare for the large number of prospective patients. This is likely to help them make advanced plans in terms of man-power and logistical requirements for a better service delivery to the satisfaction and expectations of clients.

Again, the Nigeria Medical Health Service can adapt same model for nationwide forecasting of patient attendance cases. Additionally, the Government of Nigeria can also use the findings to review its financial commitment and contribution to the various health facilities throughout the country. Lastly, this study will add to the existing literature for academic purposes.

II. REVIEW OF RELATED LITERATURE

The importance of human health in national development has made efficiency in the production of health services in the Health Care System a subject of intense research interests in the literature ([3]). This sounds reasonable because spending on health is normally regarded as productive investment. Consequently, health is a fundamental goal of development. In addition, growth in health care costs has been attributed, at least in part, to the inefficiency of health care institutions ([4]).

However, the definition of health adopted by providers and government has implication for the process, measurement and range of services offered. The World Health Organization defines health as "a state of complete physical, social and mental well-being and not merely absence of disease and infirmity". In this way, health is metabolic efficiency while sickness or ill health is metabolic inefficiency. A state of complete physical, mental, and social well-being; not just absence of disease or infirmity is a healthy status- a status in which individuals can lead social and economically productive life.

The organized provision of health care services constitutes the Health Care System. According to the [5], a health system is defined as comprising all organizations, institutions and resources that are devoted to producing health actions. The health system provides an organized manner for providing healthcare services or health actions. A health action is defined as any effort, whether in personal health care, public health services or through intersectoral initiatives focuses primarily at promoting, restoring or maintaining health.

In addition, not only does each staff category have disproportionate contribution to treatment, the weight of their decisions varies with respect to health resource usage. [6] argued that around 80 percent of decisions in health resource utilization in hospitals are made by the physician. Consequently, studies commonly categorized human inputs into input variables in attempt to measure the level of technical efficiency

In hospital literatures, capital input is taken to represent a wide range of manufactured products such as complex medical equipment, buildings, beds and vehicles employed in health care. By nature capital inputs are durable and provide services over a fairly long period of time. It is, therefore, assumed in hospital literatures that a directly proportional link exists between quantity of capital stocks and capital services ([7]). However, number of beds is the most commonly used variable in hospital efficiency studies. The use of this variable as a proxy for capital inputs has been accepted by researchers ([8])

Consumables are non-labour and non capital inputs. Drugs and medical supplies are categorized as consumables and they represent an important input in hospital health care delivery process; often consumables constitute a major share of hospital expenditure. However, few studies have employed consumables as input variables in hospital efficiency studies and none, to our knowledge, in hospital efficiency studies in developing nations. The argument according to [9] is that in most developing nations, patients often times, procure consumables from their private pockets. Therefore, using consumables as input variable in hospital efficiency studies, particularly in developing nations, will yield misleading results and faulty recommendations.

The [10] acknowledges the annual public sector budgetary allocations to health often do not match approved allocation due to bureaucracies and other barriers. Thus, private sector expenditure on health as a percentage of total health expenditures has over the years exceeded government health expenditure. The World Health Organizations' national health account (2006) showed impressive percentage for the private sector as against the public sector. According to the reports, private sector expenditures on health as percentage of total health expenditures equals 74.4 percent (2002); 72.8percent (2003) 69.6 percent (2004) and 67.6 percent (2005); this trend is indicative that out of pocket expense is still the major means of payment for the health services in Nigeria. Private health insurance is still in developmental stage with only 0.3% of the population covered ([11]).

Again, [12] researched on the topic: ―ARIMA Models for health care failure: Prediction and Comparison. They said that the number of health care failures has increased dramatically over the last twenty-two years. A common notion in economics is that some health care can become ―too big to fail.

Is this still a true statement? What is the relationship, if any, between health care sizes and health care failures? In this thesis, the proposed modeling techniques are applied to real health care failure data from the FDIC. In particular, quarterly data from 1989:Q1 to 2010:Q4 are used in the data analysis, which includes three major parts: 1) pairwise health care failure rate comparisons using the conditional test, development of the empirical recurrence rate ([13]) and the empirical recurrence rates ratio time series; and 3) the Autoregressive Integrated Moving Average (ARIMA) model selection, validation, and forecasting for the health care failures classified by the total assets.

III. ARIMA MODELING IN DISEASE SURVEILLANCE: ADVANTAGES AND DISADVANTAGES

ARMA models in [14] require large historic records of patient visits in order to begin surveillance. This is a substantial disadvantage. As can be seen from [15], in some cases long historical data are not available and not necessary. Also, combining both historical and recent trends is quite realistic. Another disadvantage of ARMA is that the corresponding detector is not sensitive to the slow growth. According to [16], outbreaks that evolve over a matter of days, for example, can often be detected with ARMA models that generate single-day predictions based on historical data. More gradually developing outbreaks are generally easier to detect by using such techniques as CuSum ([17]).

Lean [18], worked on the topic forecasting the number of patient visits to hospitals using wavelet decomposition (WD) and artificial neural network (ANN) under the framework of "decomposition and ensemble". In this model, the WD is first employed to decompose the original monthly data of the number of patient visits to hospitals into several components and one residual term. Then, the ANN as a powerful prediction tool is implemented to fit each decomposed component and generate individual prediction results. Finally, all individual prediction values are fused into the final prediction output by simple addition method. For illustration and verification, four sets of monthly series data of the number of patient visits to hospitals are used as the sample data, and the results show that the proposed model can obtain significantly more accurate forecasting results than all considered popular forecasting techniques.

[19], Employed time-series forecasting to historical medical data in order to perform prediction of early pathological signs within telehealth applications. A benchmark of state-of-the-art learning methods were applied to a set of artificial time-series data, simulating hypertensive patient profiles, based on blood pressure measurements. Results provided a fair proof of our initial hypothesis. Based on their first experimentation, they plans to further investigate these findings in real-life or lab settings with seniors, thus proving the usefulness of time-series forecasting as a monitoring tool and an early prognosis mechanism in telehealth systems.

[20] in their paper titled Time series model for forecasting the number of new admission inpatients, which aim is to explore the application of the hybrid ARIMA-NARNN model to track the trends of the new admission inpatients, which provides a methodological basis for reducing crowding. The result shows that for the monthly data, the modeling RMSE and the testing RMSE, MAE and MAPE of SARIMA-NARNN model were less than those obtained from the single SAR1MA or NARNN model, but the MAE and MAPE of modeling performance of SARIMA-NARNN model did not improve. For the daily data, all RMSE, MAE and MAPE of

NARNN model were the lowest both in modeling stage and testing stage. They concluded that hybrid model does not necessarily outperform its constituents' performances. It is worth attempting to explore the reliable model to forecast the number of new admission inpatients from different data.

IV. METHODOLOGY

Method of Data Collection

The data used in this research work is a secondary data collected from Gwarinpa National Hospital data/research department. The data covers the details of patients that visit

Estimation of the Component

Statistical estimation is concerned with using the data obtained from a random sample to obtain information about unknown population parameter.

There are two different types of estimation of population parameter, this include point estimate and interval estimate. A point estimate is an estimate of a population parameter given by a single value called a point estimate of the parameter.

Interval estimation is the use of sample data to calculate an interval of plausible values of an unknown population parameter; this is in contrast to point estimation, which gives a single value.

Interval estimate is preferable to point estimate because it indicate the precision or accuracy of an estimate.

Differencing

Differencing simply means subtracting the value of an earlier observation from the value of a later observation. Calculating differences among pairs of observations at some lag to make a non-stationary series stationary. There are possible shifts in both the mean and the dispersion over time for this series. The mean may be edging upwards, and the variability may be increasing. If the mean is changing, the trend is removed by differencing once or twice. If the variability is changing, the process may be made stationary by logarithmic transformation. Differencing the scores is the easiest way to make a non-stationary mean stationary (flat). The number of times you have to difference the scores to make the process stationary determines the value of *d*. If *d* C0, the model is already stationary and has no trend. When the series is differenced once, *d* C1 and linear trend is removed. When the difference is then differenced, *d* C2 and both linear and quadratic trend are removed. For non-stationary series, *d* values of 1 or 2 are usually adequate to make the mean stationary.

Autoregressive Moving Average (ARMA) Models

Autoregressive and Moving Average processes can be combined to obtain a very flexible class of univariate processes (proposed by Box and Jenkins), known as ARMA processes. The time series X_t is an ARMA (p, q) process, if it is stationary and

$$X_t = \alpha + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \cdots + \theta_q X_{t-q} \quad (1)$$

$$X_t = \alpha + \phi_1 X_{t-1} + \theta(L)\varepsilon_t \quad (2)$$

$$\phi(L) X_t = \theta(L)\varepsilon_t \quad (3)$$

Where $\theta, \phi, \varepsilon$ and L are as defined above with $\phi_p \neq 0$ and $\theta_q \neq 0$

An ARMA process is stationary if the roots of ϕ(L) all lie outside the unit circle and invertible if the roots of θ(L) all lie outside the unit circle.

The three terms to be estimated in the model are auto-regressive (*p*), integrated (trend—*d*), and moving average (*q*).The ARIMA (auto-regressive, integrated, moving average) model of a time series is defined by three terms (*p*, *d*, *q*). Identification of a time series is the process of finding integer, usually very small (e.g., 0, 1, or 2), values of *p*, *d*, and *q* that model the patterns in the data. When the value is 0, the element is not needed in the model. The middle element, *d*, is investigated before *p* and *q*. The goal is to determine if the process is stationary and, if not, to make it stationary before determining the values of *p* and *q*. Recall that a stationary process has a constant mean and variance over the time period of the study.

Box-Jenkins ARIMA Process

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. The Box–Jenkins methodology, named after the statisticians George Box and Gwilym Jenkins, applies ARIMA models to find the best fit of a time series to past values of this time series, in order to make forecasts. 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 remove the non-stationarity. The model is generally referred to as an ARIMA (*p, d, q*) model where *p, d*, and *q* are non-negative integers that refer to the order of the autoregressive, integrated, and moving average parts of the model respectively.

Box-Jenkins Modeling Approach

The Box-Jenkins model uses iterative three-stage modeling approach which is:

-Model identification and model selection: making sure that the variables are stationary, identifying seasonality in the dependent series (seasonally differencing it if necessary), and using plots of the autocorrelation and partial autocorrelation functions of the dependent time series to decide which (if any) autoregressive or moving average component should be used in the model.

-Parameter estimation using computation algorithms to arrive at coefficients which best fit the selected ARIMA model. The most common methods use maximum likelihood estimation or non-linear least-squares estimation.

-Model checking by testing whether the estimated model conforms to the specifications of a stationary univariate process. In particular, the residuals should be independent of each other and constant in mean and variance over time (plotting the mean and variance of residuals over time and performing a Ljung-Box test or plotting autocorrelation and partial autocorrelation of the residuals are helpful to identify misspecification). If the estimation is inadequate, we have to return to step one and attempt to build a better model.

Box-Jenkins Model Identification

Best Model Identification and Selection Criteria

The Ljung-Box statistic would be used to identify the best model. The Ljung-Box statistic, also called the modified Box-Pierce statistic, is 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

$$Q(m) = n(n+1)\sum_{j=1}^{m}\frac{r_j^2}{n-j} \tag{4}$$

Where;

n = number of usable data points after any differencing operations.

The Ljung-Box test can be defined as follows:

H_0: The data are independently distributed (i.e. the correlations in the population from which the sample is taken are 0, so that any observed correlations in the data result from randomness of the sampling process).

H_a: The data are not independently distributed.

The choice of a plausible model depends on its p-value for the modified Box-Pierce if is well above .05, indicating ―non-significance. In other words, the bigger the p-value, the better the model.

V. DATA ANALYSIS AND DISCUSSION OF THE RESULTS

DATA ANALYSIS

Unit Root Test (Test for Stationarity)

H_0: the data is non-stationary

H_1: The data is stationary

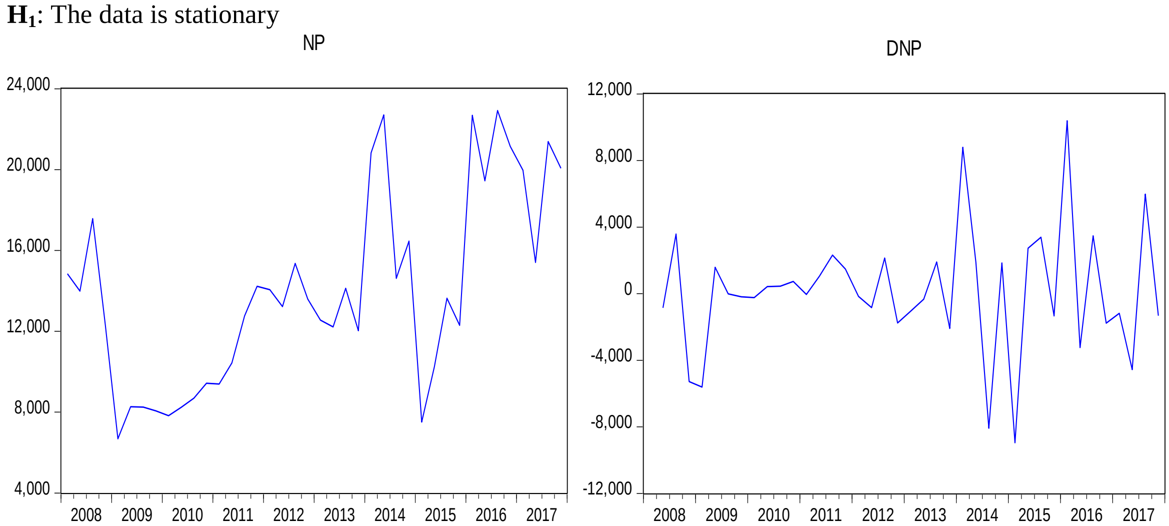


Figure1. (a) Number of patient at level (b) Number of patient at first difference

Table 4.1 Result of Augmented Dickey Fuller Unit Root Test (Test for Stationarity)

Variable	I(0): (at Level)	I(1): (at First difference)
Number of patient	Not stationary	**Stationary**

Source: Eview output

Figure 1 above is the time plot on patients that visited Gwarinpa general Hospital from 2008 first quarter to 2017 fourth quarter. Figure 1(a) above is the raw data time plot which from observation is not stationary and for a time series analysis to be run, the data set has to be stationary. Figure 1(b) above is the time plot of the number of patient at first difference. From here we can observe that the series is stationary i.e they revolve around the same mean. Augmented Dickey-Fuller test statistic also show that the data is not stationary at level with probability of 0.9090 which is greater than 0.05 level of significant for rejecting the null hypothesis (i.e the data is non stationary or has a unit root). While at first difference, Augmented Dickey-Fuller test statistic show that the data is stationary with probability statistic of 0.0006 which is less than 0.05 level of significant at 5% confidence interval of the level of accepting the null hypothesis and conclude that the data is stationary or has no unit root at first difference.

Identification of the AR and the MA process (Correlogram)
Plotting the ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function).

Sample: 2008Q1 2017Q4
Included observations: 40

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.672	0.672	19.478	0.000
		2	0.521	0.125	31.460	0.000
		3	0.353	-0.070	37.105	0.000
		4	0.159	-0.170	38.282	0.000
		5	0.165	0.195	39.592	0.000
		6	0.134	0.042	40.479	0.000
		7	0.169	0.084	41.936	0.000
		8	0.157	-0.065	43.232	0.000
		9	0.140	0.029	44.300	0.000
		10	0.057	-0.132	44.483	0.000
		11	0.027	0.060	44.527	0.000
		12	-0.006	-0.030	44.529	0.000
		13	-0.015	0.039	44.543	0.000
		14	0.032	0.023	44.608	0.000
		15	-0.046	-0.154	44.751	0.000
		16	-0.048	-0.018	44.912	0.000
		17	-0.054	0.056	45.127	0.000
		18	-0.077	0.007	45.583	0.000
		19	-0.094	-0.105	46.286	0.000
		20	-0.119	-0.047	47.474	0.001
		21	-0.146	-0.058	49.362	0.000
		22	-0.097	0.138	50.238	0.001
		23	-0.103	-0.075	51.287	0.001
		24	-0.135	-0.085	53.204	0.001

Sample: 2008Q1 2017Q4
Included observations: 39

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.302	-0.302	3.8452	0.050
		2	0.088	-0.004	4.1782	0.124
		3	-0.002	0.025	4.1784	0.243
		4	-0.335	-0.362	9.3055	0.054
		5	0.035	-0.209	9.3645	0.095
		6	-0.073	-0.119	9.6205	0.142
		7	0.038	-0.047	9.6912	0.207
		8	0.122	-0.014	10.463	0.234
		9	0.091	0.086	10.907	0.282
		10	-0.051	-0.058	11.050	0.354
		11	0.035	-0.002	11.121	0.433
		12	-0.141	-0.082	12.299	0.422
		13	-0.074	-0.092	12.639	0.476
		14	0.107	0.068	13.376	0.497
		15	-0.107	-0.053	14.145	0.515
		16	0.105	-0.065	14.915	0.531
		17	0.004	-0.067	14.916	0.602
		18	0.007	0.014	14.920	0.667
		19	0.013	-0.024	14.934	0.727
		20	-0.008	0.024	14.939	0.780
		21	-0.145	-0.167	16.814	0.722
		22	0.111	0.019	17.970	0.708
		23	0.035	0.113	18.092	0.752
		24	0.034	0.112	18.216	0.792

Figure 2(a) NP at level. 2(b) Np at first difference.

Figure 2(a) show that the series is not stationary since the ACF is descending gradually. While figure 2(b) shows that the series is stationary and it is only at lag 1 and lag 4 that it is above the error bound or 95% confidence interval. There for the only combinations we can have are AR(1), AR(4), I(1), MA(1) and MA(4), giving ARIMA(1, 1, 1), ARIMA(1, 1, 4), ARIMA(4, 1, 1), and ARIMA(4, 1, 4).

Criteria for the best model are that, the model must:

- have the best significant co-efficient
- have the least Volatility (SIGMASQ)
- Have the lowest AIC (Akaike Info Criterion) and SIC (Schwarz Criterion)
- Have the highest Adjusted R^2.

Test for Best Fit Model

Table 4.3 Result of the all the ARIMA models

	ARIMA(1,1,1)	ARIMA(1,1,4)	ARIMA(4,1,1)	ARIMA(4,1,4)
significant co-efficient	1	2	2	0
Volatility (Sigma^2)	10711311	9638950	9898818	11551431
AIC	19.29565	19.17464	19.18544	19.33535
SIC	19.46627	19.34526	19.35606	19.50597
Adjusted R^2	0.234305	0.310962	0.292386	0.174249

Source: Eview

Observing these models, ARIMA(1,1,4) have high significant co-efficient, have lowest volatility, have the lowest AIC and SIC value and have the highest Adjusted R^2. Therefore ARIMA(1,1,4) should be the best model for forecasting. To ascertain that the model is the best fit, the correlogram plot of that ARIMA model will be the determinant.

Table 4.4 Correlogram for ARIMA (1,1,4)

Sample: 2008Q1 2017Q4
Included observations: 39
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.032	-0.032	0.0440	
		2	-0.108	-0.109	0.5474	
		3	-0.092	-0.101	0.9234	0.337
		4	0.045	0.026	1.0148	0.602
		5	-0.076	-0.096	1.2875	0.732
		6	-0.130	-0.143	2.1054	0.716
		7	0.040	0.015	2.1845	0.823
		8	0.161	0.121	3.5277	0.740
		9	0.105	0.109	4.1167	0.766
		10	-0.087	-0.042	4.5307	0.806
		11	-0.009	0.012	4.5349	0.873
		12	-0.097	-0.115	5.0894	0.885
		13	-0.119	-0.130	5.9647	0.876
		14	0.078	0.107	6.3513	0.897
		15	-0.058	-0.092	6.5747	0.923
		16	0.111	0.071	7.4307	0.917
		17	-0.056	-0.088	7.6567	0.937
		18	0.009	-0.042	7.6634	0.958
		19	0.069	0.094	8.0433	0.966
		20	0.004	0.035	8.0450	0.978
		21	-0.166	-0.111	10.481	0.940
		22	0.054	0.067	10.755	0.952
		23	0.145	0.094	12.842	0.914
		24	0.044	0.052	13.048	0.932

The above table shows that all the lag structure lies between the 95% confidence interval or standard error bounds (i.e it is flat). This also means that all the information has been captured in the model. Therefore ARIMA(1,1,4) is the best fit/most ideal for predicting future occurrences.

Test for Autocorrelation

Table 4.5 Test for Autocorrelation (Ljung-Box Test for Squared residual)

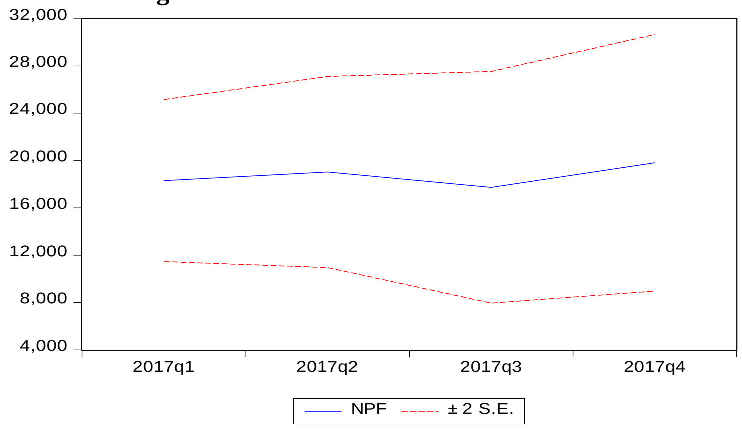
Sample: 2008Q1 2017Q4
Included observations: 39

Lag	AC	PAC	Q-Stat	Prob
1	-0.019	-0.019	0.0152	0.902
2	0.039	0.038	0.0802	0.961
3	-0.187	-0.186	1.6287	0.653
4	0.245	0.247	4.3637	0.359
5	-0.102	-0.104	4.8527	0.434
6	0.231	0.212	7.4348	0.282
7	-0.218	-0.177	9.8033	0.200
8	0.079	0.025	10.126	0.256
9	-0.184	-0.104	11.940	0.217
10	0.085	-0.061	12.337	0.263
11	-0.140	-0.011	13.458	0.264
12	0.027	-0.122	13.501	0.334
13	-0.132	0.037	14.566	0.335
14	0.097	-0.011	15.172	0.366
15	-0.223	-0.161	18.474	0.239
16	-0.011	-0.043	18.482	0.296
17	-0.123	-0.091	19.585	0.296
18	0.037	-0.055	19.687	0.351
19	-0.147	-0.096	21.425	0.314
20	0.190	0.155	24.475	0.222
21	-0.046	0.012	24.664	0.262
22	0.083	0.013	25.310	0.283
23	-0.079	0.032	25.928	0.304
24	0.169	0.032	28.980	0.221

Hypothesis

H_0: There exist no autocorrelation among the variables.
H_1: There exist autocorrelation among the variables.
Table 4.4 above shows that the model is free from autocorrelation since the probability values are greater than 0.05 at 5% level of significant.

4.2.5Forecasting



Forecast: NPF	
Actual: NP	
Forecast sample: 2017Q1 2017Q4	
Included observations: 4	
Root Mean Squared Error	2706.452
Mean Absolute Error	2298.337
Mean Abs. Percent Error	12.54433
Theil Inequality Coefficient	0.071087
Bias Proportion	0.032800
Variance Proportion	0.300895
Covariance Proportion	0.666305
Theil U2 Coefficient	0.658000
Symmetric MAPE	12.42070

Figure 3. Number of patients forecast for year 2017 from Quarter1 to Quarter4
Source: Eview output.

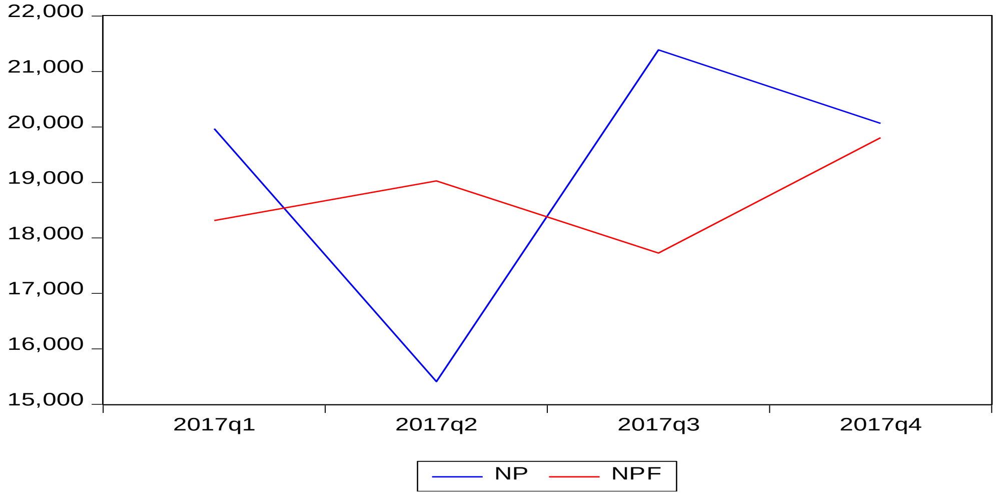


Figure4. Comparing Number of patients forecasted with the actual number of patient for year 2017
Source: Eview output.

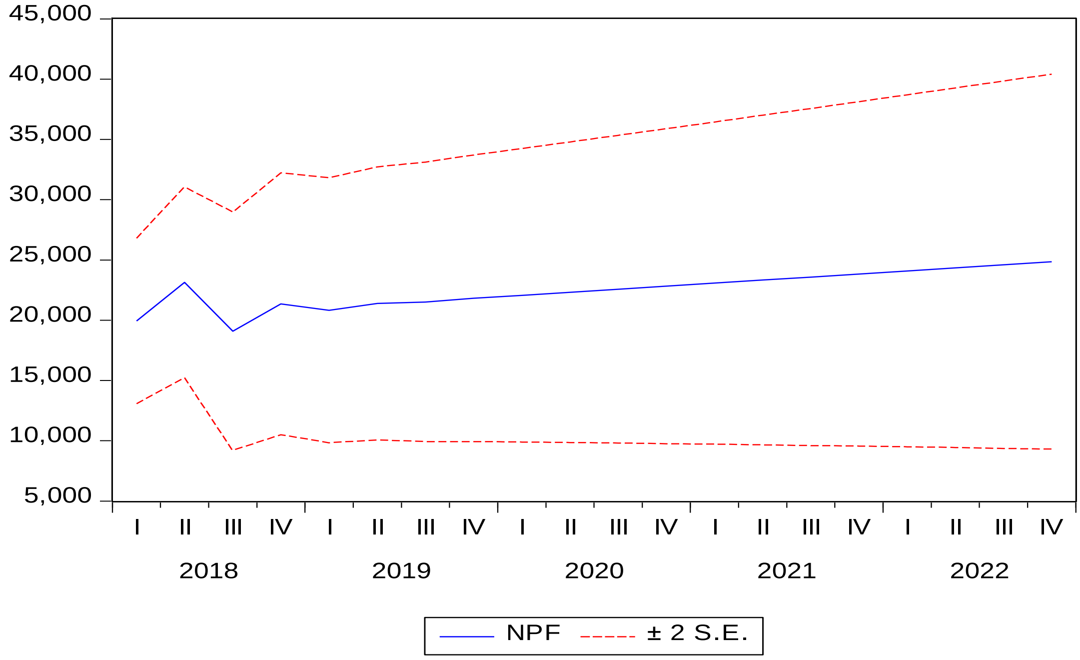


Figure 5: Number of patients forecast for year 2017Q1 to 2022Q4
Source: Eviews Output

The result above shows that the number of patients visiting Gwarinpa General Hospital will be on an increase from 2019 first quarter to 2022 fourth quarter and onward. This requires urgent intervention of government and health stack holders. The values are given below

Table 4.6 Number of patients forecasted from 2018Q1 to 2019Q4

Year	Forecast
2018Q1	19941
2018Q2	23137
2018Q3	19087
2018Q4	21361
2019Q1	20814
2019Q2	21386
2018Q2	21514
2019Q3	21818
2020Q1	22053
2020Q2	22315
2020Q3	22566
2020Q4	22821

2021Q1	23075
2021Q2	23329
2021Q3	23583
2021Q4	23837
2022Q1	24091
2022Q2	24346
2022Q3	24600
2022Q4	24854

Correlation Result

H_0: There is no correlation between death cases and number of Patients.

H_1: There is correlation between death cases and number of patients.

Table 4.7 Correlation Result

	DEATH_CASES	NP
DEATH_CASES	1.000000	0.343611
NP	0.343611	1.000000

Source: Eview output

The result of the correlation which tells us about the relationship between number of patients visiting Gwarinpa General Hospital and the number of death recorded quarterly shows that r = 0.3436, this means that there exist a weak positive relationship.

Table 4.8 Number of death cases recorded from 2008 to 2017

Years	No of Inpatients	Death Cases
2008 2009 2010 2011 2012 2013 2014 2015 2016 2017	58713 31270 34247 46810 56836 50936 74613 43714 86205 65271	26 15 25 31 70 82 48 19 36 50
TOTAL	548615	401

Here, the research shall denote the number of inpatients by X and the number of deaths cases by Y. The data used to determine the degree of relationship between the total number of inpatients and death cases.

VI. SUMMARY

This study examined and Model and Forecast the number of patient attendance at Gwarinpa General Hospital. The forecasting models ARIMA was used on a quarterly dataset between "First Quarter, 2011", to "Fourth Quarter, 2017" .The preliminary analysis of the data obtained shows that the distribution of the quarterly inflow of patient into Gwarinpa General Hospital was not stationary at level but attains stationarity at first difference which led to the use of ARIMA instead of ARMA. The Parameter of the ARMA models and Models selection, ARMA were estimated with most of the parameter significant at 1% and 5%. Significant Coefficient, Volatility, AIC, SIC and Adjusted R^2 was used to select the best model that will be used for ARIMA model because is the combination of AR and MA model. From the above criteria, ARIMA (1,1,4) was selected to be the best model since it has the highest significant coefficient, lowest volatility, smallest AIC and SIC and highest adjusted R^2. A diagnostic test also was evaluated which confirms that ARIMA (1,1,4,) is an adequate model because the residual are not dependent and the Q-Q plot are normally distributed.

VII. CONCLUSION

This study had come out with some finding in the comparing of forecasting models on Quarterly patient that visited Gwarinpa General Hospital. From the results of the forecasting models, ARIMA (1,1,4) is the adequate forecasting model in estimating number of patients that visits Gwarinpa General Hospital. Furthermore, in terms of forecasting accuracy, the forecasting model was evaluated using the correlogram also shows that all the lags falls between the error bound which supported that ARIMA(1,1,4) is appropriate for forecasting. It also shows that the data set is stationary. The correlation result shows that as population increases, there is tendency that death increase and also if population decreases, death might also reduce.

VIII. CONTRIBUTION TO KNOWLEDGE

With the recent forecast of number of patient that visit Gwarinpa General Hospital, it has bring to light the trend in which the number of patients that visit Gwarinpa General Hospital could be determined. This will help the hospital management to be able to prepare for better services so as to be able to accommodate the patients.

IX. RECOMMENDATION

Although the results of this study has proven to be statistically consistent with statistical tools model and evaluation models in treating the quarterly patients visit to Gwarinpa General Hospital. We therefore recommend that appropriate models should be used in terms of model selection and forecasting model. We also recommend that appropriate policies as forecast shows tendency of increase in number, measures should be put in place to be able to cater for those population and measures should also be on ground to reduce the death cases as population of patient visiting might increase as population of the country is on an increase. Government should also try reducing the population visiting the health care by educating the citizen of the country.

REFERENCES

[1]. Operational Guidelines. Abuja: NHIS;. National Health Insurance Scheme Nigeria. 2012 [Google Scholar]

[2]. Adeloye, D., David, R.A., Olaogun, A.A. et al. Health workforce and governance: the crisis in Nigeria. Hum Resour Health **15,** 32 (2017). https://doi.org/10.1186/s12960-017-0205-4

[3]. Hollingsworth, B. Non-Parametric and Parametric Applications Measuring Efficiency in Health Care. Health Care Management Science **6,** 203–218 (2003). https://doi.org/10.1023/A:1026255523228

[4]. Everett L. Worthington Jr. & Michael Scherer. Forgiveness is an emotion-focused coping strategy that can reduce health risks and promote health resilience: theory, review, and hypotheses, Psychology & Health, 19:3, (2004) 385-405, DOI: 10.1080/0887044042000196674

[5]. The world health report 2000 - Health systems: improving performance

[6]. Daniel Eisenberg, Sarah Ketchen Lipson and Julie Posselt. Promoting Resilience, Retention, and Mental Health. College Student Mental Health. Volume 2016, Issue 156. Pp87-95

[7]. Peacock SJ, de Silva I, Lowy FD. What determines nasal carriage of Staphylococcus aureus?. Trends Microbiol. 2001;9(12):605-610. doi:10.1016/s0966-842x(01)02254-5

[8]. Harrison JC, et al. Stress-specific activation mechanisms for the "cell integrity" MAPK pathway. J Biol Chem (2004) 279(4):2616-22

[9]. Peter Nolan PhD RGN RMN et al Violence in mental health care: The experiences of mental health nurses and psychiatrists December 2001 Journal of Advanced Nursing 30(4):934 – 941 DOI: 10.1046/j.1365-2648.1999.01163.x

[10]. Federal Ministry of Health and WHO 2002. World Bank, Nigeria: Poverty in the midst of plenty. The challenge of growth with inclusion. World Bank poverty Assessment Report No 14733-UNI, 1996

[11]. Ogunbekun I, Ogunbekun A, Orobaton N. Private Health Care in Nigeria: Walking the Tightrope. Health Policy and Planning, Oxford University Press. 1999 14(2): 81-174.

[12]. Yin Ping Chen et al. Predictive Efficiency Comparison of ARIMA-Time-Series and the Grey System GM(1,1) Forecast Model on Forecasting the Incidence Rate of Hepatitis B June 2013. DOI: 10.4028/www.scientific.net/AMR.709.836

[13]. Andrew Dean Ho. The Problem With "Proficiency": Limitations of Statistics and Policy Under No Child left behind. Educational Researcher; Aug/Sep 2008; 37, 6; ProQuest Psychology Journals.

[14]. Reis, B. Y., Pagano, M., and Mandl, K. D.(2003). Using temporal context to improve biosurveillance. Proceedings National Academy of Sciences, 100(4),1961-1965.

[15]. L. Zhou, P. Zhao, D. Wu, C. Cheng and H. Huang, "Time series model for forecasting the number of new admission inpatients," BMC Medical Informatics and Decision Making, vol. 18, pp. 1-11, 2018

[16]. Rizzo, S. L., Grigoryan, V. V., Wallstrom, G. L., Hogan, W. R., and Wagner M. M.(2005). The use of case studies for the evaluation of surveillance systems. RODS Technical Report Centre for Biomedical Informatics, University of Pittsburgh.

[17]. Hawkins, M. D., and Olwell, D. H. (1998). Cumulative sum charts and charting for quality improvement. Berlin/Heidelberg: Springer Verlag.

[18]. Lean Yu, Geye Hang, Ling Tang, Yang Zhao and K.K. Lai, Forecasting Patient Visits to Hospitals using a WD&ANN-based Decomposition and Ensemble Model, Eurasia Journal of Mathematics, Science and Technology Education, vol.13, pp. 7615-7627, 2018

[19]. Antonis S. Billis and Panagiotis D. Bamidis, (2014). Employing time-series forecasting to historical medical data…" Researchgate, January 2014.

[20]. Zhou, Lingling, et al. "Time series model for forecasting the number of new admission inpatients." BMC Medical Informatics and Decision Making, vol. 18, no. 1, 2018. Accessed 24 May 2020.

XI. APPENDIX

Dependent Variable: D(NO_OF_PATIENTS___YT_)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 08/29/19 Time: 05:55
Sample: 2 40
Included observations: 39
Convergence achieved after 18 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	254.1312	178.3336	1.425032	0.1630
AR(1)	-0.396542	0.152607	-2.598458	0.0136
MA(4)	-0.604780	0.191350	-3.160591	0.0032
SIGMASQ	9638950.	3438510.	2.803235	0.0082

R-squared	0.310962	Mean dependent var	133.8718
Adjusted R-squared	0.251902	S.D. dependent var	3789.081
S.E. of regression	3277.277	Akaike info criterion	19.17464
Sum squared resid	3.76E+08	Schwarz criterion	19.34526
Log likelihood	-369.9055	Hannan-Quinn criter.	19.23586
F-statistic	5.265158	Durbin-Watson stat	2.051939
Prob(F-statistic)	0.004195		

Inverted AR Roots	-.40			
Inverted MA Roots	.88	-.00+.88i	-.00-.88i	-.88

Dependent Variable: D(NO_OF_PATIENTS___YT_)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 08/29/19 Time: 05:57
Sample: 2 40
Included observations: 39
Convergence achieved after 21 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	234.3261	212.1119	1.104729	0.2768
AR(4)	-0.494573	0.118947	-4.157920	0.0002
MA(1)	-0.419333	0.150012	-2.795325	0.0084
SIGMASQ	9898818.	2666031.	3.712942	0.0007

R-squared	0.292386	Mean dependent var	133.8718
Adjusted R-squared	0.231733	S.D. dependent var	3789.081
S.E. of regression	3321.161	Akaike info criterion	19.18544
Sum squared resid	3.86E+08	Schwarz criterion	19.35606
Log likelihood	-370.1161	Hannan-Quinn criter.	19.24666
F-statistic	4.820657	Durbin-Watson stat	2.064249
Prob(F-statistic)	0.006511		

Inverted AR Roots	.59+.59i	.59+.59i	-.59-.59i	-.59-.59i
Inverted MA Roots	.42			

Dependent Variable: D(NO_OF_PATIENTS___YT_)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 08/29/19 Time: 05:58
Sample: 2 40
Included observations: 39
Convergence achieved after 12 iterations

Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	230.2951	344.1518	0.669167	0.5078
AR(4)	-0.036348	0.332755	-0.109233	0.9136
MA(4)	-0.476121	0.412868	-1.153205	0.2566
SIGMASQ	11551431	3482434.	3.317057	0.0021

R-squared	0.174249	Mean dependent var	133.8718
Adjusted R-squared	0.103470	S.D. dependent var	3789.081
S.E. of regression	3587.701	Akaike info criterion	19.33535
Sum squared resid	4.51E+08	Schwarz criterion	19.50597
Log likelihood	-373.0393	Hannan-Quinn criter.	19.39656
F-statistic	2.461886	Durbin-Watson stat	2.730755
Prob(F-statistic)	0.078753		

Inverted AR Roots	.31+.31i	.31+.31i	-.31-.31i	-.31-.31i
Inverted MA Roots	.83	.00-.83i	-.00+.83i	-.83

Dependent Variable: D(NO_OF_PATIENTS___YT_)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 08/29/19 Time: 05:58
Sample: 2 40
Included observations: 39
Failure to improve objective (non-zero gradients) after 14 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	236.7373	106.2466	2.228188	0.0324
AR(1)	0.528372	0.257826	2.049333	0.0480
MA(1)	-1.000000	3716.211	-0.000269	0.9998
SIGMASQ	10711311	8.95E+08	0.011970	0.9905

R-squared	0.234305	Mean dependent var	133.8718
Adjusted R-squared	0.168674	S.D. dependent var	3789.081
S.E. of regression	3454.774	Akaike info criterion	19.29565
Sum squared resid	4.18E+08	Schwarz criterion	19.46627
Log likelihood	-372.2652	Hannan-Quinn criter.	19.35687
F-statistic	3.570031	Durbin-Watson stat	2.077727
Prob(F-statistic)	0.023621		

Inverted AR Roots	.53
Inverted MA Roots	1.00

Megbatighon, John Oke, et. al. "Modeling Patient Attendance in Gwarimpa General Hospital Abuja: An ARIMA Model Approach." *International Journal of Mathematics and Statistics Invention (IJMSI)*, vol. 08(06), 2020, pp. 01-13.