Original Research Article

Modelling Dynamic Behaviour of Out-Patient Department Visits at the University of Cape Coast Hospital Using Time Series Analysis

Abstract

Background: Accurate and reliable forecasting of outpatient department visits enhances decision-making and planning for future healthcare demands and is the foundation for greater and better utilization of healthcare resources and increased levels of outpatient care and satisfaction. Though the literature has proposed several candidate models for predicting outpatient visits in some hospitals in Ghana, the model regulating outpatient visits at the University of Cape Coast Hospital (UCC) is unknown. There is therefore a need to determine the best model applicable to the specific case of UCC Hospital.

Aim: This study sought to determine and model the dynamics of outpatient visits in UCC Hospital and to project outpatient healthcare demands at the facility for the period July 2021 to July 2024.

Methods: This paper employed a monthly periodicity of 114-time series data sourced from District Health Information Management Systems Two (DHIMS 2) on outpatient department visits at UCC Hospital from January 2012 to June 2021. The autoregressive integrated moving average (ARIMA) models which are a form of the classical Box-Jenkins approach of Time Series Analysis were used to analyse the data. Analysis was performed in *EViews* 12.

Results: The study results showed twenty-five non-seasonal tentative models for the hospital and ARIMA (4, 1, 4) was selected as the best fit model with a fourth-order autoregressive and moving average terms each and one order of nonseasonal differencing. Residual analysis of the fitted model indicates that the model is adequate for forecasting. The findings revealed an overall rising trend in the incidence of outpatient department visits to the hospital over the study period with an average of 6000 visits per month. This is expected to increase to over 7000 visits per month over the next three-year period of July 2021 to June 2024 according to projections.

Conclusion: The forecast of outpatient visits in this study serves as early signals to the management of the University hospital and is intended to enhance human and material resource planning and allocation for better-quality healthcare delivery.

Keywords: Time series analysis, Modelling Dynamic behaviour, Outpatient department, Stationarity, Box-Jenkins Methods, Unit root testing, Forecasting, Outpatient visits.

Introduction

Forecasting Out-Patient Department (OPD) visits in healthcare centres have increasingly induced large interest in both theoretical and applied perspectives [1]. Accurately forecasting the number of patient visits in hospitals is key in the administration of human capital as well as essential equipment and materials resources of the hospital [2]. The OPD is the gateway to almost all of the services the hospital renders to the general public and can come under increasing pressure annually due to increasing patient volume [3]. To this end, accurately forecasting the OPD visits of the University of Cape Coast Hospital becomes very critical and beneficial. The ability to predict outpatient visits is crucial for resource planning and allotment as well as efficient scheduling of appointments in the OPD, which is aimed at avoiding overcrowding and providing high-quality patient healthcare service [2].

Time series refers to a metric sequence of a given variable recorded over regular time intervals. The values of a time series sequence can be recorded hourly, daily, weekly, monthly, quarterly, or even annually. Time series data consist of four main components: trend, seasonality, cyclical, and irregularity. Time-series techniques are invaluable tools in many fields of research including health and have been greatly exercised. One study [4] used the autoregressive integrated moving average (ARIMA) models of the Box-Jenkins methods to model the monthly OPD attendance of patients at the Cape Coast Teaching Hospital (CCTH) in the central region of Ghana. The study results revealed a non-seasonal ARIMA (2, 2, 1) as the best fit model for the hospital. Residual analysis of the model using the Ljung-Box test revealed that the model was adequate only for the forecasting purposes of the hospital. Another study [5] compared the ARIMA techniques of the Box-Jenkins methodology and exponential smoothing models to predict the number of monthly outpatient visits in a tertiary hospital in Egypt. The study results revealed that the non-seasonal ARIMA (3, 1, 3) model was more sensitive than the exponential smoothing models and hence was selected as the best fit model for the forecasting of the outpatient visits at Mansoura University Children Hospital, Egypt. Two students [6] conducted a study where they applied the ARIMA techniques in forecasting the monthly OPD attendance in Obuasi Government Hospital, Ghana. The study results revealed a non-seasonal ARIMA (2, 1, 0) model as the best fit model for forecasting the monthly OPD data for the hospital. The study recommended that its findings were beneficial for the hospital's future resource allocation and planning.

According to some studies [7, 8], client overcrowding for essential healthcare services typifies OPDs in hospitals in developing countries. To this end, overcrowding, long hours of waiting to access healthcare, coupled with inadequate staff at healthcare centres, are detrimental to client healthcare experience at the facility and represent inefficiency and loss of productivity in the economy at large [9]. Previous representative studies [4, 5, 6] have focused on modelling and predicting OPD attendance in various health facilities which are different in scope and capacity compared to the University of Cape Coast Hospital. This implies that those models may be unsuitable for application to the case of outpatient dynamics at the University of Cape Coast Hospital. Therefore, this study seeks to model and forecast the OPD attendance at the University of Cape Coast Hospital. The study is purported to aid in the proper logistical planning and allotment of the hospital's scarce resources for optimal performance.

Materials and Methods

Data Set

The research paradigm employed in this study is positivity, which emphasises the objective approach. This paradigm and approach support quantitative and secondary data research. Monthly periodicity data on out-patient department visits in UCC Hospital were collected from January 2012 to June 2021 from the District Health Information Management Systems Two (DHIMS-2). The sample size was 114 monthly observations. The Box-Jenkins methodology which this study adopted is suitable for application on a sample size of 50 observations and above [10]. The data was processed using *Eviews* 12.

The Theoretical Framework

Broadly speaking, there are five approaches to forecasting based on time series data: (1) exponential smoothing methods, (2) single-equation regression models, (3) simultaneous-equation regression models, (4) autoregressive integrated moving average (ARIMA) models, and (5) vector autoregression (VAR) models. Among these forecasting methods, two of them have become so popular in the literature. The methods include autoregressive integrated moving average (ARIMA), popularly known as the Box–Jenkins methodology [11], and vector autoregression (VAR).

The publication by Box and Jenkins [11] of *Time Series Analysis: Forecasting and Control* ushered in a new generation of relatively powerful forecasting tools for which this paper intends to employ. These methods emphasize analysing the probabilistic, or stochastic, properties of the time series on their own under the philosophy *let the data speak for themselves*. The Box-Jenkins time series models allow the series Y_t (say monthly OPD attendance) to be explained by its past, or lagged, values and stochastic error terms. For this reason, ARIMA models are sometimes called *atheoretic* models because they are not based or derived from any theory.

AR, MA, and ARIMA Modelling of Time Series Data

An Autoregressive (AR) Process

Let Y_t represent monthly OPD attendance series at time *t*. Y_t can be modelled as

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + u_t \qquad (1)$$

Where δ is the mean of *Y* and where u_t is an uncorrelated random error term with zero mean and constant variance σ^2 (i.e. *white noise*), then Y_t follows a first-order autoregressive, or AR (1), stochastic process. Here the value of *Y* at time *t* depends on its previous time period (*t-1*) and a random term u_t. Similarly, an AR (2) can be modelled as

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + u_t \qquad (2)$$

Hence, a p^{th}-order autoregressive process is modelled as

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + \cdots + \alpha_p\left(Y_{t-p} - \delta\right) + u_t \qquad (3)$$

In the AR process, only the current and previous *Y* values are involved; there are no other regressors.

The Moving Average (MA) Process

When a series is regressed against its error term and past or lagged values of the error term then the process is a moving average (MA) process. Hence, an MA(1) process is modelled as

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} \quad (4)$$

Where μ is a constant and u, as mentioned before, is the white noise stochastic error term. Equation (4) means that Y follows a first-order moving average, or an MA(1), process. Generally, a q^{th} order MA process or MA(q) can be expressed as

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \cdots + \beta_q u_{t-q} \quad (5)$$

Simply put, a moving average process is simply a linear combination of white noise error terms.

An Autoregressive and Moving Average (ARMA) Process

Sometimes a series may exhibit characteristics of both AR and MA processes together and is, therefore, an ARMA process. For instance, an ARMA (1, 1) process can be expressed as

$$Y_t = \theta + \alpha_1 Y_{t-1} + \beta_0 u_t + \beta_1 u_{t-1} \quad (6)$$

where θ represents a constant term together with one autoregressive and one moving average terms. The generalised ARMA (p, q) process can be expressed as

$$Y_t = \gamma + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \beta_0 u_t + \sum_{j=1}^{q} \beta_j u_{t-j} \quad (7)$$

where model contains p lags of the dependent variable, q lags of the error term with γ being a constant.

An Autoregressive Integrated Moving Average (ARIMA) Process

The classical Box-Jenkins [11] methods work on the assumption that the time series involved is (weakly) stationary. However, it is well known from the literature that most time series variables are non-stationary, that is, they are not *integrated;* for example, the out-patient department (OPD) attendance time series of the UCC Hospital for the period of January, 2012 to June, 2021 is non-stationary at level and has to be integrated of order 1(i.e., it is I[1]), to become stationary. Therefore, if a time series is $I(d)$, after differencing it ***d*** times, then an $I(\boldsymbol{d})$ series is obtained. Therefore, if a time series would have to be differenced ***d*** times to make it stationary and then apply the ARMA (p, q) process model to it, then the original time series is an ARIMA (p, d, q), that is, it is an autoregressive integrated moving average time series, where ***p*** denotes the number of autoregressive terms, ***d*** the number of times the series has to be differenced before it becomes stationary, and ***q*** the number of moving average terms in the series. Hence, an ARIMA (2, 1, 2) time series, for instance, has to be differenced once ($d = 1$) before it becomes stationary, and the (first-difference) stationary time series can be modeled as an ARMA (2, 2) process; that is, the series has two AR and two MA terms.

The Box–Jenkins (BJ) Methodology

ARIMA modelling has gained popularity because it is more successful in forecasting and is also considered more reliable relative to the traditional methods of forecasting. The classical Box-Jenkins [11] methodology is a four-step process for handling the ARIMA models: Identification, Estimation, Diagnostic checking, and Forecasting.

Step 1. Identification: the first process in developing the ARIMA model is to identify the parameters of the model (*p, d, q*), where ***p*** is the order of the autoregressive (AR) process, ***d*** is the number of differencing to make the series stationary (if it is not at level) and ***q*** is the order of the moving average (MA) process. This is where candidate ARIMA models can be identified and chosen for estimation and evaluation. The main tools for the identification process include the autocorrelation function (ACF), and the partial autocorrelation function (PACF) with their corresponding correlograms. It must be noted that the ACF and PACF of AR(p) and MA(q) processes have opposite markings; in the case of the AR(p) process the AC declines exponentially, whereas the PACF process cuts off to zero after the first lag. The exact opposite is true for the MA(q) process. Table 1 gives the theoretical application of the ACF and the PACF in the model identification process.

Table 1: *Theoretical Application of ACF and PACF*

Structure of model	Typical pattern of ACF	Typical pattern of PACF
AR(*p*)	Spikes decays exponentially towards zero	Significant spikes cut off to zero
MA(*q*)	Significant spikes cut off to zero	Spikes decays exponentially towards zero
ARMA(*p, q*)	Exponential decay	Exponential decay

Step 2. Estimation: After identifying the appropriate parameters for the model **(***p, d, q***)**, the next step is to estimate the parameters of the autoregressive and moving average terms included in the model using least squares methods or maximum likelihood estimators. Nonlinear (in parameter) estimation methods can be resorted to if necessary.

Step 3. Diagnostic checking: This step kicks in after estimating and evaluating candidate ARIMA models and selecting the best under certain criteria such as model with maximum number of significant coefficients, minimum level of volatility, highest R squared adjusted value, minimum Akaike Information Criteria (AIC), Bayesian Schwartz Information Criteria (SIC/BIC) or Hannan-Quinn Criteria (HQ) value. For this reason, the classical Box-Jenkins methodology is often regarded more as an art than a science [12]; considerable skill is required to choose the right ARIMA model. After fitting the appropriate ARIMA model, the goodness of fit is estimated by plotting the ACF of residuals of the fitted model. If most of the sample autocorrelation coefficients of the residuals lie within the limits $(-1.96/\sqrt{N}, +1.96/\sqrt{N})$, where N is the number of observations, then the residuals are *white noise* indicating that the model fit is appropriate. Hence, there may not be any need to look for another ARIMA model. Stability in the model can also be tested using Ljung-Box test of autocorrelations.

Step 4. Forecasting: If a candidate model successfully passes through the first three stages, that model is considered reliable and stable and hence can be used for forecasting. The objective of Box–Jenkins methodology was to identify and estimate a statistical model which could be interpreted as having generated the sample data of the series involved [13]. Alternatively, the auto-arima function of a statistical software package can be deployed to determine the best model fit for the data.

Results

Descriptive Analysis

Time plot of OPD attendance in UCC Hospital for the period January, 2012 to June, 2021is depicted in figure 1. This figure indicates an overall steady upward or increasing trend of out-patient department visits in the hospital over the study period, though suffered the heaviest deep around the period of the onset of the Coronavirus Disease 2019 (COVID-19) pandemic in Cape Coast. Tables 2&3 respectively, give values of twelve-month rolling average for the study period and beyond.

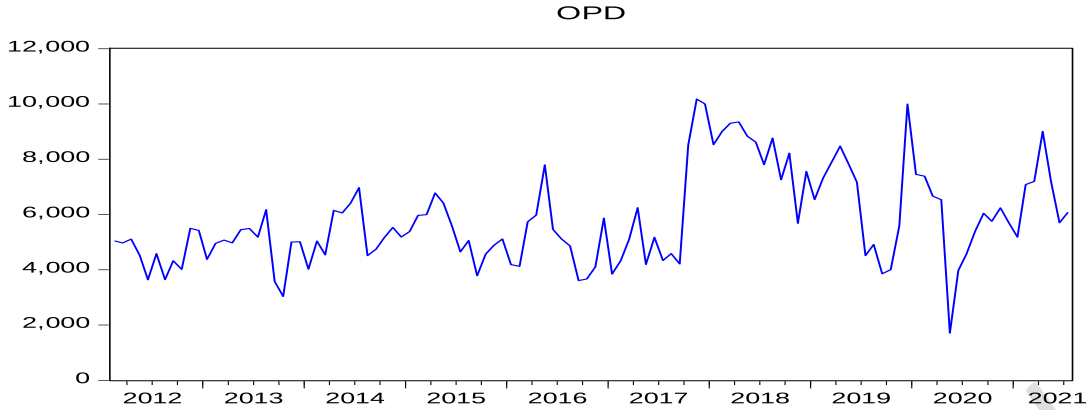

Figure 1: Time plot of UCC Hospital OPD visits over January, 2012 to June, 2021.

Table 2: Twelve month moving average values of actual OPD visits over January, 2012 to June, 2021 in UCC Hospital

Yr/ Mn	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2012												4597
2013	4589	4598	4587	4665	4819	4869	5080	5017	4936	4894	4861	4831
2014	4838	4793	4891	4941	5017	5166	5028	5125	5302	5346	5360	5474
2015	5551	5673	5725	5755	5688	5494	5540	5460	5410	5356	5351	5251
2016	5097	5076	5009	5124	5113	5151	5135	5120	5045	4980	5043	5014
2017	5030	4977	5000	4700	4676	4611	4588	4638	5042	5547	5891	6282
2018	6671	7021	7279	7666	7953	8242	8590	8844	8820	8446	8243	8077
2019	7937	7820	7747	7665	7544	7269	6949	6665	6314	6306	6509	6585
2020	6590	6488	6327	5815	5550	5556	5596	5778	5924	5977	5619	5430
2021	5404	5449	5655	6113	6256	6382						

Table 3: Twelve month moving average values of projected OPD visits over July, 2021 to June, 2024 in UCC Hospital

Year/ month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2022						6799	6881	6942	6988	7025	7056	7083
2023	7107	7129	7151	7172	7192	7212	7231	7251	7270	7290	7309	7328
2024	7347	7367	7386	7405	7424	7444						

From Table 2, the twelve-month rolling average rose to over 8000 visits per month in the second half of 2018 and slumped to over 5000 visits per month in the same period for 2020. However, it rose to over 6000 visits per month in the second quarter of 2021. This increasing trend is expected to continue into the forecast region, as can be seen in Table 3, where the rolling average hovered over 7000 visits per month in the last quarter of 2022 and onward.

Stationarity Test

The classical Box-Jenkins [11] approach works on the assumption that the time series involved is (weakly) stationary. Hence, the autocorrelogram and partial autocorrelogram functions together with Augmented Dicker-Fuller (ADF), Kwiatkowski–Phillips–Schmidt–Shin (KPSS), and Phillips–Perron (PP) unit root tests were deployed to determine the stationarity or otherwise of the out-patient department series variable in the hospital. Tables 4, 5, & 6 give the results of the ADF, KPSS, and PP unit root tests of the OPD variable respectively.

Table 4: *ADF Unit Root Test of OPD visits in UCC Hospital at 5% significance level*

OPD variable	At level		1st Difference	
Equation Form	*t-statistic*	*p-value*	*t- statistic*	*p-value*
Intercept	-4.407724	0.0005	-13.16089	0.0000
Trend & Intercept	-4.841538	0.0007	-13.10179	0.0000
None	-1.069937	0.2560	-13.21977	0.0000

Null Hypothesis: OPD/D(OPD) has a unit root

Table 5: *KPSS Unit Root Test of OPD visits in UCC Hospital at 5% significance level*

OPD variable	At level		1st Difference	
Equation Form	*LM- statistic*	*critical-value*	*LM- statistic*	*critical-value*
Intercept	0.553930	0.463000	0.052800	0.463000
Trend & Intercept	0.087746	0.146000	0.051024	0.146000

Null Hypothesis: OPD/D(OPD) is stationary

Table 6: *PP Unit Root Test of OPD visits in UCC Hospital at 5% significance level*

OPD variable	At level		1st Difference	
Equation Form	*Adj. t- statistic*	*p-value*	*Adj. t- statistic*	*p-value*
Intercept	-4.366106	0.0006	-17.04204	0.0000
Trend & Intercept	-4.858412	0.0007	-16.95654	0.0000
None	-0.644487	0.4360	-16.81909	0.0000

Null Hypothesis: OPD/D(OPD) has a unit root

From Tables 4, 5, & 6 it is clear that the three modes of unit root testing have unanimously confirmed that the OPD attendance variable at the UCC Hospital is stationary at first difference. This implies that the ARIMA form of the classical Box-Jenkins approach can be applied to the OPD data of the hospital. The ADF and PP have three forms or equations and all these must agree on the stationarity or otherwise of a particular time series variable according to the empirical literature. These forms include an intercept, trend & intercept and none and they must all agree on the stationarity or otherwise of a given series. From Tables 4-6 it is clear that the OPD attendance time series variable is integrated of order 1 i.e. $I(1)$. This means that the series is not stationary at level. Hence the OPD attendance time series variable attained stationarity at 1st difference. In addition, the ACF and PACF of the OPD attendance series of the hospital revealed many of the spikes to be outside the standard error bounds for the ACF. Also, the PACF cuts of to zero after the first significant lag. This is indicative of non-stationarity. Hence, the OPD attendance variable of the University of Cape Coast hospital is not stationary at level as confirmed by the ADF, KPSS, and PP formal unit root tests.

Model Selection Criterion and Validation

The classical Box-Jenkins approach works on the assumption that the series should be stationary. By deploying the Box-Jenkins methods on the first differenced OPD series variable and using the auto-arima function in *EViews 9* in order to find the best fitting model for the OPD attendance series of the hospital, the following tentative models in Table 7 were obtained. The system generated twenty-five tentative models and selected ARIMA (4, 1, 4) as the best fitting model for the OPD data. The results of ARIMA (4, 1, 4) provided the lowest AIC and BIC values against other competitive models such as ARIMA(2, 1, 2), ARIMA(3, 1,

2), ARIMA(3, 1, 3), ARIMA(4, 1, 2), ARIMA(2, 1, 4), ARIMA(1, 1, 1) and many other competitive models.

Table 7: *ARMA Selection Criteria Table for UCC Hospital OPD Visits*

Model	LogL	AIC*	BIC	HQ
(4,4)(0,0)	-951.909052	16.875597	17.115615	16.973007
(2,2)(0,0)	-956.329690	16.882977	17.026987	16.941423
(3,2)(0,0)	-956.101722	16.896521	17.064534	16.964708
(3,3)(0,0)	-956.099849	16.914032	17.106046	16.991960
(4,2)(0,0)	-956.101501	16.914061	17.106075	16.991989
(2,4)(0,0)	-956.103092	16.914089	17.106103	16.992017
(1,1)(0,0)	-960.594635	16.922713	17.018720	16.961677
(3,4)(0,0)	-956.097964	16.931543	17.147559	17.019212
(4,3)(0,0)	-956.098716	16.931556	17.147572	17.019225
(2,1)(0,0)	-960.359734	16.936136	17.056144	16.984840
(1,4)(0,0)	-958.407075	16.936966	17.104978	17.005153
(3,1)(0,0)	-959.419844	16.937190	17.081201	16.995636
(1,2)(0,0)	-960.424342	16.937269	17.057278	16.985974
(4,1)(0,0)	-959.027840	16.947857	17.115869	17.016044
(0,3)(0,0)	-961.220918	16.951244	17.071253	16.999949
(1,3)(0,0)	-960.932288	16.963724	17.107735	17.022170
(0,4)(0,0)	-961.209101	16.968581	17.112591	17.027026
(3,0)(0,0)	-964.372934	17.006543	17.126551	17.055248
(0,1)(0,0)	-966.732093	17.012844	17.084849	17.042067
(1,0)(0,0)	-967.007090	17.017668	17.089673	17.046891
(0,2)(0,0)	-966.008718	17.017697	17.113704	17.056661
(4,0)(0,0)	-964.371836	17.024067	17.168078	17.082513
(2,0)(0,0)	-966.992944	17.034964	17.130971	17.073928
(0,0)(0,0)	-969.875758	17.050452	17.098455	17.069934
(2,3)(0,0)	-969.668892	17.134542	17.302554	17.202729

Table 7 displays the candidate models identified and their corresponding AIC estimates with the objective of selecting the model that best fit the UCC Hospital OPD attendance series. The criterion for choosing the most appropriate model is the model with the lowest AIC value. In Table 7 the results showed that ARIMA (4, 1, 4) proved to be the most appropriate model fit because it has the lowest AIC value. Hence, diagnostics checks were performed in order to validate the selected model for forecasting (see Figure 3). Taken notice of the fact that ARIMA models are *atheoretic* models, the following estimated parameters in Table 8 for the selected model are tabulated. The estimated variance of the residual noise term is 1479628. The log likelihood is -951.909052 with AIC value given as 16.875597 as the lowest information criterion value.

Table 8: *Estimated Parameters for ARIMA (4, 1, 4)*

ARIMA(4, 1, 4)	AR(4)	MA(4)	SIGMASQ
Coefficients	*0.980571 **	*-0.909494**	*1479628**
SE	0.061321	0.124985	152516.8

*Figures in italics are significant ($P < 0.05$)

Figure 2 display a comparative graphic representation of all top twenty competitive models with ARIMA (4, 1, 4) leading as the best fit model. The model has the minimum AIC value among the top twenty auto-selected models (see Figure 2).

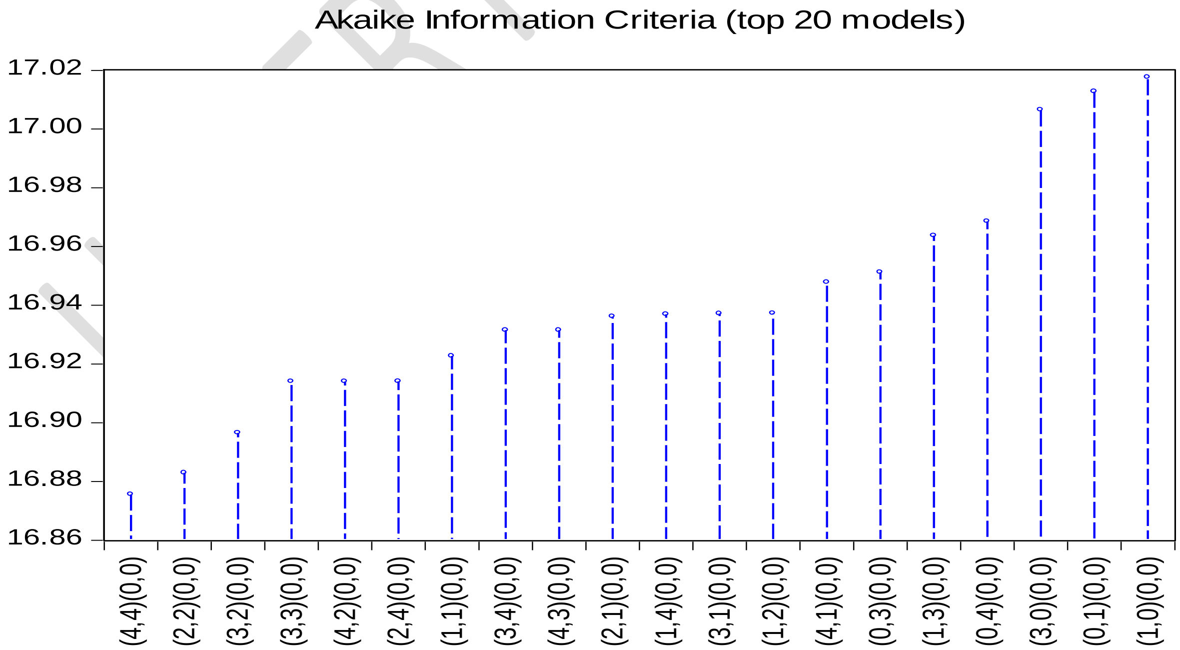

Figure 2: Comparative graphic representation of top twenty competitive models

The Ljung-Box Q-test of autocorrelation was performed to test the null hypothesis that the selected model ARIMA (4, 1, 4) do not exhibit autocorrelation in its residuals for a fixed number of lags against the alternative that some autocorrelation coefficient is non-zero. Figure 3 display the correlogram of residuals involving the ACF, PACF, and Q-Stats and the corresponding *p-values* of the selected model. From Figure 3 it is clearly indicated that the Ljung-Box Q-test gave insignificant *p-values* up to lag 36, meaning the null hypothesis is not rejected, therefore, the sample auto-correlation coefficients of the residuals lie within the limits $(-1.96/\sqrt{N}, +1.96/\sqrt{N})$ otherwise known as the standard error bounds. This implies that the errors of the selected model are *white noise* and hence model is stable and adequate for forecasting.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.108	0.108	1.3612	
		2	-0.013	-0.025	1.3823	
		3	-0.128	-0.126	3.3268	
		4	-0.088	-0.063	4.2585	
		5	-0.108	-0.099	5.6676	
		6	-0.044	-0.043	5.8985	
		7	0.117	0.108	7.5797	
		8	0.071	0.020	8.2088	
		9	0.043	0.014	8.4382	0.004
		10	0.029	0.038	8.5419	0.014
		11	0.048	0.062	8.8388	0.032
		12	-0.024	-0.002	8.9125	0.063
		13	-0.033	0.001	9.0498	0.107
		14	-0.058	-0.044	9.4859	0.148
		15	-0.071	-0.065	10.154	0.180
		16	0.057	0.073	10.597	0.226
		17	-0.013	-0.049	10.620	0.303
		18	0.042	0.008	10.856	0.369
		19	-0.103	-0.122	12.311	0.341
		20	-0.058	-0.058	12.781	0.385
		21	-0.210	-0.205	19.037	0.122
		22	-0.004	0.025	19.040	0.163
		23	0.040	-0.000	19.275	0.202
		24	0.114	0.057	21.180	0.172
		25	0.070	0.025	21.894	0.189
		26	-0.025	-0.036	21.991	0.232
		27	-0.106	-0.090	23.684	0.209
		28	-0.100	-0.008	25.218	0.193
		29	-0.052	-0.020	25.633	0.221
		30	-0.143	-0.153	28.856	0.149
		31	-0.153	-0.183	32.545	0.089
		32	-0.003	-0.040	32.546	0.114
		33	0.091	0.019	33.891	0.110
		34	0.016	-0.083	33.935	0.137
		35	0.047	-0.002	34.298	0.158
		36	0.182	0.145	39.916	0.067

Figure 3: Correlogram of Residuals for Selected model ARIMA(4, 1, 4)

Model Forecasting

After demonstrating superiority over other competitive models and being validated for forecasting, ARIMA (4, 1, 4) was subjected to forecasting the OPD series of the UCC Hospital for the next 36-month period. Figure 4 depicts forecast comparison of the selected model (red colour) against competing models (grey colour). Figure 5 displays the actual forecast over the 36-month period of July 2021 through June 2024 whereas Figure 6 depicts the trend of the entire time plot of the study period as well as the forecast period. Table 9 presents the 36-month point forecast values of OPD expected visits (health demand) in UCC Hospital over July 2021 through June 2024.

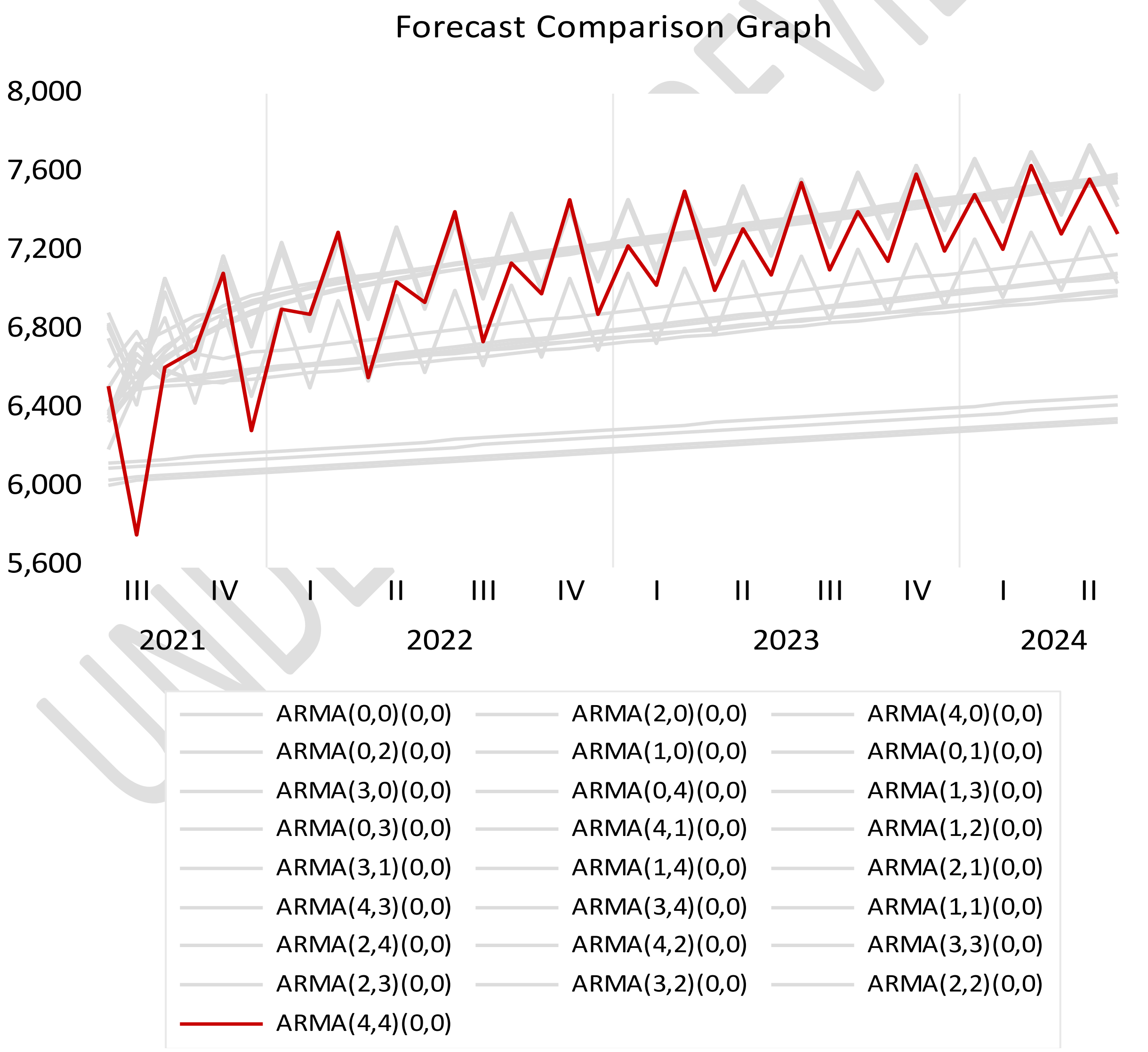

Figure 4: Forecast comparison graph of all competitive models with best model in red

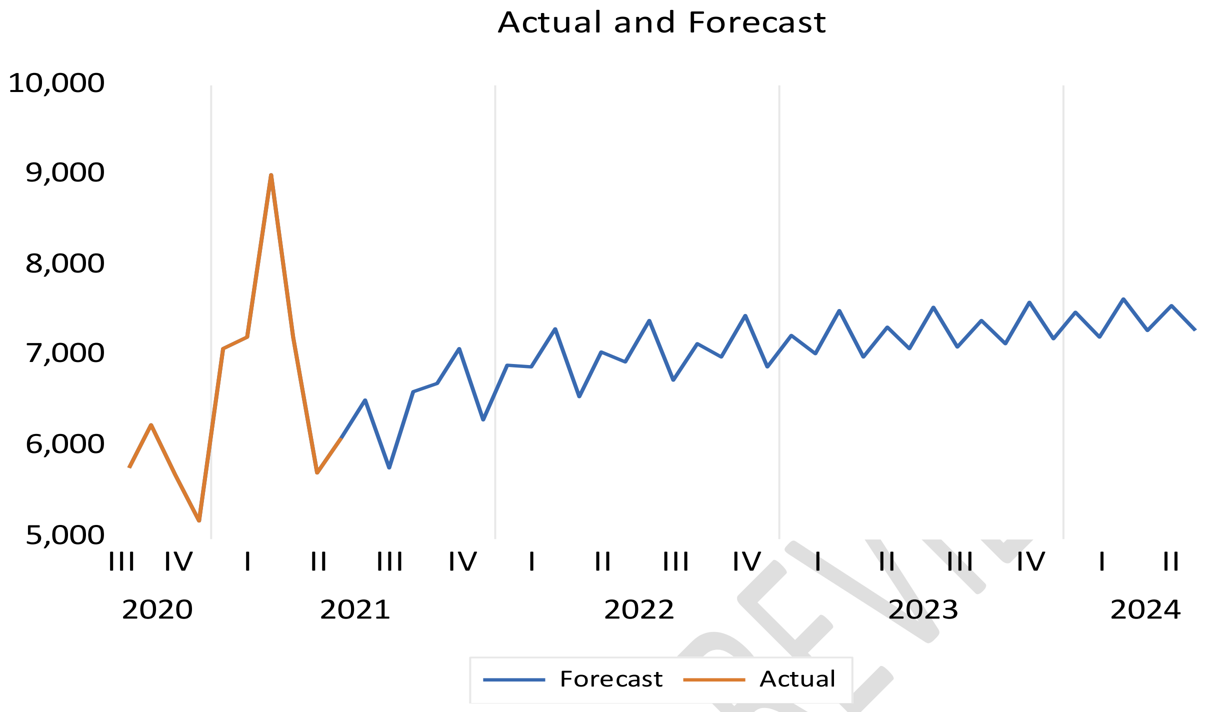

Figure 5: Forecast plot of OPD visits in UCC Hospital over July, 2021 to June, 2024

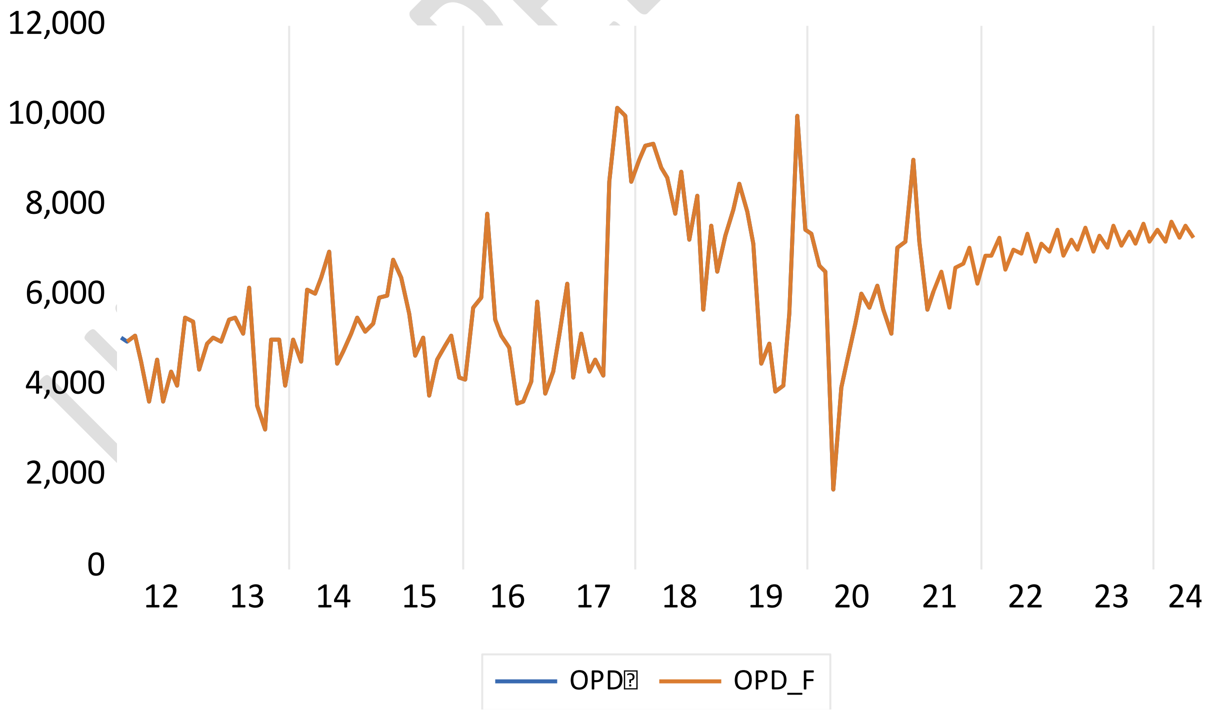

Figure 6: Actual and Forecast time plot of OPD visits in UCC Hospital over January, 2012 through June, 2024

Table 9: *Point Forecast of OPD visits in UCC Hospital over July, 2021 through June, 2024*

Yr/Mn	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2021							6510	5752	6606	6694	7081	6287
2022	6903	6873	7294	6556	7040	6937	7393	6734	7137	6979	7453	6874
2023	7225	7024	7500	6993	7311	7077	7544	7099	7397	7139	7588	7195
2024	7481	7208	7635	7283	7563	7283						

Discussion

The study results have shown an overall rising trend in outpatient department visits at UCC Hospital. This increasing trend in OPD visits is expected to continue over the next three years. A twelve-month rolling average shown in Table 2 exhibits a monthly average OPD attendance as high as over 8000 in the second halve of 2018 slumped to over 5000 visits per month on average in third and fourth quarters of 2020 and thereafter rose to over 6000 visits per month on average in the second quarter of 2021. This increasing trend in the 12-month rolling average of OPD visits is expected in the next three years with over 7000 visits per month starting from the last quarter of 2022 as displayed in Table 3 according to the selected best fit model forecast results. Thus, the University of Cape Coast Hospital will experience an overall significant increase in outpatient department visits in the next three years as projected in Table 9. The findings show an expected increase (see Figure 5 & Table 9) in the number of OPD visits (health demand) in UCC Hospital per month over the next three years. This signals that there is a huge task ahead and management as well as all staff at all levels of the University Hospital must be well prepared for this momentous task or responsibility if quality healthcare delivery (as a core mandate) must be met.

Among the twenty-five non-seasonal tentative ARIMA models formulated in Table 7 using auto-arima function in *EViews 9*, ARIMA (4, 1, 4) was selected as the best fitting model for the OPD attendance data of the University Hospital over the study period. The OPD attendance observations did not show any form of seasonality. This might apparently mean that morbidity as well as OPD attendance in the hospital do not exhibit seasonality. However, as a limitation of the study, this interpretation could be erroneous since reported visits and morbidities at the facility were not segregated. Further studies are suggested for the veracity or otherwise of this conclusion. The results of the chosen model ARIMA (4, 1, 4) gave the lowest AIC value compared to the rest of the tentative candidate models such as ARIMA (2, 1, 2), ARIMA(3, 1, 2), ARIMA(3, 1, 3), ARIMA(4, 1, 2), ARIMA(2, 1, 4), ARIMA(1, 1, 1) and many others. The various information criteria for the selected model include an AIC value of 16.875597 and a log-likelihood estimate value of -951.909052. The estimated variance of the residual noise term is 1479628. Diagnostic checks were conducted on the standardized residuals of the fitted model. The results revealed that the residuals of the model ARIMA (4, 1, 4) were *white noise*. Thus, this validated the goodness of fit of the selected model for the out-patient department attendance of the hospital. Hence, ARIMA (4, 1, 4) model have provided an adequate predictive model for the prediction of out-patient department visits of the University of Cape Coast Hospital.

The increasing or upward trend in OPD attendance revealed in this study also concurs with the findings of other studies [1, 4, 5, 6]. Similar to this study's findings was one study [4] that also had a non-seasonal model best fit for the Cape Coast Teaching Hospital (CCTH). ARIMA (2, 2, 1) was selected as the best fit model for that study with least residual noise for forecasting purposes of the hospital. The forecast results indicate that the outpatient department attendance cases will increase over the period of June, 2020 to December, 2025. As part of recommendations that study requested the CCTH health authorities to expand the OPD services in order to absorb the expected increase in OPD cases.

Doctors and paramedical staffing levels at UCC hospital therefore need to be logistically positioned if quality healthcare demand is to be adequately met without undue long waiting times which can lead to patient dissatisfaction.

Conclusions

The study has successfully identified a model that best fits the hospital's observed OPD attendance data and accurately projected estimates of OPD attendance for the hospital for the next three years. The study has revealed that there has been an overall increasing trend in outpatient visits to the University Hospital over the study period despite the many heavy shocks encountered in the system. However, the outpatient monthly visits did not exhibit any seasonality though most of these shocks were temporary. The forecast period also indicated a steady increasing trend of expected outpatient visits to the hospital. The study results showed that ARIMA (4, 1, 4) model is significantly superior to all other competitive models in the study and has remarkably outperformed them. This paper therefore proposes this model which effectively handles forecasting precision in the context of outpatient visits to the

University of Cape Coast Hospital in order to address the difficulties of complex hospital outpatient visits dynamics now and in the future.

Ethical Approval

Ethical clearance was obtained from the institutional review board of University of Cape Coast (UCCIRB). Permission was also obtained from the management of the University of Cape Coast Hospital.

Consent

As per international standard or university standard, patients' written consent has been collected and preserved by the author(s).

References

[1] Yu L, Huang G, Tang L, Zhao Y, Lai KK. Forecasting Patient Visits to Hospitals Using a WD & ANN-based Decomposition and Ensemble Model. Eurasia Journal of Mathematics, Science and Technology Education. 2017; 13 (12): 7615 – 7627.

[2] Hadavandi E, Shavandi H, Ghanbari A, Abbasian-Naghneh S. Developing a Hybrid Artificial Intelligence Model for Outpatient Visits Forecasting in Hospitals. Applied Soft Computing. 2012; 12 (2): 700 – 711

[3] Luo L, Luo L, Zhang X, He X. Hospital Daily Outpatient Visits Forecasting Using a Combinatorial Model Based on ARIMA and SES Models. BMC Heath Sciences Research. 2017; 17 (469): 1 – 13

[4] Borbor BS. Modelling and Prediction of Outpatients Department on Hospital Attendance at the Cape Coast Teaching Hospital Using the Box-Jerkins ARIMA Model. Journal of Advances in Mathematics and Computer Science. 2020; *35*(8):1-12. https://10.9734/jamcs/2020/v35i830309

[5] Elgohari H, Bakr AM, Majeed AM. (2019). Forecasting the number of outpatient visits in tertiary hospital using time series based on ARIMA and ES models. *Australian Journal of Basic and Applied Sciences.* 2019;*13*(7):70-77. https://10.22587/ajbas.2019.13.7.11

[6] Banor F, Gyan F. *Modelling Hospital Attendance in Ghana: A case study of Obuasi Government Hospital, Project work, Garden City University*. (Bachelor's thesis, Garden City University College, Ghana). 2012. Accessed on 2nd December, 2021. Available: https://www.academia.edu/6626411/Modeling_hospital_attendance_in_Ghana_A_case_of_the_Obuasi_government_hospital

[7] Bahadori M, Teymourzadeh E, Ravangard R, Raadabadi M. Factors Affecting the Overcrowding in Outpatient Healthcare. *Journal of Education and Health Promotion*. 2017; *6*: 21

[8] Yarmohammadian MH, Rezaei F, Haghshenas A, Tavakoli N. Overcrowding in Emergency Departments: A Review of Strategies to Decrease Future Challenges. *Journal of Research in Medical Science. 2017; 22*: 1 – 16.

[9] Oostrom T, Einav L, Finkelstein A. Outpatient Office Wait Times and Quality of Care for Medical Patients. *Health Affairs.2017; 24* (1): 1 – 5.

[10] Yanovitzky I, VanLear A. Time series analysis: Traditional and contemporary approaches. In A. F. Hayes, M. D. Slater, & L. B. Snyder (Eds.), *The Sage sourcebook of advanced data analysis methods for communication research*. 2008; (pp. 89–124). Sage Publications, Inc. https://doi.org/10.4135/9781452272054.n4

[11] Box PE, Jenkins GM. *Time Series Analysis: Forecasting and Control,* revised ed., Holden-Day, Oakland, California, USA, 1976: 2nd edition.

[12] Gujarati DN, Porter DC. Basic Econometrics. 5th Edition, McGraw-Hill/Irwin, Boston; 2009

[13] Pokorny M. *An Introduction to Econometrics.* Basil Blackwell, New York. 1987; p.343.