KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

A TIME SERIES ANALYSIS OF NUMBER OF PATIENTS IN FIVE DEPARTMENTS OF HO REGIONAL HOSPITAL

By

AMEXO HABEL MAWUSI

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS, KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF MSc. INDUSTRIAL MATHEMATICS

June 16, 2019

Declaration

I hereby declare that this submission is my own work towards the award of the master of science degree in industrial mathematics and that, to the best of my knowledge, it contains no material previously published by another person nor material which had been accepted for the award of any other degree of the university, except where due acknowledgment had been made in the text.

Amexo Habel Mawusi	....	....
Student	Signature	Date

Certified by:

Dr. George Parker-Lamptey	....	....
Supervisor	Signature	Date

Certified by:

Prof. Mrs. Atinuke O. Adebaji	....	....
Head of Department	Signature	Date

Dedication

I dedicate this work to my dear wife, Mrs. Rejoice Amoako and my lovely son, Caleb Yesutor Darkey. It is also dedicated to my brother, Bright Darkey and all my sisters.

Abstract

Continuous increase in human population and economy, causes a rise in healthcare demand. Overcrowding in the departments of the hospitals of Ghana specifically has become an increasingly important problem of late. The study aimed to develop a time series model for forecasting the monthly volume of patients admission to the five departments of the Ho Regional Hospital (Traflagar). To achieve the objectives of the present study, the researcher herein employed time series analysis to study the random pattern of admission into the various departments of the aforementioned hospital and as such the Box-Jenkins procedure was used to fit ARIMA and seasonal ARIMA (SARIMA) models to the data collected. Analysis of the data indicated that the various series show seasonality and hence the Seasonal ARIMA (SARIMA) models were better in fitting the admission data (having smaller mean squared errors of forecasting) than the non-seasonal ARIMA models. It was further observed that the admission to the emergency ward will fluctuate steadily from January to August. Admission will decrease in September and October and rise in November and December for the year 2019. Moreover, the admission to the male ward will have significant in March, May, June and September respectively for the 2019 year. The female ward admission for 2019 will increase but with study fluctuation between January and June. Admission reduces afterwards and an increase in October and November. In addition, the admission to the maternity ward tend to be very stable with little fluctuation. Admission fluctuates steadily for all months of the year. Finally, the forecast for children for the four years saw a significant increase in December.

Acknowledgments

I would like to express my profound gratitude to the Lord Almighty for the strength given me for the success of this work. I forever remain indebted to my supervisor and mentor, Dr. George Parker- Lamptey for his love, patience and professional directions for the success of this work. I am also very grateful to other lectures who in one way or the other helped throughout the entire programme. I am also grateful to my wife for her assistance both financially and spiritually from the start to the completion of this programme. Not forgetting Mrs. Kafui Bansah, a nurse at Ho Regional Hospital for her supports in the course of collecting data for this work.

Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

1.1 Background to the study

Persistent increase in human population results in a rise in healthcare demand. Hospital congestion in present day has become a major challenge faced by hospitals nationwide. The problem of crowding in healthcare units affects patients' satisfaction, effective nursing, treatments, waiting time and patients' length of stay. (Bahadori et al., 2017). In China, a study on overcrowding with the focus on outpatient department showed that more than ninety percent of patients who attend hospital for treatment without staying there overnight were provided care at the general health centers, which makes patients encounter an essential aspect a standard healthcare (Bao et al., 2017) and in the emergency departments (Yarmohammadian et al., 2017). Congestion in the hospital wards is an essential issue that needs to be given serious attention since this mostly affects quality of service delivery. Inadequate supply of patient beds to admit new patients would result in overcrowding. Most often than not, patient bed may not be enough to admit all patient due to the high number needing medical care. (Lingling and Alper, 2018).

Prediction of admission is an important activity which helps in hospital census, measuring the intensity of nursing care required by patients, allotment of resources and the overall management of the hospitals to improve service performance. (Lingling and Alper, 2018). Several studies on hospital management focus on the arrival of demand prediction (Mai et al., 2015), forecasting of outpatient visits (Cheng and Li, 2008), patient discharge (Zhu et al., 2017), and patient volume (Abdel-Aal and Mangoud, 1998). done on forecasting the volume of patients visiting hospital for medical attention. Forecasting new admission patients is very key that helps managers to make realistic decisions, cutting down congestion and providing better healthcare (Lingling and Alper, 2018).

Forecasting using time series method has been embraced in many areas of study such as explaining periodic pattern in the number of goods sold in a given period, projecting volume of patients visiting the hospital, making estimate of product that have been newly introduced base on units sold, to mention a few. (Song et al., 2017). Some of these methods, are suitable for solving problems that deals with forecasting linear sequence of discrete-time data. The ARIMA model has a linear property since its ability to predict future values is limited to be linear function of previous observations. However, because of the inability of the ARIMA model to capture the non-linear connections of time series in practice, there is restriction in its prediction accuracy. In a cramped situation, the forecasting accuracy of the artificial neural network (ANN) performs better due to its inherent characteristics in a nonlinear functions (Harrison and Kennedy, 2005). In recent time, the use of hybrid forecasting with combination of the auto regressive moving average (ARIMA) and the artificial neural network (ANN) model is commonly used in many studies due to its ability to predict efficiently (Zhou et al., 2014). Because of the seasonal instability of new admissions, seasonal ARIMA model is a good method that can assist in the projection of new patients admission. It is therefore crucial to make a perfect forecast for daily, weekly and monthly admissions beforehand to enable managers in deciding appropriate ways to improve on the demand for an efficient healthcare. Reis and Mandl (2003) conducted a study to detect individual and population health indicators that are discernible before confirmed diagnosis are made. They found out ARIMA model is an a perfect and dependable approach useful in forecasting medical data. The focus of the work was apply a time series method to forecast patient monthly visit to the hospital. The ARIMA model came into being by the tireless work of Box-Jenkin in 1976 and it is now very useful in many field of study including medical sciences. Initially, the model was developed to predict economic data but later became useful in medicine to project patient data.

Driven by challenges in resource scheduling and overcrowding issues face by most health centers in Ghana, we selected a very key hospital – Ho Regional Hospital (Trafalga) for our study and consider a time-series forecasting problem of monthly admissions. Ho Regional Hospital is the regional and teaching hospital. It became a teaching hospital in 2015 under the administration of His Excellency Ex. President John Dramani Mahama, to serve the University of Health and Allied Sciences. Our study enveloped all the recorded admissions experienced at the at the health unit from January to December 2006 and 2016 respectively. The hospital serve a wide range of areas in the region since it is the only hospital with modern healthcare facilities. Some of the areas it serves include Peki government hospital, Anfoega government hospital, Margaret Maquat hospital and other surrounding health units within the region.

1.2 Statement of the problem

The health status in Ghana has not changed much since independence despite heavy investment in the health service. Within the last two decades the Ghana Health Service (GHS) has embarked on far reaching reform aimed at creating the necessary environment for reversing trend. The establishment of the Ghana Health Service is seen as a major strategy for improving service delivery thereby accelerating the health status of Ghanaians (Ghana Health Service, 2002). From the GHS (2011), OPD attendance continues to increase nationally with a current OPD per capita of 0.98 and 1.09 in the year 2010 and 2011 respectively. This rise cannot, be wholly attributed to full satisfaction of clients about quality service provided at the facilities. In addition, the GHS has brought to light that the proportion of OPD attendance by clients increased from 55.81% in 2010 to 82.11% in 2011 (Arthur, 2013). This increased in patient admission placed great pressure on the health system thereby making management of resources quiet challenging.

Overcrowding in the departments of the hospitals has become an increasingly important problem in the hospital of late. This makes quality management of the hospital an issue of great concern. Crowding occurs when the need identified for urgent service outweighs the resources available for patient care. This situation happens as a result of the rapid growth of patients visit and timely decrease in health facilities (Institute of Medicine, 2006). Overcrowding reduces the ability to offer good service to patients whose conditions require urgent attention (Aliyas, 2012). Besides, long period of crowding in the emergency units may have a calamitous effects on patients hence a rise in patient mortality (Ann Emerg Med. 2012). Sun et al. (2007) conducted a study on patient admission and found out that those patients admitted on days when the emergency department is crowded experienced 5% of patient mortality with confidence interval of 95%. These among other reasons driven the researcher for the study.

1.3 Objectives of the study

Considering the problem stated, the general objective of this study is to model patient flow into the Ho Regional hospital. Specifically, the study seeks to:

1. find out the best ARIMA model that can be used to forecast the data.

2. make a four year forecast for the data.

1.4 Methodology

Secondary data on the monthly admission to the five wards were obtained from Ho Regional Hospital (Traflagar). The ten (10)- year monthly data was analyzed using R Statistical Software.

For the objectives of our study to be achieved, the researcher herein employed time series analysis to study the random pattern of admission into the various departments of the aforementioned hospital and as such the Box-Jenkins procedure was used to fit non-seasonal ARIMA and seasonal ARIMA models of the data.

Forecasting ability for both models was compared using mean square error (MSE), criteria to choose the optimal model for further forecasting. Based on the fitted models, forecast values for admission to the five wards were obtained.

1.5 Justification of the study

Describing and summarizing time series data, making predictions and fitting models with low dimension are the purposes of time series analysis. Continuous increase in the patronage of the Ho Regional Hospital calls for the need to develop a model that will assist in preparing for personnel and logistics in advance.

The findings of this study would help the hospital management to adequately prepare for the large number of prospective patients. This would help them to prepare ahead with regards to manpower and logistical needs for a better service delivery to satisfy the need of patients. Besides, the Ghana Health Service (GHS) may also use the model to plan patient's hospital attendance for the nation at large. In addition, the government of Ghana may as well use the outcomes of the study to review its financial commitment and contribution to various health facilities in the country. The result of this work would also help in the academics for decision making.

1.6 Organization of the Study

The work is ordered in five chapters. Chapter one consist background to the study, problem statement, study objective, method employed, Justification as well as study organization. The second chapter considers previous work related to the thesis. Chapter three examines the methodology employed in the study and fourth chapter deals with result obtained from. Lastly, the fifth chapter presents recommendation and the study conclusion.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

This section discusses previous works of authors relation to the study. The section enveloped previous studies and reveals author's view on problem associated with patient admission.

2.2 Patients admission in Ghana

2.2.1 Cases at the OPDs

Poor quality of service delivery in our health centers is a crucial issue that requires urgent attention. Many lives are loss as a result of poor service delivery. The main objective of the Ghana Health Service (GHS) is to improve on the quality of service delivery at all health centers.

A study conducted by Goka (2007) about diseases reported at the outpatients departments in the Accra aimed at finding the trend of various diseases reported at the Outpatient Department, forecast 2007 reported diseases and finally, to identifying the most reported disease within the period. The data used was collected from the Adabraka Polyclinic in the Greater Accra; Ghana. The data used was from 1996 to 2006. The statistical instrument used for the study was time series analysis. Result indicated that half of the cases reported was malaria at the OPD yearly. Malaria, upper respiratory tract infection and skin disease constituted the great majority (about 52.4%) of diseases reported. It was also noted that all top ten diseases exhibited upward trends. Trend analysis of these diseases yielded various forecasted values for the year 2007.

Nyako (2002) investigated into the use of clinical services in La Polyclinic, Accra. He indicated that ,to improve upon the health status of Ghana, the government has assisted in providing adequate resources to help expand the health sector. From the Ministry of Health, despite the efforts made by the government, there has not been any proper improvement in the sector as expected. However, a five-year programme was carried out with the target of realizing appropriate ways of improving on the health status of Ghana. There were series of interventions to resource and restore the La Polyclinic to help cut down patient flow in the Regional and Teaching Hospitals.

Howerver, a comparison test was carried out on OPD attendance per capita for 1999 with 1997 national target. Result showed that, attendance was low. This reveals that despite the huge investment made by the government, attendance did not improve as expected. This study was done in the Kpeshie sub District of Accra Metropolis to evaluate the level of service rendered at the aforementioned health center and to elicit patient views regarding factors that contribute to the low level of service delivery to help find out the best way this situation could be curbed. Data was obtained from patient on admission and adults from the community. Four instruments were used including exist poll interview, focus group discussion, survey from the community etc. The researcher considered 8, 100, 100 and 6 respondents. Regarding the group, interview was carried out to find out issues that affect the utilization of clinical services. Outcome from the survey revealed that, even though OPD attendance turn rather low compare with the national target, there has not been any significant deviation experienced from the utilization level as expected, which implied that utilization was now optimal. However, to determine the accurate trend, a lower series of Out Patient Unit attendance data is required.

2.3 Patients' Concern at Hospital Admission

Wendy J. et al. (2011) carried a study on patient concern at hospital admission. A survey was conducted on patient's concern regarding quality service delivery. Two health centers in the university were considered. The category of people involved were physician in the hospitals and patients on admission to ascertain their concerns with regards to service delivery. In the course of the encounter, the researcher measured the encounter length. After meeting them, patients were rated on how well their encounter before was adressed. The following rating were used: "Not at all", "Somewhat", "Mostly"or "Completely". A codebook was developed to help describe the subject matter within topics that concerns patients. The book was finally coded in 11 groupings. A 20% sample of concern was 92% on at least one code and the rest takes 79%. They used logistic regression to evaluate the relationship that exist between concerns addressed, the encounter before and the length of encounter. 109 patients were considered in all, with a concern rate of 65% standard deviation. The encounter before recorded a median of 2 of patient concern ranging from 0 to 10. 87% of patients recorded at least one concern whereas 71% recorded multiple concerns. Concerns recorded more than 3 was 28% of the total number enrolled. Concerns with regards to treatment such as medication, etc. were considered. Others includes diagnosis and logistics.

Jenkins et al. (1994) undertook a study to find out some challenges faced by the staff of a family health service center in London when handling complex admission cases. A section of the hospital staff were made to complete a self-report questionnaire to provide information on complex admissions recorded. The survey covered 47-day period with 493 questionnaire completed by 111 staffs. 171(35%) of the cases reported various challenges during admission. One very key problem reported was unavailability of beds; a hindrance to admitting new patient specifically those with at least 75% years compare to the aged below 75%.

2.4 Factors Influencing Patient Admission

Flether et al., (1977) did a retrospective analysis of factors that affect the outcome and the long period of patient's stay in a stroke restoration center. 248 patients with average age of 67, ranging 17 to 78 years who were on admission over a one and half year period in the stroke unit were considered for the study. From the study, the percentage of patients who were able to return home after an average length of hospital stay in the stroke unit was 80 as equivalent a 43-day stay. During discharge, 85 percent was adapted for walking and those who do not need help in daily living activities was 56 percent. With regards to unsuitable outcome and rise in patient length of stay include the time taken in beginning admission, severity of weakness during admissions, etc. Patient age and other factors such as hypertension and weakness were not related to the functional state of patient on discharge or length of hospital stay. From the findings, patients with severe condition showed a significant improvement when admitted, as a result, they were discharged. The above result could help to predict which patient can respond quickly to treatment in a short term despite the severity of the condition having been in the hospital for a short period.

Yingxin et al., (2001) investigated into assessing the strength of AN-DRG version 3.1 in forecasting the changes in patient's length of hospital stay and to find out other factors that affect length of stay using data collected. In all, 18 DRGs consisting 4,589 section was analyzed. They used multiple regression to model patient length of hospital stay as a function of a number of variables whose variation does not depend on that of another. In totality, only 37.6 percent variation in average length of stay could be explained. Regarding total variation, DRCs forecasted 30 percent. Other factors that affect patient length of hospital stay include age, inability to settle hospital bills, classification, source of referrals and quality of physicians to mention a few. The conclusion was that the constrain was a result of a lack of proper sign of severity within DRG.

Devos et al. (2015) used an emergency and hospital data set to analyze factors that influence patient admission to the hospital for treatment and the cost associated with accidents on roads in Belgium. They considered 4645 patients who were admitted to the emergency units of a hospital in Brussels. They employed the Logistics Regression model together with a widely used linear model to examine the probability of control admission and cost, control road users, demographic such as sex, age, income or incidence of disease and clinical characteristics including location, seriousness of injury and nature. Findings from the study showed that 20.3 percent of patients who sought medical care at the emergency unit were given admission. The probability of patients admitted, controlled by other factors appeared greater considering patients from ages 0 to 16 recorded a confidence interval of 1.74 to 3.49 and then greater or equal to 60 years compared with patients within the age group of 30 to 44 years. Based on demographic and clinical factors, motorcyclists compared to pedestrians were less likely to be admitted with 95 percent confidence interval, thus from 0.39 to 0.94. Those with fractures and injuries in the body system were considered the highest probable patients to be admitted. The generalized logistic model proved that , factors such as the old aged and men as well as socioeconomic status patients were prone to higher hospital cost.

2.5 Forecasting Patients Admission

Time series is a sequential collection of information observed overtime. The temporal order may provide additional information due to the possible serial dependence of observations (Hellstenius, 2018). The main component of time series are the variations of trend and seasonality (Gwilym, et al., 2015). Trend is the increase or decrease of data overtime and seasonality is the oscillation that occurs in example a day, weekly or monthly interval. The seasonal variations are recurring while the trend is persistent. In the study of time-series, there are two main areas. These are analysis and forecasting. With the time-series analysis, it focused on analyzing a sequence to comprehend underlying properties of the data. In time series forecasting, future samples are predicted. Time series analysis is helpful in analysing the correlation of observation overtime and the underlying features of the generative process. This information can be used to identify a suitable model of the data to generate predictions. There are two categories of time-series analysis. When data is collected on a single variable, we referred to it as univariate. When the data in multiple variables, it is referred to as multivariate. Another distinction that we can make is as to whether the model is linear or non-linear. If the model is linear, the resulting predictions are a linear combination of a previous samples. On the other hand, non-linear models are not constrained to this property. By history, time series forecasting was dominated by linear approaches such as Auto Regressive models. However, for some sequences the assumption is not appropriate and this encourage the use of non-linear models (Holger and Schreiber, 2004).

Lingling and Alper (2018) investigated into using different models to forecast daily and monthly volume of new patient admission. They developed seasonal auto regressive integrated moving average and NARNN to forecast patient visits to the departments of the hospital. By comparing the forecasting ability of the models, root mean square error and the mean absolute error were employed. A six year data was considered for analysis. For daily modelling, a nine-months data set was used, thus from 4th January to 4th September, 2016. However, the time range for testing was September to October, 2016. Result from the monthly data revealed that, root mean square error, mean absolute error and others rather performed low compared to those discovered from single SARIMA or NARNN model. There appeared no significant improvement regarding mean absolute error and mean absolute percentage error regarding their modelling ability. Regarding the daily data, the models performed poorly regarding their root mean square error, mean absolute error and absolute percentage error at the testing and modelling. They concluded that hybrid model does not automatically outweigh its constituent's performances. It is therefore very important to test for model that can accurately forecast the volume of new admission patients using different data set.

Kibaek et al. (2014) employed different forecasting approaches to study patients volume in the hospital. From observation, they compared two models namely single and the multiple variables for their study. Data from Northwester Medicine Enterprise was used for the study. A 3-year data from January 2009 to June 2012 was considered.They compared the outcomes of the two methods; univariate and multivariate based on past means. To measure the accuracy of forecast, the mean absolute percentage error (MAPE) was employed. Besides, cross and auto correlation of patients volume of hospital attendance was examined. The study outcome revealed that the forecasting method performed best compared to the historical average based method. It lessoned the mean absolute percentage error by 65.1% in a next day forecast and 8.8% mean absolute percentage error for the next month. The auto regressive moving average(ARMA) performed better than other methods in the next day forecast. Conclusion drawn from their findings revealed that forecasting strategies are useful to perfectly predict patient volume of hospital visit.

Tariq (2016) conducted a study to forecast the patient's volume at radiology department of FMH hospital, Shadman Lahore, coming for ultrasound. Time series approach was employed to examine the collected data. METHOD: Monthly data of patients coming for ultrasound at radiology department, from January 2001 to March 2015 was used for fitting the best model. Result and Conclusion: the ARIMA model for patient volume was found appropriate, after residual diagnostic checks. The ARIMA model was used to forecast the volume of patient considering a two-year data. The error predicted was 1.1%. It was concluded that the fitted ARIMA is adequate model and can be used to forecast the patient's incoming to radiology department for future planning and management.

Spencer and Stat (2007) did a study on exploring and evaluating the use of several methods to forecast patient visits to emergency unit on daily base at three different health centres. The validity of these methods were compared to the past approved method. Data on patient daily visits to the hospital was collected for the study from January 2005 to March 2007. Seasonal auto regressive integrated moving average, exponential smoothing together with artificial neural network were used for predicting patient daily visit to the hospitals. One to thirty days forecast were made beforehand. The accuracy of the forecast obtained by the method was compared with existing forecasting accuracy. Result: having performed goodness of fit test, all time series methods performed well. However, analysis made from the sample showed that time series regression models showed a little improvement with regards to post-sample forecasting accuracy in relation to multiple linear regression models but other models did not show consistency in forecasting patient volume.

Baeta, (2015) in his study employed the Vector Autoregressive (VAR) modeling approach to model and to forecast hypertension and heart disease cases in the Ho municipality. The data was obtained from the municipal hospital and it spanned from the from the first day of January to December for 2010 and 2014 respectively. The result revealed that VAR was the best to model and forecast these CVDs. Diagnostic checks revealed the model is free of serial correlations and conditional heteroscedasticity. It also passed the stability test. The model was proposed for predicting hypertension and heart disease cases in the Ho municipality for the year 2015. The study revealed that there will be slight increase in cases of both conditions of hypertension and heart disease.

Lean et al., (2017) performed a study to forecast volume of patient visit to the hospital by combining optimal Sub-Band Tree Structuring (SB-TS) also referred to as Wavelet Packet Decomposition and artificial neural network (ANN). They first and foremost employed Wavelet Packet decomposition to breakdown the actual modelling data of patient hospital attendance into many parts and one remaining term. However, the artificial neural network was applied to fit various decomposed parts and produced individual forecasting outcomes. All single forecasting values were merged into the last prediction built by adding. Four groups of monthly data of patient visits was applied as the sample for illustration and the outcome revealed that the suggested model can obtain very important accuracy in forecasting results compare to other known methods.

Schweigler et al. (2009) did a study by comparing forecasting accuracies of two models such as time series and past average models to project the capacity of bed required in the emergency unit for a short term. Past admission records on patients who visited the emergency department was used. Values were obtained from three hospitals from June 2005 through July 2006. They developed three models for each site. Hourly, daily, weekly and monthly models were developed. The models include SARIMA, sinusoidal model as well as auto regression structure. They used log likelihood and Akaike's information criterion to compare extent to which observed data matched the expected value. To examine predictive accuracies, four to twelve hours forecasted values were evaluated. The evaluation was done by comparing forecasts of the model to the true discovered values for bed occupancy using the root mean square errors. Predicted errors were as well evaluated. Outcome from the study indicated that seasonal auto regressive integrated moving average model performed better compared to historical mean. The two models had given a better forecast accuracy for four to twelve hours forecast of bed capacity in the emergency department. There had not been any different performance by AR-based models.

Farmer and Emami (1990), conducted a study using time series and regression analysis to forecast the capacity of hospital bed required for new admission patient. The average time for patient length of stay, considered patient aged 15 to 44 year group and was used to evaluate different methods suitable for forecasting mean duration values of length of stay and its later use in planning hospital bed. Outcome of the study indicated that the simple trend fitting technique had problem with specifying model error and as such put undue restriction on the considered data. In modelling the data, time series approach is considered to be more appropraite. To conclude, time series method is more recommended to model the capacity of bed unlike the simple trend fitting method.

Arful, et al., (2005) carried out a research on patient bed occupancy in Seng hospital during the outbreak of SARS. They used auto regressive integrated moving average model to project the capacity of bed required for admission. Data used for the study was on past admission records and occupancy data for separate beds. A 3-month data was used for the study. Expected outcome was measured considering the number of separated beds occupied daily and number of patient admitted.The techniques employed for the analysis include splitting the outbreak data into two for model development. To model the required number of beds, structural auto regressive equation was applied. They made estimation through the maximum likelihood approach using the Kaima filter. Result from the study indicate that auto regressive integrated moving average model (1,0,3) forecasted and described the quality of hospital bed occupied during the outbreak. The validity and training set mean absolute percentage error of 5.7 and 8.6 percent.

Neeraj, et al., (2016) investigated into the present condition of planning hospital beds and developed a model to estimate hospital bed demand in the future using past data on bed occupancy. A 3-year data was used for the analysis, thus, from January 2015 to October 2018. To estimate the quantity of bed needed for quality service at the hospital without any delay, a time series of day-to-day midnight admissions including emergency transfer admission and elective patient were considered. Result indicated that, a hospital with 1600 bed capacity, it is estimated that 100 more beds are needed to admit patient without prolonging patient waiting time. To conclude, the time spent by patient in the emergency ward would reduced.

Muge et al., (2016) studied the timely NICU census of a health center and developed models to forecast hospital admission. They tested the model at the time patient were on admission with and without the data characteristics of patient and evaluated the model accuracy compared with the fixed average census method. A five year day-to-day NICU census data was applied to developed the model. 1827 observations were made for the five years duration and 365 for one year validation. They appliedlinear and auto regression to fit model for series of seven days forecasting period and compared these models using their error criteria. Result proved that, the census recorded slightly rise in linear trend. On average, forecasting accuracy has improved hence the model is good for forecasting hospital census compared with the fixed average census approach.

Izabel et al. (2013) researched on overcrowding in the emergency department of a general hospital to develop models that could forecast patient daily visits and to obtain an optimal model based on their forecasting accuracy. Six models were developed using patient records specifically on patient daily visit to the hospital in Sao Paulo, Brazil. They used the first 33 months of patient record to build a suitable model to help forecast patient admission at the emergency units of the aforementioned hospital. The MAPE was applied to assess the models' forecasting strength using the last 3 months data. The generalized linear model, generalized equation for estimation and the SAMIRA were the developed models for the forecast. The methods were each employed with or without the mean temperature effect as the variable for prediction. Result showed that daily number of patient visiting the facility recorded 389. The data showed seasonal distribution of patient volume on weekdays, recorded the highest as compare to patient volume on weekends. There appeared small change in patient daily visits monthly. Comparing the three models, SAMIRA could not perform very well unlike GEE and GLM models with respect to their forecasting accuracy. For example, the mean absolute percentage errors obtained from the generalized estimation equation in the first month, thus, October,2012 were 11.5 and 10.8 percent . In controlling the temperature impact, the forecasting strength performed rather low. The accuracy of forecast was better in short term rather long term.

Farid et al. (2014) did a study on overcrowding in the emergency department using time series analysis to develop and project patient flow. They developed model to forecast patient daily visits to the hospital's emergency unit in Lille, France. The study demonstrated the use of time series to forecast a short term patient demand for an urgent service in the emergency department of the hospital. They used the day-to-day record of patient visit to the paediatric unit to forecast patient flow in Lille regional hospital. A one year data was used for the study, considering the month January to December 2012. Auto regressive integrated moving average was used to make analysis of GEMSA categories and the overall patient visits. Result from their study indicate that, time series analysis is the appropriate methods for forecasting.

Boutsioli (2011), examined patients' reactions to emergency healthcare using data from a general hospital in Greek over a period of four years. demand using a sample data from Greek public hospitals over the period of four years. They evaluated the output of patient admission in separate groups; emergency and elective patient. Patient demand was assessed to find out the find difference between emergency patients and forecasted emergency admission. Result revealed that production reaction to patient demand has effect on hospital costs.

Spencer and Stat (2007) researched on real-time demand forecasting in the emergency department. Hourly counts of patient arrivals, laboratory orders, radiology orders, census, patient hospital admissions, and patient discharges were obtained from a single hospital emergency department for the period 2/15/2004 to 11/30/2006. No identifiable data was retrieved. Vector auto-regression (VAR) was a time-series method used for the evolution and dependencies of variables used in the model. Data for 104 weeks were used to develop the VAR model. Out of sample forecasts were made for hourly census, laboratory orders, and radiology orders. Forecast accuracy regarding the root mean squared error (RMSE) was assessed for horizons of 1 to 24 hours in advance using the 40 weeks of data not included in the training set. Results: VAR model provided greater forecast accuracy hourly patient census when compared to the use of forecasts based on the expected (mean) patent census for a given hour of the week.

Wang-Chuan, et al., (2017) applied time series approach to model and forecast patients visit to the emergency in Southern Taiwan. The study aimed at constructing appropriate model to forecast patients monthly visit to the emergency unit. Data collected was based on patients monthly visit to the emergency department to carry out an ARIMA analysis. The model was developed based on previous emergency department visit for a period of seven years. They employed a model to fit the forecast for monthly emergency unit visit up to 2016. The final model was assessed with the aid of the mean absolute percentage error (MAPE). The result of the study showed that the ARIMA structure, provided a mean absolute percentage error of 8.91 percent observed and obtained as patients visit.

Chapter 3

METHODOLOGY

3.1 Overview

The underlying theories and methods of time series employed in the analysis of data and discussion of results are presented within this chapter.

3.2 Sampling and Data Collection

The study would have considered all the major hospital in the country but Ho Regional Hospital (Trafalgar) was sampled out for the study. Data from the Out Patient Department (OPD) of the hospital was used. We sampled out a ten year data for the study. Thus, from 2006 to 2015. Data was collected from the folders of patients who attended hospital within the aforementioned period. We collected data on number of patients admitted from five departments of the hospital, namely: Emergency ward, Maternity ward, Female ward, Male ward and Children ward

3.2.1 Ho Regional Hospital

Driven by overcrowding and resource scheduling challenges arising in healthcare system in the Volta region; Ghana, we selected a very key hospital-Ho Regional Hospital (Trafalga) for our study and consider a time-series forecasting problem of monthly admissions. Ho Regional Hospital is the regional and teaching hospital. It became a teaching hospital in 2015 under the administration of His Excellency Ex. President John Dramani Mahama, to serve the University of Health and Allied Sciences. Our study enveloped all the recorded admissions experienced at the hospital from January 2006 to December 2015. The hospital serve a wide range of areas in the region since it is the only hospital with modern healthcare facilities. Some of the areas it serves include Peki government hospital, Anfoega government hospital, Margaret Maquat hospital (Kpando) and other surrounding health units within the region.

3.3 Concept of Time Series

3.3.1 Definition:

A collection of data points sequentially according to the time of their occurrence, where the times of collection are equally spaced. For instance if $Z_1, Z_2, Z_3, ..., Z_t$ are observations of a time series, then the observations are obtained at times 1,2, and so on. The time t is usually of the frequency such as: daily, weekly, monthly, quarterly, semiannually and annually etc.
Examples of time series includes:

- the daily stock prices
- the monthly rainfall recorded at a geographical area.
- The heart rate of a patient, minute by minute.

3.3.2 Components of Time Series

Trend Component

The trend component accounts for the gradual shifting of the time series to relatively higher or lower values over a long period of time.

Seasonal Component

The seasonal component is a pattern that occurs in the series throughout a calender time with fix period of occurrence. Seasonal pattern are usually determined by seasonal factors (eg. day of the week.)

Cyclic Component

A cyclic pattern exits when data exhibits rises and falls that are not of fixed period. This cyclic component is usually associated with the seasonal component.

Irregular Component

The irregular component represent the sudden changes occurring in the time series which are unlikely to be repeated. This variation is sometimes called residual or random component.

The four components of time series can be put in two ways to give the types of the time series models:

1. Additive Model:

$$Z_t = T_t + S_t + C_t + I_t \tag{3.1}$$

2. Multiplicative Model:

$$Z_t = T_t \times S_t \times C_t \times I_t \tag{3.2}$$

where:

Z_t: represents the observed value of the time series at time t.

T_t: represents the trend component of the time series.

S_t: represents the seasonality component of the time series.

C_t: represents the cyclic component of the time series.

I_t: represents the irregular component of the time series.

3.4 Model Estimation Using Time Series

In using time series analysis in modeling the collected data, the Box-Jenkins Approach of Modeling will be employed.

3.4.1 Box-Jenkins Approach of Modeling Time Series

This methodology is a statistical concept based framework which applies well defined principles to model a wide spectrum of time series. The process is outlined as:

1. plotting the time series data.

2. Transform the data (i.e make non-stationary data stationary)

   Some test of stationarity may be used to check if a given series is stationary or not. One of the widely used test is the

   The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test:

   The KPSS test the stationarity of a time series based on the hypothesis:

   $$H_0 : \text{The time series is stationary.}$$
   $$\textit{against}$$
   $$H_1 : \text{The time series is not stationary.}$$

   This means big p-value suggests the data is stationary and does not need to be transformed.

3. Identify the dependence orders of the ARMA (p,q) model:

   This step is achieved by observing the autocorrelations and partial autocorrelations of the stationary data and making conclusions based on Table 3.1.

Summary of Stationary Time Series Models

Since the methods employed under the Box-Jenkins procedure deals with stationary series, it is prudent to look at some stationary time series models we may may consider.

Autoregressive Series

The general autoregressive process Z_t of order p satisfies the equation:

$$Z_t = \delta + \psi_1 Z_{t-1} + \psi_2 Z_{t-2} + \cdots + \psi_p Z_{t-p} + e_t = \delta + \sum_{s=1}^{p} \psi_s Z_{t-s} + e_t \quad (3.3)$$

As the name suggests, autoregressive processes regress on themselves.

The Autoregressive Process of Order One -AR(1)

The AR(1) series is defines by

$$Z_t = \delta + \psi_1 Z_{t-1} + e_t,\ e_t \sim iid(0, \sigma_e^2) \quad (3.4)$$

with the following properties:

$$\mu_{Z_t} = \frac{\delta}{1-\psi_1}$$

$$Var(Z_t) = \frac{\sigma_e^2}{1-\psi_1^2}$$

$$\gamma_h = \psi_1^h Var(Z_t) = \psi_1^h \frac{\sigma_e^2}{1-\psi_1^2}$$

$$\hat{\rho}_h = \psi_1^h$$

Subject to the constraint that e_t be independent of $Z_{t-1}, Z_{t-2}, ...$ and $\sigma_e^2 > 0$, the solution of the AR(1) defining recursion $Z_t = \psi_1 Z_{t-1} + e_t + \delta$ will be stationary iff $|\psi_1| < 1$.

The Autoregressive Process of Order two -AR(2)

The AR(2) process is given by

$$Z_t = \delta + \psi_1 Z_{t-1} + \psi_1 Z_{t-2} + e_t \tag{3.5}$$

where it is assumed that e_t is independent of $Z_{t-1}, Z_{t-2}, Z_{t-3}, ...$ The characteristic equation of the AR(2) is given by:

$$1 - \psi_1 z - \psi_2 z^2 = 0 \tag{3.6}$$

From equation (3.6), a the AR(2) process has a stationary solution iff the roots of characteristic equation are greater than 1 in absolute value. Thus the roots should lie outside the unit circle in the complex plane.

The autocovariance function at lag h for the AR(2) process is given by:

$$\gamma_h = \psi_1 \gamma_{h-1} + \psi_2 \gamma_{h-2}, \quad h = 1, 2, 3, ...$$

and the corresponding autocorrelation function given by:

$$\hat{\rho}_h = \psi_1 \hat{\rho}_{k-1} + \psi_2 \hat{\rho}_{h-2}, \quad h = 1, 2, 3, ...$$

In general, the autoregressive process of order p has it autocovariance and autocorrelation functions at lag h given by the Yule-Walker equations given below:

$$\gamma_h = \psi_1 \gamma_{h-1} + \psi_2 \gamma_{h-2} + \cdots + \psi_p \gamma_{h-p} \tag{3.7}$$

Dividing through by γ_0 results in:

$$\hat{\rho}_h = \psi_1 \hat{\rho}_{h-1} + \psi_2 \hat{\rho}_{h-2} + \cdots + \psi_p \hat{\rho}_{h-p}, h > 0 \tag{3.8}$$

3.4.2 Moving Average Processes

A time series $\{Z_t\}$ is said to be an Moving average process of order q, MA(q) if it satisfies:

$$Z_t = \mu + e_t + \phi_1 e_{t-1} + \cdots + \phi_q e_{t-q} = \mu + \sum_{s=1}^{q} \phi_s e_{t-s} + e_t \qquad (3.9)$$

where: μ is the mean of the series and
e_t is a white noise. MA models thus give values of Z_t based on a linear combination of past forecast errors.

The Moving Average of Order one-MA(1) Process

For a zero mean MA(1) process,

$$Z_t = \phi_1 e_{t-1} + e_t \qquad (3.10)$$

The variance of the series is given by: $\gamma_0 = Var(Z_t) = \sigma_e^2(1 + \phi_1)$
$\gamma_1 = Cov(Z_t, Z_{t-1}) = Cov(e_t - \phi_1 e_{t-1}, e_{t-1} - \phi_1 e_{t-2}) = -\phi_1 \sigma_e^2$. This implies $\hat{\rho}_1 = \frac{-\phi}{(1 + \phi^2)}$
$Cov(Y_t, Y_{t-2}) = Cov(e_t - \phi_1 e_{t-1}, e_{t-2} - \phi_1 e_{t-3}) = 0$. similarly $Cov(Z_t, Z_{t-k}) = 0$ for $k > 2$ since there will be no $e^{'}s$ with common subscripts between Z_t and Z_{t-k}.

The Moving Average of Order two-MA(2) Process

For a zero mean MA(2) process,

$$Z_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + e_t \qquad (3.11)$$

$$\gamma_0 = Var(Z_t) = Var(e_t - \phi_1 e_{t-1} - \phi_2 e_{t-2}) = (1 + \phi_1^2 \phi_2^2)\sigma_e^2$$

$$\gamma_1 = Cov(Z_t, Z_{t-1}) = Cov(e_t - \phi_1 e_{t-1} - \phi_2 e_{t-2}, e_{t-1} - \phi_1 e_{t-2} - \phi_2 e_{t-2})$$

$$= (-\phi_1 + \phi_1 \phi_2)\sigma_e^2$$

$$\gamma_2 = Cov(Z_t, Z_{t-2}) = Cov(e_t - \phi_1 e_{t-1} - \phi_2 e_{t-2}, e_{t-2} - \phi_1 e_{t-3} - \phi_2 e_{t-4})$$

$$= -\phi_2 \sigma_e^2$$

$$\gamma_h = 0, \ \forall h > 2$$

Therefore, for an MA(2) process,

$$\hat{\rho}_1 = \frac{-\phi_1 + \phi_1\phi_2}{1 + \phi_1^2 + \phi_2^2}$$
$$\hat{\rho}_2 = \frac{-\phi_2}{1 + \phi_1^2 + \phi_2^2}$$
$$\hat{\rho}_h = 0, \ \forall h > 2$$

The above realization about MA(1) and MA(2) processes can be generalized for an MA(q) process, such that:

$$\gamma_h = \begin{cases} (1 + \phi_1^2 + \phi_2^2 + \cdots + \phi_q^2)\sigma_e^2, & \text{if } h = 0, \\ (\phi_h + \sum_{t=1}^{q-h} \phi_t \phi_{t+h})\sigma_e^2, & \text{if } h = 1, 2, ..., q \\ 0, & \text{otherwise} \end{cases} \quad (3.12)$$

Autoregressive Moving Average Processes (ARMA)

A time series $\{Z_t\}$ is an ARMA(p,q) process if it is stationary and satisfies the equation:

$$Z_t = \sum_{s=1}^{p} \psi_s Z_{t-s} + \sum_{s=1}^{q} \phi_s e_{t-s} + \mu + e_t \quad (3.13)$$

The parameters p and q are the orders of autoregressive and the moving average models respectively.

Using the back-shift operator, B, the ARMA model can be written as

$$(1-\psi_1 B-\psi_2 B^2-\cdots-\psi_p B^p)Z_t=(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p)e_t$$
$$\psi(B)Z_t=\phi(B)e_t$$

If B is replaced with x, then $\psi(B)$ and $\phi(B)$ can be replaced with the AR and MA polynomials respectively, defined as

$$\psi(x)=1-\psi_1 x-\psi_2 x^2-\cdots-\psi_p x^p$$
$$\phi(x)=1-\phi_1 x-\phi_2 x^2-\cdots-\phi_p x^p$$

Table 3.1: Characteristics of ACF and PACF of ARMA models

	AR(p)	MA(q)	ARMA(p,q), $p,q>0$
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cutts off after lag p	Tails off	Tails off

4. Parameter estimation:

   Once the order of the model (i.e p and q) have been identified, the parameters $\psi_1,...,\psi_p$ and $\phi_1,...,\phi_q$ must be estimated. A process which is accomplished using statistical softwares.

5. Model Diagnostics :

   The viability and adequacy of the model could be assessed to decide which model best fit the data. Some of the measures used include:

   - The Akaike's Information Criterion (AIC)
   - Schwarz's Bayesian Information Criterion (BIC)

The lowest value of each of the above criteria is used to decide which model fits the data adequately. In addition, the residuals of the adequate model should be:

i. uncorrelated.

ii. normally distributed.

iii. independent.

6. Forecasting

   The most important objective of building time series model is to be able to forecast/ predict the value for that series at future times.

   The forecasting error $\varepsilon_t = Z_t - F_t$. The forecast accuracy can be measured using:

   Mean Squared Error

$$MSE = \frac{\sum(\varepsilon_t)^2}{n}$$

   The best forecasting methods is chosen to be the one that gives a smaller mean squared error.

The process flow is given in the diagram below:

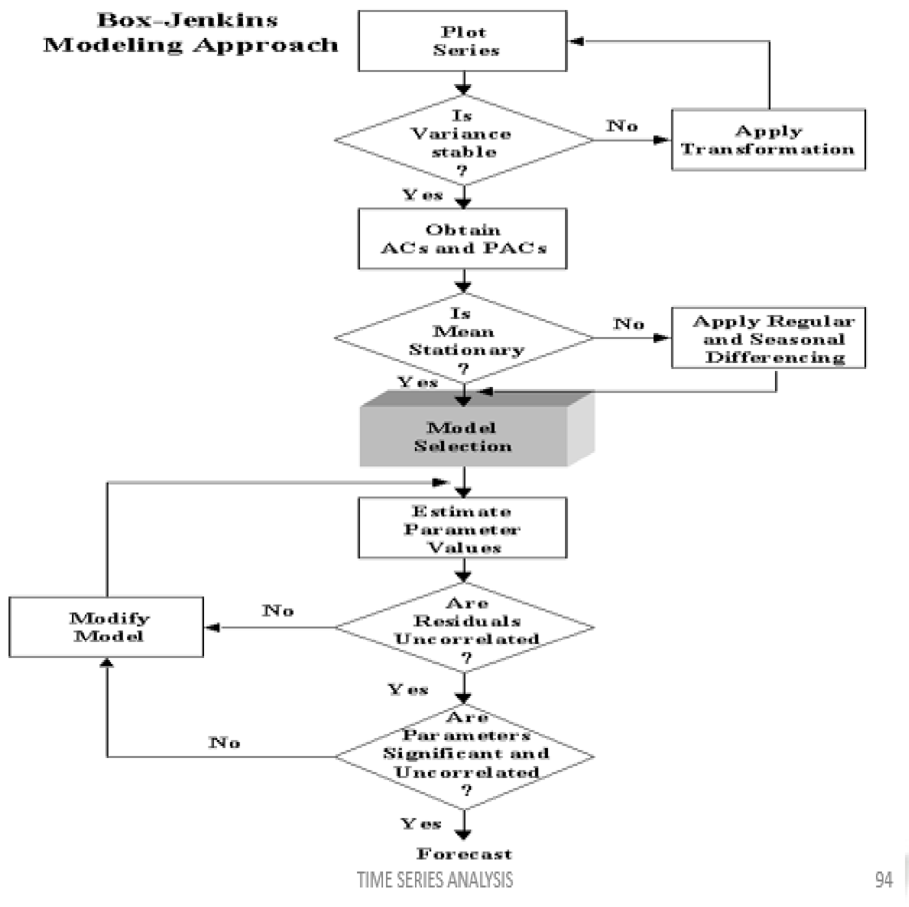

Figure 3.1: The Box-Jenkins methodology of modeling time series data

3.5 Empirical Model

The theoretical model employed in the analysis of the hospital data is the Seasonal ARIMA Model.

3.5.1 Seasonal ARIMA Model

The time series data my be affected by seasonal factors. The seasonality in the time series are patterns that occur throughout a calender time and reoccurs in periodical intervals say weekly, monthly, quarterly or annually. The seasonality parameter is usually denoted as S. That is when the seasonality occurs monthly, then $S = 12$ and $S = 4$ when the seasonality occurs quarterly.

For seasonal ARIMA model, the autoregressive and moving average terms predict Z_t using observed points and deviations at times and lags which are multiples of S. For example, when a series has quarterly seasonality, an AR (1) model will use Z_{t-4} to predict Z_t. More so, an AR(2) model will predict Z_t using Z_{t-4} and Z_{t-8}. Similarly, and MA(2) model will use e_{t-4} and e_{t-8} as predictors.

The seasonal ARIMA model comprises both non-seasonal and seasonal factors in a multiplicative model; ARIMA(p,d,q)×(P,D,Q)$_S$. For instance an ARIMA(1,0,1) series with a monthly seasonal pattern has it model written as;

$$(1 - \psi_1 B)(1 - \psi_{12} B^{12}) Z_t = (1 - \phi_1 B) e_t \tag{3.14}$$

when the AR has the seasonality component and then the equation becomes

$$(1 - \psi_1 B) Z_t = (1 - \phi_1 B)(1 - \phi_{12} B^{12}) e_t \tag{3.15}$$

when the seasonal component is on the MA part.

Chapter 4

RESULTS AND DISCUSSION

4.1 Overview

This chapter describes in detail the results of our study. It embodies the time series analysis of data collected and interpretation of the results.

4.2 Data Collection and Description

The data used for this study was obtained from the Ho Regional Hospital (Trafalgar). The monthly recorded admission of patients to the emergency, male, female, maternity and children wards constituted the data collected. An excerpt of the data is shown below.

emergency	Male	Female	Maternity	Children
24	20	35	31	15
22	15	30	34	22
29	23	28	15	10
14	31	14	10	23
18	14	20	29	13
16	23	27	24	17
25	17	31	30	20
11	31	17	33	10
35	27	15	30	19
27	33	26	30	27

Table 4.1: An excerpt of data collected from the Ho Regional Hospital (Trafalgar)

The analysis of data through time series modeling was carried out using the R-Statistical software.

4.3 Data Exploration

4.3.1 Descriptive statistics

Table 4.2: Descriptive statistics of variables in data

Statistics	Ward				
	Emergency	Male	Female	Maternity	Children
Min. Value	11	14	14	10	10
Lower Quartile	18	18	17	24	12.75
Median	24	21.5	22	28.5	16
Mean	24.92	23.02	22.39	26.69	16.44
Upper Quartile	30	29	27	31	20
Max. Value	43	34	35	35	30

4.3.2 Time plot of the various series

The time series plot of the monthly admission of patients to the five departments of the hospital are shown in figures below:

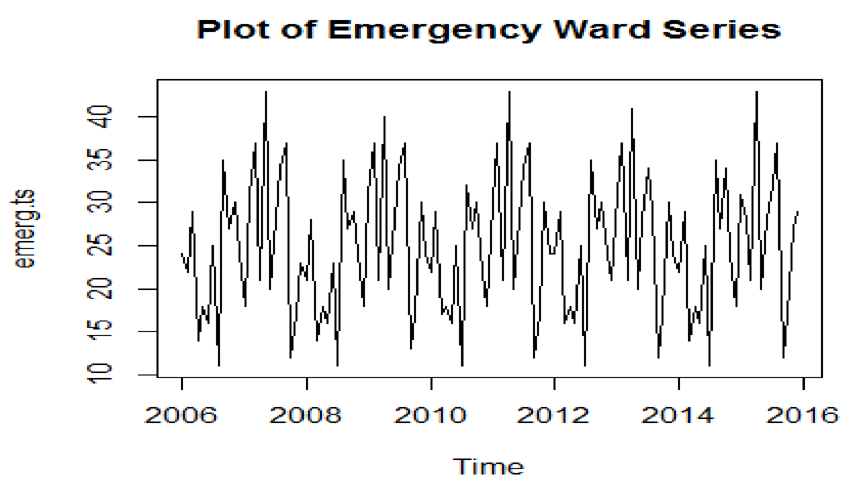

Figure 4.1: Plot of emergency ward series

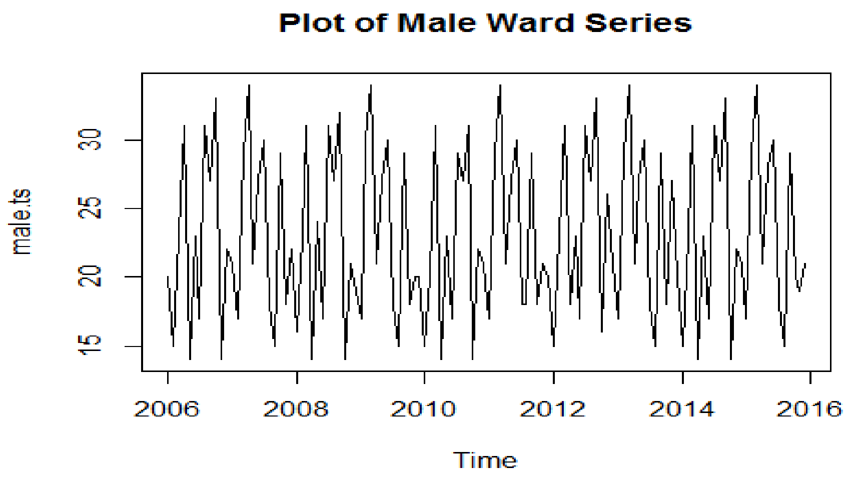

Figure 4.2: Plot of males' ward series

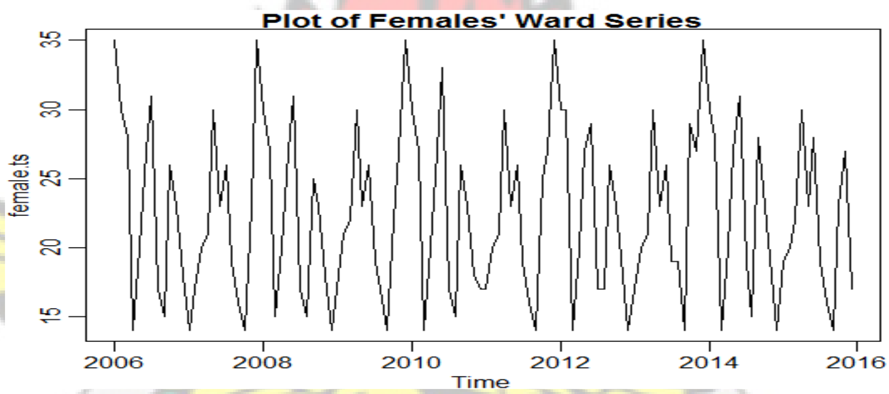

Figure 4.3: Plot of females' ward series

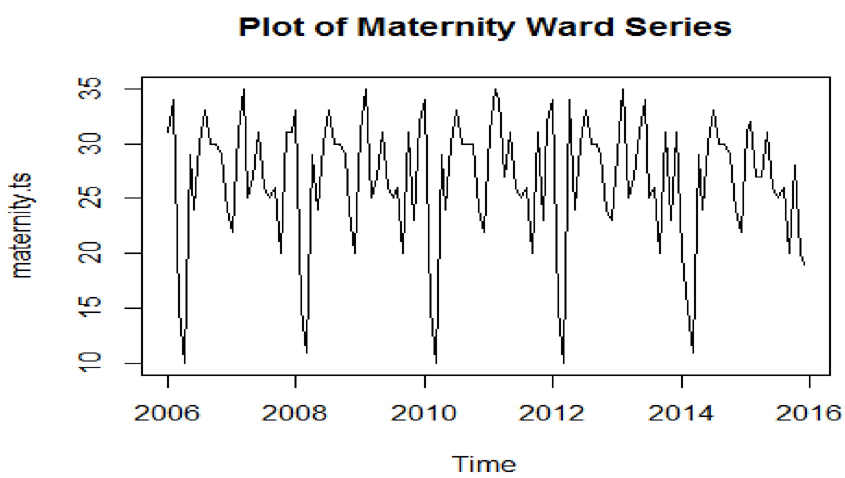

Figure 4.4: Plot of maternity ward series

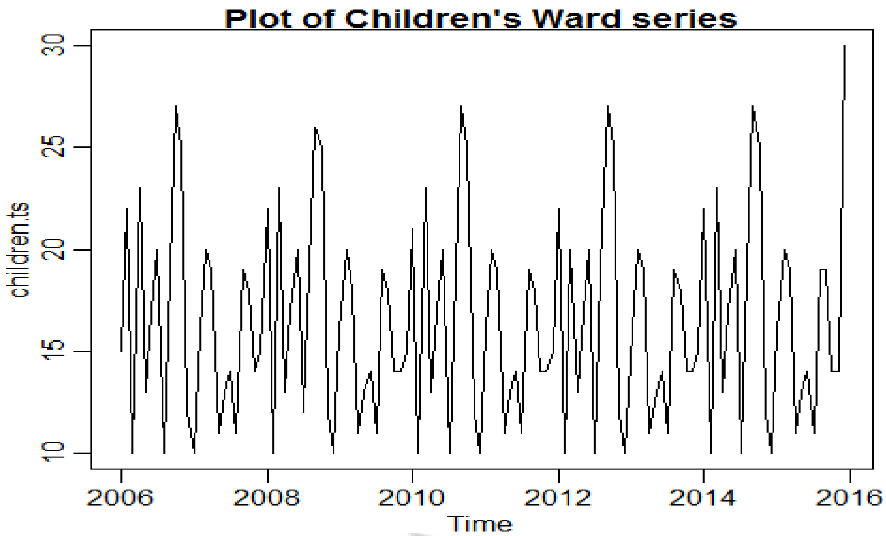

Figure 4.5: Plot of females' ward series

The plot of the raw series in figures above shows no trend and variability in the variance of the series over time. The series looks more to be stationary. However such observation could be misleading, hence the need to carry out the test of stationarity before further analysis could be done.

4.4 Time Series Analysis

Test of Stationarity

The KPSS test of stationarity was carried out and the result given below:

```
                KPSS Test for Level Stationarity

data:  emergency.ts
KPSS Level = 0.04923, Truncation lag parameter = 4, p-value = 0.1

         KPSS Test for Level Stationarity

data:  male.ts
KPSS Level = 0.036635, Truncation lag parameter = 4, p-value = 0.1
```

```
            KPSS Test for Level Stationarity


data:  female.ts
KPSS Level = 0.037401, Truncation lag parameter = 4, p-value = 0.1

           KPSS Test for Level Stationarity


data:  maternity.ts
KPSS Level = 0.052515, Truncation lag parameter = 4, p-value = 0.1

KPSS Test for Level Stationarity


data:  children.ts
KPSS Level = 0.038082, Truncation lag parameter = 4, p-value = 0.1
```

Since the p-value of the test is greater than 5% in each case, we fail to reject the null hypothesis of the KPSS test of stationarity and conclude that the series are stationary. Thus the modeling can be carried out without the need of transforming the series.

4.4.1 Model Identification

The ACF and PACF plots of the series are given below:

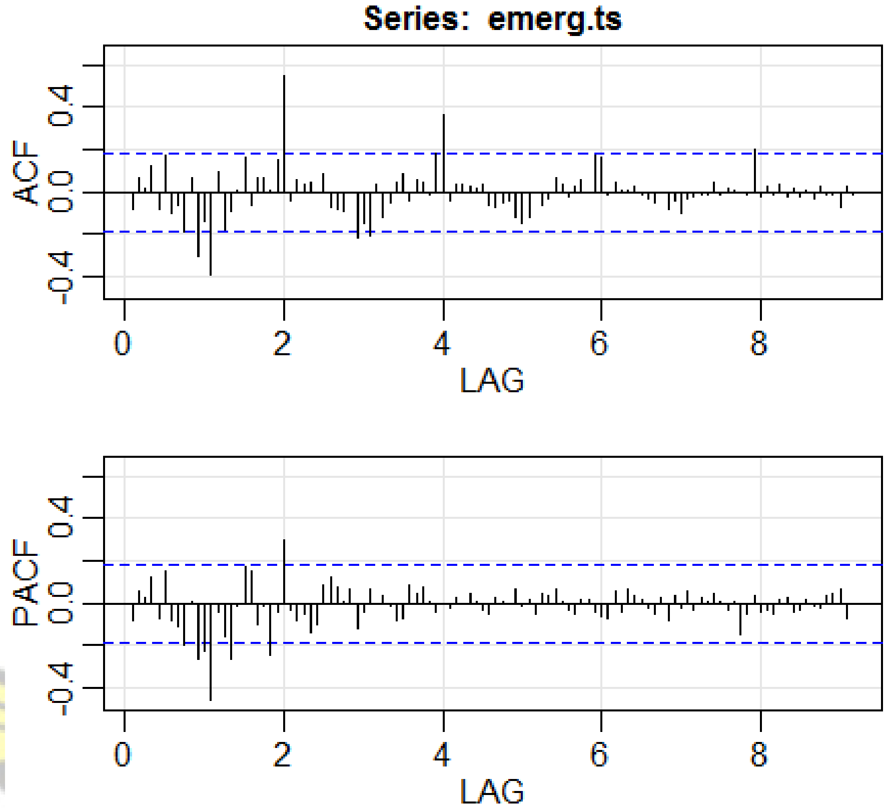

Figure 4.6: Plot of ACF and PACF of the emergency ward series

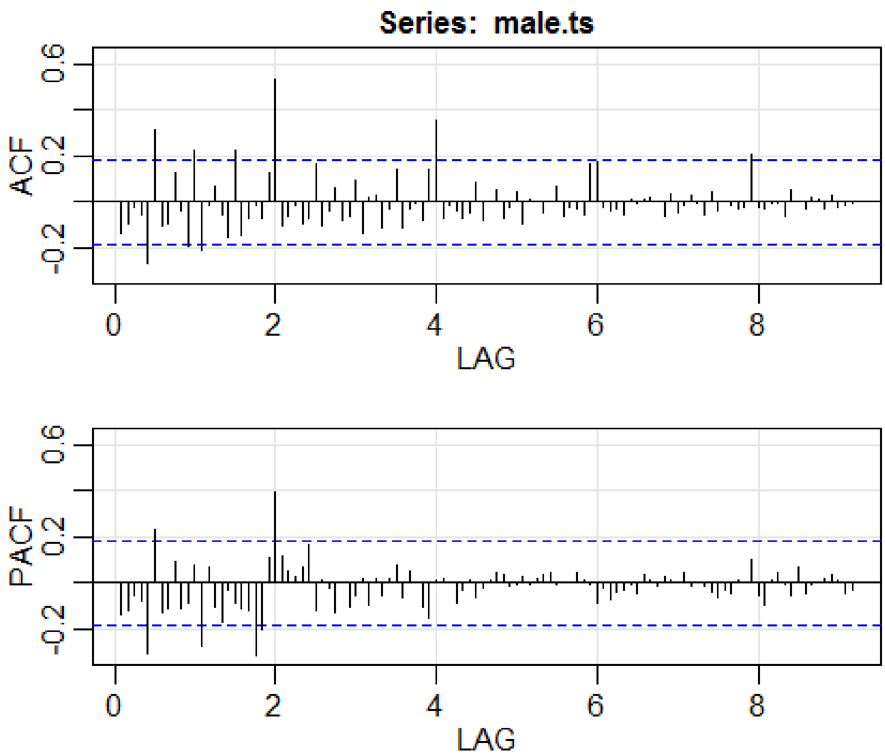

Figure 4.7: Plot of ACF and PACF of the males' ward series

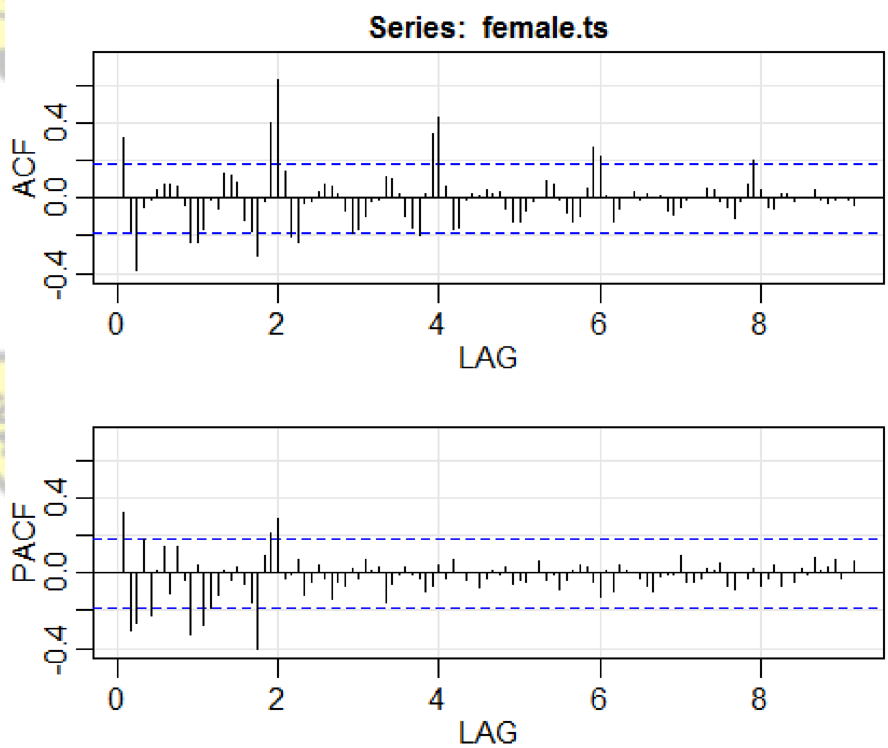

Figure 4.8: Plot of ACF and PACF of the females' ward series

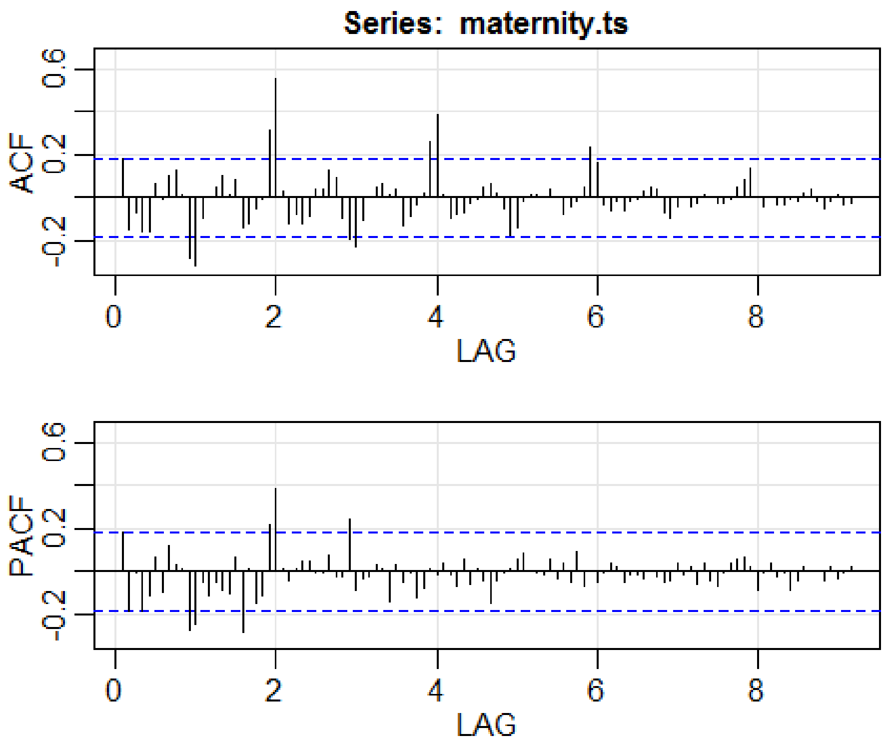

Figure 4.9: Plot of ACF and PACF of the maternity ward series

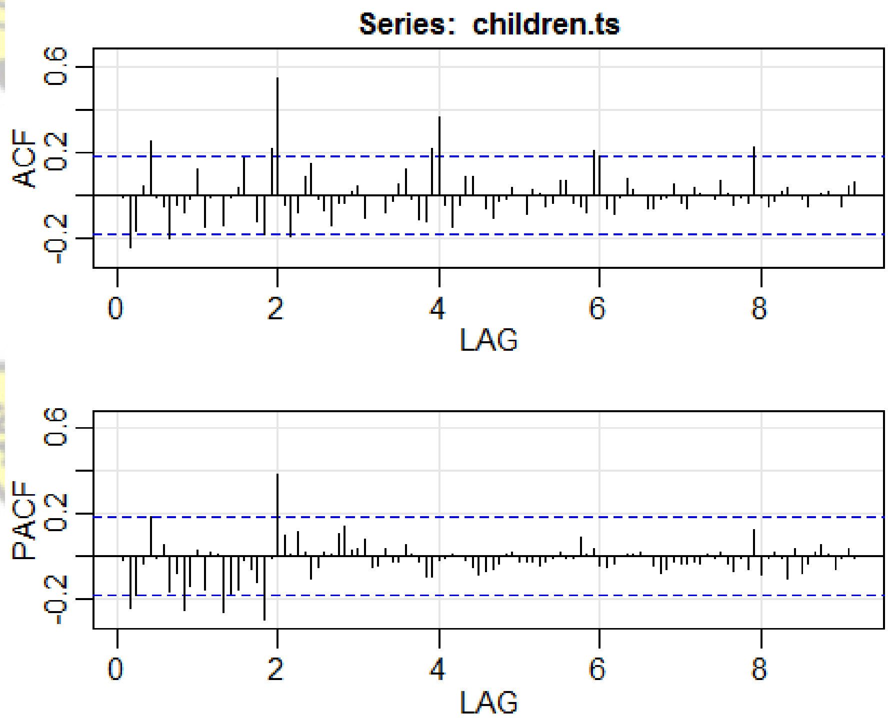

Figure 4.10: Plot of ACF and PACF of the children's ward series

Based on ACF and PACF plots in the figures above, candidate models were built and the model with the least AIC value were selected. Summary of vitals of the candidate models are given in the table below:

Table 4.3: Summary of candidate models for emergency ward series

Emergency Ward		Males' Ward		Females' Ward	
Model	AIC	Model	AIC	Model	AIC
SARIMA(9,0,9,1,0,1)$_{12}$	4.076385	SARIMA(6,0,6,1,0,2)$_{12}$	3.712863	SARIMA(5,0,3,2,0,2)$_{12}$	3.450852
SARIMA(9,0,9,1,0,0)$_{12}$	4.129498	SARIMA(6,0,6,1,0,0)$_{12}$	4.279028	SARIMA(5,0,3,2,0,0)$_{12}$	–
SARIMA(9,0,9,0,0,1)$_{12}$	4.643886	SARIMA(6,0,6,0,0,2)$_{12}$	3.892299	SARIMA(5,0,3,0,0,2)$_{12}$	–
SARIMA(0,0,9,1,0,1)$_{12}$	4.451865	SARIMA(6,0,0,1,0,2)$_{12}$	3.77958	SARIMA(5,0,0,2,0,2)$_{12}$	3.633836
SARIMA(0,0,9,1,0,0)$_{12}$	4.881737	SARIMA(6,0,0,0,0,2)$_{12}$	3.982997	SARIMA(5,0,0,2,0,0)$_{12}$	3.605603
SARIMA(0,0,9,0,0,1)$_{12}$	5.001114	SARIMA(6,0,0,1,0,0)$_{12}$	4.552253	SARIMA(5,0,0,0,0,2)$_{12}$	3.567504
SARIMA(9,0,0,1,0,1)$_{12}$	4.472599	SARIMA(0,0,6,1,0,2)$_{12}$	3.701097	SARIMA(0,0,3,2,0,2)$_{12}$	3.567596
SARIMA(9,0,0,1,0,0)$_{12}$	4.801854	SARIMA(0,0,6,1,0,0)$_{12}$	4.493007	SARIMA(0,0,3,0,0,2)$_{12}$	3.69837
SARIMA(9,0,9,0,0,1)$_{12}$	4.643886	SARIMA(0,0,6,0,0,2)$_{12}$	3.909558	SARIMA(0,0,3,2,0,0)$_{12}$	3.533272

Maternity Ward		Children's Ward	
Model	AIC	Model	AIC
SARIMA(2,0,1,2,0,2)$_{12}$	3.938806	SARIMA(2,0,1,3,0,2)$_{12}$	–
SARIMA(2,0,1,2,0,0)$_{12}$	3.916126	SARIMA(2,0,1,0,0,2)$_{12}$	3.406005
SARIMA(2,0,1,0,0,2)$_{12}$	4.170806	SARIMA(2,0,1,3,0,0)$_{12}$	–
SARIMA(0,0,1,2,0,2)$_{12}$	3.951152	SARIMA(0,0,1,3,0,2)$_{12}$	3.295412
SARIMA(0,0,1,0,0,2)$_{12}$	4.148671	SARIMA(0,0,1,0,0,2)$_{12}$	3.435317
SARIMA(0,0,1,2,0,0)$_{12}$	3.925141	SARIMA(0,0,1,3,0,0)$_{12}$	3.270434
SARIMA(2,0,0,2,0,2)$_{12}$	3.927601	SARIMA(2,0,0,3,0,2)$_{12}$	3.31333
SARIMA(2,0,0,0,0,2)$_{12}$	4.153786	SARIMA(2,0,0,0,0,2)$_{12}$	3.421727
SARIMA(2,0,0,2,0,0)$_{12}$	3.904835	ARIMA(2,0,0,3,0,0)$_{12}$	3.357231

From 4.3, it is observed that the model with the least AIC value and hence tend to adequately fit the series compared to the other candidate models is SARIMA (9,0,9,1,0,1)$_{12}$, SARIMA(0,0,6,1,0,2)$_{12}$, SARIMA(5,0,3,2,0,2)$_{12}$, SARIMA (2,0,0,2,0,0)$_{12}$, SARIMA(0,0,1,3,0,0)$_{12}$ for emergency ward, male's ward, female's ward, maternity ward and children's ward respectively.

4.4.2 Model Diagnostics

The selected models are adequate in fitting their respective series since the residuals of each model was observed to be uncorrelated, independent and normally distributed as shown by the ACF of the residuals, normal Q-Q plot and the p-values of the Ljung-Box static plot in figures below.

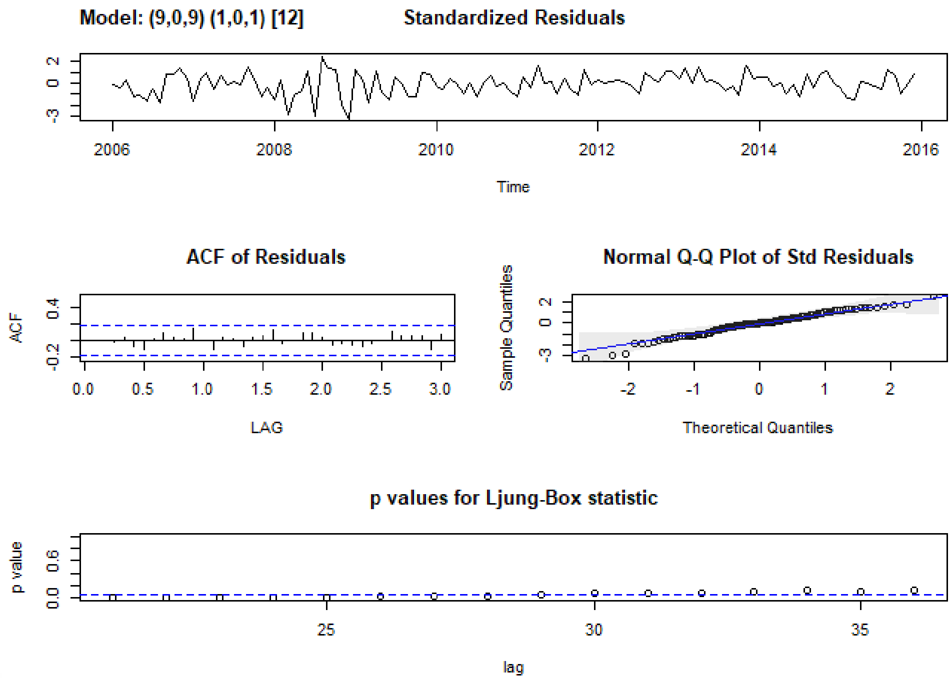

Figure 4.11: Emergency ward series best model diagnostics

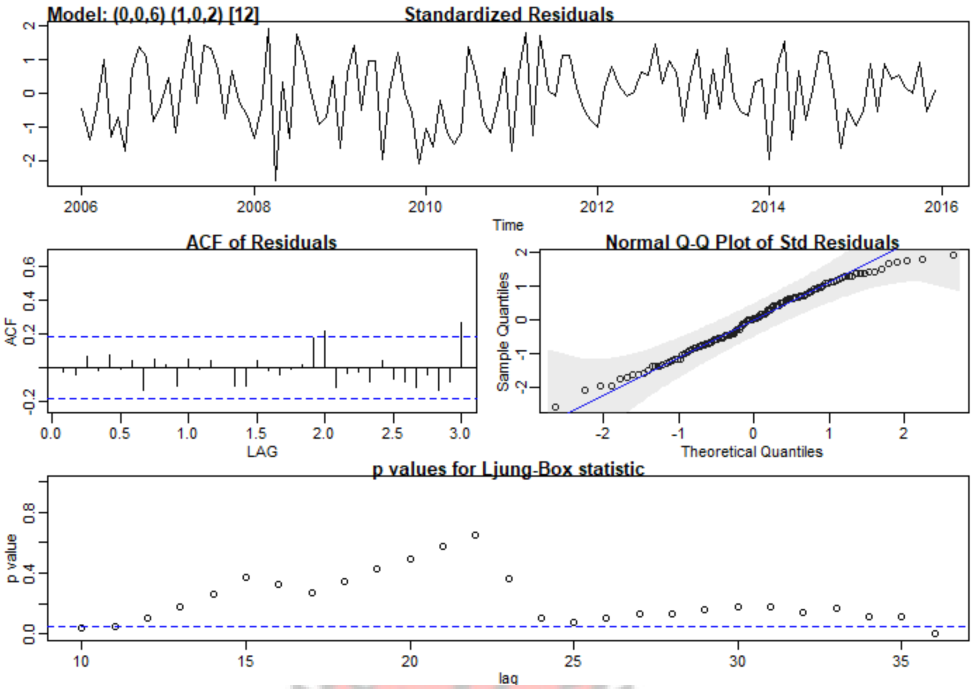

Figure 4.12: Males' ward series best model diagnostics

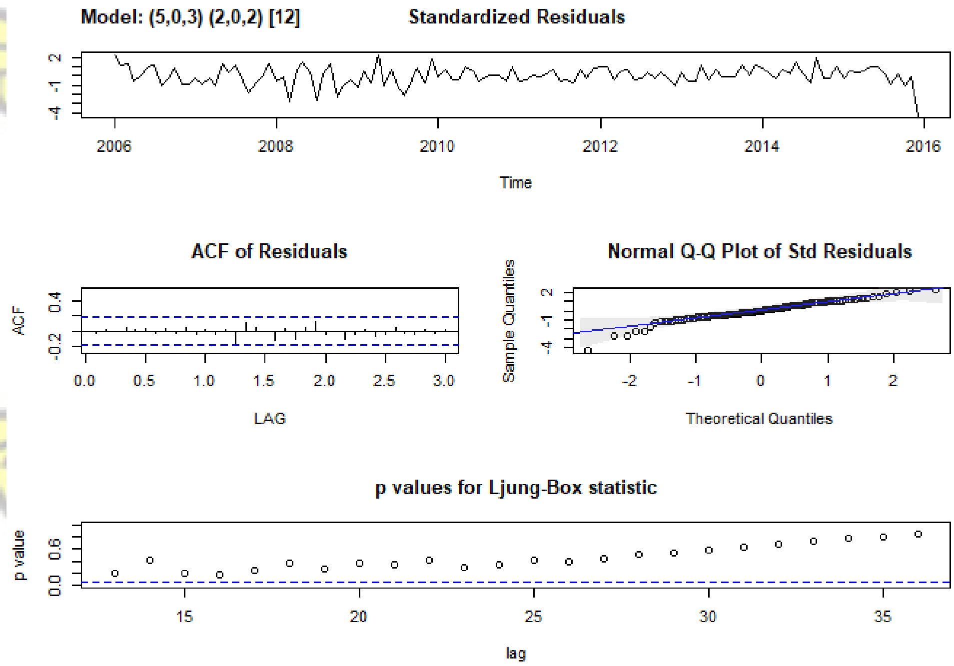

Figure 4.13: Females' ward series best model diagnostics

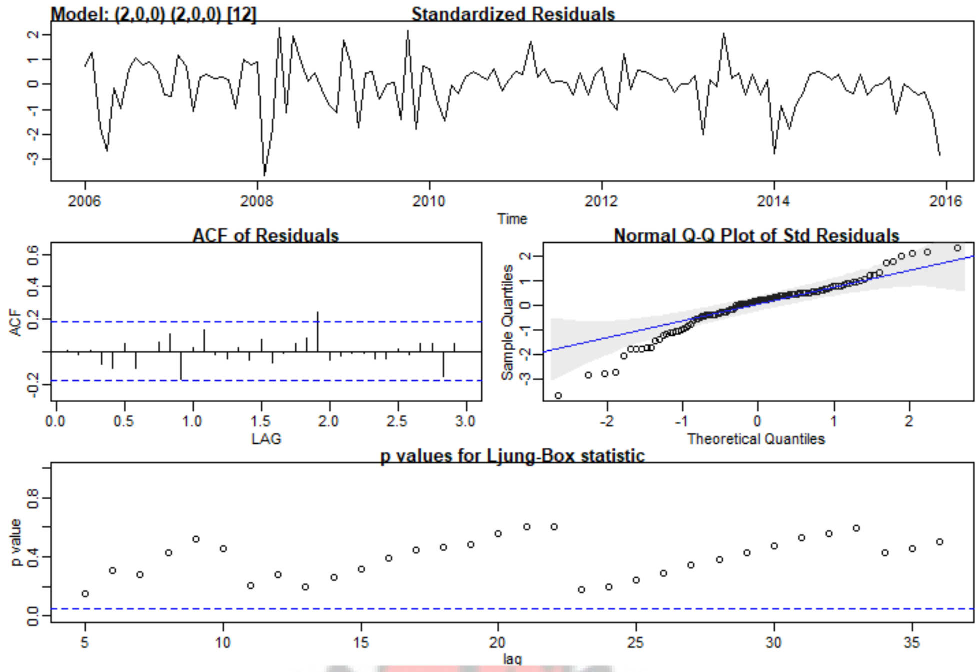

Figure 4.14: Maternity ward series best model diagnostics

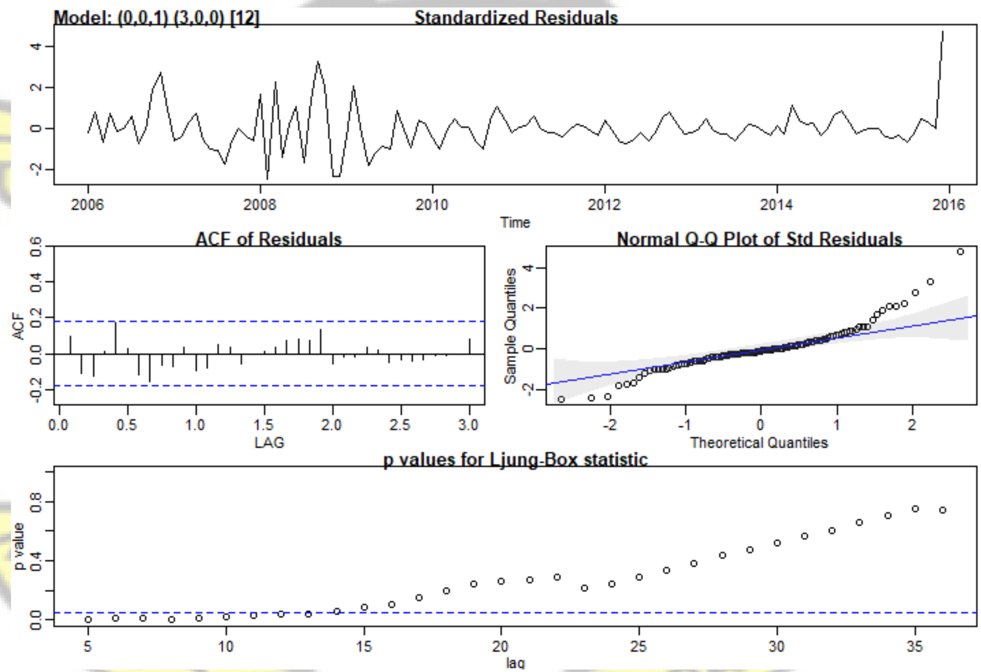

Figure 4.15: Children ward series best model diagnostics

4.4.3 Formulated Models

Emergency Ward Model

model summary

```
Coefficients:
```

ar1	ar2	ar3	ar4	ar5	ar6	ar7	ar8	ar9	ma1.
-0.64	0.28	0.66	0.61	0.14	0.029	-0.603	-0.56	-0.14	0.056

ma2	ma3	ma4	ma5	ma6	ma7	ma8	ma9	sma1	const.
-0.87	-0.69	-0.297	0.35	0.609	0.847	-0.128	-0.69	-0.143	24.911

model equation

$$\begin{aligned}(1 + 0.64B - 0.28B^2 - 0.66B^3 - 0.61B^4 - 0.14B^5 - 0.029B^6 + 0.603B^7 + 0.56B^8 + 0.14B^9)Z_t \\ = 24.911 + (1 + 0.056B - 0.87B^2 - 0.69B^3 - 0.297B^4 + 0.35B^5 + 0.609B^6 \\ +0.847B^7 - 0.128B^8 - 0.69B^9) \times (1 - 0.143B^{12})e_t\end{aligned} \tag{4.1}$$

Male Ward Model

model summary

Coefficients:

ma1	ma2	ma3	ma4	ma5	ma6	sar1	sma1
-0.265	0.144	-0.267	0.107	-0.548	0.436	-0.901	1.295

sma2	constant
0.817	23.079

model equation

$$\begin{aligned}(1 + 0.901B^{12})Z_t = 23.079 + (1 - 0.265B + 0.144B^2 - 0.267B^3 + 0.107B^4 - 0.548B^5 \\ + 0.436B^6)(1 + 1.295B^{12} + 0.817B^{24})e_t\end{aligned} \tag{4.2}$$

Female Ward Model

model summary

Coefficients:

ar1	ar2	ar3	ar4	ar5	ma1	ma2	ma3
-0.313	0.431	0.240	-0.006	0.302	0.526	-0.526	-1.00

sar1	sar2	sma1	sma2	const.
-0.069	0.696	-0.068	0.026	22.38

model equation

$$(1 + 0.313B - 0.431B^2 - 0.240B^3 + 0.006B^4 - 0.302B^5)(1 + 0.069B^{12} - 0.696B^{24})Z_t$$
$$= 22.38 + (1 + 0.526B - 0.526B^2 - 1.00B^3)(1 - 0.068B^{12} + 0.026B^{24})e_t \quad (4.3)$$

Maternity Ward Model

model summary

Coefficients:

ar1	ar2	sar1	sar2	constant
-0.0676	-0.2019	-0.1247	0.6582	26.533

model equation

$$(1 + 0.0676B + 0.2019B^2)(1 - 0.1247B^{12} + 0.6582B^{24})Z_t = 26.533 \quad (4.4)$$

Children Ward Model

model summary

Coefficients:

```
ma1      sar1       sar2       sar3     const.
-0.6397    0.2260    0.8226   -0.2856   16.4862
```

model equation

$$(1 - 0.2260B^{12} - 0.8226B^{24} + 0.2856B^{36})Z_t = 16.486 + (1 - 0.6397B)e_t \tag{4.5}$$

4.4.4 Cross Validation of forecasting models

To assess which model has the best predictive ability, the ARIMA and the Seasonal ARIMA models were used to forecast the last 12 month of the original data and the mean squared error (MSE) computed for each. The forecast and mean squared errors of the two approaches are given in the table below

Table 4.4: Cross validation of forecasting methods

Emergency Ward													
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	MSE
Actual	31	28	21	43	20	28	32	37	12	17	27	29	
SARIMA forecast	33.24	30.47	27.51	34.10	24.69	26.25	36.707	24.254	17.884	16.412	25.823	26.38	32.12983
ARIMA forecast	23.362	29.132	21.927	26.501	22.349	26.592	22.567	26.567	22.705	26.501	22.812	26.423	63.9232
Male Ward													
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	MSE
Actual	17	29	34	21	28	30	18	15	29	20	19	21	
SARIMA forecast	18.330	30.791	29.843	22.102	25.436	29.181	19.194	16.734	29.380	17.663	26.399	20.203	8.012088
ARIMA forecast	24.830	27.370	22.434	19.062	21.976	25.806	24.97	21.584	20.932	23.384	24.86	23.476	38.69303
Female Ward													
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	MSE
Actual	19	20	22	30	23	28	19	16	14	23	27	17	
SARIMA forecast	20.772	18.34	20.794	30.053	23.219	26.404	19.147	18.998	14.739	28.911	26.786	34.28	26.92726
ARIMA forecast	23.954	21.751	21.546	21.82	21.948	22.048	22.127	22.189	22.238	22.276	22.306	22.329	33.69503
Maternity Ward													
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	MSE
Actual	31	32	27	27	31	26	25	26	20	28	20	19	
SARIMA forecast	29.684	33.488	28.323	27.341	28.955	31.699	25.693	24.83	21.261	30.622	23.877	29.754	15.2898
ARIMA forecast	26.485	27.28	27.11	27.303	26.605	26.245	26.769	27.235	26.896	26.387	26.548	27.048	18.71558
Children Ward													
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	MSE
Actual	16	20	19	11	13	14	11	19	19	14	14	30	
SARIMA forecast	15.80	19.756	19.54	11.225	12.98	13.90	10.68	19.017	18.651	14.731	13.916	14.548	19.99574
ARIMA forecast	17.051	18.931	16.303	14.62	16.655	18.03	16.10	14.984	16.814	17.699	15.97	15.295	27.75494

Comparing the MSE of the two considered modeling methods, it observed that the Seasonal ARIMA (SARIMA) method gives a better forecast. Hence that method will be employed in forecasting the next 48 months; covering the years up to 2019.

4.4.5 Forecasting

Emergency Ward Series

Table 4.5: Emergency ward forecast values for for 2016–2019

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2016 forecast	20.981	27.96	18.165	18.054	16.574	27.125	11.932	31.471	24.178	37.220	20.898	18.376
2017 forecast	31.3	29.07	22.76	38.98	22.841	28.61	31.82	32.60	12.83	19.87	24.406	29.242
2018 forecast	22.15	26.323	19.68	17.012	19.06	26.34	14.019	29.162	24.87	36.83	19.84	20.601
2019 forecast	30.714	27.29	24.09	36.98	24.09	27.735	32.67	30.184	14.40	20.41	23.70	30.396

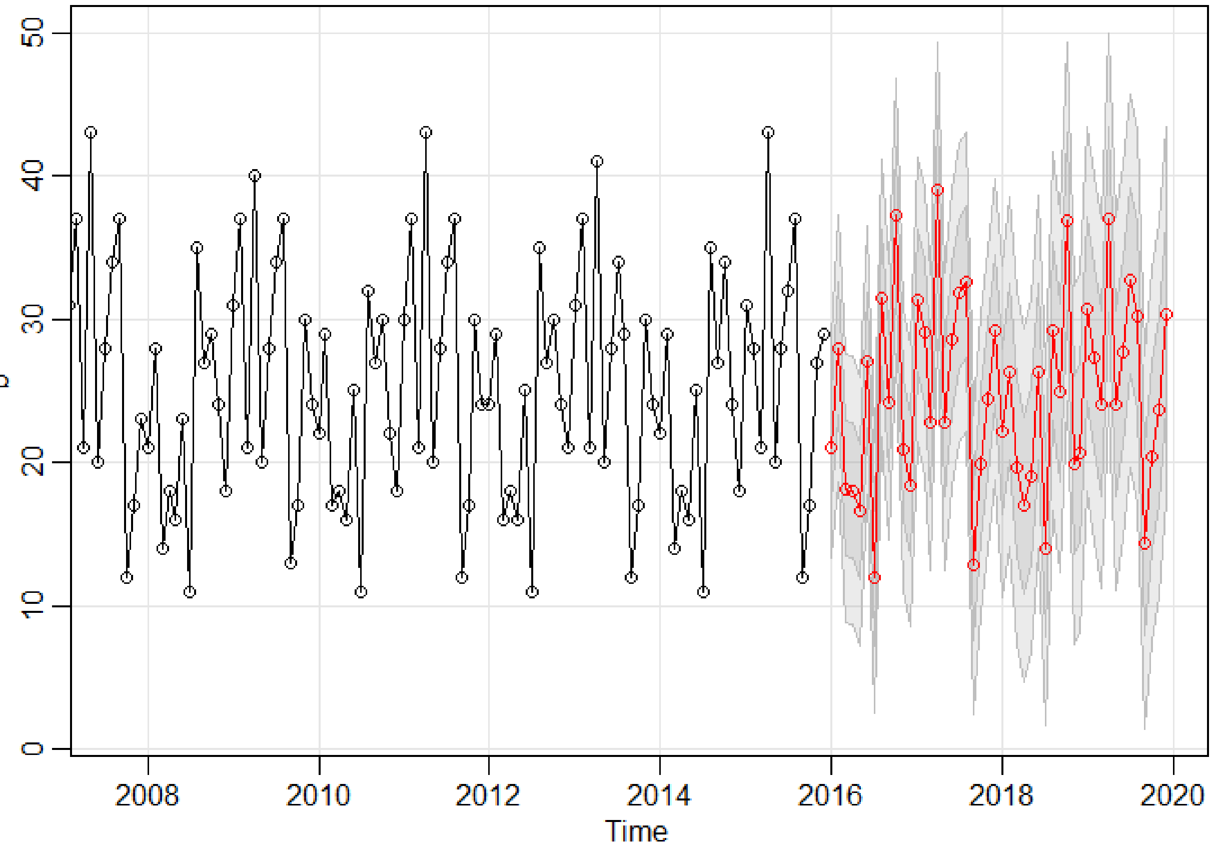

Figure 4.16: Forecast plot for emergency ward series

Male Ward Series

Table 4.6: Male ward forecast values for for 2016–2019

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2016 forecast	19.432	23.101	22.987	24.31	20.730	18.66	28.079	27.172	27.28	21.943	19.92	23.123
2017 forecast	23.011	24.74	24.96	26.341	25.31	27.983	19.763	16.48	20.77	24.39	24.125	23.014
2018 forecast	22.99	22.467	19.72	22.148	20.293	18.734	26.067	29.02	25.15	21.89	22.135	23.136
2019 forecast	23.15	23.63	26.104	23.91	25.589	26.99	20.38	17.72	21.20	24.14	23.929	23.026

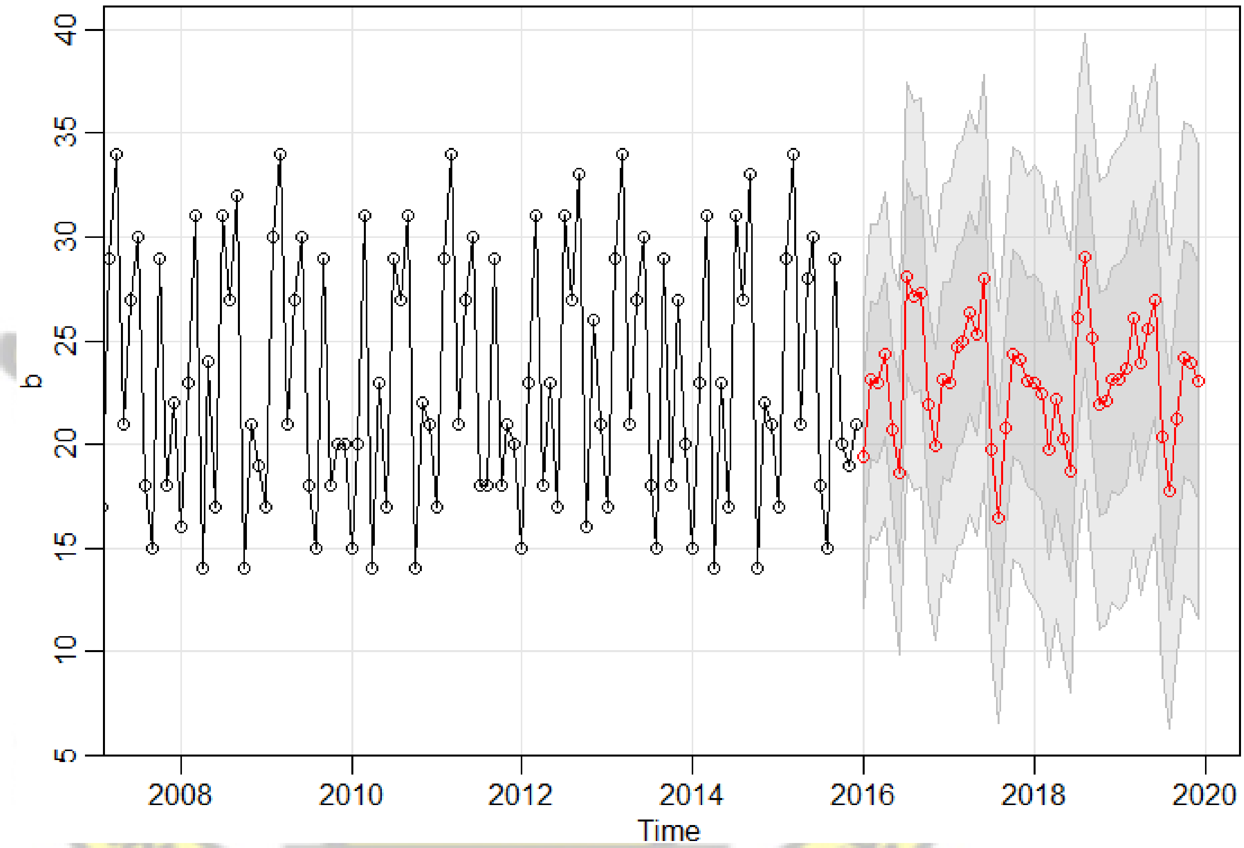

Figure 4.17: Forecast plot for male's ward series

Female Ward Series

Table 4.7: Female ward forecast values for for 2016–2019

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2016 forecast	27.00	27.158	24.05	18.165	25.058	28.719	19.561	20.71	25.49	24.427	18.761	17.90
2017 forecast	20.91	19.469	22.64	27.642	23.12	26.156	19.96	18.40	15.96	22.96	25.78	18.604
2018 forecast	25.86	25.77	23.992	18.90	24.312	26.58	20.53	21.70	24.8	23.919	19.58	19.55
2019 forecast	21.621	21.784	23.246	27.95	24.78	28.59	19.66	16.145	16.65	23.33	25.976	16.858

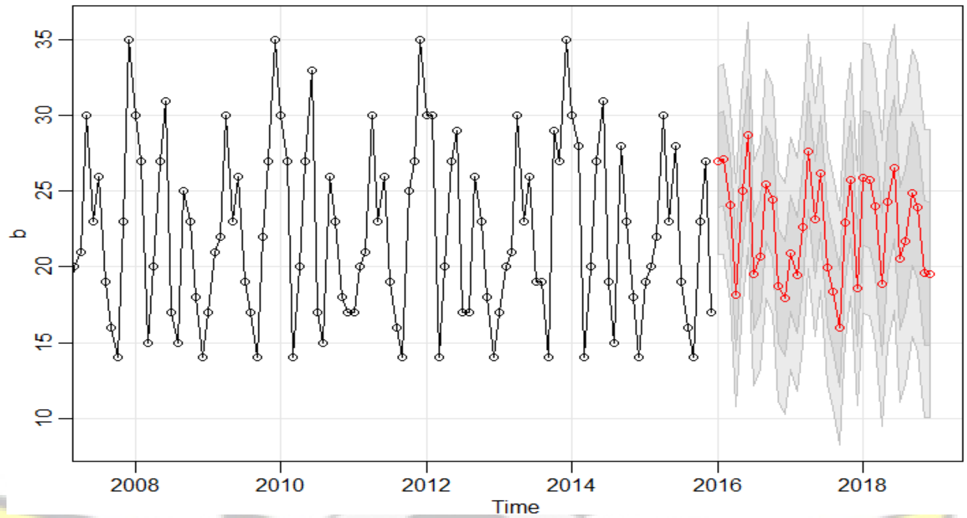

Figure 4.18: Forecast plot for females' ward series

Maternity Ward Series

Table 4.8: Maternity ward forecast values for for 2016–2019

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2016 forecast	23.334	20.37	15.77	27.70	24.431	28.95	30.95	28.869	29.636	27.97	25.679	24.48
2017 forecast	29.872	30.89	28.18	26.69	29.73	25.88	24.97	25.89	21.845	27.31	22.33	21.82
2018 forecast	24.01	21.93	19.24	27.28	24.75	28.20	29.63	28.150	29.16	27.38	26.49	25.774
2019 forecast	29.04	29.98	28.528	26.54	28.86	25.89	25.11	25.908	23.11	26.94	23.77	23.531

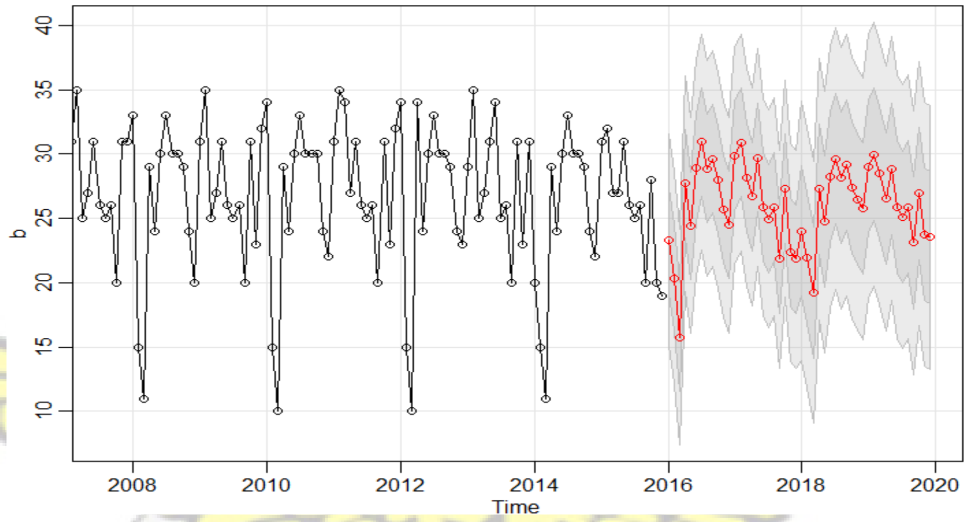

Figure 4.19: Forecast plot for maternity ward series

Children's Ward Series

Table 4.9: Forecast values for for 2016–2019

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2016 forecast	11.90	10.94	21.69	13.945	17.11	19.524	11.478	18.40	25.27	23.637	12.944	14.62
2017 forecast	13.47	19.97	17.87	12.39	13.61	14.12	12.694	18.269	17.53	13.625	14.922	29.035
2018 forecast	12.17	11.70	20.36	15.039	17.35	19.16	13.076	17.748	23.23	22.43	13.92	13.934
2019 forecast	14.34	19.86	17.013	13.519	14.13	14.27	14.02	17.69	16.36	13.43	15.63	26.76

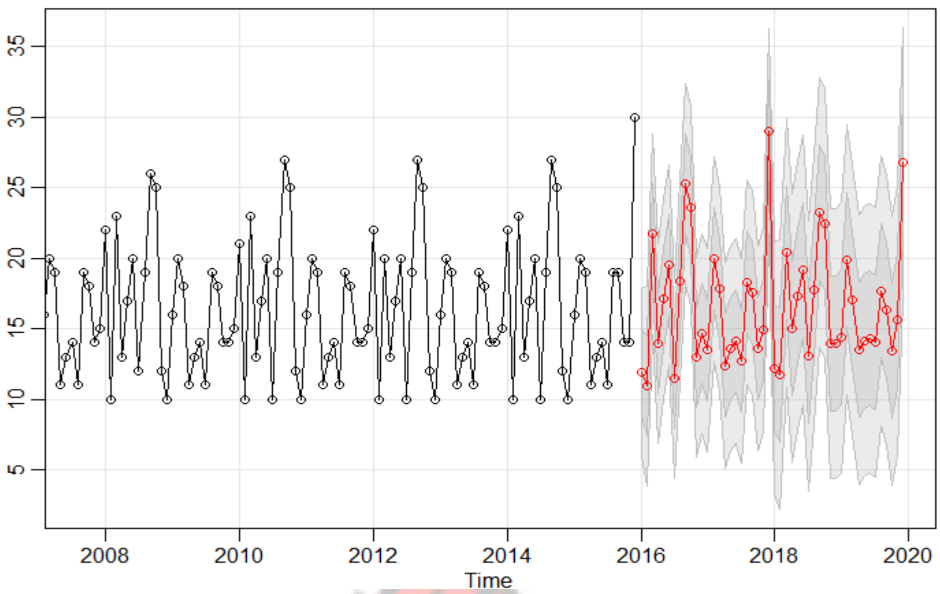

Figure 4.20: Forecast plot for children's ward series

4.5 Discussion of Results

4.5.1 Comparison to select the optimal model

Patients arrive in an unplanned manner beyond the control of the hospital, the number of hospital visits shows strong periodical, seasonal and stochastic fluctuations driven by factors such as climate, epidemic disease, the type of the day and socio demographic effects (Wargon et al., 2009). The admission data analyzed in this research work had similar characteristics hence the seasonal model, specifically the seasonal ARIMA models were the primary model used in fitting the admission data for the five departments of the hospital.

Lingling and Alper (2018) in their work concluded that the single seasonal ARIMA (SARIMA) model as a constituent of the hybrid SARIMA-NARNN was better in fitting and forecasting the monthly and daily number of newly admitted patients. This performance of the seasonal ARIMA model in terms of model fitting and forecasting ability was confirmed by the study carried out by Schweigler et al. (2009), when they compared three time series model in terms of forecasting emergency department bed occupancy. Goodness of fits were compared using log likelihood and Akaike's Information Criterion (AIC) and the root mean squared error was used to compare the model forecast. In their work, it was realized that the seasonal ARIMA model performed much better as compared to the other models. Similar observation were made in the current research work as the SARIMA models outperformed the non seasonal ARIMA model using the mean squared error as the criteria.

4.6 Interpretation of forecast of the optimal model

The observations made from the forecast of admission to the various department of the hospital are given below:

Emergency ward: In a year round of an even year, there is a steady fluctuation between January – March and a sharp decline around April. Moreover, there is a shoot in admission afterwards where the fluctuation becomes a bit steady between May and June. There is a decrease around July; from there, admission increases steadily till October, the admission then reduces afterwards till December. The odd years tend to have higher values, but follows same pattern as described for the even years.

Male: For the male series, the fluctuations tend to be minimal as well as the general admission to the male ward. It is observed that for each year, there is a steady increase in admission from January to September with significant drops occurring in April and August. Admission then rises a little around November and December.

Female: It is observed that admission in the even years tend to be very high in January and decreases steadily afterwards until it shoots up again during May. It then decreases once again till it rises again in September and fall afterwards towards December. In the odd years, there is increase from January to April. It then fluctuates evenly afterwards with significant shoot in June and November.

Maternity: Admission to this ward tend to be more higher than other wards in a year round. The forecast indicates that the admission to the ward going to be quite higher around January and February and very high between June and October. Admissions in other months are quite lower.

Children: The admission to the children's ward per the forecast, will fluctuate between 10–20 patients for most of the months while significant shoot ups occurs in the months of March and September.

Chapter 5

CONCLUSION AND RECOMMENDATION

5.1 Summary of Findings

The current study sort to investigate the admission of patients to the five department of the Trafalgar hospital by fitting time series models to the admission data to aid in forecasting the admission of patients for some future years. In so doing, it was observed that:

i. the admission series to the five departments of the hospital are stationary and required no transformation.

ii. ARIMA $(2,0,4)(1,0,2)_{12}$,ARIMA$(2,0,4)(1,1,0)_{12}$,ARIMA$(2,0,4)(1,1,0)_{12}$, ARIMA $(2,0,4)(2,0,0)_{12}$, ARIMA$(2,0,4)(1,1,0)_{12}$ were the best models for emergency ward, male's ward, female's ward, maternity ward and children's ward respectively.

iii. the various series show seasonality and hence the Seasonal ARIMA (SARIMA) models were better in fitting the admission data (having smaller mean squared errors of forecasting) than the non-seasonal ARIMA models.

iv. the admission to the maternity ward tend to be more higher than that of other wards.

5.2 Conclusion

Based on the findings made in this work, the following conclusions could be made. It is concluded that:

i. the admission to the emergency ward will fluctuate steadily between 25 to 33 from January to August. admission will decrease to about 17 patients in September and October and rises in November and December.

ii. the admission to the male ward will be 23 on the average for most months with significant peaks of 33, 27, 28 and 29 in March, May, June and September respectively for the 2019 period.

iii. The female ward admission for 2019 will increase (21–29) but with study fluctuation between January and June. Admission reduces afterwards (in the range of 16-19) and an increase to about 25 in October and November.

iv. the admission to the maternity ward tend to be very stable with little fluctuation. Admission fluctuates steadily between 24 to 29 patients for all months of the year.

v. the admission to the children's ward per the forecast, will fluctuate between 14–20 patients for most of the months while significant shoot up occurs in the month of December (28 patients).

5.3 Recommendation

It is recommended that the allocation of beds in the five department of the hospital be based on the conclusions drawn above.

REFERENCES

Abdel-Aal, R. E. and Mangoud, A. M. (1998). Modeling and forecasting monthly patient volume at a primary health care clinic using univariate time-series analysis. *Computer Methods and Programs in Biomedicine*, 56(3):235–247.

Arthur, J. (2013). Application of statistical models to outpatient department (opd) attendance data in saltpond municipal hospital of the central region of ghana.

Bahadori, M., Teymourzadeh, E., Ravangard, R., and Raadabadi, M. (2017). Factors affecting the overcrowding in outpatient healthcare. *Journal of education and health promotion*, 6.

Bao, Y., Fan, G., Zou, D., Wang, T., and Xue, D. (2017). Patient experience with outpatient encounters at public hospitals in shanghai: examining different aspects of physician services and implications of overcrowding. *PloS one*, 12(2).

Boutsioli, Z. (2011). Hospital costs and unexpected demand: the case of greece. *The Open Economics Journal*, 4:49–58.

Cheng, C. H., a. W. J. W. and Li, C. H. (2008). Forecasting the number of outpatient visits using a new fuzzy time series based on weighted-transitional matrix. *Expert Systems with Applications*, 34(4):2568–2575.

Devos, S., De Wit, L., Buyl, R., Hubloue, I., Lauwaert, D., Pien, K., and Putman, K. (2015). Factors influencing hospital admission and associated hospital costs of traffic victims admitted to an emergency department. *Journal of Transport & Health*, 2(3):406–413.

Farid, K., Harrou, F., Chaabane, S., and Tahon, C. (2014). Time series modelling and forecasting of emergency department overcrowding. *Journal of medical systems*, 38(9):107.

Farmer, R. D. and Emami, J. (1990). Models for forecasting hospital bed requirements in the acute sector. *Journal of Epidemiology & Community Health*, 44(4):307–312.

GHS (2011). Central region annual report. cape coast: Rhd. Technical report, Ghana Health Service.

Goka, F. (2007). Trend analysis of diseases reported at outpatient departments: A case study of the greater accra region. Master's thesis, University of Cape Coast, Cape Coast, Ghana.

Harrison, R. F. and Kennedy, R. L. (2005). Artificial neural network models for prediction of acute coronary syndromes using clinical data from the time of presentation. *Annals of emergency medicine*, 46(5):431–439.

Holger, K. and Schreiber, T. (2004). Nonlinear time series analysis. *Cambridge university press.*, 7.

Izabel, M., Hajat, S., and Gouveia, N. (2013). Forecasting daily emergency department visits using calendar variables and ambient temperature readings. *Academic emergency medicine*, 20(8):769–777.

Jenkins, C., Bartholomew, J., Gelder, F., and Morrell, D. (1994). Arranging hospital admission for acutely ill patients: problems encountered by general practitioners. *Br J Gen Pract,*, 44(383):251–254.

Kibaek, K., Lee, C., O'Leary, K., Rosenauer, S., and Mehrotra, S. (2014). Predicting patient volumes in hospital medicine: A comparative study of different time series forecasting methods. Technical report, Northwestern University, Illinois, USA, Scientific Report.

Lingling, G. and Alper, J. (2018). Models and strategies to integrate palliative care principles into care for people with serious illness.

Mai, Q., Aboagye-Sarfo, P., Sanfilippo, F. M., Preen, D. B., and Fatovich, D. M. (2015). Predicting the number of emergency department presentations in w estern a ustralia: A population-based time series analysis. *Emergency Medicine Australasia*, 27(1):16–21.

Nyako, L. A. (2002). An assessment of utilization of clinical services at the la polyclinic. Master's thesis, Kwame Nkrumah University of Science and Technology, Kumasi.

Reis, B. Y. and Mandl, K. D. (2003). Time series modeling for syndromic surveillance. *BMC Medical Informatics and Decision Making*, 3(1):2.

Schweigler, L. M., Desmond, J. S., McCarthy, M. L., Bukowski, K. J., Ionides, E. L., and Younger, J. G. (2009). Forecasting models of emergency department crowding. *Academic Emergency Medicine,*, 16(4):301–308.

Song, Y., Wang, Y., a. L. F., and Zhang, Y. (2017). Development of a hybrid model to predict construction and demolition waste: China as a case study. *Waste management*, 59:350–361.

Spencer, J. S. and Stat, M. (2007). Real-time demand forecasting in the emergency department. In *AMIA... Annual Symposium proceedings/AMIA Symposium. AMIA Symposium*, page 997.

Sun, Y., Heng, B. H., Seow, Y. T., and Seow, E. (2007). Forecasting daily attendances at an emergency department to aid resource planning. *BMC emergency medicine*, 9(1):1.

Tariq, I. (2016). Forecasting of monthly patient voliume at radiology department coming for ultrasound: a time serise approch.

Wendy J., D., Thomas, R. E., Lorenzetti, D. L., Spragins, and Williamson, T. (2011). Active and passive surveillance of yellow fever vaccine 17d or 17dd-associated serious adverse events: systematic review. *Vaccine*, 29(28):4544–4555.

Yarmohammadian, M. H., Rezaei, F., Haghshenas, A., and Tavakoli, N. (2017). Overcrowding in emergency departments: a review of strategies to decrease future challenges. *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences*, 22.

Zhou, L., Yu, L., Wang, Y., Lu, Z., Tian, L., Tan, L., and Liu, L. (2014). A hybrid model for predicting the prevalence of schistosomiasis in humans of qianjiang city, china. *PloS one*, 9(8).

Zhu, T., Luo, L., Zhang, X., Shi, Y., and Shen, W. (2017). Time-series approaches for forecasting the number of hospital daily discharged inpatients. *IEEE journal of biomedical and health informatics*, 21(2):515–526.