KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY, KUMASI

COLLEGE OF SCIENCE

DEPARTMENT OF MATHEMATICS

Time Series Analysis of Enrolment of Pupils in Public Second Cycle Schools in Assin South District

By

Ansah David (B. Sc. Mathematics)

A Thesis submitted to the Department of Mathematics in partial fulfillment of the requirements for the award of the degree of

MASTER OF SCIENCE

(INDUSTRIAL MATHEMATICS)

JUNE, 2014

DECLARATION

I, Ansah David hereby declare that this submission is my own work apart from the references of other people's work which has duly been acknowledged. I hereby declare that, this work has neither been presented in whole nor in part for any degree at this university or elsewhere.

Ansah David ……………………. …………………..

(Index Number: PG6316911) Signature Date

Certified by

Prof. S. K. Amponsah ……………………….. ………………….

(Supervisor) Signature Date

Certified by

Prof. S. K. Amponsah…………………… ……………………

(Head of Department) Signature Date

DEDICATION

This thesis is dedicated to my beloved wife Hawa Akyirefi and my son Fazil Ahmad Ansah.

ACKNOWLEDGEMENT

I express my sincere gratitude to my supervisor Prof. S. K. Amponsah for his contribution and guidance in making this work a success. Special thanks also go to Mr. Emmanuel Harris, a Mathematics lecturer in Kwame Nkrumah University of Science and Technology for his useful suggestions and assistance towards the completion of this work.

Mr. Desmond Koomson also needs to be acknowledged for his contribution in assisting me with the software codes to run the data. I am also grateful to Mrs. Paulina Nyankah, Headmistress of NyankumasiAhenkro Senior High School, Mr. Evans Glivey, Assistant Headmaster of NyankumasiAhenkro Senior High School, Madam Juliana Owusu, Headmistress of AssinNsuta Senior High School, Rev. Franklin K. Boadu, Headmaster of AssinManso Senior High School, Mr. Harrison Gardiner, Assistant headmaster of Assin Manso Senior High School and the headmaster of Adankwaman Senior High School for giving me the yearly enrolment data in their various schools.

Finally, my sincere gratitude goes to my wife, Hawa Akyirefi and my parents, Mr. S. K. Ansah and Madam Adjoa Amanyima for their care, love and help.

ABSTRACT

Education plays an important role in a developing country like Ghana. Education is needed to improve productivity as a result of harnessing human resources effectively. After the basic education, the pupils are enrolled in the second cycle schools through the Basic Education Certificate Examination (BECE). The study analyses the mean number of pupils enrolled in the public second cycle schools in Assin South District in the Central Region of Ghana (from 1991 to 2012). It also uses the Box-Jenkins modeling to fit a model to the enrolment of pupils in the four (4) public second cycle schools, which can be used to predict future enrolment in the district. The secondary data obtained from the four (4) public second cycle schools in the district was made up of twenty-two (22) consecutive years of number of pupils enrolled from 1991 to 2012. The Box-Jenkins ARIMA modeling was used to model the data. The appropriate ARIMA models; ARIMA (2, 2, 1), ARIMA (2, 2, 2) and ARIMA (2, 2, 3) were chosen based on the autocorrelation and partial autocorrelation functions. The estimation of the model parameters was carried out using the maximum likelihood estimation method with the R statistical software package. The ARIMA (2, 2, 1) model was selected as the best model since its AIC and BIC values were far lower than the other competing models. The T-test revealed that the MA part of the ARIMA (2, 2, 1) was statistically not significant and the final ARIMA model for forecasting was ARIMA (2,2,0). It was observed that enrolment of pupils in the public second cycle schools in the district had been increasing and it may also increase in 2013, 2014 and 2015.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

Background of the study

In a developing country like Ghana, education is needed to improve productivity as a result of harnessing human resources effectively. Education involves procedures and practices that lead to an improvement in the quality of individuals and improved social conditions as a whole. The aims of education in Ghana are to create a literate population, produce a physically healthy society and to improve the political awareness of the people. The three forms of education – formal education, informal education and non-formal education play an important role in achieving these aims of education (Omane-Akuamoah et al., 2004).

African Traditional education was the system of education for preparing Africans for meaningful life in their communities before the early Europeans introduced schools. The beginnings of school education in Ghana can be traced to the castles of the early European traders in the Gold Coast about five hundred years ago. From 1966 to 1969, the Kwapong Education Review Committee and the Mills–Odoi Commission recommended that the system of six years of primary education followed by four years of middle school course should be replaced by an integrated basic eight – year course for children between the ages of six and fourteen, and at the end of 8 – year course, pupils were to proceed secondary schools through a Common Entrance Examination to be conducted by the West African Examinations Council (WAEC)(Ampadu and Mohammed, 2006).

In March 1972, the Dzobo Committee on education was appointed to consider the proposals of the Ministry of Education. The committee came out with New Structure and Content of Education in Ghana with the aims of making education less bookish and blending academics with Vocational and Technical Skills; reducing the duration of pre-university education and making the education system relevant and adaptive to the aspirations, occupations, institutions and traditions of people. The recommendations of the educational committee could not be implemented on a national scale but some Experimental Junior Secondary Schools were designated to start the new educational programme on pilot basis. The delay continued until 1987 when the government implemented the programme but with some minor modifications. The New Educational Reforms ran from 1987 until the Free Compulsory and Universal Basic Education (FCUBE) intervention was introduced in 1996 due to certain constitutional demands (Ampadu and Mohammed, 2006).

Under the New Education Reform, the primary school course still takes six years followed by three years Junior Secondary School and three years Senior Secondary School until 2007,where the content and structure of the New Education Reform was reviewed to start from 2007 to 2015. In that review, Kindergarten education constitutes four to six years to provide the children with pre-school learning skill; six years primary school education in basic literacy, numeracy, Science and Social Studies; 3-year Junior High School education (General Curriculum for entry into Grammar, Technical, Vocational and Agricultural course). The secondary education constitutes 4-year Senior

High School education in Grammar, Technical, Vocational or Agricultural Studies after which the successful students may enter Colleges of Education, Professional Institutes, Polytechnics and Universities. In 2009, the four years duration in the Senior High School was brought back to three years (Ampadu and Mohammed, 2006).

Before the introduction of Junior Secondary School under the New Educational Reform in 1987, most students were admitted in the Senior Secondary Schools through Common Entrance Examination conducted by the West African Examinations Council (WAEC) and Middle School Leaving Certificate (MLC). Since 1990, pupils after completing Junior Secondary Schools were being admitted in the Senior Secondary Schools with the Basic Education Certificate Examination (BECE) conducted by the West African Examinations Council (WAEC). Before 2005, Heads of Public Senior High Schools were invited to decide on qualified candidates who made good grades to be enrolled in the various senior secondary schools. Since 2005, the Computerized School Selection and Placement System (CSSPS) has been operational and has gone through many phases of review taking into consideration the concerns of the public. It has become necessary to subject the enrolment of pupils in the second cycle schools in Assin South District in the central region of Ghana to a statistical study-time series analysis to examine the outcome (Ministry of Education and Ghana Education Service, 2011).

1.1 Statement of the problem

After independence, Ghana's basic and secondary school system of education created a large population of unemployed school leavers who were socially and politically becoming a danger to the society. In those days, pupils were admitted into the senior secondary schools through Common Entrance Examination or Middle School Leaving Certificate. Before 2005, heads of Public Senior Secondary Schools were invited to select qualified candidates to be enrolled in the various secondary schools in the country. The Computerized School Selection and Placement (CSSPS) had been operational since 2005 to cater for the enrolment of pupils in the public senior high schools in Ghana. After the exercise, the list of the successful candidates is sent to the respective senior high school through the district education office. The Assin South District has four senior high schools to absorb products from the surrounding junior high schools and other junior high schools in Ghana.

Since 1991, after the introduction of the new secondary school system under the 1987 New Educational Reform, the stakeholders in education in Assin South District in the Central Region of Ghana, find it difficult to predict future enrolment of pupils in the four(4) public second cycle schools in the district in order to provide infrastructure. There is no existing statistical model that can be used to predict future enrolment in the public second cycle schools in the district.Concerning the enrollment of pupils in these four senior high schools in the Assin South District, there is the need for a statistical study – time series analysis to asses and predict the enrollment in the public second cycle schools in the district.

1.2 Objectives of the study

The objectives of the study are to:

(i) estimate the mean number of pupils enrolled in the public second cycle schools in the Assin South District.

(ii) use the Box-Jenkins modeling to fit a model to the enrolment of pupils in the public second cycle schools which can be used to predict future enrolment in the Assin South District.

(iii) interpret the results of the fitted model based on the findings from the study.

1.3 Research methodology

The data for the study was a secondary data obtained from public senior high schools in the Assin South District. The data is made up of twenty-two (22) consecutive years of number of pupils enrolled in the public senior high schools in the district from 1991 to 2012.The Box-Jenkins ARIMA modeling procedure was used to model the data. The appropriate ARIMA models were chosen based on the autocorrelation and partial autocorrelation functions. The estimation of the model parameters was carried out using the maximum likelihood estimation method with the R statistical software package. The two penalty function statistics – Akaike Information Criterion (AIC) and the Schwarz Bayesian Information Criteria (BIC) were used to select the best model fitted based on the principle of parsimony.

Statistically significant parameters were maintained in the model based on the test statistics and the bounds of stationarity and invertibility. After the parameter estimation of the fitted model, diagnostic checking of the Box-Jenkins ARIMA process was carried out to examine whether the fitted model followed a white noise process by studying the autocorrelation values at a time and developing a standard error formula to test whether a particular autocorrelation value is significantly different from zero.The modified Ljung-Box Q statistic was used to test the adequacy of the model selected. Another check was done to find out whether the residuals of the model were normally distributed with constant variance and zero mean by using the residual plot against time and the Shapiro-Wilk Normality test.

1.4 Significance of the study

In Ghana as a developing country, people need education to improve their scientific understanding. Education prepares individuals to play their roles effectively in the society. People must be educated for such development in the various sectors. For the people to be educated, they must pass through the education system in Ghana – basic education, secondary education and tertiary. There is therefore the need to statistically assess the enrollment of pupils in the second cycle schools in the Assin South District so that stakeholders can use the results to implement policies.

The study can provide a model which will serve as a guide for governments to be used in the educational reforms. The model can also be used to forecast into the future the enrollment of pupils in the public second cycle schools in the Assin South District in order to provide infrastructure. More importantly, the model can serve as a document that would assess the reforms in education in Assin South District.

1.5 Scope of the study

This study was structured using the enrolment of pupils in the public second cycle schools in Assin South District in the Central Region of Ghana. The study focused on the analysis, the mean number of pupils enrolled from 1991 to 2012 and modelling an ARIMA (p, d, q) using Box-Jenkins's strategy. The data was obtained from the four (4) public second cycle schools in the district.

1.6 Limitations of the study

There were limitations in the areas of limited information and limited time period for this research. The research was conducted using secondary data, which may have some errors in its collection. Also, due to data collection problem in the nation, this study was carried out with a small sample size (a period of twenty-two years). However, these limitations do not render the findings from this research non-reliable because the researcher carefully handled these limitations with care to make sure the study objectives were achieved.

1.7 Organization of the study

The study has been divided into five chapters. The first chapter consists of introduction which includes background of the study, statement of the problem, objectives of the study, research methodology, significance of the study and the organization of the study. Chapter two constitutes the literature review on the meaning, aims and importance of education in the development of the nation. It also deals with the history of education during the colonial era and after independence. Also, the Education Acts and the various educational reforms are properly treated.

The third chapter gives the explanation of time series analysis and its illustration followed by chapter four in which the actual data analysis is dealt with. Chapter five constitutes the summary and findings from the study through which the conclusion is drawn. This is followed by the recommendations, references and appendices.

CHAPTER TWO

LITERATURE REVIEW

Meaning of education

Education developed from the Latin words "educare" which means "nurture" or "rear" and "duco" meaning "to grow". Education can be defined as the process and experiences for preparing individuals for meaningful living today and the future in their immediate environments, the community and the world at large. Education involves procedures and practices that lead to an improvement in the quality of individuals, enhance performance and improve social conditions as a whole. Education can be considered as a process. In this case education is used to refer to the experiences we encounter day by day to change our lives (Omane-Akuamoah et al., 2004).

Another context in which education is used is a part of culture. We have ways of celebrating events such as festivals, naming ceremonies and other ways of doing things in our societies. In Ghana, before the advent of school education, we had a way of preparing our young ones to fit life in our communities. Education is also seen as a subject like English, Mathematics, History, Geography and others that are studied in schools (Omane-Akuamoah et al., 2004).

2.1 The aims of education in Ghana

One of the aims of education in Ghana is to create a literate population. If many Ghanaians are able to read, write and compute, many of the problems we have now might not exist. We will be more scientific in our thinking and therefore superstitious thinking will reduce. Another aim of education in Ghana is the production of a physically healthy Ghanaian society. Some of the activities aimed at producing a physically healthy society are physical exercise and cleanliness.

Education also improves the political awareness of the people. The people must know the rights and responsibilities of the citizen for instance the importance of voting, obedience to state authority, patriotism and loyalty to Ghana (Omane-Akuamoah et al., 2004).

The development of vocational and technical skills is another aim of education in Ghana. Such skills would enable the individual to earn a living and also satisfy the manpower requirements of the country. Also, the development of morality is another aim of education in Ghana, in which through that citizens of the country would put up good behavior. The development of critical and logical thinking in the solution of everyday problems is an aim of education in Ghana. The Ghanaian society is faced with numerous problems such as financial, social, political, cultural and other aspects which education seeks to help solve through the development of critical thinking among the citizens. Through education one can earn a living. The individual acquires employable skills and with that the individual is gainfully employed and thereby earns a living. Through education, individuals become literate. Every individual wants to be able to read and write. It is education that equips the individual with such skills as reading and writing through the study of subject like English (Omane-Akuamoah et al., 2004).

Through education individuals occupy prestigious positions in the community. One would like to attain certain positions in the society that are respected. The individual will study subjects that will enable him or her ascend to such positions. Individuals are educated in order to be independent. Education equips the individual with intellectual and vocational skills that enable him or her to be gainfully employed and as such live an independent life.Education helps individuals to understand the environment better. An individual would like to understand the environment and use it to his or her advantage. The school exposes the individual to environmental issues through the study of subjects like Environmental Studies, Science, Social Studies and many more (Omane-Akuamoah et al., 2004).

2.2 The importance of education

Education has so many benefits to the members in a country. Education improves scientific understanding and reduces superstitious practices and beliefs. Education is able to do this because it helps to develop critical and analytic minds. It also enables people to find out the truth and to discard or innovate certain unproductive superstitious beliefs and practices such as "trokosi" and widowhood rites.Education also serves the purpose of socialization. One individual or generation does not have to learn everything afresh. We do not have to go through the same pains of the past. Education makes the young ones aware of what should or should not be done. It makes people part of their society. Education ensures the survival of society. Education trains people to be innovative and good leaders. Education nurtures the potentials and talents of individuals. It is the purpose of education to develop talents and strengths in the individual, and at the same time minimizing weaknesses (Omane-Akuamoah et al., 2004).

Education improves the living conditions of people in the society. The society has to do be well organized and have good and innovative leadership. Thus, individual members of society die off, yet society continues to exist because one generation prepares the next before it passes away.Education identifies social and emotional problems and initiates collective solutions. Our communities are engulfed with problems such as high birth rates, disturbing rate of HIV/AIDS infections as well as sanitation problems. Education attempts to create awareness of the problems associated with these problems through some subjects like Social and Environmental Studies. In the school, certain activities are planned to create awareness on sanitation issues (Omane-Akuamoah et al., 2004).

2.3 Functions of education

Educational sociologists have grouped the functions of education into social, cultural, economic, selection and political. The social function of education has to do with how education is able to prepare individuals to fit well into their social environment. The social function is based on a concept called socialization. Culture can be said to be the whole way of life of a people. The cultural function of education concerns with the issues of transmitting culture, preserving it the way it is and innovating of changing it (Omane-Akuamoah et al., 2004).

The economic function of education refers to the role that education plays to prepare and equip its products with the needed occupational and vocational skills and attitudes that satisfy the manpower needs of the society. The selection function of education refers to the role education plays in selecting, directing and allocating individuals to different areas and occupations. The political function of education refers to the process by which education socializes individuals politically by instilling leadership qualities, democratic principles and good citizenship in them (Omane-Akuamoah et al., 2004).

2.4 Challenges to educational effort in Ghana

It is the responsibility of governments to provide for the education of her citizens. Over the years, the government of Ghana has tried to perform this role which is the constitutional right of all Ghanaian children. The government has implemented reforms and intervention programmes aimed at improving the quality of education and increasing accessibility to basic and secondary education. In spite of all these efforts by government to improve the provision of education to Ghanaian children, there have been a lot of challenges (Omane-Akuamoah et al., 2004).

One of the challenges in education sector is financial constraints. Many schools, especially those in the rural areas still lack the basic required educational inputs like satisfactory buildings, books equipment and other relevant materials. There are still in existence a large number of untrained teachers in the system. Insufficient number of teachers has led to the recruitment of untrained or pupil teachers who may not have any professional training in the act and methods of teaching. This does not auger well for effective implementation of educational reforms at the basic level of education (Omane-Akuamoah et al., 2004).

The cost of education at the senior secondary and university levels are high and tends to discourage many economically disadvantaged children from benefiting from secondary and tertiary education. The number of senior secondary schools we have cannot absorb all senior high school graduates, and the limited number of our universities make it impossible to absorb all senior high school graduates. As a result of this, many qualified students are unable to gain admission into senior secondary schools and universities. Some of these unlucky students roam the streets without any meaningful employment (Omane-Akuamoah et al., 2004).

2.5 Forms of education

Based on structure and characteristics, education has been grouped into three forms. These are formal education, informal education and non-formal education.

2.6 Formal education

Formal education is the organized and structured learning or training that takes place in the school. The programmes and activities in formal education are carried out in formally established institutions like primary school, secondary school (junior and senior secondary schools), university and any other tertiary schools. In formal education, activities are properly organized according to the syllabus, time table and specified curriculum (Omane-Akuamoah et al., 2004).

The programme of activities is carefully and systematically done according to laid down rules. Formal education has a supervised teaching component. Activities that go on within the school are closely monitored. Formal education employs the scientific approach in teaching. Several methods are available for teaching learners. Formal education has two main types. They are general education and vocational or technical education. General education is also called liberal or grammar education. It is given to broaden the mental ability of the learners. It includes subjects like Humanities, Biological and Social Sciences and Mathematics. Vocational education is given to the individual to acquire a specific job. It is normally geared towards self-employment. Some of the subjects under this education are carpentry, electronics, catering, masonry, engineering, music, law and medicine (Omane-Akuamoah et al., 2004).

2.6.1 Informal education

Informal education is used to describe the type of education where learning or training takes place casually, at all times, throughout the person's life. Most of the times, informal learning takes place unconsciously or unintentionally as people go through all kinds of experiences in the family, with friends and at other social events.The informal education is completely unstructured. The individual is free to learn according to his or her own interest. Informal education is organized by non-institutional bodies. It does not require a campus or classroom. Learning takes place at any place. It is practically oriented, through observation and participation. Learners acquire information by watching and taking part in the activity. Informal education is also characterized by non-supervised teaching. The contents of informal education include character or moral training, vocational training, intellectual training, physical training and social training (Omane-Akuamoah et al., 2004).

2.6.2 Non-formal education

Non-formal education is an organized learning activity carried on outside the structure of the formal education system which consciously aimed at meeting specific learning needs of particular groups of children, youths or adult in the society. Non-formal education is less structured, with activities organized at any place. It does not follow rigid time tables, syllabus and methods of teaching. It often involves planned meetings and also, some evaluation is often done to ensure that learners have acquired the skills expected (Omane-Akuamoah et al., 2004).

As a result, it is not totally informal. The non-formal education has an open membership, and open to all members who need such information. For instance, information on agricultural practices is open to farmers, fishermen and herdsmen.The non-formal education is also organized to aim at the needs of particular groups of people. These may be farmers, traders, community leaders or traditional birth attendants. The contents of non-formal education include General or Basic Education, Family Improvement Education, Community Improvement Education and Occupational Education (Omane-Akuamoah et al., 2004).

2.7 African traditional education

African traditional education refers to the non-specialized, informal and non-literate form of education given in African societies intended to produce useful persons for self-sufficient communities in Africa. In reality, African traditional education was the system of education for preparing Africans for meaningful life in their communities before the early Europeans introduced formal education. With the introduction of formal education, African traditional education continues to exist alongside school education. Today, African traditional education may be explained as all the skills, knowledge, experiences and the training that individuals receive outside the classroom. In general, African traditional education had and still has the sole purpose of inducting individuals into the society with all its institutions, taboos, cultural norms, values and functions with the view of making the individuals a part of the totality of the social consciousness. It is indigenous in nature; the educational goals, methods and content information to be learnt differed from one local area to another; depending on the needs of that particular society and the demands of its environment (Ampadu and Mohammed, 2006).

African traditional education was incidental in nature; this means that it was not planned to follow a particular structure but any incident that happened presented a learning instance for the individual. African traditional education embraces all aspects of the personality of the individual – physical, economic, social, political and religious. It is also given through observation, imitation and participation alongside the elders in productive activities. The home is the school in African traditional education and the teachers are the parents and the elders in the family and the community. In African traditional education it was the duty of all elders in the family to train the young members so that they might bring honour and not disgrace to the whole culture of the people formed the curriculum of the African traditional education (Ampadu and Mohammed, 2006).

2.8 Castle school education

About five hundred years ago, European sailors started visiting the coastal towns of our country and established trade links with the local communities. The first of these early Europeans to set foot on our land were Portuguese from Portugal and they established a trade link with the people of Edina near Cape Coast. The Europeans found so much gold that they decided to set up a permanent trading. In order to protect their goods, the early settlers put up very big and strong buildings called "castles" in which they dwelt and traded with the local people. The first of these great buildings is the Elmina Castle which was completed in 1882 under the leadership of the Portuguese sailor and explorer – Don Diego de – Azambuja (Ampadu and Mohammed, 2006).

There were forts and castles at places like Axim, Dixcove, Cape Coast, Anomabo, Abandze and Accra. It was in these castles that the first schools in our country were established. The first Castle schools were the types of schools established in the castles along the coast of Ghana, which provided the western European concept of education. The education was provided in designated institutions called schools, and the education followed a specifically-programmed set of activities called a curriculum. The teachers played specialized and professional roles. The schools had specific sites where teaching and learning took place. The castle school education focused on reading, writing, arithmetic and later, religious instruction (Ampadu et al., 2006).

The first castle school was established in the Elmina Castle in 1529. The curriculum was the same in all the castle schools and the only difference was the languages. All the castles and forts that were built in the coastal towns of the Gold Coast (Ghana) had schools in them but notable among them were the Elmina Castle School, the Cape Coast Castle School and the Christiansburg Castle School. The Castle Schools in the Gold Coast were Elmina Castle School, Cape Coast Castle School and the Christiansburg Castle School at Osu in Accra (Ampadu and Mohammed, 2006).

2.9 Early Christian Missionary's contributions in education in Gold Coast

The missionary educational enterprise in the Gold Coast was a follow-up of the Castle School educational activities. This has been broken into: The Society for the Propagation of the Gospel (SPG), The Basel Missionary Society, The Wesleyan Missionary Society, The Bremen Missionary Society and the Roman Catholic Mission (Ampadu and Mohammed, 2006).

2.9.1 The Society for the Propagation of the Gospel (SPG)

The SPG was one of the earliest missions and in fact the first to have started Christian missionary works in Ghana. The SPG missionary and educational works started with the arrival of Rev. Thomas Thompson at Cape Coast in 1752. Another SPG minister who continued the works of Rev. Thomas Thompson was Rev. Philip Quaque. He worked mainly in Cape Coast where he later evolved into the Anglican Church and its areas of operation were Accra and parts of Eastern and Western regions (Ampadu and Mohammed, 2006).

2.9.2 The Basel Mission

The Basel Missionaries were invited by a Danish governor of the Christiansburg Castle. The Basel Missionary Society came to the Gold Coast at the instance of one Governor Major de-Richelieu of the Danish Castle at Christiansburg in Accra to start educational work. In 1828, they opened a school at Christiansburg and started educational activities. The Basel Mission included in their curriculum an intensive programme in agricultural and manual instruction. They opened a boys' school at Akropong in 1843 followed by a girls' school in 1847 and a teacher training college (Presbyterian Training College) at Akropong (Ampadu and Mohammed, 2006).

.

2.9.3 Wesleyan Missionary Society

The Wesleyan Missionaries from England came to the Gold Coast at the instance of a product of the Cape Coast Castle School, one William de-Graft who had settled at Dixcove as a trader. Rev. Thomas Birch Freeman spread Wesleyan educational activities beyond the coast to Ashanti region and Nigeria. The Wesleyan Mission (which later evolved into the Methodist church) became well established in Cape Coast, Anomabo and Saltpond. Rev. Thomas Birch Freeman (1809 – 1890) worked mainly in Cape Coast and later extended the Wesleyan educational and missionary activities to Mankessim, Dixcove, Accra and Kumasi.The Wesleyan mission established a Theological institute in Accra in 1842 to train teachers, but the venture was abandoned as a result of the death of the first Principal of the Institute – Rev. Samuel Shipman. Efforts were made by the early Christian missions to encourage girls to go to school. They placed emphasis on the 3Rs – reading, writing and arithmetic, promoted reading, writing and numeracy (Ampadu and Mohammed, 2006).

2.9.4 The Bremen Missionary Society

The Bremen Missionary Society came from the city of Bremen in northern Germany. The Bremen Missionary Society started missionary work in the Gold Coast in 1847 when they sent a group of missionaries led by Rev. Lorenz Wolf to start work at Peki in the present day Volta Region. The Bremen Missionary educational curriculum emphasized technical and vocational education as well as the development of the local Ewe language (Ampadu and Mohammed, 2006).

2.9.5 The Roman Catholic Mission

In 1880, a major attempt was made to revive the Catholic faith in Ghana (Gold Coast) when two French missionaries – Father Auguste Moreau and Eugene Murat arrived at Elmina and succeeded in laying a solid foundation for Catholic missionary and educational activities started at Elmina and later spread to Amissano, near Cape Coast and then Keta.The second contribution was made by Sisters of Our Lady of Apostles (OLA) and Father of the Society of African Missions (SMA) who were very instrumental in the missionary works in the Volta Region. The third dimension was the contributions of another Catholic Missionary group called “the White Fathers” (because of the long, white dresses they wore). They entered the country from Algeria through the Sahara desert and were responsible for the introduction of the Catholic faith and education in northern Ghana, especially in the upper regions (Ampadu and Mohammed, 2006).

2.10 Islamic Missionary educational activities

Islamic religion was introduced into West Africa in the 11th century AD and by the 14th century, Islam had reached modern day Ghana. The Mande Dyula traders, who came into contact with Islam, spread the religion and associated educational practices to northern Ghana down to Ashante region. The Mande traders came as far as Salaga, Bono, Techiman and eventually Kumasi to trade and brought the religion with them. Ghanaians living in those areas at that time accepted the Islamic. Islamic education is slightly different from African traditional education and the formal education that was introduced by the Christian missionaries. In Ghana, two major Islamic groups have been in existence since 1922 and they are Suni Muslims (the traditional orthodox sect) which originated from Arabia and the Ahmadiyya Movement (originated from Parkistan) (Ampadu and Mohammed, 2006).

2.10.1 Orthodox Islamic education

The orthodox Islamic education has three main levels: elementary, secondary and tertiary. Islamic education was open to all but most especially to Muslim males. It was not based on the ability to pay fees because every Muslim scholar has the pious duty to pass on what he has learnt from the Quran and other Islamic texts to the younger generation. The course content was the Tafsir and the Hadith. The instructional approaches were rote learning which involved mass recitation of aspects of the Quran, explanation of aspects of the Quran done at the higher level and the reading and commentary of the Quran (Ampadu and Mohammed, 2006).

2.10.2 Ahmadiyya Muslims and Islamic education

The Ahmadiyya Muslim education was introduced into the country in 1921 with its headquarters at Saltpond in the central region. They adopted the educational approaches of the Christian missionaries: primary level, secondary level and tertiary level. The course content was the combination of Islamic education and that of the western education and technology (Ampadu and Mohammed, 2006).

2.11 Education ordinance

An education ordinance can be described as a document backed by law stating the direction of education in a given place. Before 1844 the people of the Gold Coast lived as different groups under their tribal chiefs. On 6th March 1844 in the Cape Coast Castle, some coastal chiefs (mostly Fante Chiefs) signed the famous Bond of 1844 with the British governor then to put their people under British rule. settlements on the Gold Coast. The provisions of the Education Ordinance of 1852 were that curricula of schools were to be streamlined to include technical education, provisions were to be made for girls education, there was to be established a Bond of School visitors or supervisors, grant-in-aid was to be given only to government schools in Cape Coast and Accra and the whole educational programme was to be financed from the Poll Tax Ordinance which enjoined every man, woman and child to pay one shilling as tax (Ampadu and Mohammed, 2006).

The Education Ordinance of 1852 failed because the ordinance was premature and over-ambitious since the people of the Gold Coast Colony were not politically mature to accept the provision on taxation for education. With the failure of the 1852 ordinance, the Education Ordinance of 1882 came into being. The government was using whatever money was available to assist the missions. The aims of the 1882 Ordinance were to raise the standard of education in mission and private schools to introduce a system of "Single control of education" and create a systematic basis for providing grants-in-aid to schools. The Education Ordinance of 1882 also failed as a result of lack of funds (Ampadu and Mohammed, 2006).

The 1887 Education Ordinance was intended to amend the 1882 Ordinance. The Education Ordinance of 1887 was to be established a new Board of Education Ordinance for Gold Coast with the power to make rules for the inspection of schools and certification of teachers. The administration of all "assisted schools" placed in the hands of Manager instead of Local

School Boards. Industrial schools were to be opened where boys and girls could learn practical subjects. The factors that led to the failure of the Education Ordinance of 1887 were financial constraints and the negative effects of the "Payment by Results" system (Ampadu and Mohammed, 2006).

2.12 Governor Rodger's educational reforms of 1909

At the beginning of the 20th century conditions in the country and other political developments caused the need to reform education in the Gold Coast and these culminated in the 1909 educational reforms under Governor John Rodger.The recommendations provided under Governor Rodger's Rodger's Reforms of 1909 included: end to the "Payment by Results" system, Agricultural education emphasized and Technical education was recommended (Ampadu and Mohammed, 2006).

2.13 Governor Guggisberg's educational reforms (1919 – 1927)

Governor Gordon Guggisberg was appointed by the British government in 1919 to the Gold Coast. He set up a committee under Oman (the then Director of Education). The report of the committee included fifty-two (52) recommendations and fifty-three (53) suggestions which included the introduction of English as a subject of instruction as early as possible and the use of the vernacular as a medium of instruction, preparation of vernacular textbooks in line with the policy of old Basel mission, need for better training for teachers and greatly improved conditions of service and recommended the establishment of a secondary, boarding school for boys. Guggisberg brought about the sixteen principles of educational reforms, and among those reforms provision of secondary schools with an education standard that will fit young men and women to enter a university. The government also had ultimate control of education throughout the Gold Coast (Ampadu and Mohammed, 2006).

2.14 Developments in education under the accelerated development plan (1951)

In 1951, Dr. Kwame Nkrumah laid before parliament an Accelerated Development Plan for Education. Some of the proposals of the Accelerated Development Plan of Education were:

- A six- year basic primary course for all children of school going age at public expense
- Infant-junior schools were to be known as primary schools and senior primary schools were to be known as middle schools and were to be regarded as part of the post-primary system
- Additional day secondary schools were to be provided and certain non-assisted secondary schools were to be assisted.
- Four secondary-technical schools were to be provided including the conversion of the Government Technical Schools at Takoradi. Technical schools were to be established at Tarkwa, Accra and Sekondi-Takoradi.

  The middle schools in the northern territories were to increase in number as quickly as possible and more potential teachers were to be provided (Ampadu and Mohammed, 2006).

2.15 Status of education in Ghana at the end of the colonial period

As a result of earlier educational plans including that of Governor Guggisberg and more specifically the Accelerated Development Plan of education, enrolment levels in primary and middle schools rose considerably and by 1957, at the attainment of independence, there were well over four hundred and fifty thousand (450,000) children in primary schools and about 115,831 middle school pupils in the country. At the end of colonial period, the curriculum implemented in Gold Coast; primary, middle and secondary schools was substantially modeled on that of British schools. The subjects taught were geared towards academic and literary education. By the time of independence, Twi, Ewe and Ga were recognized by the Cambridge Examination Syndicate as examinable subjects. In some of the existing secondary schools like Mfantsipim, Adisadel and Achimota, Sixth form courses were to be taken after GCE Ordinary level (Ampadu and Mohammed, 2006).

By the end of the colonial period, there were about thirty-eight (38) secondary schools with a total enrolment of 9,860 students. All the schools were receiving government assistance. The secondary school curriculum most especially was more academic and "bookish". By the end of the colonial period, technical education had been made the responsibility of a separate department under a Chief Technical Education Officer. In addition, seven technical institutions had been opened including Tarkwa School of Mines (now University of Mining Technology). The Kumasi College of Technology was also opened in 1952 to provide courses of technological and vocational training (Ampadu and Mohammed, 2006).

In 1948, the British colonial government opened the University College of the Gold Coast (Legon) to meet the increasing demand for higher education. One great contribution the CPP government under Dr. Kwame Nkrumah made as part of the Accelerated Development Plan was the establishment of the West African Examinations

Council (WAEC) in 1952 to take over the running of the West African School Certificate Examination which was then in the hands of the Cambridge University Local Examination Syndicate (Ampadu and Mohammed, 2006).

2.16 Post-Independence developments in education

Ten years after the introduction of the Accelerated Development, the Education Act of 1961 was passed. It was a legal instrument passed in parliament stating what the government policy was regarding the development and promotion of education in republican Ghana. The act clearly spelt out the way education was to be organized (Ampadu and Mohammed, 2006).

The Education Act of 1961 was passed because the Accelerated Development Plan had run for ten years and as such there was the need to introduce innovations to tune up education to meet the philosophy and aspirations of the government and the citizens. Under the Education Act of 1961, the Minister of Education was given sweeping powers to make regulations which had the force of law, make regulations affecting the examination and appointment of teachers, the curriculum, keeping of school records and the power to close down a private institution if he is satisfied that it is dangerous or potentially dangerous to the physical or moral welfare of the pupils (Ampadu and Mohammed, 2006).

2.17 Curriculum innovations during the period: 1966 – 69

On 24th February, 1966, the CPP government, Dr. Kwame Nkrumah who introduced the Education Act of 1961 was overthrown in a military coup. This change in governance resulted in a review of the educational provisions in Ghana. The National Liberation Council (NLC) which overthrew the first republican government appointed a new Commissioner of Education who administered education in the country until 1969. The NLC government appointed two bodies, Kwapong Education Review Committee and the Mills – Odoi Commission to look into some aspects of the educational system in Ghana (Ampadu and Mohammed, 2006).

2.17.1 The Kwapong Education Review Committee

In 1966, the government set up an education review committee under the chairmanship of Professor Alex Kwapong to study the country's educational system and make recommendations for change. Some of the recommendations were:

- The introduction of textbook fee at the basic level at the rate of three cedis per pupil in the middle school.
- That the system of six years of primary education followed four years of middle school course should be replaced by an integrated basic eight-year course for children between the ages of 6 and 14, and at the end of the 8-year course, pupils were to proceed secondary schools through a Common Entrance Examination to be conducted by the West African Examinations Council (WAEC).

- That there should be considerable expansion in public secondary school system (the total number of secondary schools rose from 103 in 1966 to 108 in 1969). That the duration of secondary education was to be five years leading to school Certificate of GCE "O" Level, followed by two-years sixth form course leading to GCE "A" Level (Ampadu and Mohammed, 2006).

2.17.2 Mills – Odoi Commission

The Commission after a comprehensive study recommended that there should decentralization of the management of secondary schools and teacher training colleges from the Ministry of Education to the regional and district authorities and the establishment of a Teaching Service Division of the Public Service Commission (now Ghana Education Service) (Ampadu and Mohammed, 2006).

2.17.3 The Dzobo Committee report on education (1972)

In 1972 the democratically elected government was overthrown in another military coup that brought the National Redemption Council of Col. I. K. Acheampong to power. The changes resulted in the call for the educational change in the country. These calls culminated in the setting up of the Dzobo Educational Committee in 1972 to reform the existing structure and content of education in the country (Ampadu and Mohammed, 2006).

The education system in Ghana by the 1970s was generally accused of being too academic in content and being almost wasteful. In 1972, the Dzobo Committee on education was appointed to consider the proposals of the Ministry of Education alongside other proposals. The new structure and content of education proposed by the Dzobo Committee are given as:

- Kindergarten education: 2 years for children 4 – 6 years.
- Primary school education: 6 years for children 6 – 12 years.
- Junior Secondary School education: 3 years for students 12 -15 years.
- Lower Senior Secondary education: 2 years for students 15 -17 years.
- Upper Senior Secondary education: 2 years for students 17 -19 years.

The recommendations from the Dzobo Committee were forwarded to the Ministry of Education in 1974 and the implantation date was set for September 1975. The recommendations of the educational committee could not be implemented on a national scale. A limited number of Experimental Junior Secondary Schools were designated to start the new educational programme on pilot basis. The delay continued until 1987 when the PNDC government implemented the programme but with some minor modifications (Ampadu and Mohammed, 2006).

The 1987 New Educational Reform Programme

The objectives of the 1987 New Education Reform Programme were to reduce the duration of the pre-university education from the then seventeen years to twelve years, increase access to education at the basic and secondary school levels and to improve the quality, efficiency and relevance of pre-university education by expanding the curriculum of both primary and junior secondary school. The New Educational Reforms ran from 1987 until the FCUBE intervention was introduced in 1996 due to certain constitutional demands.Under the New Educational Reform, the introduction of Basic Education Certificate Examination (B.E.C.E.) and the Senior Secondary School Certificate Examination (S.S.S.C.E.) now the West African Senior Secondary School Certificate Examination (W.A.S.S.C.E.) was made to replace the General Certificate of Examination (G.C.E.) (Ampadu and Mohammed, 2006).

2.19 The FCUBE Educational intervention of 1995 – 2005

In 1992, Ghana returned to constitutional rule and the 1992 Constitution became the guiding document that spelt out how the country should be governed. The 1992 Constitution of Ghana under article 39(2) demanded that the government should implement a Free Compulsory and Universal Basic Education (FCUBE) Programme for Ghanaian children of school-going age by 2005.There were some challenges in the implementation of the FCUBE Programme. Access to education was still limited and so many children of school-going age especially in the northern parts of Ghana and girls were not in school. The management of schools was not efficient enough to be able to use the human and material resources available to achieve quality education and increase access to education. The quality of teaching and learning in schools was not efficient to equip school children with literacy, numeracy and problem-solving skills (Ampadu and Mohammed, 2006).

2.20 The New Educational Reform Review Programme of 2007 – 2015

The FCUBE programme has run its ten-year full term by 2005 (1995 – 2005). The government had to review education to address current needs and aspirations. A review was needed to address the weaknesses in the existing programme. The proposed structure and content of the New Education Reform Review Programme is given below:

2.20.1 Basic education (first cycle)

It comprises 2-year Kindergarten education for 4 to 6 years old children to provide them with pre-school learning skills; 6-year Primary School education in basic literacy, numeracy, Science and Social Studies; 3-year Junior High School education (General Curriculum for entry into Grammar, Technical, Vocational or Agricultural course at the Senior High School) (Ampadu and Mohammed, 2006).

2.20.2 Secondary education (second cycle)

This comprises 4-year Senior High School Education in Grammar, Technical, Vocational or Agricultural Studies; the entry into the Senior High School will be through an externally supervised examination in the core subjects of English, Mathematics, Science, Social Studies and a Ghanaian Language. After Senior High School, students who do not want to continue further could terminate and enter into apprenticeship training. The government in 2009 changed the 4-year secondary education back to 3-year programme (Ampadu and Mohammed, 2006).

2.20.3 Tertiary education (third cycle)

After the students had successfully completed the secondary education, they can continue their education in Colleges of Education (Teacher Training Colleges), Professional Institutes, Polytechnics and Universities (Ampadu and Mohammed, 2006).

2.21 Capitation grant in Ghana

In order to meet the Millennium Development Goals (MDGs) for education the government abolished all fees charged by schools and also provided schools with small grant for each pupil enrolled. The programme was first piloted with World Bank support in Ghana's most deprived districts in 2004 (USAID, 2007). The enrolment in schools then increased by 14.5 percent and for that matter, the government extended the "Capitation Grant" system nationwide in 2005 (Adamu-Issah et al., 2007). Under the capitation grant system, every public Kindergarten, Primary and Junior High School receives a grant of about 4.50 Ghana cedis per pupil per year. Schools were not allowed to charge any fees from parents (Adamu-Issah et al., 2007).

2.22 Computerized Schools Selection and Placement System (CSSPS)

Presently, Ghana has 12,630 primary schools, 5,450 junior high schools and 503 senior high schools. As at 2005, the admissions of pupils into the public second cycle schools was done through Computerized Schools Selection and Placement System (CSSPS). As part of the measures to ensure smooth placement of qualified BECE candidates, the Ghana Education Service (GES) has put in place the CSSPS which has been operational since 2005 (Ghana Education Service, 2011).The candidates must choose six (6) schools (1st – 6th) and they can only select one school from category A. The candidates can only select two schools from category B and one school cannot be chosen twice. If a candidate makes five (5) choices one each from categories A, B, C, D and T, the sixth choice must be made from category C or D (Ghana Education Service, 2011).

Candidates who desire to offer purely Technical Programme may select all their six (6) schools from category T. Candidates who have the desire to offer courses in private institutions may decide to select all their six (6) schools from category P. Candidates have the option to select their six (6) schools from mixture of categories A, B, C, D, T and P. Regardless of the categories, candidates must arrange their choices in order of preference (Ministry of Education and Ghana Education Service, 2011). Concerning the re-entry admissions, qualified BECE candidates who were unable to get admission into senior high school or Technical Institutes within the immediate past three years will be eligible for consideration for selection and placement alongside the current year's candidates. Admission of foreign candidates can also be considered. Foreign candidates refer to any Non-Ghanaian candidates seeking admission into senior high schools or Technical or Vocational Institute in Ghana or Ghanaian citizens domiciled in foreign countries and returning to Ghana to continue their education and children of Ghanaian citizens on foreign missions (Ghana Education Service, 2011).

Foreign students are required to write an aptitude test before considered for placement. The computerized system uses the total processed raw scores of six subjects instead of grades of each candidate for the selection. The subjects comprise core subjects namely; English Language, Mathematics, Integrated Science, Social Studies and two other best subjects. For Technical Institutes the subjects are English Language, Mathematics, Integrated Science, Basic Design and Technology and two other best subjects. To qualify for selection and placement a candidate's grade in any of the four core subjects are expected not to exceed 5. The minimum grade for each of the best two subjects should not exceed 6, and if added to the four core subjects, must not exceed an aggregate of 30. The selection is on merit, where the computer places all candidates into their first choice schools using the ranking order (Ghana Education Service, 2011).

The raw scores of each candidate are used to do the ranking. The ranking may displace first choice candidates with second choice candidates in that manner. This principle also applies to placement of all candidates and is on merit and not choice of candidates. A candidate whose grade for any of the core subjects exceeds 6 or has any of his or her core subjects cancelled by WAEC will be deemed not qualified for selection and placement

Ghana Education Service, 2011).Where there is a problem of non-placement after the 6th choice, such candidates are listed and the list processed by the CSSPS Secretariat as special cases. For 2012 and subsequent selection and placement years, 30% catchment area allocation is considered as the seventh choice. Placement results are released to Mobile Telephone Companies where candidates access their results by texting their index numbers plus year of completion to special codes which are issued by the Mobile Companies (Ghana Education Service, 2011).

2.23 Time series

A time series is a sequence of observations taken sequentially in time. For example, a monthly sequence of the quantity of goods shipped from factory, a weekly series of the number of road accidents and hourly observations made on the yield of a chemical process. If the set of observations is continuous, the time series is said to be continuous. Also, if the set is discrete, the time series is said to be discrete. Time series can be applied in fields like economics, business, engineering, natural sciences and social sciences. Time series analysis is concerned with techniques for the analysis of time series observations. This requires the development of stochastic and dynamic models for time series data and the use of such models in important areas of application (Box et al., 1994).

CHAPTER THREE

METHODOLOGY

3.0 Time series

Time series is a time dependent sequence, $Y_1, Y_2, Y_3, \dots \dots, Y_N$ or $\{Y_t\}, t \in N$ where $1, 2, 3, \dots \dots \dots, N$ denote time steps. The mean (μ_t) and the variance ($\sigma_t{}^2$) of the observations $Y_1, Y_2, Y_3, \dots \dots, Y_N$ or $\{Y_t\}, t \in N$ are given as $E(Y_t)$ and $E(Y_t - \mu_t)^2$ respectively.Time series can also be explained as a sequence of observations on a single entity measured at regular time intervals. For example, yearly gross national product, monthly traffic fatalities, weakly weights of a dieter and the enrolment of pupils in the second cycle schools as it is used in this study can be classified as time series.

3.1 Deterministic time series

If a time series can be expressed as a known function then, it is said to be a deterministic time series, that is, $Y_t = f(t)$. (3.1)

3.2 Stochastic time series

If a time series can be expressed as $Y_t = X(t)$,(3.2) where X is a random variable.Then $\{Y_t\}$ is a stochastic time series.

3.3 Objectives of time series analysis

There are several possible objectives in analyzing a time series. These objectives may be classified as description, explanation, prediction and control.

3.3.1 Description

When presented with a time series data the first step in the analysis is usually to plot the data to obtain simple descriptive measures of the main properties of the series such as seasonal effect and trend. Apart from trend and seasonal variation, the outlier to look for in the graph of the time series is the possible presence of turning points, where for example, an upward trend has suddenly changed to a downward trend.

3.3.2 Explanation

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

3.3.3 Prediction

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

3.3.4 Control

When a time series is generated which measures the quantity of a manufacturing process, the aim of the analysis may be to control the process. Control procedures are of several different kinds. In statistical quality control for instance, the observations are plotted on control charts and the controller takes action as a result of studying the charts, Box and Jenkins have described a more sophisticated control strategy which is based on fitting a stochastic model to the series, from which future values of the series are predicted. The values of process variables predicted by the model are taken as target values and the variables conform to the target values.

3.4 Components of a time series

Traditional methods of time series analysis are mainly concerned with decomposing the variation in a series into the various components of trend, periodic and stochastic.

3.4.1 Periodic component

If $Y_t = Y_{t+T} + e_t \forall t \in N$, (3.3)

then the time series has a periodic component of period T.

3.4.2 Trend component

If $Y_t = y + \beta_t + e_t$ (3.4)

then there exists a linear trend with the slope being β.

3.4.3 Stochastic component

If $Y_t = y_t + e_t$ (3.5)

then e_t is the stochastic component of the time series.

3.5 Stationary time series

A time series is said to be strictly stationary if the joint distribution of $X_{t1}, \ldots \ldots .., X_{tn}$ is the same as the joint distribution of $X_{t1+T}, \ldots \ldots \ldots .., X_{tn+T}$ for all $t_{1+T}, \ldots \ldots \ldots, t_{n+T}$. In order words, shifting the time origin by an amount T has no effect on the joint distributions, which must therefore depend only on the intervals between $t_1, \ldots \ldots \ldots . t_n$.

3.6 Achieving stationarity

If there is trend in the mean then differencing the time series data will remove the trend and stationarity will be achieved. For non-seasonal data, first order differencing is usually sufficient to attain stationarity, so that the new series $Y_1, Y_2, \ldots \ldots .., Y_{N-1}$ is formed from the original series$\{y_t = X_{t+1} - X_t = \nabla X_{t+1}\}$. (3.6)

First order differencing is widely used, but occasionally second-order differencing is required using the operator, ∇^2 wher $\nabla^2 X_{t+2} = \nabla X_{t+2} - \nabla X_{t+1} = X_{t+2} - 2X_{t+1} + X_t$ (3.7)

If there is a trend in variance, the series is made stationary by transforming the data as follows:

$Y_t = U_t$ (3.8)

where$U_t = \log X_t$ (3.9)

3.7 Autocorrelation function

The autocorrelation function measures the degree of correlation between neighbouring observations in a time series. The autocorrelation function at lag k is defined as

$$\rho_k = \frac{E[(Y_t - \mu_Y)(Y_{t+k} - \mu_Y)]}{[E(Y_t - \mu_Y)^2 E(Y_{t+k} - \mu_Y)^2]^{1/2}} \Rightarrow \rho_k = \frac{cov(Y_t, Y_{t+k})}{\sigma_{Y+k}} \qquad (3.10)$$

The autocorrelation coefficient is estimated from sample observations using the formula;

$$r_k = \frac{\sum_{t=2}^{n}(Y_t - \overline{Y}_t)(Y_{t+k} - \overline{Y}_{t+k})}{\sum_{t=1}^{n}(Y_t - \overline{Y}_t)^2} \qquad (3.11)$$

3.8 The sampling distribution of autocorrelation coefficients

The autocorrelation coefficients of a random data are approximately normal with $\mu_{\rho k}$ and $\sigma_{\rho k} = \frac{1}{n}$ where n is the size of the sample. Thus for a random sample of size 40 we expect $-2\sigma_k \leq r_k \leq 2\sigma_k$ for significance limits of two standard errors which is $\frac{-2}{\sqrt{40}} \leq r_k \leq \frac{2}{\sqrt{40}}$ which equals $-0.316 \leq r_k \leq 0.316$. Statistical theory teaches us that the sampling distribution of autocorrelation coefficients of random data are normal with $\mu_{\rho k} = 0$ and $\sigma_{\rho k} = \frac{1}{\sqrt{n}}$. If r_k is the estimate of ρ_k, then we can use our knowledge of normal distributions to make interval estimates of the population autocorrelation coefficient ρ_k.

3.9 Partial autocorrelation function

Partial autocorrelation function measures the degree of association between Y_t and Y_{t+k} when the effects of other time lags on Y are held constant. The partial autocorrelation function (PACF) denoted by $\{\emptyset_{kk}: k = 1, 2, \dots \dots \dots ..\}$, the set of partial autocorrelations at various lags k are defined by $\emptyset_{kk} = \frac{|P_k{}^*|}{|P_k|}$ (3.12)

Where P_k is the $k \times k$ autocorrelation matrix, and $P^*{}_k$ is P_k with the last column replaced by $[\rho_1, \rho_2, \dots \dots \dots \dots .\rho_k]^T$ and $\rho_k = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ & \vdots & & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & & \cdots & 1 \end{bmatrix}$ (3.13)

So $\emptyset_{11} = \emptyset_1 = \rho_1$ and $\emptyset_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_1 \end{vmatrix}} \Rightarrow \emptyset_{22} = \frac{\rho_2 - \rho_1{}^2}{1 - \rho_2{}^2}$ (3.14)

and estimates of $\emptyset_{kk}$ can be obtained by replacing the ρ_1 by r_1.

3.10 The sampling distribution of the partial autocorrelation coefficients

The partial autocorrelation coefficients of random data are approximately normal with $\mu_{\emptyset kk} = 0$ and $\sigma_{\emptyset kk} = \frac{1}{\sqrt{n}}$ where n is the size of the sample. Thus for a random sample of size 40, we expect $-2\sigma\emptyset_{kk} \leq r_k \leq 2\sigma\emptyset_{kk}$ for significance limits of two standard errors which is $\frac{-2}{\sqrt{40}} \leq \emptyset_{KK} \leq \frac{2}{\sqrt{40}}$ which is equal to $-0.316 \leq \emptyset_{kk} \leq 0.316$. Hence any value of $\emptyset_{kk}$ lying outside this interval is said to be significantly different from zero.

3.11 Purely random process

A discrete-time process is called a purely random process if it consists of a sequence of random variables$\{e_t\}$, which are mutually independent and identically distributed. Such a process has constant mean and variance and

$\gamma(k) = cov(e_t,\ e_{t+k}) = 0$ For $k = 1,2,3, \dots$ (3.15)

It is also strictly stationary. The autocorrelation function (acf) is given by $\rho(k) = \begin{cases} 1 \dots \dots \dots . k = 0 \\ 0 \dots \dots k = +1, +2, \end{cases}$(3.16)

where $\rho(k)$ is the autocorrelation function (acf). A purely random process is sometimes called a white noise.

3.12 An autoregressive model of order $p[AR(P)]$

An autoregressive model of order p denoted by $AR(p)$ is a special kind of regression in which the explanatory variables are past values of the process. An autoregressive model of order p is given by $Y_t = \sum_{k=1}^{p} \alpha_k Y_{t-k} + \mu + e_t$ (3.17)

where μ is the mean of the time series data and e_t is a white noise. The order of an $AR(p)$ process is determined by the partial autocorrelation function (PACF). An $AR(p)$ process has its pacf cutting off after lag p and the autocorrelation function (ACF) decays. For example, the partial autocorrelation function (PACF) of an $AR(1)$ process cuts off after lag one (1). The $AR(1)$ process is given by $Y_t = \alpha_1 Y_{t-1} + \mu + e_t$(3.18)

From the above equation, putting $\mu = 0$, we have $Y_t = \alpha_1 Y_{t-1} + e_t$.

Multiplying through by Y_{t-k}we have $Y_{t-k}Y_t = \alpha_1 Y_{t-k}Y_{t-1} + e_t Y_{t-k}$ (3.19)

$cov(Y_{t-k}, Y_t) = \alpha_1 cov(Y_{t-k}, Y_t) + cov(Y_{t-k}, e_t)$. (3.20)

But $cov(Y_{t-k}, e_t) = 0$. Since Y_{t-k} depends only on $e_{t-k}, e_{t-k-1}, \dots \dots \dots \dots \dots$

which are not correlated with e_t for $k > 0$. Hence $Y_k = \alpha_1 Y_{k-1}$. (3.21)

Dividing through by γ_k we have $\frac{\gamma_k}{\gamma_0} = \alpha_1 \frac{\gamma_{k-1}}{\gamma_0} (\rho_k = \alpha_1 \rho_{k-1}\ where\ \rho_0 = 1)$.

We have $\rho_1 = \alpha_1 \rho_0 = \alpha_1 (since\ \rho_0 = 1). \rho_1 = \alpha_1$,For $= 2 \rho_2 = \alpha_1 \rho_1 = \alpha_1(\alpha_1) = {\alpha_1}^2$,

For $k = 3$, $\rho_3 = \alpha_1 \rho_2 = \alpha_1({\alpha_1}^2) = {\alpha_1}^3$, and in general, $\rho_k = {\alpha_1}^k$ (3.22)

3.13 Yule-Walker equations

The important fact about autoregressive models is that it is possible to obtain a simple set of linear equations that expresses the parameters of the model in terms of autocorrelations and variance of the data. These linear equations are called the Yule-Walker equations. The general $AR(p)$ model is

$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + \cdots \dots \dots \dots \dots \dots + \alpha_p Y_{t-p} + e_t$ (3.23)

where Y_t is a zero mean process and e_t is a white noise process and $E(e_t, Y_{t-k}) = 0$ for $k > 0$.

$\gamma_k = E(Y_t, Y_{t-k})$.(3.24)

Considering the general $AR(p)$ we have

$$\gamma_k = E(Y_t, Y_{t-k}) = E\big[\big(\alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_p Y_{t-p} + e_t\big)Y_{t-k}\big]$$

$\gamma_k = \alpha_1 E(Y_{t-1}, Y_{t-k}) + \alpha_2 E(Y_{t-2}, Y_{t-k}) + \cdots + \alpha_p E\big(Y_{t-p}, Y_{t-k}\big)E(e_t, Y_{t-k})$.

(3.25)

The auto covariance γ_k of a stationary process is a function of the lag between observations, not their starting point. Thus $\gamma_k = E(Y_t, Y_{t-k}) = E(Y_{t+s}, Y_{t+s-k})$

(3.26)

For any value$s$, since $(t + s - k) = k$. Thus

$\gamma_k = \alpha_1 \gamma_{k-1} + \alpha_2 \gamma_{k-2} + \cdots + \alpha_p \gamma_{k-p}$(3.27)

for $k > 0$ $E(e_t, Y_{t-k}) = 0$ and $\gamma_k = \alpha_1 \gamma_{k-1} + \alpha_2 \gamma_{k-2} + \cdots + \alpha_p \gamma_{k-p}$

Dividing through by γ_0 and recalling that $\rho_k = \frac{\gamma_k}{\gamma_0}$

we obtain the Yule-Walker equations,

$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2} + \cdots + \alpha_p \rho_{k-p}$, $k > 0$. (3.28)

For k = 0 we obtain $\gamma_0 = {\sigma_Y}^2 = \alpha_1 \gamma_{-1} + \alpha_2 \gamma_{-2} + \alpha_p \gamma_{-p} + E(e_t, Y_t)$.(3.29)

But for the auto covariance $\gamma_{-s} = \gamma_s$ and

$E(e_t, Y_t) = E(\alpha_1 Y_{t-1} e_t + \alpha_2 Y_{t-2} e_t + \cdots + \alpha_p Y_{t-p} e_t + {e_t}^2) = E({e_t}^2)$

$E(e_t, Y_t) = {\sigma_e}^2.$ (3.30)

Because $E(Y_{t-k}, e_t) = 0k > 0,$

thus $\gamma_0 = {\sigma_Y}^2 = \alpha_1 Y_1 + \alpha_2 Y_2 + \cdots \alpha_p Y_p + {\sigma_e}^2$ (3.31)

and substituting $\rho_k = \frac{\gamma_k}{\gamma_0}$, we obtain $1 = \alpha_1\rho_1 + \cdots + \alpha_p\rho_p + \frac{{\sigma_e}^2}{{\sigma_Y}^2}$ or

${\sigma_e}^2 = {\sigma_Y}^2(1 - \alpha_1\rho_1 - \cdots - \alpha_{p)}.$ (3.32)

3.14 Estimating $AR(p)$ parameters using the Yule-Walker equations

The Yule-Walker equations can be used to estimate the parameters of the $AR(p)$ model. Given an $AR(2)$ process, the parameters α_1 and α_2 can be estimated. Suppose $\rho_1 = 0.5$and $\rho_2 = 0.1$, then α_1 and α_2 can be found using the Yule-Walker equations with $k = 1$ and $k = 2$, thus

$k = 1$: $\rho_1 = \alpha_1\rho_0 + \alpha_2\rho_1$ (3.33)

Thus, $\alpha_1 + 0.5\alpha_2 = 0.5$ (3.34), since $\rho_0 = 1$. For $k = 2$, we obtain

$\rho_2 = \alpha_1\rho_1 + \alpha_2\rho_{0=}0.5\alpha_1 + \alpha_2$(3.35)

Solving equations (3.34) and (3.35), we get $\alpha_1 = 0.6$ and$\alpha_2 = -0.2$. Hence the model is $Y_t = 0.6Y_{t-1} - 0.2Y_{t-2} + e_t$. (3.36)

3.15 Estimating AR (p) parameters using the method of ordinary least squares

The method of ordinary least squares can be employed to estimate the parameters of the $AR(p)$ process. In multiple regression we have

$Y = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + e_t$ (3.37)

and$\beta = (X^T X)^{-1} X^T Y$ (3.38)

where$\beta = [\beta_0,\ \beta_1, \ldots\ , \beta_k]^T$, (3.39)

$X = [X_1, X_2, \ldots, X_n]^T$ (3.40)

and $Y = [Y_1, Y_2, \ldots, Y_n]^T$. (3.41)

Similarly, with the AR process the X vector is formed using the past values of Y. For example, consider the AR (1) process:

$Y_t = \alpha_1 Y_{t-1} + e_t + \mu.$ (3.42)

Hence $Y_2 = \alpha_1 Y_1 + \mu$ (3.43)

$Y_3 = \alpha_1 Y_2 + \mu$ (3.44)

$Y_t = \alpha_1 Y_{t-1} + \mu.$ (3.45)

This equation is over determined and it is solved using the ordinary least squares method. The X vector is $(Y_1, Y_2, \ldots., Y_n)^T$ and the Y vector is $\left(Y_{1-\mu}, \ldots \ldots \ldots., Y_{t-\mu}\right)^T$. Then

$\alpha_1 = (A^T A) A^T Y.$ (3.46)

3.16 Moving Average Model of order q MA (q)

The Moving Average (MA) models provide predictions of Y based on a linear combination of past forecast errors. In particular the Moving Average (MA) model of order q is given by

$$Y_t = \sum_{k=1}^{q} \theta_k \, e_{t-k} + \mu + e_t. \qquad (3.47)$$

3.17 Autocorrelation function (ACF) of MA (q)

$$\gamma(k) = cov(Y_t, Y_{t-k})$$

$$= cov(\theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t+k-q}, \theta_1 e_{t+k-1} + \theta_2 e_{t+k-1} + \cdots + \theta_q e_{t+k-q})$$

$$= \begin{cases} 0 \ldots\ldots\ldots\ldots k > q \\ \sigma_e{}^2 \sum_{i=1}^{q-k} \theta_i \theta_{i+k} \ \ldots\ldots\ldots\ldots k = 0, 1, \ldots, q.. \\ \gamma(k) \ldots\ldots k < 0 \end{cases} \qquad (3.48)$$

Since $cov(e_s, e_t) = \begin{cases} \sigma_e{}^2 \ldots\ldots\ldots\ldots s = t \\ 0 \ldots\ldots\ldots\ldots\ldots\ldots s \neq t \end{cases}$ (3.49)

Hence the autocorrelation function (acf) of the MA(q) process is given by

$\rho(k) = 1$ for $k = 0$ (3.50)

$\rho(k) = \sum_{i=1}^{q-k} \theta_i \theta_{i+k} / \sum_{i=1}^{q} \theta_i{}^2 \ for \ k = 1, 2, 3, \ldots, q.$ (3.51)

$\rho(k) = 0 \ for \ k > q$ (3.52)

$\rho(k) = \rho(-k)\ for\ k < 0.$ (3.53)

The order of the MA(q) process is given by the autocorrelation function (ACF). The ACF cuts off after lag q and the partial autocorrelation function (PACF) decays to zero. Thus an MA (1) process cuts off after lag one. In order words the acf after lag one will not be significantly different from zero.

3.17.1 Moving average process of order one MA (1)

The MA (1) process is given by

$$\rho(k) = \begin{cases} 1\ \ldots\ldots\ldots\ldots\ldots k = 0 \\ \theta_1/(1+\theta_1{}^2)\ \ldots\ldots\ldots. k = \pm 1 \\ 0\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots otherwise \end{cases} \quad (3.54)$$

Thus, $\rho_1 = \frac{\theta_1}{1+\theta_1{}^2}$ (3.55)

$$\rho_1 + \rho_1\theta_1{}^2 - \theta_1 = 0$$

$\rho_1\theta_1{}^2 - \theta_1 + \rho_1 = 0$ (3.56)

The parameters are the roots of a quadratic. This means that we can find two MA (1) process that correspond to the same autocorrelation function (ACF). To establish a one-to-one correspondence between the acf and the model and to obtain a converging autoregressive representation, we restrict the moving average parameter such that $|\theta| <$

1. This restriction is known as the invertibility condition. Invertibility implies that the process can be written in terms of an autoregressive model.

3.18 Estimation of the model parameters of the MA (q) process

For an MA (1) process an iterative method is used since the ordinary least squares cannot be used as the residual sum of squares is not a quadratic function. The approach suggested by Box and Jenkins is used. Given the MA (1) model

$Y_t = \theta_1 e_{t-1} + \mu + e_t$ (3.57)

whereμ and θ_1 are constants and $r_1 = \frac{\theta_1}{1+{\theta_1}^2}$. (3.58)

Select suitable values of μ and θ_1 such that $\mu = \overline{Y}$ and θ_1 given by the solution of $r_1 = \frac{\theta_1}{1+{\theta_1}^2}$.

Then the corresponding residual sum of squares is calculated using $Y_t = \mu + e_t + \theta_1 e_{t-1}$ recursively in the form $e_t = Y_t - \mu - \theta_1$ with $e_0 = 0$, we have

$e_1 = Y_1 - \mu$ (3.59)

$e_2 = Y_2 - \mu - \theta_1 e_1$ (3.60)

$e_3 = Y_3 - \mu - \theta_1 e_2$ (3.61)

$e_N = Y_N - \mu - \theta_1 e_{N-1}.$ (3.62)

Then, $\sum_{t=1}^{N} e_t{}^2$ is calculated. This procedure is then repeated for other values of μ and θ_1 and the sum of squares $\sum_{t=1}^{N} e_t{}^2$ computed for a grid of points in the (μ, θ_1) plane. We then determine by inspection the least square estimates of μ and θ_1 which minimize$\sum_{t=1}^{N} e_t{}^2$.

3.19 The backward-shift operator

The backward-shift operator B operates on a series to move it back one time unit. That is

$BY_t = Y_{t-1}$ (3.63)

$B^2Y_t = Y_{t-2}$ (3.64)

$B^kY_t = Y_{t-k}$(3.65)

For example an AR (2) process can be expressed using the B operator as follows.

$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + e_t = \alpha_1 BY_t + \alpha_2 B^2 Y_t + e_t = (\alpha_1 B + \alpha_2 B^2)Y_t + e_t$(3.66)

Thus, $Y_t = \alpha_1 Y_{t-1} - \alpha_2 Y_{t-2} = e_t$ (3.67)

and$Y_t(1 - \alpha_1 B - \alpha_2 B^2) = e_t$ (3.68)

In general an AR(p) process is given by $Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i}$ (3.69)

may be expressed in terms of the B operator as follows:

$\left(1 - \alpha_1 B - \alpha_2 B^2 - \alpha_3 B^3 - \cdots - \alpha_p B^p\right) Y_t = e_t.$ (3.70)

Writing the expression in brackets as a polynomial in B, denoted AR (B), we have

$AR(B)Y_t = e_t$. (3.71)

The MA model has a similar representation. For example an MA (2) process can be expressed using the B operator as follows:

$Y_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + e_t = (\theta_1 B + \theta_2 B^2)e_t + e_t$. (3.72)

In general an MA (q) process is given by $Y_t = \sum_{i=1}^{q} \theta_i e_{t-i}$ (3.73)

may be expressed in terms of the B operator as follows:

$Y_t = \left(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q\right)e_t$. (3.74)

Similarly, writing the expression in brackets as a polynomial in B denoted MA (B), we get

$Y_t = MA(B)e_t$. (3.75)

3.20 The duality of AR and MA processes

The Random Walk process given by $Y_t = Y_{t-1} + e_t$. (3.76)

can be re-written as an infinite moving average. Indeed, consider the following moving average,

$Y_t = e_t + e_{t-1} + e_{t-2} + \cdots = \sum_{i=0}^{\infty} e_{t-i}$ (3.77)

$$= (1 + B + B^2 + \cdots)e_t = \left(\sum_{i=0}^{\infty} B^i\right)e_t. \quad (3.78)$$

Recall that $\sum_{i=0}^{\infty} y^i = 1/(1-y)$, is valid when $|y| < 1$. Hence $Y_t = \left(\frac{1}{1-B}\right) e_t$ so that

$$(1 - B)Y_t = e_t$$

$$Y_t - BY_t = e_t \quad (3.79)$$

$$Y_t - Y_{t-1} = e_t\ ; \qquad Y_t = Y_{t-1} + e_t (3.80)$$

which is a random walk process. This means that a finite autoregressive process is an infinite moving average process. The inverse is also true; that is, finite moving-average processes are infinite autoregressive processes. For example, MA (1) process is an infinite autoregressive process as shown below:

$$Y_t = e_t - \theta e_{t-1} \quad (3.81)$$

Using the B operator notation we have $Y_t = (1 - \theta_1 B)e_t$

$$\frac{Y_t}{1-\theta_1 B} = e_t \quad (3.82)$$

$$\left(1 + \theta_1 B + {\theta_2}^2 + \cdots\right)Y_t = e_t (3.83)$$

$$Y_t + \theta_1 Y_{t-1} + {\theta_2}^2 Y_{t-2} + \cdots = e_t. \quad (3.84)$$

3.21 Autoregressive Moving Average (ARMA or "Mixed" Process)

Consider the process given by;

$Y_t = \alpha_1 Y_{t-1} + \theta_1 e_{t-1} + e_t$(3.85)

which can be written as $Y_t - \alpha_1 Y_{t-1} = e_t + \theta_1 e_{t-1}$ (3.86) or

$(1 - \alpha B)Y_t = (1 - \theta B)e_t$ (3.87)

$AR(B)Y_t = MM(B)e_t$ (3.88)

This is called mixed or autoregressive moving average (ARMA) process of order (1, 1). If $|\theta| < 1$ it can be rewritten as:

$(1 - \alpha B)\left(\frac{1}{1+\theta B}\right) Y_t = e_t$ (3.89)

$(1 - \alpha B)(1 - \theta B + \theta^2 B^2 - \theta^3 B^3 + \cdots)Y_t = e_t$ (3.90)

$[(1 - \alpha + \theta)B + (\alpha\theta + \theta^2)B^2 + \cdots]Y_t = e_t.$ (3.91)

This is an infinite order AR process. This is true if $|\propto| < 1$ and $|\theta| < 1$. i.e. if the AR is stationary and MA is stationary and invertible. If there are two polynomials in B, MA (B) and AR(B), and an ARMA model, $AR(B) = MA(B)e_t.$ (3.92)

It is possible to write the model as an infinite AR process:

$\left[\frac{AR(B)}{MA(B)}\right] Y_t = e_t$(3.93)

or an infinite MA process:

$Y_t = \left[\frac{MA(B)}{AR(B)}\right] e_t$ (3.94)

and approximate either by finite process.

ARMA processes are parsimonious, however identifying them using ACF and PACF may be difficult. The condition necessary for dividing by AR (B) is that the AR process is stationary and by MA (B) is that MA process being invertible.

3.22 Autoregressive Moving Average Models (ARMA)

A more general model is a mixture of the AR (p) and MA (q) models and is called an autoregressive moving average model (ARMA) of order (p, q). The ARMA (p, q) is given by

$Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{i=1}^{q} \theta_i e_{t-i} + \mu + e_t$.(3.95)

An important characteristic of ARMA models is that both the ACF and PACF do not cut off as in AR and MA models.

3.22.1 ARMA (1, 1) model

An example of an ARMA (p, q) model is the ARMA (1, 1) model given by

$Y_t = \mu + \alpha_1 Y_{t-1} + \theta_1 e_{t-1} + e_t$. (3.96)

The ARMA (1, 1) model is stationary if $-1 < \alpha_1 < 1$ and invertible if $-1 < \theta_1 < 1$.Its theoretical autocorrelation function (acf) and partial autocorrelation function (pacf) trail off to zero in a damped exponential. In an ARMA (1, 1) model both acf and the pacf trail off to zero.

3.23 Estimating the parameters of an ARMA model

The procedure for estimating the parameters of the ARMA model is like the one for the MA model, it is an iterative method. Like the MA the residual sum of squares is calculated at every point on a suitable grid of the parameter values and the values which give the minimum sum of squares are the estimates. For an ARMA (1, 1) the model is given by

$$Y_t - \mu = \alpha_1(Y_{t-1} - \mu)e_t + \theta_1 e_{t-1}. \qquad (3.97)$$

Given N observations $Y_1, \dots, Y_N$, we guess values for μ, α_1 and θ_1. Set $e_0 = 0$ and $Y_0 = \mu$ and then calculate the residuals recursively by

$$e_1 = Y_1 - \mu \qquad (3.98)$$

$$e_2 = Y_2 - \mu - \alpha_1(Y_1 - \mu) - \theta_1 e_1. \qquad (3.99)$$

$$e_N = Y_N - \mu - \alpha_1(Y_1 - \mu) - \theta_1 e_{N-1}. \qquad (3.100)$$

The residual sum of squares $\sum_{i=1}^{N} e_i{}^2$ is calculated. Then other values of μ, α_1, θ_1 are tried until the minimum residual sum of squares is found. It has been found that most of the stationary time series occurring in practice can be fitted by AR(1), AR(2), MA(1), MA(2), ARMA(1, 1) or white noise models.

3.24 The Autoregressive Integrated Moving Average Model (ARIMA)

If a non-stationary time series which has variation in the mean is differenced to remove the variation the resulting time series is called an integrated time series. It is called an integrated model because the stationary model which is fitted to the differenced data has to be summed or integrated to provide a model for the non-stationary data. Notationally, all AR(p) and MA(q) models can be represented as ARIMA models. For example an AR(1) can be represented as ARIMA(1,0,0); that is no differencing and no MA part. The general model is ARIMA(p,d,q) where p is the order of the AR part, d is the degree of differencing and q the order of the MA part. Writing $W_t = \nabla^d Y_t = (1-B)^d Y_t$, the general ARIMA process is of the form

$$W_t = \sum_{i=1}^{p} \alpha_i W_{t-1} + \sum_{i=1}^{q} \theta_i e_{t-1} + \mu + e_t \quad (3.101)$$

3.24.1 ARIMA (1, 1, 1) process

An example of ARIMA (p,d,q) is the ARIMA(1,1,1) which has one autoregressive parameter, one level of differencing and one MA parameter is given by

$$W_t = \alpha_1 W_{t-1} + \theta_1 e_{t-1} + \mu + e_t. \quad (3.102)$$

$(1-B)Y_t = \alpha_1(1-B)Y_{t-1} + \theta_1 e_{t-1} + \mu + e_t$ (3.103)

which can be simplified further as

$Y_t - Y_{t-1} = \alpha_1 Y_{t-1} - \alpha_1 Y_{t-2} + \theta_1 e_{t-1} + \mu + e_t.$

$Y_t - Y_{t-1} = \alpha_1(Y_{t-1} - Y_{t-2}) + \theta_1 e_{t-1} + \mu + e_t.$ (3.104)

3.25 Stationarity and invertibility conditions of specific time series

In the table below, the stationarity and invertibility conditions of specific time series models and the behaviour of their theoretical acf and pacf functions are displayed.

Table 3.1. Specific time series models

ARIMA model	Stationarity condition	Invertibility condition	ACF coefficients	PACF coefficients
(1,d,0)	$-1 < \alpha_1 < 1$	None	dies down	Cuts off after lag one
(2,d,0)	$\alpha_1 + \alpha_2 < 1$ $\alpha_1 - \alpha_2 < 1$ $-1 < \alpha_2 < 1$	None	dies down	cuts off after lag two
(0,d,1)	None	$-1 < \theta_1 < 1$	cuts off after lag one	dies down
(0,d,2)	None	$\theta_1 + \theta_2 < 1$ $\theta_2 - \theta_1 < 1$ $\theta_2 < 1$	cuts off after lag two	dies down
(1,d,1)	$-1 < \alpha_1 < 1$	$-1 < \theta_1 < 1$	dies down	dies down

3.26 Seasonal Autoregressive Integrated Moving-Average Model (SARIMA)

Using the multiplicative seasonal ARIMA (SARIMA) model we have the general notation

$(p, d, q) \times (P, D, Q)s$ where p, d, q is the non-seasonal part and P, D, Q the seasonal part with p, d, q having their usual meaning and P is the order of the seasonal AR process, D the differencing of the seasonal process, Q the order of seasonal MA process of the time series and s is the order of seasonality.

For the purposes of identifying a seasonal ARIMA process, we divide the process into two parts. To identify the seasonal pattern, we ignore the non-seasonal process and determine whether the seasonality is determined by an AR or an MA process by focusing on the coefficients of the seasonal terms. Suppose that the non-seasonal part is an ARIMA (1,0,1) and the time series shows a quarterly seasonal pattern, then the complete model becomes $(1 - \alpha_1 B)(1 - \alpha_4 B^4)Y_t = (1 - \theta_1 B)e_t$. (3.105)

If seasonality is on the AR portion, or $(1 - \alpha_1 B)Y_t = (1 - \theta_1 B)(1 - \theta_4 B^4)e_t$. (3.106)

If seasonality is on the MA portion, where $(1 - \theta_4 B^4)Y_t = Y_t - \theta_4 B^4 Y_t$ (3.107)

In a similar manner $(1 - \theta_4 B^4)e_t = e_t - \theta_4 B^4 e_t$ (3.108)

$= e_t - \theta_4 e_{t-4}$ (3.109)

An MA(2) seasonal process with one level of differencing is expressed by

$(1-B)(1-B_4)Y_t = (1-\theta_1 B - \theta_2 B^2)e_t$ (3.110)

3.27 Purely seasonal models

A purely seasonal time series is one that has only seasonal AR or MA parameters. Seasonal autoregressive models are built with parameter called seasonal autoregressive parameters (SAR parameters). The SAR parameters represent autoregressive relationships that exist between time series data separated by multiples of the number of periods per season. For example, a model with one SAR parameter is written as $Y_t = \alpha_s Y_{t-s} + e_t$. (3.111)

ARIMA$(P, D, Q)^S = ARIMA(1,0,0)^S$ where s is the number of periods per season. The parameter is called the SAR parameter of order s. A general seasonal autoregressive model with p SAR parameters is written as follows:

$Y_t = \sum_{i=1}^{p} \alpha_{is} Y_{t-is} + e_t$(3.112)

whereY_{t-s} is order s, Y_{t-2s} is order $2s$,..., and Y_{t-ps} is of order ps.

3.27.1 The ARIMA$(\mathbf{2,0,0})^{\mathbf{4}}$

The model ARIMA$(2,0,0)^4$ is a quarterly seasonal autoregressive model of order two. The model is $Y_t = \alpha_4 Y_{t-4} + \alpha_8 Y_{t-8} + e_t$. (p=2, s=4); (3.113)

3.28 Seasonal Moving-Average models

Seasonal Moving Average models are built with seasonal moving-average parameters (SMA parameters). SMA parameters represent moving-average relationships that exist among the time series observations separated by a multiple of the number of periods per season. The seasonal moving-average model with Q parameters is $Y_t = \sum_{i=1}^{Q} \theta_{is} e_{t-is} + e_t$. (3.114)

3.29 The ARIMA $(\mathbf{0,0,1})^{\mathbf{4}}$

The model ARIMA$(0,0,1)^4$ is a quarterly seasonal moving average of order one, that is, it has one seasonal moving average parameter. A model with one SMA parameter is written as

$$Y_t = \theta_s e_{t-s} + e_t. \quad (3.115)$$

3.30 Mixed SAR and SMA model

A mixed SAR and SMA model is given by $Y_t = \sum_{i=1}^{p} \alpha_{is} Y_{t-is} + \sum_{i=1}^{q} \theta_{is} e_{t-is} + e_t$. (3.116)

The order of the seasonal ARMA model is expressed in terms of both PS and QS. For purely SMA models, the autocorrelations die down and partial autocorrelations cut off after one seasonal lag for an SAR (1) model. Similarly, the partial autocorrelations die down for SMA models. Also, the autocorrelations cut off after lag one lag for an SMA (1) model and after two lags for an SMA (2) model. For a mixed with one SAR and one SMA both the autocorrelation function and the partial autocorrelation function die down.

3.31 Stationarity and invertibility conditions of specific pure seasonal time series models

The stationarity and invertibility conditions of specific pure seasonal time series models and the behaviour of their theoretical acf and pacf are presented in a tabular form as:

Table 3.2. Specific pure seasonal time series models

ARIMA model	stationarity condition	invertibility condition	acf coefficients	pacf coefficients
$(1,0,0)^s$	$-1 < \alpha_s < 1$	None	dies down	cuts off after one seasonal lag
$(2,0,0)^s$	$\alpha_s + \alpha_{2s} < 1$ $\alpha_s - \alpha_{2s} < 1$ $-1 < \alpha_{2s} < 1$	None	dies down	cuts off after two seasonal lags
$(0,D,1)^s$	None	$-1 < \theta_s < 1$	cuts off after one seasonal lag	dies down
$(0,D,2)^s$	None	$\theta_s + \theta_{2s} < 1$ $\theta_{2s} - \theta_s < 1$ $\theta_{2s} < 1$	cuts off after two seasonal lags	dies down
$(1,D,1)^s$	$-1 < \alpha_s < 1$	$-1 < \theta_s < 1$	dies down	dies down

3.32 The Box-Jenkins method of modeling time series

The Box-Jenkins methodology is a statistically sophisticated way of analyzing and building a forecasting model which best represents a time series. The first stage is the identification of the appropriate ARIMA models through the study of the autocorrelation and partial autocorrelation functions. For example, if the partial autocorrelation function cuts off after lag one and the autocorrelation function decays then ARIMA is identified.

The next stage is to estimate the parameters of the ARIMA model chosen. The third stage is the diagnostic checking of the model. The Q statistic is used for the model adequacy check. If the model is not adequate then the forecaster goes to stage one to identify an alternative model and it is tested for adequacy and if adequate then the forecaster goes to the final stage of the process.

The fourth stage is where the analyst uses the model chosen to forecast and the process ends. Below are the steps for the representation of Box-Jenkins process:

Step one

Collect data for forecasting

Step two

Identify the ARIMA model

Step three

Estimate parameters in tentative ARIMA model

Step four

Do diagnostic checking to determine whether the model is adequate

If the model is adequate, go to step five. If the model is not adequate, update the ARIMA and go back to step two and repeat the process.

Step five

Use the model to forecast

3.33 Identification techniques

The first step in model identification is to determine whether or not the series is stationary. If the series is not stationary, it can be converted to a stationary series by the method of differencing. Upon obtaining a stationary and/or invertible series, we must identify the form of the model to be used. The form is established by comparing the autocorrelation and partial autocorrelation coefficients of the data to be fitted with the corresponding distributions for the various ARIMA models.

If autocorrelations trail off exponentially to zero, an AR model is identified. Similarly, if the partial autocorrelations trail off to zero, then an MA model is identified. If both autocorrelations and partial autocorrelations trail off to zero then a mixed ARMA model is indicated. The order of the AR model is indicated by the number of partial autocorrelations and the order of the MA model by the number of autocorrelations that are statistically different from zero.

Identification methods are rough procedures applied to a set of data to indicate the kind of representational model that will be further investigated. The aim here is to obtain some idea of the values p, d, and q needed in the general linear ARIMA model and to obtain initial estimates for the parameters. The task here is to identify an appropriate subclass of models from the general ARIMA family $\alpha(B)\nabla^d Y_t = \theta_0 + \theta(B)e_t$ (3.117)

which may be used to represent a given time series. The approach will be as follows;

(a) To difference Y_t as many times as is needed to produce stationarity reducing the process under study to the mixed autoregressive moving average process $\alpha(B)w_t = \theta_0 + \theta(B)e_t$(3.118) where

$w_t = (1 - B)^d Y_t = \nabla^d Y_t.$ (3.119)

(b) To identify the resulting ARMA process, the principal tools for putting (a) and (b) into effect are the sample autocorrelation function and the sample partial autocorrelation function. Apart from helping to guess the form of the model, they are used to obtain approximate estimates of the parameters of the model. These approximations are useful at the estimation stage to provide starting values for iterative procedures employed at that stage.

3.34 Use of the autocorrelation and partial autocorrelation functions in identification

A stationary mixed autoregressive-moving average process of order (p, 0, q),

$\alpha(B)Y_t = \theta(B)e_t$, (1.120)

autocorrelation function satisfies the difference equation $\alpha(B)\rho_k = 0, k > 0$. (3.121)

Also, if $\alpha(B) = \prod_{i=1}^{p}(1 - G_i B)$. (3.122).

The solution of this difference equation for the kth autocorrelation is, assuming distinct roots of the form

$$\rho_k = A_1 {G_1}^k + A_2 {G_2}^k + \cdots + A_p {G_p}^k, \; k > q - p. \qquad (3.123)$$

The stationarity requirement that the zeros of $\alpha(B)$ lie outside the unit circle implies that the roots $G_1, G_2, G_3, \ldots, G_k$ lie outside the unit circle.

Inspection of the equation$\rho_k = A_1 {G_1}^k + A_2 {G_2}^k + \cdots + A_p {G_p}^k, \; k < q - p$ (3.124)

shows that in the case of a stationary model in which none of the roots lie close to the boundary of the unit circle, the autocorrelation function will quickly die out or decay for moderate and large k. Suppose that a single real root, say G_1, approaches unity, so that $G_1 = 1 - \delta$ where δ is a small positive quantity. Then since for k large, $\rho_k = A_1(1 - k\delta)$, (3.125)

the autocorrelation function will not die out quickly and will fall off slowly and very nearly linearly. Similarly if more than one root approaches unity the autocorrelation function will decay slowly. Therefore if the autocorrelation function dies out slowly it implies there is at least a root which approaches unity. As a result failure of the estimated autocorrelation function to die out rapidly might logically suggest that the underlying stochastic process is non-stationary in Y_t but possible stationary in ∇Y_t, or in some higher difference.

It is therefore assumed that the degree of differencing necessary to achieve stationarity has been reached when the autocorrelation function of $W_t = \nabla^d Y_t$ (3.126)

dies out fairly quickly.

3.35 Identifying the resultant stationary ARMA process

The autocorrelation function of an autoregressive process of order p tails off, its partial autocorrelation function has a cut-off after lag p. For a mixed autoregressive process both the autocorrelation function and the partial autocorrelation function tails off. Furthermore, the autocorrelation function for a mixed process, containing a pth-order autoregressive component and qth-order moving average component, is a mixture of exponentials and damped sine waves after the first q-p lags. Conversely, the partial autocorrelation function for a mixed process is dominated by a mixture of exponentials and damped sine waves after the first q-p lags. For example, the partial autocorrelation function of AR (1) process cuts off after lag one and the autocorrelation function tails off or decays.

Other tools for model identification

The tools for model identification to be considered are Akaike's Information Criteria (AIC) andSchwarz's Bayesian Information Criterion (BIC).

Akaike's Information Criteria (AIC)

The Akaike's Information Criteria (AIC) which was proposed by Akaike uses the maximum likelihood method. In the interpretation of the approach, a range of potential ARMA model is estimated by maximum likelihood methods, and for each, the AIC is calculated, given by

$$\text{AIC}(p,q) = \frac{-2\ln(maximized\ likelihood)+2r}{N}$$

$$\text{AIC}(p,q) = ln\left(\acute{\sigma_e}^2\right) + r\frac{2}{n} + constant \quad (3.127)$$

Where n is the sample size or the number of observations in the historical time series data, $\acute{\sigma_e}^2$ is the maximum likelihood estimte of σ_e^2 and it is the residual or shock variance, $r = p + q + 1$ (3.128) denotes the number of parameters estimated in the model. Given two or more competing models the one with the smaller AIC values will be selected.

3.36.2 Schwarz's Bayesian Information Criterion (BIC)

Schwarz's BIC like AIC uses the maximum likelihood method. It is given by

$$\mathrm{BIC}(p,q) = ln\left(\hat{\sigma}_e{}^2\right) + \frac{\ln(\mathrm{n})}{\mathrm{n}}, \qquad (3.128)$$

where $\hat{\sigma}_e{}^2$ is the maximum likelihood estimate of $\sigma_e{}^2$, $r = p + q + 1$ denotes the number of parameters estimated in the model, including a constant term and n is the sample size or the number of observations in the time series data. The BIC imposes a greater penalty for the number of estimated model parameters than does AIC. The use of minimum BIC for model selection results in a chosen model whose number of parameters is less than that chosen under AIC.

3.37 Estimating the model's parameters

Once a model is identified the next stage of the Box-Jenkins approach is to estimate the parameters. In this study the estimation of the parameters will be done using the R statistical package. The methods normally used for estimation are the least squares estimates, the method-of-moments estimates and the maximum likelihood estimates. The two function statistics, the Akaike information criterion (AIC) and the Schwarz Bayesian Information Criterion (BIC) are used to penalize fitted models based on the principle of parsimony, which are the checks used in verifying the adequacy of the models chosen. In comparing, models with the smallest AIC and BIC are expected to have residuals which resembles a white noise process.

There will be a reported standard error for a particular parameter for each parameter estimate. A test for statistical significance can be conducted from the parameter estimate and its standard error. A t-test, which is a test of whether a parameter is significantly different from zero is used. The t-value of the test is obtained from the computation of the ratio of each parameter estimate to its standard error. For statistically significant parameters, the absolute values of the t-ratios are expected to be greater than or equal to 2. From the test, significant parameters should be maintained whilst parameters which are not significant should be removed from the model.

The estimated AR and MA parameters must conform to their respective boundary conditions as stated in Tables 3.1 and 3.2. If the AR and MA parameters do not lie within those bounds of stationality and invertibility, then the parameters of that model must be re-estimated, or a new model should be considered for estimation. If the parameters of the selected model are subjected to all the checks explained above and the result conforms to the conditions, such a model is considered for the diagnostic checking stage, else the analyst has to re-start the process from the identification stage.

3.38 Testing the model for adequacy

After identifying an appropriate method for a time series data it is very important to check that the model is adequate. The error terms e_t are examined and for model to be adequate the errors should be random. If the error terms are statistically different from zero, the model is not considered adequate. The autocorrelation function coefficients of white noise data has a sampling distribution that can be approximated by a normal curve with mean zero and standard error, $\frac{1}{\sqrt{n}}$ where n is the number of data points in the observed series. For a white noise process, 95% of all sample autocorrelation values must lie within a range; mean plus or minus 1.96 standard errors. The mean in this case is zero and the standard error is $\frac{1}{\sqrt{n}}$. it is expected that about 95% of all sample autocorrelation values to be within the range of $\pm 1.96\sqrt{n}$ The test. If this condition does not hold, then the fitted model does not follow a white noise process.

The modified version of the portmanteau test statistic, the Ljung-Box test statistic is the Q statistic given by:

$Q = n(n+2)\sum_{i=1}^{k}\frac{{r_i}^2}{n-k}$ (3.129)

which is approximately distributed as a χ^2 with $k-p-q$ degrees of freedom, where n is the length of the time series , k is the first k autocorrelations being checked, p is the order of the AR process and q is the order of the MA process and r is the estimated autocorrelation coefficient of the ith residual term.

The hypothesis to be tested is formulated in the form;

$\mathrm{H_0}$: The set of autocorrelations for residual is white noise (the model fit the data quite well)

$\mathrm{H_1}$: The set of autocorrelations for residual is different from white noise.The test statistic (Q) is compared with the chi-square distribution given as ${\chi^2}_{\alpha,(h-p-q)}$, where α is taken to be 5%, h is the maximum lag being considered, and p and q are the order of the AR and MA processes respectively.

If the calculated value of Q is greater than χ^2 for $k - p - q$ degrees of freedom and 5% significance level, then the model is considered inadequate and adequate if Q is less then χ^2 for $k - p - q$ degrees of freedom. If the model is tested inadequate then the forecaster should select an alternative model and test for the adequacy of the model.

In diagnostic checking, it is also necessary to check whether the residuals of the chosen model are normally distributed with zero mean and constant variance. This can be done through a residual plot against time, and an objective test called the Shapiro-Wilk Normality test. One can also use a residual plot of a histogram and a normal probability graph to check the normality on the series under consideration. The hypothesis of the Shapiro-Wilk test is given by:

H_0: The error terms in the model are normally distributed

H_1: The error terms in the model are different from a normal distribution.

The test is considered significant if the reported p-value is less than 0.05, otherwise it is not significant. In conclusion, if a chosen model does not satisfy these assumptions at the diagnostic stage, the analyst is advised to start the Box-Jenkins process – identification, estimation and diagnostic checking and the cycle is continued until an appropriate model is obtained. After that, the analyst may use such model to forecast (Reagan, 1984) and (Makridakis et al., 1998).

CHAPTER FOUR

DATA ANALYSIS AND RESULTS

4.0 Introduction

This chapter entails an analysis of the yearly enrolment of pupils in public second cycle schools, with data obtained from four (4) public Senior High Schools (SHS); Assin Manso Senior High, Nyankumasi Ahenkro Senior High, Assin Nsuta Senior High and Adankwaman Senior High Schools in the Assin South District between the years, 1991-2012.The statistical computing tool employed for this work is R software, and the Box Jenkins methodology of time series analysis is also used.

4.1 Preliminary Analysis

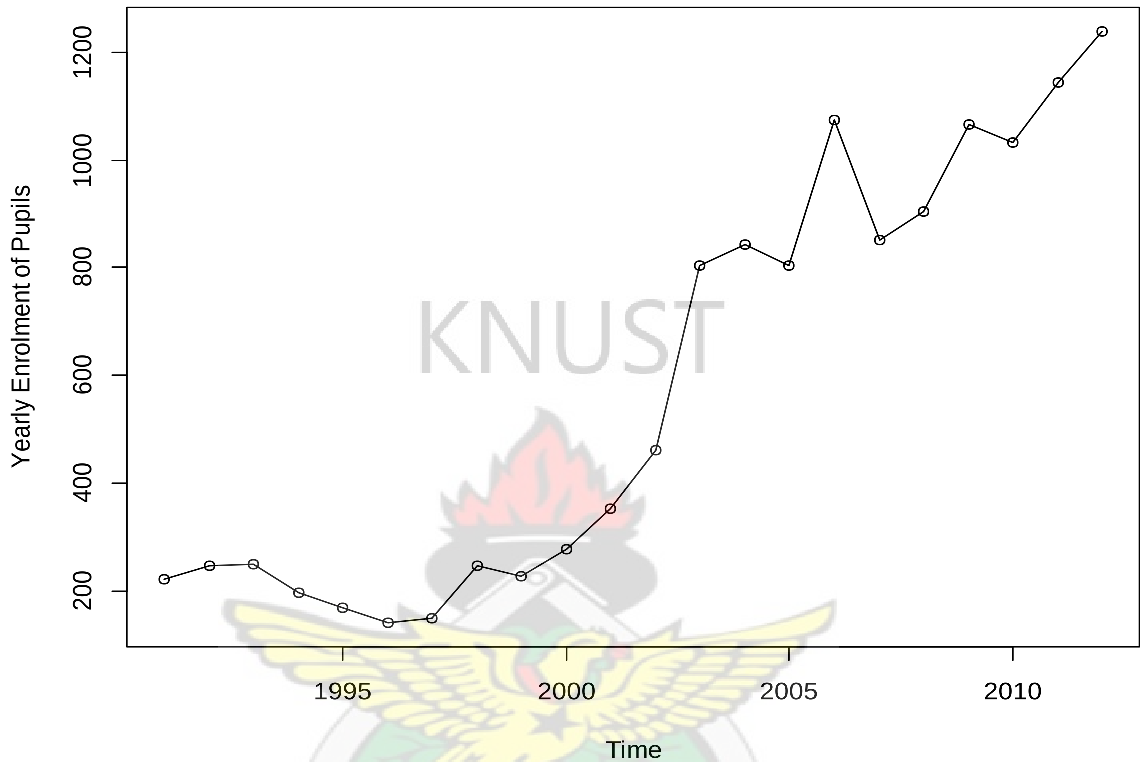

Figure 4.1 : Trend of Yearly Enrolment of Pupils In Second Cycle Public Schools, From 1991 TO 2012

Figure 4.1 shows the pattern of yearly enrolment of pupils in public second cycle schools obtained from the Assin South District, between 1991 and 2012.

Generally, we observe an upward trend throughout the entire period, but somehow stable trend from 1991 to 2000.However, after the year 2000, enrolment increases sharply from 2001 to 2003 and thereafter behaves irregularly. The maximum and minimum enrolments occur in2012 and 1996, recording totals of 1239 and 141 pupils respectively. Also, the upward linear trend that is shown from the observed figure (in Figure 4.1) shows that the data series is non-stationary in mean. The existence of seasonality is also absent from the observed pattern.

Table 4.1: Summary statistics of Enrolment Data

Minimum	1st Quartile	Median	Mean	3rd Quartile	Standard deviation	Maximum
141.0	231.0	405.5	577.0	890.8	391.932	1239.0

From Table 4.1, we observe that the minimum enrolment recorded is 141, which occurred in 1996. The maximum enrolment figure is 1239, and this also occurred in 2012.

The median enrolment figure is nowhere close to the mean enrolment figure. This may indicate a non-symmetric behaviour of the enrolment data distribution.

4.2 Achieving Stationarity

The observed data is first differenced and tested for stationarity.

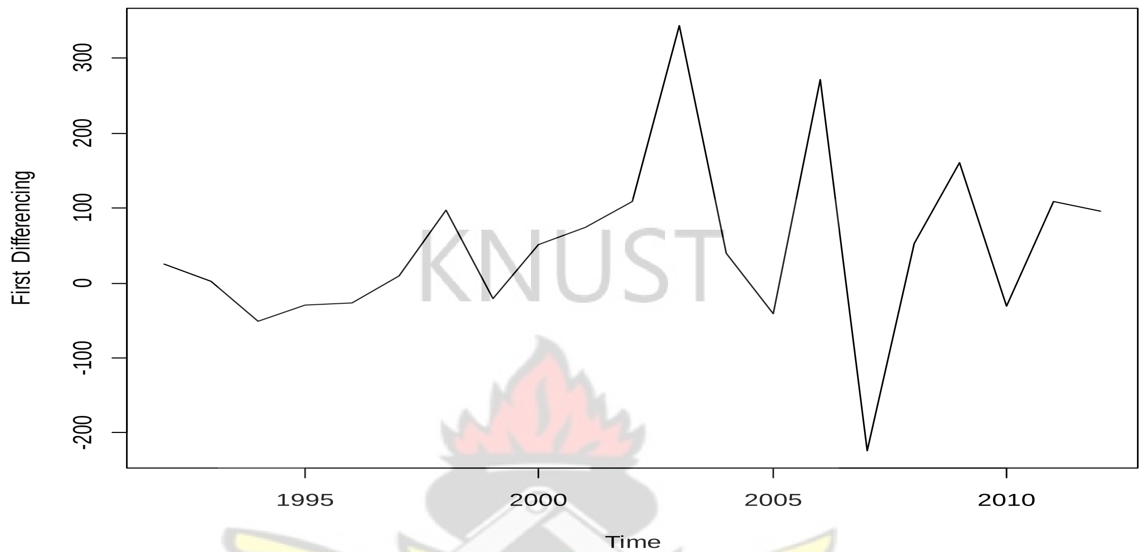

Figure 4.2 : Pattern of Enrolment Data After First Differencing

From Figure 4.2 , it can be seen that after the first differencing, the enrolment data series looks stationary but with some observations still deviating widely from the mean. It can therefore be said that our data is non-seasonal, since for non-seasonal data, first order differencing is usually sufficient to attain apparent stationarity.

We therefore perform a second differencing to observe how stable the data series will look.

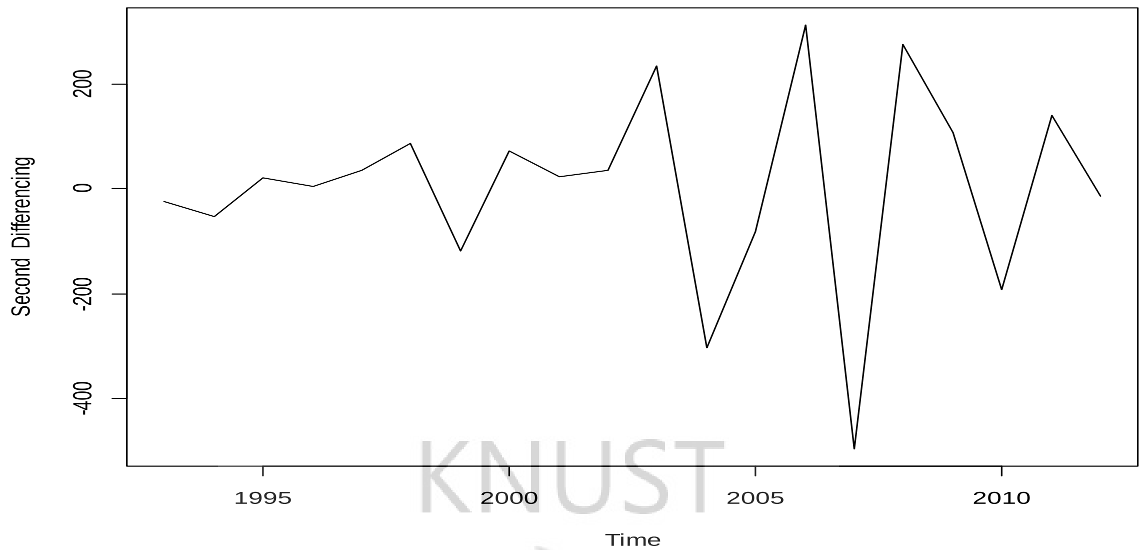

Figure 4.3 : Pattern of Enrolment Data After Second Differencing

From Figure 4.3, it can be seen that after the second differencing, the enrolment data series looks more stable as most observations beat about the mean of zero.

Furthermore, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test is computed to validate the claim that the differenced data is level or trend stationary.

4.3 KPSS Test for Level Stationarity

data: diff1

KPSS Level = 0.2155, Truncation lag parameter = 1, p-value = 0.1

Conclusion: At an α(alpha) 5% level of significance, we fail to reject the Null hypothesis that the differenced enrolment series is trend or level stationary since the p-value (0.1) > 0.05, and hence conclude that the series is stationary.

4.4 Model Identification

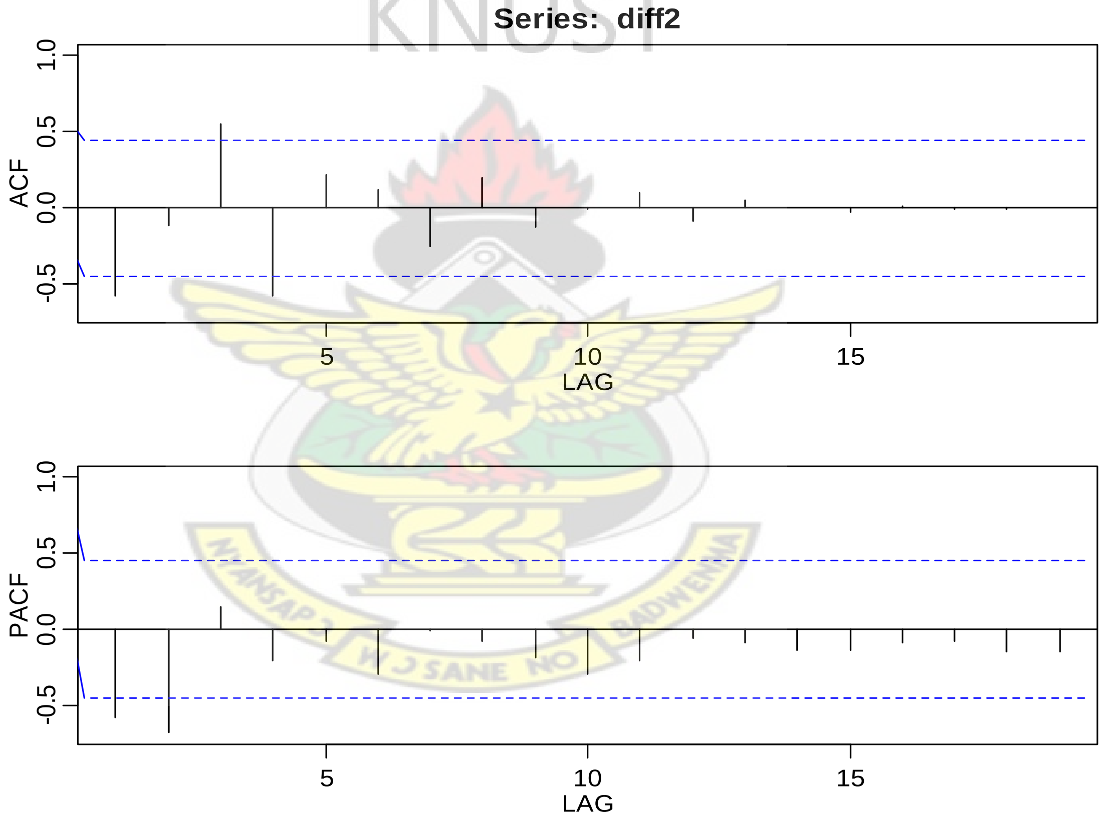

Figure 4.4: Shows the Sample ACF (top) and Sample PACF (bottom).

In order to select the appropriate model and also make more accurate forecasts, we fit several feasible ARIMA models to the observed data by making reference to the Sample

ACF and Sample PACF (in Figure 4.4 above) of the difference data. Since the data was differenced, the fitted ARIMA models would be of order (p, d=2, q).

From Figure 4.4, the sample ACF tails off to zero after lag 4, with only lags 1,3 and 4 exceeding the significant bound. The sample PACF also tails off to zero after lag 2 i.e. all the other lags fall within the significant bounds, dying down exponentially.

From the foregoing analysis, the following ARIMA (Autoregressive integrated moving average) models are therefore possible for the data series:

- ARIMA(2,2,2)
- ARIMA(2,2,1)
- ARIMA(2,2,3)

4.5 Estimation of Parameters and Diagnostic Checking

At this point we proceed to estimate the parameters and investigate whether the residuals of the selected ARIMA models are normally distributed with mean zero and constant variance, and also whether there are no correlations between successive residuals (i.e. randomness of residuals).

To check for correlations between successive residuals, we make use of a correlogram and also the Ljung-Box test to further ascertain the adequacy (randomness) of the chosen model.

Also to check whether the residuals are normally distributed with mean zero and constant variance, we make use of a normality quantile-quantile plot (q-q plot) and a histogram.

If the residuals are normally distributed, the points on the normal quantile-quantile plot should approximately be linear, with residual mean as the intercept and residual standard deviation as the slope whilst the shape of the histogram shows "a bell-like" shape.

- **ARIMA(2,2,2)**

Coefficients:

	ar1	ar2	ma1	ma2
	-1.1017	-0.6878	0.2910	-0.1464
s.e.	0.2005	0.2209	0.3514	0.4311

sigma^2 estimated as 11377: log likelihood = -122.56

AIC = 255.11   AICc = 259.4   BIC = 260.09

$AIC

[1] 10.70303

$AICc

[1] 10.96439

$BIC[1] 9.901399

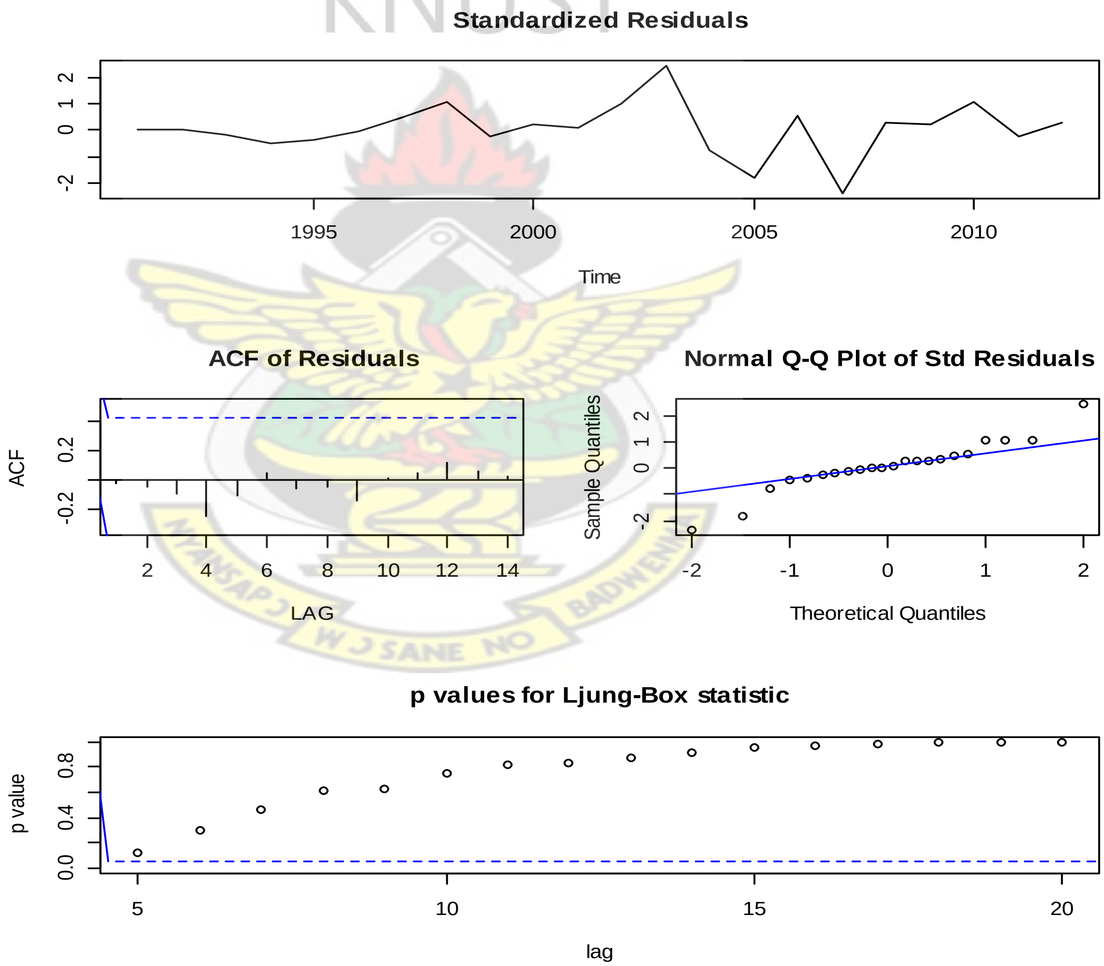

Figure 4.5: Shows the Standardized residuals (top), ACF of residuals and Normality Q-Q plots (middle) and p-values for Ljung-Box Statistic (bottom).

From Figure 4.5, the ACF of residuals shows that none of the sample autocorrelations for the lags exceed the significant bounds. This gives an indication of a white noise process.

Also, from the p-values of the Ljung-box test above all lie above α (alpha) 5% level of significance for all lags.

Hence from these deductions, we fail to reject the null hypothesis that the series of residuals exhibits no autocorrelation and conclude that there is absolutely no evidence for non-zero autocorrelations in the residuals at all lags (i.e. the residuals are independently distributed).

`

Also, from the plot in Figure 4.5,the QQ-normal plot for the residuals gives an indication of a plausible symmetric distribution since most of its residuals do not deviate that much from the line of best fit and it distribution looks approximately linear (Unimodal).

- **ARIMA(2,2,1)**

Coefficients:

	ar1	ar2	ma1
	-1.1147	-0.7406	0.3550
s.e.	0.1867	0.1337	0.3062

sigma^2 estimated as 11467: log likelihood = -122.62

AIC = 253.23   AICc = 255.9   BIC=257.22

$AIC

[1] 10.61993

$AICc

[1] 10.81779

$BIC

[1] 9.768704

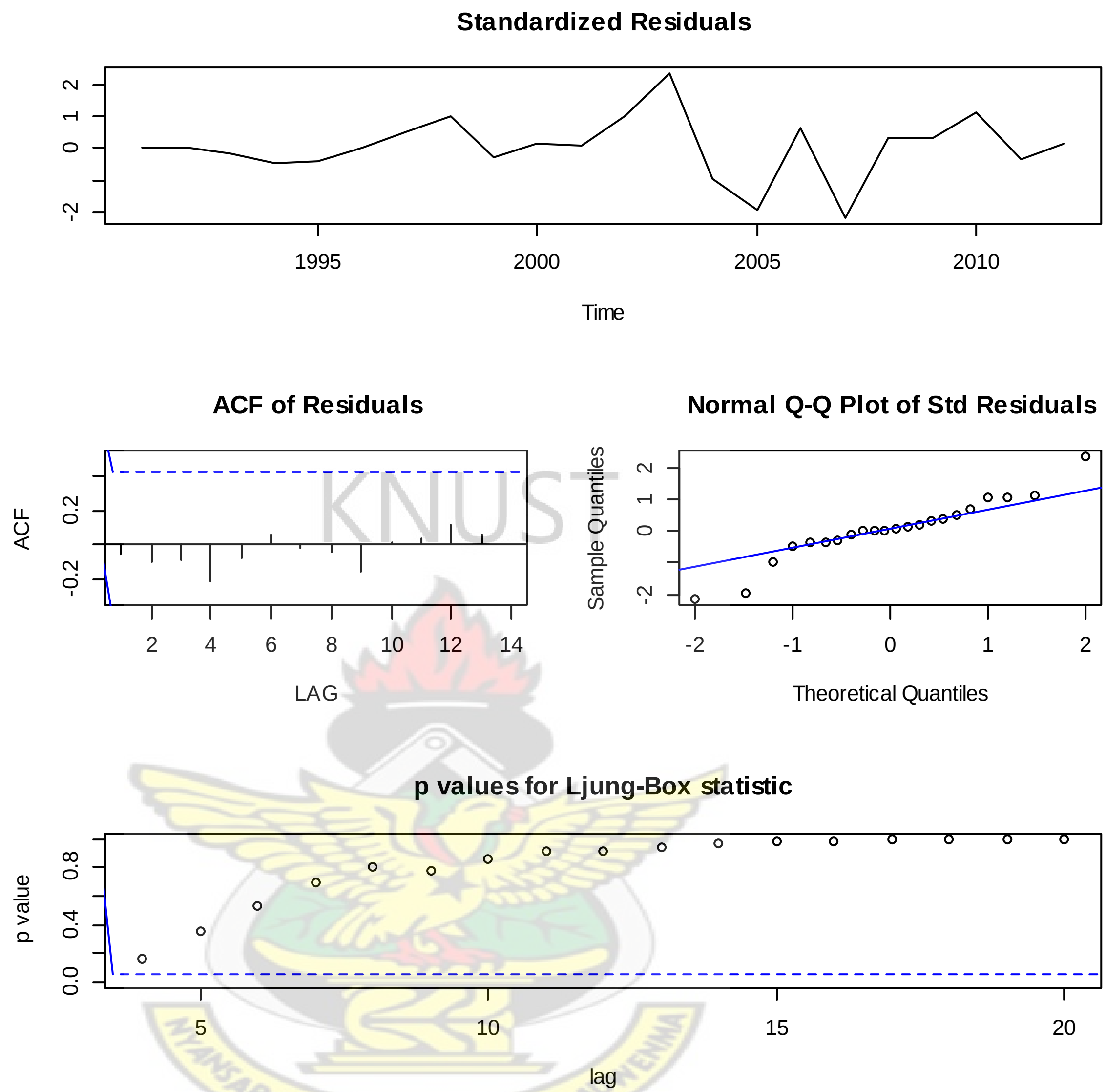

Figure 4.6: Shows the Standardized residuals (top), ACF of residuals and Normality Q-Q plots (middle) and p-values for Ljung-Box Statistic (bottom).

From Figure 4.6, the ACF of residuals shows that none of the sample autocorrelations for the lags exceed the significant bounds. This gives an indication of a white noise process.

Also, from the p-values of the Ljung-box test above all lie above α (alpha) 5% level of significance for all lags.

Hence from these deductions, we fail to reject the null hypothesis that the series of residuals exhibits no autocorrelation and conclude that there is absolutely no evidence for non-zero autocorrelations in the residuals at all lags (i.e. the residuals are independently distributed).Also, from the plot in figure 4.6,the QQ-normal plot for the residuals gives an indication of a plausible symmetric distribution since most of its residuals do not deviate that much from the line of best fit and it distribution looks approximately linear (Unimodal).

- **ARIMA(2,2,3)**

Coefficients:

	ar1	ar2	ma1	ma2	ma3
	-1.2447	-0.8412	0.3559	-0.2906	-0.3661
s.e.	0.2209	0.2139	0.3455	0.4244	0.4042

sigma^2 estimated as 10664: log likelihood=-122.13

AIC=256.25 AICc=262.72 BIC=262.23

$AIC

[1] 10.72918

$AICc

[1] 11.07463

$BIC

[1] 9.977144

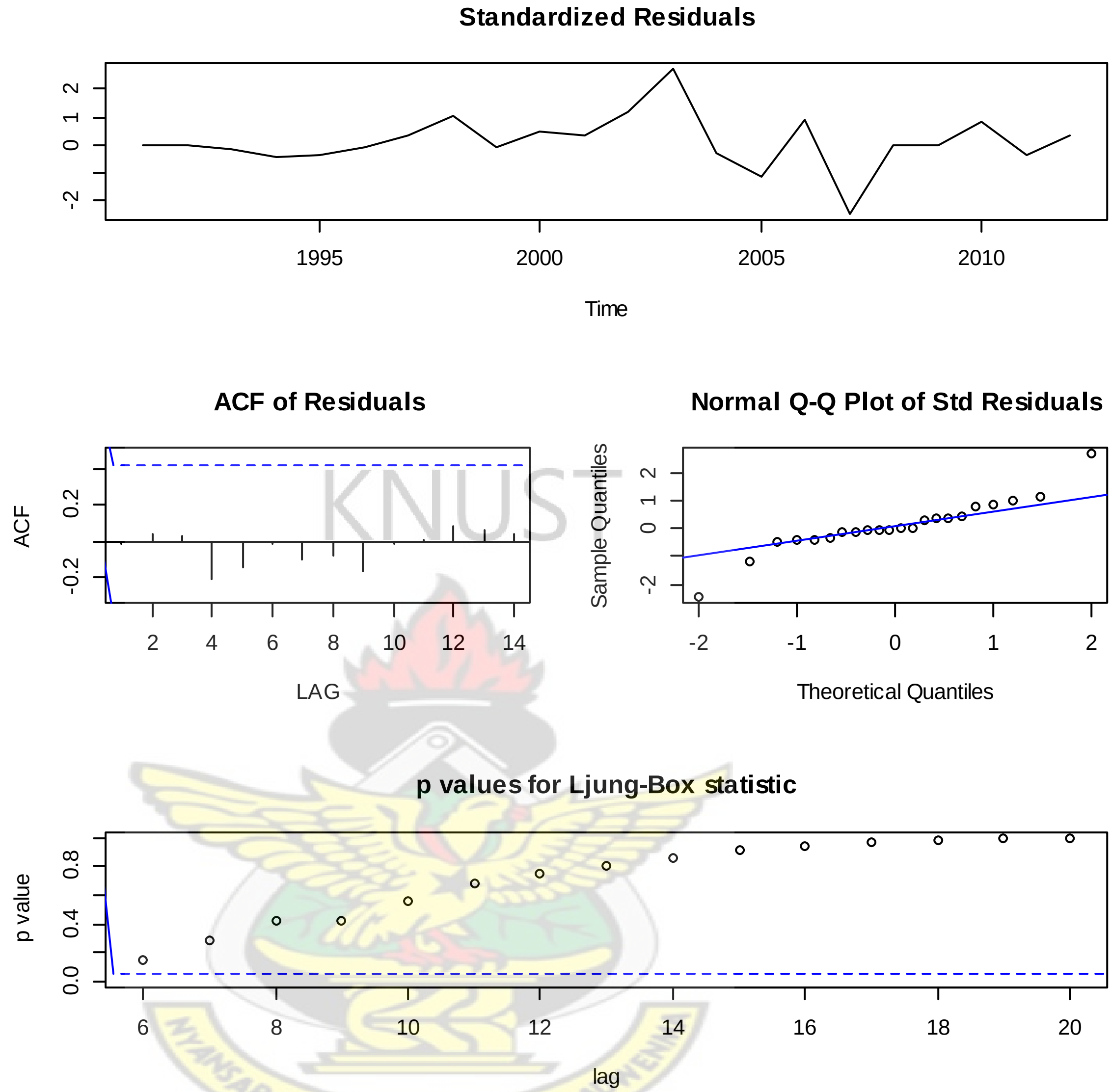

Figure 4.7: Shows the Standardized residuals (top), ACF of residuals and Normality Q-Q plots (middle) and p-values for Ljung-Box Statistic (bottom).

From Figure 4.7 , the ACF of residuals shows that none of the sample autocorrelations for the lags exceed the significant bounds. This gives an indication of a white noise process.

Also, from the p-values of the Ljung-box test above all lie above α (alpha) 5% level of significance for all lags.

Hence from these deductions, we fail to reject the null hypothesis that the series of residuals exhibits no autocorrelation and conclude that there is absolutely no evidence for non-zero autocorrelations in the residuals at all lags (i.e. the residuals are independently distributed).Also, from the plot in Figure 4.7,the QQ-normal plot for the residuals gives an indication of a plausible symmetric distribution since most of its residuals do not deviate that much from the line of best fit and it distribution looks approximately linear (Unimodal).

Table 4.2: Test on Parameter Estimates of Models

MODEL	PARAMETER	ESTIMATE	SE	$\lvert T \rvert -$ VALUE	STATISTICALLY SIGNIFICANT IF $\lvert T \rvert \geq 2$
ARIMA(2,2,2)	ar1	-1.1017	0.2005	5.4948	***Significant.***
	ar2	-0.6878	0.2209	3.1136	***Significant.***
	ma1	0.2910	0.3514	0.8281	***Non-Significant.***
	ma2	-0.1464	0.4311	0.3396	***Non-Significant.***
ARIMA(2,2,1)	ar1	-1.1147	0.1867	5.9705	***Significant.***
	ar2	-0.7406	0.1337	5.5393	***Significant.***
	ma1	0.3550	0.3062	1.1594	***Non-Significant.***

ARIMA(2,2,3)	ar1	-1.2447	0.2209	5.6347	***Significant.***
	ar2	-0.8412	0.2139	3.9327	***Significant.***
	ma1	0.3559	0.3455	1.0301	***Non-Significant.***
	ma2	-0.2906	0.4244	0.6847	***Non-Significant.***
	ma3	-0.3661	0.4042	0.9057	***Non-Significant.***

From Table 4.2, it is realized that the t-values for all MAs' across all models are statistically non-significant while that of the ARs' are statistically significant.

4.6 Model Selection

In order to select the most appropriate model for our data, we compare all competing models and select the one with the minimum AIC (Akaike Information Criterion value), Schwartz Bayesian Information Criterion (BIC) and Residual Variance. Other statistical tests like Root Mean Squared Error (RMSE), Mean Abs. Percent Error (MAPE), Bias

Proportion or Mean Forecast Error (MFE) and Mean Absolute Scaled Error are also used in testing the forecast accuracy of the fitted models.

It should be noted that a model which fits the data well does not necessarily forecast well. The best model will be the one that achieves a compromise between the various information criterion values.

Table 4.3 AIC, BIC and Residual Variance for the possible Models

Model	**Akaike Information Criterion (AIC)**	**Residual Variance**	**BIC**	**AICc**
ARIMA(2,2,2)	255.11	11377	260.09	259.4
ARIMA(2,2,1)	253.23	11467	257.22	255.9
ARIMA(2,2,3)	256.25	10664	262.23	262.72

From Table 4.3, it is clear that ARIMA(2,2,1) model is the best model for forecasting since its AIC and BICvalues are far lower than the other competing models. Also, the MA (1) parameter of the model is statistically non-significant; hence we simply ignore it since it could have a negative effect on the forecast if used for prediction. Therefore from the foregoing deductions, it is clear that ARIMA (2, 2, 0) model is the best model for forecasting.

4.7 Parameter Estimates of ARIMA (2,2,0) with drift

Coefficients:

	ar1	ar2	drift
	-0.9353	-0.6421	51.5323
s.e.	0.1593	0.1520	73995.9866

Training set error measures:

ME	RMSE	MAE	MPE	MAPE	MASE
6.6246241	104.7759157	67.4585827	1.7583785	14.0582386	0.7591802

The chosen model for the Enrolment of pupils in SHS is of the form;

$$Y_t - 2Y_{t-1} + Y_{t-2} = \alpha_1(Y_{t-1} - 2Y_{t-2} + Y_{t-3}) + \alpha_2(Y_{t-2} - 2Y_{t-3} + Y_{t-4}) + e_t + \mu \quad . \quad (4.1)$$

$$Y_t = (2 + \alpha_1)Y_{t-1} + (\alpha_2 - 2\alpha_1 - 1)Y_{t-2} + (\alpha_1 - 2\alpha_2)Y_{t-3} + \alpha_2 Y_{t-4} + e_t + \mu. \quad (4.2)$$

$$Y_t = 1.0647Y_{t-1} + 0.2285Y_{t-2} + 0.3489Y_{t-3} - 0.6421Y_{t-4} + 51.5323 + e_t \quad . \quad (4.3)$$

This indicates that the adequate fitted model is a combination of previous enrolment values and a constant.

4.8 Forecasting

We also make forecast using the best fitted model for the next three years. Below is the graph of the forecasts.

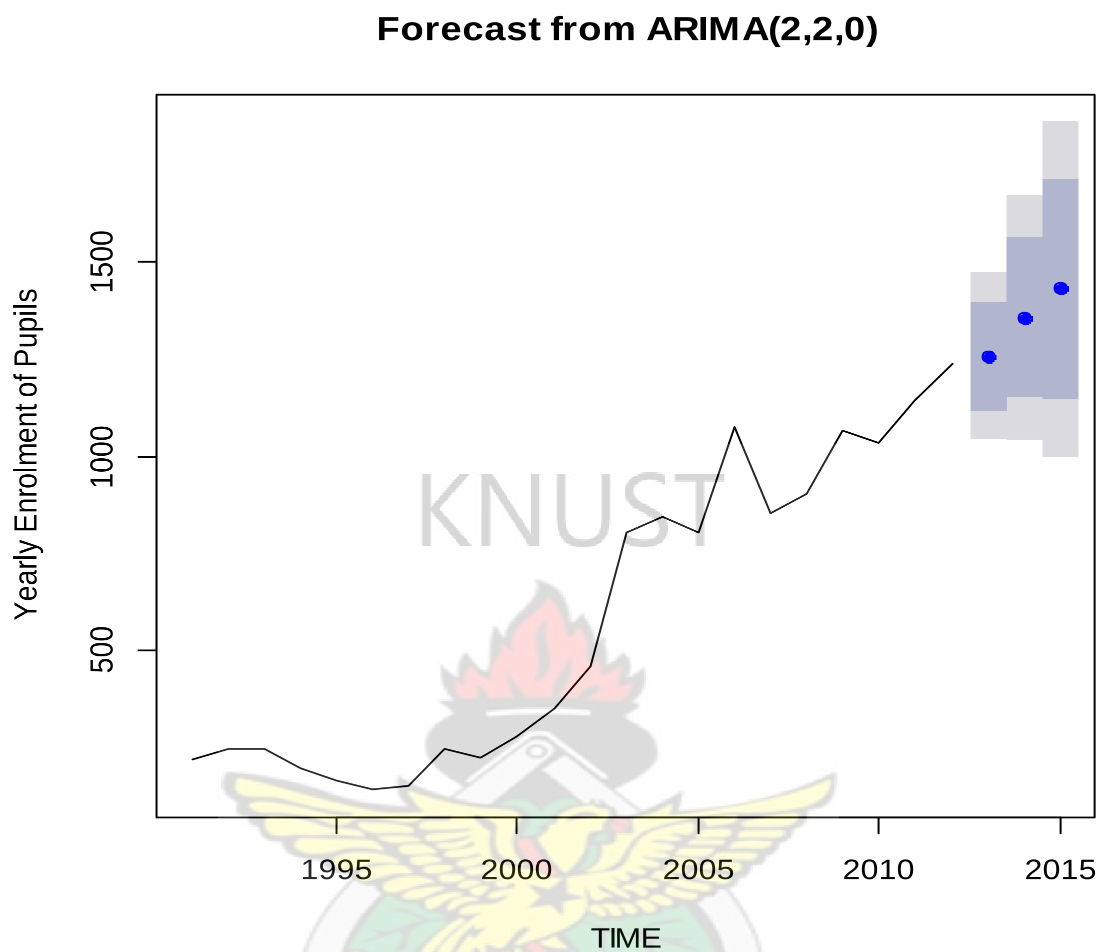

Figure 4.8: The forecasted enrolment figures of pupils are shown by the blue line, whilst the shaded areas show 80% and 95% prediction intervals respectively.

The forecasted values and standard errors are given in Table 4.4.

Table 4.4: Forecasted Values for ARIMA (2, 2,0)

YEAR	PREDICTED ENROLMENT FIGURE	STANDARD ERROR
2013	1257	110
2014	1357	161
2015	1430	219

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.0 Introduction

This chapter deals with the summary of findings concerning the enrolment of pupils in the public second cycle schools in Assin South District and recommendations for Ministry of Education and future researchers.

5.1 Summary of findings from the study

The main purpose of the study was to develop a model that can be used to predict future enrolments in the public second cycle schools in Assin South District. The yearly enrolment figures from 1991 to 2012 were used to develop the model. The mean number of pupils enrolled yearly from 1991 to 2012 was found to be 577, with standard deviation 391.932. From the study, the enrolment increased throughout the entire period, but somehow stable from 1991 to 2000. After the year 2000, enrolment increased sharply from 2001 to 2003 and thereafter the trend of enrolment behaved irregularly.

The maximum and minimum enrolments occurred in 2012 and 1996, recorded totals of 1239 and 141 pupils respectively. It was revealed from the study that the series could best be fitted with the model:

$$Y_t = 1.0647Y_{t-1} + 0.2285Y_{t-2} + 0.3489Y_{t-3} - 0.6421Y_{t-4} + 51.5323 + e_t$$

(5.1)

which gave predicted enrolment figures for 2013, 2014 and 2015 as 1257, 1357 and 1430 respectively, with the residuals, e_t being white noise. The model is a combination of previous enrolment values and a constant.

5.2 Conclusions

Generally, it was observed from the study that enrolment in the public second cycle schools in Assin South District had been increasing, even though it was somehow stable from 1991 to 2000. The senior secondary school system under the 1987 New Educational Reform was started in 1991 and since that time, there had been an upward trend in enrolment in the district.

5.3 Recommendations for Ministry of Education

After the introduction of the senior secondary school system, enrolment of pupils in the public second cycle schools in Assin South District had been increasing. It was observed from the study that the enrolment will increase in 2013, 2014 and 2015. There must be expansion in infrastructure (classrooms, dormitories, teachers' bungalows, library, computer laboratories, etc). Also, the stakeholders in education should provide logistics like vehicles, textbooks, computers and other materials needed in educational sector in the district. More teachers should be trained in the country and some should be posted to the public second cycle schools in Assin South District.

REFERENCES

Abraham, B. and Ledolter, J. (1983). *Statistical Methods for Forecasting,* New York Wiley.

Akaike, M. (1974). A new look at the statistical model identification. *I E EE Transactions in Automatic Control,* 716 – 23.

Ampadu, F. O., Mohammed, R. (2006). Trends in Education and School Management, Teacher Education Division, Ghana Education Service, Edo Printing Press, Accra.

Automatic Forecasting Systems (1986). *AUTOBJ. The Users Guide*, Hatboro, PA: Automatic Forecasting Systems Inc.

Box, G. E. P and Jenkins, G. M. (1976). *Time Series Analysis, Forecasting and Control,* San Francisco, CA: Holden-Day.

Box, G. E. P and Pierce, D. A. (1970). Distribution of the residual autocorrelations in autoregressive-integrated-moving-average time series models, *Journal of the American Statistical Association,* 65, 1509 – 26.

Box, G. E. P., Jenkins G. M., and Reinsel G. C.,(1994).Time Series Analysis: Forecasting and Control, third edition. Prentice Hall, Upper Saddle River, New Jersey.

Brandon, C., Jarrett, J. and Khumuwala, S. (1983). Revising forecasts of accounting earnings: a comparison with the Box-Jenkins method, *Management Science*, 29(2), February.

CCE – University of Cape Coast (2001). Philosophical and Social Foundations of Education. Cape Coast

Chatfield C., (1996). The Analysis of time series: an introduction, fifth edition, London: Chapman and Hall.

Cholette, P. A. (1983). ARIMA forecasting of seasonally adjusted series versus unadjusted series, presented at the annual meeting of the International Society of Forecasters, June 1983.

Cleary, J. P. and Levenbach, H. (1982). *The Professional Forecaster: The Forecasting Process Through Data Analysis.* Belmont, CA: Lifetime Learning Publications chs 17 – 24.

Cryer, J. D. (1986). *Time Series Analysis*. Boston, MA: Duxbury, chs 6, 7, 8, 9, 10.

Nsia D., (2007). Time Series Analysis and Forecasting Methods, Lecture Notes, Kwame Nkrumah University of Science and Technology, Kumasi.

Federal Reserve Bulletin, various issues. Washington, DC: US Board of Governors of the Federal Reserve System.

Foster, G. (1977). Quarterly accounting data: time series properties and predictive ability results, *Accounting Review*, 52, January, 1 -21.

GES, (2011). Computerized Schools Selection and Placement System (CSSPS). (Online) Available from: http://www.ghanaschoolsonline.com (Accessed on 15th March, 2013).

Griffin, P. A. (1977). The Time Series behaviour of quarterly earnings: preliminary evidence, *Journal of Accounting Research,* Spring, 71 – 83.

Hill G. W. and Fildes, R. (1984). The accuracy of extrapolation methods, an automatic Box-Jenkins package, S I F T, *Journal of Forecasting,* 3, 319 – 23.

Hill, G. W. and Woodworth, D. (1980). Automatic Box-Jenkins forecasting, *Journal of the Operational Research Society,* 31, 413 – 22.

Hoff, J. C. (1983). *A practical Guide to Box-Jenkins Forecasting*, Belmont, CA : Lifetime Learning Press, chs 12 -17.

Jarrett, J. E. (1989). Forecasting monthly earnings per share-time series models, *OMEGA. The International Jounal of Management Science*, 17(1) 37 – 44.

Jenkins, G. M. (1979). *Practical Experiences with Modeling and Forecasting Time Series,*St Helier, Jersey: Gwilym Jenkins and Partners.

Jones, R. H. (1980). Maximum likelihood fitting of ARIMA models to time series with missing observations, *Technometrics*, 22, 389 – 96.

Kohn, R. and Ansley, C. (1986). Estimation, Prediction and Interpolation for ARIMA models with missing data, *Journal of the American Statistical Association,* 81, 751 – 61.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models, *Biometrika*, 65, 297 – 308.

Lorek, K. MacDonald, C. L. and Patz, D. H. (1976). A comparative examination of management forecasts and Box-Jenkins forecast of earnings, *Accounting Review*, 51, April, 321 – 30.

Lorek, K. and MacDonald, C. L. (1979). Predicting annual net earnings with quarterly earnings time series models. *Journal of Accounting Research,* 17 Spring 190 – 204.

Lutekepol, H. (1984). Comparison of criteria for estimating the order of a vector autoregressive process, *Journal of Time Series Analysis*, 6 35 – 52.

Makridakis, S. and Hibbon, M. (1983). The accuracy of extrapolative methods Presented at the annual meeting of the International Society of Forecasters, Philadelphia, PA, June 1983.

Marquardt, D. W. (1963). An algorithm for least least squares estimation of nonlinear parameters, *Journal of the Society for Industrial and Applied Mathematics,* 11, 431 – 41.

Makridakis, S. and Wheelwright, S. C. (1983). *Forecasting Methods and Applications,* New York : Wiley, chs 9, 10.

Mcleod, A. I. (1978). On the distribution of residual autocorrelations in Box-Jenkins models, *Journal of the Royal Statistical Society Series B,* 40(3) 296 – 302.

McWilliam, H. O. A., Kwamena-Poh, M. A. (1975). The Development of Education in Ghana. London Longman Group Limited.

Ministry of Education (2001). Education Studies book for Basic School Teachers.Winneba.

Nelson, C. R. (1973). *Applied Time Series Analysis for Managerial Forecasting,* San Francisco, CA : Holden-Day, chs 6, 7.

O'Donovan, T. M. (1983). Short *Term Forecasting : An Introduction to the Box-Jenkins Approach,* Chichester : Wiley chs 4, 5, 6.

Omane-Akuamoah J., Obeng-Ampadu C., Asamoah D., Baffoe-Bonnie B. D., Prah D. (2004). Principles and Practices of Education For UTTDBE Programme (By Distance), Teacher Education Division, Ghana Education Service, Accra.

Pack, D. J. (1977). *A Computer Program for the Analysis of Time Series Models using the Box-Jenkins Philosophy.* Hatboro, PA: Automatic Forecasting Systems.

Pankratz, A. (1983). *Forecasting with Univariate Box-Jenkins Models*, New York Wiley.

Robert A. Yaffee., and McGee M., (2000). Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS. Academic Press Inc., San Dieg.

Shumway R. H., and Stoffer D. S., (2006). Time Series Analysis and its Applications (with R Examples), second edition. New Yorke: Springer.

Survey of Current Business, various issues, Washington, DC: Bureau of Economic Analysis, US Department of Commerce.

Texter, P. and Ord, K. (1989). Forecasting using automatic identification procedures: a comparative analysis, *International Journal of For*

APPENDIX

Table 1: Yearly enrolment of pupils in public second cycle schools in Assin South District, from 1991 to 2012.

TIME(YEARS)	YEAR OF ADMISSION	NUMBER OF PUPILS ENROLLED
1	1991	221
2	1992	246
3	1993	248
4	1994	197
5	1995	167
6	1996	141
7	1997	150
8	1998	247
9	1999	226

10	2000	277
11	2001	351
12	2002	460
13	2003	804
14	2004	844
15	2005	803
16	2006	1075
17	2007	851
18	2008	904
19	2009	1065
20	2010	1034
21	2011	1143
22	2012	1239

Source: yearly enrolment of pupils in the public second cycle schools (from 1991 to 2012) obtained from four senior high schools in Assin South District.