KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

Autoregressive Integrated Moving Average (ARIMA) Intervention Analysis Model for the Major Crimes in Ghana.

(The case of the Eastern Region)

By

BOAKYE AGYEMANG (BSc. Economics and Statistics)

A Thesis submitted to the Department of Mathematics, Kwame Nkrumah University of Science and Technology in partial fulfillment of the requirements for the degree of

MASTER OF PHILOSOPHY
(APPLIED MATHEMATICS)

Faculty of Physical Sciences, College of Science

May, 2012

DECLARATION

I hereby declare that this submission is my own original work towards the award of the M.Phil degree and that, to the best of my knowledge, it contains no material previously published by another person nor material which has been accepted for the award of any other degree of the University, except where due acknowledgement has been made in the text.

Boakye Agyemang	…………………….	……..……..…………
(Index No.: PG5070110)	Signature	Date

Certified by:

Mr. F. K. Darkwah	…………………….	……..……..…………
(Supervisor)	Signature	Date

Certified by:

Mr. F.K. Darkwah	…………………….	……..……..…………
(Head of Dept.)	Signature	Date

ACKNOWLEDGMENT

I wish to first and foremost express my profound appreciation and gratitude to the Almighty God for his immeasurable love, inexhaustible grace and divine protection and provision throughout the course of my studies in the University.

I also wish to express my deepest appreciation to my supervisor Mr. F. K. Darkwah who also doubles as Head, Mathematics Department KNUST for meticulously reading and supervising and making valuable contributions in the realization of the success of this work.

Again, I specially wish to thank the Management and members of the Staff Development and Scholarship Committee of Koforidua Polytechnic for their advice, and financial support. In the same vein, I wish to make mention of Prof. (Ing.) Reynolds Okai-Rector, Mrs. Mary Abena Agyepong-Registrar, Mr. J. M. Yeboah-Finance Officer, Mr. G. K. Abledu-Dean School of Applied Science and Technology and Mr. Okoampa Edwin-Former Head Applied Mathematics Department, all of Koforidua Polytechnic.

Furthermore, I thank all lecturers in the Mathematics Department, KNUST for the immense academic knowledge impacted especially during the course work sessions.

I also express my deepest gratitude to the Eastern Regional Police Command of the Ghana Police Service especially DSP. Yaw Ohene Takyi-Deputy Regional Crime Officer, Inspector Adjei and all Officers in the Regional CID unit for their assistance in getting access to the needed data for this research work.

Finally, to my wonderful family in the persons of my lovely wife Stella, my sweet daughter Michelle and my big sister Gifty for being solidly behind me in this journey that is worth undertaken.

DEDICATION

I graciously and dutifully dedicate this piece of work to the Almighty God who by His grace and mercies endowed me with knowledge and strength to undertake this academic journey successfully.

I again respectfully dedicate it to my lovely daughter Michelle and my niece Fosu Hemaa.

ABSTRACT

This research work sought to estimate and assess the nature of the impact of the establishment and operations of the Community Policing Unit of the Ghana Police Service as a major intervention in crime prevention and control in Ghana using ARIMA intervention analysis.

Using time series data (total monthly crime for the five major crime categories) covering the period 2000 to 2011 obtained from the Regional Crime Unit of the Eastern Regional Command of the Ghana Police Service, an impact assessment model was obtained. Empirical results from the study indicate that the pre-intervention period could best be modeled with an AR(1) error process based on which the full intervention model was obtained. The intervention event actually had an effect of reducing crime in the region by an approximate monthly reduction of 16 cases, which however was found to be abrupt but temporal and statistically significant. Its corresponding long term effect was also found to be approximately a reduction of 17 which is almost the same as the intervention effect. However, statistically insignificant rate of decay (δ) of 0.0406 resulted in the temporal nature of the duration of the effect of the intervention.

The results concluded that the overall intervention model was statistically significant based on the hypothesis tests by means of the Portomanteau test (Ljung-Box and Pierce), coupled with the analysis of the residual plots as well as the penalty statistics (as in AICs, AICcs and BICs) based on the principle of parsimony. Appropriate recommendations have been made based on the conclusions from the findings for consideration.

TABLE OF CONTENT

List of Tables

List of Figures

List of Acronyms

CHAPTER 1

INTRODUCTION

1.3 BACKGROUND TO THE STUDY

Ghana is a West African country bordering on the Gulf of Guinea and bounded by Côte d'Ivoire to the west, Burkina Faso to the north, Togo to the east, and the Atlantic Ocean to the south. It compares in size to Oregon, and its largest river is the Volta.

Called the Gold Coast, the area was first seen by Portuguese traders in 1470. They were followed by the English (1553), the Dutch (1595), and the Swedes (1640). British rule over the Gold Coast began in 1820, but it was not until after quelling the severe resistance of the Ashanti in 1901 that it was firmly established. British Togoland, formerly a colony of Germany, was incorporated into Ghana by referendum in 1956. Created as an independent country on March 6, 1957, Ghana, as the result of a plebiscite, became a republic on July 1, 1960.

Currently, Ghana has a Land area of 88,811 sq mi (230,020 sq km); total area of 92,456 sq mi (239,460 sq km) and total Population of 24,339,838 (2010 est.). Additionally, it has growth rate 1.8%; birth rate of 28.0/1000; infant mortality rate of 49.9/1000; life expectancy of 60.5; and density per sq km: 101 (infoplease, 2011).

Ghana is currently governed by constitutionally democratic elected government (National Democratic Congress) led by President John Evans Atta Mills since 2009 with a four year mandate.

Ghana has being one of the peaceful countries within sub-Saharan Africa. The serene atmosphere of the country and friendly disposition of Ghanaians offer visitors the necessary motivation to visit Ghana. For this reason, Ghana has earned global recognition and respect, and viewed as the gate-way to Africa. However, against the backdrop of a steadily improving global recognition, is the emergence of a new trend of **crimes** which are slowly but surely gaining root in the country.

The word **crime**, from the root of Latin cernō means "I decide, I give judgment". Originally the Latin word (aside from its lack of nasal aperatus) crīmen meant "charge" or "cry of distress". The Ancient Greek word krima (κρίμα), from which the Latin cognate was derived, typically referred to an intellectual mistake or an offense against the community, rather than a private or moral wrong. Individuals, human societies may each define crime and crimes differently, in different localities (state, local, international), at different time stages of the so-called "crime" (planning, disclosure, supposedly intended, supposedly prepared, incompleted, completed or futuristically proclaimed after the "crime") (Wikipedia, 2011).

Furthermore, the Criminal Code of Ghana (Act 2960), defines crime to include both the act, or actusriea and the intent to commit the act, or mensrea. Fafa, (2010) defined **crime** as the breaking of rule(s) or regulation(s) for which a governing authority (via mechanisms such as legal systems) can ultimately prescribe a conviction. Another definition of crime is a deviant behavior that violates prevailing norms or cultural standards prescribing how human beings behave in a society.

Therefore, based on the above definitions, crime can simply be defined as the breaking of rules or regulations or a deviant behavior that violates prescribed norms or values which is frowned upon by society.

Several forms of crime occur in societies which may be described as major or minor by virtue of their nature and impact. These are categorized and defined by Levitt (1996), as follows:

Motor Vehicle Theft - The theft or attempted theft of a motor vehicle.

Burglary - The unlawful entry of a structure to commit a felony or a theft. Attempted forcible entry is included.

Robbery - The taking or attempting to take anything of value from the care, custody, or control of a person or persons by force, or threat of force, or violence, and/or by putting the victim in fear.

Larceny - The unlawful taking of property from possession of another. Examples are thefts of bicycles or automobile accessories, shoplifting, pocket-picking, or the stealing of any property or article which is not taken by force and violence or by fraud. Attempted larcenies are included. Embezzlement, “con” games, forgery, and worthless checks are excluded.

Aggravated Assault - An unlawful attack by one person upon another for the purpose of inflicting severe or aggravated bodily injury. This type of assault usually is accompanied by the use of a weapon or by means likely to produce death or great bodily harm. Simple assaults are excluded (Levitt, 1996).

Murder - The unlawful killing of one human by another, especially with premeditated malice. Source: The American Heritage Dictionary of the English Language, 2000, Fourth Edition, Houghton Mifflin Company.

Forcible Rape - The carnal knowledge of an individual against his or her will. Included are rapes by force and attempts or assaults to rape.

These classifications are also consistent with the Criminal Code of Ghana, that is Act 2960.

According to Act 2960, five degrees of offenses are recognized in Ghana. Capital offenses, for which the maximum penalty is death by hanging, include murder, treason, and piracy. First-degree felonies punishable by life imprisonment are limited to manslaughter, rape, and mutiny. Second-degree felonies, punishable by ten years' imprisonment, include intentional and unlawful harm to persons, perjury, and robbery. Misdemeanors, punishable by various terms of imprisonment, include assault, theft, unlawful assembly, official corruption, and public nuisances. Increased penalties apply to individuals with a prior criminal record. Corporal punishment is not permitted. Punishments for juveniles are subject to two restrictions: no death sentence may be passed against a juvenile, and no juvenile under age seventeen may be imprisoned. Regulations and laws such as these are not applied equitably. Indeed, defendants habitually resort to one or another measure to avoid or ameliorate punishment.

All crimes either major or minor have negative effects both on individuals and societies or the nation at large. These negative effects range from destruction and lost of property, lost of innocent lives, fear and panic, security threat, budget constraints, and many more. These crimes may take place in the homes, at the work places, on the streets at public gatherings, others and range from robbery, thieving or stealing, car snatching, serial killing, cyber fraud, rape, domestic violence, illegal small scale mining (popularly known as galamsey), deforestation, child abuse, narcotics trade, physical assaults and damage to property, just to mention a few.

Crime reports have become rampant in the media and the country as a whole.

Some specific crime-related issues and cases that have caused controversies in our media and society in recent times are the Johnson Kombian who broke jail on two occasions and has been reported to have killed two policemen, the murder of an American-British missionary who was kidnapped and killed by Ghanaians in Koforidua in the Eastern Region, the alleged mass rape story at Kintampo that has caused a lot of uproar among the general public, serial killing of women, highway robberies, vehicle snatching, illicit drug and child trafficking and many more.

The Ghana Police Service in pursuance of their core mandate as well as other statutory government institutions such as the Ghana Statistical Service (GSS) and Judiciary Service collects and compiles statistics of crimes that occur in the country.

Crime statistics here refers to statistical measures of the crime in societies or simply crime data. Given that crime is secretive by nature, measurements of it are likely to be inaccurate. The above statistical data and media reportage provides convincing evidence that Ghana is not immune to crimes, especially in her vision to ensure a peaceful and safe environment to facilitate economic and social activities as a pre-requisite for making Ghana a Gateway to West Africa.

Appiahene-Gyamfi (1998**),** conducted a study at the School of Criminology Simon Fraser University Burnaby, on the topic Violent crime in Ghana: The case of robbery. However, the study was limited to the trends and patterns of robbery, and reactions to it in contemporary Ghana between 1982 and 1993. This study thus attempts to examine and analyze the five major crime statistics in Ghana namely murder, rape, defilement, robbery and the use and possession of drugs(cocaine, heroine, Indian hemp) from the year 2000 to 2011, the case of the Eastern Region with regard to trends, policies and interventions as well as their effects. In order to assess the impact of this event, ARIMA intervention analysis is applied in this study to evaluate the pattern and duration of its effects.

ARIMA is the acronym for autoregressive integrated moving average. In statistics and econometrics, and in particular in time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. These models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting).

Intervention analysis or event study is also a form of dynamic regression used to assess the impact of a special event on the time series of interest. The main focus is to estimate the dynamic effect on the mean level of the series, but other effects can also be considered. “Given a known intervention, is there evidence that change in the series of the kind expected actually occurred, and, if so, what can be said of the nature and magnitude of the change?” (Box and Tiao, 1975). Several interventions have been put in place in preventing crimes in Ghana. Some were point or discrete whilst others were continuous or permanent interventions. Some specific interventions put in place in Ghana in recent times to combat crime include the following: **Police partner transport owners in crime combat, Xmas Crime Combat, Government Ordering Assemblies to Name Streets, Ghana, Togo Police join forces to combat crime, Police launched 'operation calm life', Crime educations, and the establishment of Community Policing Unit of the Ghana Police Service** which is the main intervention being evaluated in this study.

The Community Policing Unit of the Ghana Police Service was established in June 2002 with the setting up of the administration section and the Bicycle Patrol Unit. This came as a result of the realization that there is the need for the police to collaborate with members of the communities, stakeholders, chiefs and opinion leaders in dealing with crime through effective communication. In this vein, the unit started with the bicycle patrols which was made possible when the International Criminal Investigation Training Assistance Programme (ICTIAP) under the United States Department of Justice in conjunction with International Police Mountain Bikes Association helped the Ghana Police Service in training a number of police bicycle patrols.

Finally, Community Policing Unit is an emerging concept in policing which seeks to bridge the communication and interaction gap between police institutions and the communities that they serve. It aims at encouraging the establishment of a close relation with civil societies in order to give the police an opportunity to understand and appreciate security needs and concerns of the various societies in which they operate. This method of policing leads to a situation where the police can work in partnership with local people to identify potential problems and take proactive steps in responding to them.

1.4 PROBLEM STATEMENT

The crime rates in Ghana have been very low in the past especially between the 1990,s and 2000 even compared to industrialized countries. The rate for all index offenses combined was 461.28 for Ghana, compared with 1709.88 for Japan and 4123.97 for USA (INTERPOL data for Ghana). More so, between 1996 and 2000, according to INTERPOL data, the rate of total index offenses increased from 416.32 to 461.28, an increase of 10.8%.

However, in recent times, there has being the emergence of a new trend of crimes which are slowly gaining roots in the country: serial killings, armed robbery, and cocaine trafficking. Crime reports have become rampant in the media in most parts of the country to the extent that a week or month would not end without an occurrence of crime in parts of the country especially in our major cities and on our roads.

The total crime in the Eastern region of Ghana for 2005 was 22,235 cases and that for 2006 was 23,476 cases representing 5.6% increase, which further rose from 23,476 to 23,607 in 2007 also representing 0.6% increase. This could pose serious challenges to the region and also to the entire nation if the trend of crime cases remain as it is and therefore need to be investigated statistically in order to make informed and intelligent decisions on the basis of such analysis.

Finally, the above crime cases recorded have had negative impact both on individuals, societies and subsequently on the nation at large. These negative effects ranged from destruction and lost of property, lost of innocent lives, fear and panic, security threat, budget constraints, and many more, a clear example being the murder of an American-British missionary who happens to be an investor in Ghana who was kidnapped and killed by Ghanaians in Koforidua in the Eastern Region.

1.3 OBJECTIVES OF THE STUDY

i. To model crime rate in Eastern Region of Ghana using time series analysis.

ii. To analyze the impact of the interventions in combating crime in the Eastern Region of Ghana.

iii. To evaluate the pattern and duration of the effect of the intervention in the region.

1.4 METHODOLOGY OF THE STUDY

A stochastic, time-series ARIMA model was used to analyze the dynamics of changes, variations and interruptions in crimes of Ghana through time-series data. This ARIMA model can help to perceive whether the policy interventions impact the crime rate in Ghana. The study will mainly use descriptive statistical data derived from official police records of the Ghana Police Service, Criminal Investigation Department of the Eastern Regional Command spanning from year 2000 to 2010. R and MINITAB software packages were used for the analysis. Additional information were solicited in both printed (hard copy) and electronic (soft copy) forms from the KNUST Library Kumasi specifically from the specifically from the E-Resources Center, British Council Library Kumasi, Eastern Regional Library Koforidua, Google and yahoo search engines.

1.5 JUSTIFICATION OF THE STUDY

i. The study would serve as a guide to politicians and stakeholders in making informed and intelligent policy decisions with regard to the management of crime rates especially in the Eastern part of the country that would spearhead the developmental agenda in terms of peace and total tranquility.

ii. Also, the information that would be unraveled from this study when put into effective and better utilization in the management of crime would have the economic benefit of saving cost to the region and the nation at large.

iii. The Regional Police would benefit greatly from accurate forecasts of crime within small geographic area such as in the region. Then it would be possible to target patrols to areas with forecasted crime increases.

1.6 ORGANIZATION OF THE STUDY

This thesis consists of five chapters. In this chapter, the introduction and background to the study and the objectives of this thesis have been discussed as well as and the structure of the thesis presented.

The next chapter which is the second chapter has dealt with the review of relevant literature showing the work done previously in the area of crime and its related issues, the use and application of time series intervention analysis and ARIMA models with supported relevant references.

Chapter three deals with the methodology and has discussed the ARIMA model with intervention analysis in terms of the equations involved and the procedures of model estimation.

The empirical results and analysis of model estimation using both the ARIMA model and the ARIMA-Intervention model is then explained in the fourth chapter, in which a comparative study is further conducted to assess whether the ARIMA-Intervention model is able to account for the impact of crimes in Ghana compared to the ARIMA model.

In chapter five which happens to be the last chapter, major findings are summarized and highlighted, followed by the conclusions and recommendations.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

This chapter of the research primarily deals with the review of empirical related literature of previous authors regarding time series analysis, crime, autoregressive integrated moving average (ARIMA) models and in particular intervention analysis. The literature review will basically be based on methods, data, findings and conclusions of the related subject matter.

2.2 EMPIRICAL LITERATURE REVIEW OF RELATED RESEARCH

Firstly, Andam-Akorful, e tal (2010) aimed at developing a prototype geospatial suite for the Service to help in crime mapping, analysis and monitoring as well as police asset management. To achieve this aim, the operation of the GP has been studied and an information system that suits the Ghanaian environment developed. A GIS developed with the Map Window GIS open source software allows a user to plot crime locations, perform statistical calculations and network analysis, and monitor in real time police personnel on patrol. Based on the GIS developed, it is therefore recommended that the security agency use it in their quest to combat crime in Ghana.

Secondly, Lawson (2008) stated carefully based on practical perspective that properly designed Community measures or early intervention programmes are most cost effective route to prevention than imprisonment, although it may be said that prison has an important role to play in protecting the community against the most dangerous and recalcitrant offenders and in punishing the most serious crimes, research evidence indicates the many disadvantages of over using imprisonment. Firstly, it can harm the chances of young people to have to make amends and fulfill their potentials as citizens. Secondly, by definition, prison limits the opportunities for young people to contribute to civic society and democratic life. Thirdly, most young people who are sent to prison, leave it no better equipped to fit into society than when they entered it, and indeed some young offenders leave prison a good deal worse off.

According to Fafa (2010) the rate of crime needs to be seriously checked and dealt with because this issue is creating a lot of panic and fear in the country since Ghana is known for its peace and as a nation, we must help maintain this. The government should empower the police, military and other security agencies to help curb this problem because we as Ghanaians are now living in a period of great fear and panic.

Fletcher (2010) attempted to bridge the gap in knowledge in the area of the effectiveness of CCTV in reducing antisocial behaviour with the use of both primary and secondary sources. The results were startling, showing CCTV to have little impact on the level of antisocial behaviour in one area and the opposite in another area. The interview with the Blackpool CCTV Unit and research has highlighted some interesting factors as to why these results may have occurred.

Also, Grove (2010), made two significant contributions to the advancement of knowledge within crime prevention through a comparison of systematic review and scientific realist evaluation methods for crime prevention. The first of these is to evaluate the success of repeat victimization prevention interventions. Interventions across four crime types are assessed herein, and the context-mechanisms-outcome configurations examined. The second contribution of their thesis assessed two techniques of meta-evaluation: systematic reviews and realist syntheses. Repeat victimization prevention is revealed as an effective way of reducing crime, with a need for further research to apply the principle across further crime types. A requirement is identified for a greater breadth and depth of information to be included in future crime prevention evaluations. The systematic review is shown to be a useful way of assessing the overall effectiveness of the interventions, whilst the realist synthesis fills in the detail of why some interventions work and others fail. It is concluded that both approaches to meta-evaluation have useful contributions to make, and that a 'third way' incorporating the best elements from each method should be developed.

Moreover, Appiahene-Gyamfi (2002) utilized Paul and Patricia Brantingham's pattern theory to examine the factors that accounted for the broad crime trends and patterns in Ghana. The picture that emerged from the analysis of 1980–1996 police data showed discernable patterns of crime at national and regional levels of aggregation. Crime was highest in the more developed and densely populated regions of southern Ghana. Indeed, crime rates increased from northern to southern Ghana, with a heavy concentration in Ashanti, the most populous region. Overall, crime was highest in Greater Accra except murder, which was highest in Ashanti. Assault was the highest recorded offense, followed by theft.

Amoakohene (2004) examined violence against women in Ghana and how it affects and is perceived by them. Women's perceptions of their rights, responsibilities, duties and abuses or violations are evaluated using open-ended qualitative questions in two major cities in Ghana: Accra and Kumasi. Policy responses to domestic violence are then examined by first reviewing what provisions exist in the country's constitution to address the problem and then the specific steps the government itself has taken. Civil society's response in the form of activities by non-governmental organizations is also reviewed. Finally, the effects of domestic violence on women's health and well-being are examined and suggestions for addressing the problem are made.

Furthermore, Otchere, (2007) expressed in his article titled ‘‘increased crime rate in Ghana: a call for security reform’’ that Ghana is one of the peaceful countries within sub-Saharan Africa. However, against the backdrop of a steadily improving global recognition, is the emergence of a new trend of crimes which are slowly but surely gaining root in the country: serial killings, armed robbery, and cocaine trafficking. These social vices have created a negative impact on the good name of Ghana. On the social front, there is inflation in the country and standard of living has risen. Under such conditions, the temptation for people to engage in inelegant means of livelihood is high because the good old adage states that, “an idle hand is a devil’s tool.” Therefore the spate of armed robbery, cell phone snatching, and countless minor incidents are rooted to unemployment.

The Ghana Police Administration on Tuesday, November 3, 2009 observed that the general Crime Statistics across the country has reduced by 8 percent from 1,150 cases between January and September last year to 1,054 cases the same period this year. According to the Service, the decrease is as a result of renewed and consistent effort at fighting organised crime in the country. The Greater Accra Region tops the country in terms of general crime numbers of 392. Ashanti Region however tops with murder cases of 44 so far this year. Armed Robbery, illicit Narcotics Trade, Visa Fraud are the key crime types in Ghana.

Huang (2011) focused on identifying the patterns of the non-gun related crimes in the 21 areas of Los Angeles based on exploratory data analysis, principle component analysis, cluster analysis, as well as Pearson's X^2 statistics to discover unusual crimes in areas with the following conclusions: The percentages of total crime in 21areas are almost the same but the percentages of a specific crime type in 21 areas differ a lot, and each crime type has its own pattern. BTFV is the crime type that happens most. HOM, ARSON and KID were rare. The distributions of 13 crime types' frequency vary according to areas. Most crimes are most frequent in January and less frequent in February, from PCA, part of the variance (up to 50%) of frequency of 13 crime types in 60 months can be well explained by 1 or 2 PCs, according to the area.

Again, Appiahene-Gyamfi (1998) studied and discussed the trends and patterns of robbery, and reactions to it in contemporary Ghana between 1982 and 1993. The study contends that robbery as a crime of opportunity appears to have been prevalent in pre-colonial times as well as during the subsequent period of slavery. Its trends and patterns however, have changed with the introduction of a monetary economy that has resulted in increased opportunities and targets for robbery. The descriptive statistical data derived from official police records concluded that even though the incidence and volume of robbery in Ghana is quantitatively small compared to the rates of other index offenses, and minuscule within the population at large, official reaction to it has been rather swift and merciless. No reason can be assigned to the executions other than deterrence, which raises questions as to its efficacy.

Fernandez (2005), studied crime prevention and the perception of safety in campus design focusing on the outdoor environment on a college campus. The criteria for a safe design was developed from research gathered on crime prevention and the psychological reactions of users to exterior site features as well as crimes reported on the LSU campus were compiled on a crime map and analyzed to determine whether student perceptions of unsafe and safe areas were justified. The results exposed a perceived lack of safety among users in certain areas, the evaluations of both perceived safe and unsafe areas on campus brought about a better understanding of how users see and interact in their surroundings. In order to design or improve an area many factors must be in place to make the area safe for users and deter crime while at the same time being perceived as safe by the users to the site.

Han (2009) conducted three empirical studies detecting the determinants of crime in England and Wales using time series analyses to look for cointegrating relationships between property crimes and unemployment as well as law enforcement instruments, employing panel data and corresponding techniques to control for area-specific fixed effects as well as the endogeneity of law enforcement variables and allowed crime rate to have spatial spillover effect, in other words, the crime rate in one area is affected by, in addition to its local crime-influential factors, the crime rates and crime-related factors in its neighbouring areas.

Donohue and Levitt (2011) offered evidence that legalized abortion has contributed significantly to recent crime reductions on their study titled The Impact of Legalized Abortion on Crime. Crime began to fall roughly 18 years after abortion legalization. The 5 states that allowed abortion in 1970 experienced declines earlier than the rest of the nation, which legalized in 1973 with Roe V. Wade. States with high abortion rates in the 1970s and 1980s experienced greater crime reductions in the 1990s. In high abortion states, only arrests of those born after abortion legalization fall relative to low abortion states. Legalized abortion appears to account for as much as 50 percent of the recent drop in crime.

Peng, etal (2008) used time series model of ARIMA to make short-term forecasting of property crime for one city of China. With the given data of property crime for 50 weeks, an ARIMA model is determined and the crime amount of 1 week ahead is predicted. The model fitted and forecast results were compared with the SES and HES. It showed that the ARIMA model had higher fitting and forecasting accuracy than exponential smoothing and therefore would be helpful for the local police stations and municipal governments in decision making and crime suppression.

Chen, etal (2008) applied an autoregressive integrated moving average (ARIMA) to make weekly and daily forecasting of property crime for a city of China. It is shown that the model of AR (1) is suitable for crime sample distributing by week and IMA (1, 1) by day. The mean absolute percentage error (MAPE) and magnitude relative error were taken as the error measurements for model fitting and forecasting. The results obtained proved that the model of AR (1) had higher accuracy in fitting and forecasting than BOA

(1, 1). This result could be attributed to the crime stochastic difference between day and week. When forecasting for day, the crime stochastic was strong, so it is hard to pick up the turning points. But for week, the stochastic of the crime was eliminated effectively. So, for short-term crime forecasting, it is better to make prediction for week than for day.

Mohamad, etal (2003) forecasted residential burglary in Kuala Lumpur. Compared to other crimes in Kuala Lumpur, residential burglary shows high number per year. The econometric and ARIMA model were constructed to develop the forecasting model, it supported by statistical Software to validate the forecasting model by using 2004 burglary data.

More so, Jennifer, etal (2010), applied Autoregressive Integrated Moving Average (ARIMA) with intervention model to evaluate the impact of different local, regional and global incidents of a man-made, natural and health character, in Taiwan over the last decade. The incidents used in this study are the Asian financial crisis starting in mid-1997, the September 21st earthquake in 1999, the September 11th terrorist attacks in 2001, and the outbreak of Severe Acute Respiratory Syndrome (SARS) in 2003. Empirical results revealed that the SARS illness had a significant impact, whereas the Asian economic crisis, the September 21st earthquake, and the September 11th terrorist attacks showed no significant effect on air movements.

Ledolter and Chan (1996) examined whether a significant change in fatal and major-injury accident rate can be detected following the implementation of a higher speed limit in the state of Iowa using intervention analysis. Findings from their study revealed a 20% increase in the number of state-wide fatal accidents to the speed limit change. The impact was largest on rural interstates, where the number of fatal accidents increased by 57%, implying two additional fatal accidents each quarter.

Chang, et al. (1993) proposed a multiplicative time-series model to capture the stochastic fatality pattern based on a long-term nationwide fatality data in the U.S.A, 'before' a speed limit change (January 1975 - March 1987). This long-term fatality model along with the 2-year 'after' information (April 1987 - December 1989) was then used to detect the possible impact pattern. The results of the intervention analyses indicated that the increased speed limit had significant initial impacts on highway fatalities as the nationwide level. Such impacts, however, decayed after about a 1-year ‘learning period’.

Zambon et al. (2007) assessed the effects of a demerit points system introduced in Italy in July 2003 on the prevalence of seat belt use (intermediate outcome) and the number of road traffic deaths and injuries (health outcomes) through the application of intervention analysis. The methodology was based on a pre and post-intervention regional observation study for seat belt investigation from April 2003 to October 2004, and a national time-series analysis of road traffic deaths and injuries between 1999 and 2004. The involved 19551 drivers, 19057 front passengers and 8123 rear passengers in the investigation into seat belt use, whereas 38154 fatalities and 1938550 injured subjects were examined for the time-series analysis. The findings revealed an increase in observed seat belt use of 51.8% among drivers, of 42.3% among front passengers and of 120.7% among rear passengers. It further revealed that an estimated number of 1545 deaths and 91772 injuries were prevented in 18 months after the introduction of the legislation.

Yaacob et al. (2011) investigated the effects of OPS Sikap on road accidents in Malaysia using intervention analysis by assessing the intervention effect in comparison with the standard ARIMA model, and hence to obtain the best model for forecasting purposes. The findings revealed that there was a drop in the number of road accidents during OPS Sikap II, VI, VIII, XII and XIV, but the significant reduction could only be seen after the implementation of OPS Sikap VII and XIV with an expected number of reductions by about 1,227 and 1,484 accidents associated with respective interventions.

Chung, et al. (2009) analyzed the impact of financial crisis on the manufacturing industry in china using data collected from March 2005 to November 2008 by the china statistical databases of the national bureau of statistics in china. The intervention effect of the global financial crisis that began in September 2008 on china’s manufacturing industry, as measured in this study, was temporary and abrupt. The results again indicated that china’s manufacturing industry may have to tolerate a significant negative effect cased by the global financial crisis over a period of time, with its gross industry output value declining throughout 2008 and 2009 before reaching a state of equilibrium. The study further compared the results of the ARIMA and ARIMA Intervention models, and concluded that the application of intervention analysis was appropriate for explaining the dynamics and impact of interruptions and changes of time-series in a more detailed and precise manner.

Shittu (2009), used the intervention analysis approach to model exchange rate in Nigeria in the presence of finance and political instability. Monthly exchange rate of Naira vis-à-vis US Dollar from 1970 to 2004 was used on some identified intervention variables. The result showed that most of the interventions are pulse function with gradual and linear but significant impact in the naira-dollar exchange rates.

Rock-Antoine and Hannarong (2002), investigated the impact of the North American Free Trade Agreement on both bilateral trade and income of each member country- US, Canada, and Mexico. They covered time series data before and after NAFTA was formed from 1980 to 1999. In the study, NAFTA was considered as a prolonged impulse function in international trade activities among the three trading partners by employing an intervention-function model. Findings from their study revealed that NAFTA increases bilateral trade between US-Canada and US-Mexico, and in terms of income, NAFTA benefits Canada the most “certainly”.

Abdalla (2006) measured the impact of Yazegi Company’s decision to deliver new kinds of soft drink, and the intervention impact of the Israeli constraints on sales delivery in Gaza. It was found that there was a significant impact on the company’s sales due to the company’s intervention (positive impact), and the Israeli intervention (negative impact). Also comparing the non-intervention model and the intervention model concluded that the forecasts of the intervention model were better than that of the non-intervention model, because it values were closer to the actual data and had smaller standard errors.

Min et al. (2010) used Autoregressive Integrated Moving Average (ARIMA) with integration model (also known as integration analysis) to evaluate the impact of different local, regional and global incidents of a man-made, natural and health character, in Taiwan over the last decade. The incidents used in this study are the Asian financial crisis starting in mid 1997, the September 21st earthquake in 1999, the September 11th terrorist attacks in 2001, and the outbreak of severe acute respiratory syndrome (SARS) in 2003.

Empirical results revealed that the SARS illness had a significant impact, whereas the Asian economic crisis, the September 21st earthquake and the September 11th terrorist attacks showed no significant effect on air movements.

Again, Min (2008b) assessed whether two events, the 21st September 1999 earthquake and the severe acute respiratory syndrome outbreak in 2003, had a temporary or long-term impact on the inbound tourism demand for Japan and further assessed whether intervention analysis produces better forecasts as compared with forecasts without intervention analysis. The results confirmed that the effect of both disasters on Japanese inbound tourism presented only temporarily, and the forecasting efficiency of ARIMA with intervention is superior to that of a model without intervention.

Girard (2000) used an ARIMA model with intervention analysis technique to analyze and assess the epidemiology situation of whooping-cough in England and Wales for the period 1940 to 1990. The ARIMA modeling of this illness contains intervention variables, such as the introduction of widespread vaccination in 1957 and the fall in the level of vaccination down to 30% in 1978. The results of the study confirmed the role of the intervention variables on the evolution of the morbidity due to whooping-cough, by quantifying their impact on the level of the morbidity, as well as the delay needed before they have an influence on the increase of recorded cases of whooping-cough.

Sathianandan (2006) used quarter-wise total marine fish landings in the two states namely Kerala and Karnataka during 1960-2000 for an impact study, by adopting two popular time series methods for intervention analysis. The first method is based on seasonal ARIMA modeling and the second is based on regression modeling with ARMA type errors. The analysis revealed that for Kerala the model found suitable is seasonal ARIMA type model and for Karnataka the feasible model was regression model with ARMA errors. Based on the final estimated intervention models, the effect of the interventions was estimated at 2.26 lakh tonnes and 88 thousand tones per annum respectively for Kerala and Karnataka.

Ismail, et al. (2009) assessed the impact of the first terrorist Bali's bomb that occurred on October 12th, 2002 to tourism industry in Indonesia by applying a pulse function of intervention focusing on the differential statistics that can be used to determine the order of intervention model. The study specifically focused on the derivation of some effect shapes, either temporary, gradually or permanent on the arrival of tourist into Bali. The final result from the study revealed a decreasing trend in tourist arrival in Bali, Indonesia.

Muhammad et al. (2010) attempted to model and explain the magnitude and periodic impacts of the Asian financial crisis since July 1997 and terrorist attacks referring to the Bali bombings on October 12th 2002 and October 1st 2005, respectively using monthly data comprising the number of tourist arrivals in Indonesia via Soekarno-Hatta airport are used as the data for this case study. The results showed that the Asian financial crisis and Bali bombings yield negative impacts on the number of tourist arrivals to Indonesia via Soekarno-Hatta airport. Also, the Asian financial crisis gave a negative permanent impact after seven month delay, whilst the first and second Bali bombings also yielded negative impacts which had temporary effect after six and twelve months delay respectively.

Bonham and Gangnes (1996) analyzed the effect on hotel revenues of the Hawaii room tax using time series intervention analysis. It specifies a time series model of revenue behavior that captures the long run co-integrating relationships among revenues and important income and relative price variables, as well as other short-run dynamic influences. This study estimates the effect on Hawaii hotel room revenues of the 5% Hawaii hotel room tax introduced in January 1987. This study concludes with no evidence of statistically significant tax impacts.

Gonzales, etal (1999) used univariate Box-Jenkins time-series analyses (ARIMA models), in modeling and forecasting future energy production and consumption in Asturias. Initially, each series was recorded monthly from 1980 to 1996. These data include trend and seasonal variations which allow the use of ARIMA univariate models for predictions of future behavioral patterns. The optimum forecasting models obtained for each energetic series, have a satisfactory degree of statistical validity (low approximation errors) and are suitable for use as reference inputs in a regional energetic plan for the period 1997-98.

Choueiki, etal (1997) investigated solving the short-term load forecasting (STLF) problem with artificial neural networks by conducting a fractional factorial experiment. The results were analyzed, and the factors and factor interactions that influence forecast errors were identified and quantified. From this analysis, rules were derived for building a "quasi optimal" neural network to solve the STLF problem. The quasi optimal neural network was compared to an automated Box-Jenkins seasonal ARIMA model and was found to perform better.

Abdel, etal (1997) developed autoregressive integrated moving average (ARIMA) models for the Eastern Province of Saudi Arabia using data for five (5) years and evaluated on forecasting new data for the sixth year. The optimum model derived is a multiplicative combination of seasonal and nonseasonal autoregressive parts, each being of the first order, following first differencing at both the seasonal and nonseasonal levels. Compared to regression and abductive network machine-learning models previously developed on the same data, ARIMA models require less data, have fewer coefficients, and are more accurate. The optimum ARIMA model forecasts monthly data for the evaluation year with an average percentage error of 3.8% compared to 8.1% and 5.6% for the best multiple-series regression and abductory induction mechanism (AIM) models, respectively; the mean-square forecasting error is reduced with the ARIMA model by factors of 3.2 and 1.6, respectively.

Vaziri (1997) studied the fluctuations of the Caspian Sea's mean monthly surface water level for the period of January 1986 to December 1993 using, Artificial neural network (ANN) and autoregressive integrated moving average (ARIMA) models. The results revealed that fluctuations of the Caspian Sea's surface water level have a tremendous economical, social, political, and environmental impact on the surrounding countries of Azerbaijan, Iran, Kazakhstan, Turkmenistan, and Russia. It was also found that the ANN and ARIMA prediction for the period of January to December 1993 were in good agreement with the recorded data. However, the monthly predictions for January to December 1994 reflected a continuing rise in the Caspian Sea water surface level that would have adverse impacts on the surrounding countries.

Montanari, etal (1997) applied fractionally differenced autoregressive integrated moving average (FARIMA) model to analyze hydrologic time series, specifically to the monthly and daily inflows of Lake Maggiore, Italy. In contrast to utilizing the traditional ARIMA models, this approach permits the modeling of both short- and long-term persistence in a time series.

Marshment, etal (1996) presented econometric modeling techniques for short-range intercity traffic forecasting. The Oklahoma Turnpike Authority required a modeling system that would produce forecasts of monthly volumes by vehicle type, be easily updated, allow scenario testing, and generate understandable results. An autoregressive integrated moving average (ARIMA) model and a regression model were compared using a structured comparative test. The ARIMA model identified patterns in time series data that are likely to recur, whereas the regression model used economic variables to predict changes in traffic volumes. The regression model outperformed the ARIMA model for all vehicle classes except passenger cars. The regression model was also more accurate at predicting past volumes because it included economic variables.

Shvartser, etal (1993) forecasted the hourly water demands with a model based on pattern recognition and time-series analysis is described. Three repeating segments, "rising," "oscillating," and "falling," make up the daily demand pattern. These segments were defined as successive states of a Markov process. Low-order auto-regressive integrated moving average (ARIMA) models were fitted to each segment. An hourly forecast for the next 24 hours or several days can then be produced with the ARIMA model. The forecast can be performed in real time, and the state of the system can be monitored continuously. A new demand forecast can be produced with updated information from detected deviations from the planned state. The model's process of development, application, and evaluation is demonstrated on a water system in Israel.

Ho and Xie (1997) investigated the approach to repairable system reliability forecasting based on the Autoregressive Integrated Moving Average (ARIMA). An illustrative example on a mechanical system failures was presented and comparison also made with the traditional Duane model. It is concluded that ARIMA model was a viable alternative that gives satisfactory results in terms of its predictive performance.

Hung, etal (1997) employed six methods (decomposition, Holt-Winters seasonal smoothing, univariate ARIMA, transfer function. intervention model, and vector ARIMA) of time series analyses in the modeling and forecasting of the quantities of cement demand in Taiwan in 1993, 1994, and 1995 using monthly data from January 1982 to December 1995. MAPE and MSE were used to evaluate performances of various models. The empirical evidence suggested the forecasting performances of all 6 models were less satisfactory in the year of 1994 (the year of turning point) than in the years of 1993 and 1995. For the one year forecast, the transfer function method, using the authorized square footage of building permits as an input variable, was the best model.

Bianchi, etal (1997) analyzed existing and improved methods for forecasting incoming calls to telemarketing centers for the purposes of planning and budgeting. They also analyzed the use of additive and Multiplicative versions of Holt-Winters exponentially weighted moving average models and compare it to Box-Jenkins (ARIMA) modeling with intervention analysis. They determine the forecasting accuracy of HW and ARIMA models for samples of telemarketing data and concluded that ARIMA models with intervention analysis performed better.

Sanjeev, etal (2003) examined the impact of abolished parole and reformed sentencing for all felony offenders committed on or after January 1, 1995 by the Commonwealth of Virginia, considering structural time series models as an alternative to the Box-Jenkins ARIMA models that form the standard time series approach to intervention analysis. The study revealed limited support for the deterrent impact of parole abolition and sentence reform is obtained using univariate modelling devices, even after including unemployment as an explanatory variable. Finally, the flexibility of structural time series models is illustrated by presenting a multivariate analysis that provides some additional evidence of the deterrent impact of the new legislation.

Polat (2007) generated a spatio-temporal crime prediction model by using time series forecasting with simple spatial disaggregation approach in Geographical Information Systems (GIS). The model is generated by utilizing crime data for the year 2003 in Bahçelievler and Merkez Çankaya police precincts. Methodology starts with obtaining clusters with different clustering algorithms and compared in terms of land-use and representation to select the most appropriate clustering algorithms. In order to predict crime in time dimension a time series model (ARIMA) is fitted for each week day, then the forecasted crime occurrences in time are disaggregated according to spatial crime cluster patterns. Actually, these results represent sensitive areas to crime. Also, the number of incidents predicted indicates the level of sensitivity. Higher number of incidents predicted means the area is more prone to criminal activities. The solid results of this study is to determine these areas and the level of influence. Police should utilize the model first by understanding the reason of clusters. Why the area covered by these clusters are attractive for offenders. This phase needs background information about the area. If area is known and identified in terms of land use, configuration of buildings, important organizations; it is possible to detect opportunities for crime in the crime triangle.

Rusco and Walls (2001) estimated the long-run impact of removing the export ban through the use of a time series intervention analysis. The results indicated that Alaskan crude oil prices increased between $0.98 and $1.30 on the West Coast spot market relative to prices of comparable crude oils as a result of removing the export ban. However, we find no evidence that West Coast prices for refined oil products regular unleaded gasoline, diesel fuel, and jet fuel increased as a result of lifting the ban.

Enders, etal (1990) employed intervention analysis to assess the effectiveness of four specific terrorist-thwarting policies undertaken between January 5, 1973 and April 15,1986. These policies included: (1) installation of metal detectors in airports, (2) enhanced security for U.S. embassies and personnel, (3) the legislation of the Reagan "get-tough" laws on terrorism, and (4) the U.S. retaliatory strike against Libya. The most successful policy involved metal detectors. Expenditures to secure U.S. embassies had the intended effect, but it also had the unintended effect of putting non-U.S. diplomats at somewhat greater risk. The Reagen get-tough laws were ineffective. Unfortunately, the Libyan raid had the unintended effect of increasing U.S. and U.K. attacks temporarily.

Finally, Ferrand, etal (2011) examined the impact of specific events, including branded-drug and generic entry, a black box warning, direct-to-consumer advertising

(DTCA), and new indication approval, on Medicaid spending on antidepressants. Using quarterly expenditure data for 1991-2005 from the national Medicaid pharmacy claims database maintained by the Centers for Medicare and Medicaid Services, a time-series autoregressive integrated moving average (ARIMA) intervention analysis was performed on 6 specific antidepressant drugs and on overall antidepressant spending. Twenty-nine potentially relevant interventions and their dates of occurrence were identified from the literature. Each was tested for an impact on the time series. Forecasts from the models were compared with a holdout sample of actual expenditure data. Interventions with significant impacts on Medicaid expenditures included the patent expiration of Prozac® (P<0.01) and the entry of generic paroxetine producers (P=0.04), which reduced expenditures on Prozac® and Paxil®, respectively, and the 1997 increase in DTCA (P=0.05), which increased spending on Wellbutrin®. Except for Paxil®, the ARIMA models had low prediction errors. Generic entry at the aggregate level did not lead to a reduction in overall expenditures (P>0.05), implying that the expanding market for antidepressants overwhelmed the effect of generic competition.

2.3 CONCLUDING REMARKS ON LITERATURE REVIEW

Based on the above literature reviewed by pervious authors it can be concluded that ARIMA models with and without interventions are applied not only to criminal analysis (criminology), but rather to a wide variety of applications ranging from archeology to zoology. The intervention analysis developed by Box and Tiao (1975) is found to be powerful statistical technique in evaluating or assessing the impact of events, policies or programs on time series data.

CHAPTER 3

METHODOLOGY

3.1 INTRODUCTION

This chapter attempts to examine thoroughly the basic plots, definitions and concepts of time series analysis, assumptions, conditions, principles and processes involved in the application of autoregressive integrated moving average (ARIMA) with intervention analysis developed by Box and Tiao, (1975), specifically applied in this research work.

3.2 BASIC CONCEPTS AND DEFINITIONS OF TIME SERIES

3.2.1 Basic definitions

Time series is defined as a collection of observations or measurements on quantitative variables made sequentially or in a uniform set of time period, usually daily, weekly, monthly, quarterly, annually, and so on and so forth. Examples include total monthly crime for a jurisdiction for a period of ten years, daily stock prices of a firm for a period of one year, monthly electricity consumption for a household for a period of five years, etc.

Time series analysis comprises methods or processes that break down a series into components and explainable portions that allows trends to be identified, estimates and forecasts to be made. Basically time series analysis attempts to understand the underlying context of the data points through the use of a model to forecast future values based on known past values. Such time series models include GARCH, TARCH, EGARCH, FIGARCH, CGARCH, ARIMA, etc but the main focus of this study is based on ARIMA model with intervention by Box-Jenkins and Tao, 1975 which is discussed subsequently pretty soon.

3.2.2 Time Series Graph

Time series plot is simply a graph which display observations on the y-axis against equally spaced time intervals on the x-axis. The time series plot specifically consists of:

Time scale (index, calendar, clock, or stamp column) on the x-axis; data scale on the y-axis; and lines displaying each time series as shown in the Figure 3.1 below for a given hypothetical data. The plots are usually used to: detect trends in your data over time; detect seasonality in your data; and compare trends across groups

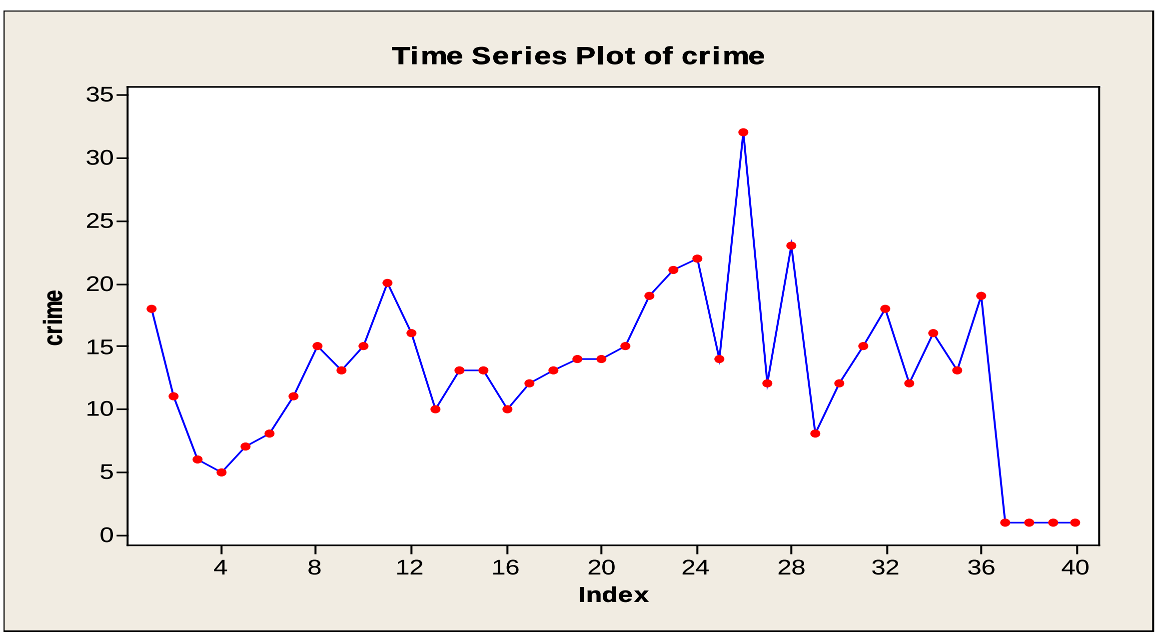

Figure 3.1 Time series plot for a hypothetical data of 40 observations

3.3 COMPONENTS OF TIME SERIES

A vital step in choosing appropriate modeling and forecasting procedure is to consider the type of data patterns exhibited from the time series graphs of the time plots. The sources of variation in terms of patterns in time series data are mostly classified into four main components. These components include seasonal variation; trend variation; cyclic changes; and the remaining “irregular” fluctuations.

3.3.1 The Trend (T)

The trend is simply the underlying long term behavior or pattern of the data or series. The Australian Bureau of Statistics (ABS, 2008) defined trend as the 'long term' movement in a time series without calendar related and irregular effects, and is a reflection of the underlying level. It is the result of influences such as population growth, price inflation and general economic changes. The following graph depicts a series in which there is an obvious upward trend over time:

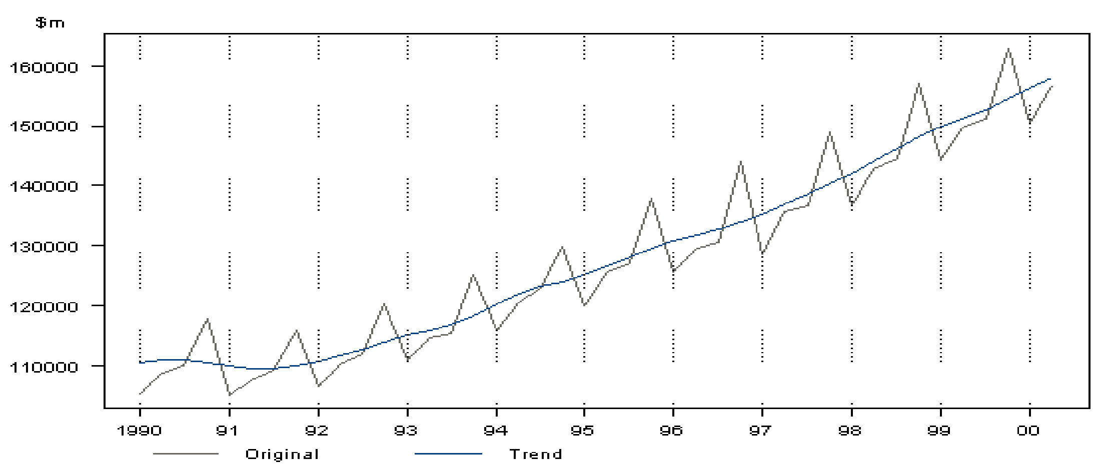

Figure 3.2: Upward trend graph of a hypothetical time series data

3.3.2 Seasonal variation(S)

A seasonal effect is a systematic and calendar related effect. Some examples include the sharp escalation in most Retail series which occurs around December in response to the Christmas period, or an increase in water consumption in summer due to warmer weather. Other seasonal effects include trading day effects (the number of working or trading days in a given month differs from year to year which will impact upon the level of activity in that month) and moving holidays (the timing of holidays such as Easter varies, so the effects of the holiday will be experienced in different periods each year).

Seasonal adjustment is the process of estimating and then removing from a time series influences that are systematic and calendar related. Observed data needs to be seasonally adjusted as seasonal effects can conceal both the true underlying movement in the series, as well as certain non-seasonal characteristics which may be of interest to analysts. Seasonality in a time series can be identified by regularly spaced peaks and troughs which have a consistent direction and approximately the same magnitude every year, relative to the trend as depicted in the figure 3.3 below.

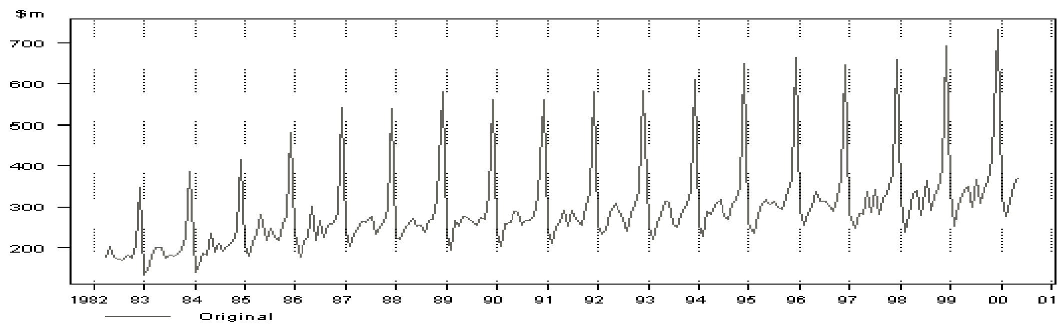

Figure 3.3 Graphical display of seasonal effect of a hypothetical data.

Other techniques that can be used in time series analysis to detect seasonality include:

1. A seasonal subseries plot is a specialized technique for showing seasonality.
2. Multiple box plots can be used as an alternative to the seasonal subseries plot to detect seasonality.
3. The autocorrelation plot can help identify seasonality.

3.3.3 Cyclical variations (C)

Cyclical variations are the short term fluctuations (rises and falls) that exist in the data that are not of a fixed period. They are usually due to unexpected or unpredictable events such as those associated with the business cycle sharp rise in inflation or stock price, etc. The main difference between the seasonal and cyclical variation is the fact that the former is of a constant length and recurs at regular intervals, while the latter varies in length. More so, the length of a cycle is averagely longer than that of seasonality with the magnitude of a cycle usually being more variable than that of seasonal variation.

3.3.4 Irregular variations (I)

The irregular component (sometimes also known as the residual) is what remains after the seasonal and trend components of a time series have been estimated and removed. It results from short term fluctuations in the series which are neither systematic nor predictable. In a highly irregular series, these fluctuations can dominate movements, which will mask the trend and seasonality. The Figure 3.4 below is a graph which is of a highly irregular hypothetical time series.

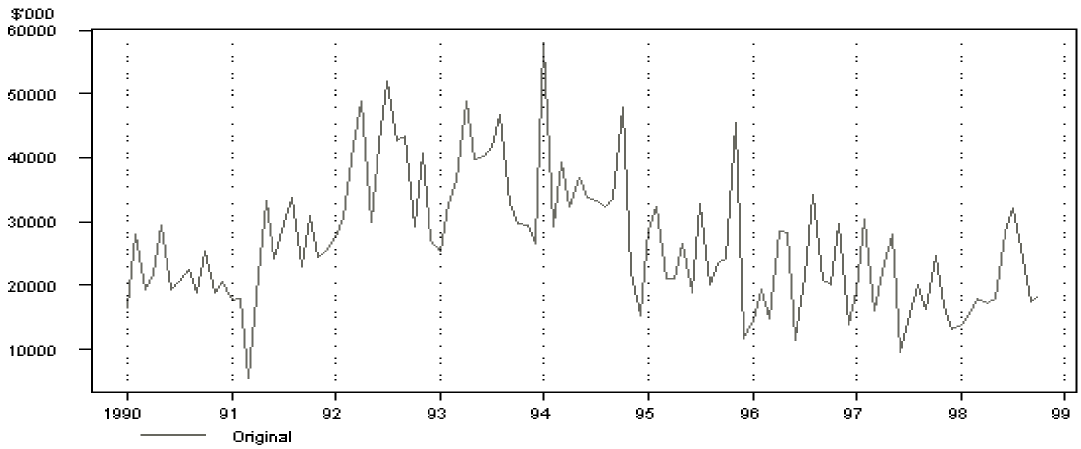

Figure 3.4 Typical irregular effect graph of a hypothetical time series data.

3.4 A COMMON ASSUMPTION IN TIME SERIES TECHNIQUES

A common assumption in many time series techniques is that the data are stationary.

A stationary process has the property that the mean, variance and autocorrelation structure do not change over time. Stationarity can be defined in precise mathematical terms as

(i) the mean $u(t) = E(y_t)$

(ii) the variance $\sigma^2(t) = var(y_t) = \gamma(0)$

(iii) The autocovariances $\gamma(t_1, t_2) = Cov(y_{t1}, y_{t2})$

hence a time series is said to be strictly stationary if the joint distribution of any set of n observations $y(t1, t2) = Cov(y_{t1}, y_{t2})$ is the same as the joint distribution of $y(t1), y(t2) \ldots\ldots y(t_n)$ for all n and k

If the time series is not stationary, we can often transform it to stationarity with one of the following techniques.

1. We can difference the data. That is, given the series Z_t, we create the new series

   $$Y_i = Z_i - Z_{i-1}$$

   The differenced data will contain one less point than the original data. Although you can difference the data more than once, one difference is usually sufficient.

2. If the data contain a trend, we can fit some type of curve to the data and then model the residuals from that fit. Since the purpose of the fit is to simply remove long term trend, a simple fit, such as a straight line, is typically used.
3. For non-constant variance, taking the logarithm or square root of the series may stabilize the variance. For negative data, you can add a suitable constant to make the entire data positive before applying the transformation. This constant can then

be subtracted from the model to obtain predicted (i.e., the fitted) values and forecasts for future points.

3.5 UNIVARIATE TIME SERIES MODELS

These are time series analysis models with only one series of observations. The monthly crime data displayed in Appendix I is a typical example of a univariate time series. Univariate time series models usually view their series as a function of its own past, random shocks and time. Some basic univariate time series models and their processes are discussed under 3.5.1 below as follows.

3.5.1 Common Approaches to Univariate Time Series

There are a number of approaches to modeling time series. A few of the most common approaches are outlined below.

a). Decomposition

One approach is to decompose the time series into a trend, seasonal, and residual component. In other words decomposition refers to separating a time series into trend, cyclical, and irregular effects. Decomposition may be linked to de-trending and de-seasonalizing data so as to leave only irregular effects, which are the main focus of time series analysis. Triple exponential smoothing is an example of this approach. Another example, called seasonal loess, is based on locally weighted least squares and is discussed by Cleveland (1993).

b). The spectral plot

Another approach, commonly used in scientific and engineering applications, is to analyze the series in the frequency domain. An example of this approach in modeling a sinusoidal type data set is shown in the beam deflection case study. The spectral plot is the primary tool for the frequency analysis of time series.

Detailed discussions of frequency-based methods are included in Bloomfield (1976), Jenkins and Watts (1968), and Chatfield (1996).

c). Autoregressive (AR) models

Another common approach for modeling univariate time series is the autoregressive (AR) model:

$$Y_i = Z_i - Z_{i-1}$$

where $\mathbf{Y_i}$ is the time series, and

$$Y_i = Z_i - Z_{i-1}$$

with $Y_i = Z_i - Z_{i-1}$ denoting the process mean.

An autoregressive model is simply a linear regression of the current value of the series against one or more prior values of the series AR(p). The value of ***p*** is called the order of the AR model. AR models can be analyzed with one of various methods, including standard linear least squares techniques. They also have a straightforward interpretation.

d). Moving Average (MA) models

Another common approach for modeling univariate time series models is the moving average (MA) model:

$$Y_i = Z_i - Z_{i-1}$$

where $\mathbf{Y_t}$ is the time series, $Y_i = Z_i - Z_{i-1}$ is the mean of the series, $Y_i = Z_i - Z_{i-1}{}_{1}$, ... , $Y_i = Z_i - Z_{i-1}{}_{q}$ are the parameters of the model MA(q). The value of **q** is called the order of the MA model. That is, a moving average model is conceptually a linear regression of the current value of the series against the white noise or random shocks of one or more prior values of the series. The random shocks at each point are assumed to come from the same distribution, typically a normal distribution, with location at zero and constant scale. The distinction in this model is that these random shocks are propagated to future values of the time series. Sometimes the ACF and PACF will suggest that a MA model would be a better model choice and sometimes both AR and MA terms should be used in the same model.

It is also important to note, however, that the error terms after the model is fit should be independent and follow the standard assumptions for a univariate process.

Box and Jenkins popularized an approach that combines the moving average and the autoregressive approaches (Box, Jenkins, and Reinsel, 1994). This resulted in autoregressive moving average model (ARMA).

The Box-Jenkins model assumes that the time series is stationary. Box and Jenkins recommend differencing non-stationary series one or more times to achieve stationarity. Doing so produces an ARIMA model, with the "I" standing for

"Integrated". This is described in detail below since it is the main method used in the analysis of data in this research.

3.6 BOX-JENKINS ARIMA PROCESS

In statistics and econometrics, and in particular in time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average(ARMA) model. The Box–Jenkins methodology, named after the statisticians George Box and Gwilym Jenkins, applies ARIMA models to find the best fit of a time series to past values of this time series, in order to make forecasts. They are applied in some cases where data show evidence of non-stationarity, where an initial differencing step (corresponding to the "integrated" part of the model) can be applied to remove the non-stationarity.

The model is generally referred to as an ARIMA(p,d,q) model where p, d, and q are non-negative integers that refer to the order of the autoregressive, integrated, and moving average parts of the model respectively.

3.6.1 Modeling approach

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

1. Model identification and model selection: making sure that the variables are stationary, identifying seasonality in the dependent series (seasonally differencing it if necessary), and using plots of the autocorrelation and partial autocorrelation functions of the dependent time series to decide which (if any)

autoregressive or moving average component should be used in the model.

2. Parameter estimation using computation algorithms to arrive at coefficients which best fit the selected ARIMA model. The most common methods use maximum likelihood estimation or non-linear least-squares estimation.
3. Model checking by testing whether the estimated model conforms to the specifications of a stationary univariate process. In particular, the residuals should be independent of each other and constant in mean and variance over time. (Plotting the mean and variance of residuals over time and performing a Ljung-Box test or plotting autocorrelation and partial autocorrelation of the residuals are helpful to identify misspecification.) If the estimation is inadequate, we have to return to step one and attempt to build a better model.

3.6.1.1 Box-Jenkins model identification

Stationarity and seasonality

The first step in developing a Box–Jenkins model is to determine if the time series is stationary and if there is any significant seasonality that needs to be modeled.

Detecting stationarity

Stationarity can be assessed from a run sequence plot. The run sequence plot should show constant location and scale. It can also be detected from an autocorrelation plot. Specifically, non-stationarity is often indicated by an autocorrelation plot with very slow decay.

Finally, unit root tests provide a more formal approach to determining the degree of differencing such as Kwiatkowski-Phillips-Schmidt-Shin (KPSS) and Phillips-Perron Unit Root Tests are carried out employing the unit root testing procedures of Hamilton (1994). The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for the null hypothesis of a level stationary against an alternative of unit root together with the Philips-Peron test for the null hypothesis of a unit root against the alternative of a stationary series.

The decision rule is that for the KPSS test if the p-value of its test statistic is greater than the critical value of say 0.05, then reject the null hypothesis of having a level stationary series and therefore conclude the alternate hypothesis that it has a unit root. The Philips-Peron Test on the other hand test for the null hypothesis of unit root against an alternative hypothesis of stationarity by rejecting the null hypothesis if its *p*-value is less than the critical value chosen.

Detecting seasonality

Seasonality (or periodicity) can usually be assessed from an autocorrelation plot, a seasonal subseries plot, or a spectral plot.

Differencing to achieve stationarity

Box and Jenkins recommend the differencing approach to achieve stationarity. However, fitting a curve and subtracting the fitted values from the original data can also be used in the context of Box–Jenkins models.

Seasonal differencing

At the model identification stage, the goal is to detect seasonality, if it exists, and to identify the order for the seasonal autoregressive and seasonal moving average terms. For many series, the period is known and a single seasonality term is sufficient. For example, for monthly data one would typically include either a seasonal AR 12 term or a seasonal MA 12 term. For Box–Jenkins models, one does not explicitly remove seasonality before fitting the model. Instead, one includes the order of the seasonal terms in the model specification to the ARIMA estimation software. However, it may be helpful to apply a seasonal difference to the data and regenerate the autocorrelation and partial autocorrelation plots. This may help in the model identification of the non-seasonal component of the model. In some cases, the seasonal differencing may remove most or all of the seasonality effect.

Identify *p* and *q*

Once stationarity and seasonality have been addressed, the next step is to identify the order (i.e., the p and q) of the autoregressive and moving average terms. These are determined by examining the values of the autocorrelations and the partialautocorrelations with their corresponding plots as explained below.

Autocorrelation and partial autocorrelation plots

The primary tools for doing this are the autocorrelation plot and the partial autocorrelation plot. The sample autocorrelation plot and the sample partial autocorrelation plot are compared to the theoretical behavior of these plots when the order is known.

Order of autoregressive process (p)

Specifically, for an AR(1) process, the sample autocorrelation function should have an exponentially decreasing appearance. However, higher-order AR processes are often a mixture of exponentially decreasing and damped sinusoidal components.

For higher-order autoregressive processes, the sample autocorrelation needs to be supplemented with a partial autocorrelation plot. The partial autocorrelation of an AR(p) process becomes zero at lag p + 1 and greater, so we examine the sample partial autocorrelation function to see if there is evidence of a departure from zero. This is usually determined by placing a 95% confidence interval on the sample partial autocorrelation plot (most software programs that generate sample autocorrelation plots will also plot this confidence interval). If the software program does not generate the confidence band, it is approximately $\pm 2/\sqrt{N}$, with N denoting the sample size.

Order of moving-average process (q)

The autocorrelation function of a MA(q) process becomes zero at lag q + 1 and greater, so we examine the sample autocorrelation function to see where it essentially becomes zero. We do this by placing the 95% confidence interval for the sample autocorrelation function on the sample autocorrelation plot. Most software that can generate the autocorrelation plot can also generate this confidence interval.

The sample partial autocorrelation function is generally not helpful for identifying the order of the moving average process.

Model Estimation

After identifying the order of the tentative model, the parameters of the model are estimated using the maximum likelihood estimation to determine the AR and MA parameters, as well as all other parameters reported in the study.

Three other penalty function statistics namely the Akaike information criteria (AIC), the Schwarz Bayesian information criteria as well as the corrected Akaike information criteria (AICc) are explained in penalizing fitted models based on the principle of parsimony. These statistics were one of the various checks used to verify the adequacy of the chosen models. Comparatively, models with the smallest AIC and BIC are deemed to have residuals which resembles a white noise process. Twice the number of estimated parameters minus two times the log likelihood gives the AIC value of a model. The BIC is computed as $-2\ln(L)+\ln(n)k,$ where L is the likelihood, n denotes the number of residuals and k is the number of free parameters

Each parameter estimate reports standard error for that particular parameter. Using the parameter estimate and its standard error, a test for statistical significance (t-value) are then conducted. For statistically significant parameters, the absolute values of the *t*-ratios are expected to be greater than 1.96 or 2 in order for the parameters to be maintained in the model whereas parameters which are not significant are trimmed or removed from the model.

Furthermore, the estimated AR and MA parameters must also conform to certain boundary condition, that is they must lie between -1 and 1. If the AR and MA parameters do not lie within those bounds of stationarity then the parameters of the model are re-estimated or if possible a different candidate model is alternatively considered for estimation. All these checks when strictly adhered to would lead to obtaining reliable results from the model.

Diagnostic checking

The diagnostic stage of the Box-Jenkins ARIMA process is to examine whether the fitted model follows a white noise process. This can be done by studying the autocorrelation values (r_k) one at a time, and to develop a standard error formula to test whether a particular r_k value is significantly different from zero. Theoretically, it is envisage that all autocorrelation coefficients for a series of random numbers must be zero. However, because of the presence of finite samples, each sample autocorrelations might not be exactly zero. The ACF coefficients of white noise data is said to have a sampling distribution that can be approximated by a normal curve with mean zero and standard error of $\frac{1}{\sqrt{n}}$, where n gives the number of data points in the observed series. For a white noise process, 95% of all sample autocorrelation values (r_k) must lie within a range specified by the mean plus or minus 1.96 standard errors. In this case, since the mean of the process is zero and the standard error is $\frac{1}{\sqrt{n}}$, one should expect about 95% of all sample autocorrelation values (r_k) to be within the range of $\pm 1.96\sqrt{n}$ or $(-1.96\sqrt{n} < r_k < 1.96\sqrt{n})$. If this condition does not hold, then the model fitted do not follow a white noise process, or the residuals are not white noise. The correlogram of the ACF would therefore show lines at the critical values of $\pm 1.96\sqrt{n}$ for easily verification.

The Ljung-Box test is a modified version of the portmanteau test statistic developed by Ljung and Box (1978) is also used. The modified Ljung- Box Q statistic tests whether the model's residuals have a mean of zero, constant variance and serially uncorrelated r_k values (a white noise check). The test statistic is given by;

$$Q = n(n+2)\sum_{k=1}^{h}\frac{r_k^{\ 2}}{(n-k)}$$ where n denote the number of data points in the series, $r^2{}_k$ is the square of the autocorrelation at lag k, and h is the maximum lag being considered. The hypothesis to be tested is formulated in the form;

H_0: The set of autocorrelations for residual is white noise (model fit data quite well)

H_1: The set of autocorrelations for residual is different from white noise

The test statistic (Q) is compared with a chi-square distribution written as $\chi^2{}_{\alpha,(h-p-q)}$, where α is taken to be 5% (0.05), h is the maximum lag being considered, and p and q are the order of the AR and MA processes respectively. The decision is to accept the null hypothesis (H_0) if $Q < \chi^2{}_{\alpha,(h-p-q)}$, and to reject the alternative hypothesis if $Q > \chi^2{}_{\alpha,(h-p-q)}$. In other words, the residuals are not white if the test statistic Q lies in the extreme 5% $(\alpha = 0.05)$ of the right-hand tail of the chi-square distribution.

3.7 THE ARIMA MODEL AND INTERVENTION ANALYSIS

A stochastic, time-series ARIMA model will be used to analyze the dynamics of changes, variations and interruptions in the crime situation in Ghana through time-series data. This ARIMA model can help to perceive whether the various crime combat interventions impacts the reduction and curbing of various forms of crime and the nature of their effects, if any.

The applications of the ARIMA model with and without intervention analysis have been widely used in different aspects, such as flexible manufacturing system scheduling and simulation (Ip, 1997; Ip et al., 1999), tourism forecasting (Cho, 2001), investigation and forecast of economic factors (Chung et al., 2008), and impact analysis on air travel demand (Lai and Lu, 2005).

3.7.1 Assumptions of Intervention Analysis Models

The fitting of an ARIMA models coupled with accurate intervention analysis model is dependent upon the fulfillment of these assumptions.

1. Stationarity is a critical assumption of time series analysis, stipulating that statistical descriptors of the time series are invariant for different ranges of the series. Weak stationarity assumes only that the mean and variance are invariant. Strict stationarity also requires that the series is normally distributed. Stationarity is tested by the following tests: Durbin-Watson, Dickey-Fuller, Augmented D-F, and Root Examination for univariate time series. There is also a test (Fountis-Dickey) for joint stationarity when modeling two time series together. Testing stationarity is a first step in time series modeling. These may

be followed by tests for normality: the normal distribution test, Jarqua-Bear, or studentized range tests.

2. Uncontrolled autocorrelation. Time series analysis requires stationarity be established through differencing or some other technique. If two variables trend upward in raw data, they will tend to correlate highly when a linear technique such as OLS (ordinary least-squares) regression is applied. For data in such series, the value of any given datum is largely determined by the value of the preceding datum in the series. This autocorrelation must be controlled before inferences may be made about correlation with other variables. Failure to control autocorrelation is vary apt to lead to spurious results, thinking there is a strong effect.

   More technically, significance tests of OLS regression estimates assume non-autocorrelation of the error terms. Error terms at sequential points in the series should constitute a random series. It is also assumed that the mean of the error terms will be zero (because estimates are half are above and half below the actual values), and the variance of the error terms will be constant throughout the time series. When, as in many time series, the value of a datum in time t largely determines the value of the subsequent datum in time t + 1, a dependency exists linking the error terms and the non-autocorrelation assumption is violated. The practical effect is that the significance of OLS estimates is computed to be far better than actual, leading the researcher to think that significant relationships exist when they do not. The Durbin-Watson test is the standard test for autocorrelation.

3. Applying Linear Techniques to Nonlinear Data. OLS regression assumes linear relationships. Applying linear techniques to nonlinear data will underestimate relationships and increase error of estimate. As with other uses of OLS regression, the linearity assumption is not violated by adding power or other nonlinear transform terms to the equation. A common test is Ramsey's RESET test, discussed in the section on data assumptions. There are a variety of other tests for linear or nonlinear dependence, including the Keenan, Luukkonen, McLeod-Li, and Hsieh tests. If non-linearity is present, it may be possible to eliminate it by double differencing or data transformation.
4. Arbitrary model lag order. Model lag order can have great effects on results. While tests exist to determine the optimal model order, these tests are purely statistical in nature. The researcher should have a theoretical basis establishing the face validity of the order of the model he or she has put forward.
5. No outliers. As in other forms of regression, outliers may affect conclusions strongly and misleadingly.
6. Random shocks. If shocks are present in the time series, they are assumed to be randomly distributed with a mean of 0 and a constant variance.
7. Uncorrelated random error. Residuals in a good time series model will be randomly distributed, exhibit a normal distribution, have non-significant autocorrelations and partial autocorrelations, and have a mean of 0 and homogeneity of variance over time. Correlated error does not bias estimates but

does inflate standard errors, making statistical inference problematic. The Durbin-Watson, Ljung-Box and a few other tests are the standard tests for correlated error.

8. Closeness of system. The system in which the input event and the impact response take place is usually assumed to be closed, apart from the noise model of the series, the only exogenous impact on the series is presumed to be that of the intervention event.

9. The temporal delimitations of the input event are presumed to be known. The time of onset, the duration, and the time of termination of the input event have to be identifiable again with the noise model that describes the preintervention series has to be stable with a mean of zero.

10. Enoughness of observations. Another assumption is that there should be enough observations in the series before and after the onset of the event for the analyst to separately model the preintervention and post intervention series by whichever parameter estimation process the analyst chooses to use. In the conventional approach where there is enough dataset prior to the intervention event, the preintervention series is modeled first, and the impact is modeled afterward.

3.7.2 Transfer Function and Univariate ARIMA Model

The ARIMA model developed by Box and Jenkins (1976) has become popular due to its advantages of power and flexibility (Pankratz, 1983). Put simply, the general transfer function of ARIMA is of the following form:

$$Yt = \Sigma[\omega (B)\ \delta (B)]B^{b}\ X_{t} + [\theta (B)\ \varphi (B)]\varepsilon_{t} \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (3.1)$$

where Yt and Xt are the output and input series respectively,

b is the time delay,

$\omega (B)\ \delta (B)$ is the polynomial of the transfer function,

$[\theta (B)\ \varphi (B)]\varepsilon_{t}$ is the noise model, and

ε_{t} is the residual, i.e. white noise.

Equation (3.1) is further simplified to obtain Equation (3.2) below:

$$Y_{t} = V (B)X_{t} + N_{t} \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots\ldots(3.2)$$

where $V (B) = \delta^{-1} (B)\omega (B)B^{b}$,

$\omega (B) = \omega_{0} - \omega_{1}B - \ldots - \omega_{s} B^{s}$,

$\delta (B) = 1 - \delta_{1}B - \ldots - \delta_{r} B^{r}$, and $N_{t} = [\theta (B)\ \varphi (B)]\varepsilon_{t}$.

The univariate ARIMA model combines three components: Autoregressive (AR), Integration (I) and Moving Averages (MA), therefore the general form of a univariate ARIMA model denoted as ARIMA(p, d, q)(P, D,Q) $_{s}$ is defined as Equation (3.2), which when further simplified results as Equation (3.3):

$$(1 - B)^{d} (1 - B^{s})^{D}Y_{t} = [\theta_{q} (B)\Theta_{Q} (B^{s}) / \varphi_{p} (B)\Phi_{p} (B^{s})]\varepsilon_{t} \quad \ldots \quad `\ldots \quad \ldots \quad (3.3)$$

$$\Delta d\Delta DsY_{t} = [\theta_{q} (B)\Theta_{Q} (B^{s}) / \varphi_{p} (B)\Phi_{p}(B^{s})]\varepsilon_{t} \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (3.4)$$

where p , d , and q are the order of the AR, I and MA terms respectively; P , D , and Q are the order of the seasonal AR, I and MA terms respectively;

$\Delta d = (1 - B)^d$ and $\Delta D_s = (1 - B^s)^D$ represent the regular and seasonal I operators respectively; $\varphi_p(B)$ and $\Phi_p(B^s)$ are the nonseasonal and seasonal AR operators

respectively; $\theta_q(B)$ and $\Theta_Q(B^s)$ are the nonseasonal and seasonal MA operators respectively; and ε_t is the disturbance or random error.

3.7.3 ARIMA Model with Intervention Analysis

As discussed by Box and Tiao (1975), an intervention model is of the general form:

$$Y_t = V(B)I_t + N_t \qquad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \qquad (3.5)$$

where I_t is an intervention or dummy variable that is defined as:

$$I_t = \begin{cases} 1, t = T \\ 0, t \neq T \end{cases} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.6)$$

In this instance the intervention input begins in 2003 ($t=T$) where it is coded as 1, and remains for just a period in the case of the pulse function, but remains as 1 for the entire presence of the intervention exercise in the case of the step function and is therefore with regard to the Community Policing Unit intervention events are formulated as; $Y_t = c + w_1 I_{1t} + \frac{\theta(L)}{\phi(L)}\varepsilon_t$, ………………………………………….(3.7) where $I_{1t} = S_t^{(2003)} = \begin{cases} 1, & t \geq 2003 \\ 0, & otherwise \end{cases}$ ……………………………………………(3.8)

c is a constant and Y_t is the level of change with respect to gains or losses made in the value of reduction. The intervention variable I_{1t} is a step function which corresponds to the community policing unit programme.

3.7.4 Procedures of Model Development

In general, the model estimation of ARIMA consists of the following three stages:

1. Unit root test and identification of the order of difference, i.e. d. This preliminary step is essential to stabilize the time-series data and reduce the residual. The Augmented Dickey-Fuller (ADF) test is often employed for the analysis of the unit root, where the null hypothesis is that the input series has a unit root.

2. Estimation and diagnosis of the parameters of transfer function, i.e. p and q. The autocorrelation function (ACF), partial autocorrelation function (PACF) and cross-autocorrelation function (CACF) are important to tentatively estimate the parameters of transfer function, while statistical measures naturally provide statistical evidence to support the determination of an appropriate transfer function.

3. Residual/noise diagnostic check. The correlogram of Q-statistics based on the ACF and PACF of the residual is generally used for residual analysis.

3.7.5 Estimation and Diagnostic checking for the Impact Assessment Models

The main parameters estimated are the impact (w_0) and the decay (δ) parameters using any chosen estimation method. The time delay is usually deduced from the spikes of the time plot such that if the size of the spikes in the time plot becomes pronounced at the second lag after the impact, then the delay time should be set to two time periods. The maximum likelihood estimation method is most preferably used to estimate the intervention parameters as well as that of the noise, N_t.

An estimated value for w_0 computed as $\frac{\sum_T Y_t I_t}{\sum_T Y_t^2}$, would apparently indicate the size of rise or drop in the level of the response series. Also, the delta (δ) parameter associated with the gradual temporary and the gradual permanent effects estimates the adjustments subsequent to the change (pulse decay) or the overall rate of increase (gradual permanent), and is expected to be within the bounds of system stability $(-1 < \delta < 1)$. Moreover, estimation of the impact parameters should reflect what is known and observed about the impact of the intervention event.

Diagnostic checks on the significance of the hypothesized parameters and the behaviour of the residuals are conducted on the tentative model. A *t*-test statistic is used to check the significance of all parameters reported from the estimated impact model. Any parameter whose *t*-value is less than 1.96 or 2 is deemed to have a non significant effect and must be quickly removed or trimmed from the model. The examination of the residuals should verify whether it follows a white noise process or otherwise. When the residuals are tested to be white noise, then the adequacy of the impact or intervention model will be fully established.

3.7.6 Forms of Intervention Models

An intervention model may be formulated as; $Y_t = \frac{w_0}{(1-\delta L)} I_{t-b} + N_t$ ....(3.9)

where I_{t-b} is the intervention indicator variable normally known as the change agent, scored 0 or 1 for the absence or presence of the intervention event and the subscript b is a possible time delay for the impact to take off. w_0 is the impact parameter which indicates the magnitude of the impact, and δ represent the decay parameter, whereas N_t is the noise model. Depending on the situation that prevails, the response series Y_t may not quickly observe the impact of the intervention event. The b index in I_{t-b} gives the number of periods delayed between the onset of a known intervention and the actual time it is impacting on the response series (Y_t). If b is assigned a value of 2, there would be exactly two time periods of delay between the intervention event I_t and the time it takes for its impact to be fully realized on the response series Y_t.

Usually, there are two major forms that characterize intervention or impact assessments. These are usually observed by the duration and nature of the impacts. Some interventions could give temporary or permanent effects with respect to the duration. The nature of impacts can also be seen as abrupt or gradual processes. Sudden and constant changes (abrupt permanent) are normally attributed to step functions; sudden and instantaneous changes (abrupt temporary) are modeled with pulse function; gradual and permanent effects are mainly modeled with step function with first-order decay rate; gradual and decaying changes (pulse decay) can also be modeled with pulse function with first-order decay rate.

1. **Abrupt Permanent and Gradual Permanent effects**

The abrupt onset and permanent duration effects are popularly called a simple step function. Step functions are mainly used to model permanent changes in the response series (Y_t). A step function with a first-order decay rate may be written as;

$$Y_t = \frac{w_0}{(1-\delta_1 L)} I_{t-b} + N_t \quad \ldots\ldots(3.10)$$

If after fitting the model in (3.10) the denominator reduces to unity, the model will then be called a simple step function with a zero-order decay, where $f(I_t) = s_t^{(T)} = w_0 I_{t-b}$. Also, if there is no time delays $(b=0)$, then $f(I_t) = s_t^{(T)} = w_0 I_t$, and the full model will now be of the form;

$Y_t = w_0 I_t + N_t$ ……………………………………………………………………......(3.11)

However, the gradual permanent effects are characterized by slow changes in the level of the series that usually result in a new permanent level. It is usually modeled with step functions with first-order decays as shown in (3.10). If the index $b = 0$ in (3.10), the gradual permanent effect model then becomes; $Y_t = \frac{w_0}{(1-\delta_1 L)} I_t + N_t$ ……….(3.12)

where $-1 < \delta_1 < 1$.

Again, if the noise model (N_t) is subtracted from the response series (Y_t), then

$$Y_t - N_t = \frac{w_0}{(1-\delta L)} I_t \text{ or } Y_t^* = \frac{w_0}{(1-\delta L)} I_t .$$

We then expand and simplify to obtain;

$$Y_t^* (1-\delta L) = w_0 I_t$$

$$Y_t^* - \delta L Y_t^* = w_0 I_t$$

$$\Rightarrow \quad Y_t^* - \delta Y_{t-1} = w_0 I_t .$$

Since the impact at t is w_0, then the impact at $t+1$ is given by;

$$Y^*{}_{t+1} = \delta Y_t^* + w_0 I_{t+1} = w_0 (1+\delta) .$$

The impact over time or the change level obtained from the gradual permanent effect is given by;

$$w_0(1+\delta+\delta^2+\delta^3+\delta^4+\delta^5+\delta^6...).I_t$$

Therefore, $Y^*_{i+n}=\sum_{k=0}^{n}\delta_1{}^k w_0$. Again, since $\delta<1$, and δ^{10} or δ^{100} is infinitesimal, then the effect as $t\to\infty$ gets very smaller. In all, the asymptotic or the long term change given by a gradual permanent effect is of the form; $\frac{w_0}{1-\delta}$ ………………………(3.13)

2. **Abrupt Temporary and Gradual Temporary effects**

Temporary effects are often modeled with pulse functions. The abrupt onset and temporary duration effects are often called the "pulse effect", simply modeled as; $Y_t=w_0 I_{t-b}+N_t$, where I_{t-b} is the intervention indicator coded 0 prior to the event and 1 at the onset, and *b* may indicate a possible time delay between Y_t and I_t. The gradual temporary effects are often modeled with pulse functions having first-order decay rates.

It is also formulated as; $Y_t=\frac{w_0}{(1-\delta_1 L)}I_t(1-L)+N_t$ ……………………………(3.14)

The various forms of the intervention models are further illustrated in the Figure 3.5 below.

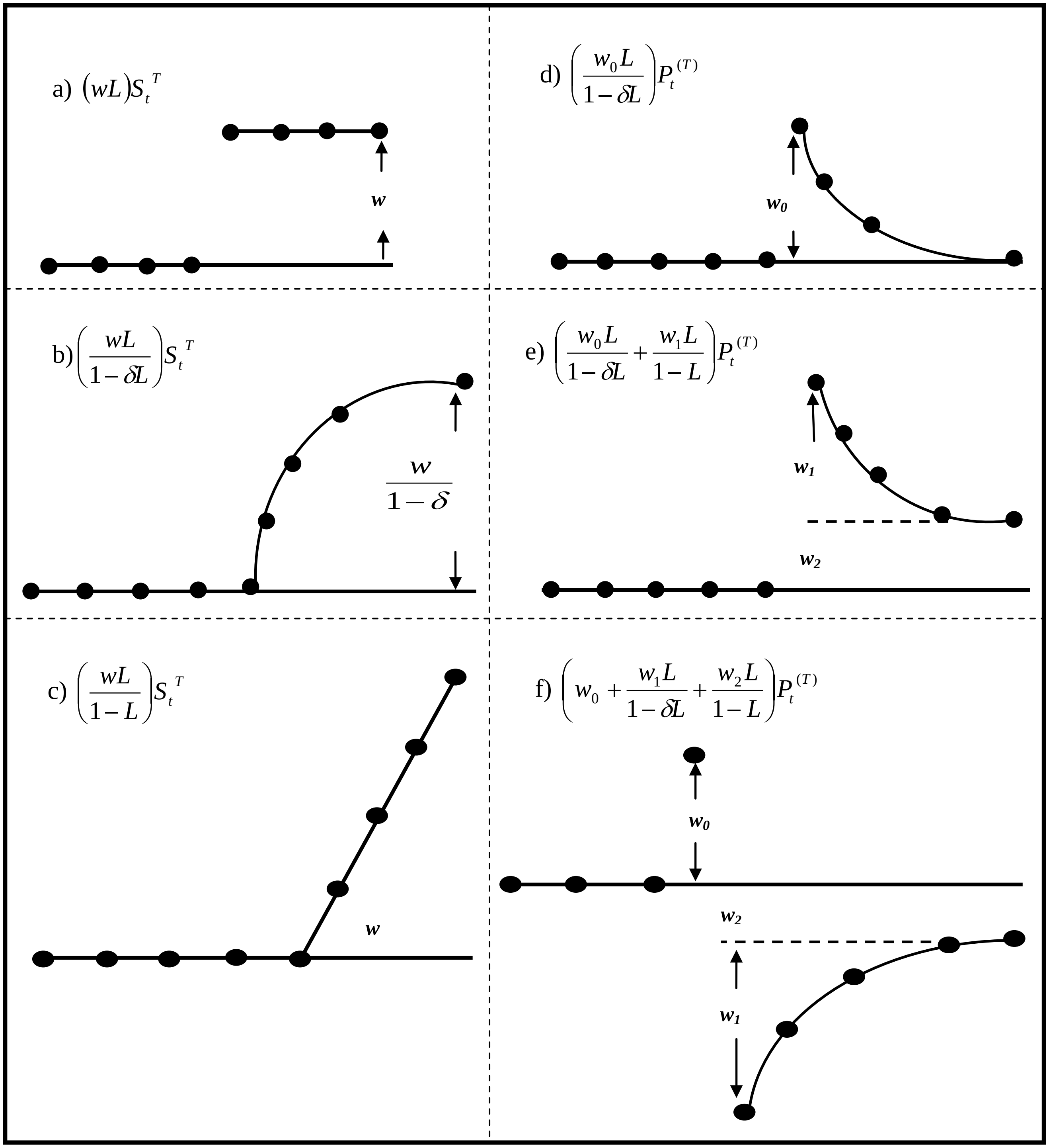

Figure 3.5 Graphical outputs for some hypothetical intervention events

From the Figure 3.5 above, the function (a) $Y_t = (wL)S_t^{(T)}$ may be used to indicate the presence of a permanent step change with unknown level of magnitude w after time T. The model (b) $Y_t = \left(\frac{wL}{1-\delta L}\right)S_t^{(T)}$ also corresponds to a gradual permanent change with decay rate δ, which later results in a long-term change in level given by $\frac{w}{1-\delta}$. Again, $Y_t = \left(\frac{w_1 L}{1-\delta L}\right)P_t^{(T)}$ represents a gradual temporary change after time T of unknown magnitude w_1 and decay rate δ.

Also, (a), (b), (c) represent step and pulse intervention inputs whilst (d), (e), (f) represent response to a step pulse input.

CHAPTER 4

DATA PRESENTATION AND ANALYSES

4.1 INTRODUCTION

This chapter deals with the analyses of the crime statistics data obtained from the Criminal Investigations Department of the Ghana Police Service Eastern Regional Command for the major crime categories as described by the Ghana Police Service namely: murder, rape, defilement, robbery and the use and possession of drugs (cocaine, heroine, Indian hemp) from year 2000 to 2011.

4.2 DISPLAY OF DATA

For the purposes of the flow of the analysis, the time series data for the five (5) major categories of crime in Ghana for the Eastern Region is displayed in Appendix I.

4.3 TIME SERIES GRAPH OF DATA

Time series plots which display observations on the y-axis against equally spaced time intervals on the x-axis used to evaluate patterns and behaviors in data over time is displayed in the Figure 4.1 below:

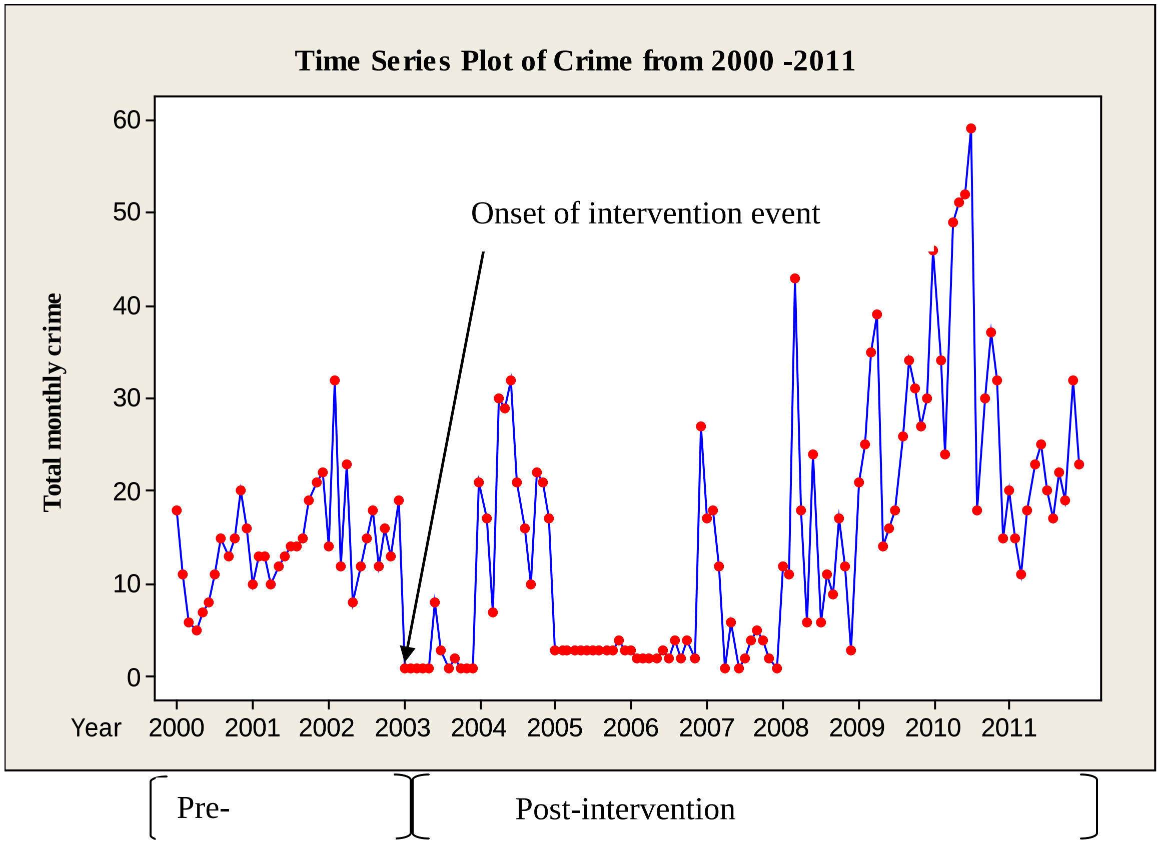

Figure 4.1: Time Series Graph of Crime from 2000 - 2011

The Figure 4.1 above indicates clearly that the occurrence of the five major crime categories in the Eastern Region of Ghana were not constant but rather varied from one year to the other as well as from one month to the other with no systematically visible pattern, structural breaks, outliers, and no identifiable trend components in the time series data or non monotonous (that is consistently increasing or decreasing).

Crime occurrence were fairly high between the years 2000 to 2002 but saw a significant decline in 2003 from about nineteen(19) to one (1) and remained fairly constant in the range of one (1) and eight (8) until 2004 which recorded a sharp spike. Perhaps, the quick decline in 2003 could be attributed to the intervention or the effect of the intervention which remains the focus of the study.

Monthly crime occurrence in the region were generally very low between the years 2005 to 2007 with only a few significant spikes (approximately four (4)) and further rose fairly up and down movements between 2008 and 2011.

In conclusion, the monthly occurrence of the major crime in the Eastern region of Ghana with regard to the plot in figure 4.1 above can be perceived to exhibit four main characteristics: a steady increase from 2000 to 2002; an apparent decrease from 2003 to 2004; an approximately constant rate between 2005 and 2007 and finally very high rapid fluctuations from 2007 to 2011.

4.4 MODELLING THE PRE-INTERVENTION SERIES (2000 -2002)

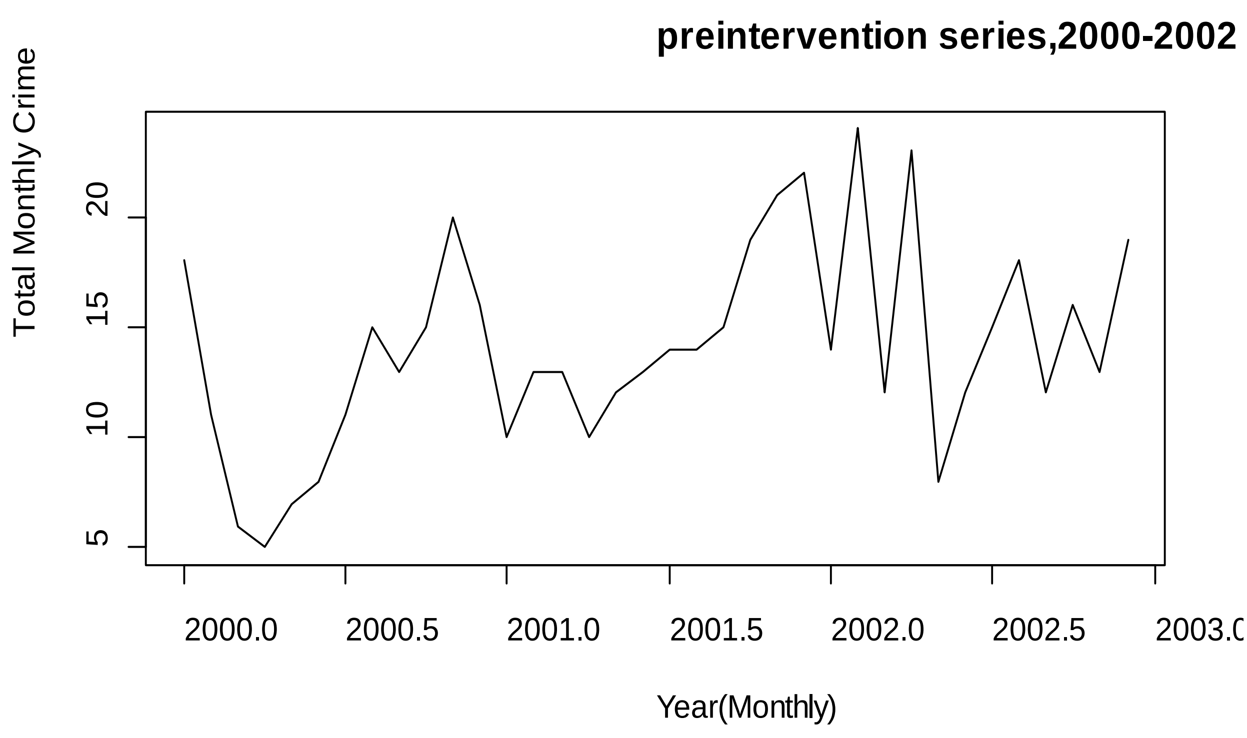

Figure 4.2: Time Series Graph of the Pre-intervention series

The Figure 4.2 above which is the plot of the preintervention series is absolutely the same as the first part of the main graph displayed in Figure 4.1 above. It indicates clearly that the monthly occurrence of the five major crime categories in the Eastern Region of Ghana saw a sharp decline in the part of year 2000 which was subsequently followed by a steady monthly increase in crime up to the end of year which saw a drastic downturn. Again, the year 2001 showed a steady increase or growth in crime and ended yet another sharp decline, with year 2002 exhibiting rapid fluctuations throughout.

Even though, the graph does not show systematically visible pattern, structural breaks, outliers, and no identifiable trend components in the time series data or non monotonous (that is consistently increasing or decreasing) but the sharp decrease that do occur at the end of each year suggest that there could be some seasonal component present that could render the series nonstationary and are investigated by the examination of the autocorrelation and partial autocorrelation functions as shown below.

4.4.1 Model Identification process for the preintervention series

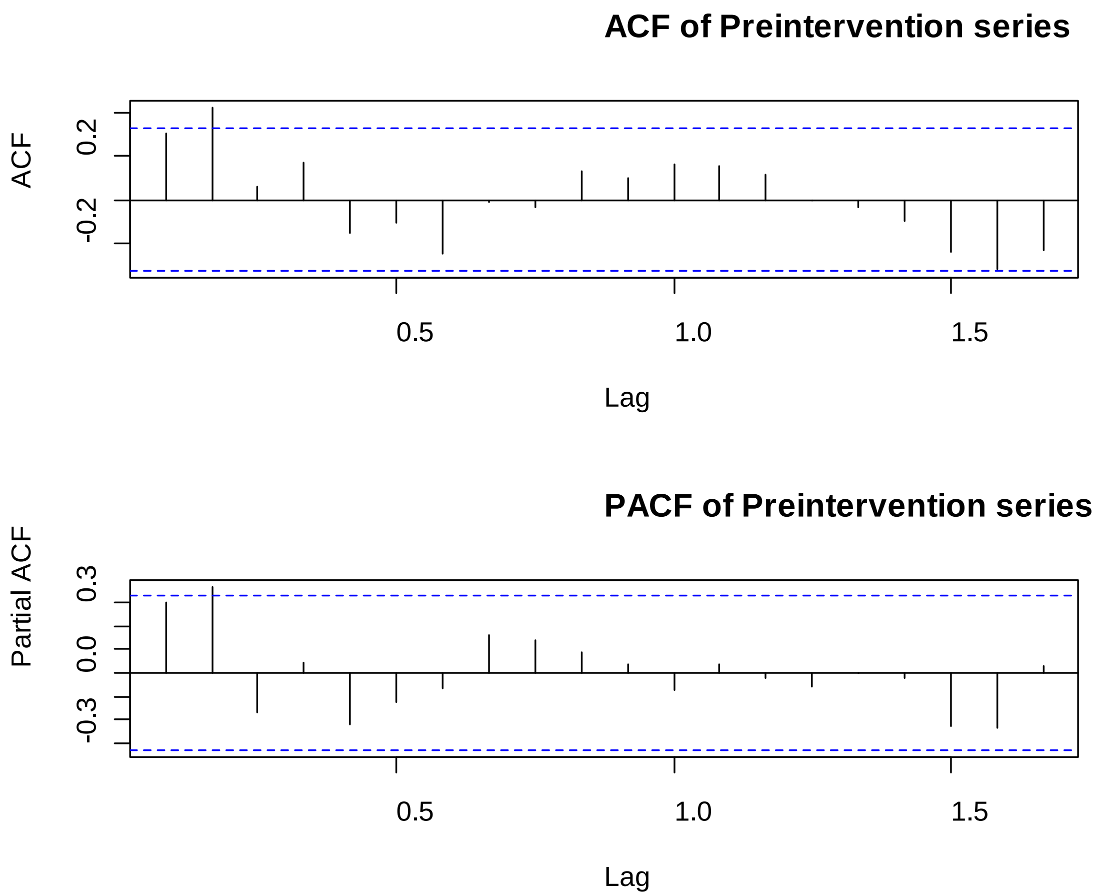

Figure 4.3 Plot of autocorrelations of the crime data (autocorrelogram)

The Figure 4.3 above (autocorrelogram) displays graphically and numerically the autocorrelation function (ACF) and the partial autocorrelation function (PACF), that is, serial correlation coefficients (and their standard errors) for consecutive lags in a specified range of lags. The ACF displays high positive spike at lag 2 which slowly die out not to or near zero with the spikes further alternating in sign resulting in a wavelike form which is typical of a non stationary series.

The PACF for the crime preintervention series is not different from that of its ACF as it can be seen from the Figure 4.3 above thereby depicting a series that is either non stationary in the mean and perhaps variance and therefore require some form of differencing or transformation.

Table 4.1 Autocorrelation Function for Pre-intervention Series

Lag	ACF	T	LBQ
1	0.302981	1.82	3.59
2	0.424782	2.34	10.85
3	0.063945	0.31	11.02
4	0.169204	0.81	12.24
5	-0.157127	-0.74	13.33
6	-0.102476	-0.48	13.81
7	-0.248941	-1.15	16.73
8	-0.012889	-0.06	16.74
9	-0.036997	-0.17	16.81
10	0.132453	0.59	17.73
11	0.102649	0.45	18.31
12	0.162011	0.71	19.81
13	0.155868	0.68	21.25
14	0.114452	0.49	22.07
15	-0.000794	-0.00	22.07
16	-0.036623	-0.16	22.16
17	-0.09947	-0.46	22.88
18	-0.244660	-1.03	27.43
19	-0.317076	-1.30	35.52
20	-0.235805	-0.93	40.27
21	-0.191450	-0.73	43.61
22	-0.152901	-0.58	45.90
23	-0.086528	-0.32	46.69
24	-0.041407	-0.15	46.88
25	0.065277	0.24	47.41
26	0.021014	0.08	47.47
27	0.045870	0.17	47.79
28	-0.064994	-0.24	48.51
29	-0.060293	-0.22	49.22
30	-0.041133	-0.15	49.61
31	-0.024601	-0.09	49.78
32	-0.065688	-0.24	51.25
33	-0.037983	-0.14	51.91
34	-0.025505	-0.09	52.35
35	0.025313	0.09	53.23

The ACF's for the pre-intervention series displayed in Table 4.1 above shows a large positive significant spike at several lag 2 with ACF of 0.424782 and a corresponding T-statistic of 2.34. In checking for stationary series using autocorrelations, one commonly used rule is that a t-statistic greater in absolute value than 2 indicates that the corresponding autocorrelation is not equal to zero, or significantly different from zero. Since the T-statistic for lag 2 is greater than 2 in absolute value it follow that the corresponding ACF of 0.424782 for lag 2 is significantly different from zero. Furthermore, the ACF's for the highest lags that is 32, 33, 34, and 35 do not tend or approximate to zero indicating a typical case of non stationary series since for a stationary process, the main feature of the correlogram is that the autocorrelations tend toward zero as the lag increases (Harvey, 1993).

Table 4.2 Partial Autocorrelation Function for the Pre-intervention Series

Lag	ACF	T
1	0.302981	1.82
2	0.366641	2.20
3	-0.163659	-0.98
4	0.043917	0.26
5	-0.215286	-1.29
6	-0.122902	-0.74
7	-0.062876	-0.38
8	0.164949	0.99
9	0.136577	0.82
10	0.089567	0.54
11	0.037114	0.22
12	-0.074951	-0.45
13	0.038885	0.23
14	-0.018178	-0.11
15	-0.058524	-0.35
16	0.001989	0.01
17	-0.022364	-0.13
18	-0.228286	-1.37
19	-0.232526	-1.40
20	0.029700	0.18
21	0.024818	0.15
22	-0.027793	-0.17
23	0.003363	0.02
24	-0.102122	-0.61
25	-0.022127	-0.13
26	-0.054717	-0.33
27	0.036894	0.22
28	-0.009447	-0.06
29	-0.015919	-0.10
30	0.112154	0.67
31	0.063253	0.38
32	0.023871	0.14
33	-0.047214	-0.28
34	-0.044384	-0.27

Again, the partial autocorrelations are examined in the same vein using the t-statistic (T) for a particular lag to test whether or not the corresponding partial autocorrelation coefficient equals zero. The post intervention series, exhibit partial autocorrelation for lag 2, with absolute T-statistic of 2.20 that is significantly different from zero since it is greater in absolute value than 2 just like that of the AFC's with the partial autocorrelations not approximating to zero as the lags increases thereby making the series nonstationary.

Finally, unit root tests provide a more formal approach to determining the degree of differencing such as Kwiatkowski-Phillips-Schmidt-Shin (KPSS) and Phillips-Perron Unit Root Tests are carried out as shown below employing the unit root testing procedures of Hamilton (1994).

Table 4.3 Unit Root and Stationarity Tests for the pre-intervention series

Summary of Test Statistic			
Test type	**Test Statistic**	**Lag Order**	**P-value**
KPSS	0.0895	1	0.02547
PPT	-516.479	3	0.066

Table 4.3 above presents the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for the null hypothesis of a level stationary against an alternative of unit root together with the Philips-Peron test for the null hypothesis of a unit root against the alternative of a stationary series.

The KPSS test statistic of 0.0895 with a *p*-value of 0.02547 which is less than the critical value of 0.05 as presented in Table 4.3 rejects the null hypothesis of having a level stationary series and therefore conclude the alternate hypothesis that it has a unit root. Philips-Peron Test on the other hand test statistic and its *p*-value fails to reject the null hypothesis of a unit root at 5% significance level, since its p-value of 0.060 was greater than 0.05.

In conclusion is clear from the time series plot of the preintervention crime series, the ACF's, PACF's with their graphical displays and the objective tests the series has to be transformed or differenced to stabilize or stationarize the data before its capability is assessed or improvements are initiated, since the tests confer non stationarity in the crime preintervention series.

4.4.2 Model Identification process for the differenced preintervention series

After the preintervention series was found to be nonstationary through the various tests, the series is transformed by differencing and the model identification performed as follows:

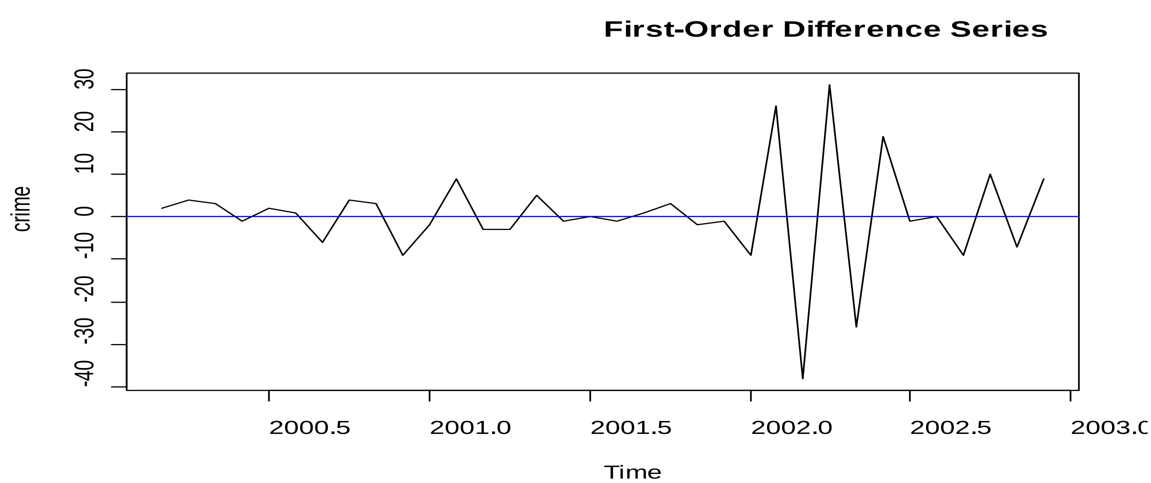

Figure 4.4 Time Series Graph First-Order Differenced Pre-intervention series

The initial step in model identification is to undertake a graphical analysis of the data that would suggest whether a series is likely to be stationary or nonstationary. The Figure 4.4 above display the time series plot of the first difference of the preintervention crime series with the mean superimposed as the blue line. The first order differenced crime series can be said to be stationary if the mean, variance, and covariance of the series remain constant or zero over time. It is clear from the Figure 4.4 that the mean is exactly zero which confers a stationary series.

The unit root test which is a formal method of testing the stationarity of a series is subsequently performed to augment the graphical analysis already performed since ignoring the problem of the unit root will cause an error with the statistical inference (Nelson and Plosser, 1982).

Table 4.4 Unit Root and Stationarity Tests for the Differenced pre-intervention series

Summary of Test Statistic			
Test type	**Test Statistic**	**Lag Order**	**P-value**
KPSS	0.5719	1	0.100
PPT	-33.9837	3	0.01

Table 4.4 above presents the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for the null hypothesis of a level stationary against an alternative of unit root together with the Philips-Peron test for the null hypothesis of a unit root against the alternative of a stationary series.

The KPSS test statistic of 0.5719 with a *p*-value of 0.100 which is greater than the critical value of 0.05 as presented in Table 4.4 do not reject the null hypothesis of having a level stationary series. Philips-Peron Test on the other hand test statistic and its *p*-value of 0.01 reject the null hypothesis of a unit root at 5% significance level, since its p-value is less than 0.05.

It therefore can be concluded the time series plot of the differenced preintervention crime series, and the objective tests indicates that the series is stationary.

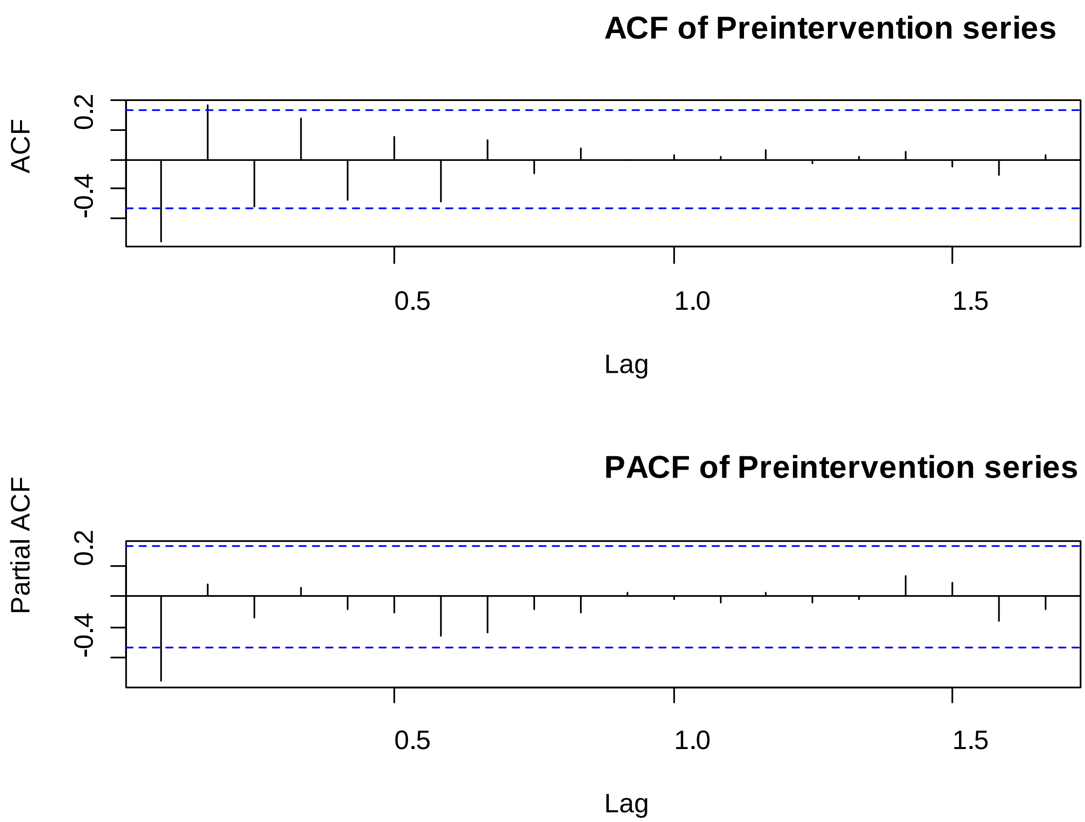

Figure 4.5 Plot of ACF and PACF of the Differenced Pre-intervention crime series

The model identification process in which the form and order of tentative models are basically selected the sample autocorrelation functions and the partial autocorrelation functions of the differenced preintervention crime series after the hypothesis tests have been proven to be stationary.

If the PACF displays a sharp cutoff while the ACF decays more slowly (i.e., has significant spikes at higher lags), we say that the series displays an ‘‘AR signature’’ The lag at which the PACF cuts off is the indicated number of AR terms.

Based on Figure 4.5 above, it can be seen that there is a slow decay in the ACF with single negative significant spike around the PACF which displays a sharp cutoff at lag 1 with ACF and T-statistic -0.554431 and -3.28 respectively. For the differenced preintervention crime series, the partial autocorrelation for lag 1 (-0.554431) is significantly different from zero (the corresponding t-statistic equals -3.28, thus it is greater in absolute value than 2). The other partial autocorrelations have small corresponding t-statistics. This pattern is typical to an autoregressive (AR) process of order one. The identified order of the model is therefore ARIMA (1,1,0) representing AR(1), I(1) and MA(0).

In order to verify the adequacy of the AR(1) process, the model is slightly over-fitted with an AR(2) process and later diagnostic checks performed to obtain the best fitted model.

4.4.3 Estimation of model parameters

At this juncture, the parameters of the selected ARIMA(1, 1, 0) and the slightly over-fitted ARIMA(2, 1, 0) models are estimated using the maximum likelihood estimation method. The results from the two estimated models are therefore displayed in Tables 4.5 and 4.6 respectively.

Table 4.5: Parameter Estimates for ARIMA(1, 1, 0) model

Model Fit Statistics			
AIC	**AICc**	**BIC**	
207.53	208.3	212.19	
Coefficients	**Estimate**	**STD Error**	**t-value**
ar1	-0.5863	0.1397	-41.968
Intercept	0.0400	0.4743	0.0843

Table 4.6: Parameter Estimates for ARIMA(2, 1, 0) model

Model Fit Statistics			
AIC	**AICc**	**BIC**	
209.49	210.83	215.71	
Coeeficients	**Estimate**	**STD Error**	**t-value**
ar1	-0.5647	0.1824	-3.0959
ar2	0.0339	0.1840	0.1842
Intercept	0.0333	0.4918	0.0677

Based on the parameters as reported in Table 4.5 above, the estimate of the ar1 coefficient (ϕ_1) of -0.5863 is found to be statistically significant since it test statistic of t-value of -4.1968 is greater than 2 in absolute terms, and is therefore maintained in the model. The estimated ar1 coefficient again strictly conforms to the bounds of parameter stationary since its value of -0.5863 lies between -1 and 1.

Again the *t*-test conducted on the ARIMA(2, 1, 0) coefficients was not statistically significant for one of the coefficients as reported in Table 4.6. The ar2 coefficient (ϕ_2) of 0.1842 is not statistically significant since its absolute value is less than 2. Therefore the null hypothesis (Ho) of parameter are or equal zero is not rejected resulting in its removal from the model.

Additionally, comparing the ar1 and ar2 models above interms of the AIC, AICc and BIC of (207.49, 208.3 and 212.19) and (209.49, 210.83 and 215.71) respectively, clearly slightly prefer ar1 to ar2 model since their estimated AIC, AICc and BIC are smaller as compared to that of the ar2.

In conclusion, based on the parameter estimates in the Tables 4.5 and 4.6 respectively chose the ar1(ARIMA 1,1,0) as the best model for the preintervention crime series with regard to the tests hypothesis coupled with the examination of the AICs, AICcs and BICs. The exact ar1 (ARIMA 1,1,0) model is thus given as $Y_t = 0.0400 - 0.5863 x_t$.

4.4.4 Model Adequacy (Diagnostic) checking of estimated models

After haven chosen the ar1 (ARIMA 1,1,0) model as the best or tentative model as opposed to the ar2 (ARIM 2,1,0) based on the conclusion under 4.4.3 above, the model adequacy is further checked to draw empirical conclusions regarding the model as good fit and for that matter its usage in estimation and forecasting. These tests are performed using the Ljung-Box Test coupled with the ACF and PACF plot of the residuals as reported in Table 4.7 and Figure 4.6 respectively below to test for correlation for the residuals:

Table 4.7: Ljung-Box Test for ARIMA (1,1,0) model

Summary of Test Statistic			
Test type	**X-squared (χ^2)**	**df**	**P-value**
Ljung-Box	185.789	20	0.5493

The hypothesis that the Ljung-Box test is:

Null hypothesis (Ho): The residuals are uncorrelated

Alternative hypothesis (H1): The residuals are correlated

The test is significant and its corresponding null hypothesis rejected if the p-value is less that chosen critical value of 0.05. From Table 4.7, the Ljung-Box test for the preintervention crime data, the chi-square statistics of 18.5789 gives a corresponding p-value of 0.5493. Because the p-value is quite large (greater than the usually chosen a-level of 0.05), the test is not significant and therefore we do not reject the null hypothesis, thus the residuals appear to be uncorrelated. This indicates that the residuals of the fitted AR(1) model are white noise, and for that matter the model fits the series quite well (the parameters of the model are significantly different from zero and the residuals are uncorrelated), so you can use this model to make forecasts.

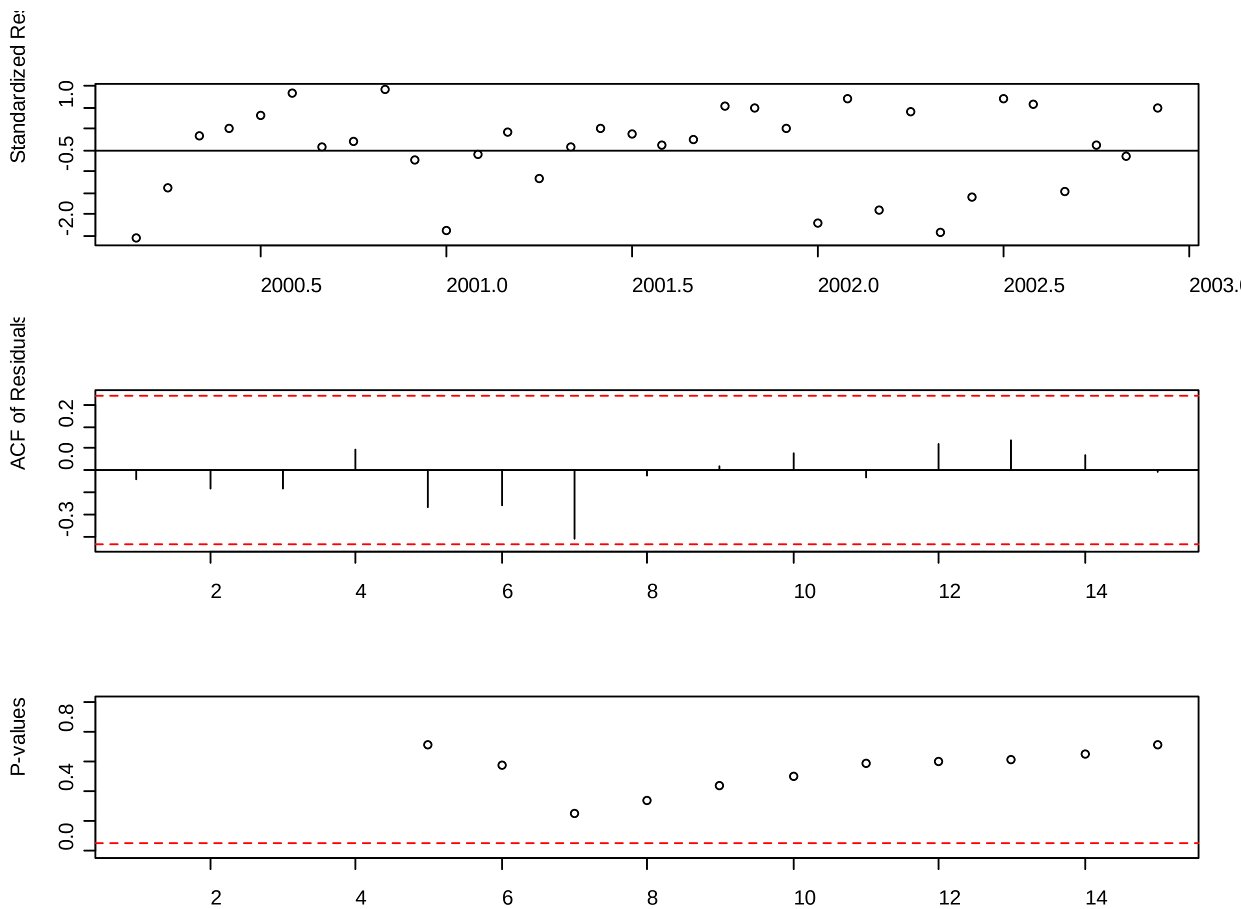

Figure 4.6 Diagnostic Residual plots of ARIMA (1, 1, 0) model

The plots Figure 4.6 comprise of the time plot of the residuals, ACF plot of the residuals and the probability plot of the residuals respectively. The time plot of the residuals clearly shows that the residuals appear to be randomly scattered about zero, no evidence exists that the error terms are correlated with one another as well as no evidence of existence of an outlier. The residuals or errors are therefore conceived of as an independently and identically distributed (i.i.d) sequence with a constant variance and a zero mean. The ACF plot of the residuals shows no evidence of a significant spike in the ACF plot (the spikes are within the confidence limits) indicating that the residuals seem to be uncorrelated. Therefore, the AR(1) model appears to fit well so you can use this model to make forecasts. Finally, the probability plot of the residuals also indicates that the individual probabilities of the residual are greater than or above 0.05 (red line). This shows that the residuals of ARIMA (1, 1, 0) is a white noise process. Thus the residual plots corroborate the conclusion of the Ljung-Box test.

4.5 ESTIMATION AND DIAGNOSTIC CHECKS FOR THE FULL INTERVENTION MODEL

The ARIMA (1, 1, 0) model of the preintervention crime series is now carried out together with the intervention function. The results of the estimated parameters are presented in Table 4.8 below with the, hypothesized intervention model parameters subsequently estimated and diagnosed as follows.

Table 4.8: Parameter Estimates for the hypothesized Intervention model

Model Fit Statistics			
AIC	**AICc**	**BIC**	
1029.8	1030.24	1044.61	
Coefficients	**Estimate**	**STD Error**	**t-value**
ar1	-0.3373	0.0789	-42.750
T1-AR1 ($\delta.$)	0.0406	0.5185	0.0783
T1-MA0 (w_0)	-16.2339	85.616	-2.961
Intercept	0.1241	0.5448	0.2277

The Table 4.8 above report the parameter estimates of the full intervention model including the penalty statistics. The coefficient of the estimated parameter for the ar1 part significantly differ from zero since its test statistic (t-value) of -4.2750 is greater than 2 in absolute terms, whilst the estimated ar1 coefficient again strictly conforms to the bounds of parameter stationary since its value of -0.3373 is found to lie between -1 and 1.

The T1-MA0(w_0) and T1-AR1($\delta.$) denotes the intervention event (ie Community Policing) and decay or reduction event respectively. The estimate of the intervention event parameter of -16.2339 is interpreted as the magnitude of the impact of the intervention event. Its negative sign as expected indicate a reduction in the series as a result of the intervention effect. Specifically, it means that the introduction of the Community Policing in Ghana has been able to reduce the monthly occurrence of major crime categories in the Eastern Region by approximately an average of 16 cases. Its corresponding t-value of -2.961 indicates that this reduction is statistically significant since it is greater than 2 in absolute value.

The decay or reduction component of 0.0406 is however not statistically significant since its corresponding test statistic of 0.0783 is less than 2 in absolute value but satisfy the condition of system stability since it lies between -1 and 1. The long term effect which is given by the relation $Longterm = \frac{w_0}{1-\partial}$ is -16.921 is not quite different from its corresponding impact parameter of -16.2339 which has already been found to be statistically significant. It therefore can be inferred from the above conclusion that the long term effect of the intervention would be significant. Additionally, it is important to state that the penalty function statistics reported in terms of AIC, AICc and BIC with corresponding values of 1029.8, 1030.24 and 1044.61 respectively penalizes the fitted model based on the principle of parsimony. Thus the full intervention model which is a step function is thus given as $Y_t = \frac{-16.2339}{(1-0.0406L)} I_t + (0.0400 - 0.5863 y_{t-1})$

Table 4.9: Za (Zivot and Adrews) Test for potential break in data

Summary of test statistic			
Test type	**P-value**	**Test statistic**	**Critical values**
Za	0.00227	-3.6517	0.01= - 5.57 0.05= -5.08 0.1= -4.82
Potential break in data at position 37			

Zivot and Andrews test in Table 4.9 reported a potential break point at position 37 in the series where the intervention event took off. This obviously indicates that the onset of the Community Policing as a crime intervention in the Eastern region of Ghana was characterized by an immediate impact, hence the significant structural break at the year of onset (year 2003).

4.5.1 Diagnostic checks for the full Intervention model

The concluding part of the full intervention model is the diagnostics which tends to establish whether full model is well fitted or otherwise. This is achieved by the Ljung-Box Test as well as the residual plots as reported in Table 4.10 and Figure 4.7 respectively.

Table 4.10: Ljung-Box Test for the Intervention model

Summary of Test Statistic			
Test type	**X-squared (χ^2)**	**df**	**P-value**
Ljung-Box	13.4239	24	0.9586

The Ljung-Box Test as presented in Table 4.10 above tests the hypothesis below:

H0: The residuals are random or white noise

Ha: The residuals are not random

The results from Table 4.10 fails to reject the null hypothesis of white noise of the residuals at 5% significant level since the p-value 0.9586 is greater than the critical value of 0.05 and therefore follows that the fitted intervention model provides a good fit for the entire crime series.

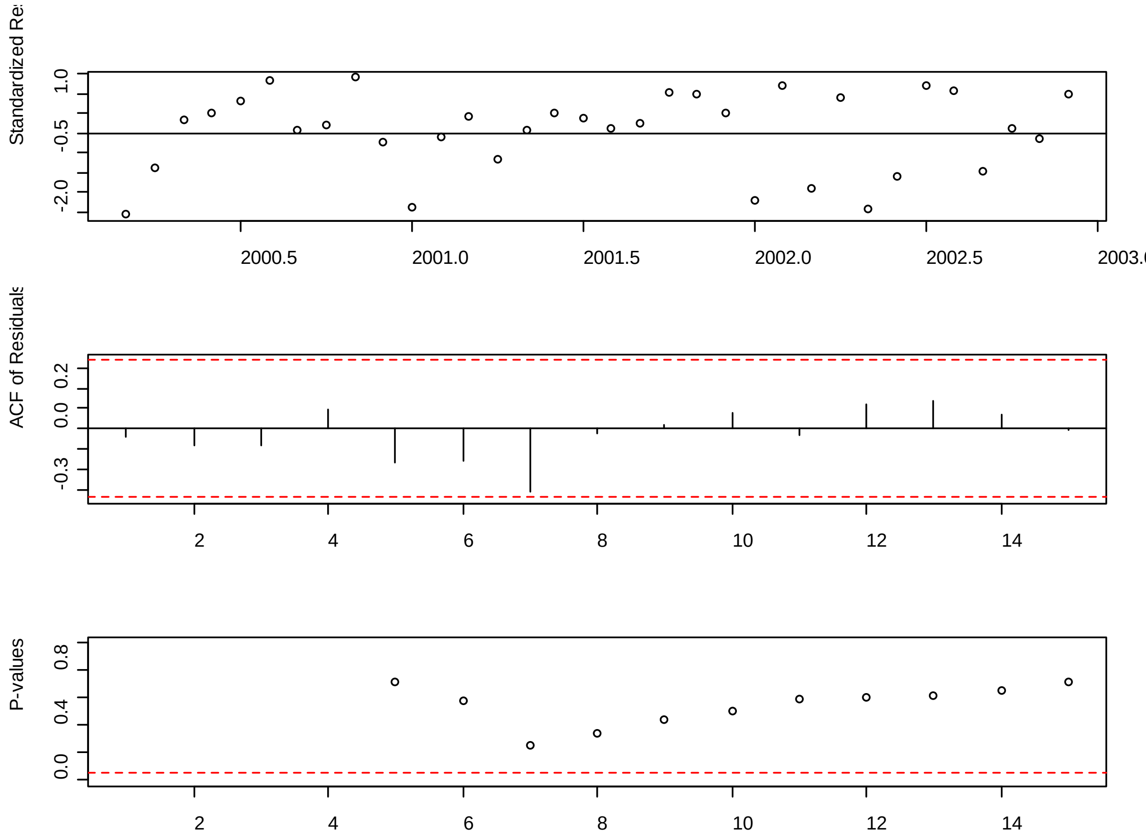

Figure 4.7: Residual plots of the full intervention model.

From the Figure 4.7 above it is clear that all the three (3) diagnostic residual plots do not show any anomalies for the fitted Intervention model. Clearly, the time plot of the residuals clearly shows that the residuals appear to be randomly scattered about zero, no evidence exists that the error terms are correlated with one another, there are no significant spikes in the ACF and PACF plots of the residuals, as well as the probability plot of the residuals also indicates that the individual probabilities of the residual are greater than or above 0.05. This shows that the residuals full intervention model is a white noise process, thereby corroborating with the conclusion of the Ljung-Box test.

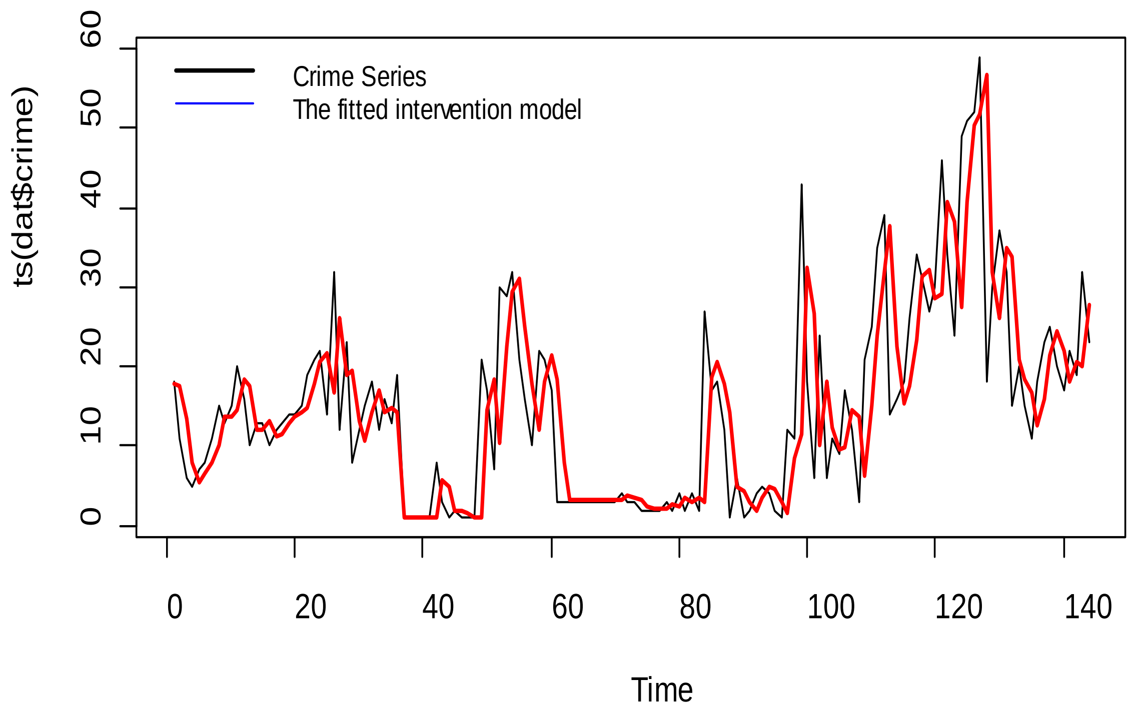

Figure 4.8: The fitted intervention model verses the actual crime series

The Figure 4.8 which shows the fitted intervention model as against the actual crime series indicates clearly that the full intervention model actually fits the data very well.

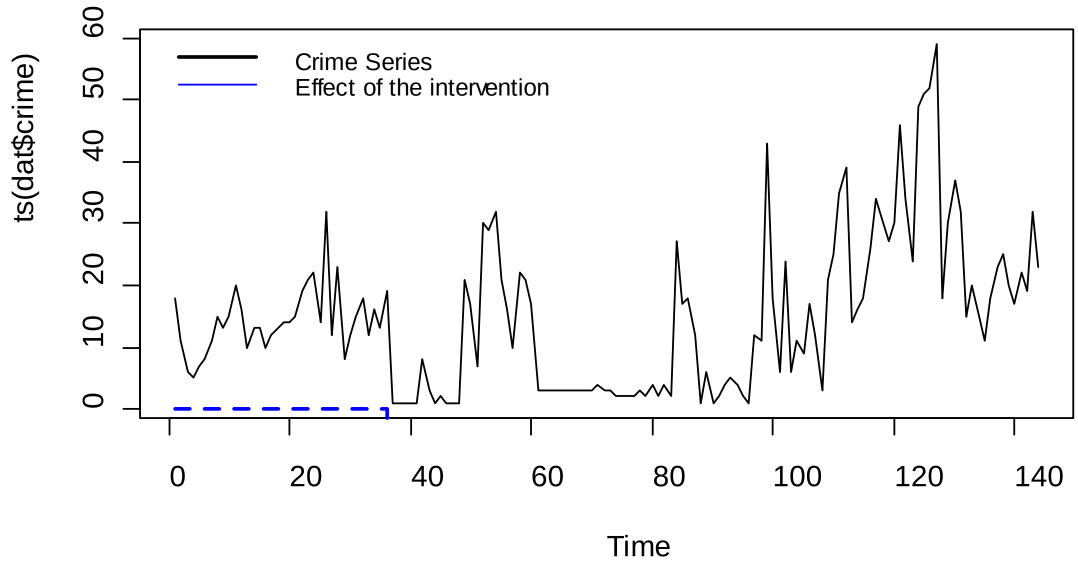

Figure 4.9: Graph of the Intervention Event

Finally, the of fitting a full intervention or impact assessment model is graphed as displayed in Figure 4.9 above with the thick black line indicating the crime series and the blue dash line representing the effect of the Community Policing event. The Community Policing programme exhibit a clear situation of a simple pulse function which tends to decay to zero and therefore representing an abrupt temporal effect of the intervention.

4.6. DISCUSSION OF RESULTS

Based on the analyses and interpretations of the appropriate secondary data on the major crime categories namely: murder, rape, defilement, robbery and the use and possession of drugs (cocaine, heroine, Indian hemp) obtained from the Regional Crime Unit of the Eastern Regional Command of the Ghana Police Service covering the period of year 2000 to 2011 so far, the results of the study are discussed as follows:

Firstly, the study revealed that the five major crime categories in the Eastern Region of Ghana were not constant but rather varied from one year to the other as well as from one month to the other with no systematically visible pattern, structural breaks, outliers, and no identifiable trend components in the time series data or non monotonous (as depicted in Figure 4.1 from the time plot of the crime series).

Secondly, from the analysis two models were developed for the crime series, the first one for the pre-intervention crime and the second one being the full intervention model. The best model for the pre-intervention crime series was an ar1(ARIMA 1,1,0) which was thus given as $Y_t = 0.0400 - 0.5863x_t$ with the residuals ε_t being white noise, whilst that of the full intervention was modeled as

$Y_t = \frac{-16.2339}{(1-0.0406L)} I_t + (0.0400 - 0.5863y_{t-1})$ also with residuals found to be white noise.

Also, empirical results further indicate that the intervention event was found to have had an impact (w_0) of reducing the crime cases over the period under study by a monthly average estimate of approximately 16 cases with a long term effect being -16.2339. The test of hypothesis on these findings failed to reject the null hypothesis of statistical significance of both the long term effect (-16.2339) and the overall effect (16 cases of reduction), as depicted in Table 4.8 above. This significant effect of the Community Policing Unit in crime prevention is consistent with the results obtained by (GSS, 2010), that the sampled general public gave regarding reasons for reporting crimes to the Police as well as crime preventing measures as revealed by the Victimization survey in Ghana which was carried out by the Ghana Statistical Service in collaboration with the United Nations Office on Drugs and Crime (UNODC) in 2009.

Again, the evaluation of the patterns and durations of the effect of the intervention revealed that the intervention effect was temporal but immediate and abrupt as it can be seen from the graph of the intervention event in Figure 4.9 above with the specific rate of decay (δ) estimated to be 0.0406 which was statistically found to be insignificant. This could be the reason why graph of the crime series as displayed in Figure 4.1 above reveals significant fluctuations in the series after the year 2008, despite the corresponding reduction effect in 2003 as a result of the intervention. The abrupt nature of the intervention were further corroborated by two significant tests, the corresponding absolute t-value of 1.9861 which is statistically significant (from Table 4.8) and most especially the Za test depicted in Table 4.9 which indicates a significant structural break in the series at position 37 which represent the point of onset of the intervention. This finding is consistent with the result that Chung, etal 2009 obtained from their study which sought to examine the impact of the global financial crisis on China's manufacturing industry. The temporary nature of the intervention effect could be attributed partly to the fact that the Community Policing Unit is a module in the National Youth Employment programme which requires the passing out of the system, recruitment and training of new members after every two year period.

Finally, the study also revealed that the full intervention model actually fits the data very well as it is clearly portrayed in Figure 4.8 above which shows the fitted intervention model as against the actual crime series. This finding is further corroborated by the hypothesis test of null hypothesis of significance (from Ljung-Box Test in Table 4.10) and subsequently by white noise of residuals from the diagnostics checks performed as portrayed in Figure 4.7 which clearly portray randomness of the standard errors of the residual, no significant spike in the residual plot of the ACFs and PACFs as well as the probability plot of residuals falling above the 0.05 line which confer significance of the model.

CHAPTER 5

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 INTRODUCTION

This final chapter of this study is based on the previous four (4) chapters and deals with the conclusions and recommendations from the study

5.2 CONCLUSIONS

Based on the objectives, as well as the literature reviewed, as well as the discussions made from the analyzes, it is desirable to draw the following conclusions:

Firstly, with regard to the first objective it can be concluded that the full intervention model was of the form $Y_t = \frac{-16.2339}{(1-0.0406L)} I_t + (0.0400 - 0.5863 y_{t-1})$ which is found to be significant and adequate by the Ljung-Box test and was further corroborated by diagnostics of the residual plots.

Secondly, it can be concluded that the introduction of the Community Policing unit of the Ghana Police Service as a crime combat intervention with regard to the Eastern region had a significant abrupt impact (w_0) of reducing the crime cases over the period under study by a monthly average estimate of approximately 16 cases with a long term effect being -16.2339.

Thirdly, the evaluation of the patterns and durations of the effect of the intervention revealed that the intervention effect was temporal with the specific rate of decay (δ) estimated to be 0.0406 which was statistically found to be insignificant.

Furthermore, it can be concluded based on the results that the full intervention model developed actually fits the data very well.

5.3 RECOMMENDATIONS

Based on the discussion of the results and conclusions drawn from the study, the following recommendations are worth considering:

Firstly, it is recommended that the Police administration should review the activities and operations of the current Community Policing Unit by restoring them back to their rightful positions in the communities rather than all of them being used as Motor Traffic Union (MTTU) officers in controlling motor traffic and apprehending offending motorists.

It is also recommended that the Bicycle Patrol Unit of the Community Policing Unit which came as a result of the realization that there is the need for the Police to collaborate with members of the communities, stake holders, chief, and opinion leaders in dealing with crime should be reinforced to help curtail community related crimes especially in the Eastern Region of Ghana.

Thirdly, it is recommended that the size of Community Police force should be increased by recruiting and training of enough officers in order to help meet the police to citizen ratio. This will go a long way to help in prevention of crimes that could be curtailed by police presence in communities.

Furthermore, it recommended that individuals who are well to do, corporate institutions, non governmental organizations and other benevolent organizations should partner the government with regard to funding, equipment and logistics by emulating exemplary style of the British High Commission who donated $50,000 worth of equipment such as handcuffs, shoulder guards, batons, baton holders, knee and elbow protective guards, shields, body armour, body bags for carrying dead bodies, helmets and groin guards to the Ghana Police Service to fully equip the Service to be able to embark on modern policing which is based on proactive policing rather than reactive policing.

Furthermore, the operations and activities of the Community Policing unit of the Ghana Police Service should be delinked from the National Youth Employment Program (NYEP) so as to curtail the interruption in the effective operation of the unit posed by recruitment and training of new personnels after every two year period.

Finally, it is strongly recommended that further intervention analysis be carried out in the other regions and possibly in the entire country with regard to the operations and activities of the Community Policing in combating the major crimes so as enable informed and intelligent decisions to be made on the basis of such analysis.

REFERENCES

Abdalla El-Habil (2006), Measuring the Impact of a New Policy on Industry in Gaza. An - Najah Univ. J. Res. (H. Sc.) Vol. 21(2).

Abdel-Aal RE, Mangoud AM. Modeling and forecasting monthly patient volume at a primary health care clinic using univariate time-series analysis. Computer Methods Programs Biomed 1998 Jun;56(3):235-47.

Amoakohene, M. I. (2004). Violence against women in Ghana: A look at women's perceptions and review of policy and social responses. Social Science & Medicine, Volume 59, Issue 11, Pages 2373-2385.

Appiahene-Gyamfi, J. (1998). Violent crime in Ghana: The case of robbery. Journal of Criminal Justice Volume 26, Issue 5, Pages 409-424.

Appiahene-Gyamfi, J. (2002). An analyses of the broad crime trends and patterns in Ghana. Journal of Criminal Justice, Volume 30, Issue 3, Pages 229-230.

Balkin, S. and J. K. Ord (2001). Assessing the impact of speed-limit increases on fatal interstate crashes, Journal of Transportation and Statistics : 1-26.

Bianchi, L., Jeffrey J. and Hanumara, R.C., (1998). Improving forecasting for telemarketing centers by ARIMA modeling with intervention. International Journal of Forecasting, Volume 14, Issue 4, Pages 497-504. http://www.sciencedirect.com/science.

Bloomfield, Peter, (1976). Fourier Analysis of Time Series: An Introduction. John Wiley and Sons, USA.

Blumstein, A. (2000), Disaggregating the Violence Trends. In Blumstein, A. and Wallman, J. (eds.), The Crime Drop in America. Cambridge University Press, New York.

Blumstein, A. and Wallman, J . (eds.), The Crime Drop in America, Cambridge University Press, New York, NY, pp. 1-12.

Bonham, S. C. and Gangnes Byron (1996). Intervention analysis with cointegrated time series: the case of the Hawaii hotel room tax. Journal of Applied Economics, 28, 1281 – 1293.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). Time Series Analysis, Forecasting and Control, 3rd ed. Prentice Hall, Englewood Clifs, New Jersey.

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

Box, G.E.P. and Tiao, G.C. (1975). Intervention analysis with applications to economic and environmental problems. Journal of American Statistical Association. 70, 70–79. http://scialert.net/fulltext/

Chang G.L, Chen C.H, and Carter E.C., (1993). Intervention analysis for the impacts of the 65 mph speed limit on rural interstate highway fatalities. J Safety Res 1993; 241:33–53.

Chatfield, C. (1996). The Analysis of Time Series, 5th ed., Chapman & Hall, New York.

Cho, V. (2001), Tourism Forecasting and its Relationship with Leading Economic Indicators. Journal of Hospitality and Tourism Research, 25(4), 399-420,ISSN: 10963480.

Chung, C.P Roy., Ip, W.H. and Chan, S.L., (2009). An ARIMA-Intervention Analysis Model for the Financial Crisis in China's Manufacturing Industry. International Journal of Engineering Business Management, Vol. 1, No. 1, pp. 15-18.

Chung, R. C. P., Ip, W. H. and Chan, S. L. (2008), Impacts of the Overheating Economy on China's Manufacturing Industry. The International Journal of Advanced Manufacturing Technology, DOI: 10.1007/s00170-008-1792-y, ISSN: 02683768(print) or 14333015(online).

Cleveland, W. S. (1993). Visualizing data, Summit. Hobert Press, New Jersey.

Cullen J. B. and Levitt S. D. (1996). "Crime, Urban Flight, and the Consequences for Cities," NBER Working Papers 5737, National Bureau of Economic Research, Inc.

Eck, J. and E. Maguire (2000), Have changes in policing reduced violent crime? An assessment of the evidence. In Blumstein, A. and Wallman, J. (eds.), The Crime Drop in America, Cambridge University Press, New York, NY, Pp. 207-265.

Enders, W., Sandler, T., and Cauley, J., (1990). Assessing the impact of terrorist-thwarting policies: An intervention time series approach. Journal of Defense Economics 2(1), 1–18.

Fafa, H. B. (2010). High crime rate in Ghana. Thursday, 11 November 2010, Jayee University College, Ghana.

Ferrand, Y., Kelton, C.M., Guo J.J., Levy M.S., and Yu, Y., (2011). Using time-series intervention analysis to understand U.S. Medicaid expenditures on antidepressant agents. Res Social Adm Pharm.7(1):64-80. http://www.ncbi.nlm.nih.gov/pubmed.

Ghana Police Service (2008). Crime Prevention. Designed and Hosted by MyGhanaonline.com.

Ghana Statistical Service (2010). Victimization Survey in Ghana (2009), Key Findings. Sprint City Co. LTD, Ghana.

Girard, D.Z., (2000). Intervention times series analysis of pertussis vaccination in England and Wales. Health Policy (2000) 54: 13-25.

Grove, A. S. (1996). Only the paranoid survive: how to exploit the crisis points that challenge every companyand career. Currency Doubleday, New York.

Hamilton, J. D. (1994), Time Series Analysis. Princeton University Press, New Jersey, USA.

Harvey, A. C., Koopman, S. J., Shephard, N. and J. A. Doornik (2000), Structural Time Series Analysis, Modeling, and Prediction using STAMP, Timberlake Consultants Press, London.

http://www.interpol.int/ Accessed, 12th September 2011.

Ip, W.H. (1997). Ruled-based ARIMA Model for FMS. Journal of Materials Processing Technology, 66(1-3), 240-243. ISSN: 09240136.

Ip, W.H., Fung, R. and Keung, K. W. (1999). An Investigation of Stochastic Analysis of Flexible Manufacturing Systems Simulation. International Journal of Advanced Manufacturing Technology, 15, 244-250, ISSN: 02683768.

Ismail, Z., Suhartono, A. Yahaya and Efendi, R., (2009). Intervention Model for Analyzing the Impact of Terrorism to Tourism Industry. Journal of Mathematics and Statistics 5 (4): 322-329, ISSN: 1549-3644.

Jenkins, G. M.and Watts,D G. (1968). Spectral Analysis and its Application. San Holden Day, Francisco.

Jennifer. C. H., Hsien-Hung, M. K. and Liu, H. H. (2010). Interventions affecting air transport passenger demand in Taiwan. African Journal of Business Management Vol. 4(10), pp. 2121-2131, ISSN: 1993-8233.

Lai S. L. and Li, W. L. (2005). Impact Analysis of September 11 on Air Travel Demand in the USA. Journal of Air Transport Management, 11, 455-458, ISSN: 09696997.

Lawson, M. (2008). Youth Crime in Ghana, Feature Article Sat, 04 Oct 2008. http://www.infoplease.com/ipa/A0107584.html. Accessed, 12th September 2011.

Ledolter, J. and Chan, K.S., (1996). Evaluating the impact of the 65mph maximum speed limit on Iowa rural interstates. American Statistician, 50, 79 – 85.

Levitt, S. D, (1996). "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation," The Quarterly Journal of Economics, MIT Press, vol. 111(2), pages 319-51.

Levitt, S. D. (1996), The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation, The Quarterly Journal of Economics, Volume 111, Issue 2, 319-351.

Ljung, G. and Box, G. (1978). "On a Measure of Lack of Fit in Time Series Models", Biometrika, 65, 297-303.

Makridakis, S., Wellwright, S. C. and Hyndman, R. J. (1998). Forecasting Methods and Applications. 3rd Edition. John Wiley, U.S.A.

Mensa-Bonsu, Henrieth, J.A.N (1999). The Annotated Criminal Code of Ghana Act 29/60. Third Edition. Black Mask Ltd, Ghana.

Min, C.H Jennifer (2008b). "Forecasting Japanese tourism demand in Taiwan using an intervention analysis", International Journal of Culture, Tourism and Hospitality Research, Vol. 2 Iss: 3, pp.197 – 216.

Min, C.H Jennifer, Hsien-Hung K., and Hsiang, L. Hsi (2010). Interventions affecting air transport passenger demand in Taiwan. African Journal of Business Management Vol. 4(10), pp. 2121-2131, ISSN: 1993-8233.

Muhammad, H. Lee, Suhartono and Bahrom Sanugi (2010). Multi Input Intervention Model for Evaluating the Impact of the Asian Crisis and Terrorist Attacks on Tourist Arrivals. MATEMATIKA, Volume 26, Number 1, 83–106.

Otchere, J. N. K. (2007). Increased Crime Rate In Ghana. Feature Article of Wednesday, 12 September 2007 San Diego State University.

Pankratz, A. (1983). Forecasting with Univariate Box-Jenkins Model: concepts and case, John Wiley and Sons, ISBN: 0471090239, New York.

R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.
Rock-Antoine M. and Hannarong Shamsub (2002). Who is benefiting the most from NAFTA? An Intervention time series analysis. Journal of economic development 69 volume 27, number 2.

Rusco W, Frank W.. and David Walls (2001). Lifting the Alaskan oil export ban: an intervention analysis. Energy Journal, vol. 22, pp. 81-94.
Sanjeev Sridharan, Westat Sunˇcica Vuji´c, Intervention time series analysis of crime rates: the impact of sentence reforms in Virginia. Tinbergen Institute Discussion Paper, Free University Amsterdam and Tinbergen Institute Siem Jan Koopman, Free University Amsterdam and Tinbergen Institute 2003.

Sathianandan, T.V, Somy Kuriakose , Mini, K. G and Joji, T.V., (2006). Impact of introduction of crafts with outboard engines on marine fish production in Kerala and Karnataka – a study through Intervention analysis. Indian Journal of Fisheries, 53 (3): pp. 271-282.
Shittu O.Ismail (2009). Modeling Exchange Rate in Nigeria in the Presence of Financial and Political Instability: An Intervention Analysis Approach. Middle Eastern Finance and Economics, ISSN: 1450 – 2889, Issue 5.

www.wikipedia.org. accessed on 15th April, 2012.

Yaacob, W.F, Husin, W.Z, Abd. Aziz, N., and Nordin, N. I., (2011). An Intervention Model of Road Accidents: The Case of OPS Sikap Intervention. Journal of Applied Sciences, 11: 1105-1112. http://scialert.net/fulltext/.

Zambon F., Fedeli U., Cristiana V., Marchesan M., Avossa F., Brocco S., and Spolaore P., (2007). Evidence-based policy on road safety: the effect of the demerit points system on seat belt use and health outcomes. Journal of Epidemiology and Community Health, Pp 877-881. http://jech.bmj.com/content/6.

APPENDIX

Appendix I: Monthly Crime statistics in the Eastern Region of Ghana from 2000-2012

	YEAR											
MONTH	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
January	18	10	14	1	21	3	3	17	12	21	46	20
February	11	13	32	1	17	3	2	18	11	25	34	15
March	6	13	12	1	7	3	2	12	43	35	24	11
April	5	10	23	1	30	3	2	1	18	39	49	18
May	7	12	8	1	29	3	2	6	6	14	51	23
June	8	13	12	8	32	3	3	1	24	16	52	25
July	11	14	15	3	21	3	2	2	6	18	59	20
August	15	14	18	1	16	3	4	4	11	26	18	17
September	13	15	12	2	10	3	2	5	9	34	30	22
October	15	19	16	1	22	3	4	4	17	31	37	19
November	20	21	13	1	21	4	2	2	12	27	32	32
December	16	22	19	1	17	3	27	1	3	30	15	23

Source: Regional CID Unit, Eastern Regional Command of Ghana Police Service (February, 2012)