Islamic University - Gaza
Higher Education Deanship
Faculty of Commerce
Department of Business Administration

الجامعة الإسلامية – غزة
عمادة الدراسات العليا
كلية التجارة
قسم إدارة الأعمال

Developing Forecasting Models for UNRWA Schools in the Gaza Strip

Submitted by

Issa Joma Hamdan

Supervisors

Prof. Dr. Yousif Hussein Ashour   Dr. Samir Khaled Safi

A Thesis Presented in Partial Fulfillment of the Requirement for
The MBA Degree

2010

Islamic University - Gaza
Higher Education Deanship
Faculty of Commerce
Department of Business Administration

الجامعة الإسلامية – غزة
عمادة الدراسات العليا
كلية التجارة
قسم إدارة الأعمال

Developing Forecasting Models for UNRWA Schools in the Gaza Strip

تطوير نماذج تنبؤ لمدارس وكالة الغوث في قطاع غزة

Submitted by

Issa Joma Hamdan

Supervisors

Prof. Dr. Yousif Hussein Ashour

Dr. Samir Khaled Safi

A Thesis Presented in Partial Fulfillment of the Requirement for
The MBA Degree

2010

Developing Forecasting Models for UNRWA Schools in the Gaza Strip

Abstract

This study aimed to develop forecasting models for number of schools, pupils, and classrooms of UNRWA in the Gaza Strip, where Gaza Strip has special circumstances, since the people has been living under strict siege for four years.

Gaza strip is the most crowded area over the world, where one and a half million people live over 365 square kilometers. In Gaza Strip, there is a crisis over every thing, even; land, water and food. So planning and forecasting are the right approach to face such obstacles and challenges.

The main approach for analysis was statistically based methods. ARIMA (Autoregressive integrated moving average) models were chosen to forecast number of schools, pupils, and classrooms of UNRWA in Gaza Strip.

To get more accurate, useful and practical results, Gaza field was divided into six areas. Each area has seven variables, namely, number of schools, classrooms for elementary stage, boy classrooms in preparatory stage, girl classrooms in preparatory stage, pupils in elementary stage, boy pupils in preparatory stage and girl pupils in preparatory stage. Indeed, 42 time series were modeled.

For each time series among 42 time series, ARIMA models with different orders, which were tentatively chosen depending on ACF and PACF plot, were applied, then the best fit model which has the best AIC, BIC and accuracy measures values, was chosen.

Results indicated that ARIMA models were the most appropriate method to fit the considered data, where the results were accurate and satisfied the required criteria.

Modeling by other forecasting methods can be conducted by future researchers in addition to studying the effects of UNRWA annual budget on the growth of number of schools and classrooms.

تطوير نماذج تنبؤ لمدارس وكالة الغوث في قطاع غزة

ملخص الرسالة

تهدف هذه الدراسة لتطوير نماذج تنبؤ لعدد المدارس والطلاب والصفوف الدراسية لوكالة الغوث في قطاع غزة. فقطاع غزة لديه خصوصية حيثُ يعيش سكانه تحت حصار إسرائيلي مشدد منذ أربع سنوات.

قطاع غزة هو المنطقة الأكثر ازدحاما في العالم حيث يعيشُ مليون ونصف من الناس علي 365 كيلو متر مربع لهذا فان كل شيء محدود الأرض والمياه وحتى الغذاء. ولهذا فالتخطيط و التنبؤ هما الاتجاه الصحيح لتخطي المعوقات والتحديات.

لقد تم استخدام المنهجية الإحصائية التحليلية في البحث، وقد تم اختيار نماذج الارتباط الذاتي والمتوسط المتحرك المتكاملة للتنبؤ بأعداد المدارس والطلاب والصفوف الدراسية لوكالة الغوث في قطاع غزة.

للحصول علي نتائج أكثر دقة وفائدة وموضوعية تم تقسيم قطاع غزة إلي ستة مناطق وكل منطقة بدورها احتوت سبعة متغيرات عدد المدارس و عدد الصفوف للمرحلة الابتدائية وعدد صفوف الإعدادية للذكور وعدد صفوف الإعدادية للإناث وعدد طلاب الابتدائية وعدد طلاب الإعدادية الذكور وعدد طلاب الإعدادية الإناث. أي أننا قسمنا البيانات إلى 42 سلسلة زمنية.

تم تطبيق مجموعة من نماذج الارتباط الذاتي والمتوسط المتحرك برتب مختلفة علي كل السلاسل الزمنية تحت الدراسة، وقد تم اختيار الرتب لهذه النماذج بفحص رسوم دالة الارتباط الذاتي والارتباط الذاتي الجزئي للسلسلة تحت الدراسة، وبعد ذلك تم اختيار أفضل نموذج والذي يتميز بأفضل قيم لمعيار الإعلام الذاتيAIC ومعيار بايزينBIC ومعايير الدقة.

أظهرت النتائج أن نماذج الارتباط الذاتي والمتوسط المتحرك المتكاملة هي الطريقة الأكثر مناسبة للبيانات تحت الدراسة، حيث أن النتائج كانت دقيقة وتحقق المعايير المطلوبة .

ننصح الباحثين في المستقبل أن يدرسوا استخدام طرق تنبؤ غير المستخدمة في هذه الدراسة بالإضافة لدراسة تأثير الموازنة السنوية للوكالة على عدد المدارس والصفوف الدراسية.

Acknowledgments

I would like to express deep sense of gratitude to all who contributed to achieve my goal and complete this Master thesis 'Developing Forecasting Models For UNRWA Schools In Gaza Strip'.

I would like to acknowledge my co-supervisors Dr. Samir Khaled Safi and Prof. Dr. Yousif Hussein Ashour, who guided me throughout the thesis, I also would like to acknowledge the examiners Dr. Salah R. Agha and Dr. Mohammad I. Megdad For their valuable suggestions and beneficial notices.

I would like to thank Islamic University staff for their assistance. I also would like to acknowledge the UNRWA Education Programme for providing me with relevant UNRWA data.

Last, but not least, thank you Allah for giving me strength to continue work under daily life pressures.

Dedication

To,
My Mother,

For her magnificent devotions to her family

My Father Soul,

For his directions and support

My Family Member,

For her support, understanding and prayers

List of Tables

Table **Title** **Page**

List of Figures

Figure **Title** **Page**

Abbreviations

ACF	Autocorrelation Function
AIC	Akiak's Information Criterion
ARIMA	Autoregressive Integrated and Moving Average
ARMA	Autoregressive and Moving Average
BIC	Bayesian's Information Criterion
C E	Number of Elementary Classrooms
C P B	Number of Preparatory Boy Classrooms
C P G	Number of Preparatory Girl Classrooms
C S	Number of Schools
Gaza	Gaza Governorate in Gaza Strip
Khan	Khanyounis Governorate in Gaza Strip
Khan E	Eastern area in Khanyounis Governorate
Middle	Middle Governorate in Gaza strip
MoEHE	Ministry of Education and Higher Education
Northern	Northern Governorate in Gaza Strip
P E	Number of Elementary Pupils
P P B	Number of Preparatory Boys Pupils
P P G	Number of Preparatory Girls Pupils
PACF	Partial Autocorrelation Function
PCBS	Palestinian Central Bureau Statistics
PHC-2007	Population, Housing and Establishment Census 2007
Rafah	Rafah Governorate in Gaza Strip
UN	United Nation Agency
UNCTAD	United Nation Conference on Trade and Development

UNHCR Office of the United Nations High Commissioner for Refugees

UNRWA The United Nations Relief and Works Agency for Palestine Refugees in the Near East

Glossary of Terms

Extrapolation Extrapolation is when the value of a variable is estimated at times which have not yet been observed. This estimate may be reasonably reliable for short times into the future, but for longer times, the estimate is liable to become less accurate.

Full double shift system Two different schools use the same school premises on full double shift (i.e. one in the morning and one in the afternoon).

Gaza Field UNRWA agency in Gaza Strip.

Model Selection Choosing the parametric family to use for estimation of parameters for proposed model.

Over-Differencing using differencing with higher order and introduces unnecessary correlations into a series and will complicate the modeling process.

Over-Fitting occurs when a statistical model describes random error or noise instead of the underlying relationship. Over-fitting generally occurs when a model is excessively complex, such as having too many degrees of freedom.

Population Growth Rate the change in population over time.

Developing Forecasting Models for UNRWA Schools in Gaza Strip

Chapter One: Research Background

1. Research Background

1.1. Introduction

"Education is our investment" that what Ms. Mahasin Muhesen, the chief of field education programme recognized educational process in Palestinian mind.

So a five year plan was developed by MoEHE in 2000 and, in early 2002, the ministry began to develop an education for all plan the five-year remained the guiding instrument for developing the Palestinian education sector (Nicoli, S., 2006, p. 8).

The Palestinian people had been living exceptional situation from 1948 till now. In year 1948, 480000 of Palestinians were uprooted from their homes and land, after that they were Palestinian refugees.

After one year, in 1949 United Nation Agency (UN) established the United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA) according to 302 (IV) resolution, to provide humanitarian relief and education, health, social services and emergency assistance (UNHCR, 2007).

Education had become an important need for Palestinian people because they lost their income resources from agriculture, trade and industry, after they were uprooted from their homes and lands, so Palestinian families paid their basic needs to save good education for their sons.

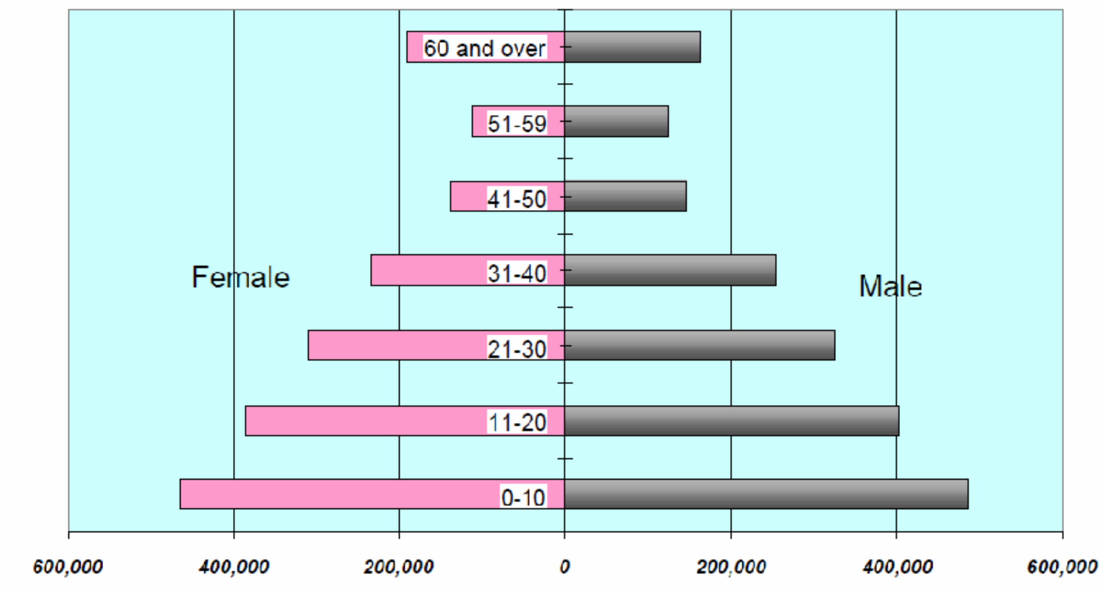
Al Zaroo and Gillian (2003, pp. 165-188) recorded responses of group interview to identify the education, the majority of participants identified education as their first priority. So in 2007, drop out rate for Palestinian schools in Gaza Strip and West bank was 0.8 % (PCBS census, 2007).

Education program is the heart of UNRWA activities, it currently accounts for over 59 % of the agency's budget and more than three quarters of its staff (UNRWA, 2007a).

Gaza Strip is a special case, now it has the highest level of density in the world. One and a half million people live within 365 square kilometers (PCBS census, 2007). At the same time, in 2007, the annual growth rate of registered Palestinian refugees in Gaza Strip was more than 3 %, around

56% of them are under 25 years of age. Fig. 1.1 explains more details, (UNRWA, 2007c).

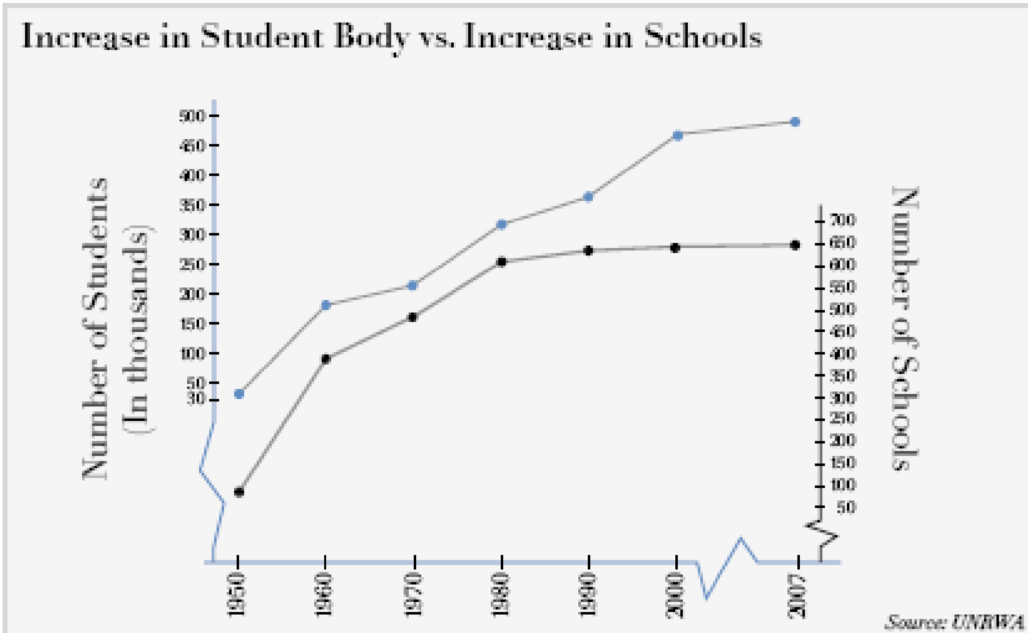
Figure 1.1 Distribution of refugee population by age in Gaza Strip



These situations pushed the managers or controllers to be efficient and to think carefully in the question "what are the necessary future needs and resources".

In fact, educational environment, in UNRWA schools in Gaza Strip, is highly insufficient, as 77% of schools work in full double shift system and the schools have the highest pupil per class ratio (27.6) comparing to other UNRWA fields. Fig. 1.2 illustrates the development for number of students and schools over time (1950-2007) (UNRWA, 2007d).

Figure 1.2 Development of UNRWA students and schools over time (1950-2007)



1.2. Research Problem

The main question that was addressed in this study was: what will be the future number of schools, pupils, and classrooms for UNRWA schools in Gaza Strip. This question included sub questions, derived from the main question

- What is the general trend for the development of the number of schools, pupils, and classrooms?
- What is the best model to describe the future development of schools, pupils, and classroom?

1.3. Research Objectives

The main objective of this study was building quantitative model which can be used to forecast the number of schools, pupils, and classrooms in UNRWA schools in Gaza Strip.

1.4. Research Variables

Dependent variables: - number of UNRWA schools, pupils, and classrooms in Gaza Field. Schools were divided only according to area, where pupils and classrooms were divided according to area for elementary stage and according to area and six for preparatory stage.

Independent variable: - the time, the time scale was on a yearly basis

1.5. Research Importance

Most competition during election campaigns concentrate on education. All individuals care about education where they will be influenced by its policy, so education is the most sensitive sector in any country.

The first step of success is developing and improving education field. For example economy one of these field, depends on education for success, so any economical development plan must be associated with educational development plan, as educational organizations are the incubator for leaders and decision makers.

Providing aiding tool, which can be used to forecast the future number of schools, pupils and classrooms, is an important issue, as the decision makers will use it to forecast future values, then take necessary actions and yield the required resources to avoid troubles and problems in education sector.

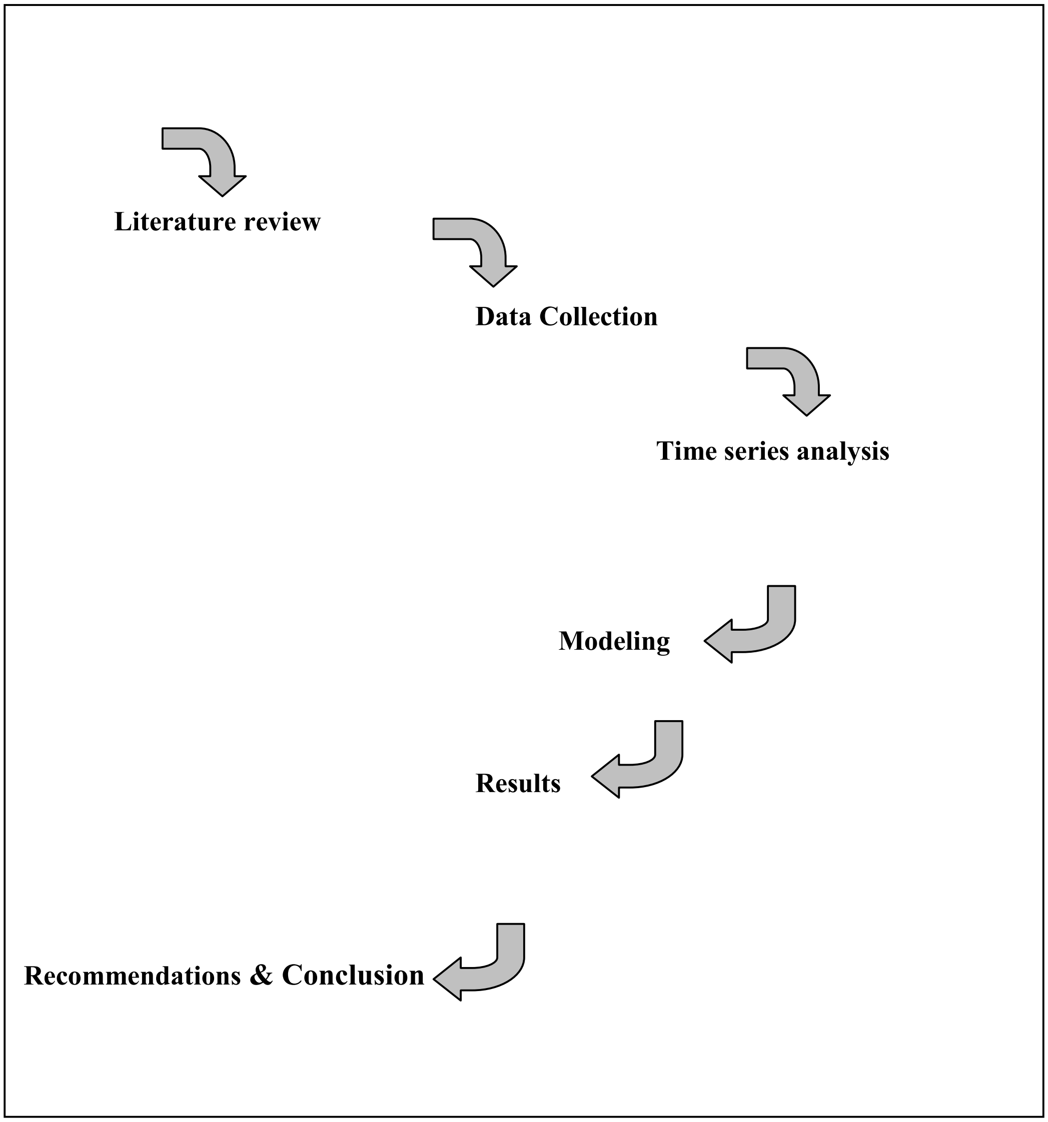
In addition to developing models to forecast number of schools, pupils and classrooms, the study will help decision makers to develop policies and strategies to improve the education process outcomes.

Forecasting models support managers in planning mission so they pay attention to develop forecasting model, as a good planning guide to a good and an efficient management.

1.6. Research Methodology

This study comprised of several stages and each stage involved sub stages, these stages are outlined in Fig. 1-3.

Figure 1.3 Stages of study



Statistical based method was used to forecast the future number of UNRWA schools, pupils and classrooms, more specifically, time series analysis.

Time series data is the data that is collected or observed over successive increments of time (Jordan, 1985, p. 377), collected from its original sources, education field office in Gaza Strip.

Model development: several approaches were used for model fitting, and then the best fit model was chosen.

Graphical approach: graphical approach was used to reveal important characteristics of the series, it was used to determine general trend of the series.

ARIMA approach: Autoregressive Integrated Moving Average (**ARIMA**) uses past value of dependent variables as input in addition to usage of moving average.

Software package tool: S**-plus and Minitab** packages were used to conduct quantitative data analysis.

1.7. Research Structure

This study was divided into five chapters as follows:
Chapter one: Research Background
Chapter two: Theory of Forecasting and Modeling
Chapter three: Time Series Models
Chapter four: Model Inference and Forecasting
Chapter five: Results and Analysis
Chapter six: Conclusions and Recommendations

1.8. Scope and Limits

This study was limited to:

- Schools, rented or owned by UNRWA in Gaza either elementary or primary level.
- Pupils, enrolled formally in UNRWA schools in Gaza
- Classroom, rented or owned by UNRWA in Gaza

Historical time series data was annual observation along 14 years. Time series data was different for schools, pupils and classrooms.

1.9. Previous Studies

In dynamic life, like nowadays, forecasting is considered a basic tool for any managerial operation, though forecasting has a great part of global searching movement.

However, there was no any study concerned of forecasting over education in Gaza Strip. So the decision makers in Gaza had been using constant increasing percentage to forecast educational variables, these forecasts lacks the accuracy and reliability.

The following studies used different forecasting techniques in education field.

1. Chen, Chau-Kuang (2008): ***An Integrated Enrollment Forecast Model***, This paper illustrated the development of the integrated enrollment forecast model for Oklahoma State University enrollment series from Fall 1962 to Fall 2004. The two best models generated by ARIMA and linear regression methods fitted the data exceptionally well with high R- squared values of 0.96 and 0.97, respectively. Both models also forecasted highly accurate Oklahoma State University enrollment with MAPE values of 2.11% and 1.62%, respectively. The best linear regression model outperformed the best ARIMA model, ARIMA (1, 1, 0), for the turning points in 1983 and 1995. On the other hand, the best ARIMA model demonstrated more accurate forecasts than the best linear regression model in years 1972 and 1989. However, there was no significant or practical mean difference in the absolute percentage errors between the two models. The integrated enrollment forecast model had demonstrated its model validity and forecasting accuracy. Hence, it can be replicated and may well be useful for estimating aggregated student enrollment in other similar institutions.

2. AlHendi (2007): ***Developing a Forecasting Model for Mobile Users in Palestine***, The main objective of this research was to build and operate a diffusion model present the Palestinian telecommunication market which experiences unusual circumstances and faces a series of obstacles and uncertainties. The author used Logistic Model to estimate the market potential of cellular mobile users.

3. Aljabre, Biome, Almhesin (2004): ***Forecasting the Future of Education in the District of Almadinah: Applying Time Series Analysis,*** This study aimed to predict number of schools, pupils, teachers and yearly expenditures for Almadina Area in the Saudi Kingdom. The study used descriptive and time series analysis to forecast values of concerned variables. The study applied linear smoothing and exponential models, and then made comparison between these methods according to Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) accuracy measures. The results

indicated that concerned variables will be increased over 10 future years with different percentages.

4. Ahlburg, McPherson, Schapiro (1994): ***Predicting Higher Education Enrollment in the United States: an Evaluation of Different Modeling Approaches,*** The purpose of this paper was to assess the state of the art in model-based enrollment prediction for U.S. higher education. The study concluded that combining the results from disaggregated forecasting models and trying alternative approaches is a much better option for predicting higher education enrollments than searching for a universal model that works for all groups at all times.

5. Bernhardt,V. (1983): ***Seattle's Small-Area Approach to Forecasting Enrollments at the School Level***, This paper described a procedure that combined forecasting of enrollments and management of facilities. The Seattle system prepared its forecasts for a relatively small local unit called the "small planning unit", and then five-year projections were prepared for each small area, and were then aggregated to prepare forecasts for large geographical areas. The procedure applied to forecast the variable was a modification of the traditional grade progression or holding power or cohort survival. This procedure had been successfully implemented for the Seattle, Washington, Public Schools.

6. Frabkel, Gerald (1982): ***projections of Education Statistics to 1990-91***, this study aimed to cover all enrollment at all educational levels, numbers of high school graduates and earners of higher education degrees, numbers of instructional staff, and educational expenditures at all levels. Projections of enrollments in elementary and secondary schools were based on a grade-retention or cohort-survival method where all enrollment data were shown by organizational level control. Higher education enrollment data were controlled for the additional variables of student age, sex, and attendance status, two or four year program.

Chapter Two: Theory of Forecasting and Modeling

2. Theory of Forecasting and Modeling

2.1. Introduction

Forecasting is a daily skill; every body in the world does it. People use their intuitive to predict, what will happen in their next days. But in business forecasting is a systematic process, controlled by procedures and controls.

Managers in all fields of business, continuously need to make decisions, these decision may lead to success or fail for their organization, may cause of loose millions of dollar. So managers use forecasting techniques to make right decision and decrease fail probability. Now a day managers recruit specialist employees to help them in forecasting process (Hanke, Reitsch, 1991, p. 698).

2.2. Forecasting

Forecasting is a hybrid science, so the forecaster needs to get many of disciplines like mathematics, statistics and social science. But the main question of forecasting how the things will develop or run in the future (Stevenson, 1989, p. 424).

Although the high technology especially in computer and programming field, forecasting still depends on personal judgment of manager, so forecasting is difficult and complicated process.

All types of forecasting depend on past data to predict future events, as more precision of historical data as higher precision of forecasting process outputs, so successful business organizations keep and organize data to analyze it in the future and get data which may guide to the right decision (Chatfield, 1996, p. 13).

2.2.1 Scientific Forecasting

It has been seen that forecasting is an essential and inherent part of business activity so there is amass need to use scientific method for forecasting to avoid or decrease the percent of error in forecasting process outputs. Now will display some of the steps which are involved in any scientific forecasting process as described by Warren Gilchrist (1976)

The first step: data Collection

Historical data, in forecasting process, look like raw materials in industry field, so before any forecasting process, the data relevant to a considered problem must be gathered. There are different tools involved in data collection process like questionnaires and records.

The second step: data reduction

The aim here is to filter and cluster relevant and important data from the vast amount of information obtained from the first step.

The third step: model construction

Now, the structure and trend of the reduced data over a range of time must be examined, the structure of data enables the forecaster to build mathematical or statistical model to describe the situation in the past. The model can be used to forecast events in the future either by values or other means.

The fourth step: model extrapolation

Any problem which seems simple may involve many of variables, so the forecaster extrapolates the right model depending on his past experience. The forecaster may ignore some variables which do not have tangible effect on the precision of the results according to his estimation, so model extrapolation process need high skill and experience.

2.3. Forecasting Methods

There are different methods for forecasting which can be classified under two main methods, qualitative methods and quantitative methods

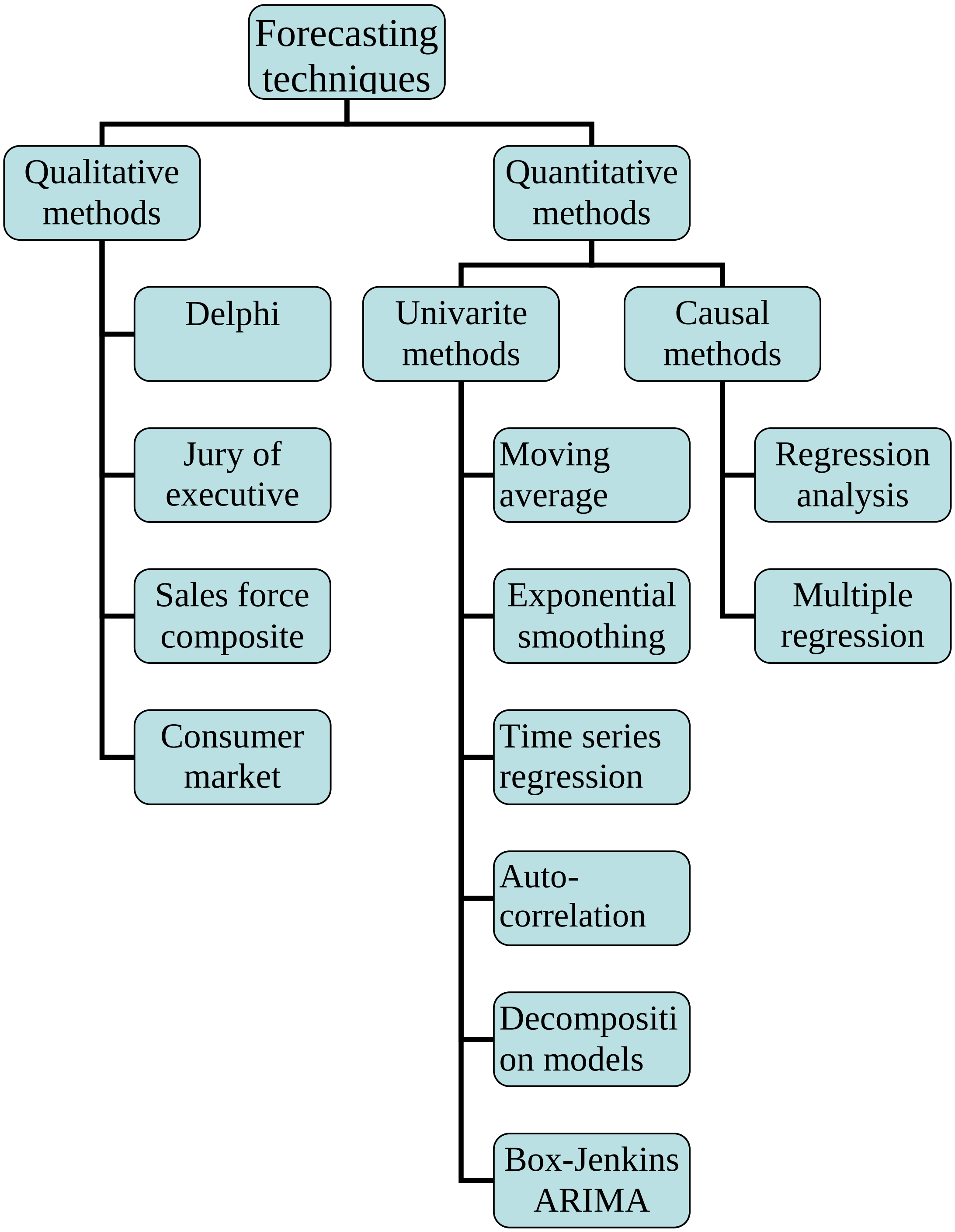
2.3.1. Qualitative Methods

Qualitative methods can simply distinguish by their outcomes, as the outcomes not numerical values but descriptive. Qualitative methods are simple and easy and the forecaster has no need for historical data. It can be used in simple problems, primary evaluation or when historical data is unavailable (Bowerman, 2005, p.7).

2.3.2. Quantitative Methods

These methods use historical data to predict the future events, so historical data must be available. The outcomes of these methods are numerical values. Quantitative methods can be classified into two main techniques Causal technique and time series technique. The figure 2.1 illustrates the main quantitative and qualitative methods.

Figure 2.1 Tree of forecasting techniques



Source:(Render, 2006, p. 151)

2.3.3. Causal Method

This method involves identifying the variables related to the variable to be predicted. After identifying the relevant variables, a statistical model that describes the relationship between the variable to be predicted and related variables, is developed. This method is complicated; however, it produces accurate forecasting (Render, 2006, p. 151).

2.3.4. Forecasting Accuracy

Each forecasting process includes some ration of uncertainty; any improvement on forecasting process will lead to more accurate results, however. Will not make the process 100% certainty.

2.4. Regression Analysis

Regression analysis technique is one of causal methods. A variable of interest called the dependent variable and other variable called independent variable.

If the independent variable is the time, the data is called time series data and the analysis called time series regression, in other hand if there is only one independent variable, the analysis called simple regression analysis otherwise called multiple regression analysis (Hanke, Reitsch, 1991, p. 450).

2.4.1. Simple Linear Regression Model

The symbol for dependent variable is Y, and symbol X for independent variable. Scatter diagram display the X-Y relationship in graphic term (Hanke, Reitsch, 1991, p. 451). The next equation used to draw straight line through the data points of the scatter diagram.

$$Y = \beta_o + \beta_1 X + \varepsilon$$

Where β_o = Y-intercept
β_1 = slope of the line
ε = random error component

One of the major aims of regression analysis is forecasting values in the future, the previous equation can be used to predict values of variable Y under changes of variable X.

2.4.2. Simple Coefficient of Determination

The simple coefficient of determination "r^2" is a measure of usefulness of the simple linear regression model (Bowerman, 2005, p.114).

The simple prediction for any dependent variable is a mean value of this dependent variable. This coefficient is a signal of the usefulness of using simple linear regression model for prediction substitute of simple prediction "mean value".

$$r^2 = 1 - \frac{SSE}{TSS}$$

Where r^2 = the coefficient of determination
TSS = total sum of squares
SSE = sum of squares error

TSS is the sum of the difference between the observed value Y_i and the mean value over the historical data $\bar{Y}$, but SSE is the sum of the difference between the observed value Y_i and the predicted value $\hat{Y}$ using the simple regression equation.

2.4.3. Simple Coefficient of Correlation

It is useful to know, that the regression equation can be used to forecast the values of variable of interest beyond the original interval, coefficient of correlation gives us a signal of the validity and precision of regression equation to forecast outside original interval. Coefficient of correlation is one measure of the relationship between two variables (Stevenson, 1989, p. 448).

$$r = \pm\sqrt{r^2}$$

Where r =the coefficient of correlation

The negative sign refers to reverse relationship between the variables, but the positive sign refers to direct relationship between the variables.

2.4.4. Multiple Regression Analysis

In many cases, independent variable Y is affected by two or more variables. However, the forecasting process will be more difficult, but can be done by the following equation (Hanke, Reitsch, 1991, p. 496).

$$Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \Lambda + \beta_k X_k + \varepsilon$$

Where Y = dependent variable
$X_{1,} X_{2,} \Lambda , X_k$ = independent variables
$\beta_{o,} \beta_{1,} \beta_{2,} \Lambda , \beta_k$ = parameters

ε = random error component

2.5. Univariate Method

Time series data are historical values of a variable that had been recorded at periodic interval (Stevenson,1989, p. 428).

When univariate techniques are used, historical data analyzed to identify a data pattern. Then, in assumption that it will remain stable in the future without impact change; this data pattern is extrapolated in order to produce forecasts (Bowerman, 2005, p. 23).

Selecting a suitable technique to use it with time series data requires knowledge of the pattern of time series. Therefore in the following pages, the time series techniques which may be used in real life will be described.

2.5.1. Averaging Models

The best usage for averaging techniques when the time series aggregate or vary around an average. Although data series may fluctuate around an average value, the averaging techniques smooth out some of the up and down fluctuations in the series. Smoothing out can be useful in many cases, as these fluctuations represent random changes in series (Stevenson, 1989, p. 702).

Moving Averages

The Moving average model uses the average of several past time periods as the forecast to the next period. In practice, the forecaster must decide how many past periods to average. A trial-and- error is often used to find the number of periods that would be best in minimizing the error (Hanke, Reitsch, 1991, p. 703).

The following equation is used to compute moving average forecasts.

$$F_{t+1} = \frac{\left(Z_t + Z_{t-1} + Z_{t-2} + \Lambda\ \Lambda\ \Lambda\ \Lambda\ + Z_{t-m+1}\right)}{m}$$

Where F_{t+1} = Forecast for time period t+1

Z_t = Z value for time period t

m = Number of terms in the moving average

2.5.2. Exponential Smoothing

The disadvantage of moving average method that each time period has the same effect on forecasting value, but in real life modern period has an affect on future forecasting more than old period.

Exponential smoothing method uses weighted average of past time series value to arrive at a smooth forecast (Hanke, Reitsch, 1991, p. 703).

Exponential smoothing provides a forecasting method that is most effective when the components (trend and seasonal factors) of the time series may be changing over time (Bowerman, 2005, p. 345).

Simple Exponential Smoothing

The next equation is a simple exponential smoothing formula which can be used when there is no trend or seasonal pattern but the mean of the time series is slowly changing over time (Chatfield, 1996, p. 103).

$$F_{t+1} = \alpha\ Z_t + (1-\alpha)\alpha\ Z_{t-1} + (1-\alpha)^2\alpha\ Z_{t-2} + \ldots + (1-\alpha)^{t-1}\alpha\ Z_1 + (1-\alpha)^t F_1$$

Where F_{t+1} = Forecast for time period t+1

Z_t = Y value for time period t

α = smoothing constant, a value between 0 and 1

F_t = average experience of the series smoothed to period t, or forecast value for period t

Holt's Trend Corrected Exponential Smoothing

When the time series display a linear trend, simple exponential smoothing formula will be inappropriate for forecasting and need adjustments to be convenient with this case.

The following formula "Holt's trend corrected exponential smoothing" can be used when both mean and growth rate are changing over time with no seasonal pattern, by adding or subtracting an amount to each forecast reflecting the linear trend (Bowerman, 2005, p. 357).

$$F_{t+1} = \alpha\ Z_t + (1-\alpha)[F_t + b_t]$$
$$b_{t+1} = \gamma[F_{t+1} - F_t] + (1-\gamma)b_t$$

Where b_t = forecast of growth rate for period t

γ = smoothing constant for growth rate

2.6. Time Series Regression Models

Such models correlate the dependent variable Z_t to the time (Shumway, 2006, p. 48). These models can be used when dependent variable in time series data remains constant over time.

2.6.1. Trend Model

Trend may be linear or quadratic or no trend or any other type. However, the trend will be constant over time. The following model can describe a time series variable Z_t over time (Curwin, Slater, 2000, p. 425).

$$Z_t = TR_t + \varepsilon_t$$

Where Z_t = the value of the time series in period t

TR_t = the trend in time period t

ε_t = the error term in time period t

In the following section, the most realistic cases of trend will be considered.

No trend

$$TR_t = \beta_o$$

The model refers that there is no long move up or move down.

2.6.1.1. Linear Trend

$$TR_t = \beta_o + \beta_1 t$$

The model implies that the variable of interest changes according to straight line. The line moves up over time, if the slope of line β_1 is greater than zero and moves down if the slope is less than zero. The most common method used to find the parameters of the line (β_o, β_1) is least square method (Curwin, Slater, 2000, p. 327).

2.6.1.2. Quadratic Trend

$$TR_t = \beta_o + \beta_1 t + \beta_2 t^2$$

The model implies that there is a quadratic change over time.

2.6.1.3. Polynomial Trend

$$TR_t = \beta_o + \beta_1 t + \beta_2 t^2 + \Lambda + \beta_p t^p$$

The model implies third order ($p = 3$) or higher ($p > 3$) polynomials model

2.6.2. The least Square Method

The least square method is a mathematical procedure to find and estimate the parameters of the line (β_o, β_1), which minimize the sum of square of error terms **SSE** (the sum of squared errors).

The method of least square determines the equation for the straight line that minimizes the sum of the squared distances between the line and the data values as measured in the Z direction (Hanke, Reitsch, 1991, p. 453).

$$e_t = Z_t - \hat{Z}_t$$

$$SSE = \sum_{t=1}^{n} e^2 = \sum_{t=1}^{t=n} (Z_t - \hat{Z}_t)^2$$

Where Z_t =observed value of time period t

$\hat{Z}_t$ =estimated value of time period t

$$\beta_1 = \frac{n\sum_{t=1}^{t=n} tZ_t - (\sum_{t=1}^{t=n} t)(\sum_{t=1}^{t=n} Z_t)}{n\sum_{t=1}^{t=n} t^2 - (\sum_{t=1}^{t=n} t)^2}$$

$$\beta_o = \frac{\sum_{t=1}^{t=n} Y_t}{n} - \frac{\beta_1 \sum_{t=1}^{t=n} t}{n}$$

Where Z_t = the value of the time series in period t

t = time period t

n = number of time period t

2.6.3. Time Series Regression Assumption (Bowerman, 2005, p. 238-242)

Under these assumptions, a model can be easily developed with acceptable accuracy without need to other restricted assumptions.

1. At any given time period t, the population of errors has a mean equal to zero.

2. **Constant variance assumption**. At any given value of time period, the population of potential of residuals has a variance that does not depend on the value of time period t. that is; the different population residuals corresponding to different values of time period have equal variances.

3. **Normality assumption**. At any given value of time period t, the population of potential residuals has a normal distribution.

4. **Independence assumption**. Any one value of residual e is statistically independent of any other value of e. That is, the value of error e corresponding to an observed value of Z_t is statistically independent of the value of the residual corresponding to any other observed value of Z_t.

,

2.7. Autocorrelation Analysis

The independence assumption says that error term occurs in a random pattern over time. When a time series regression model is built to forecast a future value of the variable of interest, it is assumed that the independency assumption is satisfied. However, in real life, this assumption is often violated in many cases and it is common for the error term ε to be auto-correlated.

Autocorrelation is the correlation between variable, lagged one or more period, and itself (Hanke, Reitsch, 1991, p. 655).

2.7.1. Positive Auto-Correlation

Positive auto-correlation exists if a positive error term in time period t tends to be followed by another positive error in time period t + k (a later time period) and if a negative error term in time period t tends to be followed by another negative error in time period t + k(a later time period) (Bowerman, 2005, p. 288).

2.7.2. Negative Auto-Correlation

negative auto-correlation exists if a positive error term in time period t tends to be followed by another negative error in time period t + k (a later time period) and if a negative error term in time period t tends to be followed by another positive error in time period t + k(a later time period) (Bowerman, 2005, p. 289).

2.7.3. The Scatter Diagram of Error Term

The scatter diagram of error term over time can be used to explain if the time series data of the variable of interest has auto-correlation or not. If the scatter diagram has a cyclical appearance, the error terms are positively autocorrelated. If the scatter diagram has alternating appearance, the error terms are negatively autocorrelated. In both cases the dependency assumption is violated(Bowerman, 2005, p. 243).

2.7.4. Order of Autocorrelation

First –order autocorrelation exists when error term ε_t , error in period *t*, is correlated to ε_{t-1}, error time in period *t-1* (Bowerman, 2005, p 291).

$$\varepsilon_t = \phi\varepsilon_{t-1} + a_t$$

Where ε_t = the error term in time period t

ϕ = the correlation coefficient between error terms separated by one time period.

$a_1, a_2, \ldots$ = "random shock" values satisfy normality assumptions.

The most common method, used to test first – order (positive or negative) autocorrelation is the **Durbin-watson test** (Bowerman, 2005, p. 246).

The **Durbin-watson test** statistic is

$$d = \frac{\sum_{t=2}^{n}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=2}^{n}\varepsilon_t{}^2}$$

Where $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are the time - ordered error times.

2.7.5. Autocorrelation Models

If the analyst ignores auto-correlation effects on time series model, the model will have inadequate predictions. In such a case the analyst should take autocorrelation problem into account by modeling auto-correlation (Hanke, Reitsch, 1991, p. 669).

In real life, the first order autocorrelation is the most encountered problem (Bowerman, 2005, p 293). So the processes to build forecasting model which avoid first–order autocorrelation and have adequate results will be explained.

$$Z_t = \beta_o + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_k x_{tk} + \varepsilon_t$$

Where $x_{t1}, x_{t2}, \ldots, x_{tk}$ = the independent variable to be observed in time period *t*

$$\varepsilon_t = \phi\varepsilon_{t-1} + a_t$$

2.8. Decomposition Methods

Time series may have seasonal effects in addition to trend. However, seasonal effects must be detected and considered into account.

These models have no theoretical basis; they are strictly an intuitive approach. However, decomposition models have been found useful when the parameters describing a time series are not changing over time (Brockwell, 2002, p. 23-31)

In this method, the time series are decomposed into several factors and basic elements.

1. Trend
2. Seasonal is a pattern of change in quarterly or monthly data that repeats it self from year to year (Hanke, Reitsch, 1991, p. 604).
3. Cyclical is the wavelike fluctuation around the trend (Hanke, Reitsch, 1991, p. 603).
4. Irregular is a measure of variability of the time series after the other components have been removed.

The time series can be adequately described and adequate results can be found, if these factors are determined.

2.8.1. Multiplicative Decomposition

The model of multiplicative decomposition can be used when the time series exhibits increasing or decreasing seasonal variation, as the amplitude of seasonal movement depend on the level of the time series (Bowerman, 2005, p. 326).

$$Z_t = TR_t x SN_t x CL_t x IR_t$$

Where TR_t = the trend component

SN_t = the seasonal component

CL_t = the cyclical component

IR_t = the irregular component

2.8.2. Additive Decomposition

The model of multiplicative decomposition can be used when the time series exhibits constant seasonal variation, as the amplitude of seasonal movements does not depend on the level of the time series (Bowerman, 2005, p. 327).

$$Z_t = TR_t + SN_t + CL_t + IR_t$$

2.9. Box-Jenkins Methods (ARIMA Models)

The Box-Jenkins uses both the autoregressive and moving average techniques for forecasting (Hanke, Reitsch, 1991, p. 717).

2.9.1. Box-Jenkins Methods Steps

Box-Jenkins methodology consists of a four-step iterative procedure (Bowerman, 2005, p. 401).

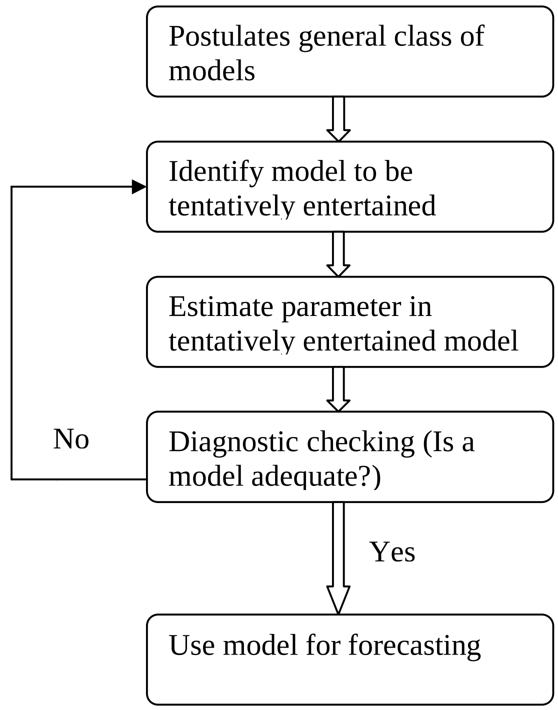
Tentative identification: historical data are used to tentatively identify an appropriate Box-Jenkins model.

Estimation: historical data are used to estimate the parameters of the tentatively identified model.

Diagnostic checking: various diagnostics are used to check the adequacy of the tentatively identified model and, if need be, to suggest an improved model.

Forecasting: once a final model is developed, it is used to forecast future time series values.

Figure 2.2 Box diagrams of Box-Jenkins procedures



Source :-(Hanke, Reitsch, 1991, p.717)

The main condition to use the Box-Jenkins methodology is that the time series of interest must be stationary

Stationary series is one whose basic statistical properties (like mean and variance) have no changes over time (Hanke, Reitsch, 1991, p. 661), and differencing can be used to transform a nonstationary time series to a stationary one.

This methodology does not assume that time series produces particular pattern. Instead of that, uses an iterative approach to identify potentially an appropriate model from general class of models. Then the selected model checked over the historical data to be sure that model is adequate (Chatfield, 1996, p. 98).

The appropriate model, whose error terms are small, will be considered randomly distributed, and independent. If the specified model is not appropriate, the process is repeated using another class of models to improve the suggested one (Bowerman, 2005, p. 400).

Box-Jenkins is a relatively accurate technique and powerful forecasting tool. However, it is quite complex and require computer analysis to perform many of computations which required identifying the appropriate model, estimating parameters and checking that the model is adequate (Bowerman, 2005, p. 402). However, ARIMA models are criticized for their black box approach that makes no attempt to discover the factors affecting the system of interest (Lim, 2002, p. 119).

2.9.2. Stationary and Nonstationary Time Series

A time series is considered stationary if the statistical properties of the time series are essentially constant over time. The easiest way to check the stationarity of a time series is a plot diagram.

If the diagram seems to fluctuate with constant variation around constant mean, then it is reasonable to believe that the time series stationary. If the diagram does not fluctuate with constant variation or do not fluctuate around a constant mean, then it is reasonable to believe that the time series is nonstationary (cryer, 2008, p. 87-92). In this study, more sophisticated methods were utilized to help in determining whether a time series is stationary or nonstationary

2.10. Choosing a Forecasting Method

In choosing a forecasting method, the forecaster must consider several factors.

2.10.1. Factors for Choosing a Forecasting Method (Bowerman, 2005, p. 19).

- The time frame
- The pattern of data
- The cost of forecasting
- The accuracy desired
- The availability of data
- The ease of operation and understanding

Box-Jenkins models can be used to identify an appropriate model from above presented methods which need combining the examination of data plots with great deal of intuition, whereas the Box-Jenkins methodology provides amore extensive collection of models and a much more systematic procedure for identifying an appropriate model (Bowerman, 2005, p. 24).

2.10.2. Discussion over Suitable Forecasting Method

According to previous review, it seems that there is not a single forecasting technique, can provide sufficient results in all fields or situations. However, it seems that ARIMA models could potentially be suitable for recent study for many reasons as follows:

A. The study has 42 time series, with several data pattern and ARIMA models don't assume that time series follow particular pattern.

B. ARIMA models provide a systemic procedure for the analysis of time series sufficiently general to handle virtually all empirically observed time series data pattern (Lim, 2002. p.129).

C. ARIMA models provide forecasts with acceptable accuracy even with comparison to other traditional methods (Lim, 2002. p.127).

D. The main factors which are affecting the behavior of variables, are uncontrollable factors at least in current time, so regression method is not usefulness, the factors are
   1. Population growth rate.
   2. Educational policy which is extremely changeable according to the changes in UNRWA administration.
   3. Annual fund which is extremely changeable according to donations and projects.

2.11. Model Building Strategy

Developing appropriate model is not easy mission, so multi-step model developing strategy will be used. There are four steps in the strategy; each of them may be repeated several times.

2.11.1. Model Identification (Brockwell, 2002, p. 14)

In this step, the time series plot is checked to determine whether there are:

A. Trend,
B. Seasonal component,
C. Any apparent sharp changes in behavior,
D. Any outlying observations.

After that, many different statistics can be computed from the data. It should be emphasized that the model, chosen at this point, is tentative and subject to revision later on the analysis. In model identification, taking into account that the chosen model should require the smallest number of parameters that will adequately represent the time series (cryer, 2008, p. 8).

2.11.2. Model Fitting

After identifying one or more model to represent the time series values, model fitting consists of finding the best possible estimates of those unknown parameters within chosen models (cryer, 2008, p. 8). Criteria such as least squares will be considered for estimation process.

2.11.3. Model Diagnostics

Model diagnostics concerns of evaluating the accuracy and the quality of the model which is identified and estimated. In diagnostics step, the model assumptions satisfaction must be checked to be sure that the model represents and fits the data adequately (Brockwell, 2002, p. 164).

If the diagnostics process proves that there are no inadequacies, the model may be completed then may be used to forecast future values. Otherwise, the first step is repeated to identify another model taking into account avoiding the inadequacies that are found in diagnostics step then complete model fitting and diagnostics. The three steps may be cycled several times, until fit model is found (Chatfield, 1996, p. 237).

Because the required calculations for each step in model building are large and intensive, statistical software like MINITAB can be used to carry out these calculations and plotting.

2.11.4. Implementation and Decision Making

The forecasts could be introduced to decision makers to use it in planning and decision making process.

Chapter Three: Time Series Models

3. Time Series Models

3.1. General Definitions

3.1.1. Residuals (Application Errors)

Application error is the difference between the value that occurs and the value that is predicted for a given time period (Stevenson, 1989, p. 456).

$$e_t = Z_t - \hat{Z}_t$$

All residuals can be calculated in one patch after applying the proposed model

Where e_t =application error, will be called residual

Z_t = actual value of variable for time period t

$\hat{Z}_t$ = predicted value for time period t

If a forecasting technique is appropriate for the problem of interest, the residuals will distribute randomly a round zero value, but if the residuals display any trend, this indicates that forecasting techniques are not appropriate.

3.1.2. Measures and Accuracy

Two methods commonly used for these purposes are the mean absolute deviation (MAD) and the mean squared error (MSE) (Render, 2006, p. 154), in addition to mean absolute percentage error (MAPE).

$$MAD = \frac{\sum_{t=1}^{n} |e_t|}{n} = \frac{\sum_{t=1}^{n} \left| Z_t - \hat{Z}_t \right|}{n}$$

$$MSE = \frac{\sum_{t=1}^{n} (e_t)^2}{n} = \frac{\sum_{t=1}^{n} (Z_t - \hat{Z}_t)^2}{n}$$

$$MAPE = \frac{\sum_{t=1}^{n} \left| \frac{z_t - \hat{z}_t}{z_t} \right|}{n} \times 100, \quad z_t \neq 0$$

3.1.3. Stationary Series

The time series will be considered stationary if it satisfies the conditions (Hamilton, 1994, p. 45)

1) $E(z_t) = \text{ constant} = \mu, \quad \forall t$

2) $\operatorname{cov}(z_t, z_s) = \begin{cases} \text{constant} = \gamma_0, \forall t, \forall s, t = s \\ f(|s-t|), \forall t, \forall s, t \neq s \end{cases}$

3.1.4. Autocovariance Function

The Autocovariance Function is defined as

$$\gamma_{t,s} = \operatorname{cov}(Z_t, Z_s), \forall t, \forall s$$
$$= E\left[(Z_t - \mu)(Z_s - \mu)\right], \forall t, \forall s$$

Let the lag k be the time period between Z_t and Z_{t-k} or Z_{t+k}, then autocovariance function is defined as

$$\gamma_k = \operatorname{cov}(Z_t, Z_{t-k}), \quad k = 0, \pm 1, \pm 2, \ldots$$
$$= E\left[(Z_t - \mu)(Z_{t-k} - \mu)\right], \quad k = 0, \pm 1, \pm 2, \ldots$$

3.1.5. Autocorrelation Function (ACF)

The Autocorrelation Function is defined as

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \quad k = 0, \pm 1, \pm 2, \ldots$$

Basic properties of autocorrelation function

1. $\rho_0 = 1$
2. $\rho_{-k} = \rho_k$
3. $|\rho_k| \leq 1$

3.1.6. Partial Autocorrelation Function (PACF)

The partial autocorrelation function at lag k is the correlation between Z_t and Z_{t-k} after removing the effect of the intervening variables $Z_{t-1}, Z_{t-2}, \ldots, Z_{t-k+1}$ which locate within (t,t-k) period, partial autocorrelation function will be donated by ϕ_{kk}, **PACF** will be calculated by iteration (Shumway, 2006, p. 106).

$$\phi_{00} = 1$$
$$\phi_{11} = \rho_1$$

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j}, \quad k = 2,3,...$$

Where $\phi_{kj} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-1}, \quad j = 1,2,...,k-1$

3.1.7. Sample Autocorrelation Function SACF

The definition of the sample autocorrelation function for an observed time series, $r_k,\ k = 0,1,2,...$ as

$$r_k = \frac{\sum_{t=1}^{n-k} (z_t - \overline{z})(z_{t+k} - \overline{z})}{\sum_{t=1}^{n} (z_t - \overline{z})^2}, \quad k = 0,1,2,...$$

Where $\overline{z} = \frac{1}{n}\sum_{t=1}^{n} z_t$, r_k is an estimator for $\hat{\rho}_k$

3.1.8. Sample Partial Autocorrelation Function SPACF

The definition of the sample partial autocorrelation function for an observed time series, $r_{kk},\ k = 0,1,2,...$ as

$$r_{00} = 1$$

$$r_{11} = r_1$$

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j}, \quad k = 2,3,...$$

Where $r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-1}, \quad j = 1,2,...,k-1$, r_{kk} is an estimator for $\hat{\phi}_{kk}$

3.1.9. White Noise Series or White Noise Process

$\{a_t\}$ a sequence of independent, identically distributed(IID) random variables, white noise process has mean zero and constant variance σ^2 (Cryer, 2008, p. 17) **.**

1) $E(a_t) = 0, \forall t$

2) $\operatorname{cov}(a_t, a_s) = \begin{cases} \sigma^2, & \forall t, \forall s, t = s \\ 0, & \forall t, \forall s, t \neq s \end{cases}$

It is denoted $a_t \approx i.i.d. \;\; N(0, \sigma^2)$

Here on, autocorrelation and partial autocorrelation function of white noise series.

Figure 3.1 ACF of white noise series

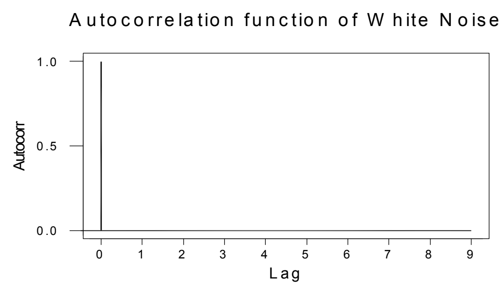
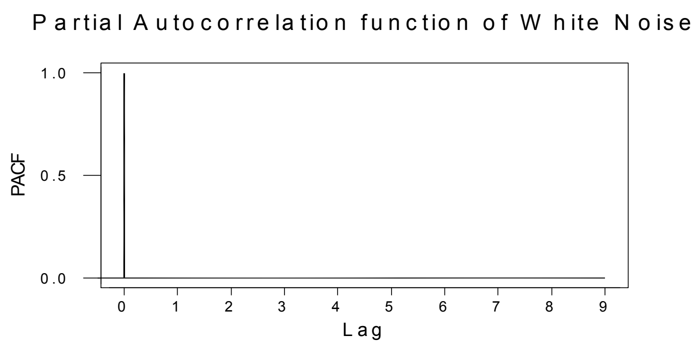


Figure 3.2 PACF of white noise series



Both Autocorrelation and partial autocorrelation functions for white noise series equal zero from lag 1.

3.2. ARMA Models

There is large family of models which is named "Autoregressive-Moving Average Models" and abbreviated by ARMA. Many of researches in different application field prove that ARMA models more fitness than other traditional methods of forecasting (Hamilton, 1994, p. 60). Some of these traditional methods mentioned in time series chapter.

Chapter Four: Model Inference and Forecasting

4. Model Inference and Forecasting

In the last chapter, the properties and plots for different types of time series were discussed, now the criteria that should be used to identify the best model will be illustrated.

4.1. Models Identification

The plot and behavior of time series data give general indications about the time series model and its specification, so the analysts often use several statistical inference methods to minimize the identification error. However, the suggested model still initiative and need diagnostics process to check its appropriateness. So the statistical inference will be considered.

4.1.1. Checking The Time Series Data for Nonstationarity in Variation

Plotting time series data then making some statistical test to check stationarity in variation, in the case of no constant variation, the analysts apply appropriate Variance stabilizing Transformation like logarithm transformation (Shumway, Stoffer, 2005, p. 63).

4.1.2. Checking The Time Series Data for Nonstationarity in Mean

(Shumway, Stoffer, 2005, p.60)

Plotting time series and its ACF and PACF to check instability of the series mean. In the case of instability, the analysts apply one degree differencing then:

1. Time series data plot or its transformation is checked.
2. ACF and PACF plot are checked.
3. If the time series still instable, higher degree differencing is applied then step 1 and step 2 are repeated.

4.1.3. Choosing The Values of p and q for A given Series

After getting stationary series, it is reasonable to determine the degree of autoregression p and the degree of moving average q.

The order p and q could be determined by comparison between sample ACF and PACF plot and ideal ACF and PACF plot. The following table summarizes the behavior of general time series models.

Table 4.1 General Behavior of the ACF and PACF for ARMA Models

	AR(p)	MA(q)	ARMA(p,q),p>0 and q>o
ACF	Tail off	Cuts off after lag q	Tail off
PACF	Cuts off after lag p	Tail off	Tail off

Source: - (Shumway, Stoffer, 2005, p. 156)

4.2. Model Estimation

This section will study the problem of estimating the parameters value of an ARMA models, based on the observed time series $\mathbf{z} = \{z_1, z_2, \ldots, z_n\}$. With assumption that a model has already been identified; that is, values for p, d, and q are specified using the methods mentioned in chapter three.

The parameters of AR model are denoted $\phi_1, \phi_2, \ldots, \phi_p$ and the parameters of MA model are denoted $\theta_1, \theta_2, \ldots, \theta_q$.

The method-of-moments estimators will be discussed firstly then the least squares estimators, and finally full maximum likelihood estimators.

4.2.1. The Method of Moments

The idea behind these estimators is that of equating population moments to sample moments then solving for the parameters in terms of the sample moments (Shumway, Stoffer, 2005, p. 122).

The simplest example of the method is to estimate a stationary process mean by a sample mean (Cryer, 2008, p. 149).

Although the method of moments can produce good estimators, they can sometimes lead to suboptimal estimators (Shumway, Stoffer, 2005, p. 122). First, the case in which the method leads to optimal (efficient) estimators, will be considered, that is, AR (p) models.

When the process is AR (p), hereby the method to estimate model parameters

The mean μ is estimated by the estimator $\bar{z}$

$$\hat{\mu} = \bar{z} = \sum_{i=1}^{n} z_i \Big/ n$$

To estimate $\phi_1, \ldots, \phi_p$, use the relation:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p},\ k > 1$$

This relation is deducted from multiplying AR (p) model equation by the term $z_{t-k} - \mu$ then taking expectation.

When $k = 1, 2, \ldots, p$, equation system is called Yule-Walker equations (Bowerman, 2005, p.150).

$$\rho_1 = \phi_1 + \phi_2 \rho_1 + \cdots + \phi_p \rho_{p-1}$$
$$\rho_2 = \phi_1 \rho_1 + \phi_2 + \cdots + \phi_p \rho_{p-2}$$
$$\vdots$$
$$\rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \cdots + \phi_p$$

Substituting ρ_k by the estimator r_k, the method of moment's parameters $\hat{\phi}_1, \ldots, \hat{\phi}_p$ will be as follow:

Transforming Yule-Walker equations to matrix form

$$\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{pmatrix} = \begin{pmatrix} 1 & r_1 & r_2 & \cdots & r_{p-2} & r_{p-1} \\ r_1 & 1 & r_1 & \cdots & r_{p-3} & r_{p-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{p-1} & r_{p-2} & r_{p-3} & \cdots & r_1 & 1 \end{pmatrix} \begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_p \end{pmatrix}$$

Solving this equation to get parameters

$$\begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_p \end{pmatrix} = \begin{pmatrix} 1 & r_1 & r_2 & \cdots & r_{p-2} & r_{p-1} \\ r_1 & 1 & r_1 & \cdots & r_{p-3} & r_{p-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{p-1} & r_{p-2} & r_{p-3} & \cdots & r_1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{pmatrix}$$

σ^2 is estimated by the equation

$$\hat{\sigma}^2 = \hat{\gamma}_0 \left(1 - \hat{\phi}_1 r_1 - \hat{\phi}_2 r_2 - \cdots \hat{\phi}_p r_p\right)$$

Where

$\hat{\gamma}_0 = \frac{1}{n}\sum_{t=1}^{n}\left(z_t - \bar{z}\right)^2$ is the sample variance

4.2.2. Moments Estimation for Autoregressive Models

4.2.2.1. AR(1) Model

For this process, $\rho_1 = \phi$. In the method of moments, ρ_1 is equated to r_1, the lag 1 sample autocorrelation. Thus, ϕ value can be estimated by a simple relation (Cryer, 2008, p. 149).

$$\hat{\phi} = r_1$$

4.2.2.2. AR(2) Model

In AR(2) case. The relationships between the parameters ϕ_1 and ϕ_2 and various moments are given by the Yule-Walker equations.

$$\rho_1 = \phi_1 + \rho_1\phi_2 \quad and \quad \rho_2 = \rho_1\phi_1 + \phi_2$$

The method of moments replaces ρ_1 by r_1 and ρ_2 by r_2 to obtain $\hat{\phi}$ = r1

$$r_1 = \phi_1 + r_1\phi_2 \quad and \quad r_2 = r_1\phi_1 + \phi_2$$

which are solved to obtain

$$\hat{\phi}_1 = \frac{r_1(1-r_2)}{1-r_1^2} \quad and \quad \hat{\phi}_2 = \frac{r_2 - r_1^2}{1-r_1^2}$$

4.2.2.3. General AR(p)

The general AR(p) case proceeds similarly. ρ_k is replaced by r_k throughout the Yule-Walker equations. Then the linear equations are solved for $\hat{\phi}_1, \hat{\phi}_2, \ldots \hat{\phi}_p$
.
The estimators which are obtained in this way are also called Yule-Walker estimators (Cryer, 2008, p. 150).

4.2.3. Moments Estimation for Moving Average Models

4.2.3.1. MA(1)

From the properties of the model, it is found that

$$\rho_1 = -\frac{\theta}{1+\theta^2}$$

When ρ_1 is equated to r_1, quadratic equation is found in θ

The solution for the quadratic equation is:-

$$\hat{\theta}_1 = \frac{-1 \pm \sqrt{1-4r_1}}{2r_1}$$

This solution give two values for the estimator $\hat{\theta}$, only the value that satisfies the condition $\left|\hat{\theta}_1\right| < 1$ is considered

4.2.3.2. MA(2)

Finding the moment of estimators for the parameters θ_1 and θ_2, the following relations will be used

$$\rho_k = \begin{cases} 1, & k=0 \\ \dfrac{-\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2}, & k=1 \\ \dfrac{-\theta_2}{1+\theta_1^2+\theta_2^2}, & k=2 \\ 0, & k>2 \end{cases}$$

Then ρ_1 is replaced by r_1 and ρ_2 by r_2 and the equations are solved to get the estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ which satisfy the conditions:-

$$\theta_2 - \theta_1 < 1, \quad \theta_2 + \theta_1 < 1, \quad \left|\theta_2\right| < 1$$

4.2.3.3. General MA(p)

For higher-order MA models, the method of moments quickly becomes complicated, so the following equations can be used

$$\rho_k = \begin{cases} \dfrac{-\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \Lambda + \theta_{q-k}\theta_q}{1+\theta_1^2 + \theta_2^2 + \Lambda + \theta_q^2} & for\, k = 1,2,\Lambda\ ,q \\ 0 \quad for \quad k > q \end{cases}$$

If ρ_k is replaced by r_k for k = 1, 2,…, q, q equations will be deducted to in q unknowns $\theta_1, \theta_2, \mathrm{K}\ \theta_q$. The resulting equations are highly nonlinear in the θ's, however, their solution would necessity be numerical (Cryer, 2008, p. 151).

4.2.4. Moments Estimation for ARMA(1,1)

For finding the moment of estimators for the parameters θ_1 and ϕ_1, the following relations can be used.

$$\rho_1 = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2 - 2\phi_1\theta_1}, \quad \rho_2 = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2 - 2\phi_1\theta_1}\phi_1$$

If ρ_1 is replaced by r_1 and ρ_2 by r_2 then the equations are solved to get the estimators $\hat{\theta}_1$ and $\hat{\phi}_1$

$$r_1 = \frac{\left(1-\hat{\phi}_1\hat{\theta}_1\right)\left(\hat{\phi}_1-\hat{\theta}_1\right)}{1+\hat{\theta}_1^2 - 2\hat{\phi}_1\hat{\theta}_1}, \quad r_2 = \frac{\left(1-\hat{\phi}_1\hat{\theta}_1\right)\left(\hat{\phi}_1-\hat{\theta}_1\right)}{1+\hat{\theta}_1^2 - 2\hat{\phi}_1\hat{\theta}_1}\hat{\phi}_1$$

By dividing r2 equation by r1 equation, the deducted equation is

$$\hat{\phi}_1 = \frac{r_2}{r_1}$$

$\hat{\theta}_1$ parameter can be found by substituting of $\hat{\phi}_1$ in r1 equation,

$$r_1 = \frac{\left(1-\frac{r_2}{r_1}\hat{\theta}_1\right)\left(\frac{r_2}{r_1}-\hat{\theta}_1\right)}{1+\hat{\theta}_1^2 - 2\frac{r_2}{r_1}\hat{\theta}_1}$$

Then quadratic equation is solved and $\hat{\theta}_1$ value which satisfy the condition $\left|\hat{\theta}_1\right| < 1$ is found

4.2.5. Conditional Least Squares Method

This method was studied in 2.6.2 section for a straight line, here after, the same principles on autoregressive, moving average and mixed models will be used.

4.2.5.1. Conditional Least Squares Estimation for ARMA(p,q)

$$\phi_p(B) z_t = \delta + \theta_q(B) a_t$$

Constant term δ can be included in a nonzero Constant mean, μ, in a stationary ARMA model

$$\phi_p(B)(z_t - \mu) = \theta_q(B) a_t$$

$$a_t = \frac{\phi_p(B)}{\theta_q(B)}(z_t - \mu)$$

The right side can be considered function of $\boldsymbol{\varphi} = \{\phi_1, \phi_2, \ldots, \phi_p\}$, $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_q\}$ and μ

$$a_t(\boldsymbol{\varphi}, \boldsymbol{\theta}, \mu) = \frac{\left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right)}{\left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_p B^p\right)}(z_t - \mu)$$

For an observed time series $\mathbf{z} = \{z_1, z_2, \ldots, z_n\}$, conditional least squares method depends on minimizing the function

$$\min_{\boldsymbol{\varphi}, \boldsymbol{\theta}, \mu} S_c(\boldsymbol{\varphi}, \boldsymbol{\theta}, \mu) = \sum_{t=p+1}^{n} a_t^2(\boldsymbol{\varphi}, \boldsymbol{\theta}, \mu | \mathbf{z})$$

Finding the required estimators needs solving the following normal equations.

$$\frac{\partial}{\partial \boldsymbol{\varphi}} S_c(\boldsymbol{\varphi},\boldsymbol{\theta},\mu)\Bigg|_{\substack{\boldsymbol{\varphi}=\hat{\boldsymbol{\varphi}}\\ \boldsymbol{\theta}=\hat{\boldsymbol{\theta}}\\ \mu=\hat{\mu}}} = \frac{\partial}{\partial \boldsymbol{\varphi}} \sum_{t=p+1}^{n} a_t^2(\boldsymbol{\varphi},\boldsymbol{\theta},\mu|\mathbf{z})\Bigg|_{\substack{\boldsymbol{\varphi}=\hat{\boldsymbol{\varphi}}\\ \boldsymbol{\theta}=\hat{\boldsymbol{\theta}}\\ \mu=\hat{\mu}}} = 0$$

$$\frac{\partial}{\partial \boldsymbol{\theta}} S_c(\boldsymbol{\varphi},\boldsymbol{\theta},\mu)\Bigg|_{\substack{\boldsymbol{\varphi}=\hat{\boldsymbol{\varphi}}\\ \boldsymbol{\theta}=\hat{\boldsymbol{\theta}}\\ \mu=\hat{\mu}}} = \frac{\partial}{\partial \boldsymbol{\theta}} \sum_{t=p+1}^{n} a_t^2(\boldsymbol{\varphi},\boldsymbol{\theta},\mu|\mathbf{z})\Bigg|_{\substack{\boldsymbol{\varphi}=\hat{\boldsymbol{\varphi}}\\ \boldsymbol{\theta}=\hat{\boldsymbol{\theta}}\\ \mu=\hat{\mu}}} = 0$$

$$\frac{\partial}{\partial \mu} S_c(\boldsymbol{\varphi},\boldsymbol{\theta},\mu)\Bigg|_{\substack{\boldsymbol{\varphi}=\hat{\boldsymbol{\varphi}}\\ \boldsymbol{\theta}=\hat{\boldsymbol{\theta}}\\ \mu=\hat{\mu}}} = \frac{\partial}{\partial \mu} \sum_{t=p+1}^{n} a_t^2(\boldsymbol{\varphi},\boldsymbol{\theta},\mu|\mathbf{z})\Bigg|_{\substack{\boldsymbol{\varphi}=\hat{\boldsymbol{\varphi}}\\ \boldsymbol{\theta}=\hat{\boldsymbol{\theta}}\\ \mu=\hat{\mu}}} = 0$$

These estimators are called conditional because they satisfy that $a_p = a_{p-1} = \cdots = a_{p+1-q} = 0$ or because the values $a_p = a_{p-1} = \ldots = a_{p+1-q}$ equal their expectation (Shumway, Stoffer, 2005, p. 128).

The variance σ^2 equals

$$\hat{\sigma}^2 = \frac{S_c(\hat{\boldsymbol{\varphi}},\hat{\boldsymbol{\theta}},\mu)}{n-(p+q+1)}$$

4.2.5.2. Conditional Least Squares Estimation for AR(1)

$$z_t = \delta + \phi_1 z_{t-1} + a_t$$

$$\therefore \delta = \mu(1-\phi_1)$$
$$\because\ z_t - \mu = \phi_1(z_{t-1}-\mu) + a_t$$

Replacing μ with its estimator $\bar{z}$ will be Simplify the derivation process

$$z_t - \bar{z} = \phi_1(z_{t-1}-\bar{z}) + a_t$$

For the observed values $\mathbf{z} = \{z_1, z_2, \ldots, z_n\}$, the error is

$$a_t(\phi_1) = (z_t - \bar{z}) - \phi_1(z_{t-1}-\bar{z}),\ t = 2,3,\cdots,n$$

Now, the sum of square errors will be minimized

$$a_t^2(\phi_1) = \left[(z_t - \bar{z}) - \phi_1(z_{t-1} - \bar{z})\right]^2, \ t = 2, 3, \ldots, n$$

$$S_c(\phi_1) = \sum_{t=2}^{n} a_t^2(\phi_1) = \sum_{t=2}^{n}\left[(z_t - \bar{z}) - \phi_1(z_{t-1} - \bar{z})\right]^2$$

Now the equation $\partial S_c / \partial \phi = 0$ will be considered for ϕ_1 parameter

$$\hat{\phi}_1 = \frac{\sum_{t=2}^{n}(z_{t-1} - \bar{z})(z_t - \bar{z})}{\sum_{t=2}^{n}(z_{t-1} - \bar{z})^2}$$

$\hat{\phi}_1$ is the conditional least squares estimation for ϕ_1 parameter and the equation $\partial S_c / \partial \mu = 0$ can be solved for μ parameter

4.2.5.3. Conditional Least Squares Estimation for MA(1)

$$Z_t = \delta + a_t - \theta_1 a_{t-1}, \quad a_t \equiv WN(0, \sigma^2)$$

$$Z_t - \mu = a_t - \theta_1 a_{t-1}$$

By replacing μ with its estimator $\bar{z}$

$$z_t - \bar{z} = a_t - \theta_1 a_{t-1}$$

Now, the series of the mean ($x_t = z_t - \bar{z}$) will be considered

$$x_t = a_t - \theta_1 a_{t-1}$$

$$a_t = x_t - \theta_1 a_{t-1}$$

For the given observations $x_1, x_2, \ldots, x_n$ and when $a_0 = 0$, the errors can be written as:

$$a_1 = x_1$$

$$a_2 = x_2 - \theta_1 a_1$$

$$a_3 = x_3 - \theta_1 a_2$$

$$\vdots$$

$$a_n = x_n - \theta_1 a_{n-1}$$

So

$$S_c(\theta_1) = \sum_{t=1}^{n} a_t^2$$

the previous equation is nonlinear and the value of θ_1 which minimize $S_c(\theta_1)$ can be found by numerical methods such as network methods or Gaussian – Newton method (Bowerman, 2005, p.157).

4.2.6. Maximum likelihood Method

Finding maximum likelihood estimates conceptually involves two steps. First, calculating maximum likelihood function. Second, finding values of parameter which maximize this function (Hamilton, 1994, p. 117).

The advantage of maximum likelihood is that all the information in data is used rather than just the first and second moments, such least squares case. Another advantage is that many large-sample results are known under very general conditions. One disadvantage is that in the first time, the joint probability density function of the process has to be applied (Cryer, 2008, p. 158).

4.2.6.1. Maximum Likelihood Estimation

For any set of observations, $\mathbf{z} = \{z_1, z_2, \ldots, z_n\}$, time series or not, the likelihood function L is defined to be the joint probability density of obtaining the data actually observed. However, it is considered as a function of the unknown parameters in the model with the observed data held fixed. For ARIMA models, L will be a function of the ϕ's, θ's, μ, and given the observations $\mathbf{z} = \{z_1, z_2, \ldots, z_n\}$ (Brockwell, 2002, p.387).

The maximum likelihood estimators are then defined as those values of the parameters for which the data actually observed are most likely, that is, the values that maximize the likelihood function (Cryer, 2008, p. 158).

This approach requires specifying particular distribution for the white noise process a_t. Typically a_t will be assumed as Gaussian white noise.

$$a_t \approx i.i.d.\ \ N(0, \sigma^2)$$

Gaussian white noises series are independent, identical, normally distributed random variables with zero means and common standard deviation σ (Hamilton, 1994, p. 25).

4.2.6.1.1. Maximum Likelihood Estimation for AR(1)

The log-Likelihood function for AR(1) model is

$$l(\phi,\mu,\sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) + \frac{1}{2}\log(1-\phi^2) - \frac{1}{2\sigma^2}S(\phi,\mu)$$

Where

$$S(\phi,\mu) = \sum_{t=2}^{n}[(z_t-\mu)-\phi(z_{t-1}-\mu)]^2 + (1-\phi^2)(z_1-\mu)^2$$

The function $S(\phi,\mu)$ is called unconditional sum-of-squares function.

For given values of ϕ and μ, $l(\phi,\mu,\sigma^2)$ can be maximized analytically with respect to σ^2 in terms of the yet-to-be determined estimators of ϕ and μ. The deducted equation is

$$\hat{\sigma}^2 = \frac{S(\hat{\phi},\hat{\mu})}{n}$$

Now, consider the estimation of ϕ and μ which is called unconditional sum of squares function $S(\phi,\mu)$

$$S(\phi,\mu) = S_c(\phi,\mu) + (1-\phi^2)(z_1-\mu)^2$$

Where $S_c(\phi,\mu)$ is the conditional sum of squares, for an observed values $\mathbf{z} = \{z_1, z_2, \ldots, z_n\}$

$$S_c(\phi,\mu) = \sum_{t=2}^{n}[(z_t-\mu)-\phi(z_{t-1}-\mu)]^2$$

4.3. Model Checking and Diagnostics

Typically, the goodness of fit for a statistical model is judged by comparing the observed values with the corresponding predicted values obtained from the fitted model. If the fitted model is appropriate, then the residuals should behave in a manner that is consistent with the model (Brockwell, 2002, p. 164). Model diagnostic includes residual analysis as well as model comparison.

In section 2.3.4, there is a definition of residual with complete details, but hereafter, with time series observations $\mathbf{z} = \{z_1, z_2, \ldots, z_n\}$.

$$e_t = z_t - \hat{z}_t$$

Residual = actual − predicted

It is noted that the residuals can be calculated as a patch after applying the model.

Diagnostic and checking process concern of residuals analysis to insure that the model satisfies its assumptions (Bowerman, 2005, p. 458), which are

1. Zero mean assumption
2. Constant variance assumption.
3. Normality assumption.
4. Independence assumption.

There are more details about these assumptions in section 2.6.3

If the model is correctly specified and the parameter estimates are reasonably close to the true values, then the residuals should have nearly the properties of white noise. They should behave roughly like independent, identically distributed normal variables with zero means and common standard deviations. Deviations from these properties can help to discover a more appropriate model (Cryer, 2008, p. 176).

4.3.1. Plot of The Residuals (Zero Mean Assumption)

The first step in diagnostic process is checking a plot of the residuals over time. If the fitted model is appropriate, the plot will be scatter around a zero line without trend. This means that the model satisfies zero mean assumption(Chatfield, 1996, p. 86)..

4.3.2. Normality Diagnostic

Normality can be checked carefully by plotting Normal Probability Plot
The straight-line pattern here supports the assumption of a normally distributed stochastic component in this model(Bowerman, 2005, p. 178). There are other tools may be used to check normality.

1. quantile-quantile (Q-Q) plot
2. Goodness of Fit Test
3. Kolmogorov-Smirnov Test

4.3.3. Independence or Correlation Diagnostic

The independence or correlation can be checked by the sample autocorrelation function ACF, this can be applied by calculating and plotting ACF for residuals then comparison between this plot and the plot of white noise series is applied. So the residuals should have nearly the properties of white noise. Independent identically distributed normally with zero means and common standard deviations (Brockwell, 2002, p. 413).

4.3.4. The Ljung-Box Test(Shumway, 2006, p. 149)

Instead of checking the residual correlations at individual lags. It is possible to consider single statistic which take into account a wide range of lags.

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

Q has an approximate chi-square distribution with (k - m) degrees of freedom

Where n = number of observation
k = number of lags
m = (p+q), number of proposed parameters in the model

4.3.5. Randomness

The randomness of residual can easily be checked by Runs test around mean or zero value (Shumway, 2006, p. 148).

4.3.6. Model Selection Criteria

These criteria give substantial signs about the goodness of model, so it can be used to check the appropriateness of proposed model

4.3.6.1. Automatic Information Criteria (AIC)

$$\text{AIC} = -2\log(\text{maximum likelihood}) + 2k$$

$$AIC(m) = n\ln\sigma_a^2 + 2m$$

Where k = (p+q), number of proposed parameters in the model

The most appropriate model has the minimum $AIC(k)$ criteria value.

4.3.6.2. Bayesian Information Criteria (BIC)

BIC = – 2log(maximum likelihood) + klog(n)

$$BIC(m) = n \ln \sigma_a^2 + m \ln(n)$$

Where n = the number of observation

k or m = (p+q)

The most appropriate model has the minimum BIC criteria value.

4.4. Forecasting

The goal of forecasting is predicting the future values of time series based on present observations and assessing the precision of these forecasts. In forecasting process, it is assumed that parameters of model are well a known and stationary case will be continued in the future.

The symbol $z_{n+1}, z_{n+2}, z_{n+3}, \ldots$ or in general $z_{n+l}, l \geq 0$ will be used for future observations and the symbol $z_n(1), z_n(2), z_n(3), \ldots$ or in general $z_n(l), l \geq 0$ for their correspondent forecasts.

4.4.1. Forecast Error

It is the difference between the future forecast $z_n(l), l \geq 0$ and the future observation $z_{n+l}, l \geq 0$. However, residuals are deducted in a patch; forecast errors deducted one by one as the time goes a head, the real values are gradually known (cryer, 2008, p. 192).

$$e_n(l) = z_{n+l} - z_n(l), \quad l \geq 0$$

4.4.2. Minimum Mean Square Error Forecasting (MMSE Forecasts) (cryer, 2008, p. 191)

$$z_n(l) = E\left(z_{n+l} \mid z_n, z_{n-1}, \ldots\right), \quad l \geq 1$$

MMSE forecasts are the conditional expectation for future observations with consideration of the observed series $\{z_1, z_2, \ldots, z_{n-1}, z_n\}$. Hereafter, the basic rule of conditional expectation.

$$1 - E\left(a_{n+j} \middle| z_n, z_{n-1}, \ldots\right) = \begin{cases} a_{n+j}, & j \le 0 \\ 0, & j > 0 \end{cases}$$

$$2 - E\left(z_{n+j} \middle| z_n, z_{n-1}, \ldots\right) = \begin{cases} z_{n+j}, & j \le 0 \\ z_n(j), & j > 0 \end{cases}$$

4.4.3. AR (1) Model Forecast

For time series observations $\{z_1, z_2, \ldots, z_{n-1}, z_n\}$, with AR (1) model

$$z_t - \mu = \phi_1(z_{t-1} - \mu) + a_t,\ \ a_t = WN(0,\sigma^2), |\phi_1| < 1, \mu \in (-\infty, \infty)$$

Now, future forecast $z_{n+1}, z_{n+2}, z_{n+3}, \ldots$ or in general form $z_{n+l},\quad l \ge 1$

$$Z_{n+\lambda} - \mu = \phi(Z_{t-\lambda-1} - \mu) + a_{t+\lambda}$$

If the conditional expectation of both sides is applied, the deducted equation will be:

$$\begin{aligned} z_n(l) &= E\left(z_{n+l} \middle| z_n, z_{n-1}, \ldots\right), \quad l \ge 1 \\ &= \mu + E\left[\left[\phi_1\left(z_{n+l-1} - \mu\right) + a_{n+l}\right] \middle| z_n, z_{n-1}, \ldots\right], \quad l \ge 1 \\ &= \mu + E\left[\phi_1\left(z_{n+l-1} - \mu\right) \middle| z_n, z_{n-1}, \ldots + a_{n+l} \middle| z_n, z_{n-1}, \ldots\right], \quad l \ge 1 \\ &= \mu + \phi_1 \mathrm{E}\left[\left(z_{n+l-1} \middle| z_n, z_{n-1}, \ldots\right) - \mu\right] + E\left[a_{n+l} \middle| z_n, z_{n-1}, \ldots\right], \quad l \ge 1 \\ z_n(l) &= \mu + \phi_1 \mathrm{E}\left[\left(z_{n+l-1} \middle| z_n, z_{n-1}, \ldots\right) - \mu\right] + E\left[a_{n+l} \middle| z_n, z_{n-1}, \ldots\right], \quad l \ge 1 \end{aligned}$$

This equation can be solved by iterating and applying condition expectation rules

$$\begin{aligned} l = 1: z_n(1) &= \mu + \phi_1 \mathrm{E}\left[\left(z_n \middle| z_n, z_{n-1}, \ldots\right) - \mu\right] + E\left[a_{n+1} \middle| z_n, z_{n-1}, \ldots\right] \\ &= \mu + \phi_1\left(z_n - \mu\right) \\ l = 2: z_n(2) &= \mu + \phi_1 \mathrm{E}\left[\left(z_{n+1} \middle| z_n, z_{n-1}, \ldots\right) - \mu\right] + E\left[a_{n+2} \middle| z_n, z_{n-1}, \ldots\right] \\ &= \mu + \phi_1\left[z_n(1) - \mu\right] \\ l = 3: z_n(3) &= \mu + \phi_1 \mathrm{E}\left[\left(z_{n+2} \middle| z_n, z_{n-1}, \ldots\right) - \mu\right] + E\left[a_{n+3} \middle| z_n, z_{n-1}, \ldots\right] \\ &= \mu + \phi_1\left[z_n(2) - \mu\right] \end{aligned}$$

The general form can be written as:-

$$Z_n(\lambda) = \mu + \phi[Z_n(\lambda - 1) - \mu] \quad for \quad \lambda \geq 1$$

This equation is **MMSE** forecast model for AR(1) Models and it illustrates that forecast process starts with the initial forecast $Z_n(1)$ then the higher forecast can be built up from $Z_n(2)$ and so on until the desired forecast $Z_n(\lambda)$ is found.

$$Z_n(\lambda) \approx \mu \quad for \quad \text{large } \lambda$$

Forecast Error for AR (1)

General Linear process for AR (1) is

$$Z_t = a_t + \phi\, a_{t-1} + \phi^2 a_{t-2} + \phi^3 a_{t-3} + \ldots$$

After some derivations, the deducted equation is

$$e_n(\lambda) = e_{n+\lambda} + \phi\ e_{n+\lambda-1} + \ldots + \phi^{\lambda-1} e_{n+1}$$

This equation can be written in the weights form

$$e_n(\lambda) = e_{n+\lambda} + \psi_1\ e_{n+\lambda-1} + \psi_2\ e_{n+\lambda-2} + \ldots + \psi_{\lambda-1}\ e_{n+1}$$

Since, $\psi_j = \phi^j, |\phi_1| < 1$

$$E(e_n(\lambda)) = 0$$

$$Var(e_n(\lambda)) = \sigma_e^2(1 + \psi_1^2 + \psi_2^2 + \ldots + \psi_{\lambda-1}^2)$$

Since, $\sum_{i=1}^{\infty} \psi_i^2 < \infty$

After finding a finite series

$$Var(e_n(\lambda)) = \sigma_e^2 \left[\frac{1 - \phi^{2\lambda}}{1 - \phi^2} \right]$$

For large λ, the variance can be calculated be the following equation

$$Var(e_n(\lambda)) \approx \frac{\sigma_e^2}{1 - \phi^2}$$

The following equation will be valid for all stationary ARMA processes

$$Var(e_n(\lambda)) \approx Var(Z_t) = \gamma_0 \quad for \quad \text{large } \lambda$$

4.4.4. AR (2) Model Forecast

$$z_t - \mu = \phi_1(z_{t-1} - \mu) + \phi_2(z_{t-2} - \mu) + a_t\,,\;\; a_t = WN(0,\sigma^2), |\phi_1| < 1, \mu \in (-\infty,\infty)$$

Where
$$\phi_2 - \phi_1 < 1$$
$$\phi_2 + \phi_1 < 1$$
$$-1 < \phi_2 < 1$$

Future forecast is

$$z_n(l) = E\left(z_{n+l} \mid z_n, z_{n-1}, \cdots\right), \quad l \geq 1$$
$$= \mu + E\Big[\big[\phi_1(z_{n+l-1} - \mu) + \phi_2(z_{n+l-2} - \mu) + a_{n+l}\big] \big| z_n, z_{n-1}, \cdots\Big], \quad l \geq 1$$
$$= \mu + E\big[\phi_1(z_{n+l-1} - \mu) \mid z_n, z_{n-1}, \cdots + \phi_2(z_{n+l-2} - \mu) \mid z_n, z_{n-1}, \cdots + a_{n+l} \mid z_n, z_{n-1}, \cdots\big], \quad l \geq 1$$
$$= \mu + \phi_1 \mathrm{E}\big[(z_{n+l-1} \mid z_n, z_{n-1}, \cdots) - \mu\big] + \phi_2 \mathrm{E}\big[(z_{n+l-2} \mid z_n, z_{n-1}, \cdots) - \mu\big] + E\big[a_{n+l} \mid z_n, z_{n-1}, \cdots\big], \quad l \geq 1$$
$$z_n(l) = \mu + \phi_1 \mathrm{E}\big[(z_{n+l-1} \mid z_n, z_{n-1}, \cdots) - \mu\big] + \phi_2 \mathrm{E}\big[(z_{n+l-2} \mid z_n, z_{n-1}, \cdots) - \mu\big] + E\big[a_{n+l} \mid z_n, z_{n-1}, \cdots\big], \quad l \geq 1$$

This equation can be solved by iterating and applying conditional expectation rules.

$$l = 1: z_n(1) = \mu + \phi_1 \mathrm{E}\big[(z_n \mid z_n, z_{n-1}, \cdots) - \mu\big] + \phi_2 \mathrm{E}\big[(z_{n-1} \mid z_n, z_{n-1}, \cdots) - \mu\big] + E\big[a_{n+1} \mid z_n, z_{n-1}, \cdots\big]$$
$$= \mu + \phi_1(z_n - \mu) + \phi_2(z_{n-1} - \mu)$$
$$l = 2: z_n(2) = \mu + \phi_1 \mathrm{E}\big[(z_{n+1} \mid z_n, z_{n-1}, \cdots) - \mu\big] + \phi_2 \mathrm{E}\big[(z_n \mid z_n, z_{n-1}, \cdots) - \mu\big] + E\big[a_{n+2} \mid z_n, z_{n-1}, \cdots\big]$$
$$= \mu + \phi_1\big[z_n(1) - \mu\big] + \phi_2(z_n - \mu)$$
$$l = 3: z_n(3) = \mu + \phi_1 \mathrm{E}\big[(z_{n+2} \mid z_n, z_{n-1}, \cdots) - \mu\big] + \phi_2 \mathrm{E}\big[(z_{n+1} \mid z_n, z_{n-1}, \cdots) - \mu\big] + E\big[a_{n+3} \mid z_n, z_{n-1}, \cdots\big]$$
$$= \mu + \phi_1\big[z_n(2) - \mu\big] + \phi_2\big[z_n(1) - \mu\big]$$
$$l = 4: z_n(4) = \mu + \phi_1 \mathrm{E}\big[(z_{n+3} \mid z_n, z_{n-1}, \cdots) - \mu\big] + \phi_2 \mathrm{E}\big[(z_{n+2} \mid z_n, z_{n-1}, \cdots) - \mu\big] + E\big[a_{n+4} \mid z_n, z_{n-1}, \cdots\big]$$
$$= \mu + \phi_1\big[z_n(3) - \mu\big] + \phi_2\big[z_n(2) - \mu\big]$$

The general form is

$$z_n(l) = \mu + \phi_1\big[z_n(l-1) - \mu\big] + \phi_2\big[z_n(l-2) - \mu\big], l \geq 1$$

4.4.5. MA (1) Model Forecast

$$Z_t = \mu + a_t - \theta_1 a_{t-1}, \quad a_t \equiv WN(0,\sigma^2)$$

By replacing t with t+1 then taking expectation, the equation will be

$$z_n(l) = E\left(z_{n+l}\,|z_n, z_{n-1}, \ldots\right), \quad l \ge 1$$
$$= \mu + E\left(a_{n+l}\,|z_n, z_{n-1}, \ldots\right) - \theta_1 E\left(a_{n+l-1}|z_n, z_{n-1}, \ldots\right), \quad l \ge 1$$

By iterating the expectation process step by step until the general equation for long time is formulated as

$$z_n(l) = \mu + E\left(a_{n+l}\,|z_n, z_{n-1}, \ldots\right) - \theta_1 E\left(a_{n+l-1}|z_n, z_{n-1}, \ldots\right), \quad l \ge 1$$
$$l = 1: \quad z_n(1) = \mu + E\left(a_{n+1}|z_n, z_{n-1}, \ldots\right) - \theta_1 E\left(a_n\,|z_n, z_{n-1}, \ldots\right)$$
$$= \mu - \theta_1 a_n$$
$$l = 2: \quad z_n(2) = \mu + E\left(a_{n+2}|z_n, z_{n-1}, \ldots\right) - \theta_1 E\left(a_{n+1}|z_n, z_{n-1}, \ldots\right)$$
$$= \mu$$
$$l = 3: \quad z_n(3) = \mu + E\left(a_{n+3}|z_n, z_{n-1}, \ldots\right) - \theta_1 E\left(a_{n+2}\,|z_n, z_{n-1}, \ldots\right)$$
$$= \mu$$

So, the general form can be written as

$$z_n(l) = \mu, \quad l \ge 2$$

MMSE forecast model for MA(1) models can be written as

$$z_n(l) = \begin{cases} \mu - \theta_1 a_n, & l = 1 \\ \mu, & l \ge 2 \end{cases}$$

4.4.6. MA (2) Model Forecast

$$Z_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}, \quad a_t \equiv WN(0, \sigma^2)$$
$$z_n(l) = E\left(z_{n+l}\,|z_n, z_{n-1}, \ldots\right), \quad l \ge 1$$
$$= \mu + E\left(a_{n+l}\,|z_n, z_{n-1}, \ldots\right) - \theta_1 E\left(a_{n+l-1}|z_n, z_{n-1}, \ldots\right) - \theta_2 E\left(a_{n+l-2}\,|z_n, z_{n-1}, \ldots\right), \quad l \ge 1$$

This equation can be solved by iterating and applying condition expectation rules.

$$z_n(l)=\mu+E\left(a_{n+l}\,|z_n,z_{n-1},\cdots\right)-\theta_1E\left(a_{n+l-1}|z_n,z_{n-1},\cdots\right)-\theta_2E\left(a_{n+l-2}\,|z_n,z_{n-1},\cdots\right),\quad l\geq 1$$

$$l=1:\quad z_n(1)=\mu+E\left(a_{n+1}|z_n,z_{n-1},\cdots\right)-\theta_1E\left(a_n\,|z_n,z_{n-1},\cdots\right)-\theta_2E\left(a_{n-1}|z_n,z_{n-1},\cdots\right)$$

$$=\mu-\theta_1a_n-\theta_2a_{n-1}$$

$$l=2:\quad z_n(2)=\mu+E\left(a_{n+2}|z_n,z_{n-1},\cdots\right)-\theta_1E\left(a_{n+1}|z_n,z_{n-1},\cdots\right)-\theta_2E\left(a_n\,|z_n,z_{n-1},\cdots\right)$$

$$=\mu-\theta_2a_n$$

$$l=3:\quad z_n(3)=\mu+E\left(a_{n+3}|z_n,z_{n-1},\cdots\right)-\theta_1E\left(a_{n+2}|z_n,z_{n-1},\cdots\right)-\theta_2E\left(a_{n+1}|z_n,z_{n-1},\cdots\right)$$

$$=\mu$$

So, that

$$z_n(l)=\mu,\quad l\geq 3$$

MMSE Forecast model for MA (2) models can be written as

$$z_n(l)=\begin{cases}\mu-\theta_1a_n-\theta_2a_{n-1}, & l=1\\ \mu-\theta_2a_n, & l=2\\ \mu, & l\geq 3\end{cases}$$

4.4.7. ARMA (p,q)

The general model for ARMA (p,q)

$$z_t=\delta+\phi_1z_{t-1}+\phi_2z_{t-2}+\cdots+\phi_pz_{t-p}+a_t-\theta_1a_{t-1}-\theta_2a_{t-2}-\cdots-\theta_qa_{t-q}$$

The forecast model can be written as

$$Z_n(\lambda)=\phi_1Z_n(\lambda-1)+\phi_2Z_n(\lambda-2)+\cdots+\phi_pZ_n(\lambda-p)+\theta_0-\theta_1E(e_{n+\lambda-1}\,|\,Z_1,Z_2,\cdots,Z_n)$$
$$-\theta_2E(e_{n+\lambda-2}\,|\,Z_1,Z_2,\cdots,Z_n)-\cdots-\theta_qE(e_{n+\lambda-q}\,|\,Z_1,Z_2,\cdots,Z_n)$$

Where $E(e_{n+j}\,|\,z_1,z_2,\ldots,z_n)=\begin{Bmatrix}0 & for\ j>0\\ e_{n+j} & for\ j\leq 0\end{Bmatrix}$

For $\lambda>q$ **,** the autoregressive part of the equation takes over, and the equation can be written as

$$Z_n(\lambda)=\phi_1Z_n(\lambda-1)+\phi_2Z_n(\lambda-2)+\cdots+\phi_pZ_n(\lambda-p)+\theta_0\ \text{ for }\ \lambda>q$$

As $\theta_0 = \mu(1-\phi_1-\phi_2-\Lambda\ -\phi_p)$, the equation can be written in terms of deviation from μ as

$$Z_n(\lambda)-\mu=\phi_1[Z_n(\lambda-1)-\mu]+\phi_2[Z_n(\lambda-2)-\mu]+\Lambda\ +\phi_p[Z_n(\lambda-p)-\mu] \text{ for } \lambda>q$$

For stationary ARMA model, $Z_n(\lambda)-\mu$ decays to zero as λ increases, and the long-term forecast will simply be the process mean μ (Cryer, 2008, p. 200).

4.4.8. ARMA(1,1) Model Forecast

$$Z_t=\mu+\phi_1(Z_{t-1}-\mu)+a_t-\theta_1 a_{t-1}, \phi_1\neq\theta_1, |\phi_1|<1$$

$$z_n(l)=E\left(z_{n+l}\,|z_n,z_{n-1},L\ \right),\quad l\geq 1$$

$$=\mu+\phi_1 E\left[\left(z_{n+l-1}-\mu\right)|z_n,z_{n-1},L\ \right]+E\left(a_{n+l}\,|z_n,z_{n-1},L\ \right)-\theta_1 E\left(a_{n+l-1}|z_n,z_{n-1},L\ \right),\quad l\geq 1$$

By iteration, the equation can be solved as follow

$$z_n(l)=\mu+\phi_1 E\left[\left(z_{n+l-1}-\mu\right)|z_n,z_{n-1},L\ \right]+E\left(a_{n+l}\,|z_n,z_{n-1},L\ \right)-\theta_1 E\left(a_{n+l-1}|z_n,z_{n-1},L\ \right),\quad l\geq 1$$

$$l=1:\quad z_n(1)=\mu+\phi_1 E\left[\left(z_n-\mu\right)|z_n,z_{n-1},L\ \right]+E\left(a_{n+1}|z_n,z_{n-1},L\ \right)-\theta_1 E\left(a_n|z_n,z_{n-1},L\ \right)$$

$$=\mu+\phi_1\left(z_n-\mu\right)-\theta_1 a_n$$

$$l=2:\quad z_n(2)=\mu+\phi_1 E\left[\left(z_{n+1}-\mu\right)|z_n,z_{n-1},L\ \right]+E\left(a_{n+2}|z_n,z_{n-1},L\ \right)-\theta_1 E\left(a_{n+1}|z_n,z_{n-1},L\ \right)$$

$$=\mu+\phi_1\left[z_n(1)-\mu\right]$$

$$l=3:\quad z_n(3)=\mu+\phi_1 E\left[\left(z_{n+2}-\mu\right)|z_n,z_{n-1},L\ \right]+E\left(a_{n+3}|z_n,z_{n-1},L\ \right)-\theta_1 E\left(a_{n+2}|z_n,z_{n-1},L\ \right)$$

$$=\mu+\phi_1\left[z_n(2)-\mu\right]$$

Where the general form is

$$z_n(l)=\mu+\phi_1\left[z_n(l-1)-\mu\right],\quad l\geq 2$$

MMSE Forecast model for ARMA (1,1) models can be written as

$$z_n(l)=\begin{cases}\mu+\phi_1\left(z_n-\mu\right)-\theta_1 a_n, & l=1\\ \mu+\phi_1\left[z_n(l-1)-\mu\right], & l\geq 2\end{cases}$$

4.4.9. Nonstationary Models

For nonstationary ARIMA models, forecasting is quite similar to forecasting of stationary ARMA models (Shumway, 2006, p. 140).

So that the model can be written as

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \cdots + \varphi_p Z_{t-p} + \varphi_{p+1} Z_{t-p-1} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$$

Where a coefficient φ equals

$$\varphi_1 = 1 + \phi_1$$
$$\varphi_j = \phi_j - \phi_{j-1} \text{ for j=1,2,…,p}$$

and

$$\varphi_{p+1} = -\phi_p$$

Thus, ARMA forecast equation can be used to forecast ARIMA models by replacing p with p+d and ϕ_j by φ_j.

4.4.10. ARIMA(1,1,1)

$$Z_t - Z_{t-1} = \phi(Z_{t-1} - Z_{t-2}) + \theta_0 + a_t - \theta a_{t-1}$$

So that

$$Z_t = (1+\phi)Z_{t-1} - \phi Z_{t-2} + \theta_0 + a_t - \theta a_{t-1}$$

Thus

$$Z_n(1) = (1+\phi)Z_t - \phi Z_{t-1} + \theta_0 - \theta a_n$$
$$Z_n(2) = (1+\phi)Z_n(1) - \phi Z_n + \theta_0$$
$$\vdots$$
$$Z_n(\lambda) = (1+\phi)Z_n(\lambda-1) - \phi Z_n(\lambda-2) + \theta_0$$

Forecast error

$$e_n(\lambda) = e_{n+\lambda} + \psi_1 e_{n+\lambda-1} + \psi_2 e_{n+\lambda-2} + \cdots + \psi_{\lambda-1} e_{n+1} \text{ for } \lambda \ge 1$$

So that

$$E(e_n(\lambda)) = 0 \text{ for } \lambda \ge 1$$

and

$$Var(e_n(\lambda)) = \sigma_e^2 \sum_{j=1}^{\lambda-1} \psi_j^2 \quad for \quad \lambda \ge 1$$

However, for nonstationary series, the ψ_j weights do not decay to zero as j increases (Cryer, 2008, p. 202).

Thus

$$\psi_j = 1 - \theta, \;\; j \ge 1 \text{ for IMA(1,1)}$$

$\psi_j = 1-\theta_2 + (1-\theta_1-\theta_2)_j$, $j \geq 1$ for IMA(2,2)

$\psi_j = (1-\phi^{j+1})/(1-\phi)$, $j \geq 1$ for ARI(1,1)

Thus, for nonstationary model the forecast error variance will increase without bound as the time λincrease. So with nonstationary series the distant future is quite uncertain (Brockwell, 2002, p. 180).

4.4.11. Forecasting Limits (Shumway, 2006, pp. 105-108)

Forecast function $z_n(l), l \geq 1$gives point forecast for time units λ which is not applicable or usefulness in statistical decision making where

$$P\left(Z_{n+m} = z_n(m)\right) = 0, \text{ for some } m > 0$$

This expression means that the forecast probability for any time unit λ equals zero, so the forecasts are absolutely uncertain.

This problem can be solved by using Interval Forecast Technique, so this technique decide that the forecast of the time unit λ will surely included in an interval for example $\lceil a,b \rceil$.

Thus

$$P\left(a \leq Z_{n+m} \leq b\right) = \left(1-\alpha\right)$$

this expression says that it is $100\times(1-\alpha)\%$ confident that the interval $\lceil a,b \rceil$ includes $Z_n(\lambda)$value, so if $\alpha = 0.05$, it is 95% confident that the interval $\lceil a,b \rceil$ includes $Z_n(\lambda)$value.

Chapter Five: Results and Analysis

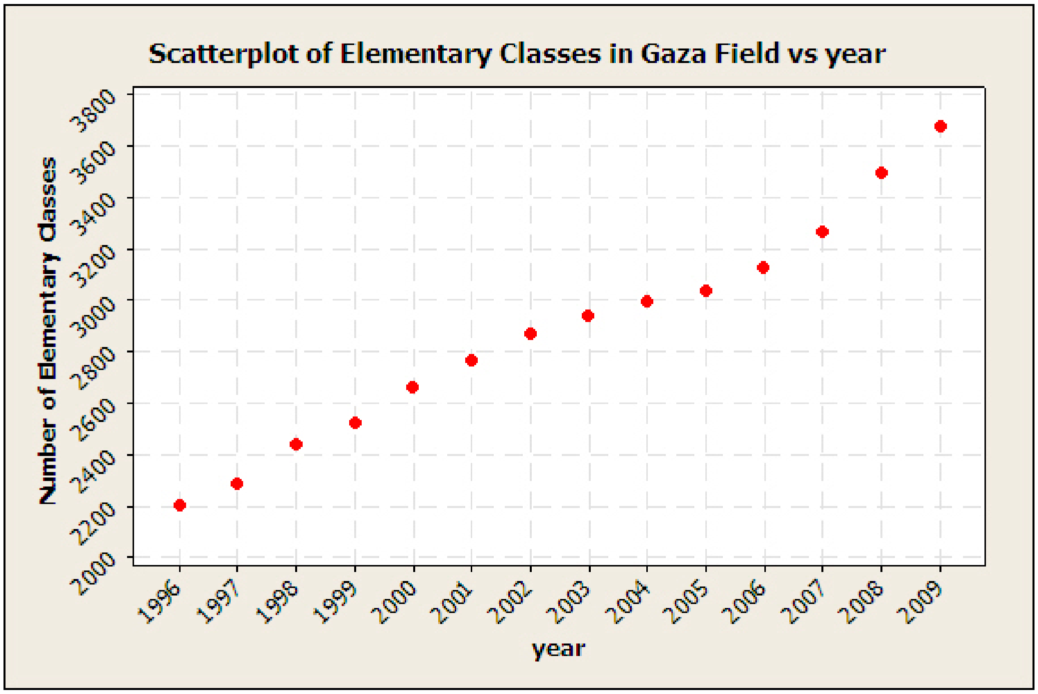
5. Results and Analysis

Now, analysis approach will be studied, so "elementary classes in Gaza Field" data was analyzed as practical case.

5.1. Scatter Time Plot

Scatter time plot was plotted by MINITAB statistical package to determine whether the time series is stationary or not.

Figure 5.1 Time series plot for elementary classes in Gaza Field



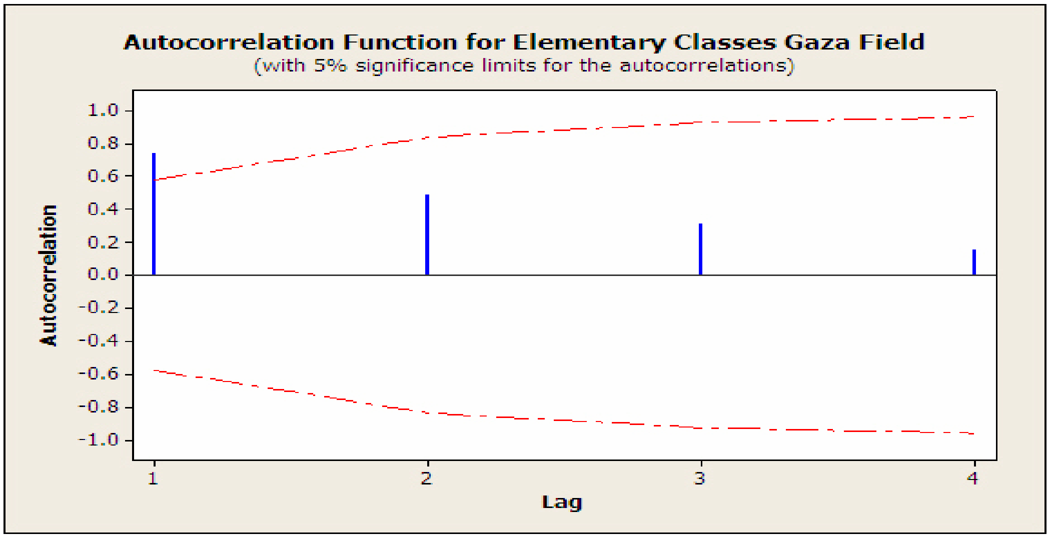
Scatter plot Analysis

Scatter plot indicated linear trend, so it was reasonable to believe that the time series is nonstationary.

5.2. **Model Identification**

autocorrelation function and partial autocorrelation functions were calculated and plotted to examine and classify the behavior of ACF and PACF then identify the model specifications. Minitab program shows the autocorrelations, associated t-statistics, and Ljung-Box Q statistics.

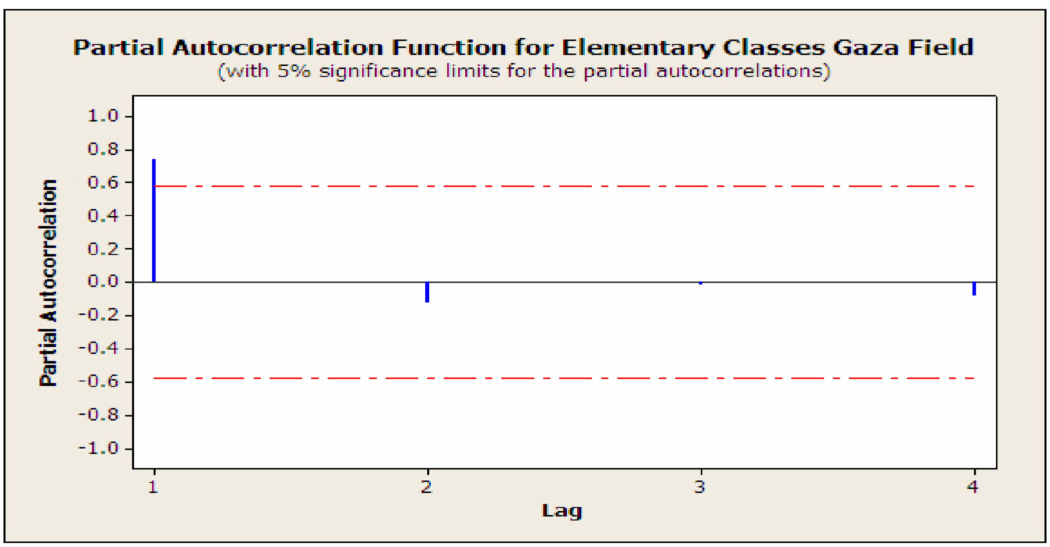
Figure 5.2 Autocorrelation function for Elementary classes in Gaza Field



Autocorrelation Function of Elementary Classes Gaza Field

Lag	ACF	T	LBQ
1	0.737522	2.76	9.37
2	0.487170	1.26	13.80
3	0.305319	0.71	15.70
4	0.151160	0.34	16.21

Figure 5.3 Partial Autocorrelation Function for Elementary classes in Gaza Field



Partial Autocorrelation Function of Elementary Classes Gaza Field

Lag	PACF	T
1	0.737522	2.76

```
  2  -0.124478  -0.47
  3  -0.015365  -0.06
  4  -0.079529  -0.30
```

5.2.1. Analysis of ACF and PACF Behavior and Model Identification

Since, autocorrelation function dies down fairly quickly after lag 1 and Partial autocorrelation cuts off after lag 1. It was concluded that the model seems to be ARIMA models with order not more than 1.

5.3. Applying ARIMA Models

MINITAB Package was used to apply consequently ARIMA models with order not more than 1 and differencing not more than 2.

ARIMA Model: Elementary Classes Gaza Field 0 1 1

```
Final Estimates of Parameters

Type         Coef  SE Coef      T      P
MA   1    -0.3546   0.2817  -1.26  0.234
Constant   112.69    19.13   5.89  0.000


Differencing: 1 regular difference
Number of observations:  Original series 14, after differencing 13
Residuals:    SS =  28658.2 (backforecasts excluded)
              MS =  2605.3  DF = 11


Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag            12  24  36  48
Chi-Square   10.4   *   *   *
DF             10   *   *   *
P-Value     0.410   *   *   *


Forecasts from period 14

                    95% Limits
Period  Forecast    Lower    Upper  Actual
    15   3793.41  3693.35  3893.48
    16   3906.10  3737.62  4074.58
    17   4018.80  3802.56  4235.03
    18   4131.49  3876.28  4386.70
    19   4244.18  3955.20  4533.15
    20   4356.87  4037.68  4676.05
    21   4469.56  4122.79  4816.33
    22   4582.25  4209.93  4954.58
    23   4694.94  4298.71  5091.17
    24   4807.63  4388.86  5226.41
```

ARIMA Model: Elementary Classes Gaza Field 0 2 1

```
Final Estimates of Parameters

Type         Coef  SE Coef     T      P
MA   1    0.3117   0.2996  1.04  0.323
Constant    7.43    10.95  0.68  0.513


Differencing: 2 regular differences
Number of observations:  Original series 14, after differencing 12
Residuals:    SS =  30134.1 (backforecasts excluded)
              MS =  3013.4  DF = 10


Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag         12  24  36  48
Chi-Square   *   *   *   *
DF           *   *   *   *
P-Value      *   *   *   *


Forecasts from period 14

                     95% Limits
Period  Forecast    Lower    Upper  Actual
    15   3866.16  3758.54  3973.77
    16   4066.74  3855.57  4277.91
    17   4274.75  3943.07  4606.42
    18   4490.18  4022.42  4957.94
    19   4713.04  4095.03  5331.06
    20   4943.33  4162.04  5724.62
    21   5181.04  4224.39  6137.69
    22   5426.18  4282.84  6569.52
    23   5678.74  4338.03  7019.46
    24   5938.74  4390.50  7486.97
```

ARIMA Model: Elementary Classes Gaza Field 1 1 0

```
Final Estimates of Parameters

Type         Coef  SE Coef     T      P
AR   1    0.5027   0.2827  1.78  0.103
Constant   57.47    13.86  4.15  0.002


Differencing: 1 regular difference
Number of observations:  Original series 14, after differencing 13
Residuals:    SS =  26757.6 (backforecasts excluded)
              MS =  2432.5  DF = 11


Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag            12  24  36  48
Chi-Square   12.0   *   *   *
DF             10   *   *   *
P-Value     0.286   *   *   *


Forecasts from period 14

                     95% Limits
Period  Forecast    Lower    Upper  Actual
```

```
    15   3818.44  3721.75  3915.13
    16   3949.02  3774.49  4123.54
    17   4072.12  3828.68  4315.56
    18   4191.47  3887.52  4495.43
    19   4308.94  3951.44  4666.43
    20   4425.45  4020.00  4830.91
    21   4541.49  4092.52  4990.47
    22   4657.29  4168.34  5146.24
    23   4772.97  4246.93  5299.01
    24   4888.59  4327.84  5449.33
```

ARIMA Model: Elementary Classes Gaza Field 1 1 1

```
Final Estimates of Parameters

Type         Coef  SE Coef     T      P
AR   1     0.5098   0.9039  0.56  0.585
MA   1     0.0082   0.9633  0.01  0.993
Constant    56.70    16.41  3.45  0.006


Differencing: 1 regular difference
Number of observations:  Original series 14, after differencing 13
Residuals:    SS =  26760.4 (backforecasts excluded)
              MS =  2676.0  DF = 10


Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag            12  24  36  48
Chi-Square   11.9   *   *   *
DF              9   *   *   *
P-Value     0.217   *   *   *


Forecasts from period 14

                     95% Limits
Period  Forecast    Lower    Upper  Actual
    15   3818.92  3717.50  3920.33
    16   3950.00  3767.04  4132.96
    17   4073.53  3818.12  4328.94
    18   4193.20  3874.01  4512.38
    19   4310.91  3935.21  4686.60
    20   4427.61  4001.24  4853.98
    21   4543.81  4071.42  5016.19
    22   4659.74  4145.09  5174.39
    23   4775.54  4221.67  5329.41
    24   4891.27  4300.71  5481.84
```

ARIMA Model: Elementary Classes Gaza Field 1 2 0

```
Final Estimates of Parameters

Type          Coef  SE Coef      T      P
AR   1     -0.4504   0.3061  -1.47  0.172
Constant     11.21    15.29   0.73  0.480


Differencing: 2 regular differences
Number of observations:  Original series 14, after differencing 12
Residuals:    SS =  27928.9 (backforecasts excluded)
              MS =  2792.9  DF = 10
```

```
Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag         12  24  36  48
Chi-Square   *   *   *   *
DF           *   *   *   *
P-Value      *   *   *   *


Forecasts from period 14

                     95% Limits
Period  Forecast    Lower    Upper  Actual
    15   3886.24  3782.63  3989.84
    16   4093.47  3902.40  4284.54
    17   4314.61  4009.02  4620.21
    18   4540.71  4107.54  4973.87
    19   4775.78  4199.73  5351.84
    20   5018.03  4286.69  5749.37
    21   5268.26  4369.57  6166.95
    22   5526.11  4449.08  6603.13
    23   5791.74  4525.91  7057.56
    24   6065.08  4600.59  7529.56
```

ARIMA Model: Elementary Classes Gaza Field 1 2 1

```
Final Estimates of Parameters

Type         Coef  SE Coef       T      P
AR   1    -0.9994   0.0303  -33.01  0.000
MA   1    -0.9079   0.1980   -4.59  0.001
Constant    13.18    25.95    0.51  0.624


Differencing: 2 regular differences
Number of observations:  Original series 14, after differencing 12
Residuals:    SS =  19965.5 (backforecasts excluded)
              MS =  2218.4  DF = 9

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag         12  24  36  48
Chi-Square   *   *   *   *
DF           *   *   *   *
P-Value      *   *   *   *


Forecasts from period 14

                     95% Limits
Period  Forecast    Lower    Upper  Actual
    15   3893.57  3801.24  3985.91
    16   4081.78  3882.84  4280.73
    17   4315.51  3981.30  4649.73
    18   4516.93  4031.22  5002.64
    19   4763.83  4105.24  5422.41
    20   4978.45  4133.93  5822.97
    21   5238.51  4190.26  6286.76
    22   5466.34  4203.13  6729.54
    23   5739.55  4245.83  7233.28
    24   5980.59  4246.36  7714.82
```

5.4. Model Fitting Criteria

to fit the best model, different information criteria was used, Akiak's Information Criteria (AIC)(see section 4.3.5.1) and Bayesian Information Criteria (BIC))(see section 4.3.5.2) in addition to accuracy measures MAPE, MAD and MSD.(see section 3.1.2)

The best fit model minimizes AIC and BIC value and has the least accuracy measures MAPE, MAD and MSD.

Table 5.1 AIC, BIC and accuracy measures values

Model	AIC	BIC	MSD	MAD	MAPE	n	σ	m
1 2 1	192.909	194.8	1663.80	32.60	1.07	12	2218.4	4
1 2 0	196.436	197.9	2327.41	42.21	1.41	12	2792.9	3
0 2 1	198.26	199.7	2511.17	45.19	1.53	12	3013.4	3
1 1 0	206.714	207.8	2058.27	37.03	1.28	13	2432.5	2
0 1 1	208.498	209.6	2204.47	37.81	1.30	13	2605.3	2
1 1 1	211.194	212.9	2058.49	37.05	1.29	13	2676	3

Note:-The table summarizes the values of AIC and BIC information criteria in addition to accuracy measures MAPE, MAD and MSD for ARIMA models.

5.5. Fitting The Likely Model

1. Table 5.1 indicated that (1 2 1) model has the least values of AIC, BIC from among models.
2. Table 5.1 indicated that (1 2 1) model has the least values of accuracy measures MAPE, MAD and MSD.
3. P-value for each coefficient tested the null hypothesis that the coefficient is equal to zero (no effect). Therefore, low p-values refer that the coefficient was a meaningful addition to the model.

From (1 2 1) model application, the estimated coefficients were produced. Table 5.2 summarizes the model results

(1 2 1) model results in table 5.2 referred to that both coefficient are significant because of their low p-values since the p-value for AR (1) coefficient is 0.000 and the p-value for MA (1) coefficient is 0.001 which means that p-value is less than 0.05.

Table 5.2 Final Estimates of Parameters for (1 2 1) model

type	coefficient	value	t-value	p-value
AR 1	$\hat{\phi}$	-0.9994	-33.01	0.000
MA 1	$\hat{\theta}$	-0.9079	-4.59	0.001

So, it was concluded that (1 2 1)model is the best fit model to represents the observations of Elementary Classes in Gaza Field time series

5.6. Diagnostic Check on The Mean of The Best Fit Model Residuals

One sample t-test was used to compute a confidence interval and perform a hypothesis test of the mean when the population standard deviation, s, is unknown.

For a two-tailed one-sample t
H_0: $\mu = 0$ versus H_1: $\mu \neq 0$

```
One-Sample T: Residuals of (1 2 1) model

Test of μ = 0 vs not = 0

N      Mean   StDev         SE Mean       T             P
12   0.8      42.6          12.3          0.06          0.950
```

Because the resulting p-value (0.95) was greater than chosen α-level 0.05, it was declared that statistical no significance and can't reject the null hypothesis, so it was concluded that data mean μ equals zero (μ=0)

5.7. Diagnostic Check on Residuals Randomness

Runs Test was used in order to check the randomness of the best fit (1 2 1) model residuals or to see if a residuals order is random.

H0: that the data are in random sequence
H: that the data are not in random sequence

```
Runs test for Residuals of (1 2 1) model

Runs above and below K = 0.790825

The observed number of runs = 8
```

```
The expected number of runs = 6.83333
7 observations above K, 5 below
* N is small, so the following approximation may be invalid.
```
P-value = 0.466

Because the resulting p-value (0.466) was greater than the alpha level of 0.05, there was sufficient evidence to conclude that the data are in random sequence.

5.8. Diagnostic Check on ACF and PACF of Residuals

MINITAB package was used to plot and calculate ACF and PACF of residuals, the graphs for the ACF and PACF of the ARIMA residuals included lines representing two standard errors to either side of zero. Values that extend beyond two standard errors are statistically significant at approximately α = 0.05, and show evidence that the model has not explained all autocorrelation in the data.

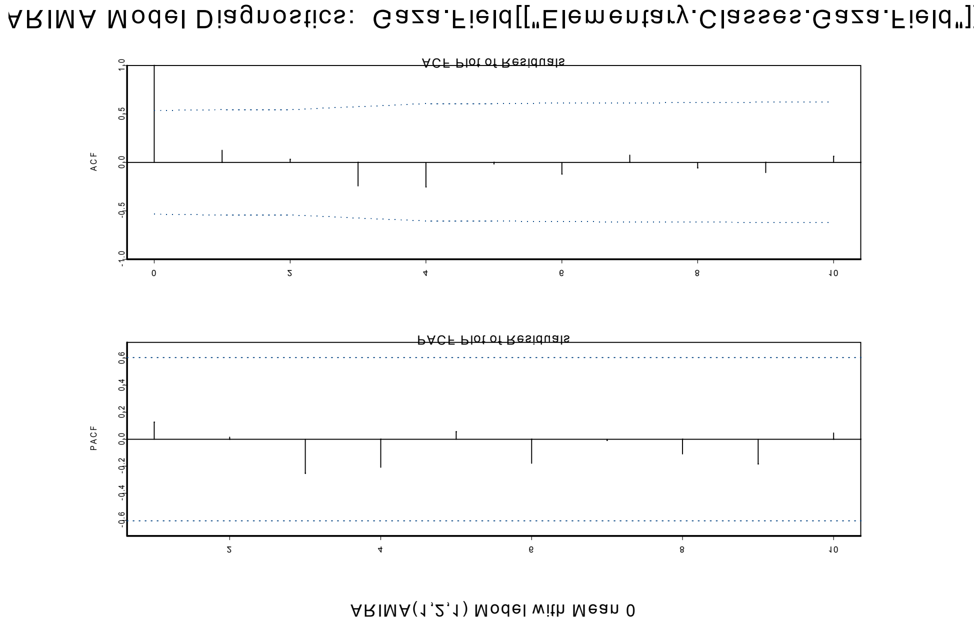
Autocorrelation for model(1 2 1) residuals

Lag	ACF	T	LBQ
1	0.023285	0.08	0.01
2	0.134194	0.46	0.31
3	-0.186103	-0.63	0.96

Partial Autocorrelation for model(1 2 1) residuals

Lag	PACF	T
1	0.023285	0.08
2	0.133725	0.46
3	-0.195524	-0.68

Figure 5.4 ACF and PACF plot for residuals of the best fit model



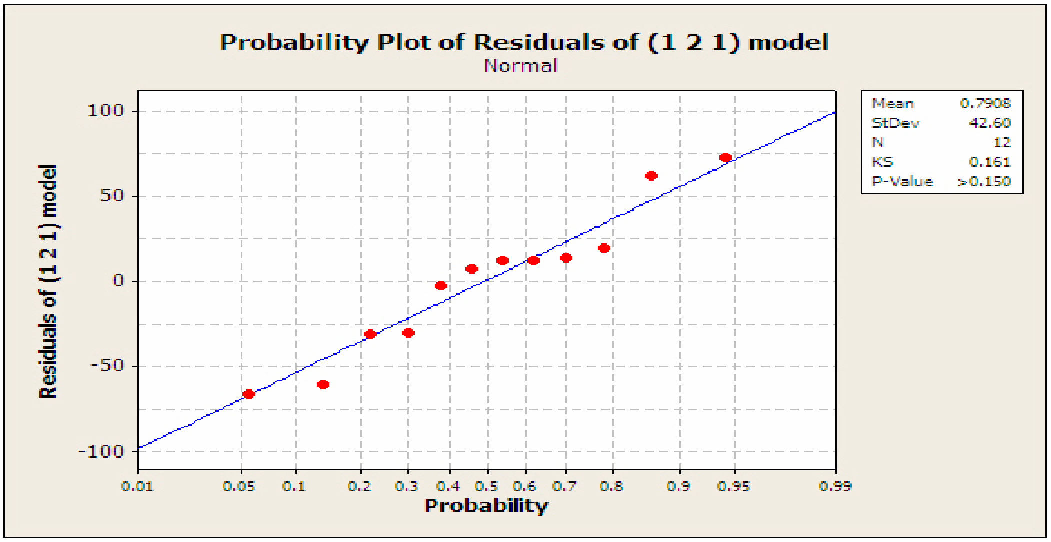
the graphs for the ACF and PACF indicated that the residuals appeared to uncorrelated. It was assumed that the spike in the ACF at lag 0 was the result of random

5.9. Diagnostic Check on Residuals Normality

5.9.1. Normal Probability Plot

For Normal Probability Plot of residuals, the points in this plot should generally form a straight line if the residuals are normally distributed. If the points on the plot depart from a straight line, the normality assumption may be invalid.

Figure 5.5 Normal probability plot for residual of the best fit model



Because the points were in a linear pattern, the plot indicated that the (1 2 1) model residuals follow normal distribution.

5.9.2. Kolmogorov-Smirnov Normality Test(KS)

This test compares the empirical cumulative distribution function of our residuals data with the expected distribution.

The hypotheses of Kolmogorov-Smirnov normality test (KS) are,

H0: data follow a normal distribution

H1: data do not follow a normal distribution

```
One sample Kolmogorov-Smirnov Test of Composite Normality

data:  Residuals.of..1.2.1..model in Gaza.Field
ks = 0.1614, p-value = 0.5
alternative hypothesis: True cdf is not the normal distn. with estimated
parameters
sample estimates:
 mean of x standard deviation of x
 0.7908252                42.59539
```

The p-value (0.5) was more than the chosen α-level (0.5), the null hypothesis can not rejected and it was concluded that the residual of (1 2 1) model are normal.

5.10. Ten-Step Ahead Forecasts

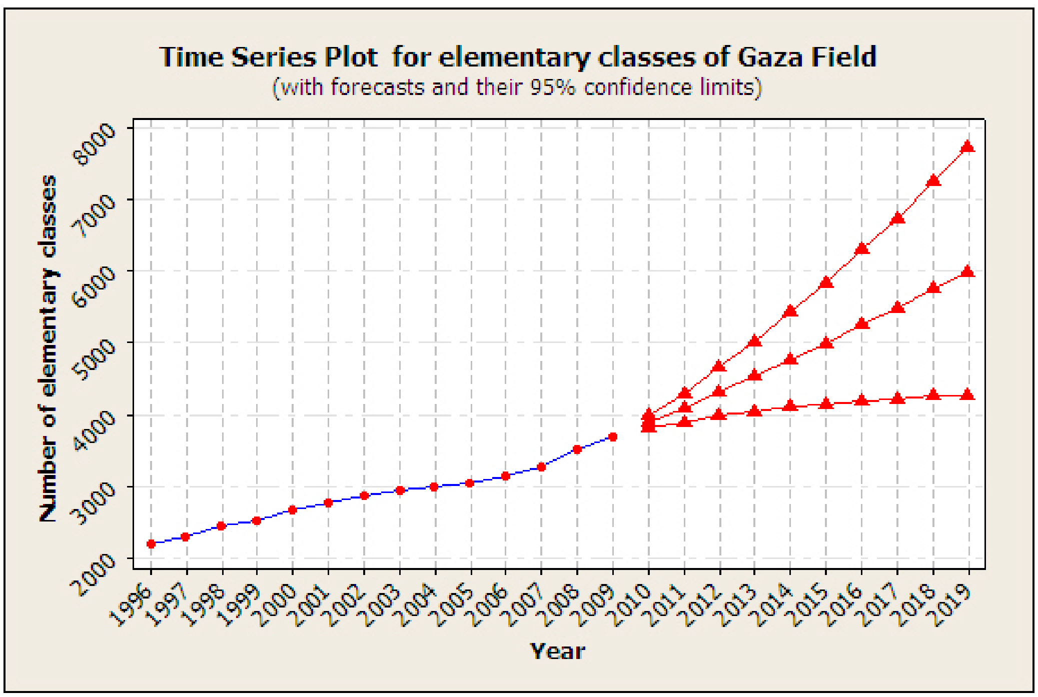
ARIMA (1 2 1) model was chosen to fit the observed data so it was used to forecast ten-step-ahead, the model gave forecasts, with 95% confidence limits

Ten-step-ahead forecasts were predicted for elementary classes in Gaza Field.

Forecasts from period 2009

		95% Limits	
Period	Forecast	Lower	Upper
2010	3893.57	3801.24	3985.91
2011	4081.78	3882.84	4280.73
2012	4315.51	3981.30	4649.73
2013	4516.93	4031.22	5002.64
2014	4763.83	4105.24	5422.41
2015	4978.45	4133.93	5822.97
2016	5238.51	4190.26	6286.76
2017	5466.34	4203.13	6729.54
2018	5739.55	4245.83	7233.28
2019	5980.59	4246.36	7714.82

Figure 5.6 Time series plot including forecasts for the best fit model



5.11. Proposed Model

From (3.2.2) section, ARIMA (1 2 1) model is,

$$\phi_1(B)z_t = \delta + \theta_1(B)a_t$$

$$(1-\phi_1 B)z_t = \delta + (1-\theta_1 B)a_t$$

$$z_t - \phi_1 z_{t-1} = \delta + a_t - \theta_1 a_{t-1}$$

$$z_t = \delta + \phi_1 z_{t-1} + a_t - \theta_1 a_{t-1},\ a_t : WN(0,\sigma^2),\ \phi_1 \neq \theta_1$$

By substitution of coefficients value from table 5-2, the ARIMA (1 2 1) model is

$$Z_t = 13.18 - 0.9994 Z_{t-1} + a_t + 0.9079 a_{t-1}$$

Where a_t is white noise series with mean equals zero and variance equals 2218.4 $a_t \equiv WN(0,2218.4)$

5.12. Summary of Proposed Models Properties

Here are summary tables for 42 proposed models; related figures will be included in indices.

Note for all tables: - C S=number of schools; C E=number of Elementary classrooms; C P B= number of Preparatory Boys classrooms; C P G= number of Preparatory Girls classrooms; P E= number of Elementary Pupils; P P B= number of Preparatory Boys Pupils; P P G= number of Preparatory Girls Pupils.

Table 5.3 Summary for number of school models

Area	Best Model	Parameters	Forecasts									
			2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Rafah	1 1 0	AR(1)=-0.2330 onstant= 1.2514	41	42	43	44	45	46	47	48	49	50
Khan	2 2 1	AR(1)= -1.1356 , AR(2)=-0.8833 MA(1)= 0.8892 constant= 0.59728	30	30	33	37	38	42	45	47	52	56
Khan\E	0 2 1	MA(1)=0.9136 constant= 0.06662	19	20	20	21	22	23	24	26	27	28
Middle	0 1 1	MA(1)=-0.5610 constant=1.4014	52	53	55	56	58	59	60	62	63	65
Gaza	1 2 1	AR(1)= -0.4411 MA(1)= 1.2096 constant= 0.61092	65	69	75	80	86	93	100	107	115	123
North	2 1 1	AR(1)= 0.7376, AR(2)=-0.8679 MA(1)= 0.9503 constant= 0.676638	38	39	40	40	40	41	42	43	43	43
Field	1 2 0	AR(1)= -0.7746 constant= 0.936	238	249	259	271	282	295	308	321	335	350

Table 5.4 Summary for number of elementary classroom models

Area	Best Model	Parameters	Forecasts									
			2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Rafah	1 1 0	AR(1)= -0.9004 constant=30.347	578	619	612	649	646	679	680	709	713	740
Khan	0 2 1	MA(1)=0.8272 constant=0.1127	348	358	368	379	389	400	411	422	432	444
Khan\E	0 1 1	MA(1)= -0.8756 constant= 12.441	368	380	393	405	418	430	443	455	468	480
Middle	1 2 0	AR(1)=-0.8820 constant=1.319	812	835	882	909	955	985	1031	1064	1109	1146
Gaza	1 1 2	AR(1)= 1.2805 MA(1)= -0.0475, MA(2)= -0.8414 constant= 18.108	993	1023	1054	1087	1120	1153	1187	1220	1253	1287
North	0 2 1	MA(1)=0.8310 constant=-1.1022	701	713	725	735	744	752	759	765	770	774
Field	1 2 1	AR(1)= -0.9994 MA(1)=-0.9079 constant=13.18	3894	4082	4316	4517	4764	4978	5239	5466	5740	5981

Table 5.5 Summary for number of boy preparatory classroom models

Area	Best Model	Parameters	Forecasts 2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Rafah	**0 1 1**	MA(1)=1.2789 constant=9.981	**189**	**199**	**209**	**219**	**229**	**239**	**249**	**259**	**269**	**279**
Khan	**1 1 1**	AR(1)=-0.5963 MA(1)=1.1685 constant=7.04158	**93**	**105**	**105**	**112**	**115**	**120**	**124**	**129**	**133**	**138**
Khan\E	**0 1 1**	MA(1)= 0.8822 constant=3.3575	**72**	**76**	**79**	**82**	**86**	**89**	**92**	**96**	**99**	**102**
Middle	**0 1 1**	MA(1)=0.9016 constant=11.3704	**228**	**239**	**250**	**262**	**273**	**284**	**296**	**307**	**319**	**330**
Gaza	**2 2 0**	AR(1)=-1.3560, AR(2)=-1.0334 constant=7.101	**315**	**317**	**361**	**413**	**417**	**485**	**523**	**543**	**626**	**649**
North	**1 1 1**	AR(1)=-0.3814 MA(1)=0.9030 constant=16.0508	**214**	**226**	**238**	**249**	**261**	**273**	**284**	**296**	**307**	**319**
Field	**1 1 0**	AR(1)= -0.6826 constant= 85.59	**1047**	**1135**	**1160**	**1228**	**1267**	**1326**	**1372**	**1426**	**1475**	**1527**

Table 5.6 Summary for number of girl preparatory classroom models

Area	Best Model	Parameters	Forecasts									
			2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Rafah	1 1 0	AR(1)= -0.6760 constant= 9.653	143	140	152	154	162	166	173	178	184	190
Khan	0 1 1	MA(1)=0.8695 constant= 3.5579	91	96	100	104	108	112	116	120	124	128
Khan\E	0 1 1	MA(1)=0.8727 constant= 2.8394	65	68	71	74	77	79	82	85	88	91
Middle	1 1 0	AR(1)=0.4473 constant=3.419	7399	7535	7699	7882	8080	8290	8507	8731	8959	9191
Gaza	0 2 2	MA(1)= 0.7929, MA(2)= -0.8824 constant=0.783	182	192	203	215	228	241	255	270	286	303
North	1 2 1	AR(1)= -1.0006 MA(1)=-0.9141 constant=-0.378	153	156	156	158	158	160	159	161	159	161
Field	0 1 1	MA(1)= 0.8570 constant=37.440	856	893	931	968	1006	1043	1080	1118	1155	1193

Table 5.7 Summary for number of elementary pupil models

Area	Best Model	Parameters	Forecasts									
			2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Rafah	1 2 0	AR(1)= -0.6610 Constant= -35.1	**24613**	**25517**	**26778**	**27768**	**28903**	**29907**	**30962**	**31948**	**32944**	**33899**
Khan	0 1 2	MA(1)= -0.4657 MA(2)= -0.8958 Constant= 223.5	**13997**	**14290**	**14514**	**14737**	**14961**	**15184**	**15408**	**15631**	**15855**	**16078**
Khan\E	0 2 1	MA(1)=0.8440 Constant=19.86	**14609**	**15220**	**15850**	**16501**	**17171**	**17861**	**18571**	**19301**	**20051**	**20821**
Middle	0 2 2	MA(1)=0.4153 MA(2)=-0.8036 Constant= -77.2	**30073**	**30885**	**31621**	**32279**	**32859**	**33363**	**33789**	**34138**	**34410**	**34605**
Gaza	2 2 2	AR(1)= 0.0694 AR(2)= 0.8300 MA(1)= 0.3895 MA(2)= -1.0661 Constant= 97.35	**43439**	**50257**	**60036**	**70149**	**82840**	**96086**	**111607**	**127844**	**146116**	**165221**
North	0 1 2	MA(1)=-0.6509 MA(2)=-1.1034 Constant=855.3	**27951**	**29157**	**30012**	**30867**	**31722**	**32578**	**33433**	**34288**	**35144**	**35999**
Field	0 1 2	MA(1)= -0.4967 MA(2)= -0.9117 Constant=3780	**153274**	**158989**	**162768**	**166548**	**170328**	**174108**	**177888**	**181668**	**185448**	**189228**

Table 5.8 Summary for number of boy preparatory pupil models

Area	Best Model	Parameters	Forecasts									
			2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Rafah	**1 1 0**	AR(1)= 0.7266 constant=41.28	**5462**	**5497**	**5564**	**5653**	**5760**	**5878**	**6006**	**6140**	**6278**	**6420**
Khan	**1 1 1**	AR(1)= -1.0048 MA(1)=-0.8653 constant=160.59	**3008**	**3169**	**3168**	**3329**	**3328**	**3490**	**3488**	**3650**	**3648**	**3811**
Khan\E	**1 1 0**	AR(1)= -0.6767 constant= 87.88	**2037**	**2259**	**2196**	**2326**	**2326**	**2414**	**2443**	**2511**	**2553**	**2613**
Middle	**2 1 1**	AR(1)= -0.2366 , AR(2)= 0.7637 MA(1)= -0.8467 constant= 54.1	**6558**	**6673**	**6513**	**6692**	**6582**	**6799**	**6718**	**6957**	**6892**	**7144**
Gaza	**2 1 2**	AR(1)= 1.7473, AR(2)= -0.9994 MA(1)= 1.6022, MA(2)= -0.6705 constant=71.557	**7813**	**8124**	**8615**	**9233**	**9896**	**10506**	**10982**	**11276**	**11384**	**11352**
North	**1 1 0**	AR(1)= 0.8467 constant= 21.36	**5986**	**5845**	**5748**	**5686**	**5656**	**5651**	**5669**	**5705**	**5757**	**5823**
Field	**1 1 0**	AR(1)=0.5181 constant=471.6	**31938**	**32697**	**33562**	**34482**	**35430**	**36393**	**37363**	**38338**	**39314**	**40291**

Table 5.9 Summary for number of girl preparatory pupil models

Area	Best Model	Parameters	Forecasts									
			2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Rafah	**0 1 1**	MA(1)= 0.8744 constant= 250.10	**6253**	**6503**	**6753**	**7003**	**7253**	**7504**	**7754**	**8004**	**8254**	**8504**
Khan	**0 1 1**	MA(1)= 0.5284 constant=106.87	**3368**	**3475**	**3582**	**3688**	**3795**	**3902**	**4009**	**4116**	**4223**	**4330**
Khan\E	**0 1 1**	MA(1)=0.8852 constant=82.85	**2555**	**2638**	**2720**	**2803**	**2886**	**2969**	**3052**	**3135**	**3217**	**3300**
Middle	**1 1 2**	AR(1)= 0.8441 MA(1)= -0.8467, MA(2)= -0.9860 constant= 31.97	**7334**	**7402**	**7491**	**7599**	**7721**	**7857**	**8003**	**8159**	**8322**	**8492**
Gaza	**1 1 1**	AR(1)= -0.6083 MA(1)=0.8362 constant=-61.63	7630	7874	8136	8395	8654	8913	9172	9432	9691	9950
North	**1 1 2**	AR(1)= 0.7747 MA(1)= 1.1702, MA(2)= -1.4066 constant= 113.025	**6348**	**6394**	**6542**	**6770**	**7059**	**7397**	**7771**	**8174**	**8599**	**9042**
Field	**1 1 0**	AR(1)=-0.2420 constant=1488.6	**32383**	**33550**	**34756**	**35953**	**37152**	**38350**	**39549**	**40747**	**41946**	**43145**

Chapter Six: Conclusions and Recommendations

6. Conclusions and Recommendations

6.1. Summary of Results and Implication

This section summarizes the findings of the research, and then provides recommendations and suggestions for future research. Now, the main findings of the thesis will be summarized according to their order in research.

6.1.1. Primary Data Analysis

The main aim of this study is to build accurate model for number of schools, pupils and classrooms for UNRWA agency in Gaza Strip. The review of forecasting techniques in chapter two indicates that there are a wide range of forecasting techniques that vary in scope, forecasting horizon, cost, accuracy, and the most difficult mission is to find the most suitable method which satisfies the study goals and aims, taking into account the study environment.

The review reveals that ARIMA models seem to be appropriate for forecasting the number of schools, pupils and classrooms for UNRWA agency in Gaza Strip. In chapter five, the properties of each time series were illustrated.

The main findings is that all the series have positive linear trend with constant variation. The implication of these findings is that these series can be easily modeled by ARIMA models without needing for logarithmic transformation which is used in variable variation situation.

The existence of autocorrelation or autocorrelation scale between observations at several lags was checked. The test results indicate that most time series autocorrelations or partial autocorrelations are not significant after lag 1. So most forecasting models were fitted with order not more than AR(2) and MA(2).

6.1.2. Modeling Procedures

Forecasting techniques, the ARIMA models, were applied with different orders based on ACF and PACF plots.

The selection of ARIMA models with different AR/AM orders were applied in chapter five section (5.2.), this approach differs from Box and Jenkins approach which is proposed by Box and Jenkins in chapter two sections (2.9.)

and ended with one selection. The approach used in this study, extends the total number of ARIMA models to at least six models.

AIC and BIC values were used to determine the best fit model among the alternatives models. These criteria can be used in several fields such as finance, econometric and other fields. In the same time, these criteria were applied in the study at fitting stage. The best fit model provided the least AIC/BIC values.

6.1.3. Evaluation of Forecasts

The previous procedure ended with two or three models then their correspondent forecasts were checked. The forecasts of ARIMA models were evaluated by three different measures namely, accuracy measures MAD, MSE and MAPE taking into consideration the forecasts horizon and the time series size.

The results from ARIMA models with better AIC/BIC values produced better accuracy measures than those with higher AIC/BIC values. At the same time, the results indicated that ARIMA models produced better accuracy measures than those with higher AR/AM order and higher differencing. Hence the over-fitting of high order ARIMA models and over-differencing was avoided.

The findings from the comparison between accuracy measures for one ARIMA model indicated that their values seemed to show similarity and equality. The implication was that ARIMA models often have the same rank from the comparison according to AIC/BIC values or accuracy measures. The selected models were ranked from the best fit model to the worse.

The findings indicated that the concerned variables for all cases increased with different patterns, this result was compatible with that from Aljabre (2004). This could be because the changes of number of schools and classrooms related to changes of number of pupils which had positive growth rate in all Arabic countries.

6.1.4. Evaluation of The Best Fit Model

T-test was applied to check zero mean hypotheses for the best fit model residuals. The findings from the t-test indicated that all the best fit models were significant at 0.05 confidence level. Though, the same findings were true for the results of Runs-test which checked that the residuals of best fit model were randomly distributed around zero mean.

The findings from the KS test, checked if the residuals of best fit model were normally distributed, indicated that the best fit model were significant at 0.05 confidence level. However, a few number of models showed clear in-normality. So, this model was rejected and other model in rank two was diagnosed. The implication was that the model with best AIC/BIC values and accuracy measures may be found not significant at 0.05 confidence level in normality test (KS).

6.2. Conclusion

As there is no single forecasting method that can forecast accurately for all situation, the concerned situation was studied and indicated that ARIMA models with different AR/AM order were the most appropriate technique. A collection of ARIMA models according to ACF and PACF plot were applied, and then the best fit model was chosen after a comparison between these models according to AIC/BIC values and accuracy measures.

Then, the residual of best fit model was diagnosed on zero mean, random distribution and normal distribution. If the best fit model passed these diagnostic stages. It will be considered a proposed model and mathematically formulated.

6.3. Recommendations

The goal of these recommendations is to provide generally policy makers and decision makers in UNRWA agency and especially in education department, the data that could be used to make efficient decisions which will guide the Palestinian refugee society to real educational improvement.

6.3.1. Recommendation for Decision Makers

In such dynamic political area, the administration needs long term plans to avoid future obstacles and to be ready to over come any crisis. So here are recommendations for policy maker in education administration in UNRWA agency.

A. Maintaining long term plans based on scientific future forecasts.

B. Monitoring the forecast model and make continuous improvement according to developments.

C. Pay attention that long term plans should not be changed according to the changes in the head of administration.

D. Searching for fund in earlier time to avoid fund crisis.

E. Taking into consideration changes in population growth rate according to political situations.

F. Maintaining detailed records for education field.

G. Pay attention that these forecasting models can be used if the situations in Gaza Strip still stable without impact changes especially in demographic structure.

6.3.2. Recommendation for Future Researches

Here are some recommendations for future research:

A. Conduct modeling by other forecasting methods.

B. Study the effects of UNRWA annual budget on the number of schools and classrooms. This factor may be the basic mover for UNRWA activities as UNRWA is humanitarian agency and its plans depend totally on donations size.

References

Books:

[1] Bowerman,o'connell,Koehler, Forecasting, time series, and regression an applied approach, fourth edition, Curt Hinrichs, 2005.

[2] Brockwell, P. J., Davis, R., A., Introduction to Time Series and Forecasting, Second Edition, Springer Science+Business Media, LLC, 2002.

[3] Chatfield, C., The Analysis of Time Series, 5th ed., Chapman & Hall, New York, NY. 1996.

[4] Cryer, J. D., Chan, K., Time Series Analysis, Second Edition, Springer Science+Business Media, LLC, 2008.

[5] Curwin, J., Slater, R., Quantitative Methods for Business Decisions, Fourth Edition, An International Thomson Publishing Company, 2000.

[6] Gilchrist, W., Statistical Forecasting, John Wiley & Sons Ltd, 1976.

[7] Hamilton, J. D., Time series analysis, Princeton university press, United State of America, 1994.

[8] Hanke J., Reitsch, A., Understanding Business Statistics, Richard D. IRWIN INC., 1991.

[9] Jordan, E., Clark, C., Introduction to Business and Economics Statistic, 7e, South-western Publishing Co.,Ohio, USA, 1985.

[10] Render, B., Stair, R., Hanna, M., Quantitative Analysis for Management, 9th ed., Allyn & Bacon, Boston, 2006.

[11] Shumway, R. H., Stoffer, D. S., Time Series Analysis and Its Applications, Second Edition, Springer Science+Business Media, LLC, 2006.

[12] Stevenson, w., J., Introduction to Management Science, Richard D. IRWIN INC., 1989.

Articles:

[1] Alzaroo, S., Gillian L., *Education in the context of Conflict and Instability: The Palestinian Case*, Social Policy and Administration, Vol. 37, No. 2, April 2003.

[2] Becker, O., Leitner J., Leopold U., *Modeling Expectation Formation Involving Several Sources of Information*, Verein fu¨r Socialpolitik and Blackwell Publishing Ltd. 2008

[3] Brooks, C., Tsolacos, S., *Forecasting models of retail rents*, ISMA Centre, Department of Economics, 1999

[4] Davidson R., Zhao H., Kumar V., *Quantitative Model to Forecast Changes in Hurricane Vulnerability of Regional Building Inventory*, Journal of Infrastructure Systems, Vol. 9, No. 2, 2003

[5] Desai, A., *Teaching quantitative and analytical methods*, Association for public Policy Analysis and Management,2006

[6] Eriksson, A., Preve, D., Yu, J., *Forecasting Realized Volatility using a Nonnegative Semiparametric Model*, Uppsala University, 2007

[7] Figlewski S., *Forecasting volatility*, BlackWell Publishers, 1997

[8] Gamboa, J., *Development of Tax Forecasting Models: Corporate and Individual Income Taxes*, Philippine Institute for Development Studies,2002

[9] Goldstone, J., *Using Quantitative and Qualitative Models to Forecast instability*, United States Institute of Peace, 2008
International Journal, vol. 31, issue 1, 2000.

[10] Karanta, I., *Expert system in forecast model building*,VTT Information Technology,

[11] Komaki, T., Penzer J., *Estimation of time-varying price elasticity in 1970-1997 Japanese raw milk supply by structural time-series model*, London School of Economics, 2002.

[12] Lendasse, A., Verleysen, M., de Bodt E., Cottrell, M., Grégoire P., *Forecasting Time-Series by Kohonen Classification*, European Symposium on Artificial Neural Networks,1998

[13] Lim, J. Y., *An Evaluating of Alternative Forecasting Models for Natural Rubber Prices*, school of Economics and Finance, Curtin University of Technology, 2002.

[14] Mariano R. S., Preve, D., *Statistical Tests for Multiple Forecast Comparison*, Uppsala University, 2007.

[15] Nicoli, S., *Fragmented foundation: education and chronic crisis in the Occupied Palestinian Territory*, International Institute for Educational Planning and Save Children UK, 2006.

[16] Preve, D., *Point Estimation in a Nonnegative First-Order Autoregression*, Uppsala University, 2007.

[17] Preve, D., *Robust Point Estimation in a Nonnegative Autoregression*, Uppsala University, 2007.

[18] Quoreshi, A. S., *Time series modeling of high frequency stock transaction data*, Umeå Economic Studies, 2006.

[19] Reade, J., *Modeling and forecasting football attendances*, The Oxonomics Society Published by Blackwell Publishing Ltd, 2007.

[20] Sarang D. N., *Exploratory Investigation of Sales Forecasting Process and Sales Forecasting System Case Study of Three Companies*, Laxmidhar M, Jönköping University, 2006.

[21] Scargle, J., *Bayesian Estimation of Time Series Lags and Structure*, NASA Ames Research Center, 2001.

[22] Singh, Sammer, *Pattern modeling in time-series forecasting*,

[23] WANG D. H., YAO Q., *Estimating GARCH models: when to use what?*, London School of Economics, 2007

Arabic Articles

[1] الجابري، بيومي، المحيسن، *استشراف مستقبل التعليم بمنطقة المدينة المنورة: تطبيق السلاسل الزمنية، 2005.*

[2] الجراح، ريم، علي، *توفيق احد نماذج BOX-JENKINS للسلاسل الزمنية واستخدامه للتنبؤ بحجم الاصابة المتوقعة لسرطان الثدي وسرطان الرئة في محافظة نينوي للفترة من 1976 إلى 2000*، تنمية الرافدين 67(24)2002، ص ص [265-276].

[3] العاني، احمد، حسين، استخدام نماذج *ARIMA* في التنبؤ الاقتصادي، 2004.

[4] محمد، منعم، عزيز، *دراسة تحليلية لتقدير نموذج الانحدار لبيانات تعاني من مشكلتي الارتباط* و التعدد الخطي معا، تنمية الرافدين 67(24)2002، ص ص [277-289].

Arabic Electronic Books

[1] بري، عدنان، *طرق التنبؤ الإحصائي*، 2002، متاح على الصفحة: http://www.abarry.net/or/or221book1.pdf

Reports:

[1] PCBS, PHC-2007, 2009.

[2] UNHCR,(2007a), The United Nations and Palestinian Refugees, Available [online]: http://www.un.org/unrwa/publications/pubs07/UN&PR_en.pdf[09.6.2008]

[3] United nation conference on trade and development, population and demographic development in the west bank and Gaza Strip until 1990, report (UNCTAD/ECDC/SEU/1), 1991.

[4] UNRWA,(2007a), unrwa in 2006, Available [online]: http://www.un.org/unrwa/publications/pubs07/unrwa_in_2006.pdf[09.6.2008].

[5] UNRWA, (2007b), General Fund Appeal 2008- 2009, Available [online]: http://www.un.org/unrwa/publications/pubs07/GF20082009.pdf[09.6.2008].

[6] UNRWA, (2007c), Annual Growth Rate of Registered Palestine Refugees and Female Percentage, Available [online]:http://www.un.org /unrwa/publications/pdf/figures.pdf[09.6.2008].

[7] UNRWA, (2007d), education, Available [online]: http://www.un.org /unrwa/publications/pdf/education.pdf [09.6.2008].

Internet Websites

[1] http://www.itl.nist.gov/div898/handbook

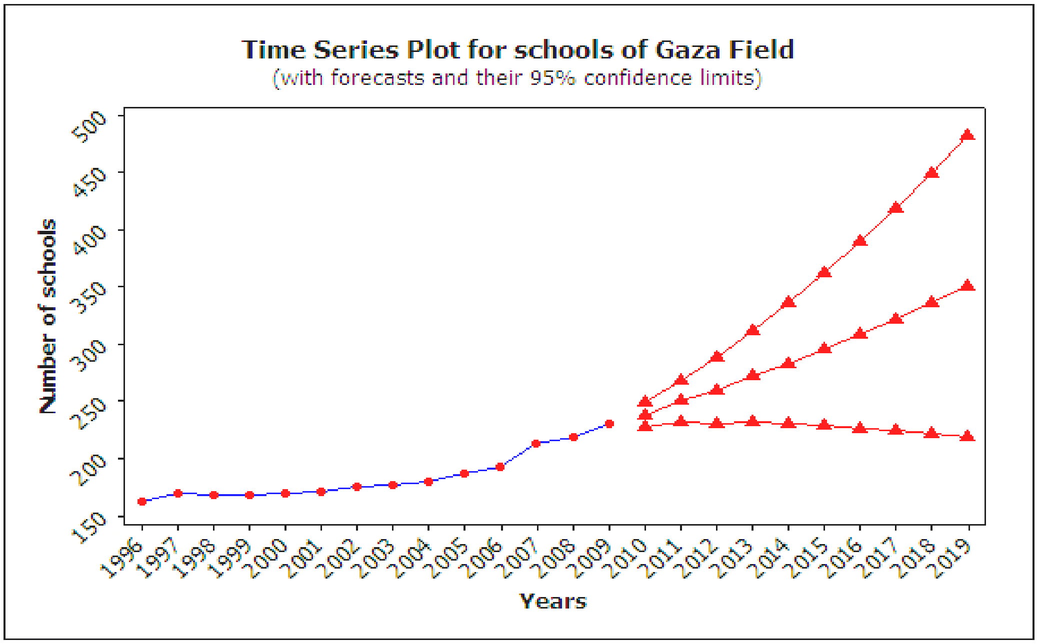
[2] http://www.pcbs.gov.ps

[3] http://www.un.org/unrwa/puplications

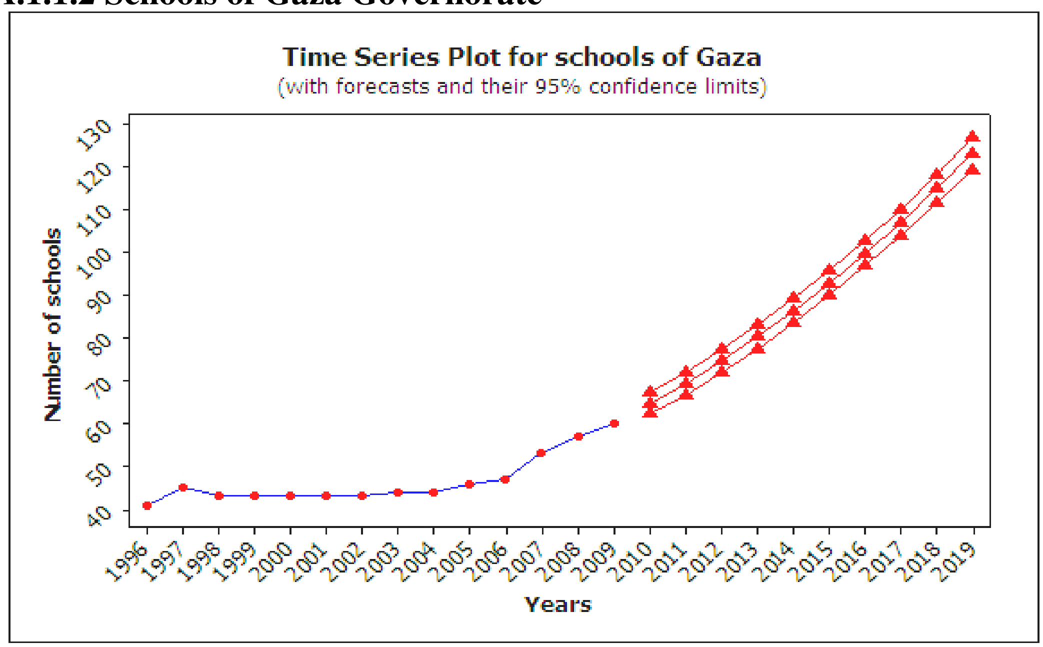
Appendix 1: Plots including 10 forecasts for the variables of study

A.1.1 Time series plot for number of schools

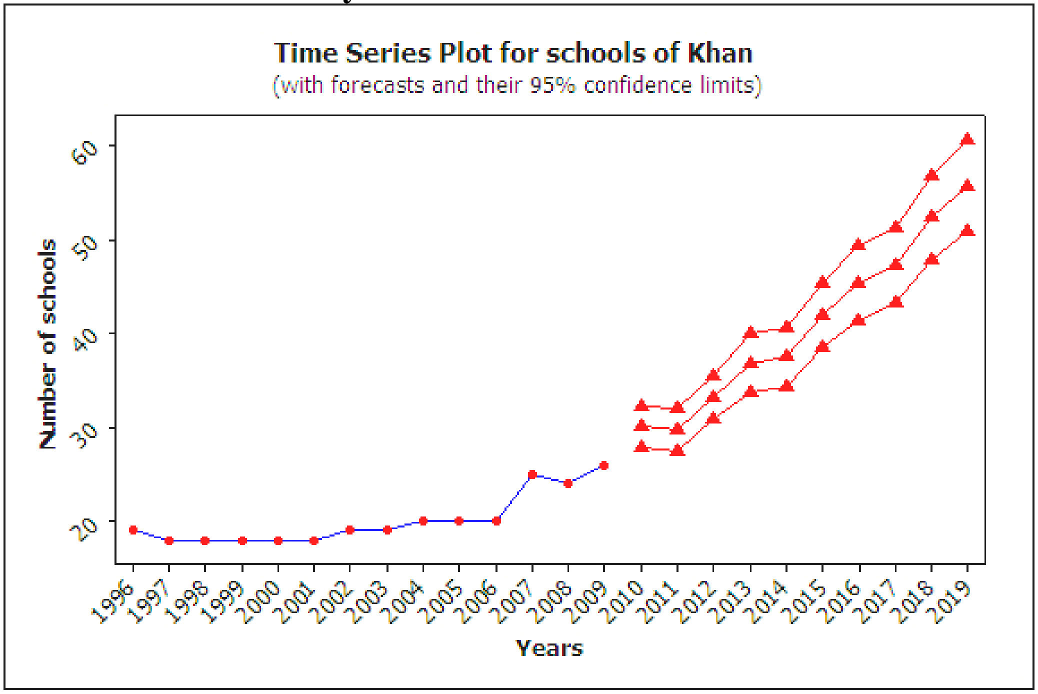
A.1.1.1 Schools of Gaza Field



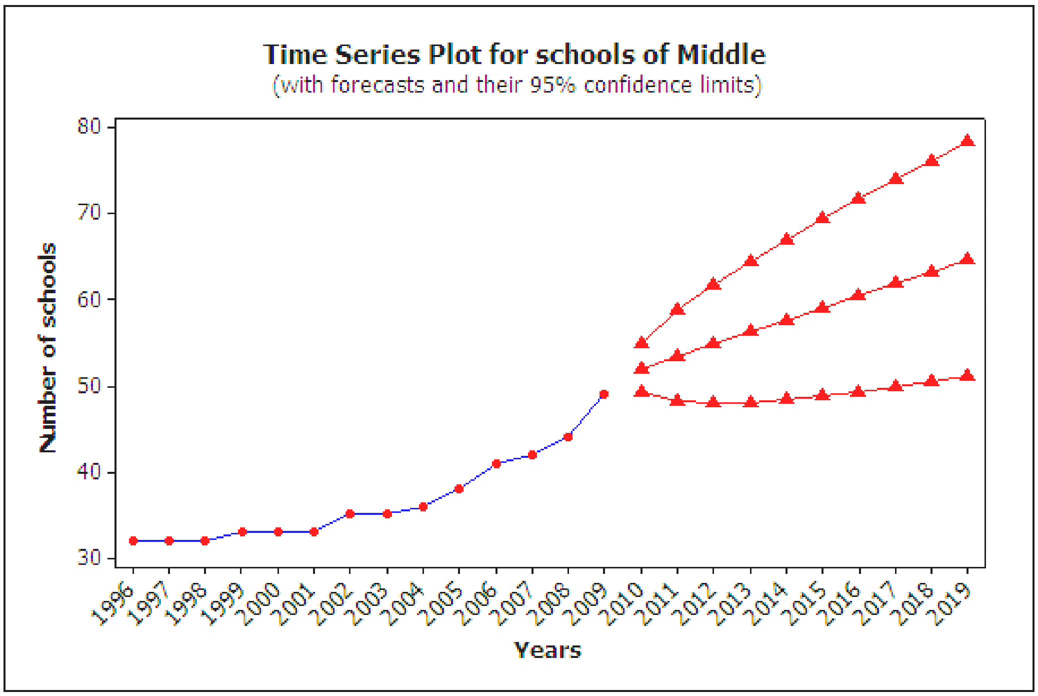
A.1.1.2 Schools of Gaza Governorate



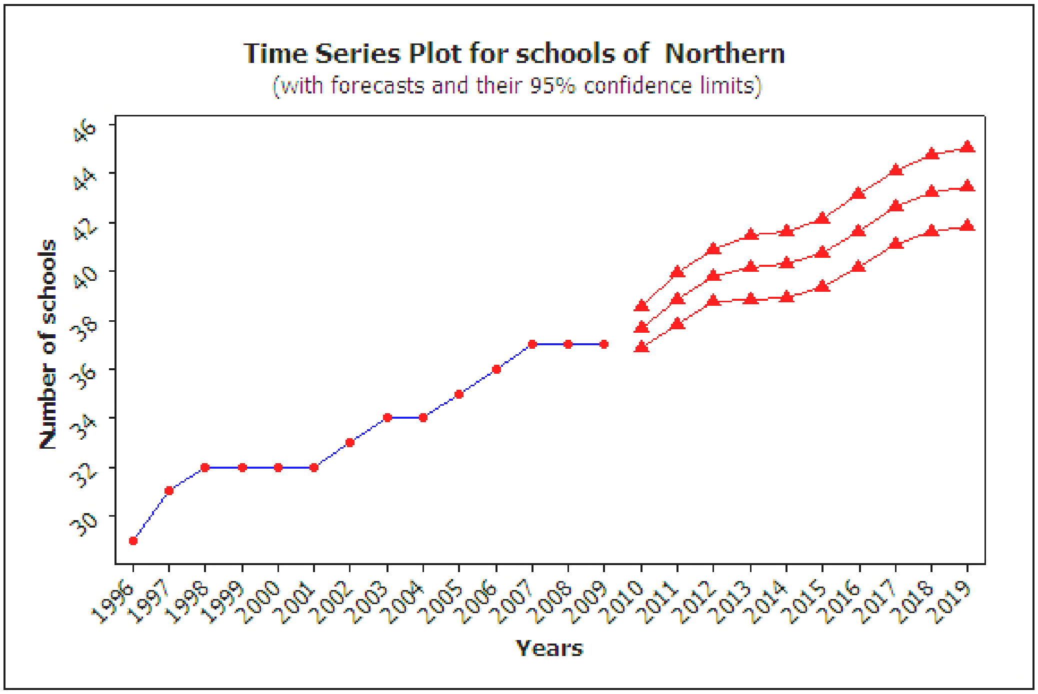
A.1.1.3 Schools of Khanyounis Western area



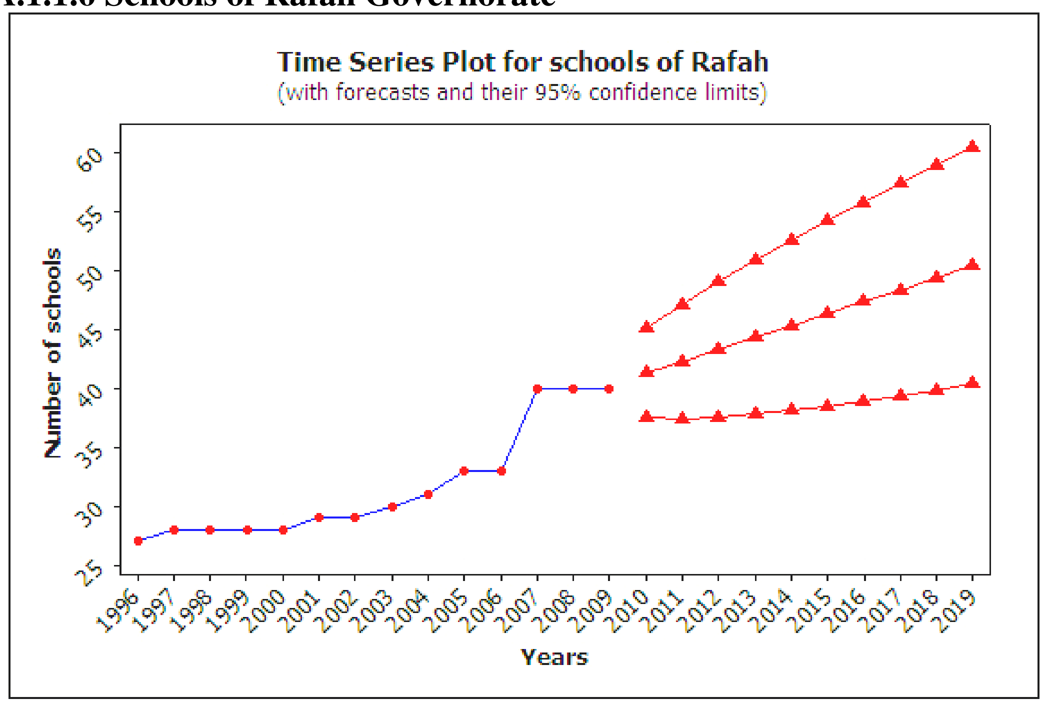
A.1.1.4 Schools of Middle Governorate



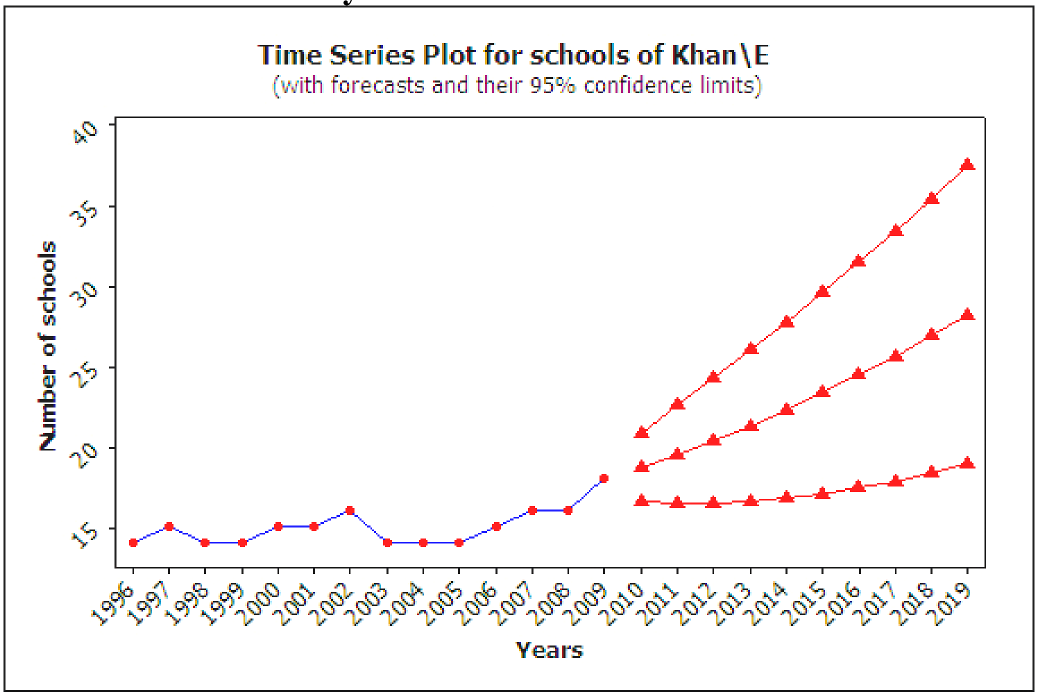
A.1.1.5 Schools of Northern Governorate



A.1.1.6 Schools of Rafah Governorate

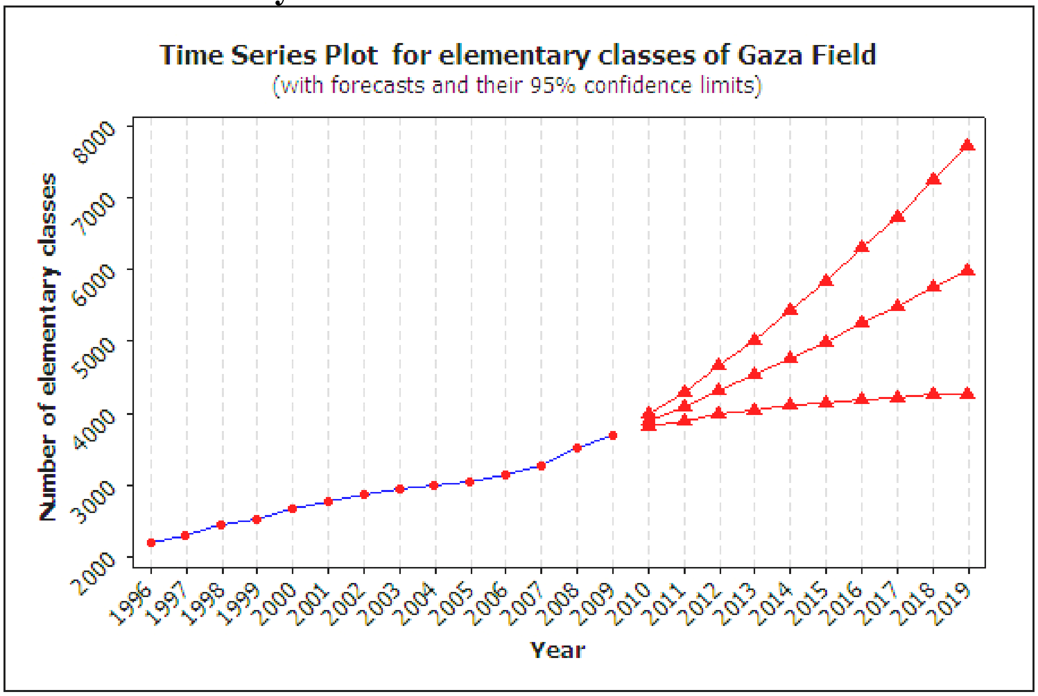


A.1.1.7 Schools of Khanyounis Eastern area

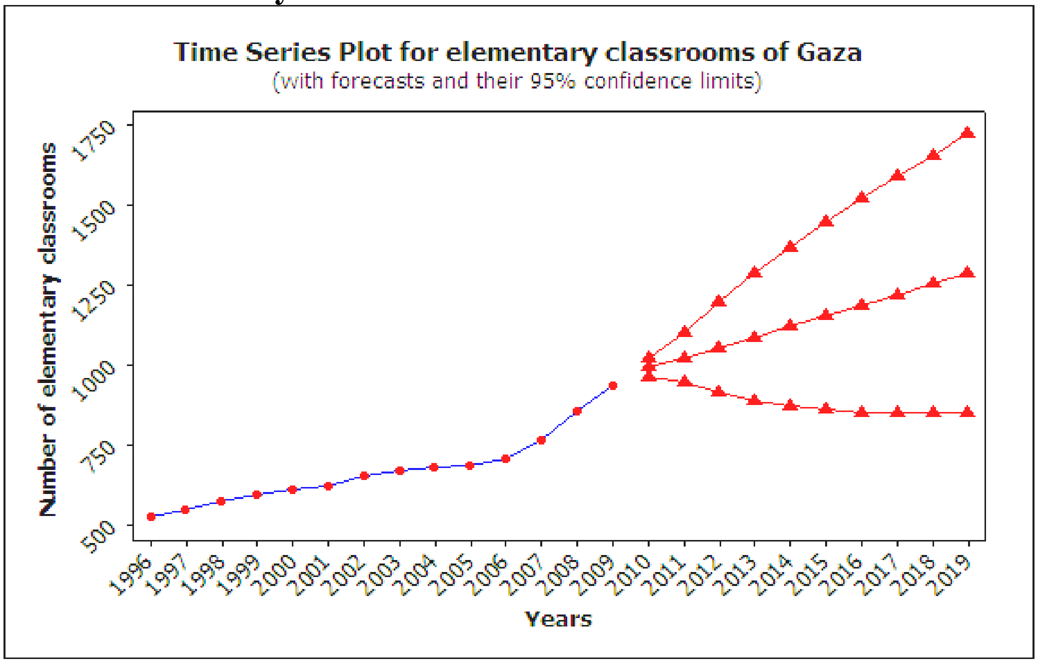


A.1.2 Time series plot for number of elementary classes

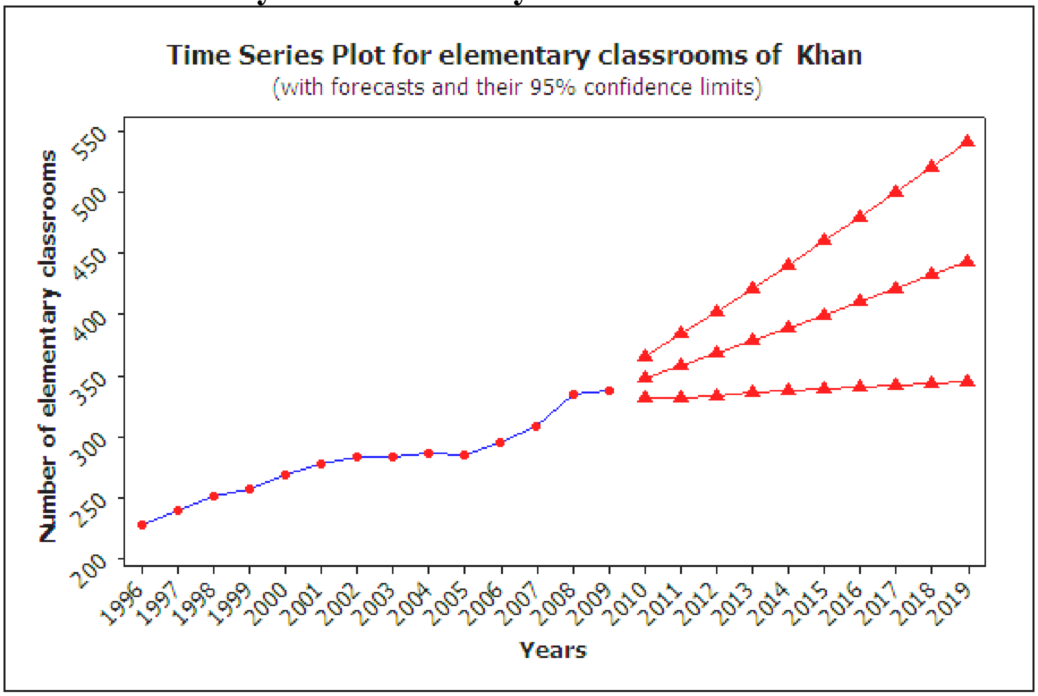
A.1.2.1 Elementary classes of Gaza Field



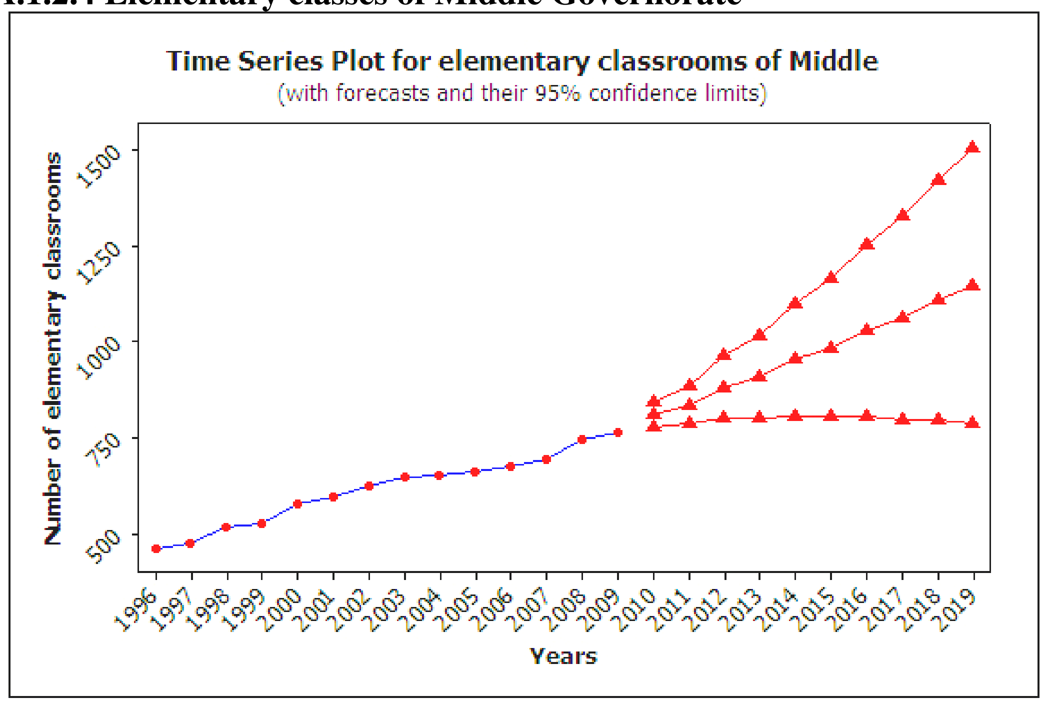
A.1.2.2 Elementary classes of Gaza Governorate



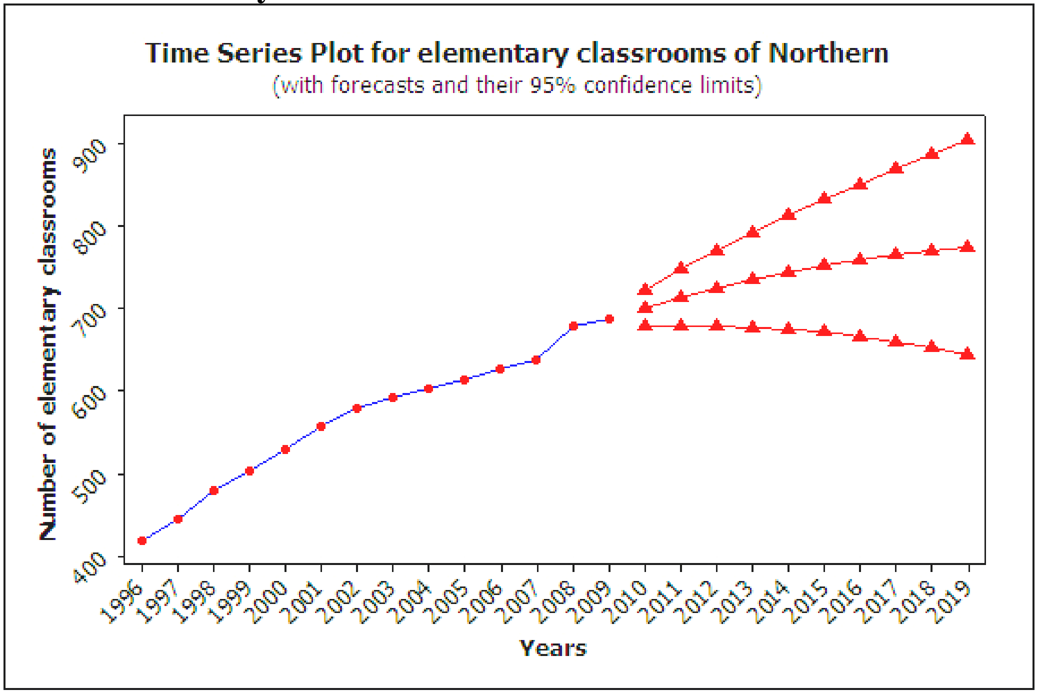
A.1.2.3 Elementary classes of Khanyounis Western area



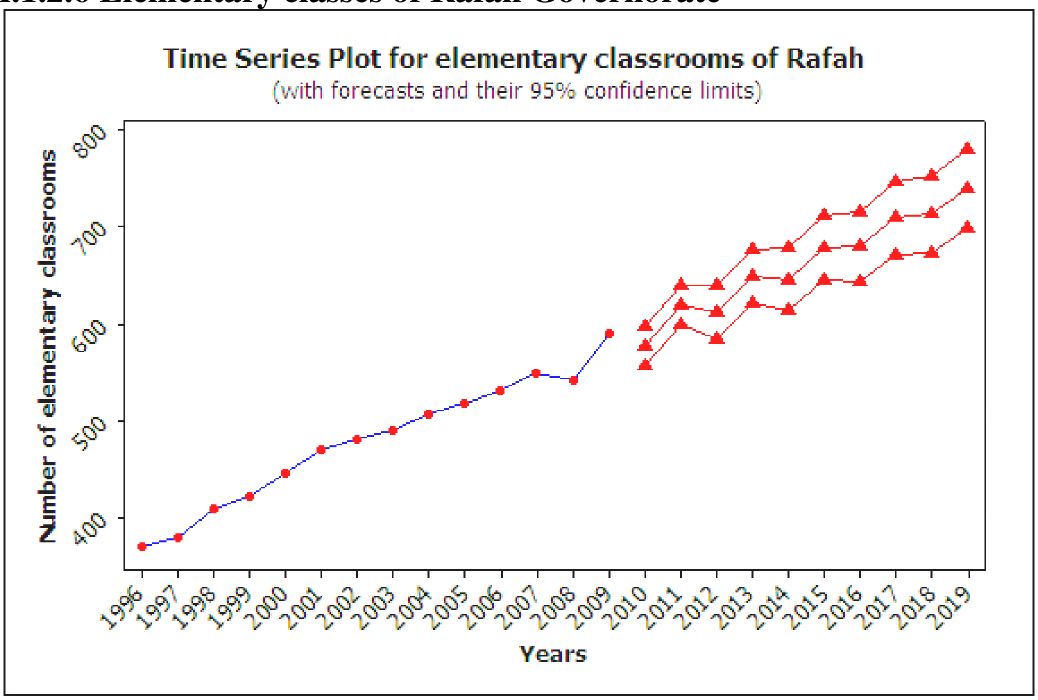
A.1.2.4 Elementary classes of Middle Governorate



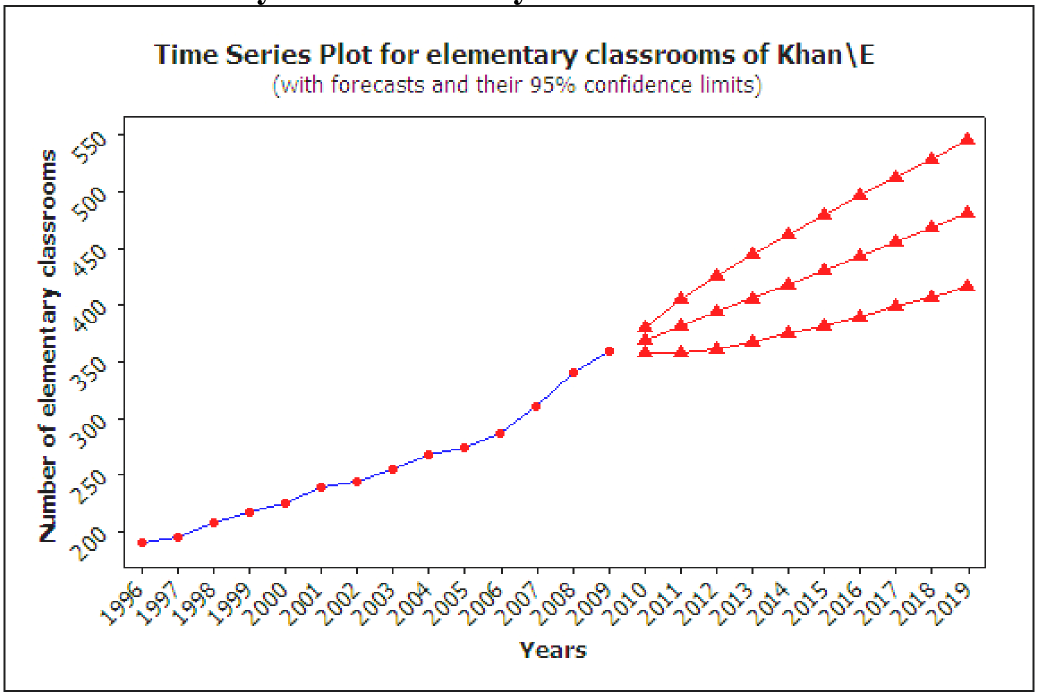
A.1.2.5 Elementary classes of Northern Governorate



A.1.2.6 Elementary classes of Rafah Governorate

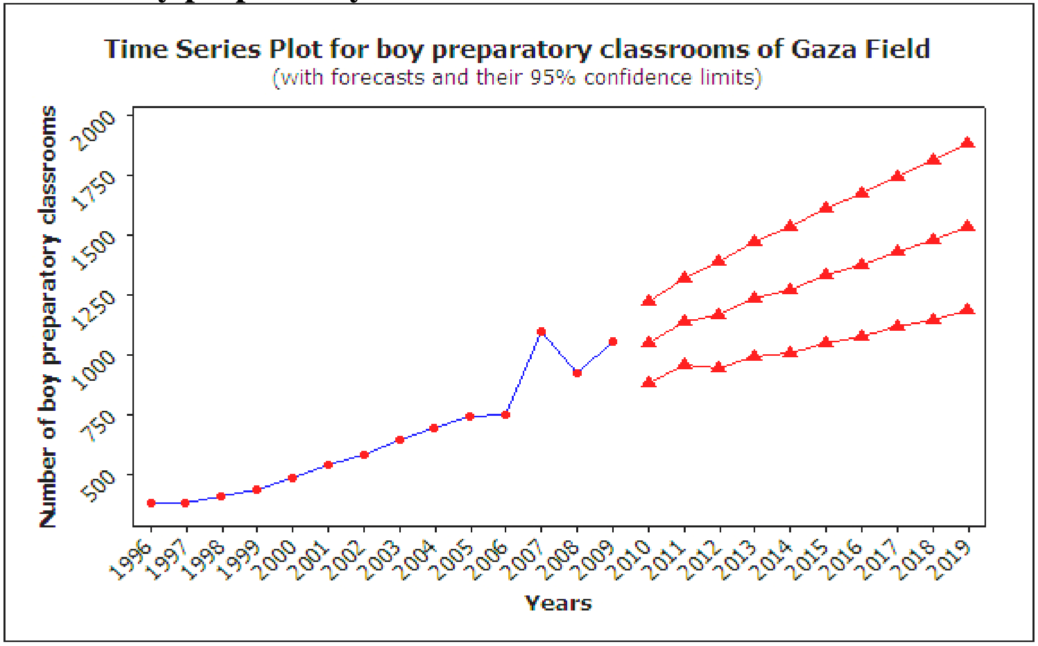


A.1.2.7 Elementary classes of Khanyounis Eastern area

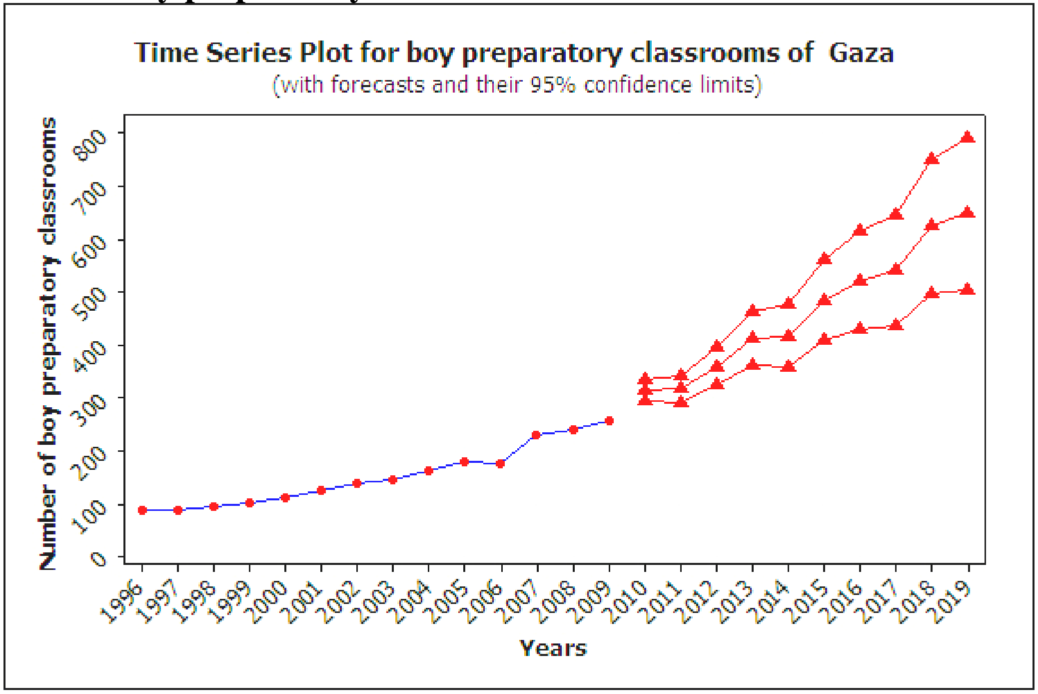


A.1.3 Time series plot for number of boy preparatory classrooms

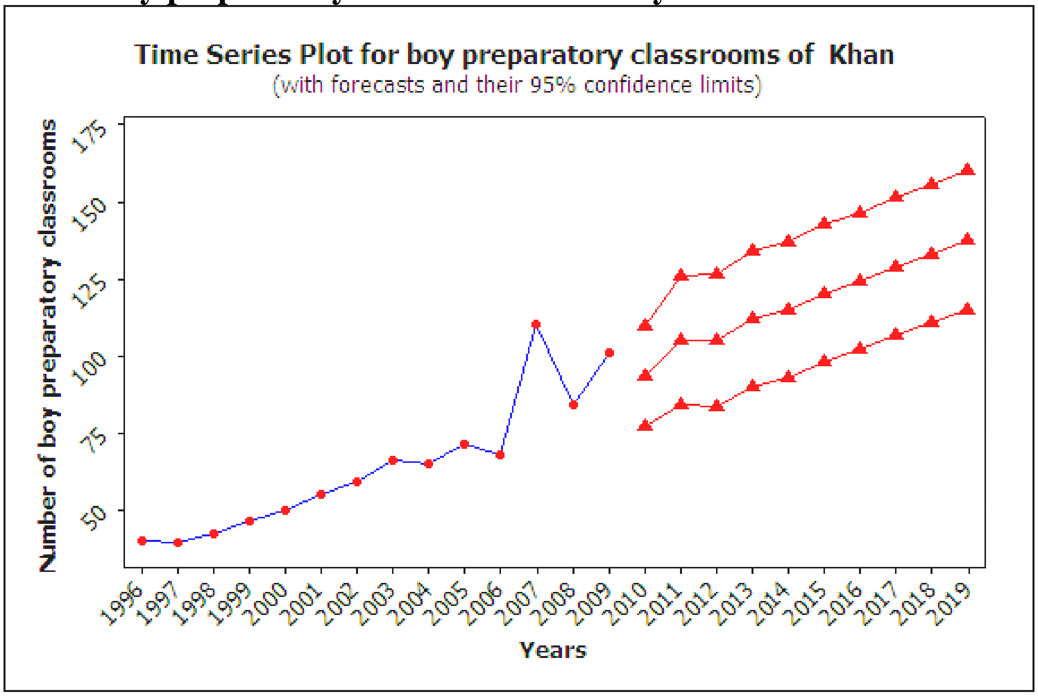
A.1.3.1 Boy preparatory classrooms of Gaza Field



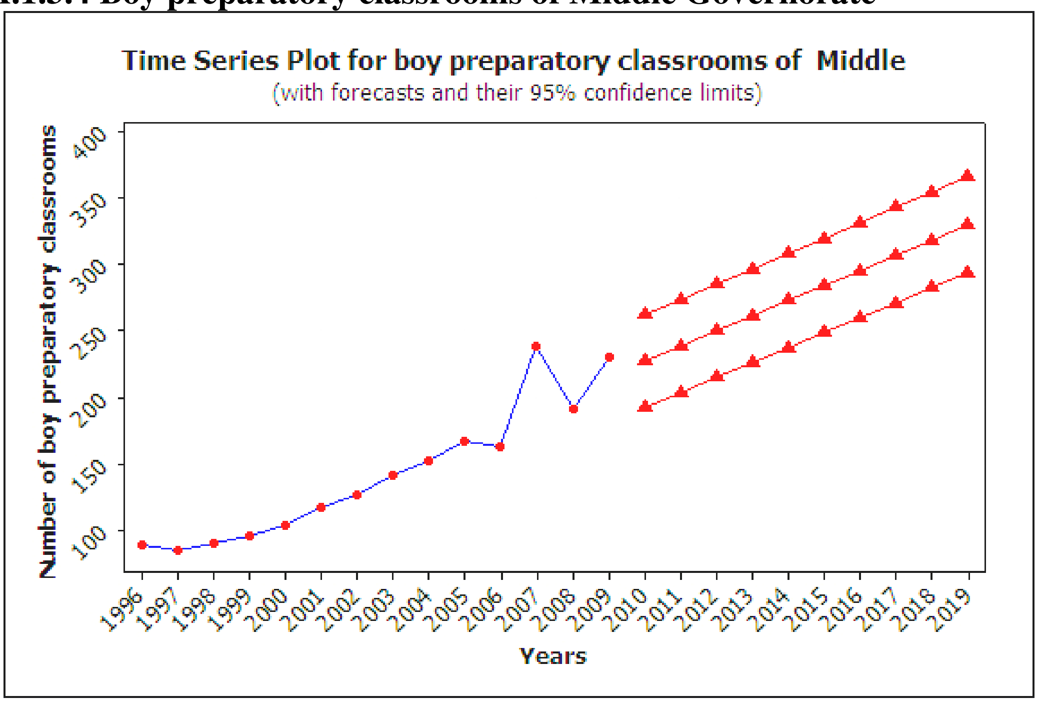
A.1.3.2 Boy preparatory classrooms of Gaza Governorate



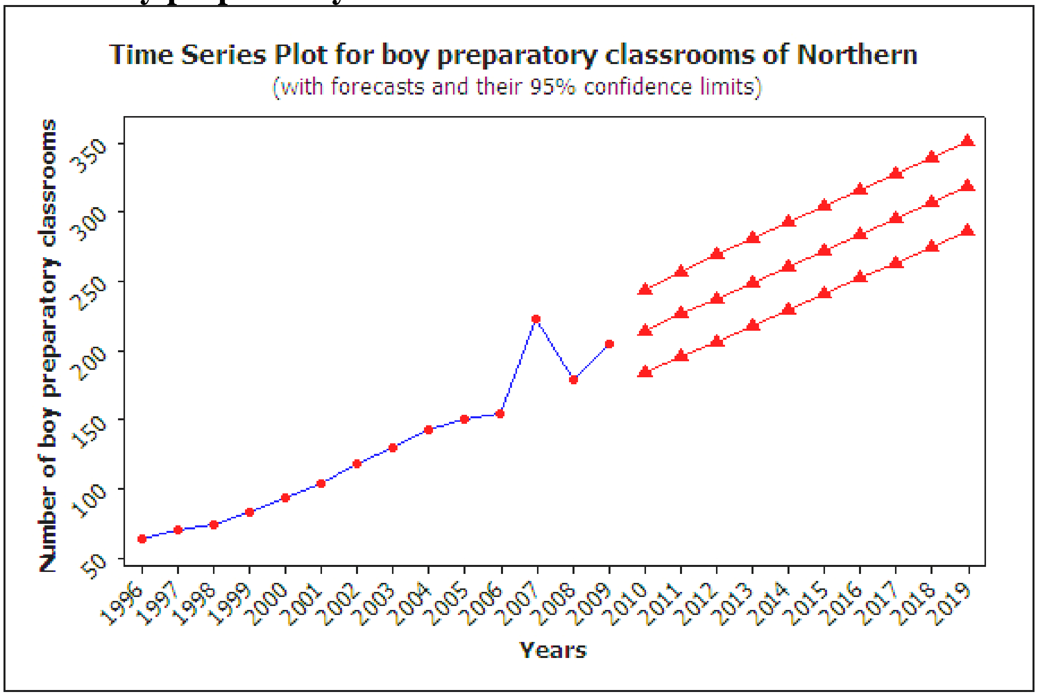
A.1.3.3 Boy preparatory classrooms of Khanyounis Western area



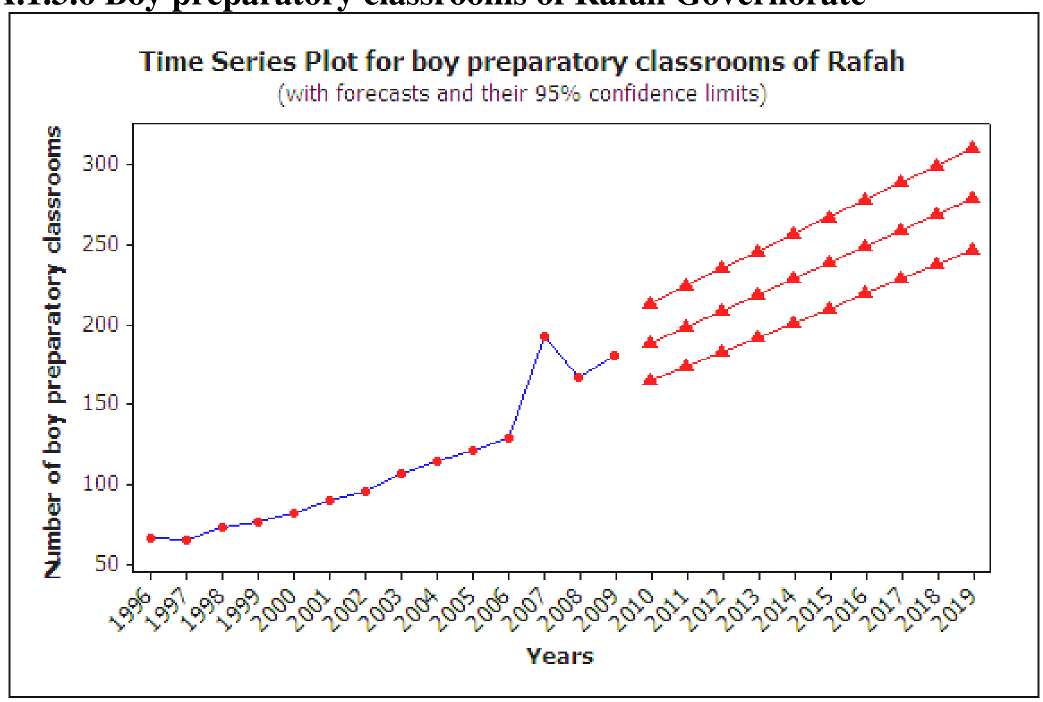
A.1.3.4 Boy preparatory classrooms of Middle Governorate



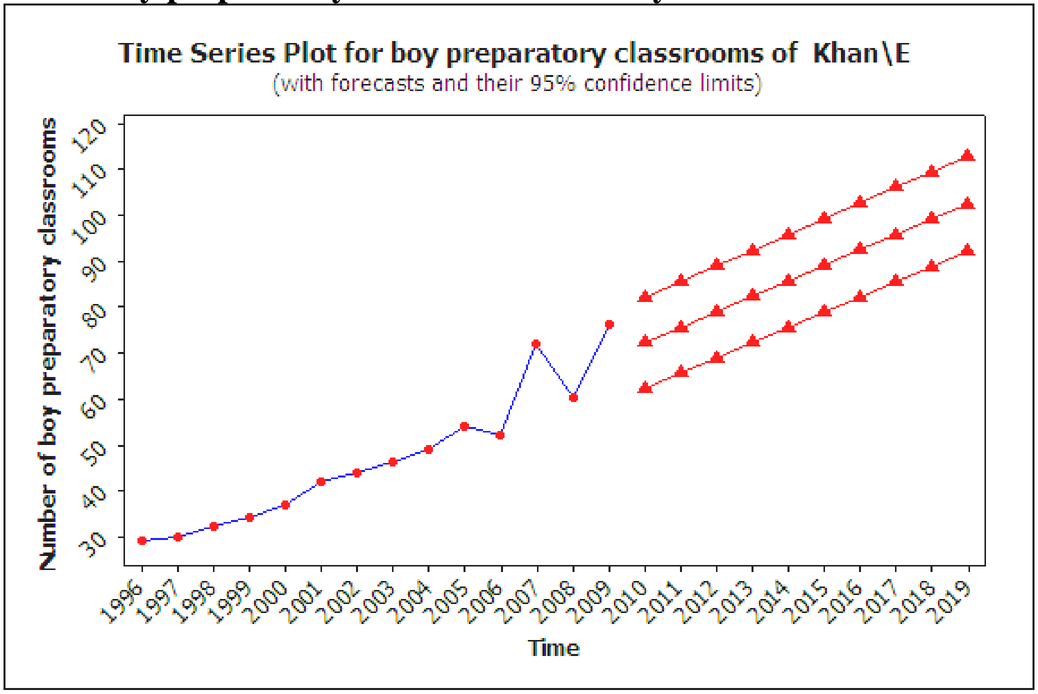
A.1.3.5 Boy preparatory classrooms of Northern Governorate



A.1.3.6 Boy preparatory classrooms of Rafah Governorate

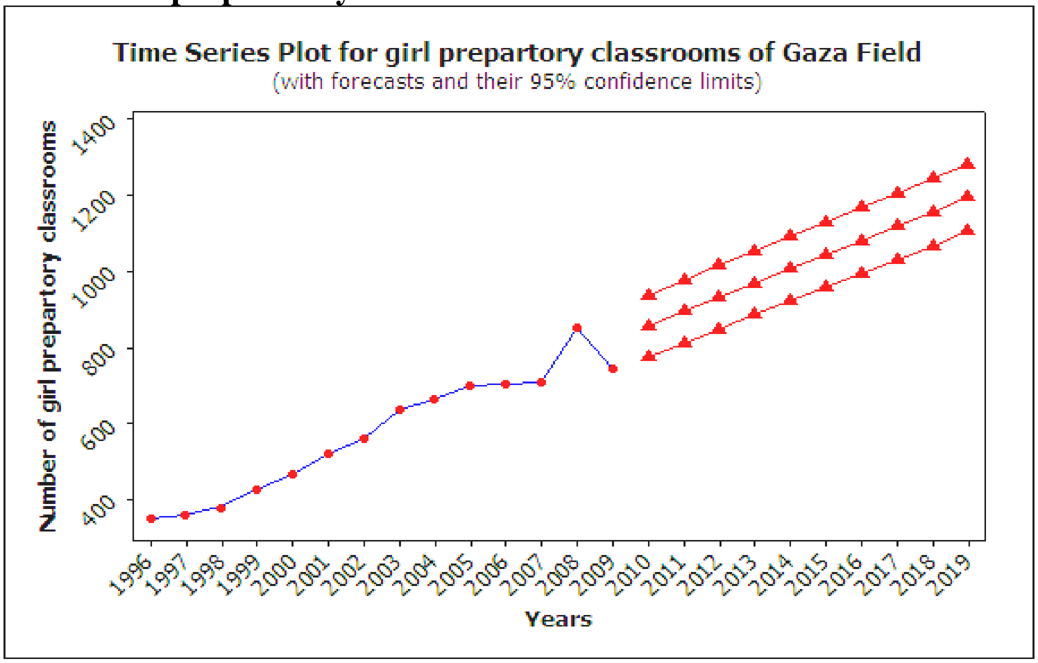


A.1.3.7 Boy preparatory classrooms of Khanyounis Eastern area

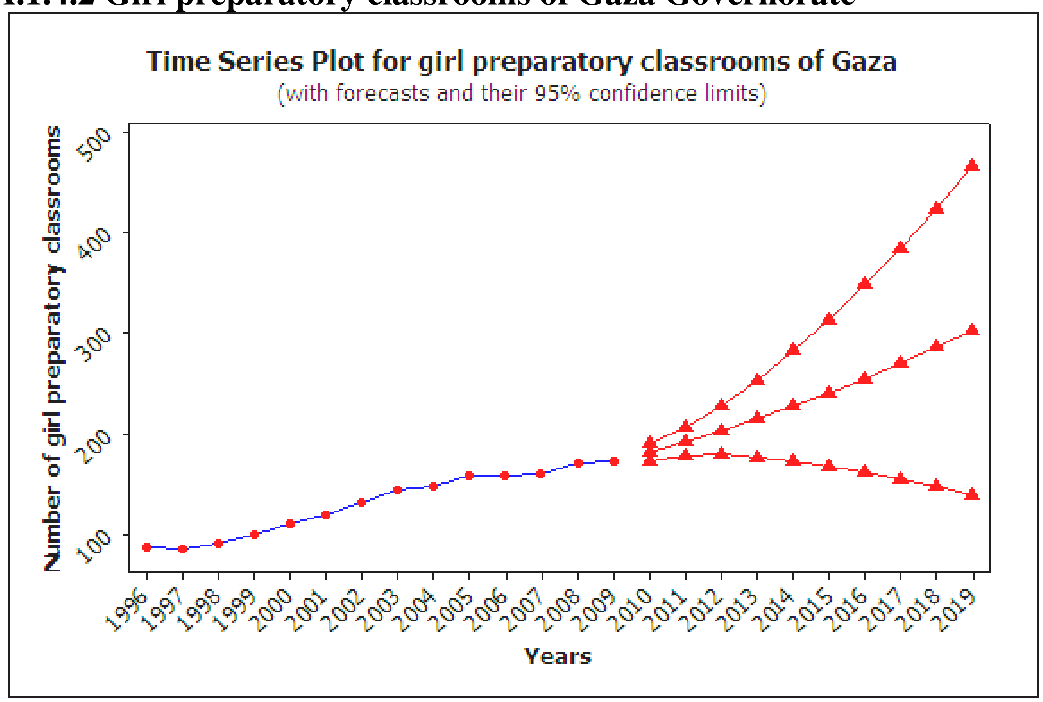


A.1.4 Time series plot for number of girl preparatory classrooms

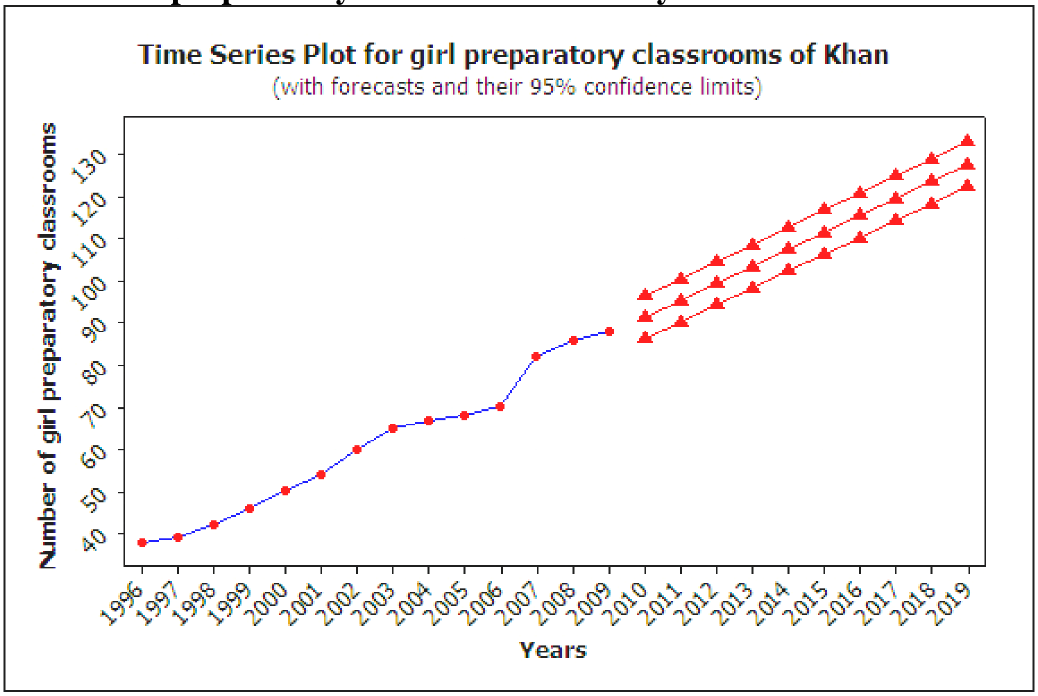
A.1.4.1 Girl preparatory classrooms of Gaza Field



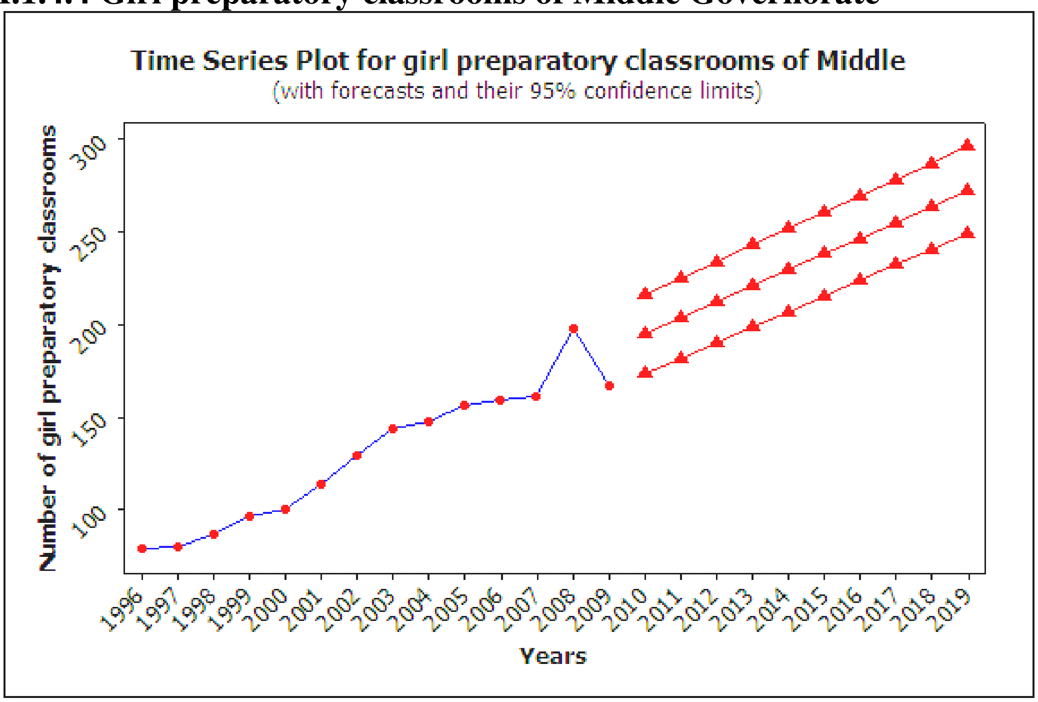
A.1.4.2 Girl preparatory classrooms of Gaza Governorate



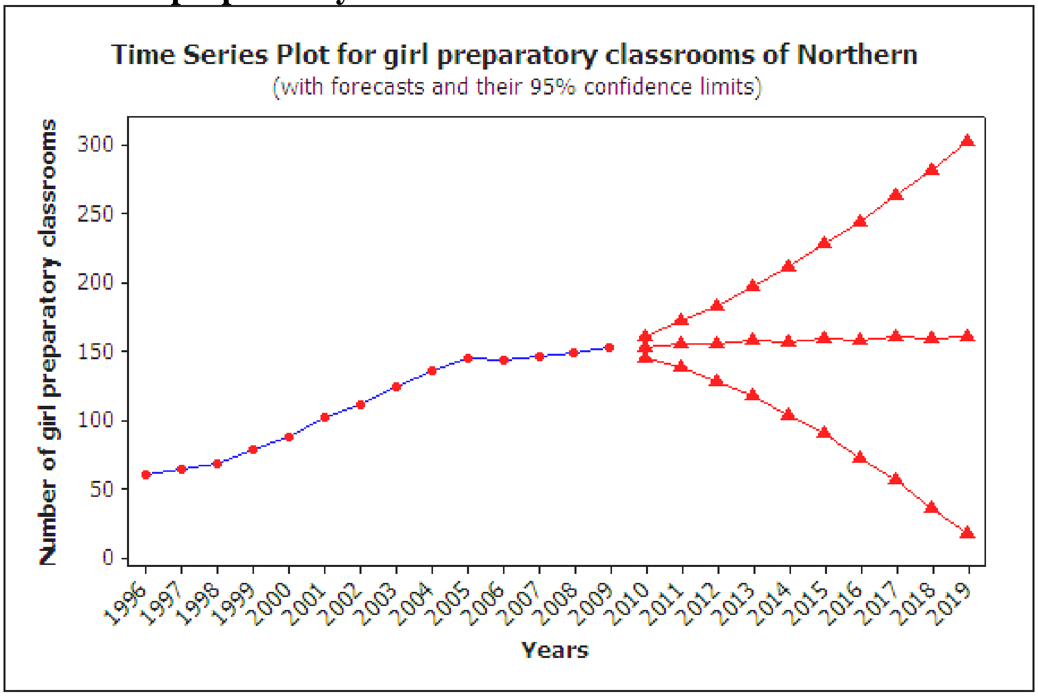
A.1.4.3 Girl preparatory classrooms of Khanyounis Western area



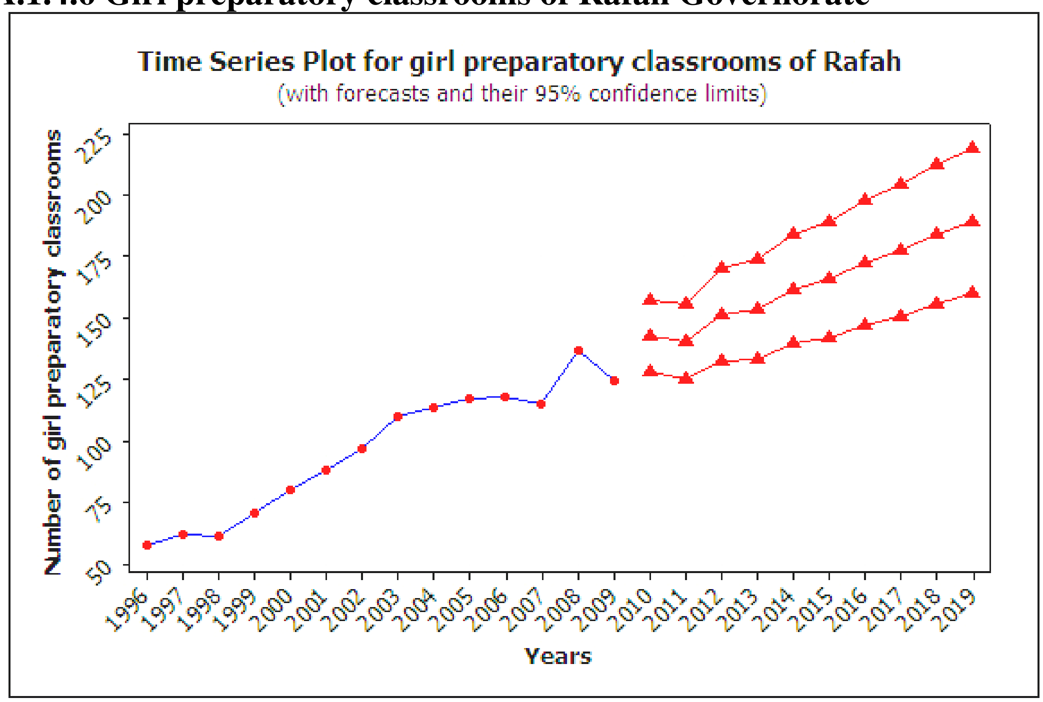
A.1.4.4 Girl preparatory classrooms of Middle Governorate



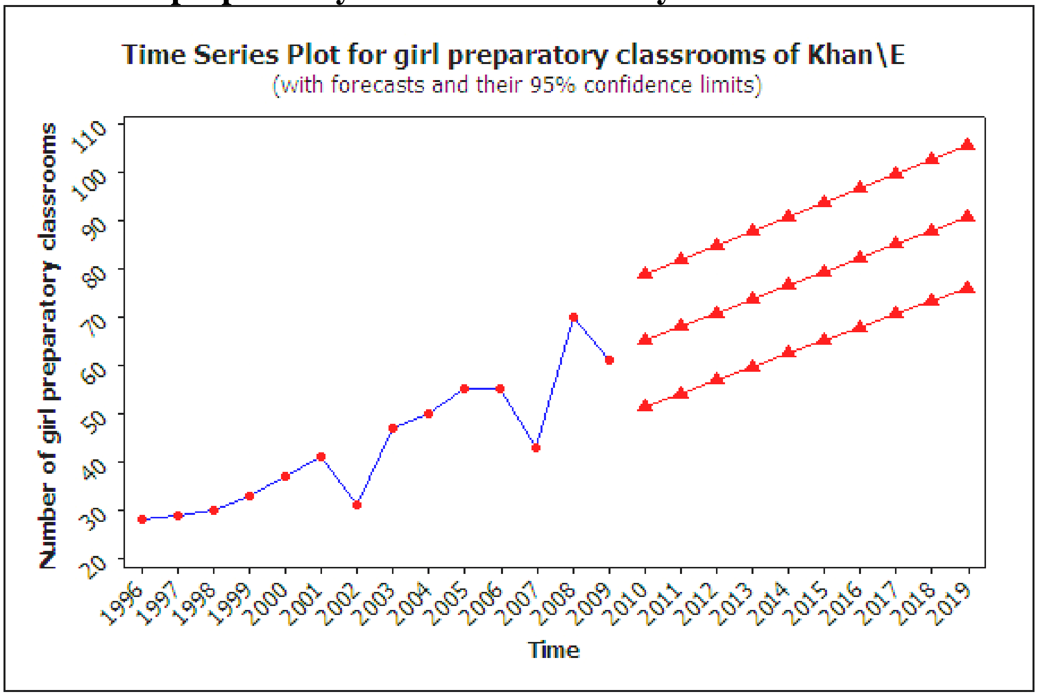
A.1.4.5 Girl preparatory classrooms of Northern Governorate



A.1.4.6 Girl preparatory classrooms of Rafah Governorate

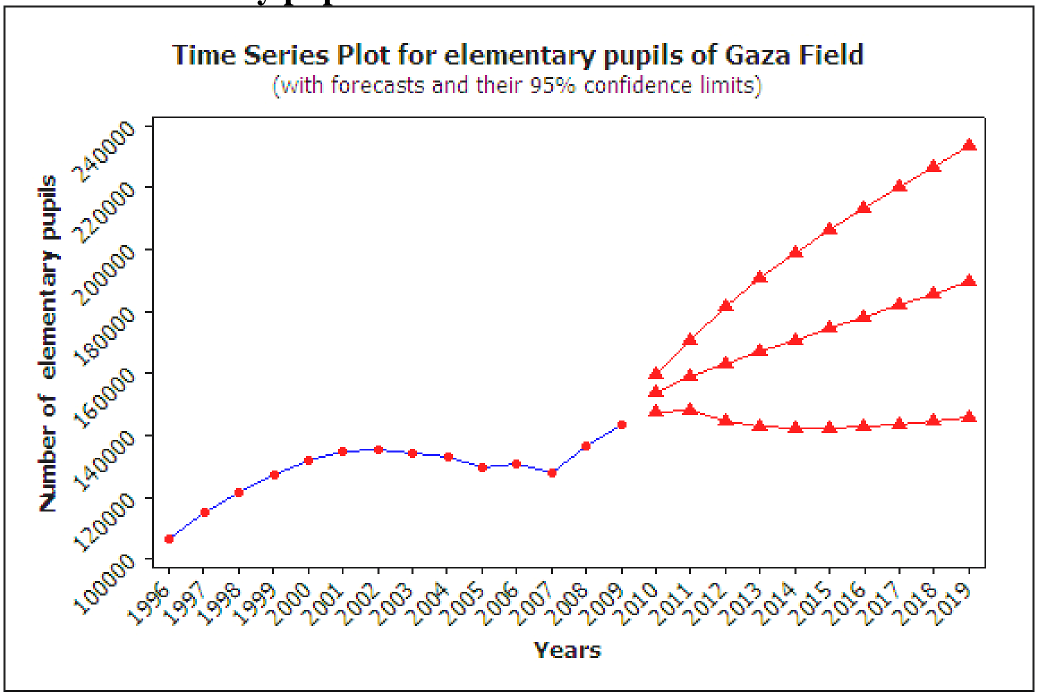


A.1.4.7 Girl preparatory classrooms of Khanyounis Eastern area

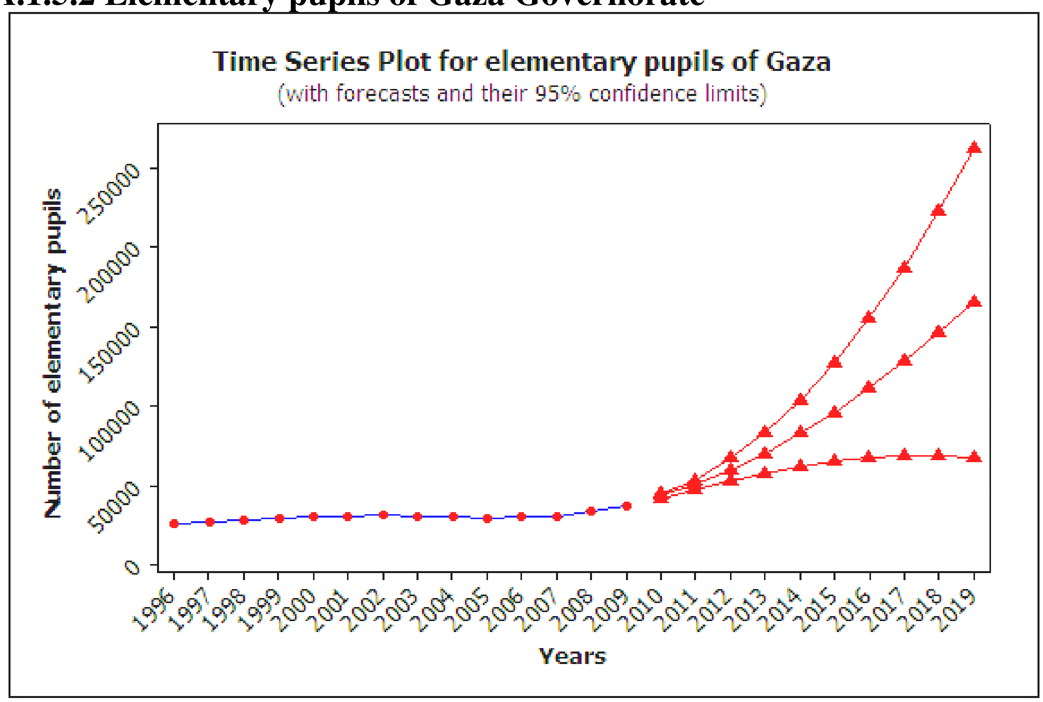


A.1.5 Time series plot for number of elementary pupils

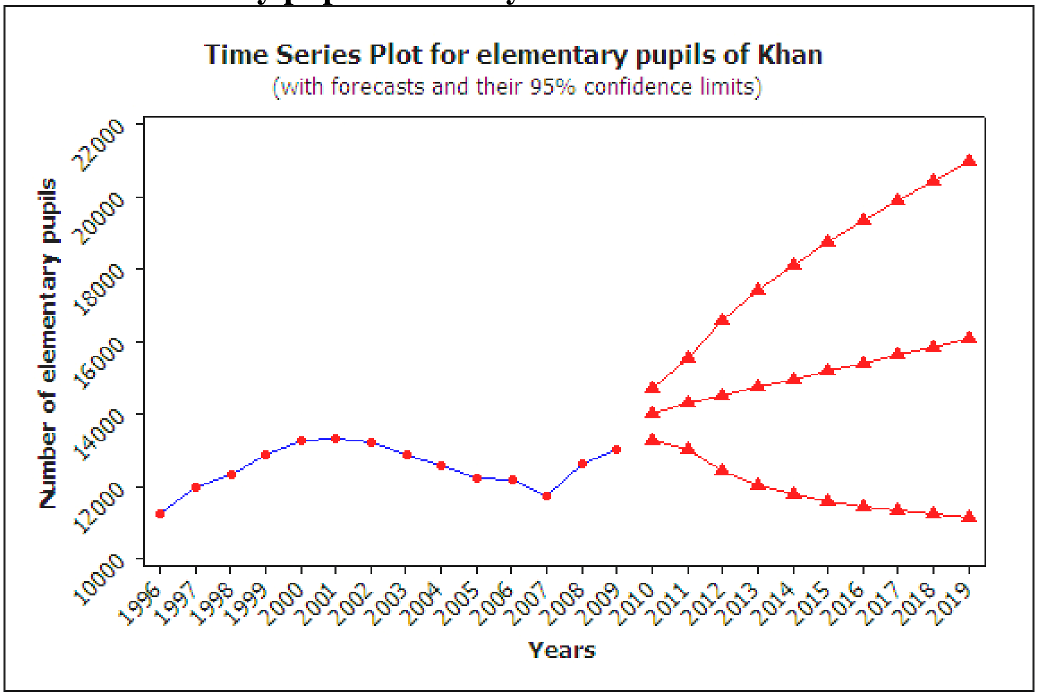
A.1.5.1 Elementary pupils of Gaza Field



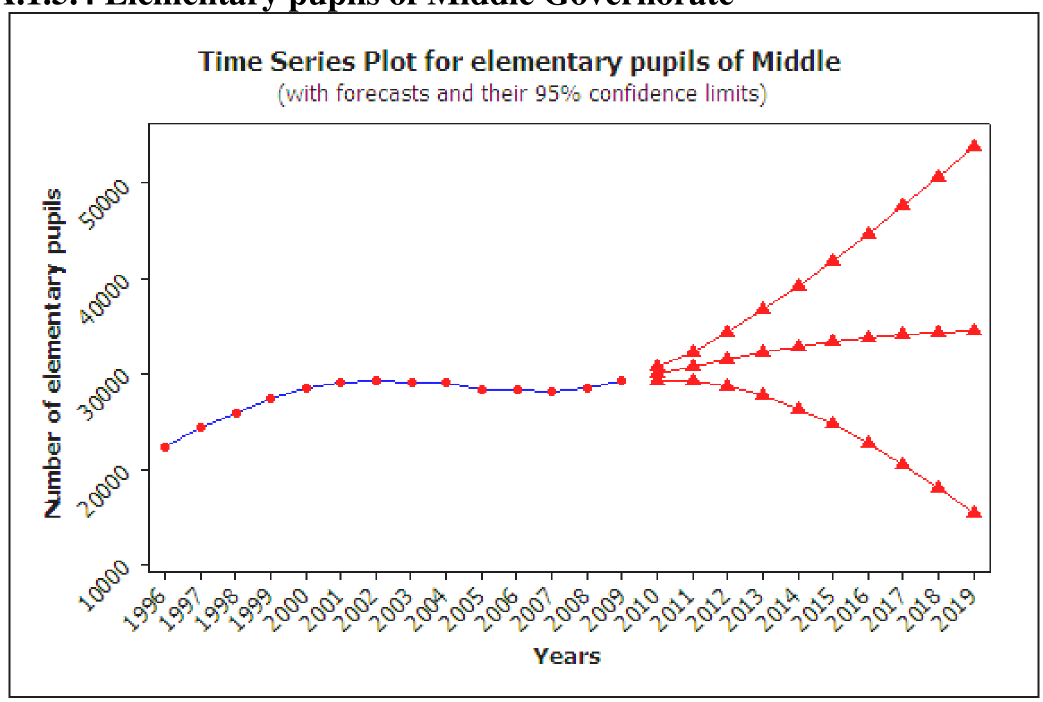
A.1.5.2 Elementary pupils of Gaza Governorate



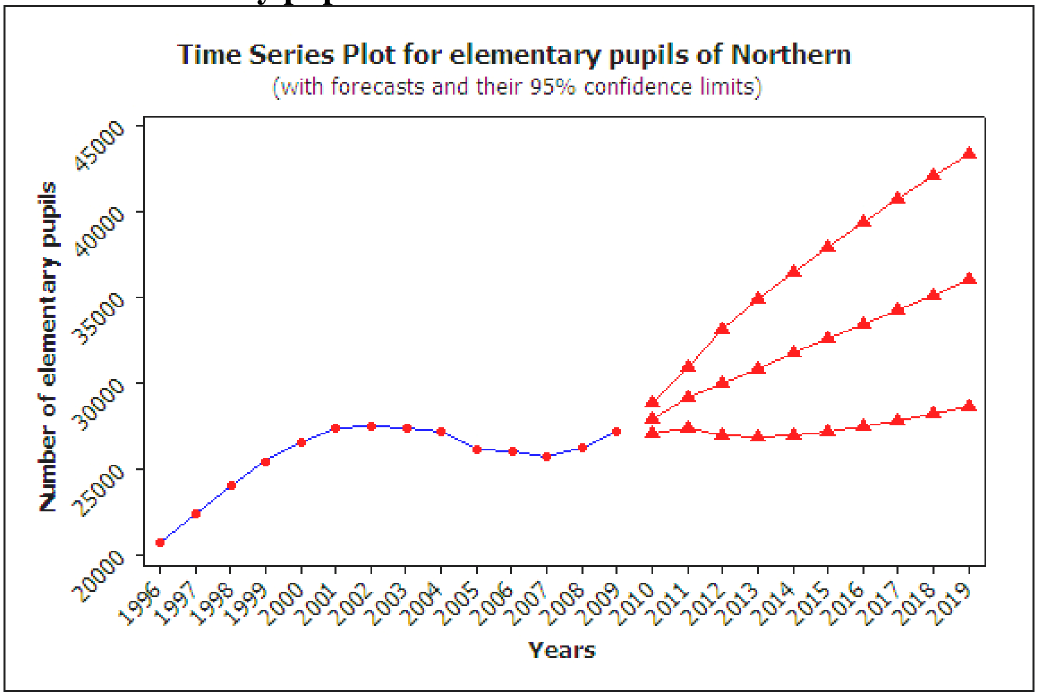
A.1.5.3 Elementary pupils of Khanyounis Western area



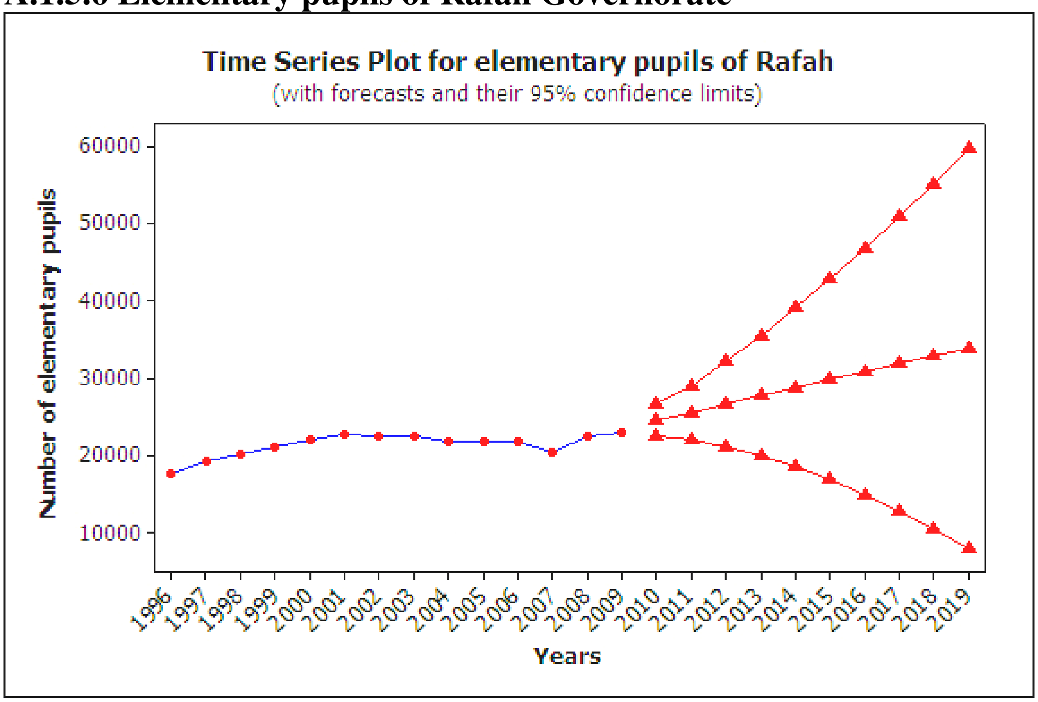
A.1.5.4 Elementary pupils of Middle Governorate



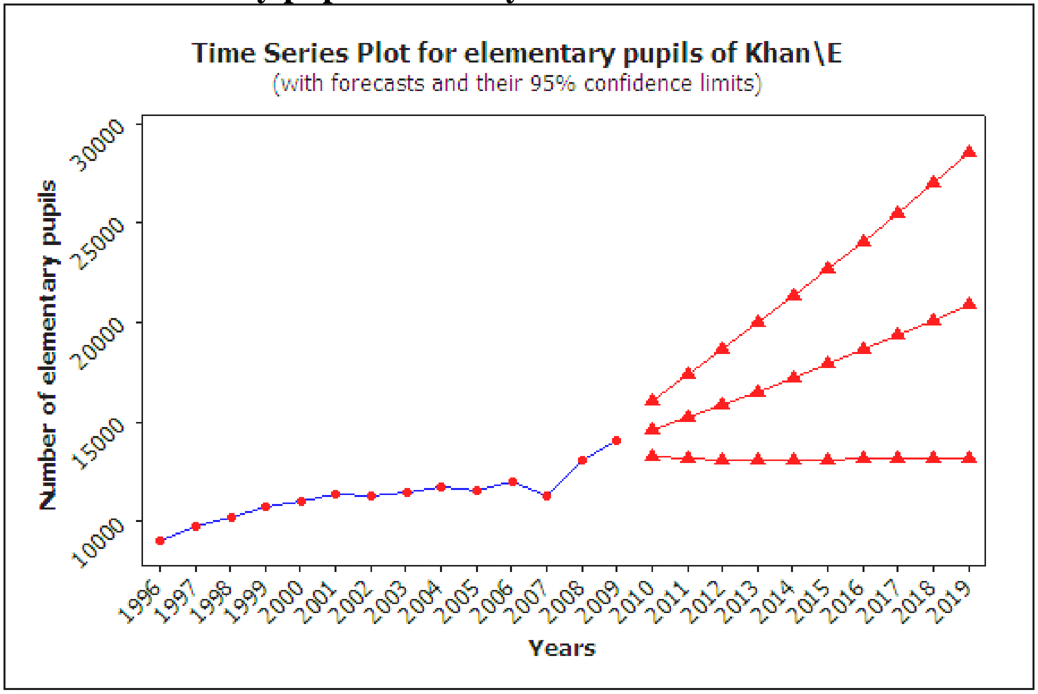
A.1.5.5 Elementary pupils of Northern Governorate



A.1.5.6 Elementary pupils of Rafah Governorate

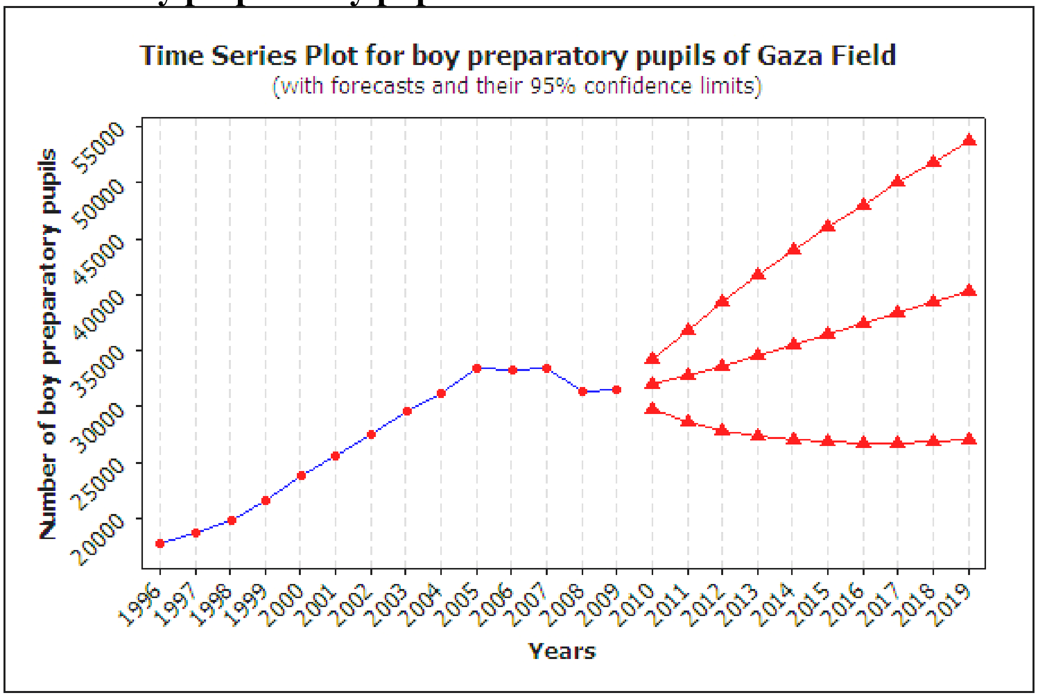


A.1.5.7 Elementary pupils of Khanyounis Eastern area

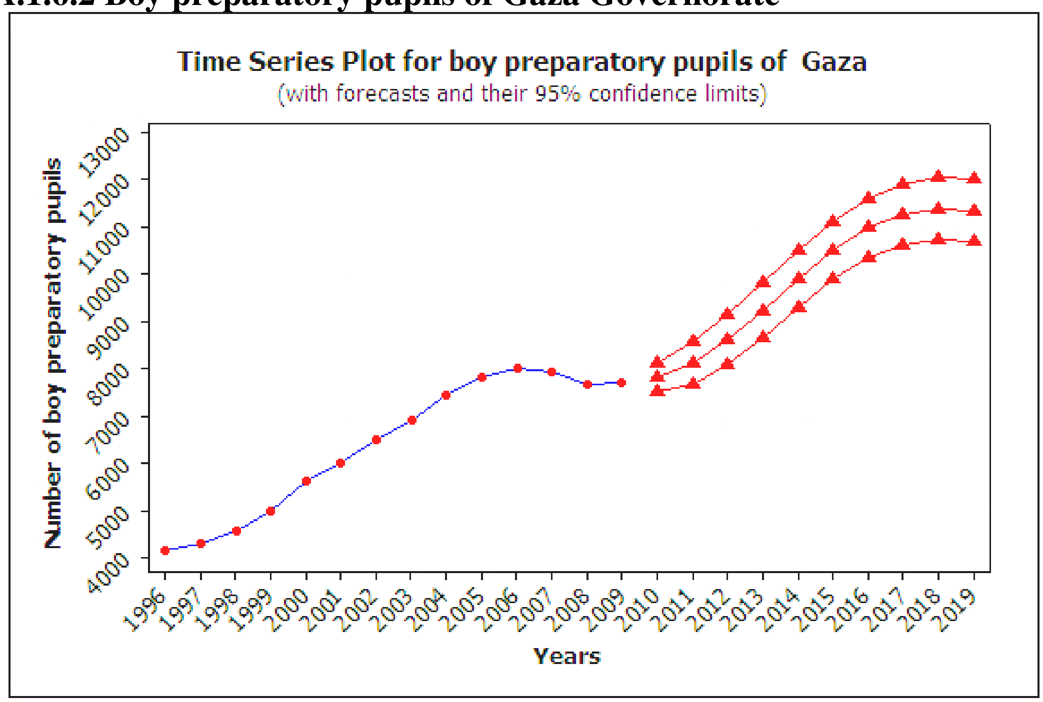


A.1.6 Time series plot for number of boy preparatory pupils

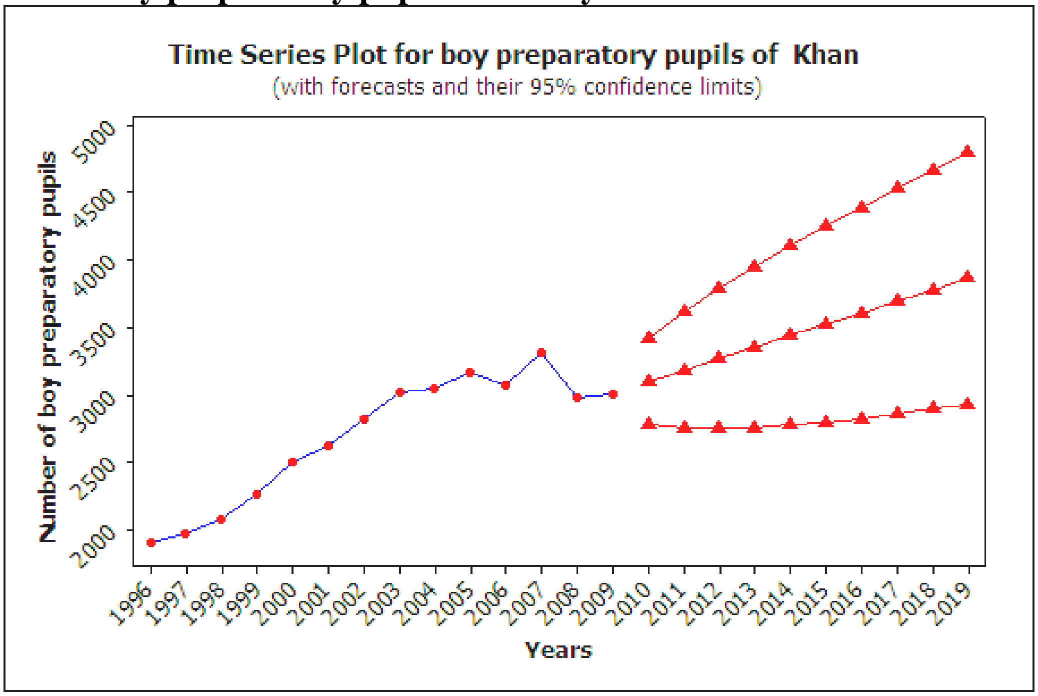
A.1.6.1 Boy preparatory pupils of Gaza Field



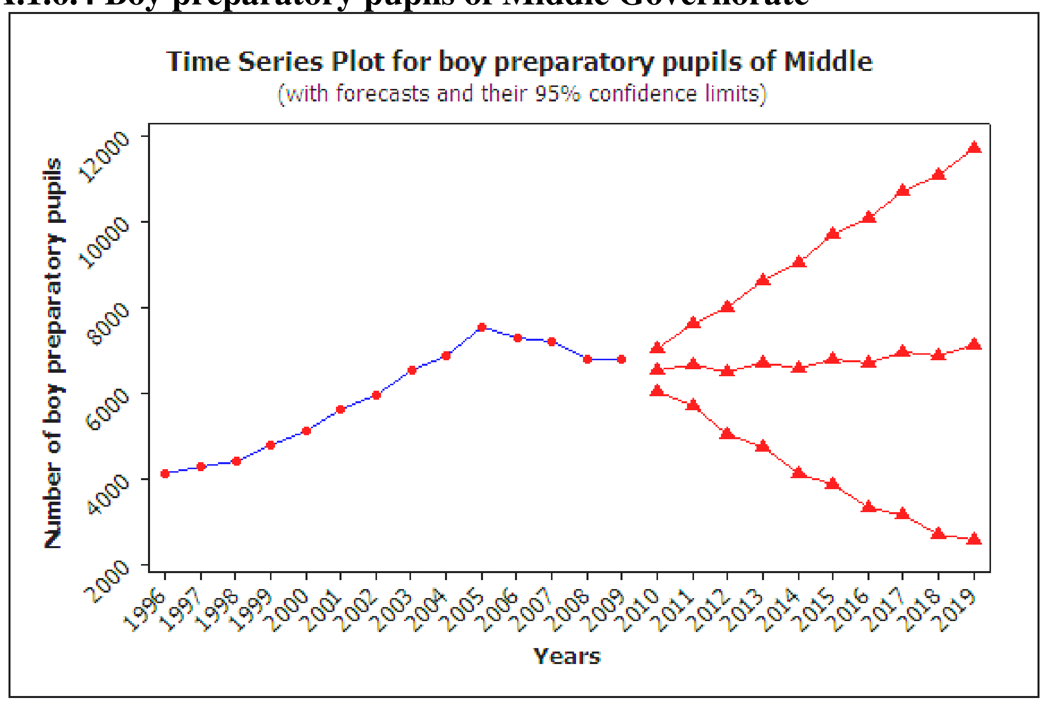
A.1.6.2 Boy preparatory pupils of Gaza Governorate



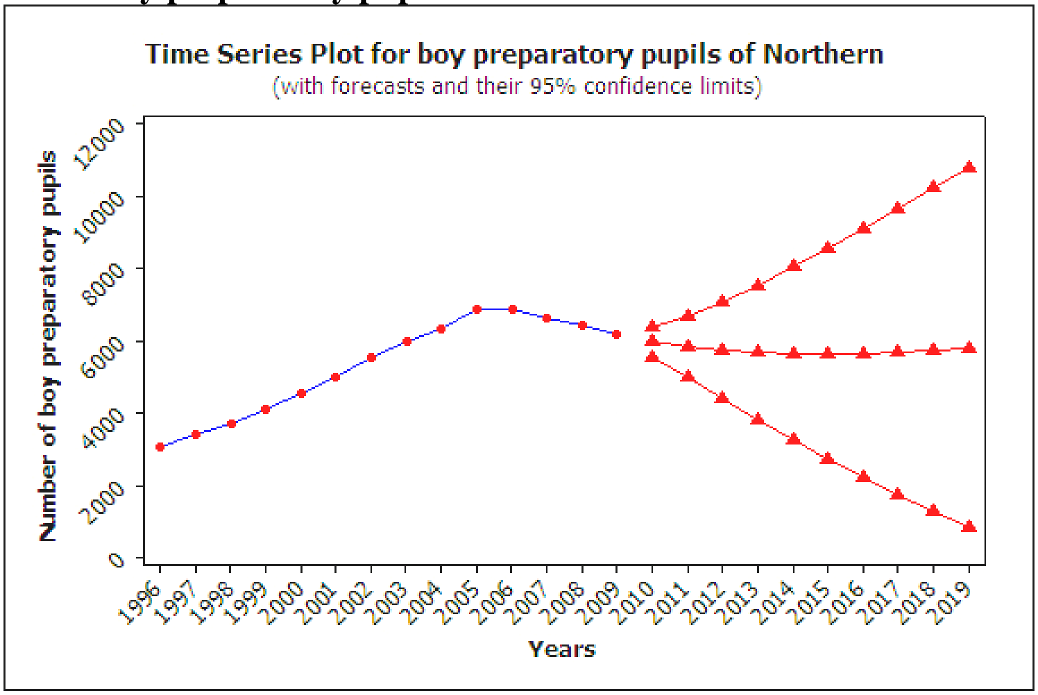
A.1.6.3 Boy preparatory pupils of Khanyounis Western area



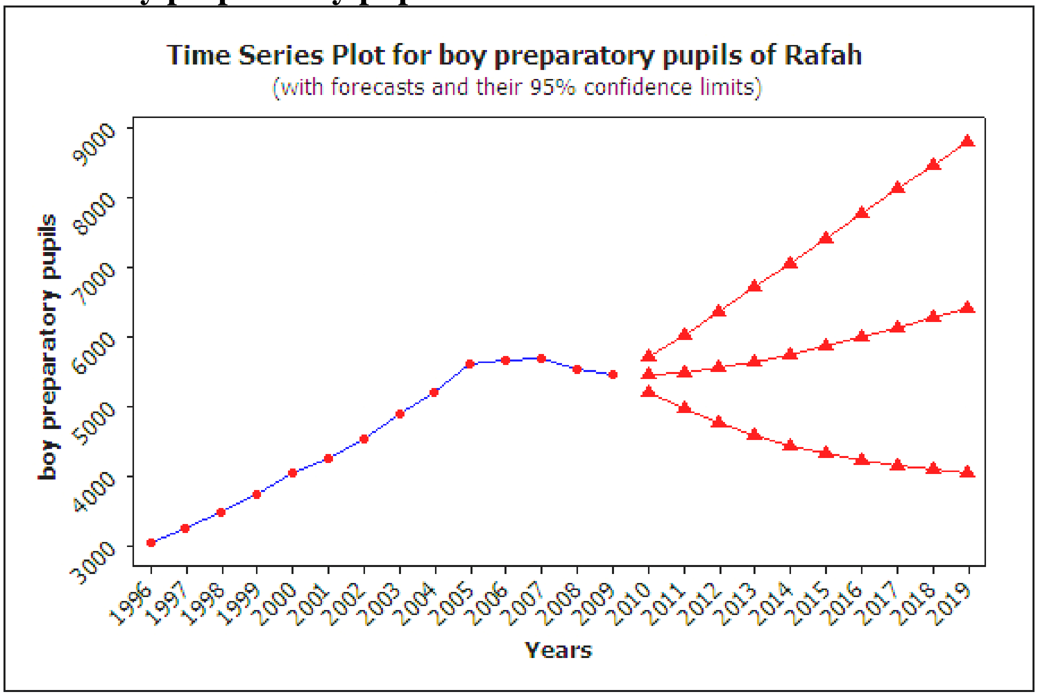
A.1.6.4 Boy preparatory pupils of Middle Governorate



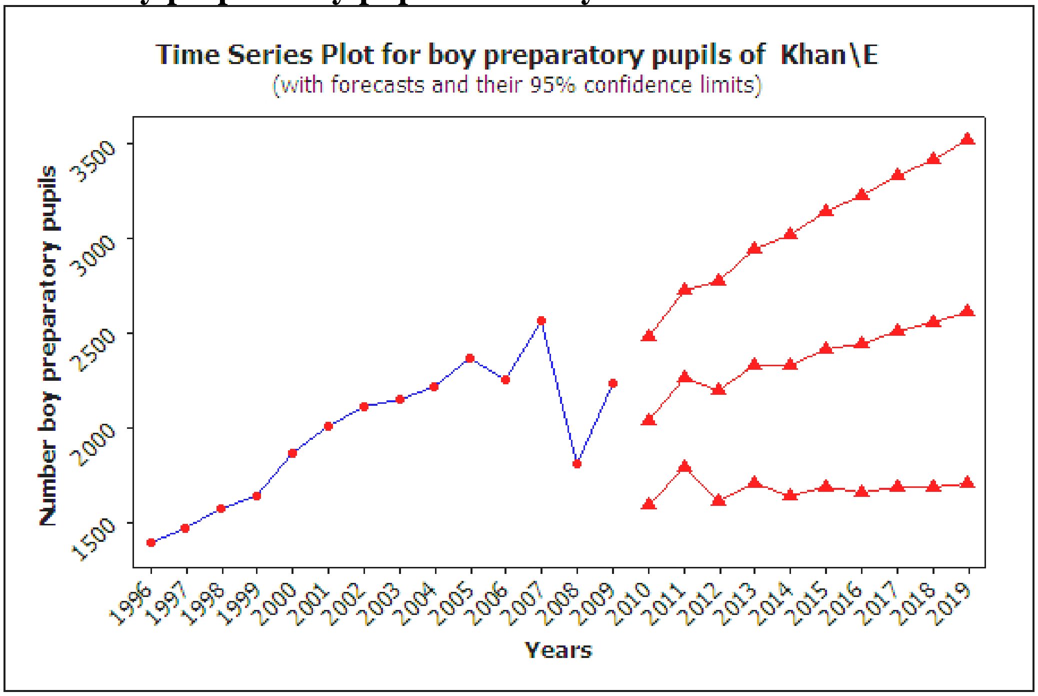
A.1.6.5 Boy preparatory pupils of Northern Governorate



A.1.6.6 Boy preparatory pupils of Rafah Governorate

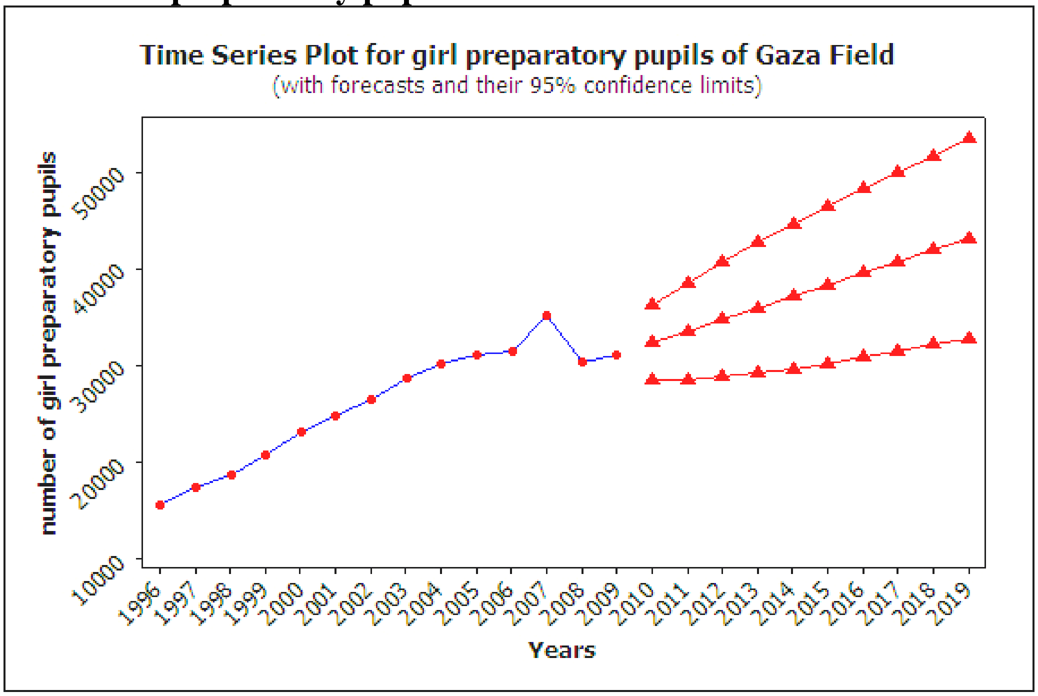


A.1.6.7 Boy preparatory pupils of Khanyounis Eastern area

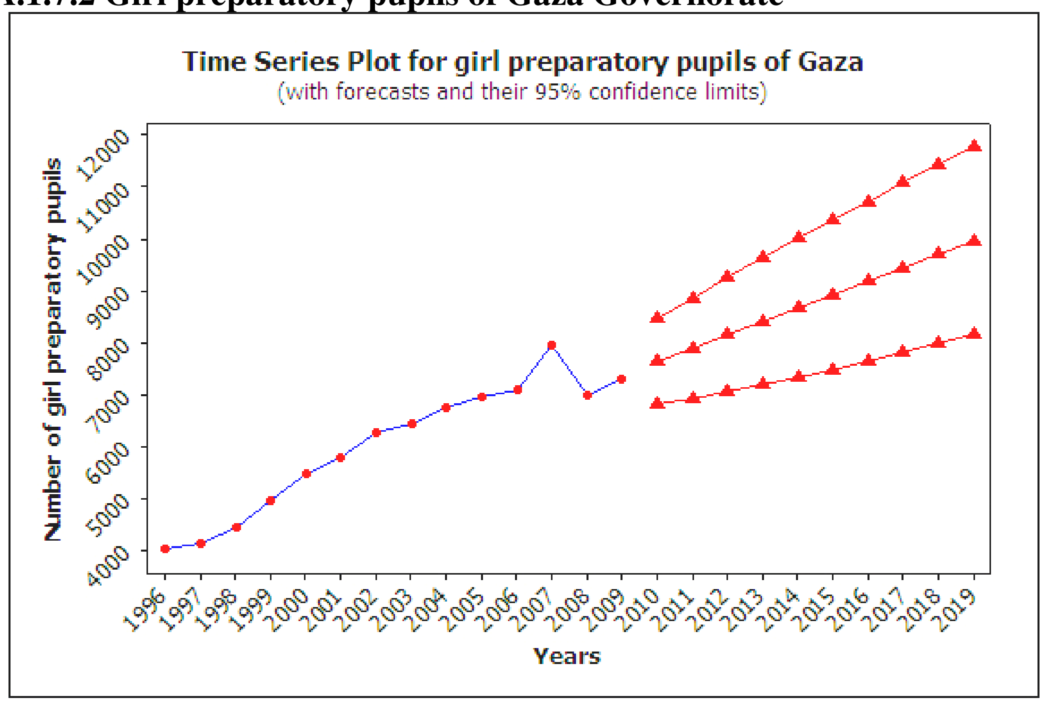


A.1.7 Time series plot for number of girl preparatory pupils

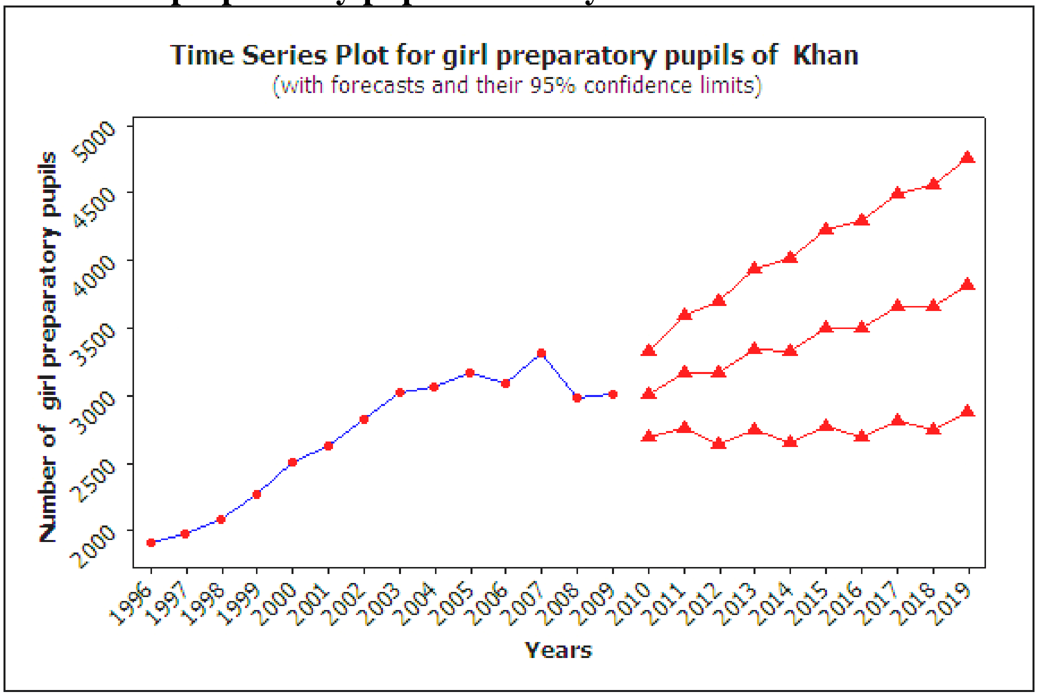
A.1.7.1 Girl preparatory pupils of Gaza Field



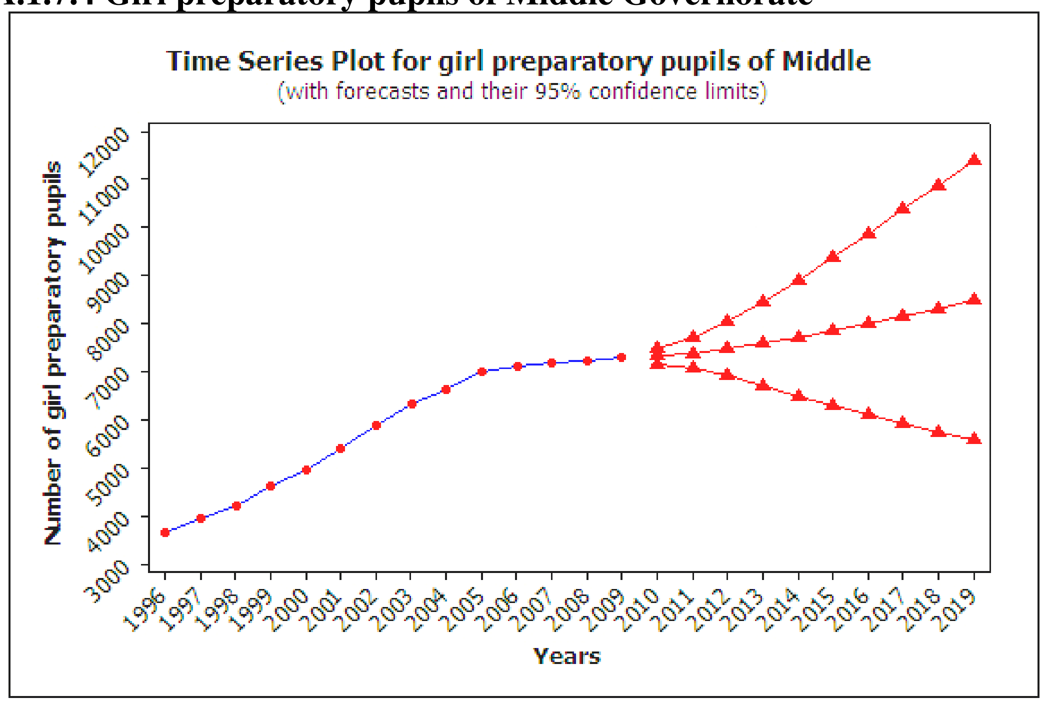
A.1.7.2 Girl preparatory pupils of Gaza Governorate



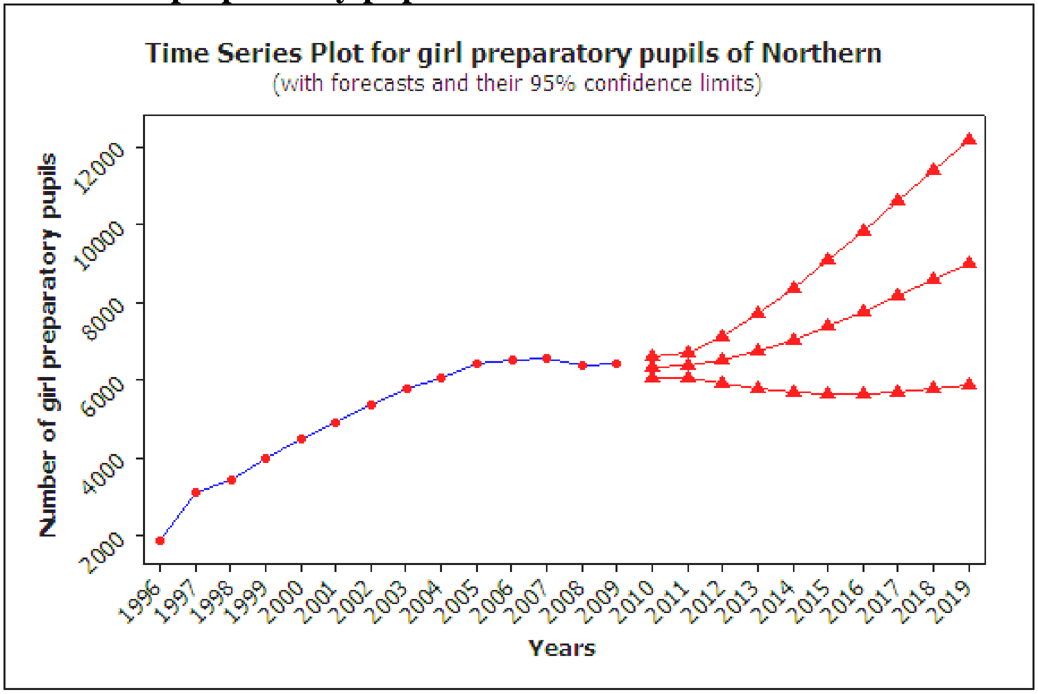
A.1.7.3 Girl preparatory pupils of Khanyounis Western area



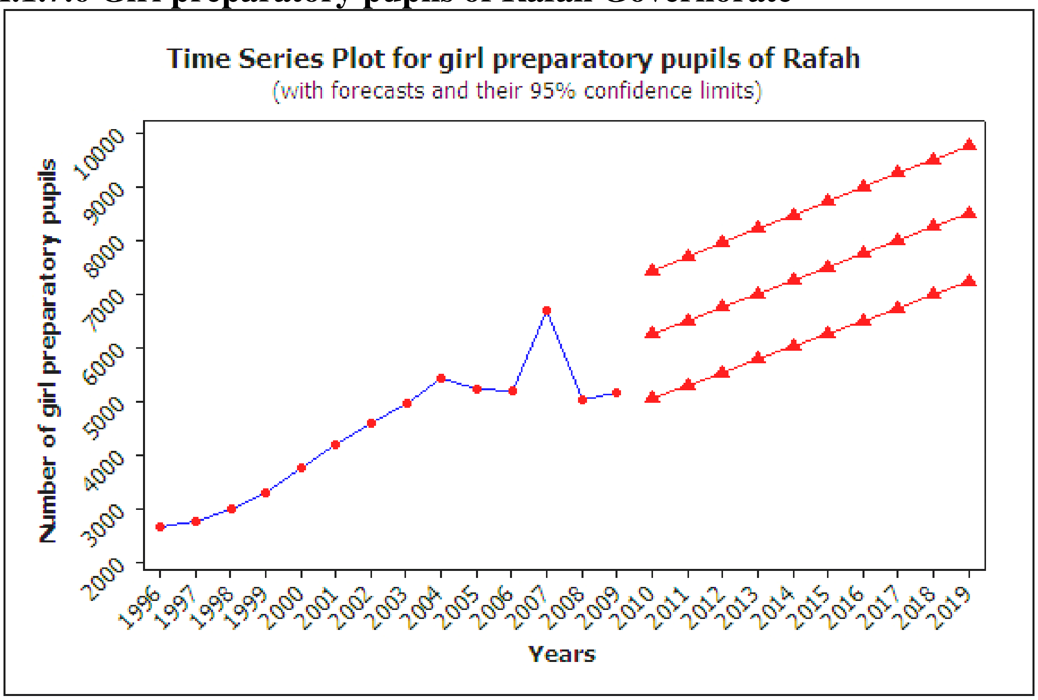
A.1.7.4 Girl preparatory pupils of Middle Governorate



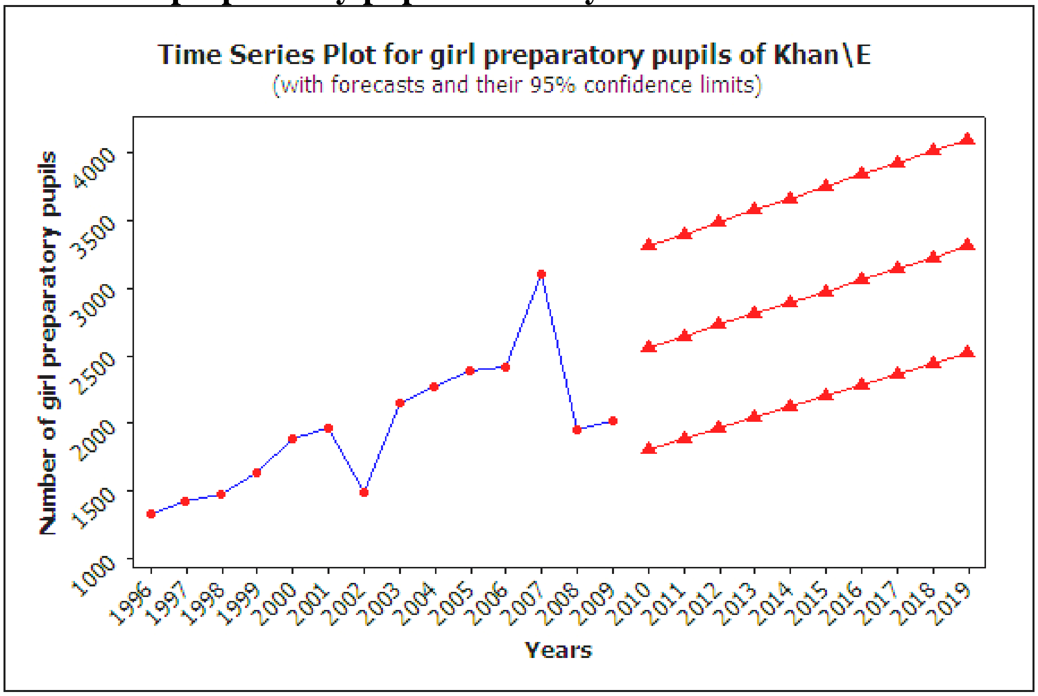
A.1.7.5 Girl preparatory pupils of Northern Governorate



A.1.7.6 Girl preparatory pupils of Rafah Governorate



A.1.7.7 Girl preparatory pupils of Khanyounis Eastern area



Appendix 2: Statistical review

A.2.1 Autocorrelation

For ACF the distance between the lines and zero for the i^{th} autocorrelation are determined by the following formula:

$$2\sqrt{1+2\sum_{k=1}^{i-1}\frac{r_k^2}{\sqrt{n}}}$$

Where n = the number of observations in the series, and r_k = the k_i autocorrelation.

A.2.2 Ljung-Box Q statistic

Use to test whether a series of observations over time are random and independent. If observations are not independent, one observation may be correlated with another observation k time units later, a relationship called autocorrelation. Autocorrelation can impair the accuracy of a time-based predictive model, such as time series plot, and lead to misinterpretation of the data.

LBQ is also used to evaluate assumptions after fitting a time series model, such as ARIMA, to ensure that the residuals are independent. The Ljung-Box Q (LBQ) statistic can be used to test the null hypothesis that the autocorrelations for all lags up to lag k equal zero.

$$Q = n(n+2)\left(\frac{\hat{r}_1^2}{n-1}+\frac{\hat{r}_2^2}{n-2}+\Lambda+\frac{\hat{r}_k^2}{n-k}\right)$$

Where n=the number of data
K=the lag number

Appendix 3: Historical data

A.3.1 Historical data of areas

A.3.1.1 Gaza Field

year	C S	C E	C P B	C P G	P E	P P B	P P G
1996	162	2198	376	351	106420	17747	15448
1997	169	2281	379	360	114864	18693	17312
1998	167	2437	407	378	121395	19830	18667
1999	168	2520	435	424	127148	21540	20736
2000	169	2658	479	465	131768	23763	23049
2001	170	2764	534	518	134569	25498	24890
2002	175	2866	581	561	134849	27490	26463
2003	176	2938	637	634	133866	29535	28698
2004	179	2994	687	661	132907	31126	30138
2005	186	3033	741	699	129102	33409	31151
2006	192	3127	744	704	130467	33196	31448
2007	213	3263	1092	708	127457	33373	35197
2008	218	3498	920	848	136369	31222	30402
2009	230	3673	1050	745	143305	31383	31052

A.3.1.2 Gaza Governorate

year	C S	C E	C P B	C P G	P E	P P B	P P G
1996	41	528	88	87	25505	4168	4032
1997	45	546	89	85	27078	4300	4136
1998	43	574	95	90	28500	4552	4441
1999	43	594	100	99	29655	5000	4945
2000	43	611	113	110	30464	5622	5472
2001	43	624	125	120	30647	5989	5782
2002	43	654	137	132	30994	6489	6248
2003	44	669	146	144	30699	6909	6441
2004	44	680	162	147	30416	7457	6725
2005	46	683	178	158	29240	7822	6937
2006	47	709	177	158	30030	8013	7077
2007	53	764	230	160	30013	7919	7942
2008	57	855	239	170	33328	7662	6994
2009	60	936	257	172	36660	7690	7283

A.3.1.3 Khanyounis Western area

year	C S	C E	C P B	C P G	P E	P P B	P P G
1996	19	227	40	38	11239	1903	1856
1997	18	239	39	39	11998	1974	1927
1998	18	251	42	42	12351	2075	2069
1999	18	257	46	46	12860	2264	2257
2000	18	269	50	50	13264	2497	2475
2001	18	277	55	54	13331	2623	2623
2002	19	284	59	60	13217	2822	2852
2003	19	284	66	65	12876	3019	3000
2004	20	286	65	67	12602	3053	3025
2005	20	285	71	68	12231	3164	3107
2006	20	295	68	70	12179	3077	3120
2007	25	308	110	82	11734	3308	3701
2008	24	335	84	86	12612	2982	3035
2009	26	338	101	72	13040	3008	3067

A.3.1.4 Middle Governorate

year	C S	C E	C P B	C P G	P E	P P B	P P G
1996	32	463	89	79	22383	4117	3681
1997	32	477	86	80	24474	4286	3945
1998	32	516	91	86	26024	4406	4218
1999	33	528	96	96	27383	4778	4637
2000	33	578	104	100	28534	5128	4968
2001	33	597	118	113	29043	5610	5405
2002	35	624	127	129	29349	5970	5900
2003	35	648	142	143	29119	6533	6354
2004	36	651	153	147	29186	6862	6627
2005	38	661	167	156	28294	7540	7023
2006	41	678	164	159	28373	7293	7114
2007	42	695	239	161	28289	7208	7216
2008	44	746	191	198	28657	6798	7023
2009	49	764	230	167	29289	6802	7094

A.3.1.5 Northern Governorate

year	C S	C E	C P B	C P G	P E	P P B	P P G
1996	29	419	64	61	20696	3098	1887
1997	31	445	70	65	22411	3398	3120
1998	32	479	74	69	24008	3724	3457
1999	32	502	83	79	25446	4116	3983
2000	32	529	93	88	26549	4581	4488
2001	32	558	104	102	27409	4998	4934
2002	33	580	118	112	27453	5555	5393
2003	34	593	130	125	27343	6022	5789
2004	34	603	143	136	27141	6340	6064
2005	35	613	150	145	26085	6907	6461
2006	36	627	154	144	25982	6889	6514
2007	37	637	223	147	25693	6668	6562
2008	37	679	179	179	26238	6428	6376
2009	37	687	205	148	27183	6177	6439

A.3.1.6 Rafah Governorate

year	C S	C E	C P B	C P G	P E	P P B	P P G
1996	27	371	66	58	17597	3064	2663
1997	28	380	65	62	19220	3268	2764
1998	28	410	73	61	20332	3501	3006
1999	28	423	76	71	21134	3745	3286
2000	28	447	82	80	22018	4065	3770
2001	29	470	90	88	22779	4268	4187
2002	29	481	96	97	22551	4545	4589
2003	30	490	107	110	22443	4902	4970
2004	31	507	115	114	21845	5201	5433
2005	33	518	121	117	21750	5610	5240
2006	33	532	129	118	21945	5669	5211
2007	40	549	193	115	20478	5709	6684
2008	40	543	167	137	22462	5540	5029
2009	40	590	181	125	23115	5471	5159

A.3.1.7 Khanyounis Eastern area

year	C S	C E	C P B	C P G	P E	P P B	P P G
1996	14	190	29	28	9000	1397	1329
1997	15	194	30	29	9683	1467	1420
1998	14	207	32	30	10180	1572	1476
1999	14	216	34	33	10670	1637	1628
2000	15	224	37	37	10939	1870	1876
2001	15	238	42	41	11360	2010	1959
2002	16	243	44	31	11285	2109	1481
2003	14	254	46	47	11386	2150	2144
2004	14	267	49	50	11717	2213	2264
2005	14	273	54	55	11502	2366	2383
2006	15	286	52	55	11958	2255	2412
2007	16	310	72	43	11250	2561	3092
2008	16	340	60	70	13072	1812	1945
2009	18	358	76	61	14018	2235	2010

Note. C S=number of schools; C E=number of Elementary classrooms; C P B= number of Preparatory Boys classrooms; C P G= number of Preparatory Girls classrooms; P E= number of Elementary Pupils; P P B= number of Preparatory Boys Pupils; P P G= number of Preparatory Girls Pupils