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
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. قطاع غزة لديه خصوصية حيث يعيش سكانه تحت حصار إسرائيلي شديد منذ أربع سنوات.

البحث

المنهجية الإحصائية التحليلية في البحث، وقد تم اختيار نماذج الارتباط الذاتي لرك المتكاملة للتنبؤ بأعداد المدارس والطلاب والصفوف الدراسية لوكالة الغوث.

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

تم تطبيق مجموعة من نماذج الزمنية تحت الدراسة، وقد تم اختيار النتائج لهذه النماذج بفحص

، وبعد ذلك تم اختيار أفضل نموذج والذي يتميز بفضل قيم لمعيار الإعلام الذاتي BIC ومعيار بارزين AIC

، مع معيار الصلابة

، وأظهرت هذه الدراسة أن النتائج كانت دقيقة وتحقق المعايير المطلوبة.

نتوصى بالباحثين في المستقبل بدراسات طرق تنبؤ غير المستخدمة في هذه الدراسة تأثير السنوية لوكالة على عدد المدارس والصفوف الدراسية. }

Developing Forecasting Models for UNRWA Schools in the Gaza Strip
Developing Forecasting Models for UNRWA Schools in the Gaza Strip

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.
Developing Forecasting Models for UNRWA Schools in the Gaza Strip

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>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 4.1</td>
<td>General Behavior of the ACF and PACF for ARMA Models</td>
<td>34</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>AIC, BIC and accuracy measures values</td>
<td>62</td>
</tr>
<tr>
<td>Table 5.2</td>
<td>Final Estimates of Parameters for (1 2 1) model</td>
<td>63</td>
</tr>
<tr>
<td>Table 5.3</td>
<td>Summary for number of school models</td>
<td>69</td>
</tr>
<tr>
<td>Table 5.4</td>
<td>Summary for number of elementary classroom models</td>
<td>70</td>
</tr>
<tr>
<td>Table 5.5</td>
<td>Summary for number of boy preparatory classroom models</td>
<td>71</td>
</tr>
<tr>
<td>Table 5.6</td>
<td>Summary for number of girl preparatory classroom models</td>
<td>72</td>
</tr>
<tr>
<td>Table 5.7</td>
<td>Summary for number of elementary pupil models</td>
<td>73</td>
</tr>
<tr>
<td>Table 5.8</td>
<td>Summary for number of boy preparatory pupil models</td>
<td>74</td>
</tr>
<tr>
<td>Table 5.9</td>
<td>Summary for number of girl preparatory pupil models</td>
<td>75</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Distribution of refugee population by age in Gaza Strip</td>
<td>2</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Development of unrwa’s students and schools over time</td>
<td>3</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Stages of study</td>
<td>5</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Tree of forecasting techniques</td>
<td>12</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Box diagram of Box-Jenkins procedures</td>
<td>23</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>ACF of white noise series</td>
<td>31</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>PACF of white noise series</td>
<td>31</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>Time series plot for elementary classes in Gaza Field</td>
<td>56</td>
</tr>
<tr>
<td>Figure 5.2</td>
<td>Autocorrelation function for Elementary classes in Gaza Field</td>
<td>57</td>
</tr>
<tr>
<td>Figure 5.3</td>
<td>Partial Autocorrelation Function for Elementary classes</td>
<td>57</td>
</tr>
<tr>
<td>Figure 5.4</td>
<td>ACF and PACF plot for residuals of the best fit model</td>
<td>65</td>
</tr>
<tr>
<td>Figure 5.5</td>
<td>Normal probability plot for residual of the best fit model</td>
<td>66</td>
</tr>
<tr>
<td>Figure 5.6</td>
<td>Time series plot including forecasts for the best fit model</td>
<td>67</td>
</tr>
</tbody>
</table>
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>Autocorrelation Function</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike’s Information Criterion</td>
</tr>
<tr>
<td>ARIMA</td>
<td>Autoregressive Integrated and Moving Average</td>
</tr>
<tr>
<td>ARMA</td>
<td>Autoregressive and Moving Average</td>
</tr>
<tr>
<td>BIC</td>
<td>Bayesian’s Information Criterion</td>
</tr>
<tr>
<td>C E</td>
<td>Number of Elementary Classrooms</td>
</tr>
<tr>
<td>C P B</td>
<td>Number of Preparatory Boy Classrooms</td>
</tr>
<tr>
<td>C P G</td>
<td>Number of Preparatory Girl Classrooms</td>
</tr>
<tr>
<td>C S</td>
<td>Number of Schools</td>
</tr>
<tr>
<td>Gaza</td>
<td>Gaza Governorate in Gaza Strip</td>
</tr>
<tr>
<td>Khan</td>
<td>Khanyounis Governorate in Gaza Strip</td>
</tr>
<tr>
<td>Khan E</td>
<td>Eastern area in Khanyounis Governorate</td>
</tr>
<tr>
<td>Middle</td>
<td>Middle Governorate in Gaza Strip</td>
</tr>
<tr>
<td>MoEHE</td>
<td>Ministry of Education and Higher Education</td>
</tr>
<tr>
<td>Northern</td>
<td>Northern Governorate in Gaza Strip</td>
</tr>
<tr>
<td>P E</td>
<td>Number of Elementary Pupils</td>
</tr>
<tr>
<td>P P B</td>
<td>Number of Preparatory Boys Pupils</td>
</tr>
<tr>
<td>P P G</td>
<td>Number of Preparatory Girls Pupils</td>
</tr>
<tr>
<td>PACF</td>
<td>Partial Autocorrelation Function</td>
</tr>
<tr>
<td>PCB</td>
<td>Palestinian Central Bureau Statistics</td>
</tr>
<tr>
<td>PHC-2007</td>
<td>Population, Housing and Establishment Census 2007</td>
</tr>
<tr>
<td>Rafah</td>
<td>Rafah Governorate in Gaza Strip</td>
</tr>
<tr>
<td>UN</td>
<td>United Nation Agency</td>
</tr>
<tr>
<td>UNCTAD</td>
<td>United Nation Conference on Trade and Development</td>
</tr>
</tbody>
</table>
**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

Abstract

Acknowledgments

Dedication

List of Tables

List of Figures

Abbreviations

Glossary of Terms

Chapters

1. Research Background

1.1. Introduction

1.2. Research Problem

1.3. Research Objectives

1.4. Research Variables

1.5. Research Importance

1.6. Research Methodology

1.7. Research Structure

1.8. Scope and Limits

1.9. Previous Studies

2. Theory of Forecasting and Modeling
Developing Forecasting Models for UNRWA Schools in the Gaza Strip

2.1. Introduction .................................................................................................................. 10
2.2. Forecasting ...................................................................................................................... 10
  2.2.1 Scientific Forecasting ............................................................................................... 10
2.3. Forecasting Methods .................................................................................................... 11
  2.3.1 Qualitative Methods ............................................................................................... 11
  2.3.2 Quantitative Methods ............................................................................................. 11
  2.3.3 Causal Method ........................................................................................................ 12
  2.3.4 Forecasting Accuracy .............................................................................................. 13
2.4. Regression Analysis ...................................................................................................... 13
  2.4.1 Simple Linear Regression Model .................................................................. 13
  2.4.2 Simple Coefficient of Determination ............................................................... 13
  2.4.3 Simple Coefficient of Correlation .................................................................. 14
  2.4.4 Multiple Regression Analysis ........................................................................... 14
2.5. Univariate Method ....................................................................................................... 15
  2.5.1 Averaging Models ................................................................................................. 15
  2.5.2 Exponential Smoothing ...................................................................................... 15
2.6. Time Series Regression Models .................................................................................. 17
  2.6.1 Trend Model ........................................................................................................... 17
    2.6.1.1 Linear Trend .................................................................................................. 17
    2.6.1.2 Quadratic Trend ............................................................................................ 17
    2.6.1.3 Polynomial Trend ......................................................................................... 17
  2.6.2 The least Square Method .................................................................................... 18
  2.6.3 Time Series Regression Assumption .................................................................... 18
2.7. Autocorrelation Analysis ............................................................................................ 19
  2.7.1 Positive Auto-Correlation .................................................................................. 19
  2.7.2 Negative Auto-Correlation .................................................................................. 19
  2.7.3 The Scatter Diagram of Error Term ..................................................................... 19
  2.7.4 Order of Autocorrelation ..................................................................................... 20
2.7.5. Autocorrelation Models ................................................................. 20
2.8. Decomposition Methods ................................................................. 21
2.8.1. Multiplicative Decomposition ....................................................... 21
2.8.2. Additive Decomposition ............................................................... 21
2.9. Box-Jenkins Methods (ARIMA Models) ........................................... 22
2.9.1. Box-Jenkins Methods Steps ......................................................... 22
2.9.2. Stationary and Nonstationary Time Series ..................................... 24
2.10. Choosing a Forecasting Method .................................................... 24
2.10.2. Discussion over Suitable Forecasting Method .............................. 25
2.11. Model Building Strategy ............................................................... 25
2.11.1. Model Identification ................................................................. 25
2.11.2. Model Fitting ............................................................................... 26
2.11.3. Model Diagnostics ..................................................................... 26
2.11.4. Implementation and Decision Making ......................................... 26

3. Time Series Models ...........................................................................28

3.1. General Definitions .......................................................................... 28
3.1.1. Residuals (Application Errors) ..................................................... 28
3.1.2. Measures and Accuracy ............................................................... 28
3.1.3. Stationary Series .......................................................................... 29
3.1.4 Autocovariance Function ............................................................... 29
3.1.5. Autocorrelation Function (ACF) ................................................... 29
3.1.6. Partial Autocorrelation Function (PACF) ...................................... 29
3.1.7. Sample Autocorrelation Function SACF ...................................... 30
3.1.8. Sample Partial Autocorrelation Function SPACF .......................... 30
3.1.9. White Noise Series or White Noise Process ................................. 30
3.2. ARMA Models ................................................................................ 31

4. Model Inference and Forecasting .......................................................33
4.1. Models Identification ................................................................. 33
4.1.1. Checking The Time Series for Nonstationarity in Variation ....... 33
4.1.2. Checking The Time Series Data for Nonstationarity in Mean .... 33
4.1.3. Choosing The Values of P and q for A given Series ................ 33
4.2. Model Estimation ........................................................................ 34
4.2.1. The Method of Moments .......................................................... 34
4.2.2. Moments Estimation for Autoregressive Models .................... 36
4.2.2.1. AR(1) Model ..................................................................... 36
4.2.2.2. AR(2) Model ..................................................................... 36
4.2.2.3. General AR(p) .................................................................... 36
4.2.3. Moments Estimation for Moving Average Models ................. 36
4.2.3.1. MA(1) .............................................................................. 36
4.2.3.2. MA(2) .............................................................................. 37
4.2.3.3. General MA(p) ................................................................. 37
4.2.4. Moments Estimation for ARMA(1,1) .................................... 38
4.2.5. Conditional Least Squares Method ......................................... 39
4.2.5.1. Conditional Least Squares Estimation for ARMA(p,q) ....... 39
4.2.5.2. Conditional Least Squares Estimation for AR(1) ............... 40
4.2.5.3. Conditional Least Squares Estimation for MA(1) ............... 41
4.2.6. Maximum likelihood Method .................................................. 42
4.2.6.1. Maximum Likelihood Estimation ....................................... 42
4.2.6.1.1. Maximum Likelihood Estimation for AR(1) .................. 43
4.3. Model Checking and Diagnostics .................................................. 43
4.3.1. Plot of The Residuals (Zero Mean Assumption) .................... 44
4.3.2. Normality Diagnostic ............................................................... 44
4.3.3. Independence or Correlation Diagnostic ............................. 45
4.3.4. The Ljung-Box Test ................................................................. 45
4.3.5. Randomness .......................................................................... 45
Developing Forecasting Models for UNRWA Schools in the Gaza Strip

4.3.6. Model Selection Criteria ................................................................. 45
  4.3.6.1. Automatic Information Criteria (AIC) ........................................ 45
  4.3.6.2. Bayesian Information Criteria (BIC) ........................................ 46
4.4. Forecasting .......................................................................................... 46
  4.4.1. Forecast Error .................................................................................. 46
  4.4.2. Minimum Mean Square Error Forecasting (MMSE Forecasts) ....... 46
  4.4.3. AR (1) Model Forecast ................................................................. 47
  4.4.4. AR (2) Model Forecast ................................................................. 49
  4.4.5. MA (1) Model Forecast ................................................................. 49
  4.4.6. MA (2) Model Forecast ................................................................. 50
  4.4.7. ARMA (p,q) .................................................................................. 51
  4.4.8. ARMA(1,1) Model Forecast .......................................................... 52
  4.4.9. Nonstationary Models .................................................................... 52
  4.4.10. ARIMA(1,1,1) ............................................................................ 53
  4.4.11. Forecasting Limits ........................................................................ 54

5. Results and Analysis .............................................................................. 56
  5.1. Scatter Time Plot ................................................................................ 56
  5.2. Model Identification .......................................................................... 56
    5.2.1. Analysis of ACF and PACF Behavior .......................................... 58
  5.3. Applying ARIMA Models ................................................................. 58
  5.4. Model Fitting Criteria ........................................................................ 62
  5.5. Fitting The Likely Model ................................................................... 62
  5.6. Diagnostic Check on The Mean of The Best Fit Model .................... 63
  5.7. Diagnostic Check on Residuals Randomness .................................... 63
  5.8. Diagnostic Check on ACF and PACF of Residuals ......................... 64
  5.9. Diagnostic Check on Residuals Normality ....................................... 65
    5.9.1. Normal Probability Plot .............................................................. 65
A.2.1 Autocorrelation................................................................. 114
A.2.2 Ljung-Box Q statistic.......................................................... 114

Appendix 3: Historical data...................................................... 115
A.3.1 Historical data of areas .................................................... 115
Chapter One: Research Background

1.1. Introduction ......................................................................................... 1
1.2. Research Problem .................................................................................. 3
1.3. Research Objectives ............................................................................... 3
1.4. Research Variables ............................................................................... 3
1.5. Research Importance ........................................................................... 4
1.6. Research Methodology ......................................................................... 4
1.7. Research Structure ............................................................................... 6
1.8. Scope and Limits ................................................................................... 6
1.9. Previous Studies ................................................................................... 6
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.
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.
Developing Forecasting Models for UNRWA Schools in Gaza Strip  Chapter 1 Research Background

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 lack 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

2.1. Introduction ........................................................................................................ 10
2.2. Forecasting ........................................................................................................ 10
2.3. Forecasting Methods ....................................................................................... 11
2.4. Regression Analysis ......................................................................................... 13
2.5. Univariate Method .......................................................................................... 15
2.6. Time Series Regression Models ..................................................................... 17
2.7. Autocorrelation Analysis ............................................................................... 19
2.8. Decomposition Methods ............................................................................... 21
2.9. Box-Jenkins Methods (ARIMA Models) ..................................................... 22
2.10. Choosing a Forecasting Method ................................................................... 24
2.11. Model Building Strategy ............................................................................. 25
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.

\[ Y = \beta_o + \beta_1 X + \epsilon \]

Where  
\( \beta_o = \) Y-intercept  
\( \beta_1 = \) slope of the line  
\( \epsilon = \) 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).

\[ Y = \beta_o + \beta_1X_1 + \beta_2X_2 + \Lambda + \beta_kX_k + \varepsilon \]

Where
- \( Y \) = dependent variable
- \( X_1, X_2, \Lambda, X_k \) = independent variables
- \( \beta_o, \beta_1, \beta_2, \Lambda, \beta_k \) = parameters
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.

\[
F_{t+1} = \frac{Z_t + Z_{t-1} + Z_{t-2} + \ldots + Z_{t-m+1}}{m}
\]

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).

\[
F_{t+1} = \alpha Z_t + (1-\alpha)\alpha Z_{t-1} + (1-\alpha)^2 \alpha Z_{t-2} + K + (1-\alpha)^{-1} \alpha Z_t + (1-\alpha)^t F_t
\]

Where

- \( F_{t+1} \) = Forecast for time period \( t+1 \)
- \( Z_t \) = Y value for time period \( t \)
- \( \alpha \) = 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 \( t \)
- \( \gamma \) = 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 \( t \)
- \( \varepsilon_t \) = 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 \( \beta_1 \) 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 \((\beta_o, \beta_1)\) 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 \((\beta_0, \beta_1)\), which minimize the sum of square of error terms \(\text{SSE}\) (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).

\[ e_t = Z_t - \hat{Z}_t \]
\[ \text{SSE} = \sum_{t=1}^{n} e_t^2 = \sum_{t=1}^{n} (Z_t - \hat{Z}_t)^2 \]

Where \(Z_t\) = observed value of time period \(t\)
\(\hat{Z}_t\) = estimated value of time period \(t\)

\[ \beta_1 = \frac{n \sum_{t=1}^{n} tZ_t - (\sum_{t=1}^{n} t)(\sum_{t=1}^{n} Z_t)}{n \sum_{t=1}^{n} t^2 - (\sum_{t=1}^{n} t)^2} \]
\[ \beta_0 = \frac{\sum_{t=1}^{n} t^2 \beta_1 + \sum_{t=1}^{n} Y_t}{n} \]

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 \( e \) 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 \( \varepsilon_t \), error in period \( t \), is correlated to \( \varepsilon_{t-1} \), error time in period \( t-1 \) (Bowerman, 2005, p 291).

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

Where \( \varepsilon_t \) = the error term in time period \( t \)
\( \phi \) = 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.

\[
Z_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_k x_{tk} + \varepsilon_t
\]

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 itself 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).

\[ Z_t = TR_t \times SN_t \times CL_t \times IR_t \]

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).
Developing Forecasting Models for UNRWA Schools in Gaza Strip  Chapter 2 Theory of Forecasting

\[ Z_t = TR_t + SN_t + CL_t + IR_t \]

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

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 (cyrer, 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 a more 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
   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

3.1. General Definitions ..........................................................28
3.2. ARMA Models ..............................................................31
Developing Forecasting Models for UNRWA Schools in Gaza Strip
Chapter 3 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 around 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).

\[
MAD = \frac{1}{n} \sum_{t=1}^{n} |e_t| = \frac{1}{n} \sum_{t=1}^{n} |Z_t - \hat{Z}_t|
\]

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

\[
MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{Z_t - \hat{Z}_t}{Z_t} \right| \times 100 \%
\]

\( Z_t \neq 0 \)


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

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

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

3.1.4. Autocovariance Function
The Autocovariance Function is defined as

\[ \gamma_{t,s} = \text{cov}(Z_t, Z_s), \ \forall t, \forall s \]

\[ = E[(Z_t - \mu)(Z_s - \mu)], \ \forall t, \forall s \]

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

\[ \gamma_k = \text{cov}(Z_t, Z_{t-k}), \ \ k = 0, \pm 1, \pm 2, L \]

\[ = E[(Z_t - \mu)(Z_{t-k} - \mu)], \ \ k = 0, \pm 1, \pm 2, L \]

3.1.5. Autocorrelation Function (ACF)
The Autocorrelation Function is defined as

\[ \rho_k = \frac{\gamma_k}{\gamma_0}, \ \ k = 0, \pm 1, \pm 2, L \]

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 \( Z_{t-1}, Z_{t-2}, ..., Z_{t-k+1} \) which locate within \( (t, t-k) \) period, partial autocorrelation function will be donated by \( \phi_k \), PACF will be calculated by iteration (Shumway, 2006, p. 106).

\[ \phi_0 = 1 \]

\[ \phi_1 = \rho_1 \]
Developing Forecasting Models for UNRWA Schools in Gaza Strip  Chapter 3 Time Series Models

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

Where \( \phi_{kj} = \phi_{k-1,j} - \phi_{kk} \phi_{k-1,k-1} \), \( j = 1, 2, ..., k-1 \)

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

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

Where \( \bar{z} = \frac{1}{n} \sum_{t=1}^{n} z_t \), \( r_k \) is an estimator for \( \rho_k \)

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

\[ r_{00} = 1 \]
\[ r_{11} = r_i \]
\[ r_k = \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j} \]
\[ r_{kk} = \frac{1}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} \], \quad k = 2, 3, ...

Where \( r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-1} \), \( j = 1, 2, ..., k-1 \), \( r_{kk} \) is an estimator for \( \phi_{kk} \)

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 \( \sigma^2 \) (Cryer, 2008, p. 17).

1) \( E(a_t) = 0, \forall t \)
2) \( \text{cov}(a_t, a_s) = \begin{cases} \sigma^2, & \forall t, \forall s, t = s \\ 0, & \forall t, \forall s, t \neq s \end{cases} \)
Developing Forecasting Models for UNRWA Schools in Gaza Strip  Chapter 3 Time Series Models

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

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

**Figure 3.1 ACF of white noise series**

![ACF of white noise series](image)

**Figure 3.2 PACF of white noise series**

![PACF of white noise series](image)

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

4.1. Models Identification .................................................. 33
4.2. Model Estimation .......................................................... 34
4.3. Model Checking and Diagnostics ..................................... 43
4.4. Forecasting ................................................................. 46
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

<table>
<thead>
<tr>
<th></th>
<th>AR(p)</th>
<th>MA(q)</th>
<th>ARMA(p,q), p&gt;0 and q&gt;0</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>Tail off</td>
<td>Cuts off after lag q</td>
<td>Tail off</td>
</tr>
<tr>
<td>PACF</td>
<td>Cuts off after lag p</td>
<td>Tail off</td>
<td>Tail off</td>
</tr>
</tbody>
</table>

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 \( z = \{z_1, z_2, \ldots, z_n\} \). 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 \( \mu \) is estimated by the estimator \( \bar{z} \)
\[
\hat{\mu} = \bar{z} = \frac{1}{n} \sum_{i=1}^{n} z_i
\]

To estimate \( \phi_1, \ldots, \phi_p \), use the relation:
\[
\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p}, \quad k > 1
\]

This relation is deduced 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).
\[
\begin{align*}
\rho_1 &= \phi_1 + \phi_2 \rho_1 + \phi_p \rho_{p-1} \\
\rho_2 &= \phi_1 \rho_1 + \phi_2 + \phi_p \rho_{p-2} \\
& \vdots \\
\rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \cdots + \phi_p
\end{align*}
\]

Substituting \( \rho_k \) by the estimator \( \hat{\rho}_k \), the method of moment’s parameters \( \hat{\phi}_1, \ldots, \hat{\phi}_p \) will be as follow:

Transforming Yule-Walker equations to matrix form
\[
\begin{pmatrix}
r_1 \\
r_2 \\
\vdots \\
r_p
\end{pmatrix} =
\begin{pmatrix}
1 & r_1 & r_2 & \cdots & r_{p-2} & r_{p-1} \\
r_1 & 1 & r_1 & \cdots & r_{p-3} & r_{p-2} \\
& & \ddots & \ddots & \ddots & \ddots \\
r_{p-1} & r_{p-2} & r_{p-3} & \cdots & 1 & r_1
\end{pmatrix}
\begin{pmatrix}
\hat{\phi}_1 \\
\hat{\phi}_2 \\
\vdots \\
\hat{\phi}_p
\end{pmatrix}
\]

Solving this equation to get parameters
\[
\begin{pmatrix}
\hat{\phi}_1 \\
\hat{\phi}_2 \\
\vdots \\
\hat{\phi}_p
\end{pmatrix} = \begin{pmatrix}
1 & r_1 & r_2 & \cdots & r_{p-2} & r_{p-1} \\
r_1 & 1 & r_1 & \cdots & r_{p-3} & r_{p-2} \\
& & \ddots & \ddots & \ddots & \ddots \\
r_{p-1} & r_{p-2} & r_{p-3} & \cdots & 1 & r_1
\end{pmatrix}^{-1}
\begin{pmatrix}
r_1 \\
r_2 \\
\vdots \\
r_p
\end{pmatrix}
\]

\( \sigma^2 \) is estimated by the equation
\[
\hat{\sigma}^2 = \gamma_0 \left( 1 - \hat{\phi}_1 r_1 - \hat{\phi}_2 r_2 - \cdots - \hat{\phi}_p r_p \right)
\]
Where
\[ \hat{\gamma}_0 = \frac{1}{n} \sum_{i=1}^{n} (z_i - \bar{z})^2 \]
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, \( \rho_1 \) is equated to \( r_1 \), the lag 1 sample autocorrelation. Thus, \( \phi \) 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 \( \phi_1 \) and \( \phi_2 \) and various moments are given by the Yule-Walker equations.
\[ \rho_1 = \phi_1 + \rho_1 \phi_2 \quad \text{and} \quad \rho_2 = \rho_1 \phi_1 + \phi_2 \]
The method of moments replaces \( \rho_1 \) by \( r_1 \) and \( \rho_2 \) by \( r_2 \) to obtain \( \hat{\phi} = r_1 \)
\[ r_1 = \phi_1 + r_1 \phi_2 \quad \text{and} \quad r_2 = r_1 \phi_1 + \phi_2 \]
which are solved to obtain
\[ \hat{\phi}_1 = \frac{r_1 (1-r_2)}{1-r_1^2} \quad \text{and} \quad \hat{\phi}_2 = \frac{r_2 - r_1^2}{1-r_1^2} \]

4.2.2.3. General AR(p)
The general AR(p) case proceeds similarly. \( \rho_k \) 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 \( \rho_1 \) is equated to \( r_1 \), quadratic equation is found in \( \theta \)

The solution for the quadratic equation is:

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

This solution gives two values for the estimator \( \hat{\theta} \), only the value that satisfies the condition \(|\hat{\theta}| < 1\) is considered

### 4.2.3.2. MA(2)

Finding the moment of estimators for the parameters \( \theta_1 \) and \( \theta_2 \), the following relations will be used

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

Then \( \rho_1 \) is replaced by \( r_1 \) and \( \rho_2 \) by \( r_2 \) and the equations are solved to get the estimators \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \) which satisfy the conditions:

\[ \theta_2 - \theta_1 < 1, \quad \theta_2 + \theta_1 < 1, \quad |\theta_2| < 1 \]

### 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
Developing Forecasting Models for UNRWA Schools in Gaza Strip .......... Chapter 4 Model Inference

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

If \( \rho_k \) is replaced by \( r_k \) for \( k = 1, 2, \ldots, q \), \( q \) equations will be deducted to \( q \) unknowns \( \theta_1, \theta_2, \Lambda, \theta_q \). The resulting equations are highly nonlinear in the \( \theta \)'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 \( \theta_1 \) and \( \phi_1 \), 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 \( \rho_1 \) is replaced by \( r_1 \) and \( \rho_2 \) by \( r_2 \) then the equations are solved to get the estimators \( \hat{\theta}_1 \) and \( \hat{\phi}_1 \)

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

By dividing \( r_2 \) equation by \( r_1 \) 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 \( r_1 \) equation,

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

Then quadratic equation is solved and \( \hat{\theta}_1 \) value which satisfy the condition \( |\hat{\theta}_1| < 1 \) 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 \( \delta \) can be included in a nonzero Constant mean, \( \mu \), 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 \( \phi = \{ \phi_1, \phi_2, \ldots, \phi_p \} \), \( \theta = \{ \theta_1, \theta_2, \ldots, \theta_q \} \)

and

\[ a_t(\phi, \theta, \mu) = \frac{1 - \phi_1B - \phi_2B^2 - \ldots - \phi_pB^p}{1 - \theta_1B - \theta_2B^2 - \ldots - \theta_qB^q}(z_t - \mu) \]

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

\[ \min_{\phi, \theta, \mu} S_c(\phi, \theta, \mu) = \sum_{i=p+1}^{n} a_i^2(\phi, \theta, \mu | z) \]

Finding the required estimators needs solving the following normal equations.
These estimators are called conditional because they satisfy that $a_p = a_{p-1} = L = a_{p+1-q} = 0$ or because the values $a_p = a_{p-1} = K = a_{p+1-q}$ equal their expectation (Shumway, Stoffer, 2005, p. 128).

The variance $\sigma^2$ equals

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

### 4.2.5.2. Conditional Least Squares Estimation for AR(1)

$$z_t = \delta + \phi_1z_{t-1} + a_t$$

$$\therefore \delta = \mu(1-\phi_1)$$

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

Replacing $\mu$ 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 $z = \{z_1, z_2, K, z_n\}$, the error is

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

Now, the sum of square errors will be minimized
Developing Forecasting Models for UNRWA Schools in Gaza Strip .......... Chapter 4 Model Inference

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

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

Now the equation \( \partial S_c / \partial \phi = 0 \) will be considered for \( \phi_1 \) 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 \( \phi_1 \) parameter and the equation \( \partial S_c / \partial \mu = 0 \) can be solved for \( \mu \) parameter

4.2.5.3. Conditional Least Squares Estimation for MA(1)

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

\[ Z_t - \mu = a_t - \theta a_{t-1} \]

By replacing \( \mu \) with its estimator \( \bar{z} \)

\[ z_t - \bar{z} = a_t - \theta a_{t-1} \]

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

\[ x_t = a_t - \theta a_{t-1} \]

\[ a_t = x_t - \theta a_{t-1} \]

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

\[ a_1 = x_1 \]

\[ a_2 = x_2 - \theta a_1 \]

\[ a_3 = x_3 - \theta a_2 \]

\[ M \]

\[ a_n = x_n - \theta a_{n-1} \]

So

\[ S_c(\theta_1) = \sum_{t=1}^{n} a_t^2 \]
the previous equation is nonlinear and the value of \( \theta_i \) which minimize \( S_c(\theta_i) \) 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, \( z = \{z_1, z_2, \ldots, z_n\} \), 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 \( \phi \) s, \( \theta \) s, \( \mu \), and given the observations
\[ 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.

\[ a_t \approx i.i.d. \quad N(0, \sigma^2) \]

Gaussian white noises series are independent, identical, normally distributed random variables with zero means and common standard deviation \( \sigma \) (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 \( \phi \) and \( \mu \), \( l(\phi, \mu, \sigma^2) \) can be maximized analytically with respect to \( \sigma^2 \) in terms of the yet-to-be determined estimators of \( \phi \) and \( \mu \). The deducted equation is

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

Now, consider the estimation of \( \phi \) and \( \mu \) 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 \( z = \{z_1, z_2, \ldots, z_n\} \)

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

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 \( z = \{z_1, z_2, \ldots, z_n\} \).
Developing Forecasting Models for UNRWA Schools in Gaza Strip ........ Chapter 4 Model Inference

\[ e_t = z_t - \hat{z}_t \]

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.
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)

\[ AIC = -2 \log (\text{maximum likelihood}) + 2k \]
\[ AIC(m) = n \ln \sigma_a^2 + 2m \]

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)

\[ BIC = -2 \log(\text{maximum likelihood}) + k \log(n) \]

\[ BIC(m) = n \ln \sigma^2 + m \ln(n) \]

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

4.4.1. Forecast Error

It is the difference between the future forecast \( z_n(1), 1 \geq 0 \) and the future observation \( z_{n+1}, 1 \geq 0 \). 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).

\[ e_n(1) = z_{n+1} - z_n(1), \quad 1 \geq 0 \]

4.4.2. Minimum Mean Square Error Forecasting (MMSE Forecasts)

(Cryer, 2008, p. 191)

\[ z_n(1) = E(z_{n+1} | z_n, z_{n-1}, \ldots), \quad 1 \geq 1 \]

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 - \( E(a_{n+j}|z_n, z_{n-1}, L) = \begin{cases} a_{n+j}, & j \leq 0 \\ 0, & j > 0 \end{cases} \)

2 - \( E(z_{n+j}|z_n, z_{n-1}, L) = \begin{cases} z_{n+j}, & j \leq 0 \\ z_n(j), & j > 0 \end{cases} \)

4.4.3. AR (1) Model Forecast

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

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

Now, future forecast \( z_{n+1}, z_{n+2}, z_{n+3}, \ldots \) or in general form \( z_{n+k}, \ k \geq 1 \)

\[
Z_{n+k} - \mu = \phi(Z_{t-1} - \mu) + a_{t+k}
\]

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

\[
z_n(1) = E(\{z_n, z_{n-1}, L\}, 1 \geq 1)
= \mu + E(\phi_1(z_{n+1} - \mu) + a_{n+1}|z_n, z_{n-1}, L), 1 \geq 1
= \mu + E(\phi_1(z_{n+1} - \mu) + a_{n+1}|z_n, z_{n-1}, L) + E(a_{n+1}|z_n, z_{n-1}, L), 1 \geq 1
= \mu + \phi_1(E(z_{n+1} - \mu) + E(a_{n+1}|z_n, z_{n-1}, L)), 1 \geq 1
\]

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

\[
1 = 1: z_n(1) = \mu + \phi_1(E(z_n - \mu) + E(a_{n+1}|z_n, z_{n-1}, L))
\]

\[
1 = 2: z_n(2) = \mu + \phi_1(z_n(1) - \mu)
\]

\[
1 = 3: z_n(3) = \mu + \phi_1(z_n(2) - \mu)
\]
The general form can be written as:

\[ Z_n(\lambda) = \mu + \phi[Z_{n}(\lambda-1) - \mu] \quad \text{for} \quad \lambda \geq 1 \]

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.

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

**Forecast Error for AR (1)**

General Linear process for AR (1) is

\[ Z_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \phi^3 a_{t-3} + \Lambda \]

After some derivations, the deducted equation is

\[ e_n(\lambda) = e_{n+\lambda} + \phi \cdot e_{n+\lambda-1} + K + \phi^{\lambda-1} e_{n+1} \]

This equation can be written in the weights form

\[ e_n(\lambda) = e_{n+\lambda} + \psi_1 \cdot e_{n+\lambda-1} + \psi_2 \cdot e_{n+\lambda-2} + K + \psi_{\lambda-1} \cdot e_{n+1} \]

Since, \( |\phi| < 1 \)

\[ E(e_n(\lambda)) = 0 \]

\[ \text{Var}(e_n(\lambda)) = \sigma_e^2 (1 + \psi_1^2 + \psi_2^2 + K + \psi_{\lambda-1}^2) \]

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

After finding a finite series

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

For large \( \lambda \), the variance can be calculated be the following equation

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

The following equation will be valid for all stationary ARMA processes

\[ \text{Var}(e_n(\lambda)) \approx \text{Var}(Z_t) = \gamma_0 \quad \text{for} \quad \text{large} \quad \lambda \]
4.4.4. AR (2) Model Forecast

\[ z_t - \mu = \phi_1 (z_{t-1} - \mu) + \phi_2 (z_{t-2} - \mu) + a_t, \quad a_t \sim WN(0, \sigma^2), |\phi_i| < 1, \mu \in (-\infty, \infty) \]

\[ \phi_2 - \phi_1 < 1 \]

Where \( \phi_2 + \phi_1 < 1 \)

\[ -1 < \phi_2 < 1 \]

Future forecast is

\[ z_n(1) = E(z_{n+1}|z_n, z_{n-1}, L), \quad 1 \geq 1 \]

\[ = \mu + E\left\{ \phi_1 (z_{n+1} - \mu) + \phi_2 (z_{n+1} - \mu) + a_{n+1}\right\}z_n, z_{n-1}, L), \quad 1 \geq 1 \]

\[ = \mu + E\left\{ \phi_1 (z_{n+1} - \mu)|z_n, z_{n-1}, L + \phi_2 (z_{n+1} - \mu)|z_n, z_{n-1}, L + a_{n+1}|z_n, z_{n-1}, L\right\}, \quad 1 \geq 1 \]

\[ = \mu + \phi_1 E\left( z_{n+1} | z_n, z_{n-1}, L \right) - \mu \right\} + \phi_2 E\left( z_{n+1} | z_n, z_{n-1}, L \right) - \mu \right\} + E\left[ a_{n+1} | z_n, z_{n-1}, L \right], \quad 1 \geq 1 \]

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

\[ 1 = 1: z_n(1) = \mu + \phi_1 E\left( z_n | z_n, z_{n-1}, L \right) - \mu \right\} + \phi_2 E\left( z_n | z_n, z_{n-1}, L \right) - \mu \right\} + E\left[ a_{n+1} | z_n, z_{n-1}, L \right] \]

\[ = \mu + \phi_1 \left( z_n - \mu \right) + \phi_2 \left( z_n - \mu \right) \]

\[ 1 = 2: z_n(2) = \mu + \phi_1 E\left( z_n | z_n, z_{n-1}, L \right) - \mu \right\} + \phi_2 E\left( z_n | z_n, z_{n-1}, L \right) - \mu \right\} + E\left[ a_{n+2} | z_n, z_{n-1}, L \right] \]

\[ = \mu + \phi_1 \left( z_n(1) - \mu \right) + \phi_2 \left( z_n - \mu \right) \]

\[ 1 = 3: z_n(3) = \mu + \phi_1 E\left( z_n | z_n, z_{n-1}, L \right) - \mu \right\} + \phi_2 E\left( z_n | z_n, z_{n-1}, L \right) - \mu \right\} + E\left[ a_{n+3} | z_n, z_{n-1}, L \right] \]

\[ = \mu + \phi_1 \left( z_n(2) - \mu \right) + \phi_2 \left( z_n(1) - \mu \right) \]

\[ 1 = 4: z_n(4) = \mu + \phi_1 E\left( z_n | z_n, z_{n-1}, L \right) - \mu \right\} + \phi_2 E\left( z_n | z_n, z_{n-1}, L \right) - \mu \right\} + E\left[ a_{n+4} | z_n, z_{n-1}, L \right] \]

\[ = \mu + \phi_1 \left( z_n(3) - \mu \right) + \phi_2 \left( z_n(2) - \mu \right) \]

The general form is

\[ z_n(1) = \mu + \phi_1 \left( z_n(1 - 1) - \mu \right) + \phi_2 \left( z_n(1 - 2) - \mu \right), 1 \geq 1 \]

4.4.5. MA (1) Model Forecast

\[ Z_t = \mu + a_t - \theta_t a_{t-1}, \quad a_t \sim WN(0, \sigma^2) \]

By replacing \( t \) with \( t+1 \) then taking expectation, the equation will be
Developing Forecasting Models for UNRWA Schools in Gaza Strip  Chapter 4  Model Inference

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

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

\[ z_n(1) = \mu + E(a_{n+1} | z_n, z_{n-1}, L) - \theta_1 E(a_{n+1-1} | z_n, z_{n-1}, L), \quad 1 \geq 1 \]
\[ 1 = 1: \quad z_n(1) = \mu + E(a_{n+1} | z_n, z_{n-1}, L) - \theta_1 E(a_n | z_n, z_{n-1}, L) \]
\[ = \mu - \theta_1 a_n \]
\[ 1 = 2: \quad z_n(2) = \mu + E(a_{n+2} | z_n, z_{n-1}, L) - \theta_1 E(a_{n+1} | z_n, z_{n-1}, L) \]
\[ = \mu \]
\[ 1 = 3: \quad z_n(3) = \mu + E(a_{n+3} | z_n, z_{n-1}, L) - \theta_1 E(a_{n+2} | z_n, z_{n-1}, L) \]
\[ = \mu \]

So, the general form can be written as

\[ z_n(1) = \mu, \quad 1 \geq 2 \]

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

\[ z_n(1) = \begin{cases} 
\mu - \theta_1 a_n, & 1 = 1 \\
\mu, & 1 \geq 2 
\end{cases} \]

4.4.6. MA (2) Model Forecast

\[ Z_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}, \quad a_t \equiv WN(0, \sigma^2) \]

\[ z_n(1) = E(z_{n+1} | z_n, z_{n-1}, L), \quad 1 \geq 1 \]
\[ = \mu + E(a_{n+1} | z_n, z_{n-1}, L) - \theta_1 E(a_{n+1-1} | z_n, z_{n-1}, L) - \theta_2 E(a_{n+1-2} | z_n, z_{n-1}, L), \quad 1 \geq 1 \]

This equation can be solved by iterating and applying condition expectation rules.
Developing Forecasting Models for UNRWA Schools in Gaza Strip ........... Chapter 4 Model Inference

\[ z_n(1) = \mu + E(a_{n+1} \mid z_n, z_{n-1}, L) - \theta_1 E(a_{n+1} \mid z_n, z_{n-1}, L) - \theta_2 E(a_{n+1} \mid z_n, z_{n-1}, L), \quad 1 \geq 1 \]

1 = 1: \[ z_n(1) = \mu + E(a_{n+1} \mid z_n, z_{n-1}, L) - \theta_1 E(a_n \mid z_n, z_{n-1}, L) - \theta_2 E(a_n \mid z_n, z_{n-1}, L) \]

= \mu - \theta_1 a_n - \theta_2 a_{n-1}

1 = 2: \[ z_n(2) = \mu + E(a_{n+2} \mid z_n, z_{n-1}, L) - \theta_1 E(a_{n+1} \mid z_n, z_{n-1}, L) - \theta_2 E(a_n \mid z_n, z_{n-1}, L) \]

= \mu - \theta_2 a_n

1 = 3: \[ z_n(3) = \mu + E(a_{n+3} \mid z_n, z_{n-1}, L) - \theta_1 E(a_{n+2} \mid z_n, z_{n-1}, L) - \theta_2 E(a_{n+1} \mid z_n, z_{n-1}, L) \]

= \mu

So, that

\[ z_n(1) = \mu, \quad 1 \geq 3 \]

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

\[
\begin{cases}
\mu - \theta_1 a_n - \theta_2 a_{n-1}, & 1 = 1 \\
\mu - \theta_2 a_n, & 1 = 2 \\
\mu, & 1 \geq 3
\end{cases}
\]

**4.4.7. ARMA (p,q)**

The general model for ARMA (p,q)

\[
z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + L + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - L - \theta_q a_{t-q}
\]

**The forecast model can be written as**

\[
Z_n(\lambda) = \phi_1 Z_n(\lambda - 1) + \phi_2 Z_n(\lambda - 2) + \Lambda + \phi_p Z_n(\lambda - p) + \theta_0 - \theta_1 E(e_{n+\lambda-1} \mid Z_1, Z_2, \Lambda, Z_n)
\]

\[- \theta_2 E(e_{n+\lambda-2} \mid Z_1, Z_2, \Lambda, Z_n) - \Lambda - \theta_q E(e_{n+\lambda-q} \mid Z_1, Z_2, \Lambda, Z_n) - \theta_1 E(e_{n+\lambda-1} \mid Z_1, Z_2, \Lambda, Z_n) - \theta_2 E(e_{n+\lambda-2} \mid Z_1, Z_2, \Lambda, Z_n) - \Lambda - \theta_q E(e_{n+\lambda-q} \mid Z_1, Z_2, \Lambda, Z_n)
\]

Where \( E(e_{n+j} \mid z_1, z_2, K, z_n) = \begin{cases} 0 \text{ for } j > 0 \\ e_{n+j} \text{ for } j \leq 0 \end{cases} \)

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

\[ Z_n(\lambda) = \phi_1 Z_n(\lambda - 1) + \phi_2 Z_n(\lambda - 2) + \Lambda + \phi_p Z_n(\lambda - p) + \theta_0 \text{ for } \lambda > q \]
As \( \theta_0 = \mu (1 - \phi_1 - \phi_2 - \lambda - \phi_p) \), the equation can be written in terms of deviation from \( \mu \) as

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

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

### 4.4.8. ARMA(1,1) Model Forecast

\[
Z_i = \mu + \phi (Z_{i-1} - \mu) + a_i - \theta, a_{i-1}, \phi \neq \theta, \phi_i < 1
\]

\[
z_n (1) = \mu + \phi E[(z_{n+1} - \mu)|z_n, z_{n-1}, \Omega], \quad 1 \geq 1
\]

By iteration, the equation can be solved as follow

\[
z_n (1) = \mu + \phi E[(z_{n+1} - \mu)|z_n, z_{n-1}, \Omega] + E(a_{n+1}|z_n, z_{n-1}, \Omega) - \theta E(a_{n+1}|z_n, z_{n-1}, \Omega), \quad 1 \geq 1
\]

1 = 1: \[
z_n (1) = \mu + \phi E[(z_{n+1} - \mu)|z_n, z_{n-1}, \Omega] + E(a_{n+1}|z_n, z_{n-1}, \Omega) - \theta E(a_{n+1}|z_n, z_{n-1}, \Omega)
\]

1 = 2: \[
z_n (2) = \mu + \phi E[(z_{n+2} - \mu)|z_n, z_{n-1}, \Omega] + E(a_{n+2}|z_n, z_{n-1}, \Omega) - \theta E(a_{n+2}|z_n, z_{n-1}, \Omega)
\]

1 = 3: \[
z_n (3) = \mu + \phi E[(z_{n+3} - \mu)|z_n, z_{n-1}, \Omega] + E(a_{n+3}|z_n, z_{n-1}, \Omega) - \theta E(a_{n+3}|z_n, z_{n-1}, \Omega)
\]

Where the general form is

\[
z_n (1) = \mu + \phi E[(z_{n+1} - \mu)|z_n, z_{n-1}, \Omega], \quad 1 \geq 2
\]

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

\[
z_n (1) = \begin{cases} 
\mu + \phi E[(z_{n+1} - \mu)|z_n, z_{n-1}, \Omega], & 1 = 1 \\
\mu + \phi E[(z_{n+1} - \mu)|z_n, z_{n-1}, \Omega], & 1 \geq 2
\end{cases}
\]

### 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

\[ Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \Lambda + \phi_p Z_{t-p} + \phi_{p+1} Z_{t-p-1} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \Lambda - \theta_q a_{t-q} \]

Where a coefficient \( \phi \) equals

\[
\phi_i = 1 + \phi_i
\]

\[
\phi_j = \phi_j - \phi_{j-1} \text{ for } j=1,2,\ldots,p
\]

and

\[
\phi_{p+1} = -\phi_p
\]

Thus, ARMA forecast equation can be used to forecast ARIMA models by replacing \( p \) with \( p+d \) and \( \phi_j \) by \( \varphi_j \).

### 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

\[ Z_n(1) = (1 + \phi)Z_t - \phi Z_{t-1} + \theta_0 - \theta a_t \]

\[ Z_n(2) = (1 + \phi)Z_n(1) - \phi Z_n + \theta_0 \]

\[ \vdots \]

\[ Z_n(\lambda) = (1 + \phi)Z_n(\lambda-1) - \phi Z_n(\lambda-2) + \theta_0 \]

Forecast error

\[ e_n(\lambda) = e_{n+\lambda} + \psi_1 e_{n+\lambda-1} + \psi_2 e_{n+\lambda-2} + \Lambda + \psi_{\lambda-1} e_{n+1} \text{ for } \lambda \geq 1 \]

So that

\[ E(e_n(\lambda)) = 0 \text{ for } \lambda \geq 1 \]

and

\[ Var(e_n(\lambda)) = \sigma^2 \sum_{j=1}^{\lambda-1} \psi^2 \text{ for } \lambda \geq 1 \]

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

Thus

\[ \psi_j = 1 - \theta, \ j \geq 1 \text{ for } IMA(1,1) \]
\[ \psi_j = 1 - \theta_2 + (1 - \theta_1 - \theta_2) j, \quad j \geq 1 \] for IMA(2,2)
\[ \psi_j = (1 - \phi^{j+1})/(1 - \phi), \quad j \geq 1 \] for ARI(1,1)

Thus, for nonstationary model the forecast error variance will increase without bound as the time \( \lambda \) increase. 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 \( z_n(l), 1 \geq 1 \) gives point forecast for time units \( \lambda \) which is not applicable or usefulness in statistical decision making where
\[
P(Z_{n+m} = z_n(m)) = 0, \text{ for some } m > 0
\]

This expression means that the forecast probability for any time unit \( \lambda \) 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 \( \lambda \) will surely included in an interval for example \([a,b]\).

Thus
\[
P(a \leq Z_{n+m} \leq b) = (1 - \alpha)
\]

this expression says that it is \( 100 \times (1 - \alpha) \)% confident that the interval \([a,b]\) includes \( Z_n(\lambda)\) value, so if \( \alpha = 0.05 \), it is 95% confident that the interval \([a,b]\) includes \( Z_n(\lambda)\) value.
Chapter Five: Results and Analysis

5.1. Scatter Time Plot ........................................................................................................56
5.2. Model Identification ....................................................................................................56
5.3. Applying ARIMA Models ............................................................................................58
5.4. Model Fitting Criteria .................................................................................................62
5.5. Fitting The Likely Model ............................................................................................62
5.6. Diagnostic Check on The Mean of The Best Fit Model ..............................................63
5.7. Diagnostic Check on Residuals Randomness ..............................................................63
5.8. Diagnostic Check on ACF and PACF of Residuals ......................................................64
5.9. Diagnostic Check on Residuals Normality ..................................................................65
5.10. Ten-Step Ahead Forecasts .........................................................................................66
5.11. Proposed Model .........................................................................................................68
5.12. Summary of Proposed Models Properties ..................................................................68
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

<table>
<thead>
<tr>
<th>Lag</th>
<th>ACF</th>
<th>T</th>
<th>LBQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.737522</td>
<td>2.76</td>
<td>9.37</td>
</tr>
<tr>
<td>2</td>
<td>0.487170</td>
<td>1.26</td>
<td>13.80</td>
</tr>
<tr>
<td>3</td>
<td>0.305319</td>
<td>0.71</td>
<td>15.70</td>
</tr>
<tr>
<td>4</td>
<td>0.151160</td>
<td>0.34</td>
<td>16.21</td>
</tr>
</tbody>
</table>

Figure 5.3 Partial Autocorrelation Function for Elementary classes in Gaza Field

<table>
<thead>
<tr>
<th>Lag</th>
<th>PACF</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.737522</td>
<td>2.76</td>
</tr>
</tbody>
</table>
Developing Forecasting Models for UNRWA Schools in Gaza Strip  Chapter 5  Analysis and Results

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

<table>
<thead>
<tr>
<th>Type</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA 1</td>
<td>-0.3546</td>
<td>0.2817</td>
<td>-1.26</td>
<td>0.234</td>
</tr>
<tr>
<td>Constant</td>
<td>112.69</td>
<td>19.13</td>
<td>5.89</td>
<td>0.000</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Lag</th>
<th>Chi-Square</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>24</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>36</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>48</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.410</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Forecasts from period 14

<table>
<thead>
<tr>
<th>Period</th>
<th>Forecast</th>
<th>Lower</th>
<th>Upper</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>3793.41</td>
<td>3693.35</td>
<td>3893.48</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>3906.10</td>
<td>3737.62</td>
<td>4074.58</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>4018.80</td>
<td>3802.56</td>
<td>4235.03</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>4131.49</td>
<td>3876.28</td>
<td>4386.70</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>4244.18</td>
<td>3955.20</td>
<td>4533.15</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>4356.87</td>
<td>4037.68</td>
<td>4676.05</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>4469.56</td>
<td>4122.79</td>
<td>4816.33</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>4582.25</td>
<td>4209.93</td>
<td>4954.58</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>4694.94</td>
<td>4298.71</td>
<td>5091.17</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>4807.63</td>
<td>4388.86</td>
<td>5226.41</td>
<td></td>
</tr>
</tbody>
</table>

ARIMA Model: Elementary Classes Gaza Field 0 2 1
Developing Forecasting Models for UNRWA Schools in Gaza Strip  Chapter 5  Analysis and Results

Final Estimates of Parameters

<table>
<thead>
<tr>
<th>Type</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA</td>
<td>0.3117</td>
<td>0.2996</td>
<td>1.04</td>
<td>0.323</td>
</tr>
<tr>
<td>Constant</td>
<td>7.43</td>
<td>10.95</td>
<td>0.68</td>
<td>0.513</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Lag</th>
<th>Chi-Square</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>24</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>36</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>48</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Forecasts from period 14

<table>
<thead>
<tr>
<th>Period</th>
<th>Forecast</th>
<th>Lower</th>
<th>Upper</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>3866.16</td>
<td>3758.54</td>
<td>3973.77</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>4066.74</td>
<td>3855.57</td>
<td>4277.91</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>4274.75</td>
<td>3943.07</td>
<td>4606.42</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>4490.18</td>
<td>4022.42</td>
<td>4957.94</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>4713.04</td>
<td>4095.03</td>
<td>5331.06</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>4943.33</td>
<td>4162.04</td>
<td>5724.62</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>5181.04</td>
<td>4224.39</td>
<td>6137.69</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>5426.18</td>
<td>4282.84</td>
<td>6569.52</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>5678.74</td>
<td>4338.03</td>
<td>7019.46</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>5938.74</td>
<td>4390.50</td>
<td>7486.97</td>
<td></td>
</tr>
</tbody>
</table>

ARIMA Model: Elementary Classes Gaza Field 1 1 0

Final Estimates of Parameters

<table>
<thead>
<tr>
<th>Type</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR</td>
<td>0.5027</td>
<td>0.2827</td>
<td>1.78</td>
<td>0.103</td>
</tr>
<tr>
<td>Constant</td>
<td>57.47</td>
<td>13.86</td>
<td>4.15</td>
<td>0.002</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Lag</th>
<th>Chi-Square</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>12.0</td>
<td>10</td>
<td>0.286</td>
</tr>
<tr>
<td>24</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>36</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>48</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Forecasts from period 14

<table>
<thead>
<tr>
<th>Period</th>
<th>Forecast</th>
<th>Lower</th>
<th>Upper</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Developing Forecasting Models for UNRWA Schools in Gaza Strip  Chapter 5  Analysis and Results

<table>
<thead>
<tr>
<th>Period</th>
<th>Forecast</th>
<th>Lower</th>
<th>Upper</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>3818.92</td>
<td>3721.75</td>
<td>3915.13</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>3949.02</td>
<td>3774.49</td>
<td>4123.54</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>4072.12</td>
<td>3828.68</td>
<td>4315.56</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>4191.47</td>
<td>3887.52</td>
<td>4495.43</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>4308.94</td>
<td>3951.44</td>
<td>4666.43</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>4425.45</td>
<td>4020.00</td>
<td>4830.91</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>4541.49</td>
<td>4092.52</td>
<td>4990.47</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>4657.29</td>
<td>4168.34</td>
<td>5146.24</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>4772.97</td>
<td>4246.93</td>
<td>5299.01</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>4888.59</td>
<td>4327.84</td>
<td>5449.33</td>
<td></td>
</tr>
</tbody>
</table>

ARIMA Model: Elementary Classes Gaza Field 1 1 1

Final Estimates of Parameters

<table>
<thead>
<tr>
<th>Type</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR 1</td>
<td>0.5098</td>
<td>0.9039</td>
<td>0.56</td>
<td>0.585</td>
</tr>
<tr>
<td>MA 1</td>
<td>0.0082</td>
<td>0.9633</td>
<td>0.01</td>
<td>0.993</td>
</tr>
<tr>
<td>Constant</td>
<td>56.70</td>
<td>16.41</td>
<td>3.45</td>
<td>0.006</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Lag</th>
<th>Chi-Square</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>*</td>
<td>9</td>
<td>*</td>
</tr>
<tr>
<td>24</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>36</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>48</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Forecasts from period 14

<table>
<thead>
<tr>
<th>Period</th>
<th>Forecast</th>
<th>Lower</th>
<th>Upper</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>3818.92</td>
<td>3721.75</td>
<td>3915.13</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>3949.02</td>
<td>3774.49</td>
<td>4123.54</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>4072.12</td>
<td>3828.68</td>
<td>4315.56</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>4191.47</td>
<td>3887.52</td>
<td>4495.43</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>4308.94</td>
<td>3951.44</td>
<td>4666.43</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>4425.45</td>
<td>4020.00</td>
<td>4830.91</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>4541.49</td>
<td>4092.52</td>
<td>4990.47</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>4657.29</td>
<td>4168.34</td>
<td>5146.24</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>4772.97</td>
<td>4246.93</td>
<td>5299.01</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>4888.59</td>
<td>4327.84</td>
<td>5449.33</td>
<td></td>
</tr>
</tbody>
</table>

ARIMA Model: Elementary Classes Gaza Field 1 2 0

Final Estimates of Parameters

<table>
<thead>
<tr>
<th>Type</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR 1</td>
<td>-0.4504</td>
<td>0.3061</td>
<td>-1.47</td>
<td>0.172</td>
</tr>
<tr>
<td>Constant</td>
<td>11.21</td>
<td>15.29</td>
<td>0.73</td>
<td>0.480</td>
</tr>
</tbody>
</table>

Differencing: 2 regular differences

Number of observations: Original series 14, after differencing 12

Residuals: SS = 27928.9 (backforecasts excluded)
MS = 2792.9 DF = 10

---

60
Developing Forecasting Models for UNRWA Schools in Gaza Strip  Chapter 5 Analysis and Results

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

<table>
<thead>
<tr>
<th>Lag</th>
<th>Chi-Square</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>24</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>36</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>48</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Forecasts from period 14

<table>
<thead>
<tr>
<th>Period</th>
<th>Forecast</th>
<th>Lower</th>
<th>Upper</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>3893.57</td>
<td>3801.24</td>
<td>3985.91</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>4093.47</td>
<td>3902.40</td>
<td>4284.54</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>4314.61</td>
<td>4009.02</td>
<td>4620.21</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>4540.71</td>
<td>4107.54</td>
<td>4973.87</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>4775.78</td>
<td>4199.73</td>
<td>5351.84</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>5018.03</td>
<td>4286.69</td>
<td>5749.37</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>5268.26</td>
<td>4369.57</td>
<td>6166.95</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>5526.11</td>
<td>4449.08</td>
<td>6603.13</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>5791.74</td>
<td>4525.91</td>
<td>7057.56</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>6065.08</td>
<td>4600.59</td>
<td>7529.56</td>
<td></td>
</tr>
</tbody>
</table>

ARIMA Model: Elementary Classes Gaza Field 1 2 1

Final Estimates of Parameters

<table>
<thead>
<tr>
<th>Type</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR</td>
<td>-0.9994</td>
<td>0.0303</td>
<td>-33.01</td>
<td>0.000</td>
</tr>
<tr>
<td>MA</td>
<td>-0.9079</td>
<td>0.1980</td>
<td>-4.59</td>
<td>0.001</td>
</tr>
<tr>
<td>Constant</td>
<td>13.18</td>
<td>25.95</td>
<td>0.51</td>
<td>0.624</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Lag</th>
<th>Chi-Square</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>24</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>36</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>48</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Forecasts from period 14

<table>
<thead>
<tr>
<th>Period</th>
<th>Forecast</th>
<th>Lower</th>
<th>Upper</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>3893.57</td>
<td>3801.24</td>
<td>3985.91</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>4093.47</td>
<td>3902.40</td>
<td>4284.54</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>4314.61</td>
<td>4009.02</td>
<td>4620.21</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>4540.71</td>
<td>4107.54</td>
<td>4973.87</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>4775.78</td>
<td>4199.73</td>
<td>5351.84</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>5018.03</td>
<td>4286.69</td>
<td>5749.37</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>5268.26</td>
<td>4369.57</td>
<td>6166.95</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>5526.11</td>
<td>4449.08</td>
<td>6603.13</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>5791.74</td>
<td>4525.91</td>
<td>7057.56</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>6065.08</td>
<td>4600.59</td>
<td>7529.56</td>
<td></td>
</tr>
</tbody>
</table>
5.4. Model Fitting Criteria

to fit the best model, different information criteria was used, Akaike’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**

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

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

<table>
<thead>
<tr>
<th>type</th>
<th>coefficient</th>
<th>value</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR 1</td>
<td>$\hat{\phi}$</td>
<td>$-0.9994$</td>
<td>-33.01</td>
<td>0.000</td>
</tr>
<tr>
<td>MA 1</td>
<td>$\hat{\theta}$</td>
<td>$-0.9079$</td>
<td>-4.59</td>
<td>0.001</td>
</tr>
</tbody>
</table>

So, it was concluded that (1 2 1) model is the best fit model to represent 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

$H_0: \mu = 0 \ \text{versus} \ \ H_1: \mu \neq 0$

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

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
<th>SE Mean</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>0.8</td>
<td>42.6</td>
<td>12.3</td>
<td>0.06</td>
<td>0.950</td>
</tr>
</tbody>
</table>

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

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.

$H_0$: 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 $\alpha = 0.05$, and show evidence that the model has not explained all autocorrelation in the data.

**Autocorrelation for model(1 2 1) residuals**

<table>
<thead>
<tr>
<th>Lag</th>
<th>ACF</th>
<th>T</th>
<th>LBQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.023285</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td>2</td>
<td>0.134194</td>
<td>0.46</td>
<td>0.31</td>
</tr>
<tr>
<td>3</td>
<td>-0.186103</td>
<td>-0.63</td>
<td>0.96</td>
</tr>
</tbody>
</table>

**Partial Autocorrelation for model(1 2 1) residuals**

<table>
<thead>
<tr>
<th>Lag</th>
<th>PACF</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.023285</td>
<td>0.08</td>
</tr>
<tr>
<td>2</td>
<td>0.133725</td>
<td>0.46</td>
</tr>
<tr>
<td>3</td>
<td>-0.195524</td>
<td>-0.68</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Period</th>
<th>Forecast</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>3893.57</td>
<td>3801.24</td>
<td>3985.91</td>
</tr>
<tr>
<td>2011</td>
<td>4081.78</td>
<td>3882.84</td>
<td>4280.73</td>
</tr>
<tr>
<td>2012</td>
<td>4315.51</td>
<td>3981.30</td>
<td>4649.73</td>
</tr>
<tr>
<td>2013</td>
<td>4516.93</td>
<td>4031.22</td>
<td>5002.64</td>
</tr>
<tr>
<td>2014</td>
<td>4763.83</td>
<td>4105.24</td>
<td>5422.41</td>
</tr>
<tr>
<td>2015</td>
<td>4978.45</td>
<td>4133.93</td>
<td>5822.97</td>
</tr>
<tr>
<td>2016</td>
<td>5238.51</td>
<td>4190.26</td>
<td>6286.76</td>
</tr>
<tr>
<td>2017</td>
<td>5466.34</td>
<td>4203.13</td>
<td>6729.54</td>
</tr>
<tr>
<td>2018</td>
<td>5739.55</td>
<td>4245.83</td>
<td>7233.28</td>
</tr>
<tr>
<td>2019</td>
<td>5980.59</td>
<td>4246.36</td>
<td>7714.82</td>
</tr>
</tbody>
</table>

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,

\[ \phi_1(B)z_t = \delta + \theta_1(B)a_t \]
\[ (1 - \phi_1B)z_t = \delta + (1 - \theta_1B)a_t \]
\[ z_t - \phi_1z_{t-1} = \delta + a_t - \theta_1a_{t-1} \]
\[ z_t = \delta + \phi_1z_{t-1} + a_t - \theta_1a_{t-1}, a_t : WN(0, \sigma^2), \phi_1 \neq \theta_1 \]

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

\[ Z_t = 13.18 - 0.9994Z_{t-1} + a_t + 0.9079a_{t-1} \]

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>
<thead>
<tr>
<th>Area</th>
<th>Best Model</th>
<th>Parameters</th>
<th>Forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rafah</td>
<td>1 1 0</td>
<td>AR(1)=-0.2330, Constant= 1.2514</td>
<td>41</td>
</tr>
<tr>
<td>Khan</td>
<td>2 2 1</td>
<td>AR(1)=-1.1356, AR(2)=-0.8833, MA(1)= 0.8892, Constant= 0.59728</td>
<td>30</td>
</tr>
<tr>
<td>Khan\E</td>
<td>0 2 1</td>
<td>MA(1)=0.9136, Constant= 0.06662</td>
<td>19</td>
</tr>
<tr>
<td>Middle</td>
<td>0 1 1</td>
<td>MA(1)=-0.5610, Constant=1.4014</td>
<td>52</td>
</tr>
<tr>
<td>Gaza</td>
<td>1 2 1</td>
<td>AR(1)=-0.4411, MA(1)= 1.2096, Constant= 0.61092</td>
<td>65</td>
</tr>
<tr>
<td>North</td>
<td>2 1 1</td>
<td>AR(1)=0.7376, AR(2)=-0.8679, MA(1)= 0.9503, Constant= 0.676638</td>
<td>38</td>
</tr>
<tr>
<td>Field</td>
<td>1 2 0</td>
<td>AR(1)=-0.7746, Constant= 0.936</td>
<td>238</td>
</tr>
</tbody>
</table>
### Table 5.4 Summary for number of elementary classroom models

<table>
<thead>
<tr>
<th>Area</th>
<th>Best Model</th>
<th>Parameters</th>
<th>Forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AR(1)= -0.9004, constant=30.347</td>
<td>578 619 612 649 646 679 680 709 713 740</td>
</tr>
<tr>
<td>Rafah</td>
<td>1 1 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khan</td>
<td>0 2 1</td>
<td>MA(1)=0.8272, constant=0.1127</td>
<td>348 358 368 379 389 400 411 422 432 444</td>
</tr>
<tr>
<td>Khan\E</td>
<td>0 1 1</td>
<td>MA(1)= -0.8756, constant= 12.441</td>
<td>368 380 393 405 418 430 443 455 468 480</td>
</tr>
<tr>
<td>Middle</td>
<td>1 2 0</td>
<td>AR(1)= -0.8820, constant=1.319</td>
<td>812 835 882 909 955 985 1031 1064 1109 1146</td>
</tr>
<tr>
<td>Gaza</td>
<td>1 1 2</td>
<td>AR(1)=1.2805, MA(1)= -0.0475, MA(2)= -0.8414, constant= 18.108</td>
<td>993 1023 1054 1087 1120 1153 1187 1220 1253 1287</td>
</tr>
<tr>
<td>North</td>
<td>0 2 1</td>
<td>MA(1)=0.8310, constant=-1.1022</td>
<td>701 713 725 735 744 752 759 765 770 774</td>
</tr>
<tr>
<td>Field</td>
<td>1 2 1</td>
<td>AR(1)= -0.9994, MA(1)=0.9079, constant=13.18</td>
<td>3894 4082 4316 4517 4764 4978 5239 5466 5740 5981</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5.5 Summary for number of boy preparatory classroom models

<table>
<thead>
<tr>
<th>Area</th>
<th>Best Model</th>
<th>Parameters</th>
<th>Forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rafah</td>
<td>0 1 1</td>
<td>MA(1)=1.2789, constant=9.981</td>
<td>189</td>
</tr>
<tr>
<td>Khan</td>
<td>1 1 1</td>
<td>AR(1)=-0.5963, MA(1)=1.1685, constant=7.04158</td>
<td>93</td>
</tr>
<tr>
<td>Khan\E</td>
<td>0 1 1</td>
<td>MA(1)=0.8822, constant=3.3575</td>
<td>72</td>
</tr>
<tr>
<td>Middle</td>
<td>0 1 1</td>
<td>MA(1)=0.9016, constant=11.3704</td>
<td>228</td>
</tr>
<tr>
<td>Gaza</td>
<td>2 2 0</td>
<td>AR(1)=-1.3560, AR(2)=-1.0334, constant=7.101</td>
<td>315</td>
</tr>
<tr>
<td>North</td>
<td>1 1 1</td>
<td>AR(1)=-0.3814, MA(1)=0.9030, constant=16.0508</td>
<td>214</td>
</tr>
<tr>
<td>Field</td>
<td>1 1 0</td>
<td>AR(1)=-0.6826, constant=85.59</td>
<td>1047</td>
</tr>
</tbody>
</table>
### Developing Forecasting Models for UNRWA Schools in Gaza Strip — Chapter 5: Analysis and Results

#### Table 5.6 Summary for number of girl preparatory classroom models

<table>
<thead>
<tr>
<th>Area</th>
<th>Best Model</th>
<th>Parameters</th>
<th>Forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rafah</td>
<td>1 1 0</td>
<td>AR(1)=-0.6760, constant= 9.653</td>
<td>143</td>
</tr>
<tr>
<td>Khan</td>
<td>0 1 1</td>
<td>MA(1)=0.8695, constant= 3.5579</td>
<td>91</td>
</tr>
<tr>
<td>Khan\E</td>
<td>0 1 1</td>
<td>MA(1)=0.8727, constant= 2.8394</td>
<td>65</td>
</tr>
<tr>
<td>Middle</td>
<td>1 1 0</td>
<td>AR(1)=0.4473, constant=3.419</td>
<td>7399</td>
</tr>
<tr>
<td>Gaza</td>
<td>0 2 2</td>
<td>MA(1)=0.7929, MA(2)= -0.8824, constant=0.783</td>
<td>182</td>
</tr>
<tr>
<td>North</td>
<td>1 2 1</td>
<td>AR(1)= -1.0006, MA(1)=-0.9141, constant=-0.378</td>
<td>153</td>
</tr>
<tr>
<td>Field</td>
<td>0 1 1</td>
<td>MA(1)= 0.8570, constant=37.440</td>
<td>856</td>
</tr>
</tbody>
</table>
### Table 5.7 Summary for number of elementary pupil models

<table>
<thead>
<tr>
<th>Area</th>
<th>Best Model</th>
<th>Parameters</th>
<th>Forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rafah</td>
<td>1 2 0</td>
<td>AR(1)= -0.6610</td>
<td>24613</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Constant= -35.1</td>
<td></td>
</tr>
<tr>
<td>Khan</td>
<td>0 1 2</td>
<td>MA(1)= -0.4657</td>
<td>13997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MA(2)= -0.8958</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Constant= 223.5</td>
<td></td>
</tr>
<tr>
<td>Khan\E</td>
<td>0 2 1</td>
<td>MA(1)=0.8440</td>
<td>14609</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Constant=19.86</td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>0 2 2</td>
<td>MA(1)=0.4153</td>
<td>30073</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MA(2)=0.8036</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Constant= -77.2</td>
<td></td>
</tr>
<tr>
<td>Gaza</td>
<td>2 2 2</td>
<td>AR(1)= 0.0694</td>
<td>43439</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AR(2)= 0.8300</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MA(1)= 0.3895</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MA(2)= -1.0661</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Constant= 97.35</td>
<td></td>
</tr>
<tr>
<td>North</td>
<td>0 1 2</td>
<td>MA(1)= -0.6509</td>
<td>27951</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MA(2)= -1.1034</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Constant=855.3</td>
<td></td>
</tr>
<tr>
<td>Field</td>
<td>0 1 2</td>
<td>MA(1)= -0.4967</td>
<td>153274</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MA(2)= -0.9117</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Constant=3780</td>
<td></td>
</tr>
</tbody>
</table>
### Developing Forecasting Models for UNRWA Schools in Gaza Strip - Chapter 5. Analysis and Results

**Table 5.8 Summary for number of boy preparatory pupil models**

<table>
<thead>
<tr>
<th>Area</th>
<th>Best Model</th>
<th>Parameters</th>
<th>Forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rafah</td>
<td>1 1 0</td>
<td>AR(1)=0.7266, constant=41.28</td>
<td>5462</td>
</tr>
<tr>
<td>Khan</td>
<td>1 1 1</td>
<td>AR(1)=-1.0048, MA(1)=-0.8653, constant=160.59</td>
<td>3008</td>
</tr>
<tr>
<td>Khan\E</td>
<td>1 1 0</td>
<td>AR(1)=-0.6767, constant=87.88</td>
<td>2037</td>
</tr>
<tr>
<td>Middle</td>
<td>2 1 1</td>
<td>AR(1)=-0.2366, AR(2)=0.7637, MA(1)=0.8467, constant=54.1</td>
<td>6558</td>
</tr>
<tr>
<td>Gaza</td>
<td>2 1 2</td>
<td>AR(1)=1.7473, AR(2)=-0.9994, MA(1)=1.6022, MA(2)=0.6705, constant=71.557</td>
<td>7813</td>
</tr>
<tr>
<td>North</td>
<td>1 1 0</td>
<td>AR(1)=0.8467, constant=21.36</td>
<td>5986</td>
</tr>
<tr>
<td>Field</td>
<td>1 1 0</td>
<td>AR(1)=0.5181, constant=471.6</td>
<td>31938</td>
</tr>
</tbody>
</table>
### Developing Forecasting Models for UNRWA Schools in Gaza Strip Chapter5 Analysis and Results

Table 5.9 Summary for number of girl preparatory pupil models

<table>
<thead>
<tr>
<th>Area</th>
<th>Best Model</th>
<th>Parameters</th>
<th>Forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rafah</td>
<td>0 1 1</td>
<td>MA(1)= 0.8744, constant= 250.10</td>
<td>6253 6503 6753 7003 7253 7504 7754 8004 8254 8504</td>
</tr>
<tr>
<td>Khan</td>
<td>0 1 1</td>
<td>MA(1)= 0.5284, constant=106.87</td>
<td>3368 3475 3582 3688 3795 3902 4009 4116 4223 4330</td>
</tr>
<tr>
<td>Khan\E</td>
<td>0 1 1</td>
<td>MA(1)=0.8852, constant=82.85</td>
<td>2555 2638 2720 2803 2886 2969 3052 3135 3217 3300</td>
</tr>
<tr>
<td>Middle</td>
<td>1 1 2</td>
<td>AR(1)= 0.8441, MA(1)= -0.8467, MA(2)= -0.9860</td>
<td>7334 7402 7491 7599 7721 7857 8003 8159 8322 8492</td>
</tr>
<tr>
<td></td>
<td></td>
<td>constant= 31.97</td>
<td></td>
</tr>
<tr>
<td>Gaza</td>
<td>1 1 1</td>
<td>AR(1)= -0.6083, MA(1)= 0.8562, constant= -61.63</td>
<td>7630 7874 8136 8395 8654 8913 9172 9432 9691 9950</td>
</tr>
<tr>
<td>North</td>
<td>1 1 2</td>
<td>AR(1)= 0.7747, MA(1)= 1.1702, MA(2)= -1.4066</td>
<td>6348 6394 6542 6770 7059 7397 7771 8174 8599 9042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>constant= 113.025</td>
<td></td>
</tr>
<tr>
<td>Field</td>
<td>1 1 0</td>
<td>AR(1)= -0.2420, constant=1488.6</td>
<td>32383 33550 34756 35953 37152 38350 39549 40747 41946 43145</td>
</tr>
</tbody>
</table>
Chapter Six: Conclusions and Recommendations

6.1. Summary of Results and Implication ........................................... 77
6.2. Conclusion .................................................................................. 79
6.3. Recommendations ...................................................................... 79
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/ARMA 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 non-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 overcome 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:


Articles:


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


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


---

83

**Arabic Articles**


[3] العاني، احمد، حسين، ARIMA.


**Arabic Electronic Books**


**Reports:**


**Internet Websites**


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

![Time Series Plot for schools of Khan](image)

(with forecasts and their 95% confidence limits)

A.1.1.4 Schools of Middle Governorate

![Time Series Plot for schools of Middle](image)

(with forecasts and their 95% confidence limits)
A.1.1.5 Schools of Northern Governorate

Time Series Plot for schools of Northern
(with forecasts and their 95% confidence limits)

A.1.1.6 Schools of Rafah Governorate

Time Series Plot for schools of Rafah
(with forecasts and their 95% confidence limits)
A.1.1.7 Schools of Khanyounis Eastern area

Time Series Plot for schools of Khan\E
(with forecasts and their 95% confidence limits)

Number of schools

Years

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

Time Series Plot for elementary classrooms of Khan\E
(with forecasts and their 95% confidence limits)
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

Time Series Plot for boy preparatory classrooms of Khanyounis Eastern area
(with forecasts and their 95% confidence limits)

Number of boy preparatory classrooms

Time

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

Time Series Plot for elementary pupils of Northern
(with forecasts and their 95% confidence limits)

A.1.5.6 Elementary pupils of Rafah Governorate

Time Series Plot for elementary pupils of Rafah
(with forecasts and their 95% confidence limits)
A.1.5.7 Elementary pupils of Khanyounis Eastern area

Time Series Plot for elementary pupils of Khan\E
(with forecasts and their 95% confidence limits)
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

Time Series Plot for boy preparatory pupils of Northern
(with forecasts and their 95% confidence limits)

A.1.6.6 Boy preparatory pupils of Rafah Governorate

Time Series Plot for boy preparatory pupils of Rafah
(with forecasts and their 95% confidence limits)
A.1.6.7 Boy preparatory pupils of Khanyounis Eastern area

Time Series Plot for boy preparatory pupils of Khan\'E
(with forecasts and their 95% confidence limits)
A.1.7 Time series plot for number of girl preparatory pupils

A.1.7.1 Girl preparatory pupils of Gaza Field

[Time Series Plot for girl preparatory pupils of Gaza Field]

(with forecasts and their 95% confidence limits)

A.1.7.2 Girl preparatory pupils of Gaza Governorate

[Time Series Plot for girl preparatory pupils of Gaza]

(with forecasts and their 95% confidence limits)
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

![Graph showing the number of girl preparatory pupils of Northern Governorate with forecasts and their 95% confidence limits.]

A.1.7.6 Girl preparatory pupils of Rafah Governorate

![Graph showing the number of girl preparatory pupils of Rafah Governorate with forecasts and their 95% confidence limits.]

[Diagram images are not provided in the text version.]

112
A.1.7.7 Girl preparatory pupils of Khanyounis Eastern area

Time Series Plot for girl preparatory pupils of Khanyounis Eastern area (with forecasts and their 95% confidence limits)
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:

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

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.

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

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

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

A.3.1.2 Gaza Governorate

<table>
<thead>
<tr>
<th>year</th>
<th>C S</th>
<th>C E</th>
<th>C P B</th>
<th>C P G</th>
<th>P E</th>
<th>P P B</th>
<th>P P G</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>41</td>
<td>528</td>
<td>88</td>
<td>87</td>
<td>25505</td>
<td>4168</td>
<td>4032</td>
</tr>
<tr>
<td>1997</td>
<td>45</td>
<td>546</td>
<td>89</td>
<td>85</td>
<td>27078</td>
<td>4300</td>
<td>4136</td>
</tr>
<tr>
<td>1998</td>
<td>43</td>
<td>574</td>
<td>95</td>
<td>90</td>
<td>28500</td>
<td>4552</td>
<td>4441</td>
</tr>
<tr>
<td>1999</td>
<td>43</td>
<td>594</td>
<td>100</td>
<td>99</td>
<td>29655</td>
<td>5000</td>
<td>4945</td>
</tr>
<tr>
<td>2000</td>
<td>43</td>
<td>611</td>
<td>113</td>
<td>110</td>
<td>30464</td>
<td>5622</td>
<td>5472</td>
</tr>
<tr>
<td>2001</td>
<td>43</td>
<td>624</td>
<td>125</td>
<td>120</td>
<td>30647</td>
<td>5989</td>
<td>5782</td>
</tr>
<tr>
<td>2002</td>
<td>43</td>
<td>654</td>
<td>137</td>
<td>132</td>
<td>30994</td>
<td>6489</td>
<td>6248</td>
</tr>
<tr>
<td>2003</td>
<td>44</td>
<td>669</td>
<td>146</td>
<td>144</td>
<td>30699</td>
<td>6909</td>
<td>6441</td>
</tr>
<tr>
<td>2004</td>
<td>44</td>
<td>680</td>
<td>162</td>
<td>147</td>
<td>30416</td>
<td>7457</td>
<td>6725</td>
</tr>
<tr>
<td>2005</td>
<td>46</td>
<td>683</td>
<td>178</td>
<td>158</td>
<td>29240</td>
<td>7822</td>
<td>6937</td>
</tr>
<tr>
<td>2006</td>
<td>47</td>
<td>709</td>
<td>177</td>
<td>158</td>
<td>30030</td>
<td>8013</td>
<td>7077</td>
</tr>
<tr>
<td>2007</td>
<td>53</td>
<td>764</td>
<td>230</td>
<td>160</td>
<td>30013</td>
<td>7919</td>
<td>7942</td>
</tr>
<tr>
<td>2008</td>
<td>57</td>
<td>855</td>
<td>239</td>
<td>170</td>
<td>33328</td>
<td>7662</td>
<td>6994</td>
</tr>
<tr>
<td>2009</td>
<td>60</td>
<td>936</td>
<td>257</td>
<td>172</td>
<td>36660</td>
<td>7690</td>
<td>7283</td>
</tr>
</tbody>
</table>
### A.3.1.3 Khanyounis Western area

<table>
<thead>
<tr>
<th>Year</th>
<th>CS</th>
<th>CE</th>
<th>CPB</th>
<th>CPG</th>
<th>PE</th>
<th>PPB</th>
<th>PPG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>19</td>
<td>227</td>
<td>40</td>
<td>38</td>
<td>11239</td>
<td>1903</td>
<td>1856</td>
</tr>
<tr>
<td>1997</td>
<td>18</td>
<td>239</td>
<td>39</td>
<td>39</td>
<td>11998</td>
<td>1974</td>
<td>1927</td>
</tr>
<tr>
<td>1998</td>
<td>18</td>
<td>251</td>
<td>42</td>
<td>42</td>
<td>12351</td>
<td>2075</td>
<td>2069</td>
</tr>
<tr>
<td>1999</td>
<td>18</td>
<td>257</td>
<td>46</td>
<td>46</td>
<td>12860</td>
<td>2264</td>
<td>2257</td>
</tr>
<tr>
<td>2000</td>
<td>18</td>
<td>269</td>
<td>50</td>
<td>50</td>
<td>13264</td>
<td>2497</td>
<td>2475</td>
</tr>
<tr>
<td>2001</td>
<td>18</td>
<td>277</td>
<td>55</td>
<td>54</td>
<td>13331</td>
<td>2623</td>
<td>2623</td>
</tr>
<tr>
<td>2002</td>
<td>19</td>
<td>284</td>
<td>59</td>
<td>60</td>
<td>13217</td>
<td>2822</td>
<td>2852</td>
</tr>
<tr>
<td>2003</td>
<td>19</td>
<td>284</td>
<td>66</td>
<td>67</td>
<td>12602</td>
<td>3053</td>
<td>3025</td>
</tr>
<tr>
<td>2004</td>
<td>20</td>
<td>286</td>
<td>65</td>
<td>67</td>
<td>12602</td>
<td>3053</td>
<td>3025</td>
</tr>
<tr>
<td>2005</td>
<td>20</td>
<td>285</td>
<td>71</td>
<td>68</td>
<td>12231</td>
<td>3164</td>
<td>3107</td>
</tr>
<tr>
<td>2006</td>
<td>20</td>
<td>295</td>
<td>68</td>
<td>70</td>
<td>12179</td>
<td>3077</td>
<td>3120</td>
</tr>
<tr>
<td>2007</td>
<td>25</td>
<td>308</td>
<td>110</td>
<td>82</td>
<td>11734</td>
<td>3308</td>
<td>3701</td>
</tr>
<tr>
<td>2008</td>
<td>24</td>
<td>335</td>
<td>84</td>
<td>86</td>
<td>12612</td>
<td>2982</td>
<td>3035</td>
</tr>
<tr>
<td>2009</td>
<td>26</td>
<td>338</td>
<td>101</td>
<td>72</td>
<td>13040</td>
<td>3008</td>
<td>3067</td>
</tr>
</tbody>
</table>

### A.3.1.4 Middle Governorate

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

116
### A.3.1.5 Northern Governorate

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

### A.3.1.6 Rafah Governorate

<table>
<thead>
<tr>
<th>Year</th>
<th>C S</th>
<th>C E</th>
<th>C P B</th>
<th>C P G</th>
<th>P E</th>
<th>P P B</th>
<th>P P G</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>27</td>
<td>371</td>
<td>66</td>
<td>58</td>
<td>17597</td>
<td>3064</td>
<td>2663</td>
</tr>
<tr>
<td>1997</td>
<td>28</td>
<td>380</td>
<td>65</td>
<td>62</td>
<td>19220</td>
<td>3268</td>
<td>2764</td>
</tr>
<tr>
<td>1998</td>
<td>28</td>
<td>410</td>
<td>73</td>
<td>61</td>
<td>20332</td>
<td>3501</td>
<td>3006</td>
</tr>
<tr>
<td>1999</td>
<td>28</td>
<td>423</td>
<td>76</td>
<td>71</td>
<td>21134</td>
<td>3745</td>
<td>3286</td>
</tr>
<tr>
<td>2000</td>
<td>28</td>
<td>447</td>
<td>82</td>
<td>80</td>
<td>22018</td>
<td>4065</td>
<td>3770</td>
</tr>
<tr>
<td>2001</td>
<td>29</td>
<td>470</td>
<td>90</td>
<td>88</td>
<td>22779</td>
<td>4268</td>
<td>4187</td>
</tr>
<tr>
<td>2002</td>
<td>29</td>
<td>481</td>
<td>96</td>
<td>97</td>
<td>22551</td>
<td>4545</td>
<td>4589</td>
</tr>
<tr>
<td>2003</td>
<td>30</td>
<td>490</td>
<td>107</td>
<td>110</td>
<td>22443</td>
<td>4902</td>
<td>4970</td>
</tr>
<tr>
<td>2004</td>
<td>31</td>
<td>507</td>
<td>115</td>
<td>114</td>
<td>21845</td>
<td>5201</td>
<td>5433</td>
</tr>
<tr>
<td>2005</td>
<td>33</td>
<td>518</td>
<td>121</td>
<td>117</td>
<td>21750</td>
<td>5610</td>
<td>5240</td>
</tr>
<tr>
<td>2006</td>
<td>33</td>
<td>532</td>
<td>129</td>
<td>118</td>
<td>21945</td>
<td>5669</td>
<td>5211</td>
</tr>
<tr>
<td>2007</td>
<td>40</td>
<td>549</td>
<td>193</td>
<td>115</td>
<td>20478</td>
<td>5709</td>
<td>6684</td>
</tr>
<tr>
<td>2008</td>
<td>40</td>
<td>543</td>
<td>167</td>
<td>137</td>
<td>22462</td>
<td>5540</td>
<td>5029</td>
</tr>
<tr>
<td>2009</td>
<td>40</td>
<td>590</td>
<td>181</td>
<td>125</td>
<td>23115</td>
<td>5471</td>
<td>5159</td>
</tr>
</tbody>
</table>
### A.3.1.7 Khanyounis Eastern area

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

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