YANGON UNIVERSITY OF ECONOMICS
DEPARTMENT OF STATISTICS

A SEASONAL TIME SERIES MODEL OF PRODUCTION TRANSFORMERS IN HITACHI SOE ELECTRIC AND MACHINERY CO., LTD (2013-2017)

BY
YU WAH HLAING
M.Econ (Statistics)
Roll No.3

May, 2018

YANGON UNIVERSITY OF ECONOMICS
DEPARTMENT OF STATISTICS

A SEASONAL TIME SERIES MODEL OF PRODUCTION TRANSFORMERS IN HITACHI SOE ELECTRIC AND MACHINERY CO., LTD (2013-2017)

This thesis is submitted as a partial fulfillment towards
the Degree of Master of Economics (Statistics)

BY
YU WAH HLAING
M.Econ (Statistics)
Roll No.3

May, 2018

YANGON UNIVERSITY OF ECONOMICS
DEPARTMENT OF STATISTICS

A SEASONAL TIME SERIES MODEL OF PRODUCTION TRANSFORMERS IN HITACHI SOE ELECTRIC AND MACHINERY CO., LTD (2013-2017)

This thesis is submitted to the Board of Examination as partial fulfillment of the requirement for degree of M.Econ (Statistics).

Approved by the Board of Examiners

Submitted by

Ma Yu Wah Hlaing
Roll No. 3
M.Econ (Statistics)

Supervised by

U Win Min Thant
Lecturer
Department of Statistics
Yangon University of Economics

May, 2018

YANGON UNIVERSITY OF ECONOMICS

DEPARTMENT OF STATISTICS

This is to certify that thesis entitled **"A SEASONAL TIME SERIES MODEL OF PRODUCTION TRANSFORMERS IN HITACHI SOE ELECTRIC AND MACHINERY CO., LTD (2013-2017)"**, submitted as a partial fulfillment towards the requirements of the degree of Master of Economics (Statistics) has been accepted by the Board of Examiners.

BOARD OF EXAMINERS

(Chairman)
Professor Dr. Tun Aung
Pro- Rector
Yangon University of Economics

(Chief Examiner)
Prof. Dr. Maw Maw Khin
Professor and Head
Department of Statistics
Yangon University of Economics

(External Examiner)
Dr. Than Pe
Lecturer (Retd.)
School of Computing and Mathematics Science
University of Western Sydney, Australia

(Examiner)
Prof. Dr. Mya Thandar
Professor
Department of Statistics
Yangon University of Economics

(External Examiner)
Daw Tin Tin Yu
Associate Professor (Retd.)
Department of Statistics
Yangon University of Economics

(External Examiner)
Daw Aye Aye Than
Associate Professor (Retd.)
Department of Statistics
Yangon University of Economics

May, 2018

ABSTRACT

Many economic time series exhibit seasonal behaviour. The estimation of seasonal variation is important problem in time series analysis. Consequently, seasonal variations are needed to determine and seasonal adjustments are needed in forecasting. In this thesis, the production series for transformers for the period of January 2013 to December 2017 are studied.

Stochastic models for monthly production series are found by using Box-Jenkins model building approach. Basic statistical characteristics for the production series are first investigated and statistical test for seasonality is applied to each series to confirm the existence of seasonality. Seasonal variation of the production series for transformers from January 2013 to December 2017 are measured by using Ratio to Moving Average Method. Suitable stochastic models for monthly production series are found by following the three stages of model building, namely, identification, estimation and diagnostic checking. Whenever needed, computer programs for the systematic development of the model building procedure are developed. It is found that ARIMA (1, 0, 0) x $(0,1,0)_{12}$, ARIMA (1, 0, 0) x $(1,1,0)_{12}$ ARIMA (1, 1, 0) x $(1,1,0)_{12}$models are suitable for our series. Forecasting is very important in future decisions making. The forecast based on the fitted model were also validated in this thesis in order to support future decision making for planning purpose.

ACKNOWLEDGEMENTS

First of all, I would like to convey my deepest gratitude to Professor Dr. Tun Aung, Pro-Rector of Yangon University of Economics, for his kind permissions for preparation of this thesis.

I would like to express my sincere appreciation to Professor Dr. Maw Maw Khin, Head of the Department of Statistics, Yangon University of Economics, for her permission to write the thesis in this field of the study, and for providing valuable advices and suggestions to improve this thesis.

I would like to mention my gratitude to Professor Dr. Mya Thandar, Department of Statistics, Yangon University of Economics, for her valuable advices and suggestions to improve this thesis.

I would like to acknowledge Dr. Than Pe, Lecturer (Rtd), School of Computing and Mathematics Science, University of Western Sydney, Australia, Daw Tin Tin Yu, Associate Professor (Rtd), Department of Statistics, Yangon University of Economics and Daw Aye Aye Than, Associate Professor (Rtd), Department of Statistics, Yangon University of Economics for their valuable comments and suggestions in preparing this thesis.

In addition, I am profoundly gratefully to all teachers from the Department of Statistics, especially to my supervisor, U Win Min Thant, Lecturer, Department of Statistics, Yangon University of Economics, for his constant guidance, in filling encouragement and valuable advices throughout the preparation of this thesis.

I would like to express my special deep-felt thanks to responsible persons of Hitachi Soe Electric and Machinery Co., Ltd for their kind permission to study the production of the company. I also want to say my thankfulness to U Mya Thaw, Factory Manager and authorized persons of production department.

Finally, I would like to convey my deepest gratitude to my parents for allowing and supporting me to study in the course of Master of Economics (Statistics) in Yangon University of Economics.

CONTENTS

CHAPTER III THEORETICAL BACKGROUND

CHAPTER IV RESULTS AND FINDINGS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ACF	=	Autocorrelation function
ANOVA	=	Analysis of variance
AR	=	Autoregressive
ARIMA	=	Autoregressive integrated moving average
ARMA	=	Autoregressive moving average
BODS	=	Board of directors
EMS	=	Environmental Management System
GMP	=	Good Manufacturing Practice
HIES	=	Hitachi Industrial Equipment Systems
HISEM	=	Hitachi Soe Electric and Machinery Co., Ltd
HR	=	Human Resources
ISO	=	International Organization for Standardization
KVA	=	Kilo Volt Ampere
MA	=	Moving average
PACF	=	Partial autocorrelation function
QC	=	Quality Control
QMS	=	Quality Management System
SACF	=	Sample autocorrelation function
SAR	=	Seasonal autoregressive
SSE	=	Error sum of square
SSM	=	Sum of square for months
SST	=	Total sum of square
SSY	=	Sum of square for year

CHAPTER I
INTRODUCTION

1.1 Rationale of the Study

Manufacturing is the production of merchandise for use or sale using labour and machines, tools, chemical and biological processing, or formulation. In Myanmar's manufacturing sector, the new regulation enacted recently as well as lifted from the sanction list make the companies require to be more competitive and overcome the new challenges in many different ways.

The production of transformers from HISEM Co., Ltd are performed as follow: the drawing design is calculated according to IEC 76/2000 and Core Coil and Tank Design Extrusion have been produced form with Auto CAD and sent to every production department. In addition to routine test, each new design of transformer power rating is finalized with Temperature High Test, Voltage Lighting Evolved Test, and Portable Sound Test. After testing the passing transformers are tested final QC by Research and Development Department. After that, they are packed and sent to the customers as the fixed date.

The products are distributed by the agents in Yangon, the upper Myanmar and the lower Myanmar. In distributing the products, the price has to be fixed and the system of distributing needs to be made in terms of the instruction of General Manager. To have more customer satisfaction, HISEM Co., Ltd give one to five years guarantee for the transformers.

The production of HISEM Co., Ltd is directly proportional to the demand. That is why the transformers are produced according to customer's orders. The production of transformers is increasing year by year because of transformers are produced according to international norms, because of producing in local area and the prices cheap, they send the transformers in time according their requirement.

Among the variety of transformers in HISEM Co., Ltd, four products with different system voltages were selected. Because of these products are highly production in every year.

1.2 Objectives of the Study

The objectives of this thesis are-

1. To examine the seasonality in the number of production transformer time series of HISEM Co., Ltd and to find out the seasonal indexes.
2. To construct a stochastic seasonal time series model and to obtain the forecast values for the number of production transformer time series of HISEM Co., Ltd.

1.3 Method of Study

An analytical method with the support of tables, figures, graphs and plots has been extensively used in this study. This method is observed to be more suitable to the nature and characteristics of the observed data series. More emphasis is put on analytical method of time series analysis and forecasting for the analysis of the data on the number of production transformer series of HISEM Co., Ltd.

1.4 Scope and Limitations of the Study

This study is based on the available information from Hitachi Soe Electric and Machinery Co., Ltd, literatures studies and statistical records from various publications for statistical analysis. Data used in this study were obtained from authorized persons of HISEM Co., Ltd and monthly production record. The data for the study were covered for the period January 2013 to December 2017.

1.5 Organization of the Study

This thesis consists of five chapters. Chapter I introduces rationale of the study, objectives of the study, method of study, scope and limitations of the study as well as organization of the study. Chapter II presents the profile of HISEM Co., Ltd and basic statistical characteristics of production series of HISEM Co., Ltd. Chapter III is concerned with time series, measuring seasonal variation in a time series by traditional method and the Box-Jenkins seasonal ARIMA models. Chapter IV includes the results and findings of seasonality in total number of production series HISEM Co., Ltd using traditional approach and the Box-Jenkins approach. Chapter V highlights conclusion, suggestions and further research problems in the case of seasonality in time series.

CHAPTER II

AN OVERVIEW OF HITACHI SOE ELECTRIC AND MACHINERY CO., LTD

2.1 Introduction

Monthly time series over the years display variations over the months as well as variations over the years. Monthly production series of HISEM Co., Ltd for the years 2013 to 2017 are shown in Appendix A.

In this chapter, profile of HISEM Co., Ltd and basic statistical characteristics of production series of HISEM Co., Ltd will be investigated.

2.2 Profile of Hitachi Soe Electric and Machinery Co., Ltd

Soe Electric and Machinery established in 1993 is a major power and distribution transformer manufacturer, with its Head office & Factory in Yangon, Myanmar. Head office located at Building No.1, Aung Chan Thar Housing Estate,East Shwegonedine Rd., Bahan Tsp and Factory situated at Plot No.472, 23rd Quarter, No.(1) Industrial Zone, Dagon Myothit(South), Yangon.First activity is specified for the scope of Head Office as 'Sales and Marketing of Electrical Transformers'. Second activity is specified for the scope of Factory's activity as 'Manufacture, Maintenance and Repair Services of Electrical Transformers'. The plant area site of factory is 40,000m^2.

SEM has branch offices in Naypyitaw and Mandalay for customer to provide sales & service. It also has sister company in Singapore called Soe Trading Co., Ltd, which has done the trading business on behalf of SEM. SEM holds a large share of the Myanmar market for distribution transformers in particular. SEM received a Gold Medal for successful achievement in producing distribution transformers in Industrial Fair 1996, Yangon, Myanmar and Certificate of Honour from Ministry of Electric Power for successful major repair achievement of 47 MVA 33/11 KV power transformer in 2000 and a Gold Medal for outstanding product of 10 MVA 33/11 KV Power Transformer in Myanmar Industrial Exhibition 2003, Mandalay, Myanmar.

SEM also acquired Certificate of ISO 9001:2000 for QMS in 2005. SEM has been the first electrical transformer company, which started practicing of ISO 14001:2004 EMS in 2013 and acquired the EMS certificate in 2014.SEM and HIES technically collaborated for amorphous transformers to produce lower no load losses transformer in 2013. Finally, Hitachi Group and SEM made the joint ventures in 2015 to fulfill the customers' current needs.

Hitachi Soe Electric and Machinery Co., Ltd (HISEM) formed by merging the Hitachi technological innovation of HIES and 23 years of electrical transformer manufacturing experience of SEM. The paid up capital of HISEM is USD 45 millions. As for distribution capacity will be 8000 pcs per year and that for power transformers capacity will be 800,000 KVA per year with different system voltages. The rated power has been limited from 50 KVA to 30,000 KVA. And then, rated voltage classes are 6.6 KV, 11 KV, 33 KV and 66 KV. As per HR data, SEM has manpower over 550 staffs. Normal working hours are 8:30 AM to 5:00 PM at HISEM Co., Ltd. The company is closed on the Sundays.

The outline of business are Manufacturing, Installation, Leasing, Maintenance, Repair & Sales of Electrical Transformers switchgear and transformer related accessories. HISEM achieved Asean outstanding engineering achievement award for year 2015 on the role in the local design and manufacturing of appropriated technology products in Myanmar. HISEM joined with SMBC Bank, BTMU Bank, KBZ Bank, CB Bank to satisfy their customer's payment for transformers. HISEM is mainly supply to Government Tender Project, System. Improvement Project, Industrial Zone, Construction.To maintain a competitive edge in such an environment HISEM continuously tries to improve the quality of what they offer to customers.

2.2.1 Vision, Goal, Mission and Quality Management System of Hitachi Soe Electric and Machinery Co., Ltd

(a) Vision and Goal

HISEM keeps the vision of endeavouring to be the leading joint venture transformer manufacturing private company for electrical transformers and related accessories for the best of our customers. HISEM continuously improve innovative techniques to meet demand and satisfaction of the customer with respect to time frame. HISEM work attitude is 'To make Tomorrow Better than Today'. HISEM is the unique source for superior transformers combined with high quality, competitive price and shorter lead times.

(b) Mission

The mission of the company are- conformity with said national and international quality and standards, manufacturing environment friendly transformer with bare minimum loss, performing to meet customer demand and satisfaction in full and fast manner, developing human resources in Myanmar, expanding both domestic and export business becoming globally standardized management & compliant company.

(c) Quality Management System

HISEM established and followed International Quality in Product, Service as well as Environment by complying with ISO 9001, ISO 14001 and GMP.

(1) Product Quality and Customer Satisfaction Policy

HISEM is committed to pursue excellence with quality and standard corresponding to customer needs. HISEM also committed to provide sincere service and conduct business in compliance with laws and relevant environment, emphasize conservation of energy and efficient use of natural resources on behalf of customer.

(2) GMP Policy

HISEM is committed to provide a safe and clean work environment to all its employees. Adopting Good Manufacturing Practices is seen as a way of life throughout the entire company.

(3) Environmental Policy

HISEM recognizes its responsibility towards the care of the environment and is committed to the avoidance and reduction of waste and pollution within the factory and, continual improvement of its Environmental Management System. HISEM will continually seek to reduce the usage of resources within the factory and minimize discharges that may pollute the environment.

HISEM is committed to comply with any and all applicable legislation and regulations with respect to the environment.

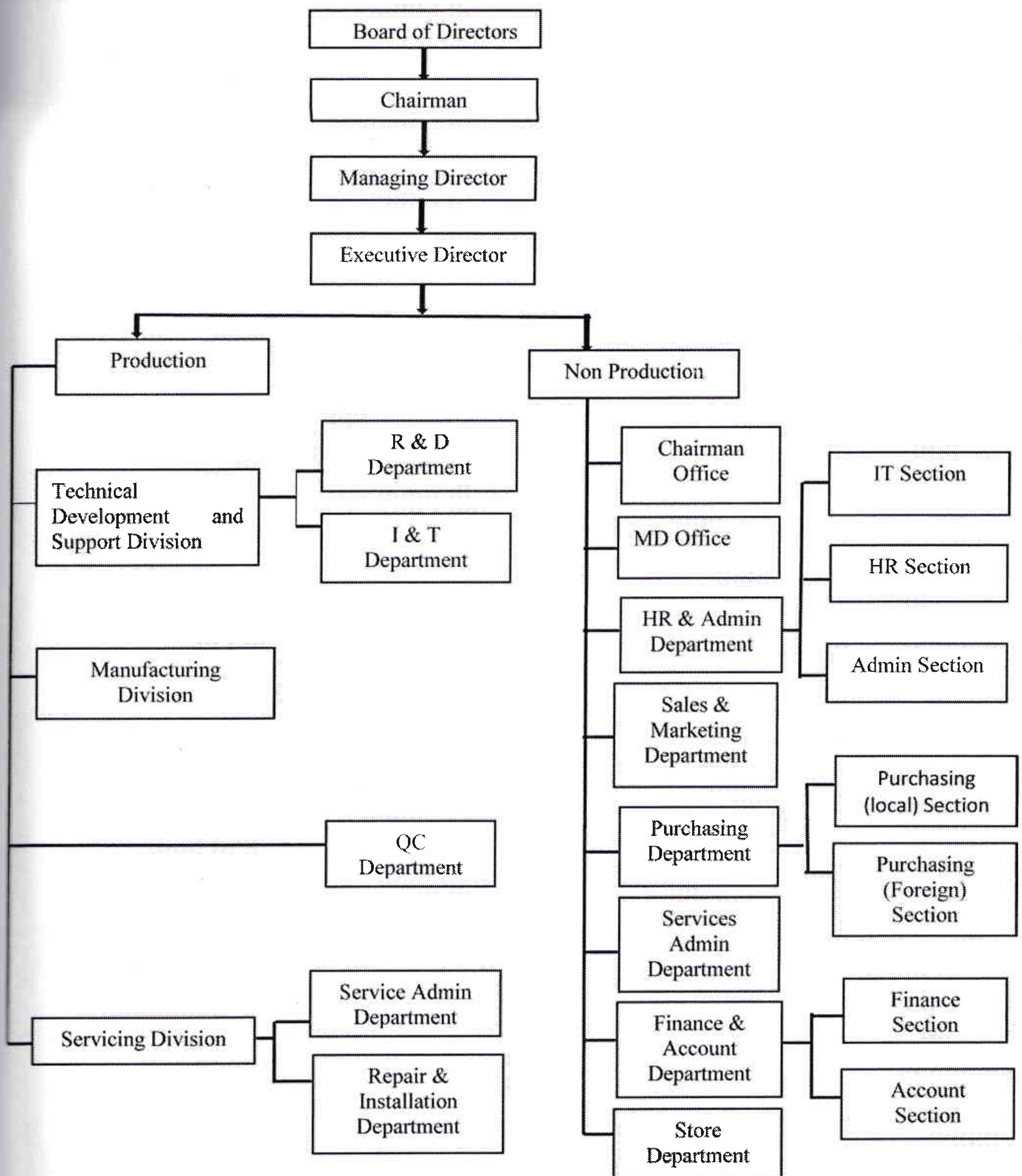
To ensure the effectiveness of its Environmental Management System, HISEM has adopted the ISO 14001:2004 and ISO 14001:2015 Standard.

The above policy provides a framework for the setting and reviewing of environmental objectives and targets is implemented and maintained within the HISEM's factory and communicated to all employees.

2.2.2 Organizational Structure of Hitachi Soe Electric and Machinery Co., Ltd

The organizational structure of Hitachi Soe Electric and Machinery Co., Ltd is shown in Figure 2.1

Figure 2.1 Organization Structure



As shown in Figure 2.1, corporate organizational structure includes; Board of Directors are formed as steering committee and two departments are comprised under Executive Director. These departments are Production department and Non Production department. There are four departments under Production department, namely, Technical Development and Support Division, Manufacturing Division, QC department and Servicing Division. Similarly, six departments are comprised under Non Production department. These are HR & Admin department, Sales & Marketing department, Purchasing department, Service Admin department, Finance & Account department and Store department. Board of director (BODs) are making important decisions concerns with production plan, budget and targets set by the BODs. The chairman is making principal decisions regarding the inventory level for each product item. He is responsible for the overall business management and supervises all the other employees. Front line managers, supervisors and staffs are the employees lie on the working floor and deal with routine jobs and upcoming problems involved in daily basic functions.

2.3 Basic Statistical Characteristics

In this section, some basic statistics of production series are presented in order to be able to see their significant variations in a summarized form.

The statistical measure used are the mean, the variance, coefficient of variation, maximum and minimum production of transformers. These values are calculated from the monthly series for each month (January to December) over a number of years and for each year over a number of months.

To calculate these values, we define y_{ij} as the value of the random variables y during j^{th}month of i^{th} year and compute,

$$\bar{y}_{i.} = \frac{1}{k}\sum_{j=1}^{k} y_{ij} \qquad ; i = 1, 2, \ldots, 5$$

= the mean value for i^{th}year

$$\bar{y}_{.j} = \frac{1}{n}\sum_{i=1}^{n} y_{ij} \qquad ; j = 1, 2, \ldots, 12$$

= the mean value for j^{th} month

$$V_i = \frac{1}{k-1}\sum_{j=1}^{k}(y_{ij} - \bar{y}_{i.})^2 \qquad ; i = 1, 2, \ldots, 5$$

= the variance for i^{th}year

$$V_j = \frac{1}{n-1}\sum_{i=1}^{n}(y_{ij} - \bar{y}_{.j})^2 \quad ; j = 1, 2, \ldots, 12$$

= the variance for j^{th} month

$$C.V\,(i) = \frac{\sqrt{V_i}}{\bar{y}_{i.}} \times 100$$

= the coefficient of variation for i^{th}year

$$C.V\,(j) = \frac{\sqrt{V_j}}{\bar{y}_{.j}} \times 100$$

= the coefficient of variation for j^{th} month

These values enable us to compare the statistical characteristics from month to month and from year to year.

2.3.1 Production Series for 100 Kilo Volt Ampere

The monthly data of production series for 100 Kilo Volt Ampere are collected for 5 years, from 2013 to 2017 and presented in Appendix A. Basic statistical characteristics of this series are investigated from two aspects. Firstly, the basic statistics for each month over a number of years (5 years) are computed. This enable us to see the pattern clearly from January to December throughout the year of the means and variances. Secondly, the basic statistics for each year over a number of months (12 months) are computed. The pattern over the years of the means and variances can be seen clearly from these.

Some basic statistics of the production series for 100 Kilo Volt Ampere are computed for each month and presented in Table (2.1).

Table (2.1)

Basic statistical characteristics for each month: Production series for 100 KVA

Month	Mean (Pcs)	Variance $(Pcs)^2$	C.V	Maximum (Pcs)	Minimum (Pcs)
January	27	35.2	21.97	38	20
February	30	29.2	18.01	39	22
March	29	18	14.63	35	24
April	27.2	13.76	13.64	31	21
May	30	33.6	19.32	36	22
June	31.6	53.04	23.05	38	19
July	26.8	16.56	15.18	30	19
August	30.4	29.44	17.85	36	22
September	27.4	29.44	19.80	33	19
October	33.8	49.76	20.87	41	22
November	35.4	45.44	19.04	42	23
December	34.4	39.04	18.16	42	26

From Table (2.1), it can be seen that the monthly mean values vary from month to month for this series. For instance, January to May, July and September have the means which are less than the overall mean 30.25(pcs). The monthly mean is highest in November with 35.4(pcs) and the lowest in July with 26.8 (pcs). The variance for each of the months vary from 13.76 $(pcs)^2$ to 53.04 $(pcs)^2$ and the coefficient of variations vary from 13.64 percent to 23.05 percent. The coefficient of variations for June is found to be largest (23.05 %). The maximum value for each month is the lowest in July and the highest in November and December. The minimum value for each month is the lowest in June, July and September and the highest in December. When the mean value for the month is large, the maximum value and the minimum value of the series are also large. For the hold observed records, the minimum value of the production series for 100 KVA is 19 (pcs), which occurs in June, July and September, 2013. Similarly, the maximum value is 42 (pcs), which occurs in November, 2016 and December, 2015.

The yearly mean value, the variance, the coefficient of variations, maximum and minimum over the twelve months for each year from 2013 to 2017 of the production series for 100 KVA are presented in Table (2.2).

Table (2.2)

Basic statistical characteristics for each year: Production series for 100 KVA

Year	Mean (Pcs)	Variance $(Pcs)^2$	C.V	Maximum (Pcs)	Minimum (Pcs)
2013	21.75	4.69	9.95	26	19
2014	27.67	12.22	12.64	34	22
2015	35.42	11.74	9.68	42	30
2016	33.17	28.97	16.23	42	25
2017	33.25	24.35	14.84	40	26

From Table (2.2), it can be seen that the yearly means vary from 21.75(pcs) in 2013 to 35.42(pcs) in 2015. The variance of each year varies from $4.69(pcs)^2$ to $28.97(pcs)^2$and coefficient of variation for each year lies between 9.68 percent and 16.23 percent.

2.3.2 Production Series for 160 Kilo Volt Ampere

The monthly data of production series for 160 Kilo Volt Ampere are collected for 5 years, from 2013 to 2017 and presented in Appendix A. Basic statistical characteristics of this series are computed in the same way as in production series for 100 KVA.

For each month these statistical characteristics are computed and presented in Table (2.3).

Table (2.3)

Basic statistical characteristics for each month: Production series for 160 KVA

Month	Mean (Pcs)	Variance $(Pcs)^2$	C.V	Maximum (Pcs)	Minimum (Pcs)
January	22	10.56	14.64	26	18
February	24.2	41.36	26.58	31	12
March	23.2	17.36	17.96	30	17
April	23.8	21.76	19.60	30	17
May	24	31.6	23.42	34	17
June	26.6	59.44	28.98	39	18
July	23	22.8	20.76	28	14
August	24.6	89.04	38.36	40	13
September	24.2	61.76	32.47	34	15
October	28.4	40.24	22.34	39	19
November	29.6	54.64	24.97	40	20
December	30.2	48.56	23.07	38	19

From Table (2.3), it can be seen that the monthly mean values vary from month to month for this series. For instance, January to May, July to September have the means which are less than the overall mean 25.33(pcs). The monthly mean is highest in December with 30.2(pcs) and the lowest in January with 22 (pcs). The variance for each of the months vary from 10.56 $(pcs)^2$ to 89.04 $(pcs)^2$ and the coefficient of variations vary from 14.64 percent to 38.36 percent. The coefficient of variations for August is found to be largest (38.36 %). The maximum value for each month is the lowest in January and the highest in August and November. The minimum value for each month is the lowest in February and the highest in November. When the mean value for the month is large, the maximum value and the minimum value of the series are also large. For the hold observed records, the minimum value of the production series for 160 KVA is 12 (pcs), which occurs in February, 2013. Similarly, the maximum value is 40 (pcs), which occurs in August and November, 2017.

The yearly mean value, the variance, the coefficient of variations, maximum and minimum over the twelve months for each year from 2013 to 2017 of the production series for 160 KVA are presented in Table (2.4).

Table (2.4)

Basic statistical characteristics for each year: Production series for 160 KVA

Year	Mean (Pcs)	Variance $(Pcs)^2$	C.V	Maximum (Pcs)	Minimum (Pcs)
2013	16.58	5.91	14.66	20	12
2014	25.08	30.91	22.16	37	16
2015	27.58	17.58	15.20	35	21
2016	24	7.83	11.66	28	19
2017	33.42	30.24	16.46	40	24

From Table (2.4), it can be seen that the yearly means vary from 16.58(pcs) in 2013 to 33.42(pcs) in 2017. The variance of each year varies from 5.91 $(pcs)^2$ to 30.91 $(pcs)^2$ and coefficient of variation for each year lies between 11.66 percent and 22.16 percent.

2.3.3 Production Series for 400 Kilo Volt Ampere

The monthly data of production series for 400 Kilo Volt Ampere are collected for 5 years, from 2013 to 2017 and presented in Appendix A. Basic statistical characteristics of this series are computed in the same way as in production series for 100 KVA.

For each month these statistical characteristics are computed and presented in Table (2.5).

Table (2.5)

Basic statistical characteristics for each month: Production series for 400 KVA

Month	Mean (Pcs)	Variance $(Pcs)^2$	C.V	Maximum (Pcs)	Minimum (Pcs)
January	22	5.84	10.79	25	18
February	28.8	4.96	7.73	32	25
March	27.2	12.16	12.82	33	23
April	24.2	24.56	20.48	30	18
May	26.4	60.24	29.40	36	15
June	28.4	86.64	32.77	38	13
July	25.6	35.44	23.25	34	17
August	26	67.6	31.62	35	12
September	25	69.6	33.37	37	14
October	32	25.2	15.69	38	27
November	34.8	29.36	15.57	42	29
December	31.8	21.76	14.67	38	25

From Table (2.5), it can be seen that the monthly mean values vary from month to month for this series. For instance, January, March to May and July to September have the means which are less than the overall mean 27.72(pcs). The monthly mean is highest in November with 34.8(pcs) and the lowest in January with 22 (pcs). The variance for each of the months vary from 4.96 $(pcs)^2$ to 86.64$(pcs)^2$and the coefficient of variations vary from 7.73 percent to 33.37 percent. The coefficient of variations for September is found to be largest (33.37 %). The maximum value for each month is the lowest in January and the highest in November. The minimum value for each month is the lowest in August and the highest in November. When the mean value for the month is large, the maximum value and the minimum value of the series are also large. For the hold observed records, the minimum value of the production series for 400 KVA is 12 (pcs), which occurs in August, 2013. Similarly, the maximum value is 42 (pcs), which occurs in November, 2017.

The yearly mean value, the variance, the coefficient of variations, maximum and minimum over the twelve months for each year from 2013 to 2017 of the production series for 400 KVA are presented in Table (2.6).

Table (2.6)

Basic statistical characteristics for each year: Production series for 400 KVA

Year	Mean (Pcs)	Variance $(Pcs)^2$	C.V	Maximum (Pcs)	Minimum (Pcs)
2013	20.42	42.41	31.90	30	12
2014	24.17	7.14	11.06	29	19
2015	27.33	16.56	14.89	34	22
2016	32.83	23.64	14.81	40	25
2017	33.83	25.47	14.92	42	22

From Table (2.6), it can be seen that the yearly means vary from 20.42(pcs) in 2013 to 33.83(pcs) in 2017. The variance of each year varies from 7.14$(pcs)^2$to 42.41$(pcs)^2$and coefficient of variation for each year lies between 11.06 percent and 31.90 percent.

2.3.4 Production Series for 2000 Kilo Volt Ampere

The monthly data of production series for 2000 Kilo Volt Ampere are collected for 5 years, from 2013 to 2017 and presented in Appendix A. Basic statistical characteristics of this series are computed in the same way as in production series for 100 KVA.

For each month these statistical characteristics are computed and presented in Table (2.7).

Table (2.7)

Basic statistical characteristics for each month: Production series for 2000 KVA

Month	Mean (Pcs)	Variance $(Pcs)^2$	C.V	Maximum (Pcs)	Minimum (Pcs)
January	23	11.36	14.78	29	19
February	21.2	4.56	10.07	25	19
March	21.6	13.04	16.72	26	17
April	22.2	5.36	10.43	26	19
May	23.4	17.84	18.05	28	17
June	22.4	23.04	21.43	30	15
July	22.2	12.56	15.96	28	18
August	20.6	31.04	27.05	25	10
September	24.2	31.36	23.14	32	15
October	26.2	7.36	10.35	29	22
November	32.6	16.24	12.36	36	25
December	27.4	8.64	10.73	30	22

From Table (2.7), it can be seen that the monthly mean values vary from month to month for this series. For instance, February, March and August have the means which are less than the overall mean 21.62(pcs). The monthly mean is highest in November with 32.6(pcs) and the lowest in August with 20.6 (pcs). The variance for each of the months vary from 4.56 $(pcs)^2$to 31.36$(pcs)^2$and the coefficient of variations vary from 10.07 percent to 27.05 percent. The coefficient of variations for August is found to be largest (27.05 %). The maximum value for each month is the lowest in February and August and the highest in November. The minimum value for each month is the lowest in August and the highest in November. When the mean value for the month is large, the maximum value and the minimum value of the series are also large. For the hold observed records, the minimum value of the production series for 400 KVA is 10 (pcs), which occurs in August, 2013. Similarly, the maximum value is 36 (pcs), which occurs in November, 2014.

The yearly mean value, the variance, the coefficient of variations, maximum and minimum over the twelve months for each year from 2013 to 2017 of the production series for 2000 KVA are presented in Table (2.8).

Table (2.8)

Basic statistical characteristics for each year: Production series for 2000 KVA

Year	Mean (Pcs)	Variance $(Pcs)^2$	C.V	Maximum (Pcs)	Minimum (Pcs)
2013	18.67	19.22	23.49	27	10
2014	23.00	22.67	20.70	36	18
2015	25.50	13.08	14.18	32	19
2016	25.50	15.25	15.31	35	20
2017	26.83	16.47	15.13	35	21

From Table (2.8), it can be seen that the yearly means vary from 18.67(pcs) in 2013 to 26.83(pcs) in 2017. The variance of each year varies from 13.08$(pcs)^2$to 22.67$(pcs)^2$and coefficient of variation for each year lies between 14.18 percent and 23.49 percent.

CHAPTER III
THEORETICAL BACKGROUND

3.1 Time Series

A time series is a set of observation measured at successive points in time or over successive periods of time. A time series is a sequence of value of some variable or composite of variables, taken at successive time periods. There are various objectives for studying time series. They include the understanding and description of the generating mechanism, the forecasting of future values, and optimal control of a system. Managers and social scientists often deal with processes that vary over time observations. A time sequence on such a process is called a time series. Time series are analysed to understand, describe, control and predict the underlying process. A time series is a sequence of n observations Y_1, Y_2,..., Y_t,..., Y_n on a process at equally spaced points in time. A time series consists of a series of observations on a variable of interest collected sequentially in time. The analysis of time is a necessary technique in many areas such as industrial research, economics, marketing, physical and chemical sciences, etc. One of the important aspects of such a series is the dependence structure of adjacent observation; for the satisfactory analysis of the series, it is necessary to construct an appropriate stochastic model which can further be used in various ways, depending on the field of applications.

In time series analysis, it is usually assumed that conditions are the same during the period for which the time series is analysed. Sometimes, conditions may change and time series observed during a certain period covers the changing conditions. The changes may be due to interventions introduced intentionally or unintentionally. It is usually desired to assertion the effect of interventions from the time series data. In such a case, direct use of conventional statistical time series covers the time period with the same condition. For example, a change in the government policy can be taken as an intervention which can change the level of an economic indicator. A time series is a collection of observations of well-defined data items obtained through repeated measurements over time. For example, measuring the value of retail sales each month of the year would comprise a time series. This is because sales revenue is well defined and consistently measured at equally spaced intervals. Data collected irregularly or only once are not time series.

3.2 Components of a Time Series

Time series as defined are discrete time series, because the observations pertain to separated points in time. There are also continuous time series, where the variable is measured continuously over time. The analysis of a time series usually involves a study of the components of the time series, such as the trend, cyclical, seasonal and irregular components. A model represents the underlying process that generates a time series. Various models exist to describe a time series. In a model the actual observations may be considered as a result of combining two components, a 'true' process or 'signal' and the random process or 'noise'. In other models, usually in economics and business, the observations are described by components, such as 'trend' cycle', 'seasonal' and 'irregular'. The first component is unpredictable variation, which can be described by a probability distribution with zero mean. In either case, the process that generates the observation can be described in terms of a set of significant pattern in time, plus an unpredictable random element.

3.2.1 Trend Component

The trend is the long- term movement in a time series. The trend component describes the net influence of long term factors. Generally, these factors include: (a) changes in the size, demographic characteristics, and geographic distribution of the population, (b) technological improvements, (c) economic development and (d) gradual shifts in habits and attitudes. Since these effects tend to operate fairly gradually and in one direction over long periods of time, the trend component usually is modelled by a smooth, continuous curve spanning the entire time series. The curve employed is called the trend curve. A major use of trend analysis is for long-term forecasting. The trend may either be an upward trend or downward trend.

3.2.2 Cyclical Component

Many variables exhibit a tendency to fluctuate above and below the long-term trend over a long period of time. These fluctuations are called cyclical fluctuations or business cycles. Typically, the cyclical component contains cycle of expansion and contraction that are of uneven duration and amplitude. Some of the factor leadings to cyclical movements in business and economic time series include buildups and depletions of inventories, shifts in rate of capital expenditure by businesses, year-to-year variations in harvests, and changes in governmental monetary and fiscal policy.

Cyclical movements are studied for information on changes in rate of current activity. This information is used for assessing current conditions and for making short-term forecast.

3.2.3 Seasonal Component

The seasonal component describes effects that occur regularly over a period of a year, month, quarter, week or day. Seasonal effects, generally, are associated with the calendar or the clock. Seasonal effects tend to recur fairly systematically. Consequently, the pattern of movement in the seasonal components tends to be more regular than the pattern in the cyclical component and therefore is more predictable, although sometimes the seasonal pattern undergoes gradual modification. Seasonal movements are measured so that seasonal effects can be taken into account in evaluating past and current activities, as well as incorporated into forecast of future activity.

3.2.4 Irregular Component

The irregular component describes residual movements that remain after the other components have been taken into account. Irregular movements reflect effects of unique and nonrecurring factors, such as strikes, unusual weather conditions, and international arises. In some business and economic time series, the cyclical component is itself so irregular that any breakdown into separate cyclical and irregular component would be arbitrary. In such cases, a combined cyclical irregular component is often developed.

3.3 Time Series Models

It is convenient to represent the series as a sum of these four components and one of the objectives may be to break the series of the down into its components, for individual study. However, in so doing, a model is imposed on the situation. It may be reasonable to suppose, that trends are due to permanent forces operating uniformly in more or less the same direction that short- term fluctuations about these long movements are in same direction.

A mathematical model of a time series may be expressed in functional form. The relationship is usually described by one of two models: the multiplicative model and the additive model.

3.3.1 Additive Time Series Model

In an additive time series model, the value of dependent variable Y can be represented as the sum of four components. Thus, the additive model takes the form

$$Y = T + S + C + I$$

Where,

Y =observed value of the variable of interest

T = trend component

S = seasonal component

C = cyclical component

I = irregular component

In the additive model, each of the four components is measured in the same units as the dependent variable Y and the components S, C and I are measured as deviation from the trend value T.

Choose the additive model when the magnitude of the seasonal pattern in the data does not depend on the magnitude of the data. In other words, the magnitude of the seasonal pattern does not change as the series goes up or down.

3.3.2 Multiplicative Time Series Models

In a multiplicative time series model the value of dependent variable Y can be represented as the product of four components. Thus, the multiplicative model takes the form

$$Y = T \times S \times C \times I$$

Where,

Y =observed value of the variable of interest

T =trend component

S = seasonal component

C = cyclical component

I = irregular component

In the multiplicative model, the trend component is expressed in the same unit of measure as the dependent variable Y. The other three components are expressed as percentage deviations from the trend.

Choose the multiplicative model when the magnitude of the seasonal pattern in the data depends on the magnitude of the data. In other words, the magnitude of the seasonal pattern increases as the data values increase, and decrease as the data values decrease.

In the additive model, the deviations from the trend are measured in absolute terms. In the multiplicative model, the deviations from the trend are measured in percentages.

3.4 Test of Seasonality

In the study of seasonality, seasonal variation for each month of the year is usually considered. The following model for the randomized complete block design (Daniel, W.W and Terre, T.C., 1992) will be used in testing seasonality in monthly tourist time series.

$$y_{ij} = \mu + \beta_i + \gamma_j + e_{ij} \ ; 1 \leq i \leq n, 1 \leq j \leq k$$

Where y_{ij} is a typical value from the overall population,

μ is an known constant,

β_i represents a yearly effect, reflecting the fact that the experimental unit fell in the i^{th} year,

γ_j represents a monthly effect, reflecting the fact that the experimental unit received the j^{th} month and

e_{ij} is a residual component representing all sources of variation other than months and years.

One make three assumptions when use the randomized complete block design. (a) Each observed y_{ij} constitutes an independent random variable of size 1 from one of the kn populations represented. (b) Each of these kn populations is normally distributed with mean μ_{ij} and the same variance 6^2. The e_{ij} are independently and normally distributed with mean 0 and variance 6^2. (c) The block and treatment effects are additive. To state this assumption another way, one say that there is no interaction between months and years.

In general, one test

H_0: There is no seasonality.

H_1: There is a seasonality.

In other words, one test the null hypothesis that the monthly means are all equal or equivalently, which mean that there are no differences in monthly effects.

To analyse the data, the needed quantities are the total sum of squares SST, the sum of squares for months SSM, the sum of squares for years SSY and the error sum of squares SSE. When these sum of squares are divided by the appropriate degree of freedom, one have the mean squares necessary for computing the F statistic. For monthly production in HISEM Co., Ltd during (2013-2017) data k=12 and n=5 years. The degree of freedom are computed as follows:

Total = Months + Years + Error

(kn-1) = (k-1)+ (n-1) + (n-1)(k-1)

Where k= months, n= years

The degrees of freedom for error can be found the following:

(kn-1)-(k-1)-(n-1) = kn-1-k+1-n+1

= kn-k-n+1

= k(n-1)-(n-1)

= (k-1)(n-1)

Short-cut formulas for computing the required sum of squares are as follows:

$$SSM = \sum_{j=1}^{k} \frac{y_{.j}^2}{n} - C \quad ; y_{.j}=\sum_{i=1}^{n} y_{ij}$$

$$SSY = \sum_{I=1}^{N} \frac{y_{I.}^2}{K} - C \quad ; y_{I.}=\sum_{j=1}^{k} y_{ij}$$

$$SST = \sum_{i=1}^{n} \sum_{j=1}^{j} y_{ij}^2 - C$$

$$SSE = SST - (SSM + SSY)$$

Where $C = \frac{y_{..}^2}{nk}$ $\quad ;y_{..} = \sum_{i=1}^{n} \sum_{j=1}^{j} y_{ij}$

The results of the calculations for the randomized complete block design are presented in the following analysis of variance (ANOVA) Table.

ANOVA Table for a Two-Way Analysis of Variance

Source	S.S	D.F	M.S	F-Ratio
Between Months	SSM	k-1	MSM = SSM/k-1	F_1 = MSM/MSE
Between Years	SSY	n-1	MSY = SSY/n-1	F_2 = MSY/MSE
Error	SSE	(n-1)(k-1)	MSE = SSE/(n-1)(k-1)	
Total	SST	kn-1		

The computed ratios F_1 with critical values $K_1 = F_{\alpha,(k-1),(n-1)(k-1)}$ is then compared. If this ratios are equal to or exceed the critical values, reject the null hypothesis.

3.5 Method of Finding Seasonal Variation

Seasonal variation is measured in terms of an index, called a seasonal index. It is an average that can be used to compare an actual observation relative to what it would be if there were no seasonal variation. An index value is attached to each period of the time series within a year. This implies that if monthly data are considered there are 12 separate seasonal indices, one for each month. There exists different methods for measuring the seasonal variation of a time series. The methods have been developed to meet different objectives of estimating seasonal and the assumed models of the time series. The seasonal pattern itself is important in the application of these methods since most of the methods assume that the seasonal pattern is constant or stable.

In finding the index of seasonal variation as seasonal measures, it should be noted that the index must

(a) Measure all the variation in the series that is seasonal in character, and

(b) Measure nothing but the seasonal variation

A seasonal index thus consists of a series of percentage figures, averaging 100, which shows the relative level of the series for the various months, quarters or weeks of the year. An index of seasonal variation can be constructed by expressing each item in the time series as a percent of the average monthly or quarterly value for the year.

There are many different methods for computation of seasonal index, some of which are quite accurate and some of which are only appropriate. The following methods will be discussed in this section.

(1) Average percentage method

(2) Ratio to moving average method

(3) Link relatives method

(4) Ratio to trend method

These methods have been developed to meet different objectives of estimating seasonals and under the assumed models of the time series. The seasonal pattern itself is important in the application of these methods since most of the methods assume that the seasonal pattern is constant or stable. Of these methods, the Ratio to Moving Average method and the Link Relatives method are simple and which are the most widely used.

Average Percentage Method

In this method the data for each month are expressed as percentages of the average for the year. The percentages for corresponding months of different years are then averaged using either a mean or median. If the mean is used, it is best to avoid extreme values which may occur.

The resulting 1200 percentages give the seasonal index. If their mean is not 100 % (i.e., if the sum is not 1200). These should be adjusted.

Ratio to Moving Average Method

The measurement of seasonal variation by using the ratio-to-moving-average method provides an index to measure the degree of the seasonal variation in a time series. The index is based on a mean of 100, with the degree of seasonality measured by variations away from the base. The following are the steps for the computation of the seasonal index by the Ratio to Moving Average method. (Steiner, 1956)

(1) Find the twelve months centered moving averages. This is equivalent to a moving average of thirteen months with weights $\frac{1}{24}$ (1,2,2,…,2,2,1).

By finding twelve months centered moving averages, we eliminate the seasonality, since the seasonal pattern is periodic with a period of twelve months. Also it will eliminate the random components or irregular movements. Therefore, the centered twelve month moving averages are the approximates of trend and cyclical components.

(2) Compute the ratio to moving average values, that is, the original data is divided by its approximate moving average value. There, the first and last six months may not be obtained.

By this step, the trend and cyclical components are removed from the original data and the ratios are the values due to seasonal and random components. They are called specific seasonals. (Steiner, 1956).

(3) Compute the averages of these ratios referring to the same months.

These averages are the crude seasonal index values.

This step involves two different purposes: the elimination of the random components and averaging the seasonal relatives referring to the same months.

(4) Adjust the crude seasonal index.

In multiplicative model, the total seasonal index values have to be equal to twelve (or 1200 percent) for monthly series. Therefore, the crude seasonal index is adjusted to get a total of twelve (or 1200 percent).

Link Relatives Method

The same assumptions as in ratio to moving average method have to be made to compute the seasonal index by the link relatives' method. The following are the steps for the computation of the seasonal index by the link relatives' method.

(1) Find the link relatives' value.

This is to divide the current value by the previous values. Then, the first one may not be obtained. These values show the relative changes of the consecutive values.

(2) Find the averages of the link relatives values referring to the same months

These averages show the average changes in consecutive months within the whole period of twelve months.

(3) Compute the chain relative values by assuming that the chain relative value of the first month is unity.

The chain relative value for the current month is the product of the chain relative value of the previous month and the average of link relatives for the current month. These chain relative values constitute seasonal pattern and the trend within a year.

(4) Determine the trend component within the year and adjust for the trend

To determine the trend component within a year, the chain relative value of the first month is computed, that is, the product of the chain value of the last month and the average of the link relatives for the first month is computed and the difference between the chain relative value and the setting value unity is found. This difference is regarded as the trend for twelve months. By dividing this value by twelve, the difference for a month is obtained, which is assumed to be the coefficient of linear trend and denoted by Δ (delta). If value $(i-1)\Delta$, $i=1,2,\ldots,12$ are subtracted from the corresponding chain relative values. Similarly, if the delta is negative, there exists a downward trend and the respective trend values $(i-1)\,|\Delta|$, $i=1,2,\ldots,12$ are added to the corresponding chain relative values. After the adjustment, the adjusted chain relative values are regarded as the crude seasonal index.

(5) Adjust the crude seasonal index

The crude seasonal index is adjusted to get a total of twelve (or 1200 percent) and the seasonal index is obtained.

Ratio to Trend Method

The following are the steps to compute seasonal index by the method of ratio to trend.

(1) Compute monthly trend values by the method of least squares.

(2) Express each original value as the percentage of the corresponding trend value.

(3) Find out the mean percentage for each month.

(4) Values obtained in (3) above give seasonal variations. Seasonal index can be calculated from these mean percentages by expressing them as percentage of their own average.

3.6 The Box – Jenkins Methodology

The Box – Jenkins methodology has been expressed steps for model identification, methods of the estimation of the parameters in the ARIMA models, diagnostic checking and forecasting.

3.6.1 Model Identification

Consider the general ARIMA (p, d, q) model

$$(1-\phi_1 B-...-\phi_p B^p)(1-B)\dot{Z}_t=\theta_0+(1-\theta_1 B-...-\theta_q B^q)a_t$$

Model identification refers to the methodology in identifying the required transformations such as variance stabilizing transformation and differencing transformations, the decision to include the deterministic parameter θ_0 when $d \geq 1$ and the proper order of p and q for the model.

The following useful steps are used to identify a tentative model.

Step 1. Plot the time series data and choose proper transformations. In any time series analysis, the first step is to plot the data. One usually gets a good idea about whether the series contains a trend, seasonality, outliers, non-constant variance and other non- normal and non-stationary phenomena. This understanding often provides a basis for postulation a possible data transformation.

In time series analysis, the most commonly used transformations are variance- stabilizing transformations and differencing. Since differencing may create some negative values, one should always apply variance stabilizing transformations before taking differences. A series with non-constant variance often needs a logarithmic transformation. More generally, to stabilize the variance, one can apply Box- Cox's power transformation.

Step2. Compute and examine the sample ACF and the sample PACF of the original series to further confirm a necessary degree of differencing. Some general rules are:

1. If the sample ACF decays very slowly and the sample PACF cuts off after lag 1 it indicates that differencing is needed. Try taking the first differencing $(1-B)Z$,
2. More generally, to remove non-stationary that one may need to consider a higher order differing $(1-B)^d Z,$ for $d > 1$. In most cases, d is either, 0,1 or 2. Some authors argue that the consequences of unnecessary differencing are much less serious than those of under different.

Step3. Compute and examine the sample ACF and PACF of the properly transformed and differenced series to identify the orders of p and q, where p is the highest order in AR polynomial $(1-\phi_1 B-...\phi_p B^p)$ and q is the highest order in MA polynomial $(1-\theta_1 B-...\theta_q B^q)$. Usually the needed orders of these p and q are less than or equal to 3.

It is useful and interesting to note that a strong duality exists between the AR and MA model in terms of their ACFs and PACFs. To build a reasonable ARIMA model, one need a minimum of n = 50 observations and the number of sample ACF and PACF to be calculated should be about $\frac{n}{4}$, although occasionally for data of good quality one may be able to identify and adequate model with a smaller sample size. To identify the order p and q by matching patterns in the sample ACF and PACF with the theoretical pattern of known model.

Table (3.1)

Characteristics Behaviour of ACF, PACF for AR, MA and ARMA Process

Process	Autocorrelation	Partial Autocorrelation
AR (p)	Infinite (damped exponentials and / or damped since waves). Tail off according to $\rho_j=\phi_1\rho_{j-1}+\phi_1\rho_{j-2}+...+\phi_p\rho_{j-p}$	Finite Spike at lag 1 through p, then cut off
MA(q)	Finite Spike at lag 1 through q, then cuts off	Infinite (dominated by damped exponentials and / or damped sine waves) Tail off
ARMA(p,q)	Infinte (damped expoentials and / or damped sine waves after first q-p lags). Irregular pattern at lag 1 through q, then tails off according to $\rho_j=\phi_1\rho_{j-1}+\phi_1\rho_{j-2}+...+\phi_p\rho_{j-p}$	Infinite (dominated by damped exponentials and^ or damped sine waves after first q-p lags) Tail off

Source: Univariate and Multivariate Methods (William W.S.Wei)

Step 4. Test the deterministic trend term θ_0 when d > 0 for nonstationary model, $\phi_p(B)(1-B)^d Z_t = \theta_0 + \theta_q(B)a_t$, where the parameter θ_0 is usually omitted so that it is capable of representing series **with** random changes in the level, slope or trend. However, the **differenced** series contains a deterministic trend mean, one can test for its **inclusion** by comparing the sample mean $\bar{W}$ of the differenced series ***W_t= (1-B)^d*** *Z* with its approximate standard error $S_{\bar{W}}$.

To derive $S_{\bar{W}}$

$\lim_{n\to\infty} nVar(\bar{W}) = \sum_{j=\infty}^{\infty} \gamma_j$, and hence,

$$\sigma_{\bar{W}}^2 = \frac{\gamma_O}{\text{n}} \to \sum{}_j^\infty = \gamma_j = \frac{1}{n}\sum{}_j^\infty = -\infty\gamma_j = \frac{1}{n}\gamma(1) \tag{3.1}$$

Where, $\gamma(B) = \sum_{-\infty}^{\infty}\gamma_K B^K = \sigma_a^2\psi(\text{B})\psi(\text{B})^{-1}$ is the autocovariance generating function and r(1) is its value at b = 1. Thus, the variance and hence the standard error for $\bar{W}$ is model dependent. For the ARIMA (1, d, 0) model, $(1-\phi_1 B)\text{W}_t = a_t$

$$(1-\phi_1\text{B})\,(1\text{-B})\,^\text{d}\dot{\text{Z}}_\text{t} = \text{a}_\text{t}$$

$$(1-\phi_1\text{B})\ \text{W}_\text{t} = \text{a}_\text{t};\text{W}_\text{t} = \frac{1}{(1\text{-}\phi_1\text{B})}\text{a}_\text{t}$$

MA representation, $\text{Z}_\text{t} = \psi(\text{B})\text{a}_\text{t}$

$$\psi(\text{B}) = \frac{1}{(1\text{-}\phi_1\text{B})}$$

Autocovariance generating function is

$$\gamma(\text{B}) = \sigma_\text{a}^2\psi(\text{B})\psi(\text{B})^{-1} = \frac{\sigma_a^2}{(1-\phi_1\text{B})(1-\phi_1\text{B}^{-1})}$$

Where, B = 1, $\gamma(1) = \dfrac{\sigma_a^2}{(1-\phi_1\text{B})^2}$

$$\sigma_{\bar{\text{w}}}^2 = \frac{\sigma_a^2}{\text{n}(1-\phi_1)^2}$$

$$= \frac{\sigma_\text{w}^2(1-\phi_1^2)}{\text{n}(1-\phi_1)^2} \qquad (\because \sigma_w^2 = \frac{\sigma_a^2}{(1-\phi_1)^2})$$

$$= \frac{\sigma_w^2}{n}\left[\frac{1+\phi_1}{1-\phi_1}\right]$$

$$= \frac{\sigma_w^2}{n}\left[\frac{1+\rho_1}{1-\rho_1}\right] \qquad (\because \phi_1 = \rho_1) \tag{3.2}$$

The required standard error is

$$S_{\bar{W}} = \sqrt{\frac{\gamma_0}{n}\left[\frac{1+\hat{\rho}_1}{1-\hat{\rho}_1}\right]} \tag{3.3}$$

Expression of $S_{\bar{W}}$ for other models can be derived similarly. However, at the model identification phase, since the underlying model is unknown, most available software use the approximation.

$$S_{\bar{W}} = \left[\frac{\hat{\gamma}_0}{n}(1+2\hat{\rho}_1+2\hat{\rho}_2+\ldots+2\hat{\rho}_k\right]^{1/2} \tag{3.4}$$

Where, $\hat{\gamma}_0$ is the sample variance and $\hat{\rho}_1, \hat{\rho}_2, \ldots, \hat{\rho}_k$ are the first k significance sample autocorrelation function of (W_t).

Under null hypothesis $\rho_k = 0$; for $k \geq 1$

$$S_{\bar{W}} = \sqrt{\frac{\gamma 0}{n}} \tag{3.5}$$

Alternatively, one can include θ_0 initially and discard it at the final model estimation if the preliminary estimation result is not significant.

3.6.2 Parameter Estimation

After a model is identified for a given time series it is important to obtain efficient estimates of the parameters. To obtain the estimate of parameters $\phi_1, \phi_2, \ldots, \phi_p \theta_1, \theta_2, \ldots, \theta_q$, one may use the least squares method since t_i can be proved that the least squares estimates are approximately maximum likelihood estimates in ARIMA models. If the least squares, method is used, to choose those value of ϕ'_s and θ'_s of the parameter set which minimize the sum of squared error $\sum_{t=1}^{n} a_t^2$ obtained from the observed time series.

There arises two difficulties in estimation stage:

(i) The equation involve unknown starting values,

(ii) The sum of squared errors function is in general nonlinear in the coefficients to be estimated.

There are two approaches to (i)

(a) The unknown starting values are simply replaced by some appropriately assumed values and estimation is conditional on these assumed starting values.

(b) The estimation is based on estimated starting values from the sample data. This unconditional approach is more efficient than the conditional approach. For long series, the difference between the results obtained by the two approaches is negligible.

Conditional Maximum Likelihood Estimation

For general stationary ARMA (p,q) model,

$$\dot{Z}_t = \phi_1 \dot{Z}_{t-1} + \ldots + \phi_1 \dot{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \tag{3.6}$$

Where, $\dot{Z}_t = \dot{Z}_t - \mu$ and $\{a_t\}$ are independent identically distributed (i.i.d), N(0, σ_a^2) white noise,

Joint probability density of a = $(a_1, a_2, \ldots, a_n)'$ is given by

$$p(a/\phi, \mu, \theta, \sigma_a^2) = (2\pi\sigma_a^2)^{-n/2} \exp\left(-\frac{1}{2\sigma_a^2}\sum_{t=1}^{n}\sigma_a^2\right) \tag{3.7}$$

Rewriting equation (3.6) as

$$a_1 = \theta_1 a_{t-1} + \ldots + \theta_q a_{t-1} + Z_t - \phi_1 \dot{Z}_{t-1} - \ldots - \phi_p \dot{Z}_{t-p} \tag{3.8}$$

one can note down the likelihood function of the parameter $(\phi, \mu, \theta, \sigma_a^2)$.

Let Z = $(Z_1, Z_2, \ldots, Z_n)'$ and assume that initial conditions

$Z_* = (Z_{1-p}, Z_{-1}, Z_0)'$ and $a_* = (a_{1-p}, a_{-1}, a_0)'$

The conditional log likelihood function,

$$\text{LnL}_* (\phi, \mu, \theta, \sigma_a^2) = -\frac{n}{2}\text{Ln}2\pi\sigma_a^2 - \frac{S_*(\phi, \mu, \theta)}{2\sigma_a^2} \tag{3.9}$$

Where,

$$S_*(\phi, \mu, \theta) = \sum_{t=1}^{n} a_t^2(\phi, \mu, \theta / Z_*, a_*, Z) \tag{3.10}$$

is the conditional sum of squares function.

The quantities of $\hat{\phi}, \hat{\mu}$, and $\hat{\theta}$, which maximize equation (3.9) are called the conditional maximum likelihood estimators.

Since Ln L* $(\phi, \mu, \theta, \sigma_a^2)$ involves the data only through S* $(\phi, \mu, \theta,)$ are the same as the conditional least squares obtained from minimizing the conditional sum of square function S* $(\phi, \mu, \theta,)$, which does not contain the parmeter σ_a^2.

There are a few alternatives for specifying the initital condition Z* and a*, based on the assumptions that $\{Z_t\}$ is stationary and $\{a_t\}$ is as series of i.i.d, $N(0, \sigma_a^2)$.

The unknown Z_t by the sample mean $\bar{Z}$ and unknown a_t by its expected value of 0, and also assume $a_p = a_{p-1} = \ldots = a_{p+1-q} = 0$ and by using equation (3.6) calculate a_t for $t \geq (P+1)$

The conditional sum of square equation (3.10) become

$$S_*(\phi, \mu, \theta) = \sum_{t=p+1}^{n} a_t^2(\phi, \mu, \theta / Z) \tag{3.11}$$

Which is also the form used by most computer programs.After obtaining the parameter estimates $\hat{\phi}, \hat{\mu}$ and $\hat{\theta}$, the estimates σ_a^2 of σ_a^2 is calculated from

$$\sigma_a^2 = \frac{S*(\hat{\phi}, \hat{\mu}, \hat{\theta})}{d.f}$$

If equation (3.11) is used to calculate the sum of squares, d.f = (n-p) – (p+q+l)

d.f = n – (2p+q+1)

where, the number of degree of freedom d.f equals the number of terms used in the sum of S* (ϕ, μ, θ) minus the number of parameters estimated.

Unconditional Maximum likelihood Estiation and Backcasting Method

One of the most important functions of a time series model is to forecast the unknown future value. And then one can back forecast or back cast the unknown value Z* = $(Z_1\text{-p}, \ldots .Z_1, Z_0)'$ and a* = $(a_1\text{-q}, \ldots .a_1, a_0)'$ needed in the computation of the sum of squares and likelihood functions.

Any ARMA model can be written in either the forward form.

$$(1-\phi B- \ldots -\phi_p B^p)Z_t = (1-\theta_1 B- \ldots -\theta_q B^q)a_t \tag{3.13}$$

Where, $B^j Z_t = Z_{t-j}$

or the backward form,

$$(1-\phi F- ...-\phi_p F^p)Z_t = (1-\theta_1 F ...-\theta_q F^q)a_t \quad (3.14)$$

Where, $F^j Z_t = Z_{t-j}$

Because of the stationary equation (3.13) and equation (3.14) should have exactly the same auto covariance structure. $\{a_t\}$ is a white noise with men zero and constant variance σ_a^2.$\{e_t\}$ is also a white noise with mean zero and constant variance σ_e^2.

The forward from equation (3.13) is used to forecast the unknown future values Z_{n+j} for j>0 base on the data $(Z_1, Z_2, \ldots, Z_n)$. The backward form equation (3.14) is also used to backcast the unknown past value Z_j and hence compute a_j for $j \leq 0$ based on the data $(Z_n, Z_{n-1}, \ldots, Z_1)$. Estimation, Box and Jenkins (1976) suggest the following unconditional log likelihood function.

$$\text{Ln L}(\phi, \mu, \theta, \sigma_a^2) = -\frac{n}{2}\text{Ln } 2\pi\sigma_a^2 - \frac{S(\phi,\mu,\theta)}{2\pi\sigma_a^2} \quad (3.15)$$

Where, $S(\phi, \mu, \theta)$ is the conditional sum of square function given by,

$$S(\phi, \mu, \theta) = \sum_{t=-\infty}^{n} [E(a_t / \phi, \mu, \theta, Z)]^2 \quad (3.16)$$

The quantities $\hat{\phi}$, $\hat{\mu}$ and $\hat{\theta}$ that maximize equation (3.15) are called unconditional maximum likelihood estimators. Ln L $(\phi, \mu, \theta, a_a^2)$ involves the data only through $S(\phi, \mu, \theta)$, these unconditional maximum likelihood estimators are equivalent to the unconditional least square estimators obtained by minimizing $S(\phi, \mu, \theta)$.

In practice, equation (3.16) is approximated by a finite form,

$$S(\phi, \mu, \theta) = \sum_{t=-M}^{n} [E(a_t / \phi, \mu, \theta, Z)]^2$$

where, M is a sufficiently large integer such that the backcast increment

$$|E(Z_1/ \phi, \mu, \theta, Z) - E(Z_{t-1}/ \phi, \mu, \theta, Z)| < \varepsilon \text{ for } t \leq -(M+1).$$

Condition expectation, $E(Z_t/ \phi, \mu, \theta, Z)$ and hence $E(a_1 / \phi, \mu, \theta, Z)$ is negligible for $t \leq -(M+1)$. After obtaining the parameter estimate $\hat{\phi}$, $\hat{\mu}$ and $\hat{\theta}$, the estimate $\hat{a}_a^2$ of a_a^2 can be calculated as

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

For efficiency, the use of backcasts for parameter estimation is important for seasonal models that are close to be non-stationary, that the series are relatively short.

3.6.3 Diagnostic Checking

Time series model building is an iterative process. It starts with model identification and parameter estimation. After parameter estimation, one has to assess model adequacy by checking whether the model assumptions are satisfied. The basic assumption is that the $\{a_t\}$ are white noise. The a_t's are uncorrelated random shocks with zero mean and constant variance. For any estimated model, the residuals a_t's are estimates of these unobserved white noise a_t's. Hence, model diagnostic checking is accomplished through a careful analysis of the residual series $(\hat{a}t)$. Because this residual series is the product of parameter estimation, the model diagnostic checking is usually contained in the estimation phase of a time series package.

(1) To check whether the errors are normally distributed, one can construct a histogram of the standardized residuals $\frac{\hat{a}t}{\hat{a}_a^2}$ and compare it with the standard normal distribution using the chi-square goodness of fit test.

(2) To check whether the variance is constant, one can examine the plot of residuals or evaluate the effect of different λ value via Box-Cox method.

(3) To check whether the residuals are approximately white noise, one can compute the sample ACF and sample PACF (or IACF) of the residuals to see whether they do not form any pattern and are all statistically insignificant.

Another useful test is the portmanteau Lack of fit test. This test uses the entire residual sample ACF's to check null hypothesis.

Hypothesis H_0 : $\rho_1 = \rho_2 = \ldots = \rho_k = 0$

The residual are not autocorrelated.

H_1 : The residual are autocorrelated.

Test statistics : $Q = n(n+2)\sum_{k=1}^{k} (n-k)^{-1} \hat{\rho}_k^2$

Critical value : $K = \chi^2_{(\alpha, k-m)}$

Decision Rule : Q > K ; Reject H_0

Otherwise ; Accept H_0

Where, m = the number of parameter estimated in the model. Based on the residual results, if the model is inadequate, a new model can be easily derived.

3.6.4 Minimum Mean Square Error Forecasts

In forecasting, one objective is to produce an optimum forecast that has no error or as little error as possible, which leads us to the minimum mean square error forecast. This forecast will produce and optimum future value with the minimum error in terms of the mean square error criterion.

Minimum Mean Square Error Forecasts for ARIMA models

Consider the general nonstationary ARIMA (p, d, q) model with d ≠ 0, i.e,

$$\phi(B)(1-B)^d Z_t = \theta(B)a_t \tag{3.17}$$

Because the model is stationary, a moving average representation,

$$Z_t = \psi(B)a_t$$

$$= a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots, \tag{3.18}$$

Where

$$\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j = \frac{\theta(B)}{\phi(B)} \tag{3.19}$$

And $\psi_0 = 1$. For t = n + 1,

$$Z_{n+1} = \sum_{j=0}^{\infty} \psi_j a_{n+l-j} \tag{3.20}$$

Suppose that at time t = n one have the observations $Z_n, Z_{n-1}, Z_{n-2}, \ldots, Z_1$ and wish to forecast l-step ahead of future value Z_{n+l} as a linear combination of the observation $Z_n, Z_{n-1}, Z_{n-2}, \ldots$. Because Z_t for t = n, n -1, n-2, ... can all be written in the form of (3.18), Let the minimum mean square error forecast $\hat{Z}_n(l)$ or Z_{n+l} be

$$\hat{Z}(l) = \psi_1^* a_n + \psi_{1+1}^* a_{n-1} + \psi_{1+2}^* a_{n-2} + \ldots \tag{3.21}$$

Where the ψ_j^* are to be determined. The mean square error of the forecast is

$$E(Z_{n+1} - \hat{Z}_n(l))^2 = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 + \sigma_a^2 \sum_{j=0}^{\infty} \left[\psi_{l+j} - \psi_{l+j}^*\right],$$

Which is easily seen to be minimized when $\psi^{*}_{l+j} = \psi_{l+j}$. Hence,

$$\hat{Z}_n(l) = \psi_1 a_n + \psi_{l+1} a_{n-1} + \psi_{l+2} a_{n-2} + ... \qquad (3.22)$$

Using (3.20) and that

$$E(a_{n+j} / Z_n, Z_{n-1}, ...) = \{^{0}_{a_{n+j}} \, j > 0,$$

$$E(a_{n+j} / Z_n, Z_{n-1}, ...) = \psi a_n + \psi_{l+1} + a_{n-1} + \psi_{l+2} + a_{n-2} + ...$$

Thus, the minimum mean square error forecast of Z_{n+l} is given by its conditional expectation. That is,

$$Z_n(l) = E(Z_{n+1} / Z_n, Z_{n-1}, ...) \qquad (3.23)$$

$\hat{Z}_n(l)$ is usually read as the l-step ahead forecast of Z_{n+l} at the forecast origin n. The forecast error is

$$e_n(l) = Z_{n+1} - \hat{Z}_n(l) = \sum_{j=0}^{l-1} \psi_j a_{n+l-j} \qquad (3.24)$$

Because E (e_n(1)| Zt, t ≤ n) = 0, the forecast is unbiased with the error variance

$$Var(e_n(l)) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 \qquad (3.25)$$

For a normal process, the (1- α) 100% forecast limits are

$$\hat{Z}_n(l) \pm N_{\alpha/2} \left[1 + \sum_{j=1}^{l-1} \psi_j^2 \right]^{1/2} \sigma_a, \qquad (3.26)$$

Where $N_{\alpha/2}$ is the standard normal deviate such that $P(N > N_{\alpha/2}) = \alpha/2$.

There forecast error $e_n(l)$ as shown (3.24) is a linear combination of the future random shocks entering the system after time n. Specifically, the one-step ahead forecast error is

$$e_n(l) = Z_{n+l} - \hat{Z}_n(l) = a_{n+l} \qquad (3.27)$$

Thus, the one-step ahead forecast errors are independent, which implies that $\hat{Z}_n(l)$ is indeed the best forecast of Z_{n+l}. Otherwise, if one-step ahead forecast errors are correlated, then one can construct a forecast $\hat{a}_{n+1}$ of a_{n+1} from the available errors a_n, a_{n-1}, a_{n-2}, ... and hence improve the forecast of Z_{n+l} by simple using

$\hat{Z}_n(l)+\hat{a}_{n+l}$ as the forecast. The forecast error for longer lead times, however, are correlated. This correlation is true for the forecast errors

$$e_n(l)=Z_{n+1}-\hat{Z}_n(l)=a_{n+l},\psi_1 a_{n+l-1}+\psi_{l-1}a_{n+1} \tag{3.28}$$

And $e_{n-j}(l)=Z_{n+l-j}-\hat{Z}_{n-j}(l)=a_{n+l-j}+\psi_1 a_{n+l-1}+....+\psi_{l-1}a_{n-j+1}$ (3.29)

Which are made at the same lead time l but different origins n and n-j for $j<l$. It is also true for the forecast errors for different lead time made from the same time origin.

Minimum Mean Square Error Forecasts for ARIMA Models

Consider the general nonstationary ARIMA (p, d, q) model with d ≠ 0, i.e.,

$$\phi(B)(1-B)^d Z_t=\theta(B)a_t \tag{3.30}$$

Where ϕ (B) = (1 - ϕ_1 B - .... - ϕ_p B^{ϕ}) is a stationary AR operator and θ(B) = (1-θ_1B - ... - $\theta_q B^q$) is an invertible MA operator, respectively. Although for this process the mean and the second-order moments such as the variance and the autocovariance functions vary over time. The complete evolution of the process is completely determined by a finite number of fixed parameters. The forecast of the process as the estimation of a function of these parameters and obtain the minimum mean square error forecast using a Bayseina argument. Using this approach with respect to the mean square error criterion, which corresponds to a squared los function, when the series is known up to time n, the optimal forecast of Z_{n+l} is given by its conditional expectation E (Z_{n+l}| Z_n, Z_{n-1}....). The minimum mean square error forecast for the stationary ARMA model discussed is, of course, a special square case of the forecast for the ARIMA (p,d,q) model with d = 0.

To derive the variance of the forecast for the general ARIMA model, the model at time $t+l$ in an AR representation that exists because the model is invertible.

Thus, $$\pi(B)Z_{t+l}=a_{t+l} \tag{3.31}$$

Where $$\pi(B)=1-\sum_{j=1}^{\infty}\pi_j B^j=\frac{\phi(B)(1-B)^d}{\theta(B)} \tag{3.32}$$

Equivalently, $$Z_{t+l}=\sum_{j=1}^{\infty}\pi_j Z_{t+l-j}+a_{t+l} \tag{3.33}$$

Apply the operator, $1+1+\Psi_1 B+...+\Psi_{l+j}B^{l-1}$

to (3.33) and obtain $$\sum_{j=0}^{\infty}\sum_{k=0}^{l-1} \pi_j \Psi_k Z_{t+l-j-k} + \sum_{j=0}^{l-1} \Psi_k a_{t+l-k} = 0 \tag{3.34}$$

where $\pi_0 = -1$ and $\Psi_0 = 1$. It can be shown that

$$\sum_{j=0}^{\infty}\sum_{k=0}^{l-1} \pi_j \Psi_k z_{t+l-j-k} = \pi_0 Z_{t+l} + \sum_{m=1}^{\infty}\sum_{i=0}^{l-1} \pi_{l-1+j-i} \Psi_i Z_{t-j+1} \tag{3.35}$$

Choosing Ψ weights $\sum_{i=0}^{m} \pi_{m-i} \Psi_i = 0$, form $= 1, 2, \ldots\ldots, l-1$ (3.36)

$$Z_{t+l} = \sum_{j=1}^{\infty} \pi_j^{(l)} + \sum_{i=0}^{l-1} \Psi_t a_{t+l-i} \tag{3.37}$$

Where, $$\pi_j^{(l)} = \sum_{i=0}^{l-1} \pi_{l-1+j-1} \Psi_t \tag{3.38}$$

Thus, given Z_t for $t \leq n$

$$\hat{Z}_n(l) = E(Z_{n+1} / Z_t, t \leq n)$$

$$= \sum_{j=1}^{\infty} \pi_j(l) Z_{n-j+1} \tag{3.39}$$

Because $E\,(a_{n+j} / Z_t, t \leq n) = 0$ for $j > 0$. The forecast error is

$$e_n(l) = Z_{n+l} - \hat{Z}_n(l)$$

$$= \sum_{j=0}^{l-1} \Psi_j a_{n+l-j} \tag{3.40}$$

Where the Ψ_j weights, by (3.36) can be calculated recursively from the π_j weights as follow.

$$\Psi_j = \sum_{i=0}^{j-1} \pi_{j-i} \Psi_i \qquad j = 1, \ldots, l-1. \tag{3.41}$$

3.7 Seasonal Time Series Models

In this section, seasonal time series are discussed. These models were developed by Box and Jenkins (1976) and have been successfully applied to many time series with seasonal variation.

3.7.1 The Seasonal Autoregressive Process of Order P, SAR (P)

The seasonal autoregressive process of order P(1) if s is the number or observation per seasonal period then the order of the AR process is an integer multiple of s and (2) the non-zero coefficents are those with subscripts that are an integer multiple of s.

The SAR (P) model is

$$Z_t = Ø_s Z_{t-s} + Ø_{2s} Z_{t-2s} + \cdots + Ø_{ps} Z_{t-ps} + a_t \tag{3.42}$$

Where, P is the largest multiple of s presented in the model. To provide special notation for the seasonal model, and so if we let

$$Ø_{js} = \Phi_{js} \tag{3.43}$$

So that Equation (3.42) become,

$$\dot{Z}_t = \Phi_1 \dot{Z}_{t-1} + \Phi_2 \dot{Z}_{t-2s} + \cdots + \Phi_p \dot{Z}_{t-ps} + a_t \tag{3.44}$$

referred to as seasonal AR process of order P. the seasonal autoregressive model in equation (3.44) expresses the current value of the process Z_t as finite weighted sum of P previous values Z_{t-s}, Z_{t-2s}, ..., Z_{t-ps} of the process plus random shock a_t.

Here,	$E[a_t]$	$= 0$	for all t.
	$V[a_t]$	$= E[a_t^2] = \sigma_a^2$	for all t, and
	$Cov[a_t, a_{t'}]$	$= E[a_t, a_{t'}] = 0$	for all $t \neq t'$

The Autoregressive Function of SAR (P) Process is

$$\gamma_k = \Phi_1 \gamma_{k-s} + \Phi_2 \gamma_{k-2s} + \cdots + \Phi_p \gamma_{k-ps}\,;\; k = 1,2,\ldots,Ps \tag{3.45}$$

The autocorrelation function (ACF) satisfies the difference equation.

$$\rho_k = \Phi_1 \rho_s + \Phi_2 \gamma_{2s} + \cdots + \Phi_p \rho_{ps}\,;\; k = 1,2,\ldots,Ps \tag{3.46}$$

The autocorrelation function (ACF) will be non-zero only lags that are integer multiples of s. The autocorrelation at seasonal lags persists indefinitely, although with declining intensity.

The First Order Seasonal Autoregressive SAR (1) Process

Consider the SAR (1) model (P=1)

$$\dot{Z}_t = \Phi_1 Z_{t-s} + a_t$$

Where, a_t 's are random shocks satisfying with usual assumptions.

The autocorrelation function of the SAR (1) process is obtained by substituting P=1 in Equation (3.45).

The Autocovariance Function is

$$\gamma_k = \Phi_1 \gamma_{k-s} \qquad ; k = 1,2,\ldots,Ps$$

The autocovariance function of the SAR (1) process is

$$\gamma_k = \begin{cases} \dfrac{\sigma_a^2}{1-\Phi_1^2} & ; k = 0 \\ \Phi_1^k \gamma_0 & ; k = s, 2s, 3s, \ldots \\ 0 & ; k = 0, s, 2s, 3s, \ldots \end{cases}$$

The autocorrelation function of the SAR (1) process is,

$$\rho_k = \begin{cases} 1 & ; k = 0 \\ \Phi_1^k & ; k = s, 2s, 3s, \ldots \\ 0 & ; k \neq 0, s, 2s, 3s, \ldots \end{cases}$$

Therefore, the autocovariance and the autocorrelation are non-zero at lags that are integer multiples of s.

The Second Order Seasonal Autoregressive SAR (2) Process

Consider SAR (2) model (P=2)

$$\dot{Z}_t = \Phi_1 \dot{Z}_{t-s} + \Phi_2 \dot{Z}_{t-2s} + a_t$$

where, $a_{t'}$ are random shocks satisfying with usual assumptions. The autocovariance function of the SAR (2) model is obtained by substituting P=2 in Equation (3.45).

The Autocovariance Function is

$$\gamma_k = \Phi_1 \gamma_{k-s} + \Phi_2 \gamma_{k-2s} \qquad ; k = 1,2,\ldots,Ps$$

Therefore, the autocovarinance function of the SAR (2) process is

$$\gamma_k = \begin{cases} \left(\dfrac{1-\Phi_2}{1+\Phi_2}\right)\left(\dfrac{\sigma_a^2}{(1-\Phi_2)^{2-}\Phi_1^2}\right) & ;k = 0 \\ \left(\dfrac{\Phi_1}{1-\Phi_1}\right)\gamma_0 & ;k = s \\ \left(\dfrac{\Phi_1^2}{1-\Phi_2} + \Phi_2\right)\gamma_0 & ;k = 2s \\ \Phi_1 \gamma_{k-s} + \Phi_2 \gamma_{k-2s} & ;k = 3s, 4s, \ldots \\ 0 & ;k = 0, s, 2s, 3s, \ldots \end{cases}$$

The autocorrelation function of the SAR (1) process is,

$$\rho_k = \begin{cases} 1 & ;k=0 \\ \dfrac{\Phi_1}{1-\Phi_2} & ;k=s \\ \dfrac{\Phi_1^{\,2}}{1-\Phi_2} + \Phi_2 & ;k=2s \\ \Phi_1\rho_{k-s} + \Phi_2\rho_{k-2s} & ;k=3s, 4s, \ldots \\ 0 & ;k \neq 0, s, 2s, 3s, \ldots \end{cases}$$

Therefore, the autocovariance and the autocorrelation function are non-zero at lags that are integer multiple of s.

3.7.2 General Multiplicative Seasonal Models

The fundamental fact about seasonal time series with period s, is that observation which are is intervals apart are similar. Therefore, one might expect that the operation $B^sX_t = X_{t-s}$ would play a particularly important role in the analysis of seasonal series, and furthermore, since nonstationarity is to be expected in the series $X_t, X_{t-s}, X_{t-2s}, \ldots.$, the simplifying operation $\nabla_s X_t = (1-B^s)X_t = X_t - X_{t-s}$ might be useful.

The seasonal effect implies that an observation for a particular quarter, say, second quarter, is related to the observation for second quarters of previous years. Suppose the t^th^ observation X_t is for the second quarters. We might be able to link this observation X_t to observations in second quarters of previous years by a model of the form

$$\Phi(B^S)\nabla_S^D X_t = \Theta(B^S)\alpha_t$$

Where s = 12, for monthly series and s = 4 for quarterly series. $\nabla_s = 1 - B^S$ and $\phi(B^S)$, $\Phi(B^S)$ are polynomials in B^S of degrees P and Q, respectively, and satisfying invertibility conditions. Similarly, a model

$$\Phi(B^S)\nabla_S^D X_{t-1} = \Theta(B^S)\alpha_{t-1}$$

might be used to link the current behavior for first quarter with previous first quarter observations, and so on, for each of the first quarters. Moreover, it would usually be reasonable to assume that the parameters Φ and Θ contained in these monthly models would be approximately the same for each quarter. Now the error components α_1, α_{t-1}, ... in these models would not in general be uncorrelated. For example, the value in last quarter, 2000, while related to previous last quarter values, would also be related to value in first, second and third quarters of 2000 etc. Thus we would expect that would be related to α_{t-1} and α_{t-2} etc. Therefore, to take care the such relationships, we introduce a second model

$$\phi(B)\nabla^d \alpha_t = \theta(B)a_t$$

where no a_t is a white noise process, and ϕ (B) and θ (B) are polynomials in B of degrees p and q, respectively, and satisfying stationary and invertible conditions and $\nabla = \nabla_1 = 1\text{-}B.$

We finally obtain a general multiplicative model

$$\phi(B)\Phi_P(B^s)\nabla^d \nabla_S^D X_t = \theta_q(B)\Theta_Q(B^s)\, a_t$$

Where for this particular example, s = 12 for monthly series and s = 4 for quarterly series. Also the subscripts p, P, q, Q have been added to remind the orders of the various operators. The resulting multiplicative process will be said to be of order (p,d,q) x $(P,D,Q)_s$. A similar argument can be used to obtain models with three or more periodic components to take care of multiple seasonalities.

3.7.3 ACF and PACF for Seasonal Models

In identifying seasonal time series, the standard ACF analysis is still the most useful method. ACF and PACF for seasonal models are more complicated. In general, the season and nonseasonal autoregressive components have their PACF cutting off at the seasonal and nonseasonal lags. On the other hand, the seasonal and nonseasonal moving average components produce PACF which shows exponential decays and / or damped sine waves at the seasonal and nonseasonal lags.

3.7.3 Model Building and Forecasting for Seasonal Models

For identification, estimation and diagnostic checking of model of seasonal data, no new principles are needed to do this, but merely and application of procedure and ideas we have already discussed in detail for non-seasonal data.

The most important aspect in the conduct of time series analysis is the use of past and present data or available observations to predict future values. The use available observations at time "t" to predict or forecast values at some future time "t+L" can serve many purpose of economic and business planning. The short term forecasts will be performed by using adequately fitted model based on the results of model building. The actual values will also be applied for the validation of the forecast values of forecasting period.

Because seasonal models are special forms of the ARIMA described in sections 3.6, the model identification, parameter estimation, diagnostic checking and forecasting for these models follow the same general methods introduced in section 3.6.1, 3.6.2, 3.6.3 and 3.6.4.

CHAPTER IV

RESULTS AND FINDINGS

The seasonal variation of production series in HISEM Co., Ltd from January 2013 to December 2017 is measured by seasonal index. The analysis is done by the Ratio to Moving Average method.

4.1 Test of Seasonality

Test of seasonality for monthly production series from January 2013 to December 2017 are calculated in the following.

Production series for 100 Kilo Volt Ampere

The result of calculation for testing the seasonality in the production series for 100 KVA (2013-2017) are shown in Table (4.1).

Table (4.1)

ANOVA Table for Production Series for 100 KVA (2013-2017)

Source of variation	Sum of square	Degree of Freedom	Mean Square Error	F-Ratio
Between Months	498.85	11	45.35	4.1151
Between Years	1477.5	4	369.375	
Error	484.9	44	11.0205	
Total	2461.25	59		

At 5% level of significance, the critical value K= $F_{(0.05,11,44)}$ is 2.01. Since the computed F-value = 4.1151 is greater than K= 2.01, it can be calculated that the monthly data of production series for 100 KVA exists seasonality.

Production series for 160 Kilo Volt Ampere

The result of calculation for testing the seasonality in the production series for 160 KVA (2013-2017) are shown in Table (4.2).

Table (4.2)

ANOVA Table for Production Series for 160 KVA (2013-2017)

Source of variation	Sum of square	Degree of Freedom	Mean Square Error	F-Ratio
Between Months	399.73	11	36.34	2.2529
Between Years	1785.67	4	446.42	
Error	709.93	44	16.13	
Total	2895.33	59		

At 5% level of significance, the critical value K= $F_{(0.05,11,44)}$ is 2.01. Since the computed F-value = 2.2529 is greater than K= 2.01, it can be calculated that the monthly data of production series for 160 KVA exists seasonality.

Production series for 400 Kilo Volt Ampere

The result of calculation for testing the seasonality in the production series for 400 KVA (2013-2017) are shown in Table (4.3).

Table (4.3)

ANOVA Table for Production Series for 400 KVA (2013-2017)

Source of variation	Sum of square	Degree of Freedom	Mean Square Error	F-Ratio
Between Months	721.38	11	65.58	4.36
Between Years	1555.6	4	388.9	
Error	661.20	44	15.03	
Total	2938.18	59		

At 5% level of significance, the critical value K= $F_{(0.05,11,44)}$ is 2.01. Since the computed F-value = 4.36 is greater than K= 2.01, it can be calculated that the monthly data of production series for 400 KVA exists seasonality.

Production series for 2000 Kilo Volt Ampere

The result of calculation for testing the seasonality in the production series for 2000 KVA (2013-2017) are shown in Table (4.4).

Table (4.4)

ANOVA Table for Production Series for 2000 KVA (2013-2017)

Source of variation	Sum of square	Degree of Freedom	Mean Square Error	F-Ratio
Between Months	631.4	11	57.4	6.18
Between Years	503.07	4	125.77	
Error	408.93	44	9.29	
Total	1543.4	59		

At 5% level of significance, the critical value K= $F_{(0.05,11,44)}$ is 2.01. Since the computed F-value = 6.18 is greater than K= 2.01, it can be calculated that the monthly data of production series for 2000 KVA exists seasonality.

4.2 Seasonal Variation

The seasonal variation of monthly production series from 2013 to 2017 are computed by the ratio to moving averages method.

Production Series for 100 Kilo Volt Ampere (2013-2017)

The seasonal variation of monthly production series from 2013 to 2017 is computed by the ratio to moving average method under multiplicative decomposition of time series. The series consists of 60 observations and it was shown in Appendix A. The results of seasonal index are shown in Table (4.5). The lowest value of seasonal index is in September and the highest is in November. The months of February, June, October and December have the larger seasonal indexes than other months. The peak period is November with the seasonal index of production series as 113 % while September the lowest month with 87%.

Table (4.5)

Seasonal Indexes for production series for 100 KVA by using the Ratio to Moving Average Method (2013-2017)

Month	Seasonal Index
January	95
February	105
March	99
April	92
May	98
June	108
July	89
August	99
September	87
October	106
November	113
December	109

Production Series for 160 Kilo Volt Ampere (2013-2017)

The seasonal variation of monthly production series from 2013 to 2017 is computed by the ratio to moving average method under multiplicative decomposition of time series. The series consists of 60 observations and it was shown in Appendix A. The results of seasonal index are shown in Table (4.6). The lowest value of seasonal index is in August and the highest is in December. The months of February, June, October and November have the larger seasonal indexes than other months. The peak period is December with the seasonal index of production series as 114 % while August the lowest month with 87%.

Table (4.6)

Seasonal Indexes for production series for 160 KVA by using the Ratio to Moving Average Method (2013-2017)

Month	Seasonal Index
January	93
February	108
March	98
April	97
May	96
June	105
July	93
August	87
September	92
October	107
November	110
December	114

Production Series for 400 Kilo Volt Ampere (2013-2017)

The seasonal variation of monthly production series from 2013 to 2017 is computed by the ratio to moving average method under multiplicative decomposition of time series. The series consists of 60 observations and it was shown in Appendix A. The results of seasonal index are shown in Table (4.7). The lowest value of seasonal index is in September and the highest is in November. The months of February, March, June, October and December have the larger seasonal indexes than other months. The peak period is November with the seasonal index of production series as 124 % while September the lowest month with 82%.

Table (4.7)

Seasonal Indexes for production series for 400 KVA by using the Ratio to Moving Average Method (2013-2017)

Month	Seasonal Index
January	86
February	103
March	100
April	88
May	99
June	109
July	90
August	90
September	82
October	116
November	124
December	113

Production Series for 2000 Kilo Volt Ampere (2013-2017)

The seasonal variation of monthly production series from 2013 to 2017 is computed by the ratio to moving average method under multiplicative decomposition of time series. The series consists of 60 observations and it was shown in Appendix A. The results of seasonal index are shown in Table (4.8). The lowest value of seasonal index is in August and the highest is in November. The months of January, May, June, October and December have the larger seasonal indexes than other months. The peak period is November with the seasonal index of production series as 136% while August the lowest month with 82%.

Table (4.8)

Seasonal Indexes for production series for 2000 KVA by using the Ratio to Moving Average Method (2013-2017)

Month	Seasonal Index
January	101
February	89
March	92
April	93
May	100
June	96
July	95
August	82
September	94
October	109
November	136
December	113

4.3 The Box-Jenkins Seasonal ARIMA Model of Production Series for 100 KVA

The monthly data of production series for 100 KVA covers 5 years, from January, 2013 to December 2017. The series consists of 60 observations.

4.3.1 Identification

For the identification of the order p and q, the (autocorrelation function) ACF and (partial autocorrelation function) PACF of the number of production series for 100 KVA are computed and plotted as shown in the following Tables and Figures.

Table (4.9)

Estimated Autocorrelation Function for the original series of Production for 100 Kilo Volt Ampere

$\hat{\rho}_k$ for $\{Z_t\}$ $\bar{Z}$ =30.25 S_z =6.459 n=60

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.643	.503	.376	.378	.307	.390	.252	.193	.134	.268	.294	.345
S.E	.126	.125	.124	.123	.122	.120	.119	.118	.117	.116	.115	.114
13-24	.161	.081	.045	.053	.024	.052	-.071	-.062	-.067	-.067	-.108	-.102
S.E	.112	.111	.110	.109	.108	.106	.105	.104	.102	.101	.100	.098

Figure (4.1)

Sample Autocorrelation Function for Monthly Production Series for 100 Kilo Volt Ampere

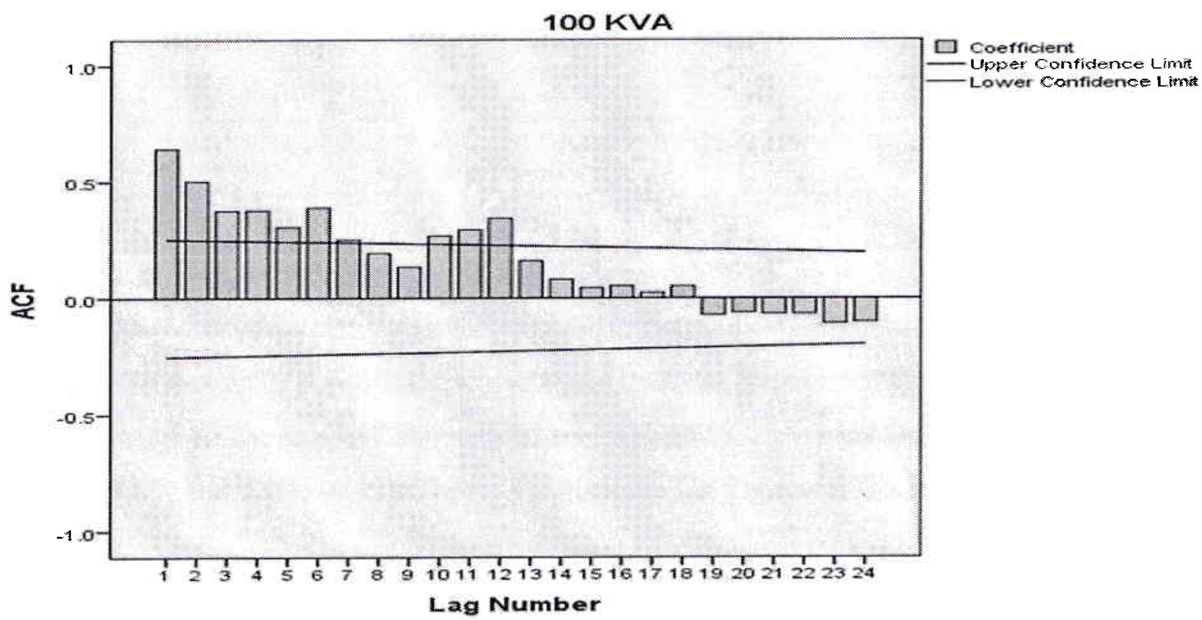


Table (4.10)

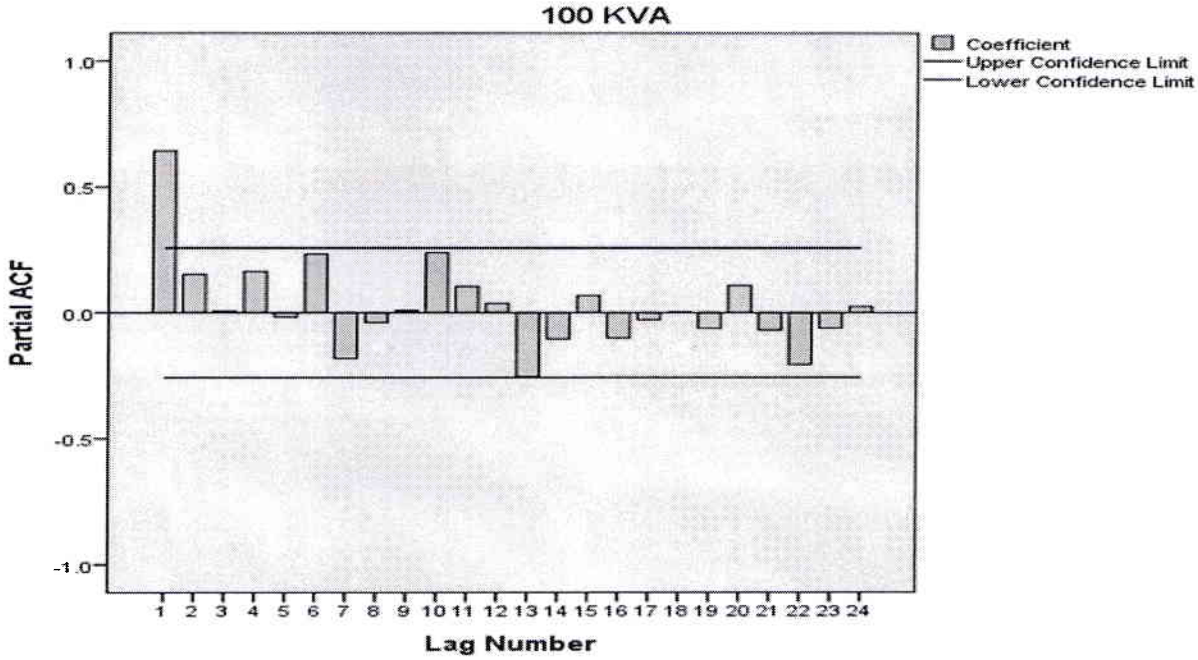
Estimated Partial Autocorrelation Function for the original series of Production for 100 Kilo Volt Ampere

$\hat{\emptyset}_{kk}$ for $\{Z_t\}$ $\bar{Z} = 30.25$ $S_z = 6.459$ n= 60

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.643	.153	.007	.165	-.017	.233	-.182	-.038	.010	.239	.106	.038
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129
13-24	-.254	-.103	.069	-.101	-.028	.001	-.063	.109	-.070	-.206	-.061	.024
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129

Figure (4.2)

Sample Partial Autocorrelation Function for Monthly Production Series for 100 Kilo Volt Ampere



The sample ACF decays slowly and the sample PACF has a single large spike at lag 1. These values indicated that the series is nonstationary and that differencing is called for. To remove nonstationary, the series is seasonal differenced and the sample ACF and PACF of the seasonal differenced series (1-B^{12})Z_t were computed as shown in Table (4.11) and Table (4.12). They were displayed in Figure (4.3) and Figure (4.4).

Table (4.11)

Estimated Autocorrelation Function for Seasonal First Difference Series of Production for 100 Kilo Volt Ampere

$\hat{\rho}_k$ for $\{W_t = (1 - B^{12})Z_t\}$ $\bar{W} = 2.87$ $S_w = 5.659$ n=48

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.424	.266	.299	.297	.069	.163	.248	.018	-.059	.188	.224	-.123
S.E	.140	.138	.137	.135	.134	.132	.131	.129	.127	.126	.124	.122
13-24	.028	.017	-.029	-.041	.060	.021	-.159	-.065	.041	-.071	-.219	-.211
S.E	.121	.119	.117	.115	.114	.112	.110	.108	.106	.104	.102	.100

Figure (4.3)

Sample Autocorrelation Function for Seasonal First Difference Series of Production Series for 100 Kilo Volt Ampere

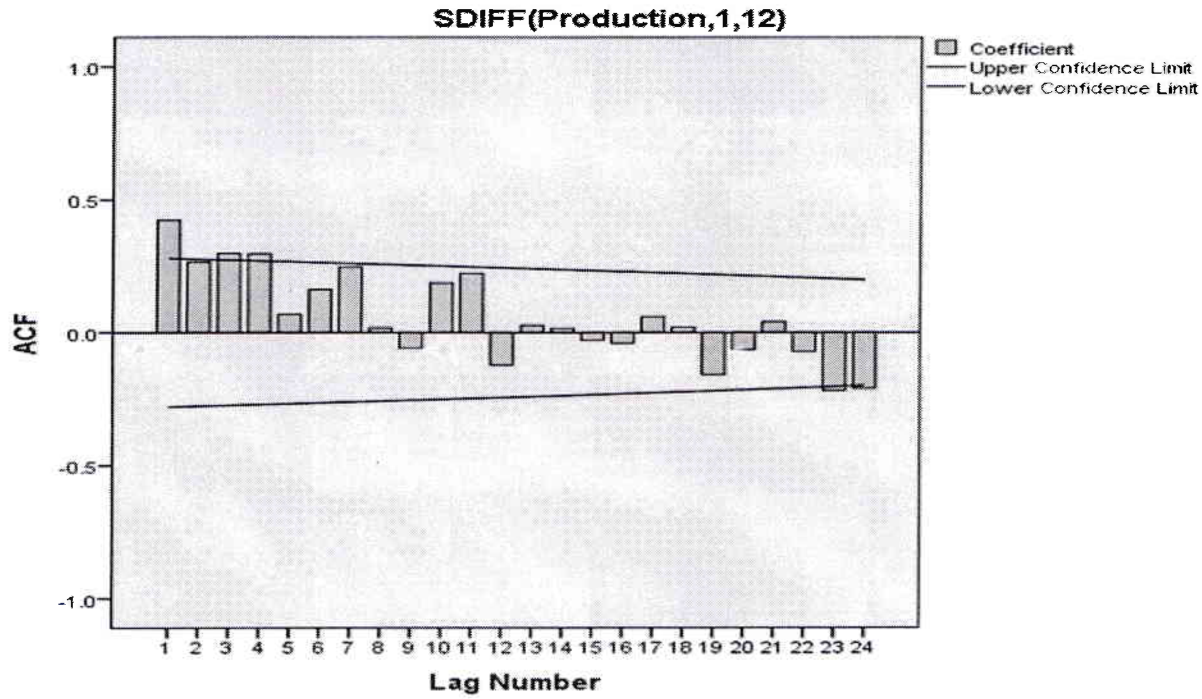


Table (4.12)

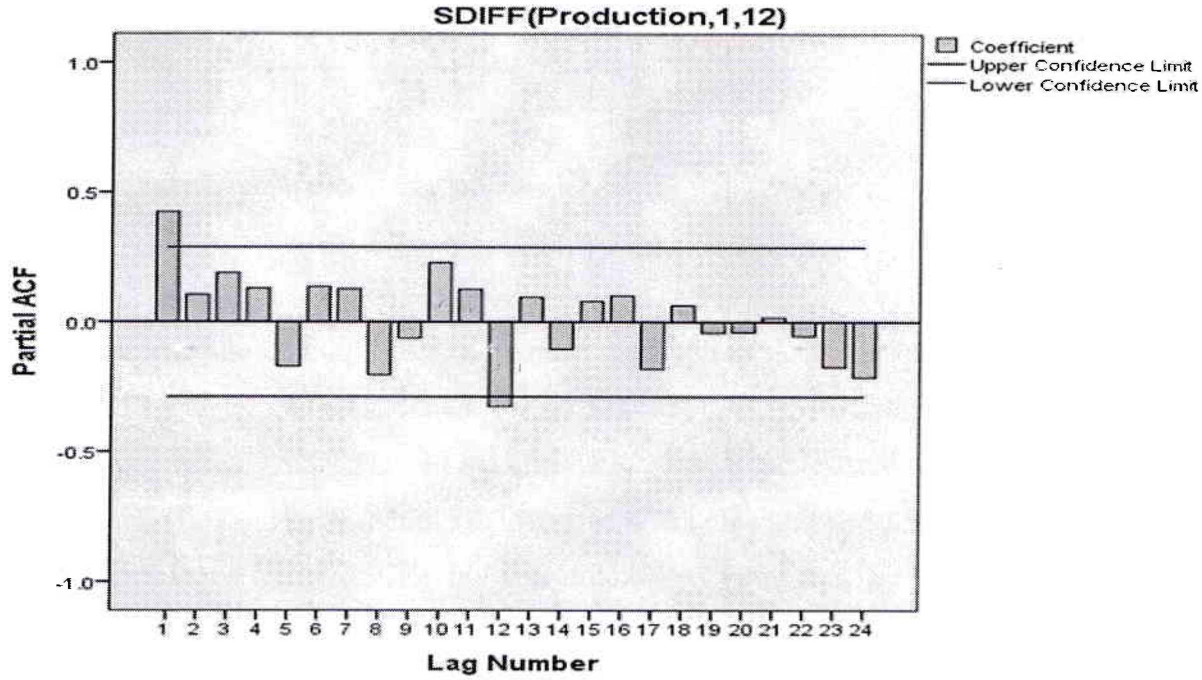
Estimated Partial Autocorrelation Function for Seasonal First Difference Series of Production for 100 Kilo Volt Ampere

$\hat{\emptyset}_{kk}$ for $\{W_t = (1 - B^{12})Z_t\}$ $\overline{W} = 2.87$ $S_w = 5.659$ n=48

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.424	.105	.189	.129	-.172	.135	.128	-.205	-.062	.229	.127	-.327
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144
13-24	.096	-.105	.079	.102	-.183	.064	-.041	-.040	.017	-.054	-.174	-.213
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144

Figure (4.4)

Sample Partial Autocorrelation Function for Seasonal First Difference Series of Production Series for 100 Kilo Volt Ampere



The sample ACF is slowly decay and the sample PACF cuts off after lag 1 because none of the sample PACF value is significant expect that lag 12.

The suggested series (1-B^{12})Z_t might be described by SAR (1) process as a tentative model for the series

Since $\bar{W}$ = 2.87, S_w=5.659, n= 48

The t value of t = $\frac{\bar{W}}{S_w/\sqrt{n}} = \frac{2.87}{5.659/\sqrt{48}} = 3.5136$

Which is significant and thus deterministic trend θ_0 is needed. Hence, the tentative model for the series following SAR(1) process:

$$(1-\Phi B^{12})Z_t = \theta_0 + a_t$$

4.3.2 Parameter Estimation for SAR (1) model

Using SAR (1) model, the estimated parameters with their statistics were shown in Table (4.13). According to this table, the estimated parameter of Φ is 0.525, since their p-value is 0.000, there is evidence to reject the null hypothesis: Φ = 0.

Table (4.13)

Estimated Parameters and Model Statistics for SAR(1) Model of Production Series for 100 KVA

	Estimate	SE	t	Sig.
Constant	29.413	1.318	22.321	0.000
Φ	0.525	0.122	4.315	0.000

The following estimated model was obtained

$$(1\text{-}0.525B^{12})Z_t = 29.413 + a_t$$

(0.122) (1.318)

The estimation of the SAR (1) model of production series for 100 KVA give $\theta_0 = 29.413$ with estimated standard error 1.318 and $\Phi = 0.525$ with the estimated standard error 0.122. Under the null hypothesis H_0: $\Phi = 0$ the test statistics t is 4.315 with p-value is 0.000. Hence, there is evidence to reject the null hypothesis.

Moreover, the sample ACFs and the sample PACFs of residual for the above tentative model were shown in Table (4.14) and (4.15), respectively. They were showed in Figure (4.5) and (4.6).

Table (4.14)

Estimated Autocorrelation Function of Residual for SAR(1) Model of Production Series for 100 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.620	.483	.411	.372	.253	.287	.220	.091	.001	.136	.123	-.008
S.E	.129	.172	.193	.207	.218	.223	.229	.232	.233	.233	.234	.235
13-24	-.031	-.076	-.102	-.108	-.108	-.097	-.234	-.190	-.165	-.190	-.231	-.223
S.E	.235	.236	.236	.237	.237	.238	.239	.243	.245	.247	.249	.253

Figure (4.5)

Sample Autocorrelation Function of Residual values for SAR(1) Model of Production Series for 100 KVA

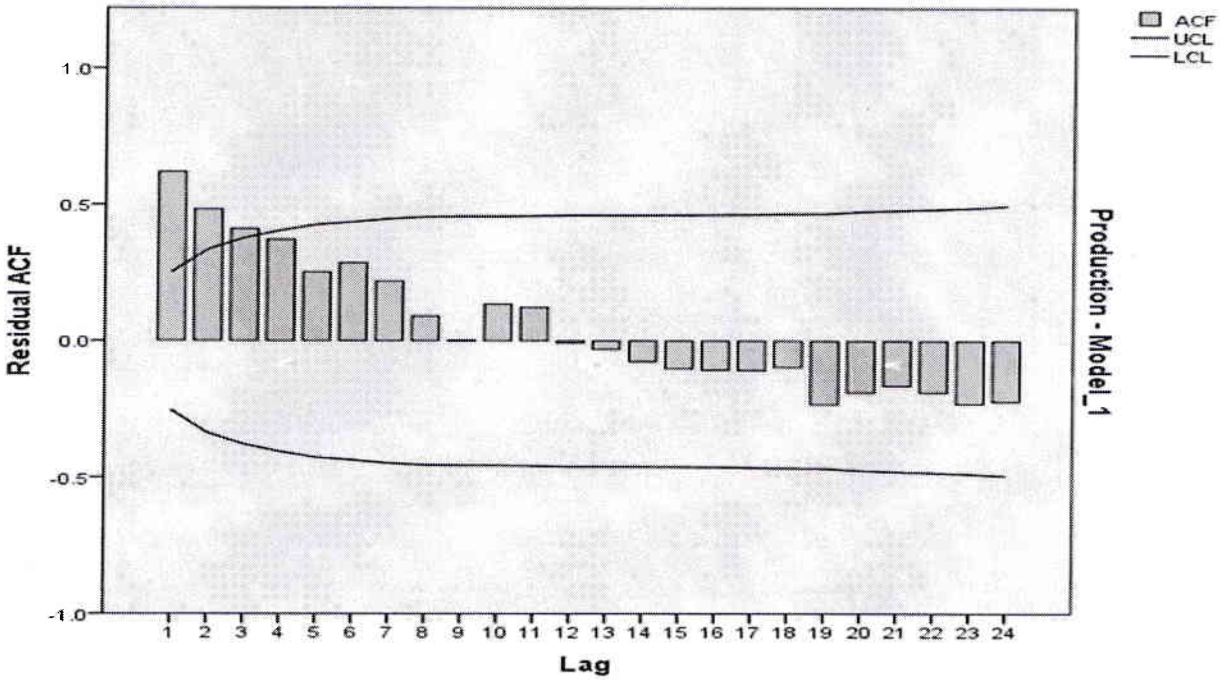


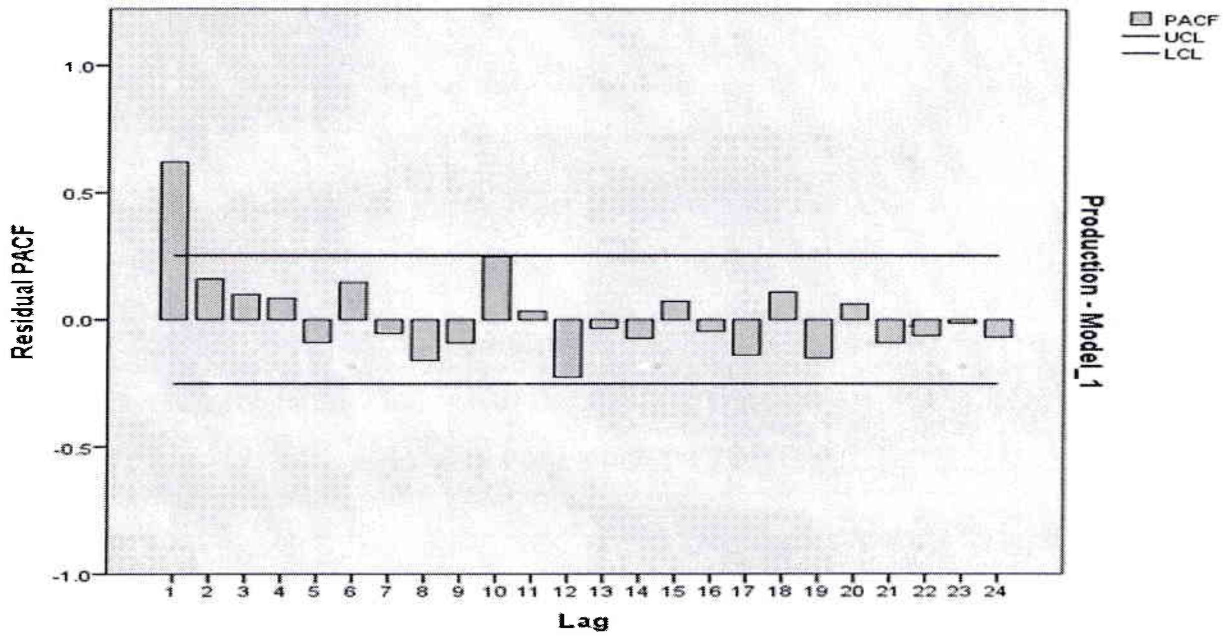
Table (4.15)

Estimated Partial Autocorrelation Function of Residual for SAR(1) Model of Production Series for 100 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.620	.160	.100	.085	-.089	.147	-.053	-.160	-.092	.249	.034	-.226
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129
13-24	-.034	-.073	.073	-.045	-.139	.110	-.152	.063	-.091	-.062	-.012	-.068
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129

Figure (4.6)

Sample Partial Autocorrelation Function of Residual values for SAR(1) Model of Production Series for 100 KVA



The Sample ACF was exponential decay and the sample PACF cuts off after lag 1 and these model exhibit a pattern. So, the residual series are not white noise process. Since, another tentative seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ model considered, that is

$$(1\text{-}\Phi B^{12})Z_t = \theta_0 + (1 - \emptyset B)a_t$$

Using multiplicative seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ model, the estimated parameters with their statistics were shown in Table (4.16). According to this table, the estimated parameter of $\emptyset$ is 0.568 since their p-value is 0.000, there is evidence to reject the null hypothesis: $\emptyset = 0$ and the estimated parameter of Φ is -0.378, since their p-value is 0.011, there is evidence to reject the null hypothesis: $\Phi = 0$.

Table (4.16)

Estimated Parameters and Model Statistics for seasonal ARIMA (1, 0, 0) x $(1, 1, 0)_{12}$ Model of Production for 100 KVA

	Estimate	SE	t	Sig.
Constant	2.880	1.217	2.367	0.022
$\emptyset$	0.568	0.131	4.350	0.000
Φ	-0.378	0.143	-2.642	0.011

4.3.3 Diagnostic Checking

To check model adequacy, in Table (4.17) and Table(4.18) was shown the residual ACF and PACF of the modified model. They were shown in Figure (4.7) and (4.8), along with the confidence interval.

$$\gamma_k(\hat{a}_t) \pm 2\widehat{S.E}[\gamma_k(\hat{a}_t)]$$

Where,

$$\widehat{S.E}[\gamma_k(\hat{a}_t)] = \frac{1}{\sqrt{n}}$$

Table (4.17)

Estimated Autocorrelation Function of Residual for seasonal ARIMA (1, 0, 0) x $(1, 1, 0)_{12}$ Model of Production for 100 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	-.048	.014	.000	.207	-.096	.133	.161	-.047	-.198	.229	.256	-.051
S.E	.144	.145	.145	.145	.151	.152	.154	.158	.158	.163	.170	.178
13-24	.037	-.029	-.046	-.006	.035	.093	-.180	-.062	.077	.081	-.066	-.166
S.E	.178	.178	.178	.178	.178	.179	.180	.183	.184	.184	.185	.186

Figure (4.7)

Sample Autocorrelation Function of Residual values for seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ Model of Production for 100 KVA

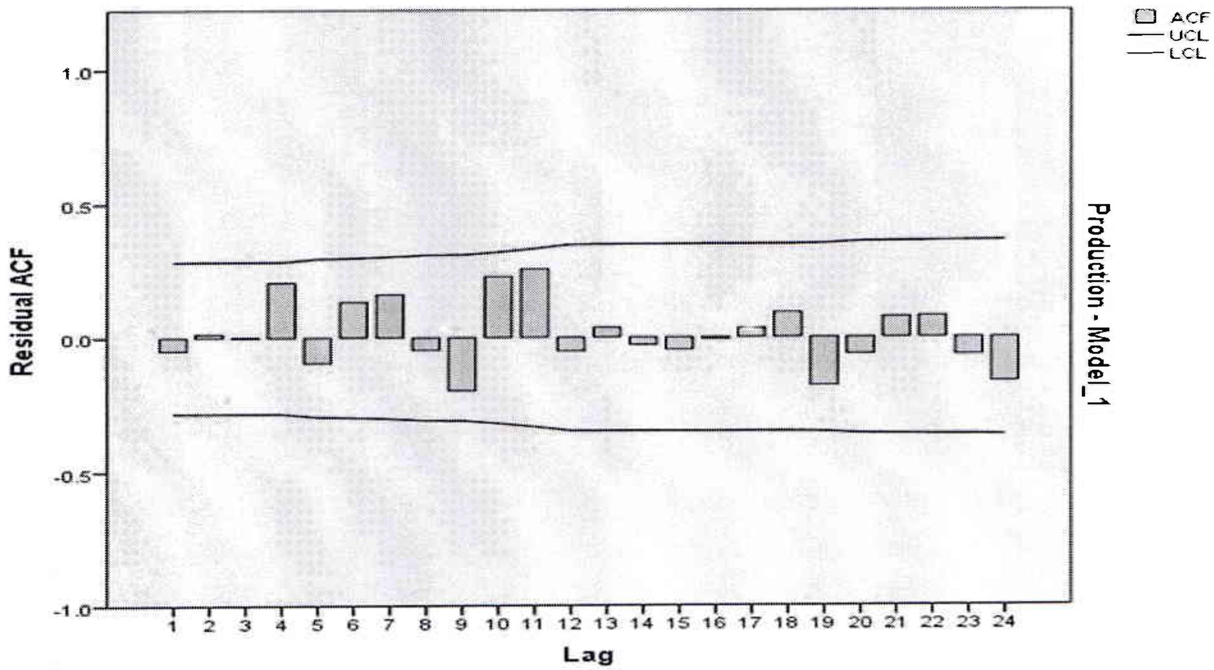


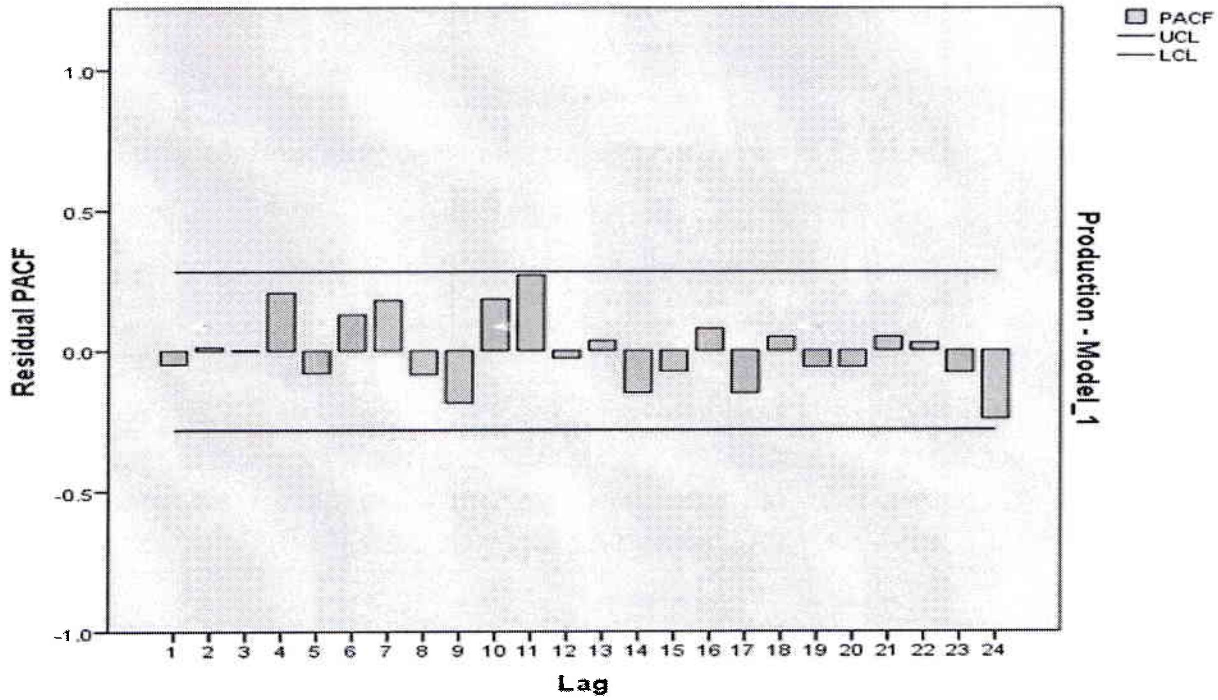
Table (4.18)

Estimated Partial Autocorrelation Function of Residual for seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ Model of Production for 100 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	-.048	.012	.001	.207	-.079	.128	.179	-.084	-.186	.184	.270	-.026
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144
13-24	.035	-.150	-.073	.078	-.151	.049	-.058	-.058	.049	.027	-.077	-.244
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144

Figure (4.8)

Sample Partial Autocorrelation Function of Residual values for seasonal ARIMA (1, 0, 0) x $(1, 1, 0)_{12}$ Model of Production for 100 KVA



Values of the residual ACF of seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ are all small and exhibit no patterns. And, the values of residual PACF of modified model lie inside the confidence limits. This suggested that this model adequate. Hence, the autocorrelation of $\hat{a_t}$ can be taken as significant different from zero.

An overall check is performed by using the test statistic,

$$Q = n \sum_{k=1}^{k} \gamma_k^2 (\hat{a}_t)$$

As the result of p value, the observed value of Q is 16.864 and it is not significant at 5 % significant level p-value is 0.394.

Thus, the fitted seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ model is judged adequate for the series.

4.4 The Box-Jenkins Seasonal ARIMA Model of Production Series for 160 KVA

The monthly data of production series for 160 KVA covers 5 years, from January, 2013 to December 2017. The series consists of 60 observations.

4.4.1 Identification

For the identification of the order p and q, the (autocorrelation function) ACF and (partial autocorrelation function) PACF of the number of production series for 160 KVA are computed and plotted as shown in the following Tables and Figures.

Table (4.19)

Estimated Autocorrelation Function for the original series of Production for 160 Kilo Volt Ampere

$\hat{\rho}_k$ for $\{Z_t\}$ $\bar{Z}$ =25.33 S_z =7.005 n=60

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.656	.587	.450	.466	.354	.308	.165	.149	.098	.081	.019	-.048
S.E	.126	.125	.124	.123	.122	.120	.119	.118	.117	.116	.115	.114
13-24	-.019	-.065	-.076	-.077	-.026	.025	.039	.030	.034	.015	.065	.087
S.E	.112	.111	.110	.109	.108	.106	.105	.104	.102	.101	.100	.098

Figure (4.9)

Sample Autocorrelation Function for Monthly Production Series for 160 Kilo Volt Ampere

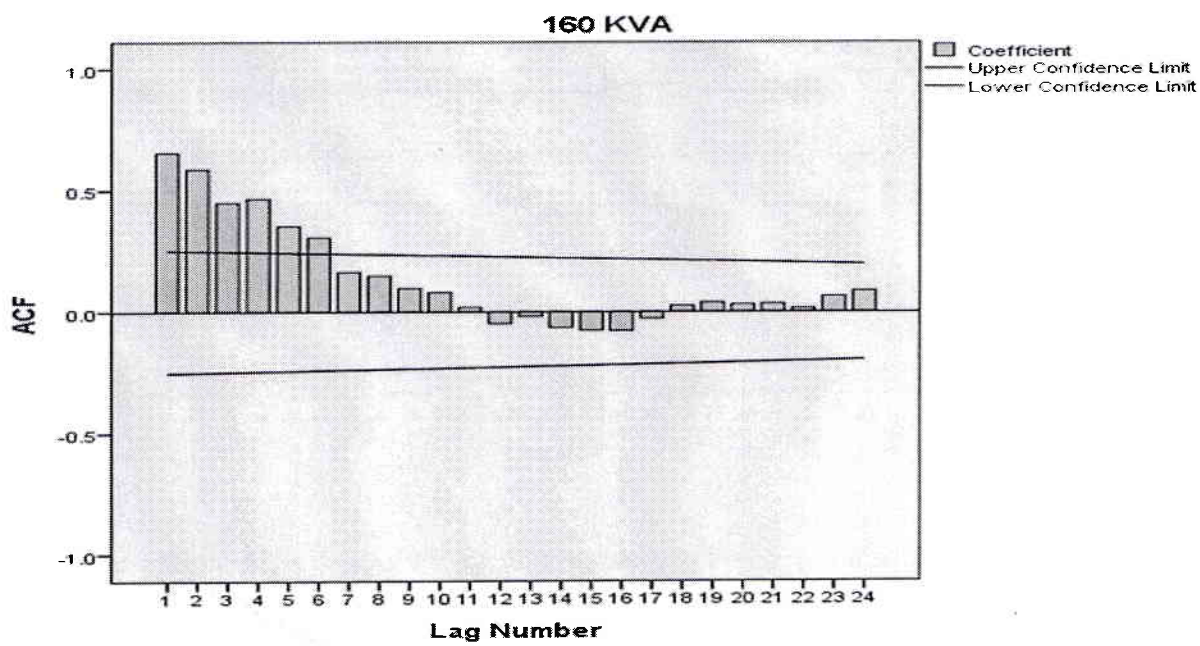


Table (4.21)

Estimated Autocorrelation Function for Seasonal First Difference Series of Production for 160 Kilo Volt Ampere

$\hat{\rho}_k$ for $\{W_t = (1 - B^{12})Z_t\}$ $\bar{W} = 4.21$ $S_w = 8.098$ n=48

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.426	.412	.286	.312	.239	.112	-.069	-.025	-.042	-.077	-.194	-.442
S.E	.140	.138	.137	.135	.134	.132	.131	.129	.127	.126	.124	.122
13-24	-.152	-.237	-.132	-.154	-.100	-.107	-.025	-.115	-.010	-.125	-.103	-.105
S.E	.121	.119	.117	.115	.114	.112	.110	.108	.106	.104	.102	.100

Figure (4.11)

Sample Autocorrelation Function for Seasonal First Difference Series of Production Series for 160 Kilo Volt Ampere

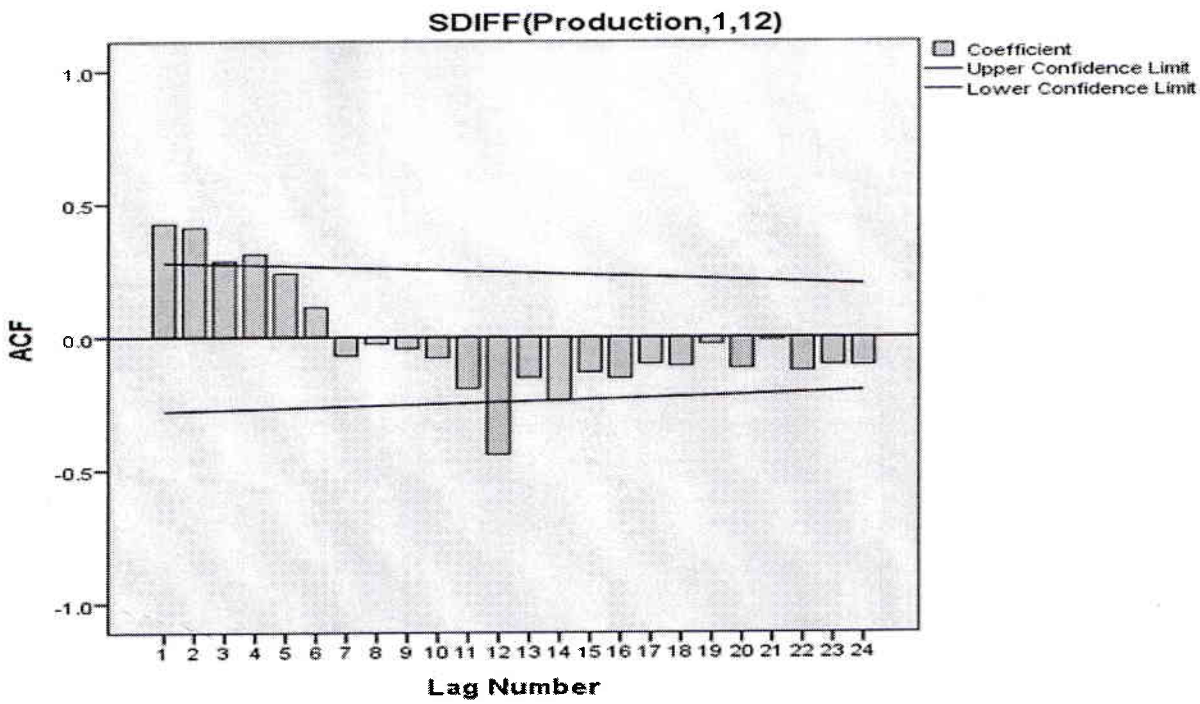


Table (4.22)

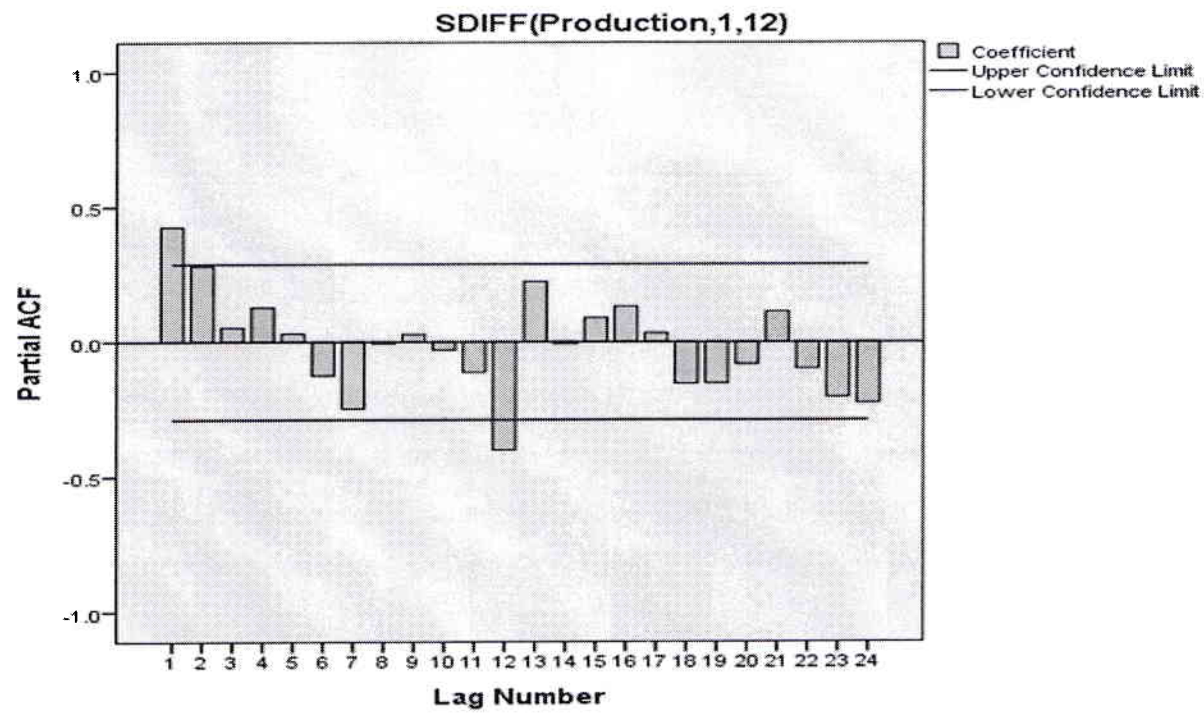
Estimated Partial Autocorrelation Function for Seasonal First Difference Series of Production for 160 Kilo Volt Ampere

$\hat{\emptyset}_{kk}$ for $\{W_t = (1 - B^{12})Z_t\}$ $\bar{W} = 4.21$ $S_w = 8.098$ n=48

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.426	.282	.053	.129	.031	-.124	-.247	-.006	.029	-.029	-.113	-.400
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144
13-24	.225	-.008	.090	.132	.035	-.154	-.152	-.081	.114	-.097	-.203	-.225
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144

Figure (4.12)

Sample Partial Autocorrelation Function for Seasonal First Difference Series of Production Series for 160 Kilo Volt Ampere



The sample ACF is damped sine wave and the sample PACF cuts off after lag 1 because none of the sample PACF value is significant expect that lag 12.

The suggested series (1-B^{12})Z_t might be described by SAR (1) process as a tentative model for the series

Since $\bar{W} = 4.21$, S_w=8.098, n= 48

The t value of $t = \frac{\bar{W}}{S_w/\sqrt{n}} = \frac{4.21}{8.098/\sqrt{48}} = 3.6018$

Which is significant and thus deterministic trend θ_0 is needed. Hence, the tentative model for the series following SAR(1) process:

$$(1\text{-}\Phi B^{12})Z_t = \theta_0 + a_t$$

4.4.2 Parameter Estimation for SAR (1) model

Using SAR (1) model, the estimated parameters with their statistics were shown in Table (4.23). According to this table, the estimated parameter of Φ is -0.103, since their p-value is 0.559, there is no evidence to reject the null hypothesis: $\Phi = 0$.

Table (4.23)

Estimated Parameters and Model Statistics for SAR(1) Model of Production Series for 160 KVA

	Estimate	SE	t	Sig.
Constant	25.344	0.874	29.002	0.000
Φ	-0.103	0.175	-0.587	0.559

The following estimated model was obtained

$$(1+ 0.103B^{12})Z_t = 25.344 + a_t$$

$$(0.175) \qquad\qquad (0.874)$$

The estimation of the SAR (1) model of production series for 160 KVA give $\theta_0 = 25.344$ with estimated standard error 0.874 and $\Phi = -0.103$ with the estimated standard error 0.175. Under the null hypothesis H_0: $\Phi = 0$ the test statistics t is -0.587 with p-value is 0.559. Hence, there is no evidence to reject the null hypothesis.

Moreover, the sample ACFs and the sample PACFs of residual for the above tentative model were shown in Table(4.24) and (4.25), respectively. They were showed in Figure (4.13) and (4.14).

Table (4.24)

Estimated Autocorrelation Function of Residual for SAR(1) Model of Production Series for 160 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.661	.591	.453	.473	.363	.328	.192	.179	.127	.115	.065	.023
S.E	.129	.177	.207	.223	.239	.248	.255	.258	.260	.261	.262	.262
13-24	.022	-.031	-.058	-.059	-.012	.042	.043	.041	.039	.027	.078	.098
S.E	.262	.262	.262	.262	.262	.262	.263	.263	.263	.263	.263	.263

Figure (4.13)

Sample Autocorrelation Function of Residual values for SAR(1) Model of Production Series for 160 KVA

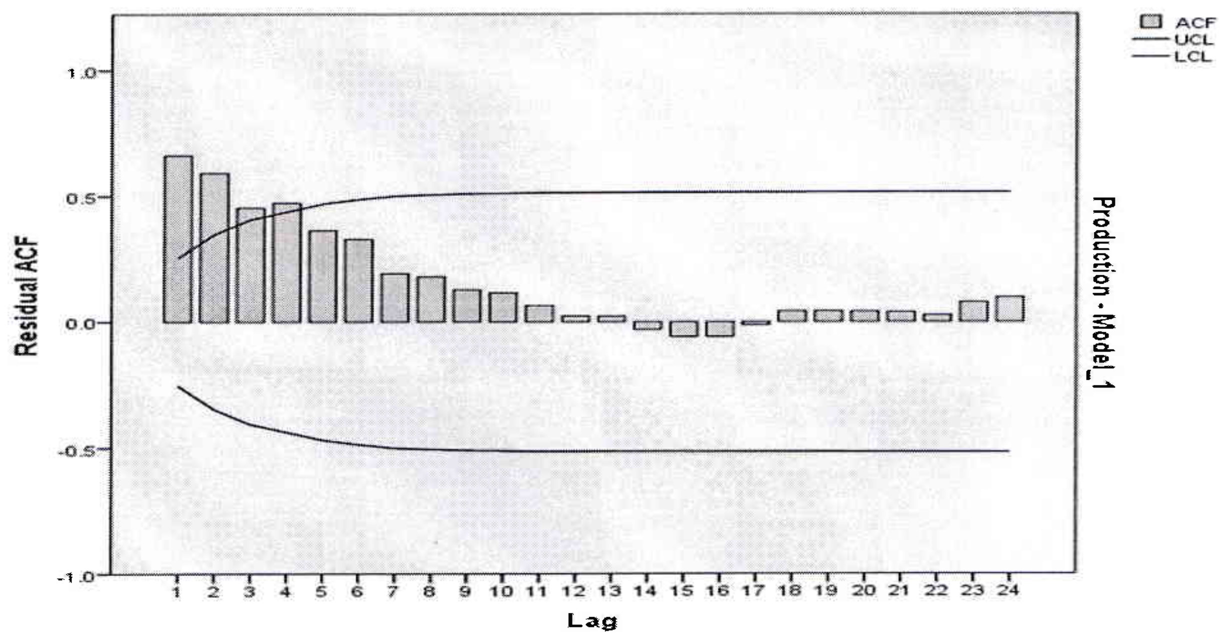


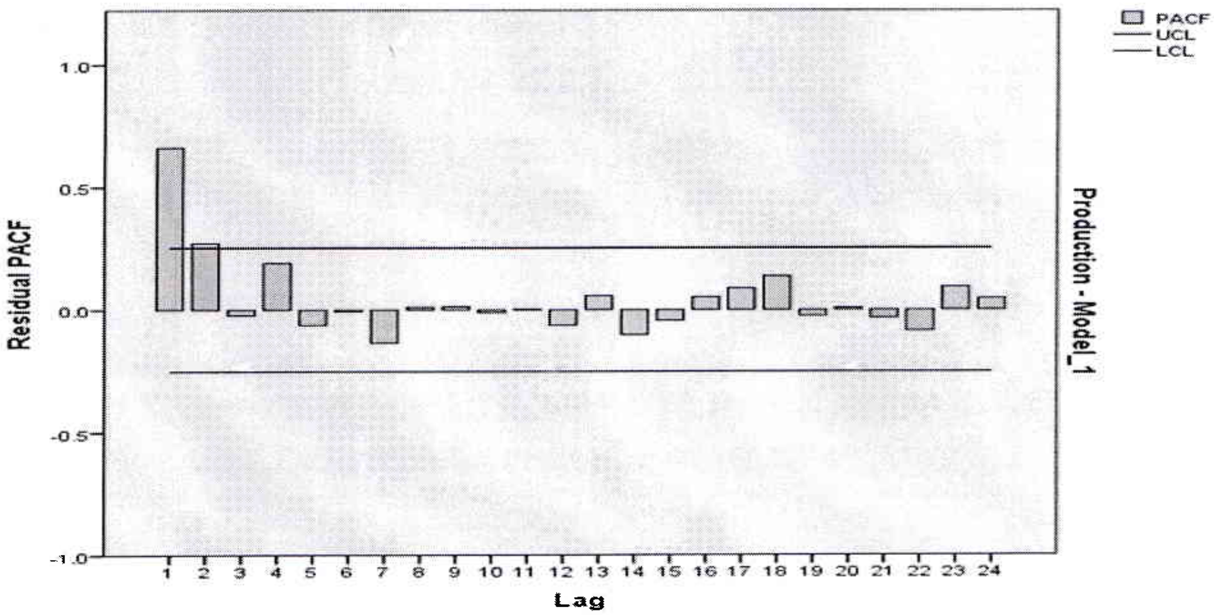
Table (4.25)

Estimated Partial Autocorrelation Function of Residual for SAR(1) Model of Production Series for 160 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.661	.273	-.023	.192	-.062	-.007	-.135	.011	.013	-.011	.003	-.062
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129
13-24	.058	-.102	-.045	.052	.088	.138	-.024	.008	-.034	-.086	.094	.047
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129

Figure (4.14)

Sample Partial Autocorrelation Function of Residual values for SAR(1) Model of Production Series for 160 KVA



The Sample ACF was exponential decay and the sample PACF cuts off after lag 1 and these model exhibit a pattern. So, the residual series are not white noise process. Since, another tentative seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ model considered, that is

$$(1\text{-}\Phi B^{12})Z_t = \theta_0 + (1 - \emptyset B)a_t$$

Using multiplicative seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ model, the estimated parameters with their statistics were shown in Table (4.26). According to this table, the estimated parameter of $\emptyset$ is 0.451 since their p-value is 0.002, there is evidence to reject the null hypothesis: $\emptyset = 0$ and the estimated parameter of Φ is -0.651, since their p-value is 0.000, there is evidence to reject the null hypothesis: $\Phi = 0$.

Table (4.26)

Estimated Parameters and Model Statistics for seasonal ARIMA (1, 0, 0) x $(1, 1, 0)_{12}$Model of Production for 160 KVA

	Estimate	SE	t	Sig.
Constant	3.063	1.002	3.057	0.004
$\emptyset$	0.451	0.135	3.331	0.002
Φ	-0.651	0.129	-5.068	0.000

4.4.3 Diagnostic Checking

To check model adequacy, in Table (4.27) and Table (4.28) was shown the residual ACF and PACF of the modified model. They were shown in Figure (4.15) and (4.16), along with the confidence interval.

$$\gamma_k(\hat{a}_t) \pm 2\widehat{S.E}[\gamma_k(\hat{a}_t)]$$

Where,

$$\widehat{S.E}[\gamma_k(\hat{a}_t)] = \frac{1}{\sqrt{n}}$$

Table (4.27)

Estimated Autocorrelation Function of Residual for seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ Model of Production for 160 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	-.117	.210	.035	.234	.141	.021	-.069	.052	.122	.033	-.061	-.214
S.E	.144	.146	.152	.153	.160	.163	.163	.163	.164	.165	.166	.166
13-24	.099	-.159	-.109	-.102	-.013	-.046	-.014	-.182	.167	-.160	.025	-.225
S.E	.172	.173	.176	.177	.178	.178	.179	.179	.183	.186	.189	.189

Figure (4.15)

Sample Autocorrelation Function of Residual values for seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ Model of Production for 160 KVA

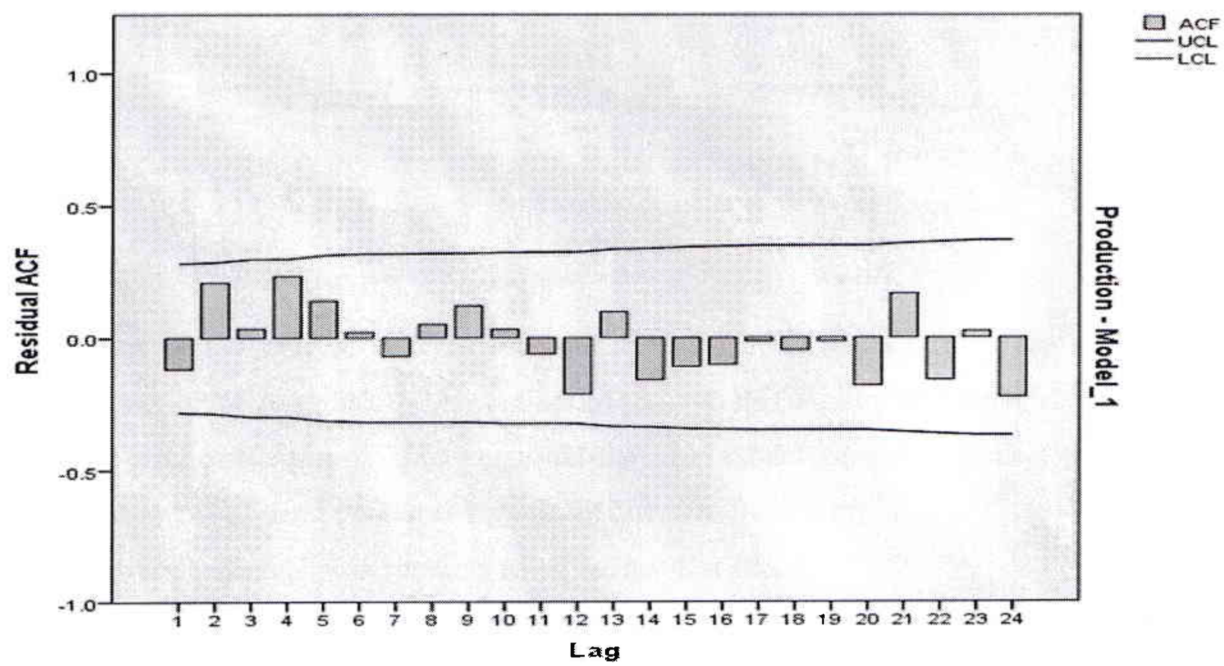


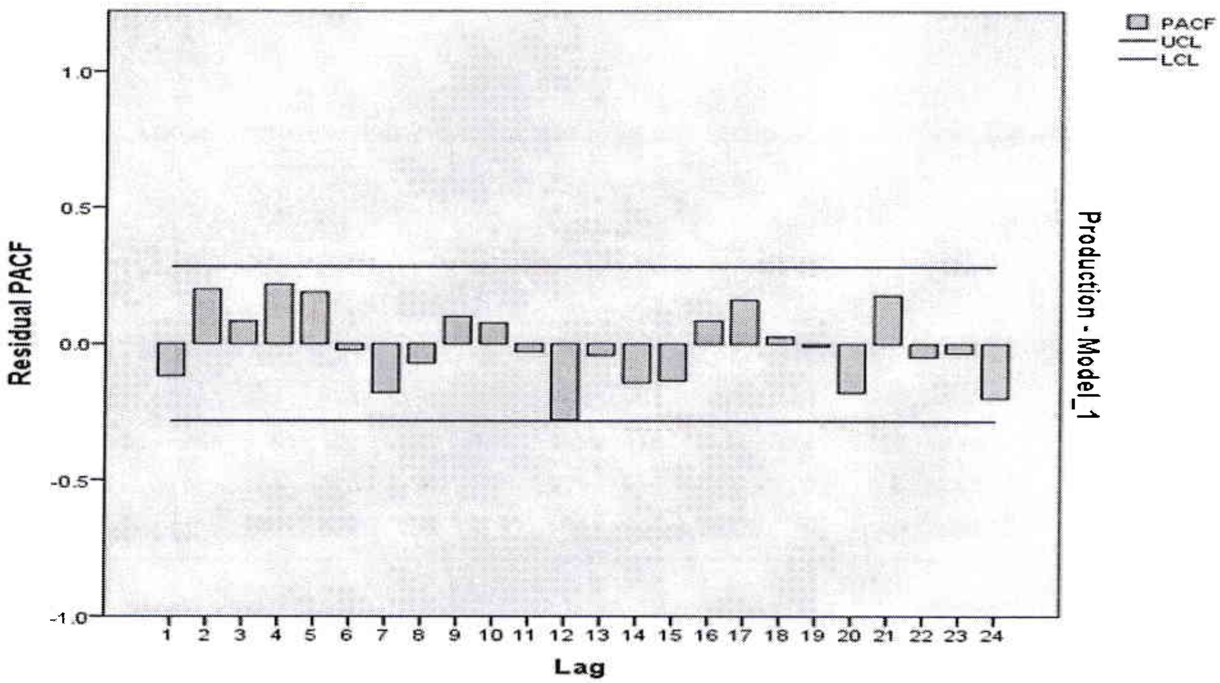
Table (4.28)

Estimated Partial Autocorrelation Function of Residual for seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ Model of Production for 160 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	-.117	.199	.083	.217	.188	-.022	-.179	-.070	.100	.077	-.029	-.278
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144
13-24	-.041	-.143	-.135	.085	.161	.027	-.009	-.178	.178	-.048	-.033	-.198
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144

Figure (4.16)

Sample Partial Autocorrelation Function of Residual values for seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ Model of Production for 160 KVA



Values of the residual ACF of seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ are all small and exhibit no patterns. And, the values of residual PACF of modified model lie inside the confidence limits. This suggested that this model adequate. Hence, the autocorrelation of $\hat{a}_t$ can be taken as significant different from zero.

An overall check is performed by using the test statistic,

$$Q = n \sum_{k=1}^{k} \gamma_k^2 (\hat{a}_t)$$

As the result of p value, the observed value of Q is 16.184 and it is not significant at 5 % significant level p-value is 0.440

Thus, the fitted seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ model is judged adequate for the series.

4.5 The Box-Jenkins Seasonal ARIMA Model of Production Series for 400 KVA

The monthly data of production series for 400 KVA covers 5 years, from January, 2013 to December 2017. The series consists of 60 observations.

4.5.1 Identification

For the identification of the order p and q, the (autocorrelation function) ACF and (partial autocorrelation function) PACF of the number of production series for 400 KVA are computed and plotted as shown in the following Tables and Figures.

Table (4.29)

Estimated Autocorrelation Function for the original series of Production for 400 Kilo Volt Ampere

$\hat{\rho}_k$ for $\{Z_t\}$ $\bar{Z}$ =27.72 S_z =7.057 n=60

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.675	.533	.457	.428	.370	.327	.298	.307	.239	.241	.270	.415
S.E	.126	.125	.124	.123	.122	.120	.119	.118	.117	.116	.115	.114
13-24	.279	.151	.201	.172	.097	.074	.021	.054	-.020	-.052	.008	.043
S.E	.112	.111	.110	.109	.108	.106	.105	.104	.102	.101	.100	.098

Figure (4.17)

Sample Autocorrelation Function for Monthly Production Series for 400 Kilo Volt Ampere

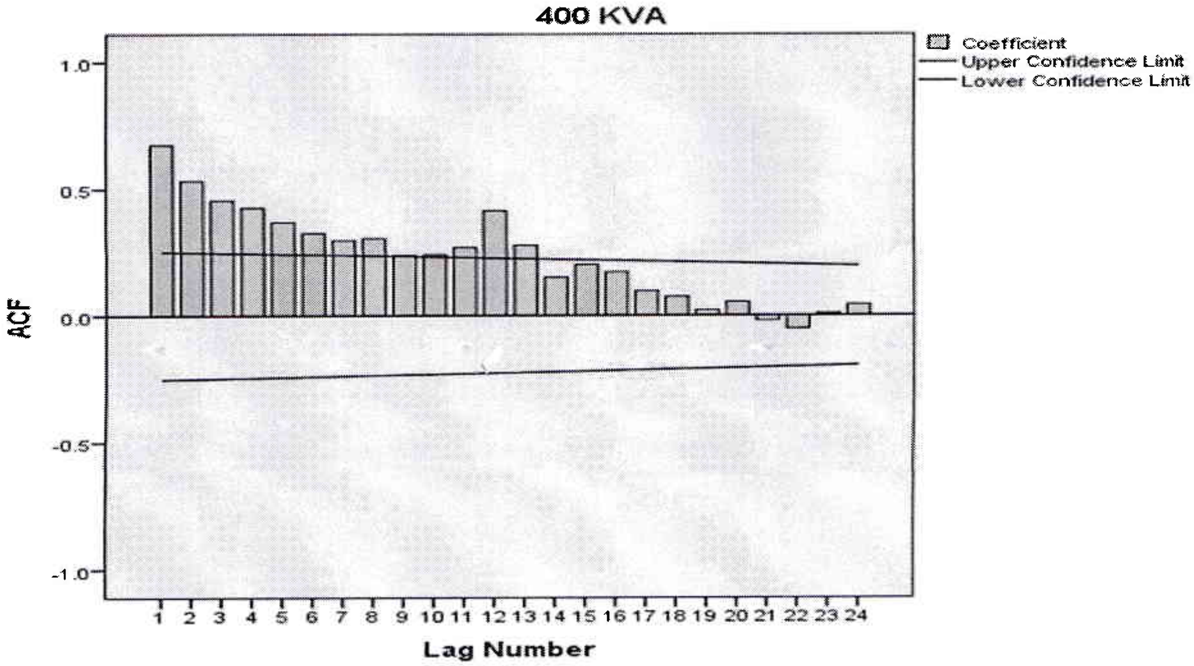


Table (4.30)

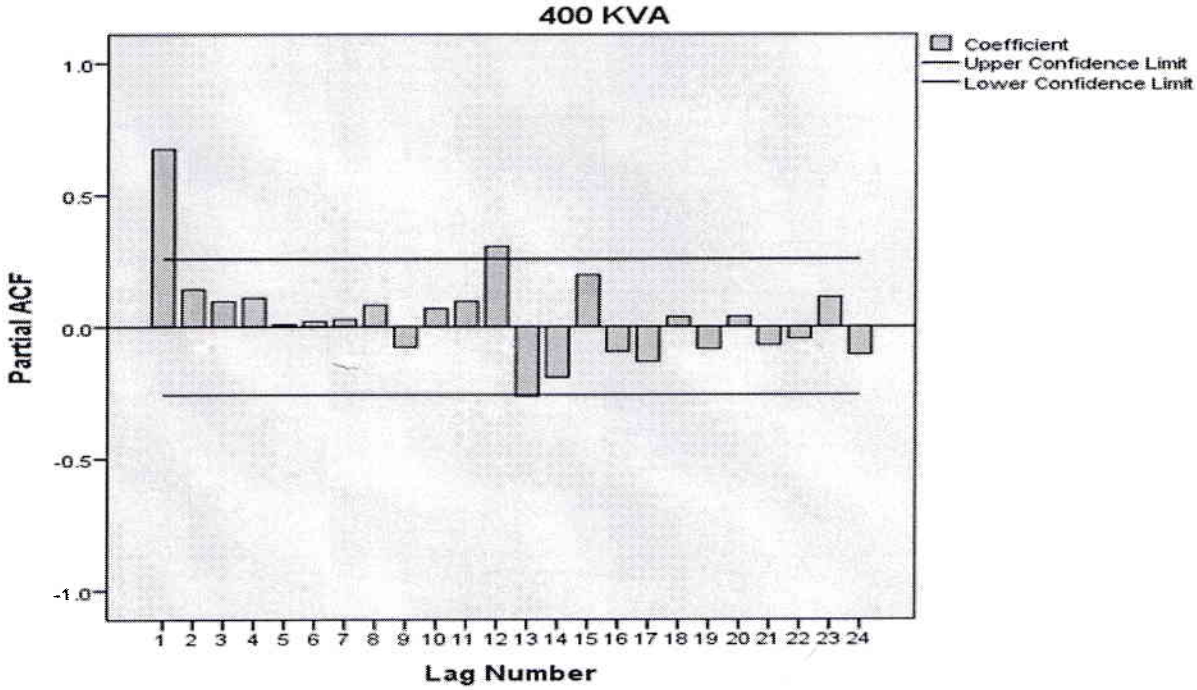
Estimated Partial Autocorrelation Function for the original series of Production for 400 Kilo Volt Ampere

$\widehat{\emptyset}_{kk}$ for $\{Z_t\}$ $\bar{Z}$ =27.72 S_z =7.057 n=60

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.675	.143	.097	.110	.010	.021	.028	.082	-.077	.070	.097	.306
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129
13-24	-.265	-.192	.198	-.095	-.133	.037	-.085	.038	-.069	-.044	.113	-.105
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129

Figure (4.18)

Sample Partial Autocorrelation Function for Monthly Production Series for 400 Kilo Volt Ampere



The sample ACF decays slowly and the sample PACF has a single large spike at lag 1. These values indicated that the series is nonstationary and that differencing is called for. To remove nonstationary, the series is seasonal differenced and the sample ACF and PACF of the seasonal differenced series (1-B^{12})Z_t were computed as shown in Table (4.31) and Table (4.32). They were displayed in Figure (4.19) and Figure (4.20).

Table (4.31)

Estimated Autocorrelation Function for Seasonal First Difference Series of Production for 400 Kilo Volt Ampere

$\hat{\rho}_k$ for $\{W_t = (1 - B^{12})Z_t\}$ $\bar{W} = 3.35$ $S_w = 4.601$ n=48

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.342	.210	-.155	-.182	-.078	-.140	-.030	-.015	.034	.066	-.209	-.276
S.E	.140	.138	.137	.135	.134	.132	.131	.129	.127	.126	.124	.122
13-24	-.151	-.028	.149	-.019	-.062	-.156	-.174	-.074	-.051	-.050	.172	.120
S.E	.121	.119	.117	.115	.114	.112	.110	.108	.106	.104	.102	.100

Figure (4.19)

Sample Autocorrelation Function for Seasonal First Difference Series of Production Series for 400 Kilo Volt Ampere

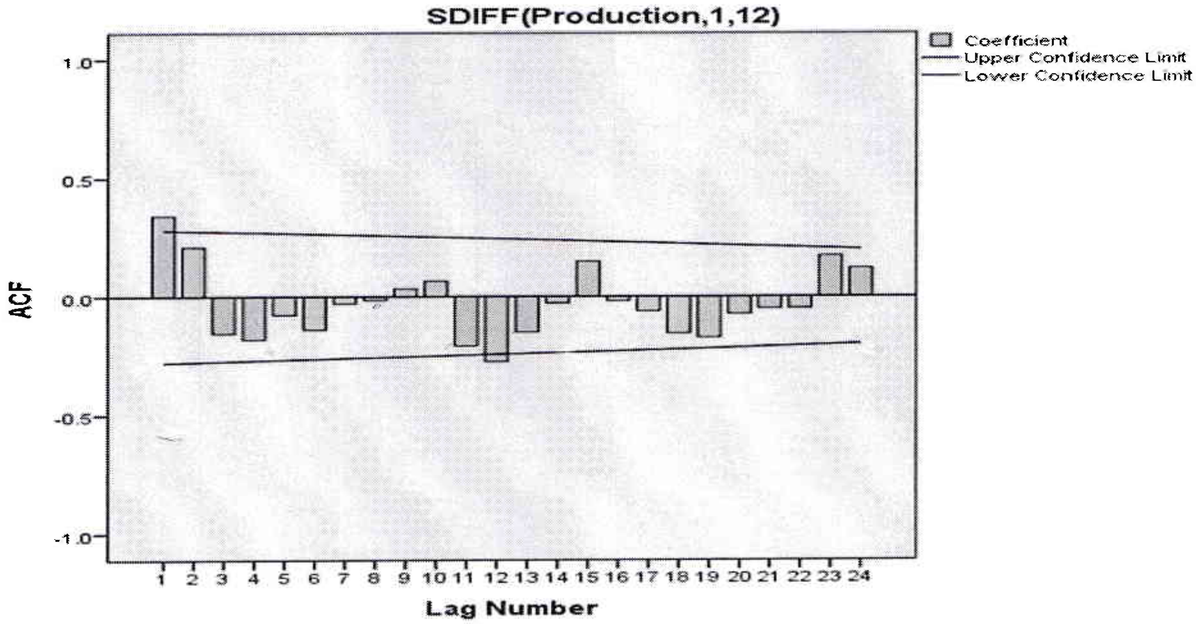


Table (4.32)

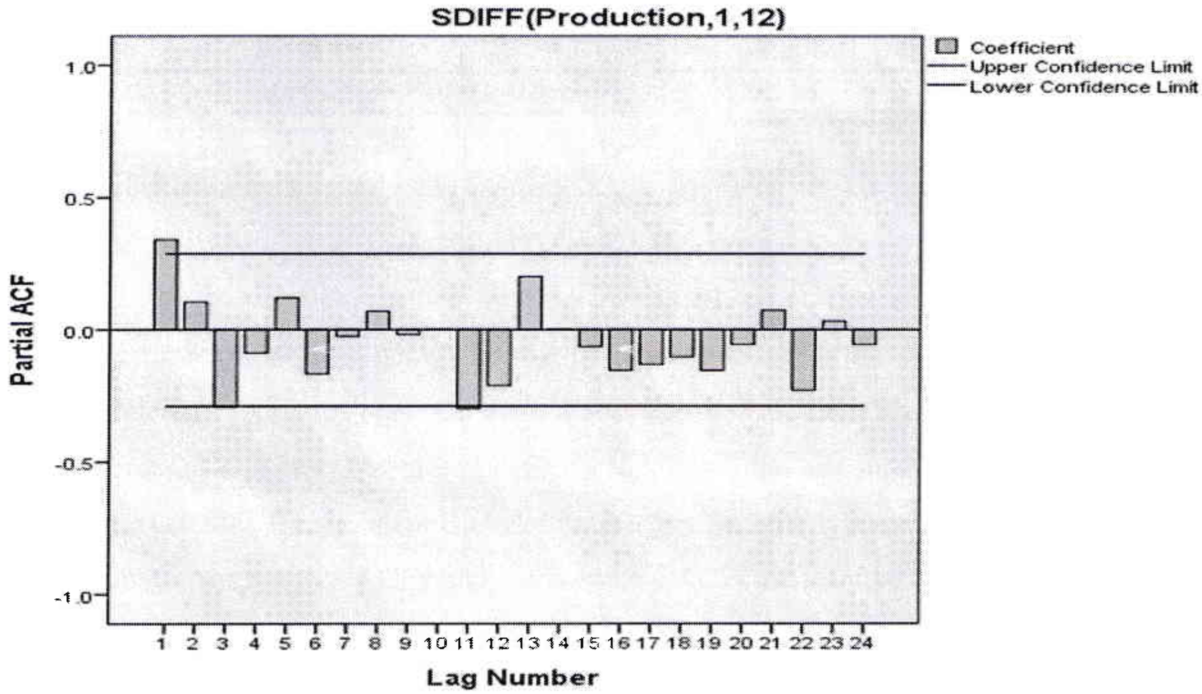
Estimated Partial Autocorrelation Function for Seasonal First Difference Series of Production for 400 Kilo Volt Ampere

$\widehat{\emptyset}_{kk}$ for $\{W_t = (1 - B^{12})Z_t\}$ $\overline{W} = 3.35$ $S_w = 4.601$ n=48

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.342	.106	-.293	-.088	.122	-.167	-.025	.071	-.019	.002	-.299	-.211
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144
13-24	.202	.003	-.064	-.154	-.132	-.103	-.154	-.056	.074	-.229	.034	-.055
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144

Figure (4.20)

Sample Partial Autocorrelation Function for Seasonal First Difference Series of Production Series for 400 Kilo Volt Ampere



The sample ACF is damped sine wave and the sample PACF cuts off after lag 1 because none of the sample PACF values is significant expect that lag 3 and 11

The suggested series (1-B^{12})Z_t might be described by SAR (1) process as a tentative model for the series

Since $\bar{W} = 3.35$, S_w=4.601, n= 48

The t value of t = $\frac{\bar{W}}{S_w/\sqrt{n}} = \frac{3.35}{4.601/\sqrt{48}} = 5.0444$

Which is significant and thus deterministic trend θ_0 is needed. Hence, the tentative model for the series following SAR(1) process:

$$(1-\Phi B^{12})Z_t = \theta_0 + a_t$$

4.5.2 Parameter Estimation for SAR (1) model

Using SAR (1) model, the estimated parameters with their statistics were shown in Table (4.33). According to this table, the estimated parameter of Φ is 0.708, since their p-value is 0.000, there is evidence to reject the null hypothesis: Φ = 0.

Table (4.33)

Estimated Parameters and Model Statistics for SAR(1) Model of Production Series for 400 KVA

	Estimate	SE	t	Sig.
Constant	27.454	1.912	14.358	0.000
Φ	0.708	0.109	6.513	0.000

The following estimated model was obtained

$$(1\text{-}0.708B^{12})Z_t = 27.454 + a_t$$

(0.109)   (1.912)

The estimation of the SAR (1) model of production series for 400 KVA give $\theta_0 = 27.454$ with estimated standard error 1.912 and $\Phi = 0.708$ with the estimated standard error 0.109. Under the null hypothesis H_0: $\Phi = 0$ the test statistics t is 6.513 with p-value is 0.000. Hence, there is evidence to reject the null hypothesis.

Moreover, the sample ACFs and the sample PACFs of residual for the above tentative model were shown in Table (4.34) and (4.35), respectively. They were showed in Figure (4.21) and (4.22).

Table (4.34)

Estimated Autocorrelation Function of Residual for SAR (1) Model of Production Series for 400 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.653	.542	.372	.286	.280	.194	.185	.143	.056	.037	-.076	-.102
S.E	.129	.176	.202	.213	.219	.225	.228	.230	.232	.232	.232	.232
13-24	-.038	-.018	.106	.054	.020	-.001	-.054	-.011	-.015	-.019	.084	.052
S.E	.233	.233	.233	.234	.234	.234	.234	.235	.235	.235	.235	.235

Figure (4.21)

Sample Autocorrelation Function of Residual values for SAR(1) Model of Production Series for 400 KVA

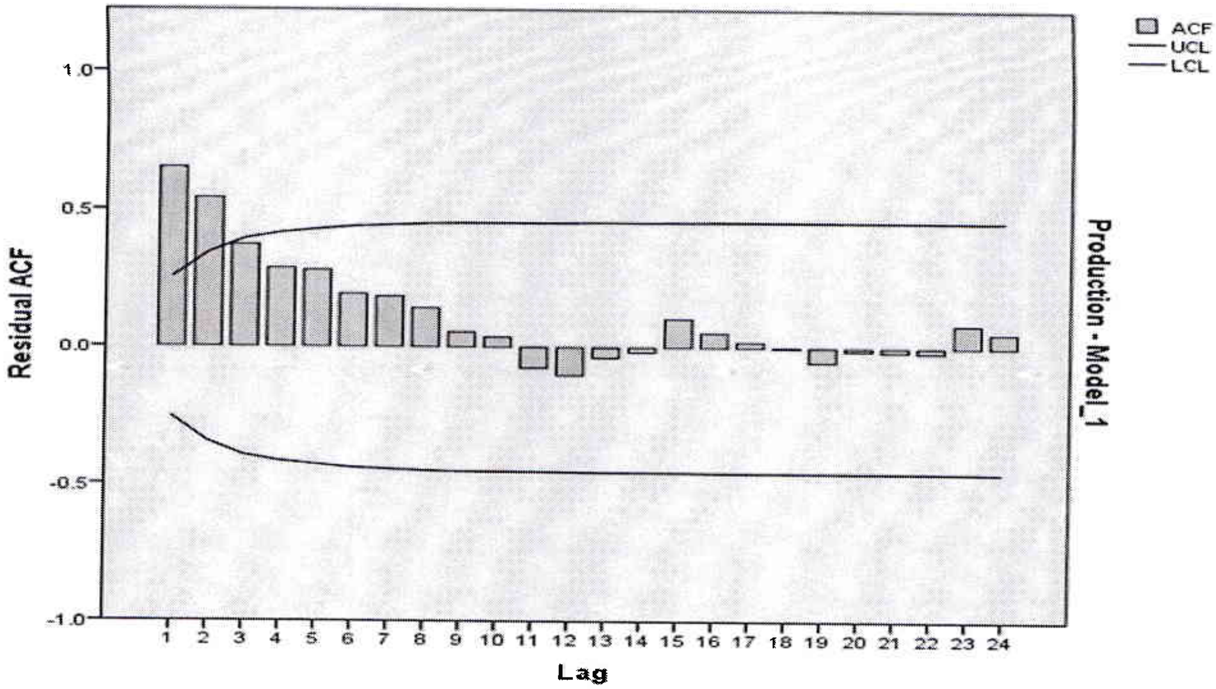


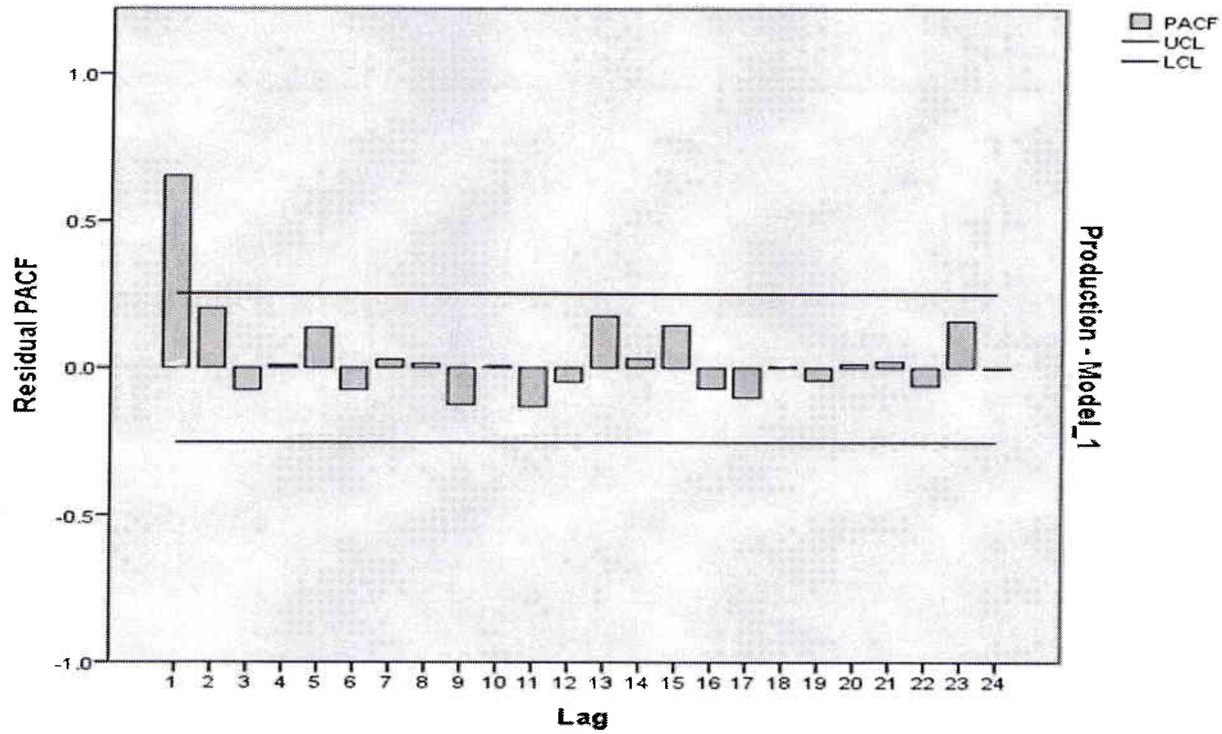
Table (4.35)

Estimated Partial Autocorrelation Function of Residual for SAR(1) Model of Production Series for 400 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.653	.202	-.076	.009	.137	-.075	.028	.014	-.125	.006	-.131	-.047
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129
13-24	.177	.032	.146	-.069	-.101	.005	-.042	.013	.023	-.061	.161	-.003
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129

Figure (4.22)

Sample Partial Autocorrelation Function of Residual values for SAR (1) Model of Production Series for 400 KVA



The Sample ACF was exponential decay and the sample PACF cuts off after lag 1 and these model exhibit a pattern. So, the residual series are not white noise process. Since, another tentative seasonal ARIMA (1, 0, 0) x $(0,1,0)_{12}$ model considered, that is

$$(1\text{-}B^{12})Z_t = \theta_0 + (1 - \emptyset B)a_t$$

Using multiplicative seasonal ARIMA (1, 0, 0) x $(0,1,0)_{12}$ model, the estimated parameters with their statistics were shown in Table (4.36).

Table (4.36)

Estimated Parameters and Model Statistics for seasonal ARIMA (1, 0, 0) x $(0, 1, 0)_{12}$ Model of Production for 400 KVA

	Estimate	SE	t	Sig.
Constant	3.314	0.952	3.482	0.001
$\emptyset$	0.345	0.140	2.460	0.018

The following estimated model was obtained

$$(1\text{-}B^{12})Z_t = 3.314 + (1 - 0.345B)a_t$$

(0.952) (0.140)

According to Table (4.36), it can be seen that the estimated parameter of $\emptyset$ is 0.345. Since their p-value of 0.018, there is evidence to reject the null hypothesis: $\emptyset = 0$.

4.5.3 Diagnostic Checking

To check model adequacy, in Table (4.37) and Table (4.38) was shown the residual ACF and PACF of the modified model. They were shown in Figure (4.23) and (4.24), along with the confidence interval.

$$\gamma_k(\hat{a}_t) \pm 2\widehat{S.E}[\gamma_k(\hat{a}_t)]$$

Where,

$$\widehat{S.E}[\gamma_k(\hat{a}_t)] = \frac{1}{\sqrt{n}}$$

Table (4.37)

Estimated Autocorrelation Function of Residual for seasonal ARIMA (1, 0, 0) x $(0,1,0)_{12}$ Model of Production for 400 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	-.030	.195	-.200	-.147	.041	-.139	.029	-.019	.028	.153	-.187	-.193
S.E	.144	.144	.150	.155	.158	.158	.161	.161	.161	.161	.164	.169
13-24	-.080	-.036	.203	-.070	-.012	-.111	-.138	-.024	-.026	-.091	.197	-.020
S.E	.173	.174	.174	.179	.179	.179	.181	.183	.183	.183	.184	.188

Figure (4.23)

Sample Autocorrelation Function of Residual values for seasonal ARIMA (1, 0, 0) x $(0,1,0)_{12}$ Model of Production for 400 KVA

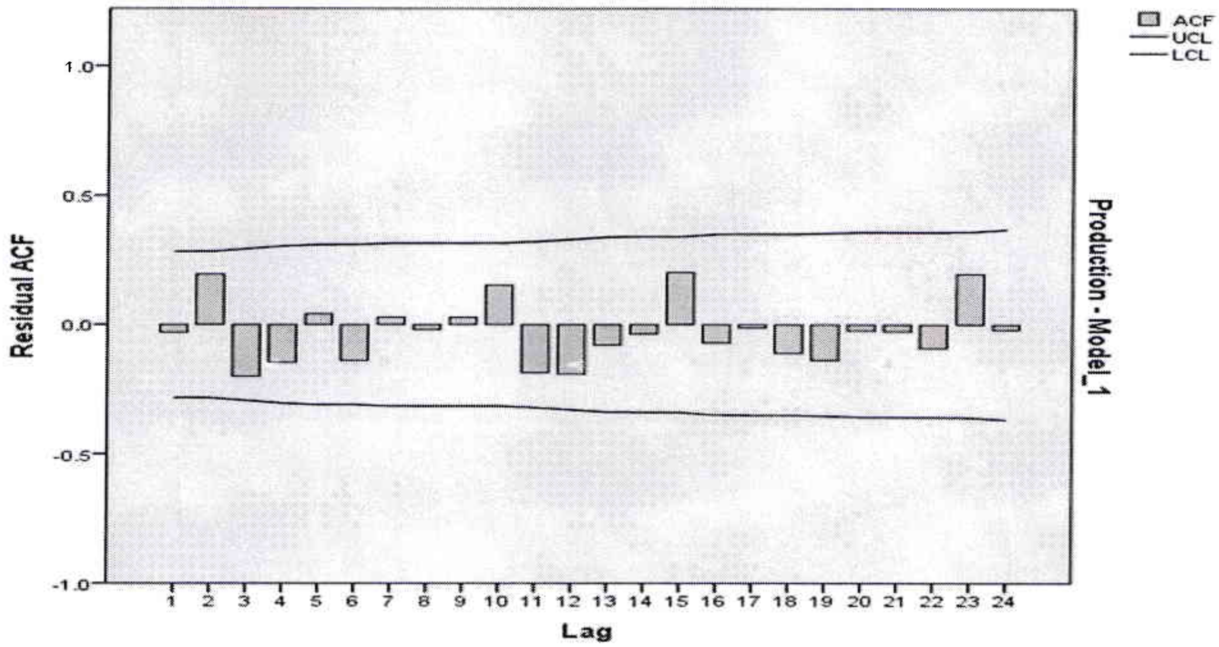


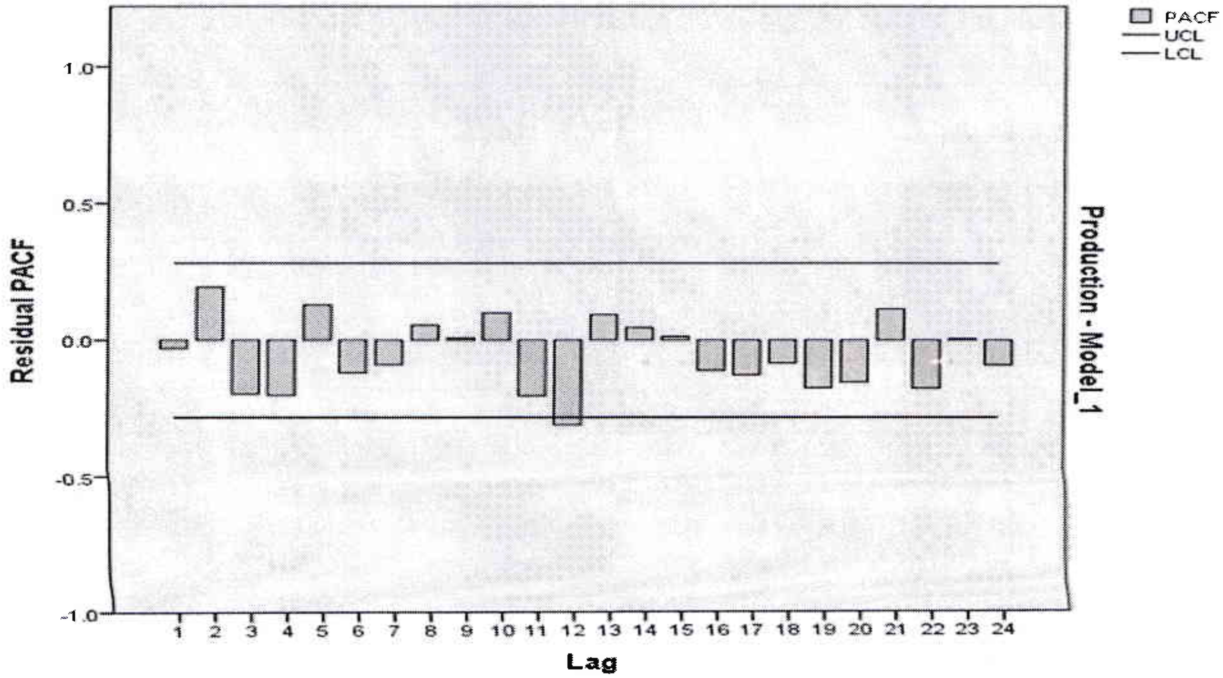
Table (4.38)

Estimated Partial Autocorrelation Function of Residual for seasonal ARIMA (1, 0, 0) x $(0,1,0)_{12}$ Model of Production for 400 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	-.030	.194	-.197	-.202	.130	-.120	-.090	.056	.008	.100	-.204	-.310
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144
13-24	.093	.046	.013	-.111	-.128	-.085	-.175	-.154	.115	-.176	.005	-.092
S.E	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144	.144

Figure (4.24)

Sample Partial Autocorrelation Function of Residual values for seasonal ARIMA $(1, 0, 0) \times (0, 1, 0)_{12}$ Model of Production for 400 KVA



Values of the residual ACF of seasonal ARIMA (1, 0, 0) x $(0,1,0)_{12}$ are all small and exhibit no patterns. And, the values of residual PACF of modified model lie inside the confidence limits except at lag 12. This suggested that this model adequate. Hence, the autocorrelation of $\hat{a}_t$ can be taken as significant different from zero.

An overall check is performed by using the test statistic,

$$Q = n \sum_{k=1}^{k} \gamma_k^2 (\hat{\hat{a}}_t)$$

As the result of p value, the observed value of Q is 17.754and it is not significant at 5 % significant level p-value is 0.405

Thus, the fitted seasonal ARIMA (1, 0, 0) x $(0,1,0)_{12}$ model is judged adequate for the series.

4.6 The Box-Jenkins Seasonal ARIMA Model of Production Series for 2000 KVA

The monthly data of production series for 2000 KVA covers 5 years, from January, 2013 to December 2017. The series consists of 60 observations.

4.6.1 Identification

For the identification of the order p and q, the (autocorrelation function) ACF and (partial autocorrelation function) PACF of the number of production series for 2000 KVA are computed and plotted as shown in the following Tables and Figures.

Table (4.39)

Estimated Autocorrelation Function for the original series of Production for 2000 Kilo Volt Ampere

$\hat{\rho}_k$ for $\{Z_t\}$ $\bar{Z}$ =23.90 S_z =5.115 n=60

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.524	.388	.155	.133	.122	.215	.127	.007	.051	.169	.302	.385
S.E	.126	.125	.124	.123	.122	.120	.119	.118	.117	.116	.115	.114
13-24	.250	.023	-.109	.021	-.014	.103	.003	-.060	-.041	.003	.154	.146
S.E	.112	.111	.110	.109	.108	.106	.105	.104	.102	.101	.100	.098

Figure (4.25)

Sample Autocorrelation Function for Monthly Production Series for 2000 Kilo Volt Ampere

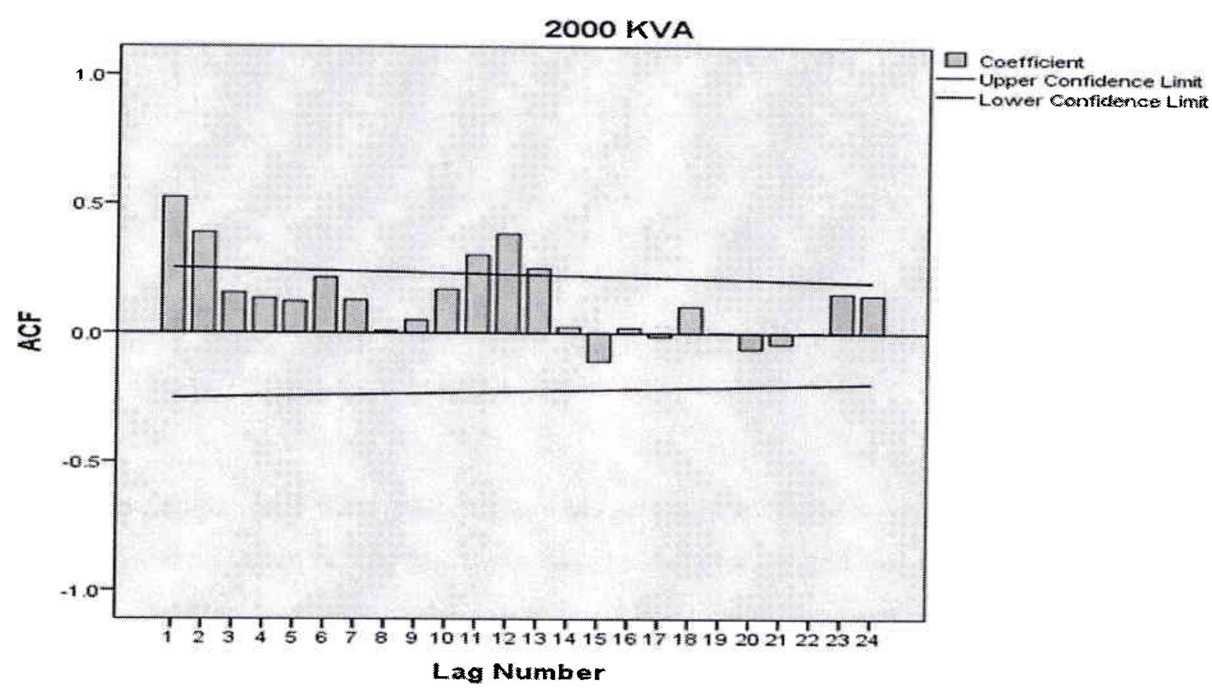


Table (4.40)

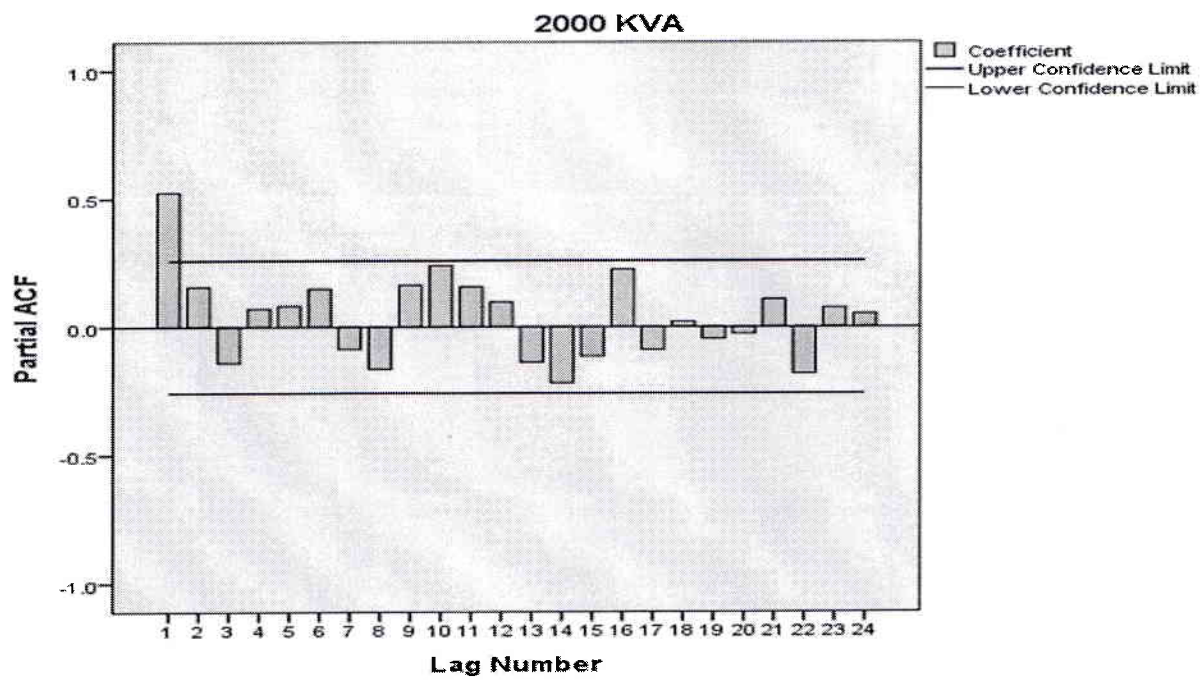
Estimated Partial Autocorrelation Function for the original series of Production for 2000 Kilo Volt Ampere

$\widehat{\emptyset}_{kk}$ for $\{Z_t\}$ $\bar{Z}$ =23.90 S_z =5.115 n=60

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.524	.157	-.140	.072	.083	.149	-.087	-.164	.164	.240	.157	.098
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129
13-24	-.137	-.218	-.114	.225	-.089	.021	-.046	-.026	.109	-.181	.076	.051
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129

Figure (4.26)

Sample Partial Autocorrelation Function for Monthly Production Series for 2000 Kilo Volt Ampere



The sample ACF decays slowly and the sample PACF has a single large spike at lag 1. These values indicated that the series is nonstationary and that differencing is called for. To remove nonstationary, the series is seasonal differenced and the sample ACF and PACF of non-seasonal differencing and seasonal differencing were computed as shown in Table (4.41) and Table (4.42). They were displayed in Figure (4.27) and Figure (4.28).

Table (4.41)

Estimated Autocorrelation Function for Non-Seasonal and Seasonal First Difference Series of Production for 2000 Kilo Volt Ampere

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	-.625	.248	-.105	.045	-.060	-.063	.107	-.117	.055	.032	.167	-.278
S.E	.141	.140	.138	.137	.135	.133	.132	.130	.128	.127	.125	.123
13-24	.137	.005	-.127	.145	-.153	.149	-.052	-.075	.226	-.282	.171	-.092
S.E	.122	.120	.118	.116	.114	.112	.110	.108	.106	.104	.102	.100

Figure (4.27)

Sample Autocorrelation Function for Non-Seasonal and Seasonal First Difference Series of Production Series for 2000 Kilo Volt Ampere

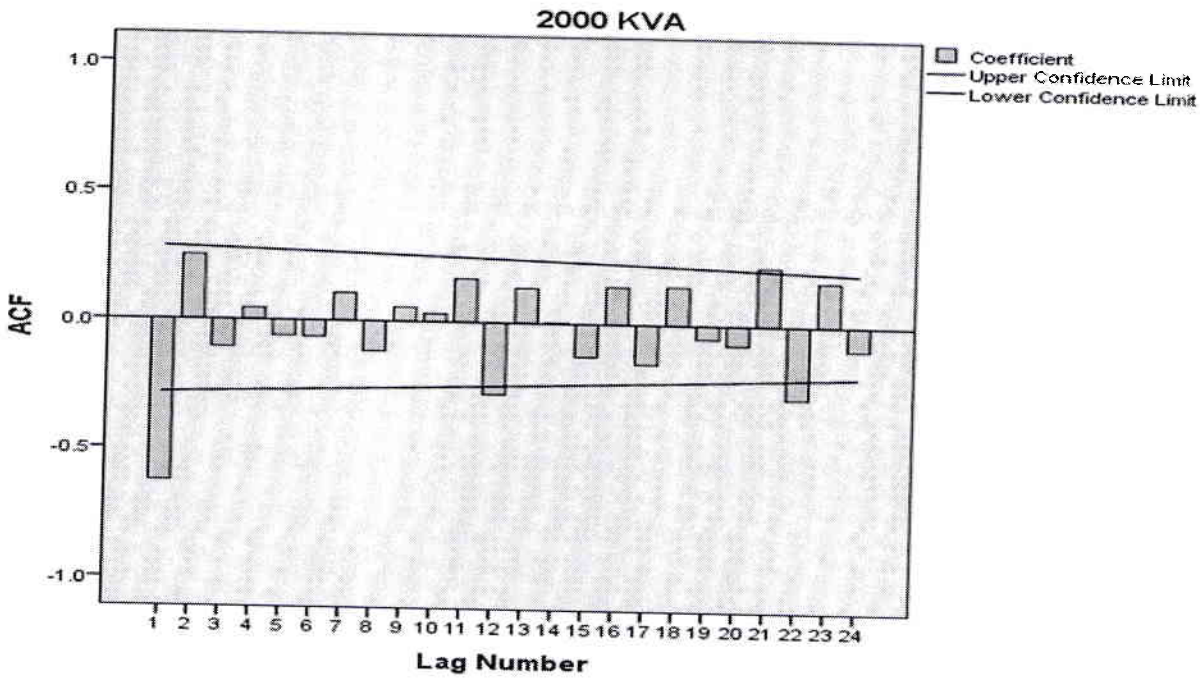


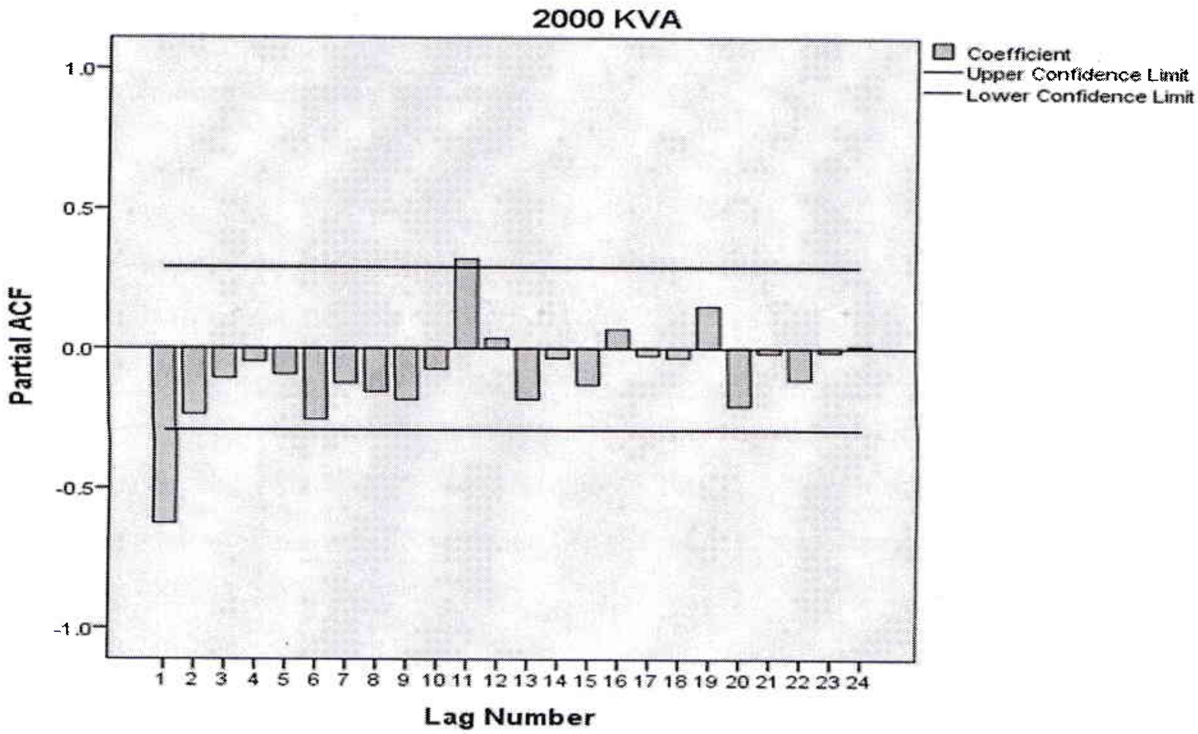
Table (4.42)

Estimated Partial Autocorrelation Function for Non-Seasonal and Seasonal First Difference Series of Production for 2000 Kilo Volt Ampere

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	-.625	-.235	-.106	-.046	-.092	-.253	-.123	-.154	-.182	-.073	.322	.036
S.E	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146
13-24	-.181	-.034	-.129	.069	-.025	-.032	.151	-.205	-.016	-.112	-.013	.011
S.E	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146

Figure (4.28)

Sample Partial Autocorrelation Function for Non-Seasonal and Seasonal First Difference Series of Production Series for 2000 Kilo Volt Ampere



The sample ACF is tails off and the sample PACF cuts off after lag 1 because none of the sample PACF values is significant expect that lag 6 and 11.

The suggested series (1-B^{12})Z_t might be described by SAR (1) process as a tentative model for the series

$$(1\text{-}\Phi B^{12})Z_t = \theta_0 + a_t$$

4.6.2 Parameter Estimation for SAR (1) model

Using SAR (1) model, the estimated parameters with their statistics were shown in Table (4.43). According to this table, the estimated parameter of Φ is 0.589, since their p-value is 0.000, there is evidence to reject the null hypothesis: $\Phi = 0$.

Table (4.43)

Estimated Parameters and Model Statistics for SAR(1) Model of Production Series for 2000 KVA

	Estimate	SE	t	Sig.
Constant	23.470	1.118	20.995	0.000
Φ	0.589	0.117	5.026	0.000

The following estimated model was obtained

$$(1\text{-}0.589B^{12})Z_t = 23.470 + a_t$$

(0.177) (1.118)

The estimation of the SAR (1) model of production series for 2000 KVA give $\theta_0 = 23.470$ with estimated standard error 1.118 and $\Phi = 0.589$ with the estimated standard error 0.117. Under the null hypothesis H_0: $\Phi = 0$ the test statistics t is 5.026 with p-value is 0.000. Hence, there is evidence to reject the null hypothesis.

Moreover, the sample ACFs and the sample PACFs of residual for the above tentative model were shown in Table (4.44) and (4.45), respectively. They were showed in Figure (4.29) and (4.30).

Table (4.44)

Estimated Autocorrelation Function of Residual for SAR (1) Model of Production Series for 2000 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.400	.439	.245	.155	.086	.096	.102	-.042	.006	.020	.024	-.119
S.E	.129	.148	.169	.174	.177	.177	.178	.179	.179	.179	.179	.179
13-24	-.006	-.085	-.149	.068	-.010	.097	-.008	-.034	-.002	-.133	.027	-.039
S.E	.181	.181	.181	.183	.184	.184	.185	.185	.185	.185	.186	.186

Figure (4.29)

Sample Autocorrelation Function of Residual values for SAR(1) Model of Production Series for 2000 KVA

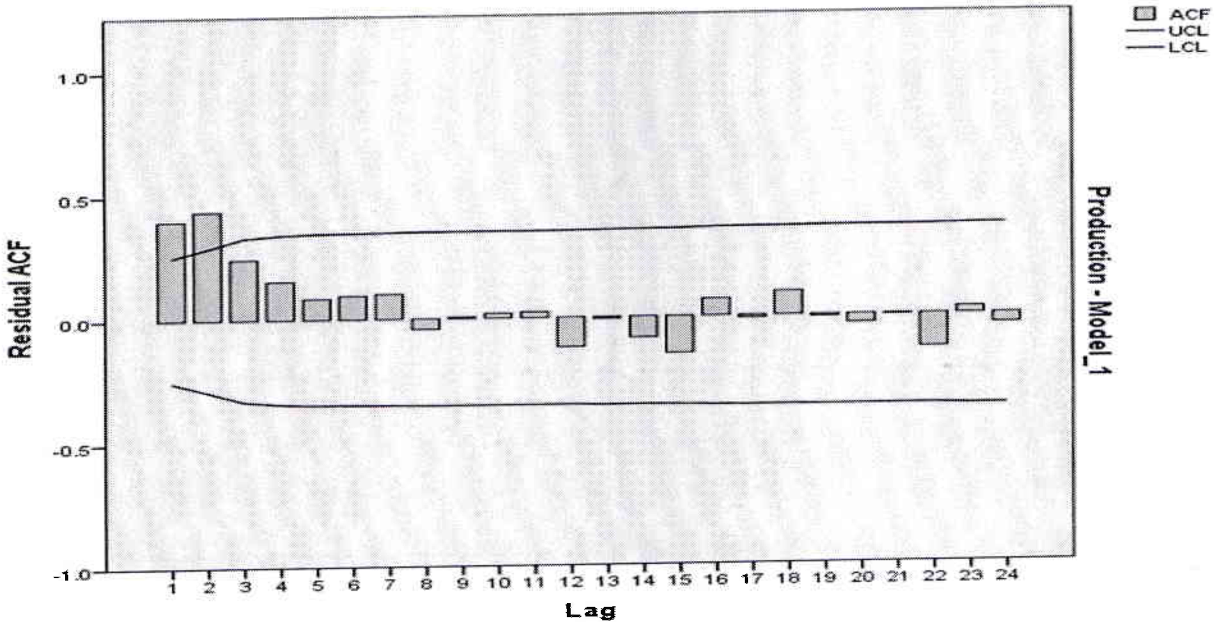


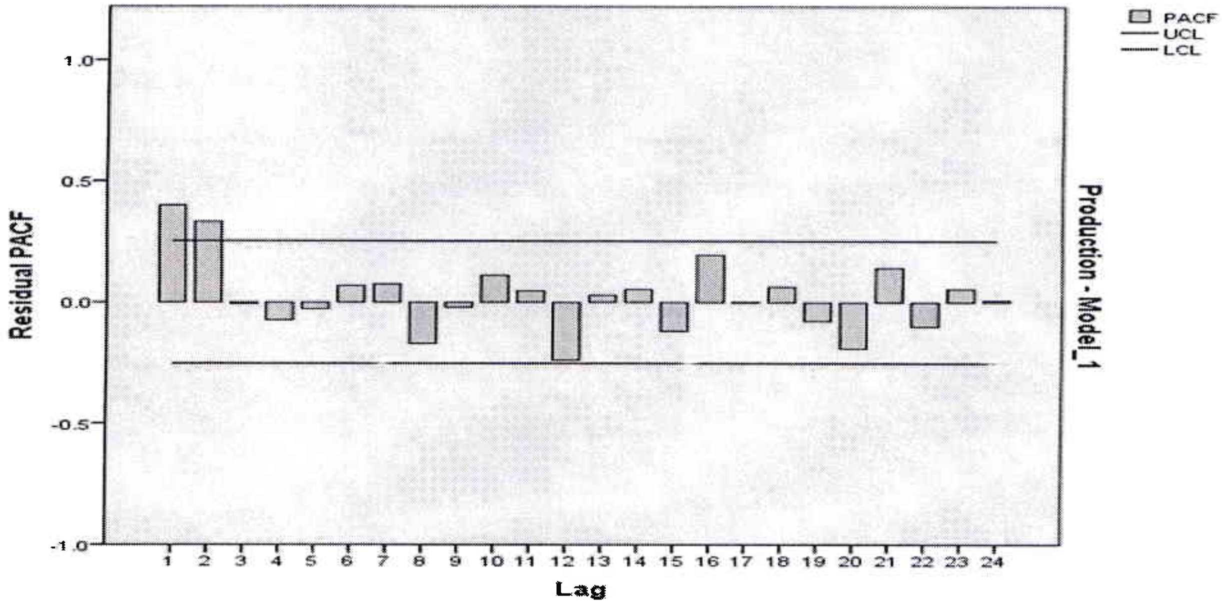
Table (4.45)

Estimated Partial Autocorrelation Function of Residual for SAR(1) Model of Production Series for 2000 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	.400	.332	-.007	-.073	-.027	.069	.076	-.171	-.021	.113	.047	-.240
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129
13-24	.030	.054	-.119	.196	-.001	.065	-.077	-.192	.142	-.101	.055	.008
S.E	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129	.129

Figure (4.30)

Sample Partial Autocorrelation Function of Residual values for SAR (1) Model of Production Series for 2000 KVA



The Sample ACF was exponential decay and the sample PACF cuts off after lag 2 and these model exhibit a pattern.So, the residual series are not white noise process. Since, another tentative seasonal ARIMA (1, 1, 0) x $(1,1,0)_{12}$ model considered, that is

$$(1\text{-}\emptyset B)(1\text{-}\Phi B^{12})(1\text{-}B)(1\text{-}B^{12})X_t = \mu + a_t$$

Using multiplicative seasonal ARIMA (1, 1, 0) x $(1,1,0)_{12}$ model, the estimated parameters with their statistics were shown in Table (4.46).

Table (4.46)

Estimated Parameters and Model Statistics for seasonal ARIMA (1, 1, 0) x $(1,1,0)_{12}$ Model of Production for 2000 KVA

	Estimate	SE	t	Sig.
Constant	-0.049	0.269	-0.181	0.857
Ø	-0.669	0.110	-6.077	0.000
Φ	-0.514	0.131	-3.919	0.000

The following estimated model was obtained

$$(1\text{-}0.669B)(1\text{-}0.514B^{12})(1\text{-}B)(1\text{-}B^{12})X_t = \text{-}0.049 + a_t$$

(0.110) (0.131) (0.269)

According to Table (4.46), the estimated parameters of Ø and Φ are -0.669 and -0.514, respectively. Since the p-value less than $\alpha = 0.05$, the parameters values are significant at 5% level.

4.6.3 Diagnostic Checking

To check model adequacy, in Table (4.47) and Table (4.48) was shown the residual ACF and PACF of the modified model. They were shown in Figure (4.31) and (4.32), along with the confidence interval.

$$\gamma_k(\hat{a}_t) \pm 2\widehat{S.E}[\gamma_k(\hat{a}_t)]$$

Where,

$$\widehat{S.E}[\gamma_k(\hat{a}_t)] = \frac{1}{\sqrt{n}}$$

Table (4.47)

Estimated Autocorrelation Function of Residual for seasonal ARIMA (1, 1, 0) x $(1,1,0)_{12}$ Model of Production for 2000 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	-.096	-.085	.027	-.115	-.223	-.091	-.038	-.183	.169	.169	.134	.064
S.E	.146	.147	.148	.148	.150	.157	.158	.158	.163	.167	.170	.172
13-24	.032	-.109	-.075	.009	-.196	.128	-.003	-.098	.156	-.028	.046	-.057
S.E	.173	.173	.174	.175	.175	.180	.182	.182	.183	.186	.186	.186

Figure (4.31)

Sample Autocorrelation Function of Residual values for seasonal ARIMA (1, 1, 0) x $(1, 1, 0)_{12}$ Model of Production for 2000 KVA

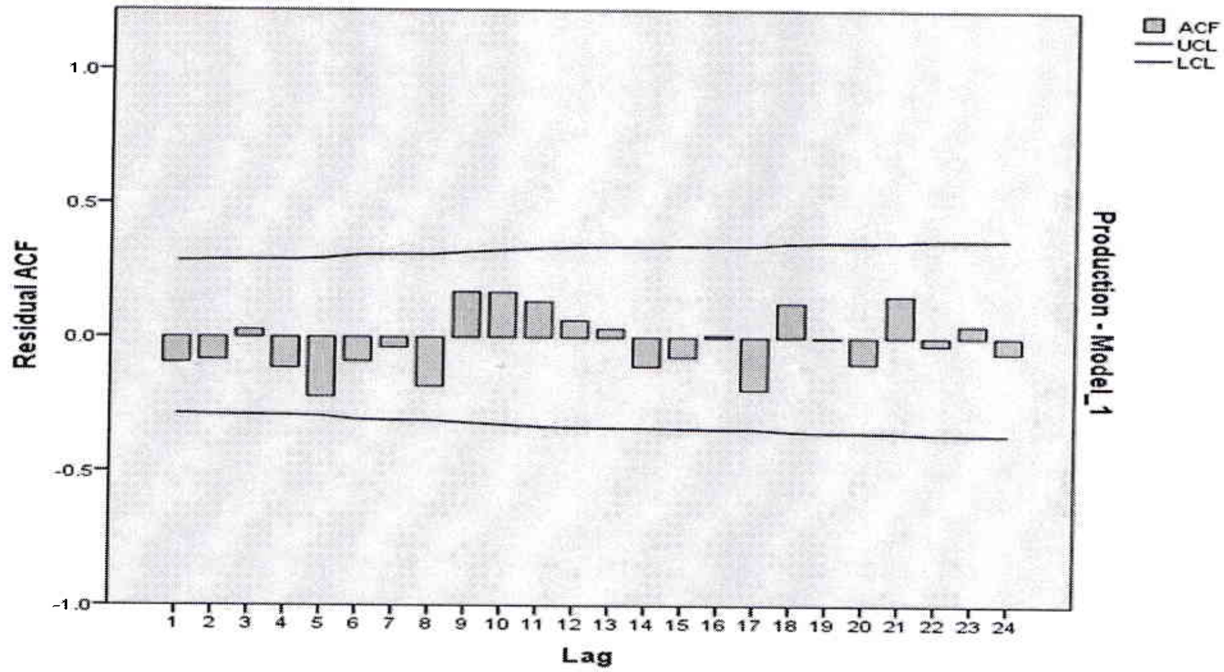


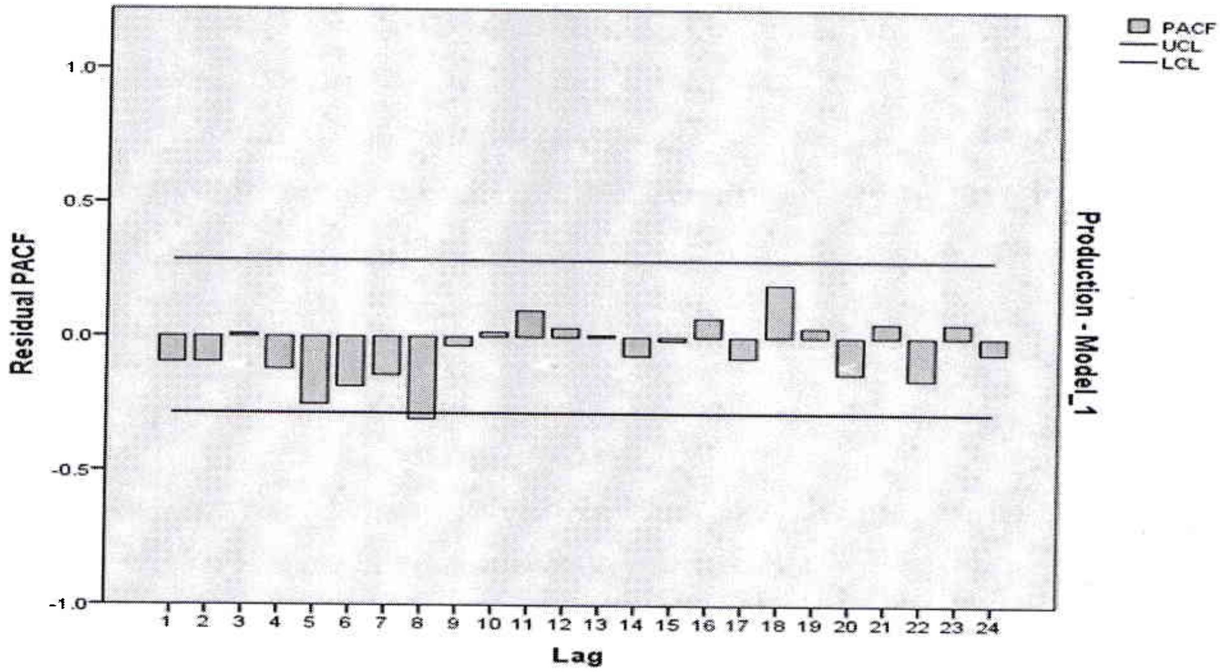
Table (4.48)

Estimated Partial Autocorrelation Function of Residual for seasonal ARIMA (1, 1, 0) x $(1, 1, 0)_{12}$ Model of Production for 2000 KVA

Lag k	1	2	3	4	5	6	7	8	9	10	11	12
1-12	-.096	-.095	.009	-.122	-.252	-.186	-.143	-.308	-.033	.017	.098	.035
S.E	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146
13-24	.007	-.070	-.012	.072	-.078	.196	.035	-.137	.053	-.160	.052	-.059
S.E	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146	.146

Figure (4.32)

Sample Partial Autocorrelation Function of Residual values for seasonal ARIMA $(1, 1, 0) \times (1, 1, 0)_{12}$ Model of Production for 2000 KVA



Values of the residual ACF of seasonal ARIMA (1, 1, 0) x $(1,1,0)_{12}$ are all small and exhibit no patterns. And, the values of residual PACF of modified model lie inside the confidence limits except at lag 8. This suggested that this model adequate. Hence, the autocorrelation of $\widehat{a_t}$ can be taken as significant different from zero.

An overall check is performed by using the test statistic,

$$Q = n \sum_{k=1}^{k} \gamma_k^2 (\hat{a}_t)$$

As the result of p value, the observed value of Q is 17.322 and it is not significant at 5 % significant level p-value is 0.365

Thus, the fitted seasonal ARIMA (1, 1, 0) x $(1,1,0)_{12}$ model is judged adequate for the series.

4.7 Forecasting

The models of the production series for 100 Kilo Volt Ampere, 160 Kilo Volt Ampere ,400 Kilo Volt Ampere and 2000 Kilo Volt Ampere have been identified, estimated and checked for adequacy. The accepted models will be used to forecast the values for January to December of 2018.

Production Series for 100 Kilo Volt Ampere

Since the model seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ is adequate, this model can be used to forecast the future value for production of 100 KVA.

The forecasts for January to December, 2018 are as shown in Table (4.49)

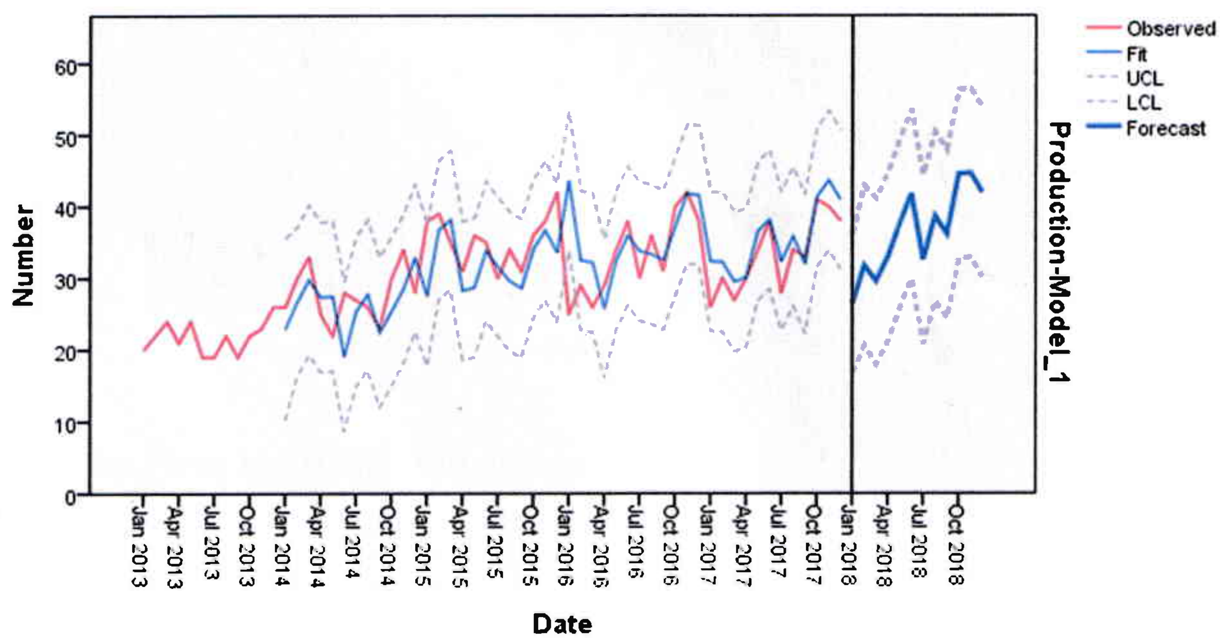
Table (4.49)

The Forecast for January to December, 2018 of Production Series for 100 KVA

Jan	26	May	38	Sep	36
Feb	32	Jun	42	Oct	45
Mar	30	Jul	33	Nov	45
Apr	33	Aug	39	Dec	42

Figure (4.33)

The Actual, Fitted and Forecast Values with 95% Confidence Limits for the Number of Production Series for 100 KVA



Production Series for 160 Kilo Volt Ampere

Since the model seasonal ARIMA (1, 0, 0) x $(1,1,0)_{12}$ is adequate, this model can be used to forecast the future value for production of 160 KVA.

The forecasts for January to December, 2018 are as shown in Table (4.50)

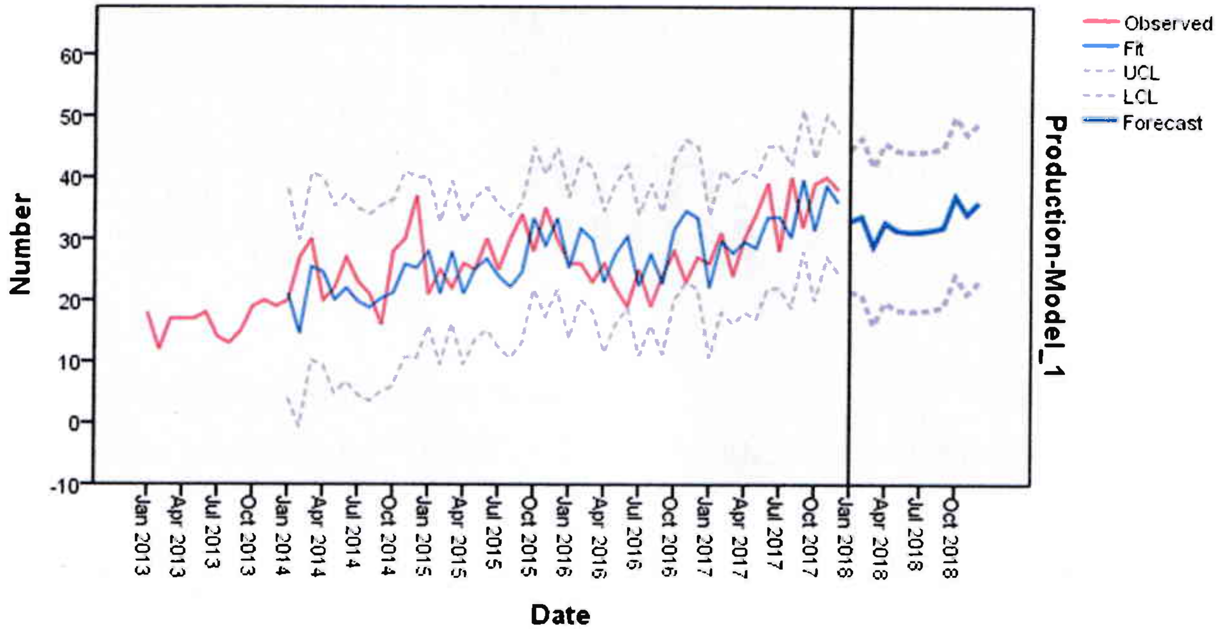
Table (4.50)

The Forecast for January to December, 2018 of Production Series for 160 KVA

Jan	33	May	31	Sep	32
Feb	34	Jun	31	Oct	37
Mar	29	Jul	31	Nov	34
Apr	33	Aug	31	Dec	36

Figure (4.34)

The Actual, Fitted and Forecast Values with 95% Confidence Limits for the Number of Production Series for 160 KVA



Production Series for 400 Kilo Volt Ampere

Since the model seasonal ARIMA (1, 0, 0) x $(0,1,0)_{12}$ is adequate, this model can be used to forecast the future value for production of 400 KVA.

The forecasts for January to December, 2018 are as shown in Table (4.51)

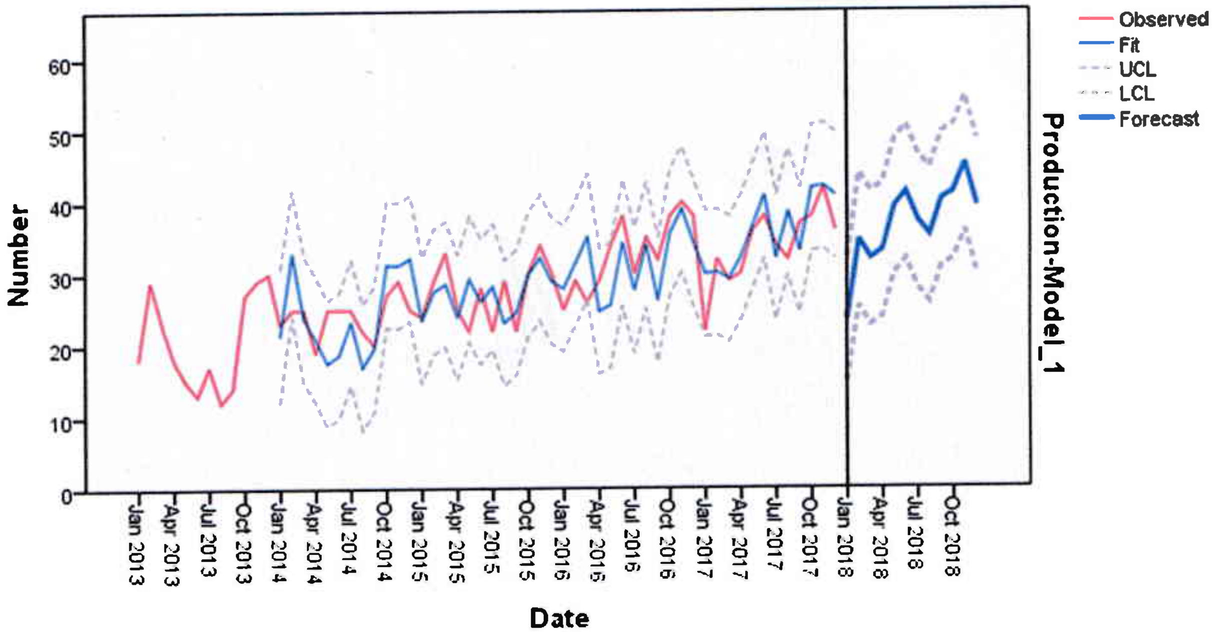
Table (4.51)

The Forecast for January to December, 2018 of Production Series for 400 KVA

Jan	23	May	39	Sep	40
Feb	35	Jun	41	Oct	41
Mar	32	Jul	37	Nov	45
Apr	33	Aug	35	Dec	39

Figure (4.35)

The Actual, Fitted and Forecast Values with 95% Confidence Limits for the Number of Production Series for 400 KVA



Production Series for 2000 Kilo Volt Ampere

Since the model seasonal ARIMA (1, 1, 0) x $(1,1,0)_{12}$ is adequate, this model can be used to forecast the future value for production of 2000 KVA.

The forecasts for January to December, 2018 are as shown in Table (4.52)

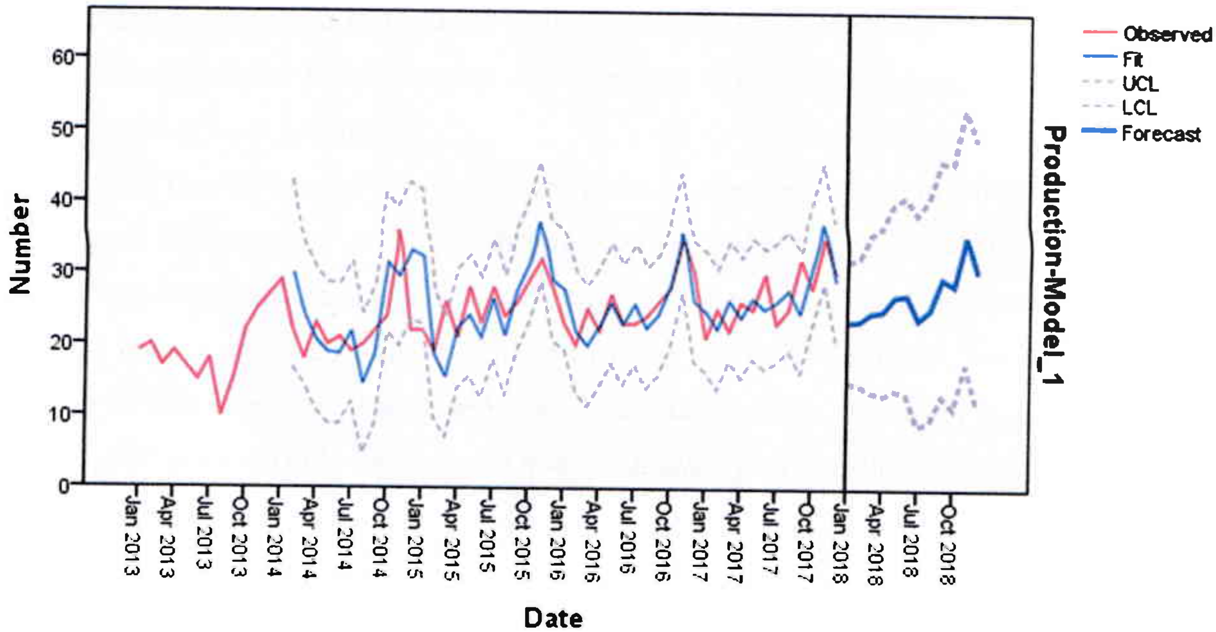
Table (4.52)

The Forecast for January to December, 2018 of Production Series for 2000 KVA

Jan	23	May	27	Sep	29
Feb	23	Jun	27	Oct	28
Mar	25	Jul	24	Nov	35
Apr	25	Aug	25	Dec	30

Figure (4.36)

The Actual, Fitted and Forecast Values with 95% Confidence Limits for the Number of Production Series for 2000 KVA



CHAPTER V

CONCLUSION

In this thesis, the basic statistical characteristics of some monthly production series of transformers such as 100 Kilo Volt Ampere,160 Kilo Volt Ampere, 400 Kilo Volt Ampere, 2000 Kilo Volt Ampere series and the model building procedures for these series have been presented.

Many time series data have important seasonal components and it is necessary to measure the seasonal variation. In time series analysis, one study the four components: trend, seasonal, cyclical and random existed in the time series model. One may test the seasonality for these data by using the ANOVA Table. Trend and cyclical components are represented by deterministic time functions, seasonal component of seasonal indexes and random components by its statistical properties.

The number of production series are gradually increasing during the period of 5 years (from studying period 2013 to 2017). It is found that the total number of production increased during winter seasons such as October, November and December.

A seasonal index may be computed for the purpose of studying the seasonal movement itself, the objective being to avoid or minimize the consequences of the seasonal changes, in order to smooth out the seasonal fluctuations. In this thesis, the Ratio to Moving Average Method is used to measure the seasonal index. From production series for 100 KVA, it has been found that the lowest value of seasonal index occurs in September and the highest value of seasonal index are observed in November. From production series for 160 KVA, it has been found that the lowest value of seasonal index occurs in August and the highest value of seasonal index are observed in December. From production series for 400 KVA, it has been found that the lowest value of seasonal index occurs in September and the highest value of seasonal index are observed in November. From production series for 2000 KVA, it has been found that the lowest value of seasonal index occurs in August and the highest value of seasonal index are observed in November.

In addition, the Box- Jenkins method was utilized in modelling and forecasting the number of production series of transformers. The multiplicative seasonal ARIMA $(1, 0, 0) \times (1,1,0)_{12}$, ARIMA $(1, 0, 0) \times (0,1,0)_{12}$ and ARIMA $(1, 1, 0) \times (1,1,0)_{12}$ models were found to be adequate for the observed data series.

Based on the best fitted model, monthly production series for transformers are forecasted for future periods of 2018. The forecast value obtained by using fitted model was generally considered to be reliable. So, the forecast values can be applied in a variety of future planning purpose which are important for the production of transformers in HISEM Co., Ltd. Finally, it is recommended that measuring seasonal variation, seasonal model building and forecasting should be updated regularly in order to give better estimates or forecasts for the number of production.

REFERENCES

1. Anderson D.R, Sweeney D.J, Williams T.A: "Statistics for business and Economics", 8^{th} Edition; South-Western.
2. Box, G. E. P., Jenkins, G. M., "Time Series Analysis: Forecasting and Control", San Francisco, Diisseldorf, Johannesburg, Londom, Panama, Singapore, Sydney. Toranto.
3. Daniel, W.W. and Terrell, J.C, (1992), "Business Statistics for Management and Economics", sixth edition, Houghton Mifflin Co., Boston.
4. May Thu Soe, (2011), "A Seasonal Time Series Model of Tourist Arrivals in Myanmar", M.Econ (Thesis), Department of Statistics, Yangon University of Economics.
5. Mya Thandar, (1997), "Seasonal Models for Monthly Transport Time Series of Myanmar", M.Econ (Thesis), Department of Statistics, Yangon University of Economics.
6. Myint Moe Moe Khin, (1998), "A Comparative Study of Some Estimation Methods on Seasonal Variation", M.Econ (Thesis), Department of Statistics, Yangon University of Economics.
7. Soe Thu Zar Tint, (2010), "Time Series Forecasting Using Holt-Winters Exponential Smoothing", M.Econ (Thesis), Department of Statistics, Yangon University of Economics.
8. Steiner, P.O., (1956), "An Introduction to the Analysis of Time Series", Preliminary edition, Rinehart, New York.
9. Thandar, (2017), "Effect of Seasonality on Tourism Demand in Myanmar", M.Econ (Thesis), Department of Statistics, Yangon University of Economics.
10. Wei, W.S., (2006), "Time Series Analysis: Univariate and Multivariate Methods", 2^{nd} Edition; Pearson Education, Inc.

APPENDIX

Appendix A

Production Series for 100 KVA(2013-2017)

Month	2013	2014	2015	2016	2017
Jan	20	26	38	25	26
Feb	22	30	39	29	30
March	24	33	35	26	27
Apr	21	25	31	29	30
May	24	22	36	34	34
Jun	19	28	35	38	38
Jul	19	27	30	30	28
Aug	22	26	34	36	34
Sep	19	23	31	31	33
Oct	22	30	36	40	41
Nov	23	34	38	42	40
Dec	26	28	42	38	38

Sources: Hitachi Soe Electric and Machinery Co., Ltd

Production Series for 160 KVA(2013-2017)

Month	2013	2014	2015	2016	2017
Jan	18	20	21	26	26
Feb	12	27	25	26	31
March	17	30	22	23	24
Apr	17	20	26	26	30
May	17	22	25	22	34
Jun	18	27	30	19	39
Jul	14	23	25	25	28
Aug	13	21	30	19	40
Sep	15	16	34	24	32
Oct	19	28	28	28	39
Nov	20	30	35	23	40
Dec	19	37	30	27	38

Sources: Hitachi Soe Electric and Machinery Co., Ltd

Production Series for 400 KVA(2013-2017)

Month	2013	2014	2015	2016	2017
Jan	18	23	24	25	22
Feb	29	25	29	29	32
March	23	25	33	26	29
Apr	18	19	25	29	30
May	15	25	22	34	36
Jun	13	25	28	38	38
Jul	17	25	22	30	34
Aug	12	22	29	35	32
Sep	14	20	22	32	37
Oct	27	27	30	38	38
Nov	29	29	34	40	42
Dec	30	25	30	38	36

Sources: Hitachi Soe Electric and Machinery Co., Ltd

Production Series for 2000 KVA(2013-2017)

Month	2013	2014	2015	2016	2017
Jan	19	29	22	23	21
Feb	20	22	19	20	25
March	17	18	26	25	22
Apr	19	23	21	22	26
May	17	20	28	27	25
Jun	15	21	23	23	30
Jul	18	19	28	23	23
Aug	10	20	24	24	25
Sep	15	22	26	26	32
Oct	22	24	29	28	28
Nov	25	36	32	35	35
Dec	27	22	28	30	30

Sources: Hitachi Soe Electric and Machinery Co., Ltd

Appendix B

Source: Hitachi Soe Electric and Machinery Co.,Ltd (2018)