



وزارة التعليم العالي والبحث العلمي

جامعة بابل

كلية الهندسة

قسم الهندسة المدنية

النماذج التصادفية لبعض خواص الماء في منطقة

الفرات الأوسط في العراق

رسالة مقدمة إلى كلية الهندسة في جامعة بابل كجزء من متطلبات

نيل درجة الماجستير في علوم الهندسة المدنية

من قبل

زيد عبد الرضا عبد

دبلوم عالي هندسة مدنية

آذار

2007

الخلاصة

لقد تم في هذه الدراسة استخدام طريقتين من طرق تحليل السلاسل الزمنية لغرض دراسة المعدلات الشهرية لبعض الخواص الفيزيائية والكيميائية لمياه النهر في كل من محافظات بابل والديوانية والنجف. تتضمن الطريقة الأولى تحليل السلسلة الزمنية إلى مركبة الاتجاه (Trend) و مركبة الدورية (Seasonality) والمركبة المتبقية (Residual). وقد استخدمت المركبة المتبقية لغرض بناء النموذج التنبئي من نوع (ARMA). أما طريقة (Box-Jenkins) للنموذج الفصلي فقد تمت باستخدام نفس السلسلة الزمنية (الأصلية) مباشرة لغرض بناء النموذج, وقد تم مقارنة النتائج المستحصلة من كلا الطريقتين ولوحظ بان طريقة (Box-Jenkins) للنموذج الفصلي ولأغلب الحالات أفضل من الطريقة الأولى (ARMA) في تمثيل البيانات الأصلية، والتي هي حسب علم المؤلف الطريقة الوحيدة المستعملة في الدراسات السابقة المشابهة لهذه الدراسة في العراق، لذا نوصي باستخدام هذه الطريقة (طريقة (Box-Jenkins) للنموذج الفصلي) في الدراسات المستقبلية. وقد استخدمت الطريقتان (طريقة (ARMA) وطريقة (Box-Jenkins)) لغرض إيجاد القيم المستقبلية للخواص الفيزيائية والكيميائية ولعشر سنوات تبدأ من نهاية شهر كانون الأول لسنة 2001 وقد بينت النتائج بان بعض هذه القيم تتجاوز حدود المواصفات العراقية وهي (Mg) في محطة الهندية و (T.A & Mg) في محطة الديوانية، بينما بعض القيم تبقى ضمن حدود المواصفات العراقية وهي (Ca, Cl, EC, T.H) في محطة الهندية و (T.A, Ca, Cl, EC, T.H, Mg) في محطة الكوفة و (EC, Ca, Cl) في محطة الديوانية.

MINISTRY OF HIGHER EDUCATION
AND SCIENTIFIC RESEARCH
UNIVERSITY OF BABYLON
COLLEGE OF ENGINEERING
DEPARTMENT OF CIVIL ENGINEERING



STOCHASTIC MODELS OF SOME PROPERTIES OF
WATER IN THE MIDDLE OF
EUPHRATES REGION
IN IRAQ

A THESIS
SUBMITTED TO THE COLLEGE OF ENGINEERING
OF THE UNIVERSITY OF BABYLON IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER
OF SCIENCE IN CIVIL
ENGINEERING

BY
ZAID ABED AL-RIDAH ABED

(H. Dip. Sc.)

MAY

2007

We certify that preparation of this thesis was carried out under our supervision at University of Babylon in partial fulfillment of the requirement for the degree of Master of Science in Civil Engineering.

Signature :

Name : Asst. Prof. Dr.

Salah Tawfeek

(supervisor)

Date : / /2007

Signature :

Name : Asst. Prof. Dr.

Mohammad A. Muslim

(supervisor)

Date : / /2007

CERTIFICATION

We certify as an examine committee that we have read this thesis entitled (***Stochastic Models of Some Properties of Water in the Middle of Euphrates Region in Iraq***) and examined the student (***Zaid Abed AL-Ridah Abed***), in its content and what related to it, and found it meets the standard of a thesis for the degree of Master of Science in Civil Engineering (Environmental Engineering).

Signature :

**Name: Asst. Prof. Dr. Saad A. Madhi
(Member)**

Date : / /2007

Signature :

**Name : Asst. Prof. Dr. Alaa H. Al-Fatlawi
(Member)**

Date : / /2007

Signature :

**Name: Asst. Prof. Dr. Kareem Radhi Al-Murshidi
(Chairman)**

Date : / /2007

Approval of Civil Engineering Department

Signature :

Name : Asst. Prof. Dr. Ammar Y. Ali

Head of Civil Engineering Department .

Date : / /2007

Approval of Deanery of the College of Engineering .

Signature :

Name : Prof. Dr. Abd Al-Wahid K. Rajih

Deane of the College of Engineering.

Date : / /2007



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

﴿ وَلَهُ ۖ الْمَشْرِقُ وَالْمَغْرِبُ فَأَيْنَمَا تُوْكَفَتُمْ وَجْهُ اللَّهِ
إِنَّ اللَّهَ وَاسِعٌ عَلِيمٌ ﴾ البقرة - سُورَةُ الْبَقَرَةِ

صَلَّى اللَّهُ
عَلَيْهِمُ الْعَظِيمِ



Acknowledgement

Pries is to ALLAH and to his prophet Mohammad. This research is completed under their benediction.

Cordial thanks and deepest gratitude to my supervisor Asst. Prof. Dr. Salah Tawfeek for his valuable guidance and his forceful encouragement throughout the preparation of this work.

I would like to express my sincere gratitude to my supervisor Asst. Prof. Dr. Mohammad A. Muslim for his guidance and kindness throughout this work.

I am grateful to the Babylon, Najaf, and Diwaniya Water Office, especially to the staff lab members.

I am also record my thanks to University of Babylon, College of Engineering and Civil Engineering Department staff for all facilities they rendered me especially Miss Eman AL-Mousawi.

I record my sincere gratitude to my family for their encouragement and support during the study.



Zaid
2007

Abstract

In this study two different methods of time series analyses are applied to monthly records of some physical and chemical properties of river water in Babylon, Najaf and Diwaniya governorates. The first method involves decomposition of the historical series into trend, seasonality and residual components. Stochastic ARMA models are then fitted to the residuals. In the second method, Box-Jenkins seasonal models are directly fitted to the same observed (original) time series and these models are compared with the ARMA models .It is found that, in most cases, the Box-Jenkins models fit the observed series better than the ARMA models which are, to the researcher best knowledge, the only models used in the previous similar studies in Iraq. Therefore, it is recommended that the Box –Jenkins seasonal models should also be considered in future studies. Forecasts of the future values of the physical and chemical properties of these waters for ten years starting from the last observation in December 2001 are also made by using both ARMA and Box –Jenkins models. It is found that, some of these values exceed the Iraqi Standard limits i.e. Mg at AL-Hindya station and T.A &Mg at Diwaniya station, whereas some values stay within the standard limits i.e. (Ca, Cl, EC & T.H at AL-Hindya station), (T.A, Ca, Cl, EC, T.H & Mg at Kufa station) and (Ca, Cl & EC at Diwaniya station).

Table of Contents

Subject	Page
Acknowledgment	I
Abstract	II
Table of contents	III
Table of figures	VI
Table of symbols	IX
Table of tables	XII
Chapter One :Introduction	2
1.1 Prelude	2
1.2 Description of study area and data :	4
1.3 Objectives of the study	5
Chapter Two :Stochastic Methods	8
2.1 Components of time series	8
2.2 Stationarity	9
2.3 Homogeneity of time series	10
2.4 Trend and jump detection	10
2.5 Removal of trend and jump component	12
2.6 Detection of periodic or seasonal component	13
2.7 Removal of periodic or seasonal component	14
2.8 Stochastic component	17
2.9 ARMA(p,q) stochastic models	17
(i)Autoregressive model AR(p) or ARMA(p,0)	18
(ii) Moving average model MA(q) or ARMA(0,q)	18

Subject	Page
(iii) Mixed models	19
2.10 The Box-Jenkins seasonal model	20
2.11 ARMA stochastic model building	21
1- Model identification	21
2-Initial estimation of model parameters	22
3-Final estimation of model parameters	23
4- Model diagnostic checking	26
Port Manteau lack of fit test	26
Residual autocorrelation function (RACF) test	27
2.12 Forecasting	27
2.13 Generation of synthetic series	30
Chapter Three :Literature Review	32
3.1 Some studies on time series from other countries	32
3.2 Previous application of time series methods to data from Iraq	39
Chapter Four: Application of ARMA Model	43
4.1 ARMA model	43
4.2 Test and removal of nonhomogeneity	45
4.3 Transforming data to normal distribution	51
4.4 Detection and removal of periodic components	63
4.4 Stochastic model	71
4.4.1 Model identification	71
4.4.2 Estimation of model parameters	71

Subject	Page
4.4.3 Diagnostic checking	80
Chapter Five :Fitting Box – Jenkins Seasonal Model	87
5.1 Identification of models	87
5.2 Estimation of parameters	90
5.3 Diagnostic check	104
5.4 Confidence region	104
Chapter Six :Forecasting	108
6.1 Forecasting using ARMA models	108
6.2 Forecasting using Box-Jenkins seasonal model	112
6.3 Mean absolute relative errors	117
6.4 Comparison between ARMA and Box-Jenkins models	117
6.5 Comparison of forecasted data with the river standards	121
6.6 Comparison between the present study and previous studies in Iraq	122
Chapter Seven :Conclusions and Recommendations	127
7.1 Conclusions	127
7.2 Recommendations	128
References	130
Appendix (A)	135
Appendix (B)	138

Table of Figures

Fig. No.	Description	Page
1.1	Map of studying area	6
4.1	Steps of ARMA method	44
4.2	Annual means of observed AL-Hindya data before removing variation in mean form 1992 to 2001.	46
4.3	Annual standard deviations of observed AL-Hindya data before removing variation in standard deviation form 1992 to 2001 .	47
4.4	Annual means of homogenous data of AL-Hindya station after removing variation in mean from 1992 to 2001 .	50
4.5	Annual standard deviation of homogenous data of AL-Hindya station after removing variation in standard deviation from 1992 to 2001 .	50
4.6	Normal distribution test of AL-Hindya station parameters .	62
4.7	Correlogram of Normalized series of AL-Hindya Station.	64
4.8	Periodic component of monthly mean of AL-Hindya station.	69
4.9	4.9 Periodic component of monthly standard deviation of AL-Hindya station.	70
4.10	ACF & PACF of T.A, Ca and Cl of AL-Hindya station.	72
4.11	ACF&PACF of EC, T.H, Mg and Tur of AL-Hindya station.	73

Fig. No.	Description	Page
4.12	Autocorrelogram of residual series of AL-Hindya station parameters.	84
5.1	Monthly observed values of T.A, Ca ,Cl and EC of AL-Hindya station from 1992 to 2001.	88
5.2	Monthly observed values of T.H , Mg and Tur of AL-Hindya station from 1992 to 2001.	89
5.3	ACF&PACF T.A, Ca and Cl of AL-Hindya station .	91
5.4	ACF&PACF EC, T.H, Mg and Cl of AL-Hindya station.	92
5.5	Confidence region of model parameters of EC of Kufa station	105
6.1	Probability limits of forecasted values obtained by ARMA model of T.A ,Ca ,Cl &EC of AL-Hindya station from 2002 to 2004 .	113
6.2	Probability limits of forecasted values obtained by ARMA model of T.H, Mg, Tur of AL-Hindya station from 2002to 2004 .	114
6.3	Probability limits of forecasted values obtained by Box-Jenkins seasonal model of T.A ,Ca ,Cl &EC of AL-Hindya station from 2002 to 2004	115
6.4	Probability limits of forecasted values obtained by Box-Jenkins seasonal model of T.H, Mg, Tur of AL-Hindya station from 2002 to 2004)	116
6.5	Comparison between observed series and the series generated by (I-ARMA, II-Box-Jenkins) methods for T.A ,Ca ,Cl &EC at AL-Hindya station	119

Fig. No.	Description	Page
6.6	Comparison between observed series and the series generated by (I-ARMA, II-Box-Jenkins) methods for T.H, Mg, Tur at AL-Hindya station	120
6.7	Comparison of 7-years(2005-2011)forecast of T.A, Ca, Cl &EC for AL-Hindya station with river standard	123
6.8	Comparison of 7-years(2005-2011)forecast of T.H ,Mg & Tur for AL-Hindya station with river standard	124
Note : Other figures are given in appendix (B)		

Table of Symbols

Symbol	Definition
A, B, A_1, B_1	Regression coefficients
ACF	Autocorrelation function
AR(p)	Autoregressive model of order (p)
ARIMA (p,d,q)	Autoregressive integrated moving average model of order (p) , (q) and degree of differencing (d)
A_h, B_h	Fourier coefficients
a_t	The shock (independent part of stochastic process)
AIC	Akaike information criterion
c_k	Autocovariance coefficient at lag k
c_o	Autocovariance coefficient at lag zero
C_k	Coefficients of kurtosis
C_s	Coefficients of skewness
Ca	Calcium
Cl	Chloride
EC	Electrical conductivity
$F_{i,j}$	Transformed series at year (i) and month (j)
J_t	Jump component
$l(\phi, \theta, \sigma_a)$	Log likelihood of ARIMA model
N	No. of observations
M_2	Mean of second sub-sample
MARE%	Mean absolute relative error
Mg	Magnesium
n_1, n_2	Number of years of first and second sub-sample respectively
PACF	Partial autocorrelation function
P_h	Variance explained by (h) harmonics
P_t	Periodic or seasonal component

Symbol	Definition
p_n	Number of turning points
$\bar{p}_n$	Expected number of turning points
$\text{Var } \bar{p}_n$	Variance of $\bar{p}_n$
$p_{\max}, p_{\min}$	Critical values of harmonically variance
p, q	Order of AR, MA respectively
Q	Statistics
r_k	Correlation coefficient at lag k
Sd_2	Standard deviation of second sub-sample
S_j	Annual Sd at year j
$S(\phi, \theta)$	Unconditional sum of squares
$S^*(\phi, \theta)$	Conditional sum of squares of ARIMA models
$S(\beta)$	Quadratic surface
SS	Sum of squares
$S(\hat{\beta})$	Minimum sum of squares
T.A	Total Alkalinity
T.H	Total Hardness
Tur	Turbidity
W_t	Stochastic component at time t
X_t	Time series observation at time t
$X_{i,j}$	Original or homogenous series at month (j) and year (i)
$\bar{X}_i$	Annual mean at year i
$\bar{X}$	Over all mean
$Y_{i,j}$	Homogeneous series at month (j) and year (i)
$Y'_{i,j}$	The series after removing seasonal component at month (j) and year (i)
Z	Standard normal variate

Symbol	Definition
$\hat{Z}_t(\ell)$	Forecast series at origin time t and lead time ℓ
β	Backward shift operator
∇	Back difference coefficient
σ_a	Residual variance
σ_j	The standard deviation at season j
μ_j	Harmonically fitted means at season j
$\hat{\sigma}_i$	Harmonically fitted variances at season j
∇_s	Seasonal difference at season s
$\phi, \Phi, \theta, \Theta$	Parameters of the model
λ	Power transformation constant
$\chi^2_{\xi}(k)$	The significant point expected by a proportion ξ of the χ^2 distribution, having k degree of freedom

Table of Tables

Table No.	Description	Page
1.1	List of water quality parameter which are used in the analysis	4
4.1, 4.2 & 4.3	Results of nonhomogeneity test of AL-Hindya , Kufa and Diwaniya station respectively	48
4.4	Regression coefficients before removing nonhomogeneity.	49
4.5	Results of nonhomogeneity test after removal of nonhomogeneity	51
4.6, 4.7 & 4.8	The values of the first four moments of homogenous series of AL-Hindya , Kufa and Diwaniya station respectively	52
4.9, 4.10 & 4.11	Effect of (λ) values the first four moments of series in AL-Hindya , Kufa and Diwaniya station respectively	54-59
4.12, 4.13 & 4.14	Values of statistical parameters of normally distributed series of AL-Hindya , Kufa and Diwaniya station respectively	59-60
4.15, 4.16 & 4.17	Comparison among the three transformation procedure of the data in AL-Hindya , Kufa and Diwaniya station respectively	61
4.18, 4.19 & 4.20	Results of harmonic analysis for monthly means of AL-Hindya , Kufa and Diwaniya station respectively	65-66
4.21, 4.22 & 4.23	Results of harmonic analysis of monthly standard deviation at AL-Hindya , Kufa and Diwaniya station respectively	67-68
4.24	Number of significant harmonics of mean and standard deviation of AL-Hindya, Kufa and Diwaniya station respectively	68

Table No.	Description	Page
4.25, 4.26 & 4.27	Initial estimation of model parameter of AL-Hindya , Kufa and Diwaniya station respectively	74-76
4.28	Conditional sum of squares of Ca of AL-Hindya station where the model is AR(1) and $\phi=0.609$.	77
4.29, 4.30 &4.31	Final estimation of model parameter of AL-Hindya , Kufa and Diwaniya station respectively	78-80
4.32	AIC value of selected model of AL-Hindya , Kufa and Diwaniya respectively .	81
4.33	Diagnostic check results with 95% confidence limits and (M-p-q) degree of freedom , where M=48	83
4.34	Final model types of each water quality parameter of AL-Hindya , Kufa and Diwaniya stations.	85
5.1& 5.2	Calculation of sum of squares of EC of Kufa station with ($\theta = 0.1$, $\Theta = 0.1$) and ($\theta = 0.55$, $\Theta = 0.895$) respectively	94-101
5.3	Estimated models parameters and sum of square	102
5.4	Seasonal model equations of the three stations.	103
5.5	Diagnostic check results of the three stations with k = 48 & degree of freedom (k –q – Q) at 95% confidence limits .	104
6.1, 6.2 &6.3	ARMA models equations of AL-Hindya , Kufa and Diwaniya station parameters respectively	108-110
6.4	Calculation of forecasting values of Ca of AL-Hindya station by using ARMA model.	111
6.5	The calculation of MARE% for the Ca parameter at Kufa station.	117
6.6	MARE% results of the three stations	116
6.7	Methods and models used for 7 years forecast	118

Table No.	Description	Page
6.8	Allowable limits of water quality parameters in surface water body used as drinking water source according to Iraqi specification.	121
6.9	Future values of T.A of AL-Hindya station from 2005 to 2011.	122
6.10	Comparison between the present study and the previous studies.	125

1-t-test

The test for homogeneity in mean of two sub-samples is given as follows (AL-Mousawi , 2003):

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S * \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}}} \quad \dots\dots\dots A(1)$$

Where:

$$S = \sqrt{\frac{\sum_{i=1}^{n_1} (X_i - \bar{X}_1)^2 + \sum_{j=1}^{n_2} (X_j - \bar{X}_1)^2}{n_1 + n_2 - 2}} \quad \dots\dots\dots A(2)$$

Where :

$\bar{X}_1, \bar{X}_2$:Means of first and second sub –samples respectively .

n_1, n_2 :Number of years in the first and second sub –samples respectively.

X_i, X_j :Annual values of the first and second sub –samples respectively.

If the calculated value of **t** by equation A(1) is less than the critical value of **t** at a certain probability level then, there is no significant difference in the mean and vise versa .

2-f-test

The test for homogeneity in the variance of two sub-samples is given as follows (AL-Mousawi , 2003):

$$f = \frac{s_1^2}{s_2^2} \dots\dots\dots A(3)$$

With $V_1=n_1-1$, $V_2=n_2-1$

Where :

s_1, s_2 :Standard deviation of the first and second sub –samples respectively and s_1 the value of the largest one .

V_1, V_2 : Degree of freedom of the first and second sub –samples respectively.

If the calculated value of f by equation A(3) is less than the critical value of f at a certain probability level then there is no significant difference in the variance and vise versa .

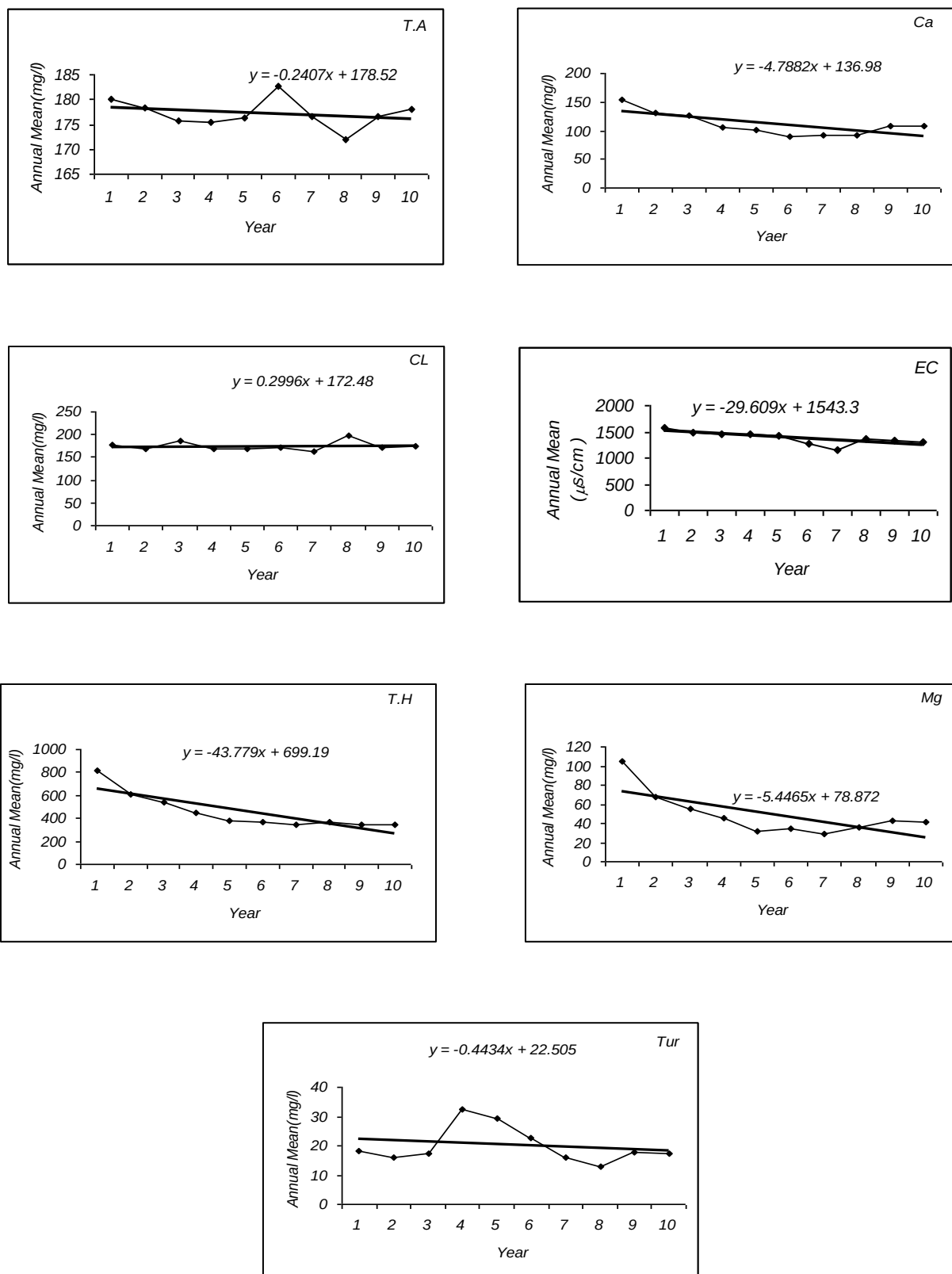


Fig. B.1:Annual means of observed Kufa data before removing variation in mean from 1992to2001 .

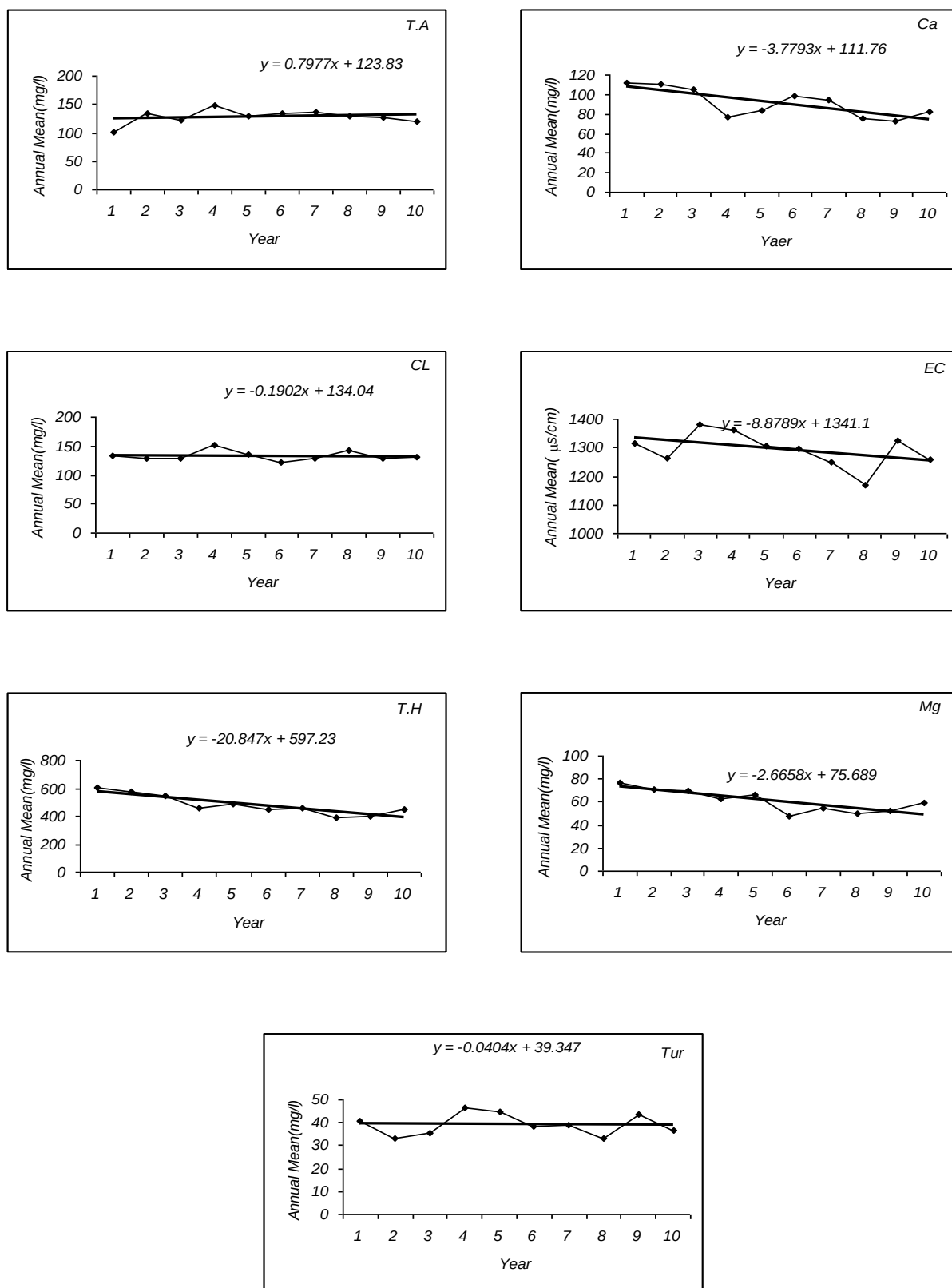


Fig. B.2: Annual means of observed Diwaniya data before removing variation in mean from 1992 to 2001 .

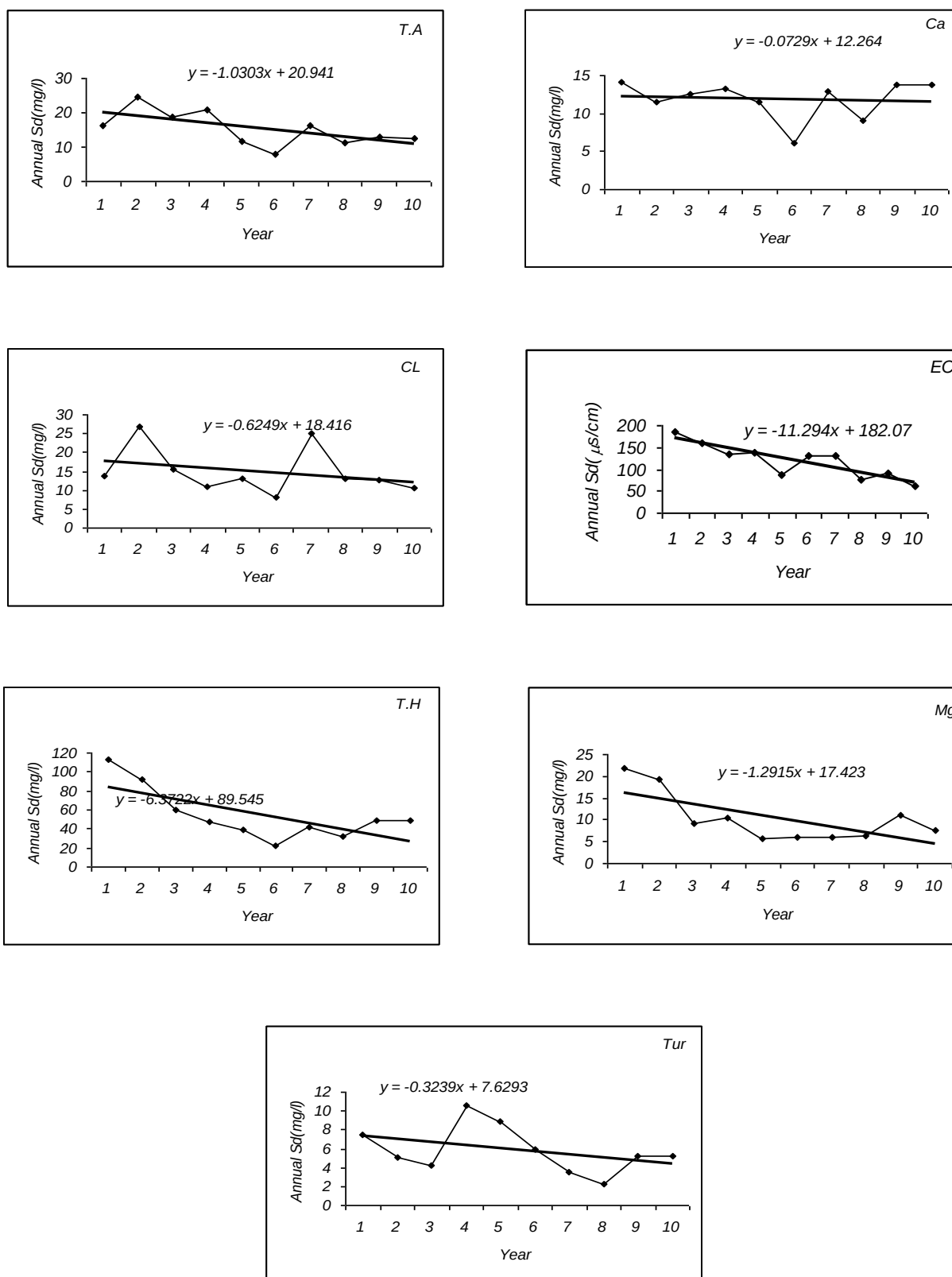


Fig. B.3: Annual standard deviations of observed Kufa data before removing variation in standard deviation from 1992 to 2001.

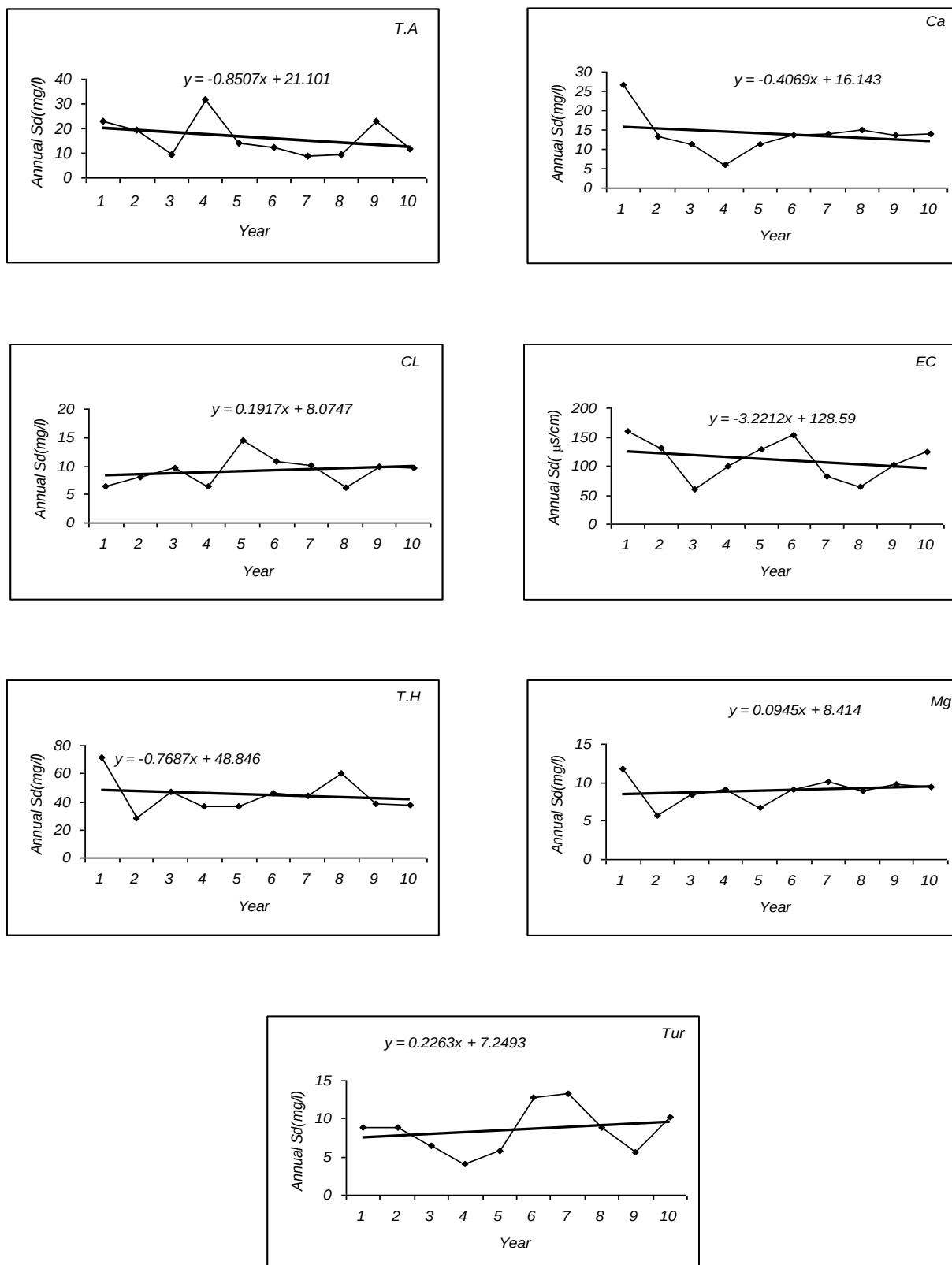


Fig. B.4: Annual standard deviations of observed Diwaniya data standard deviation from 1992 to 2001.

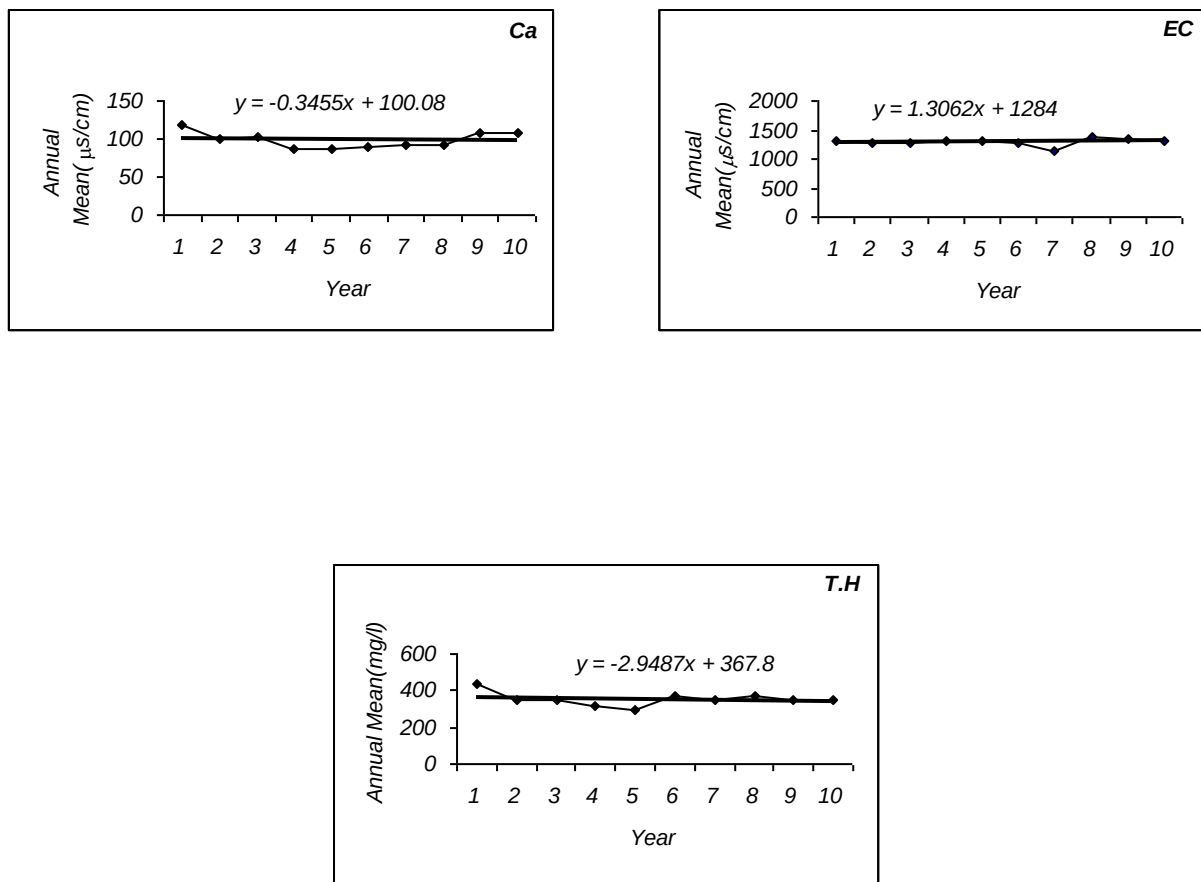


Fig. B.5: Annual means of homogenous data of Kufa station after removing variation in mean from 1992 to 2001.

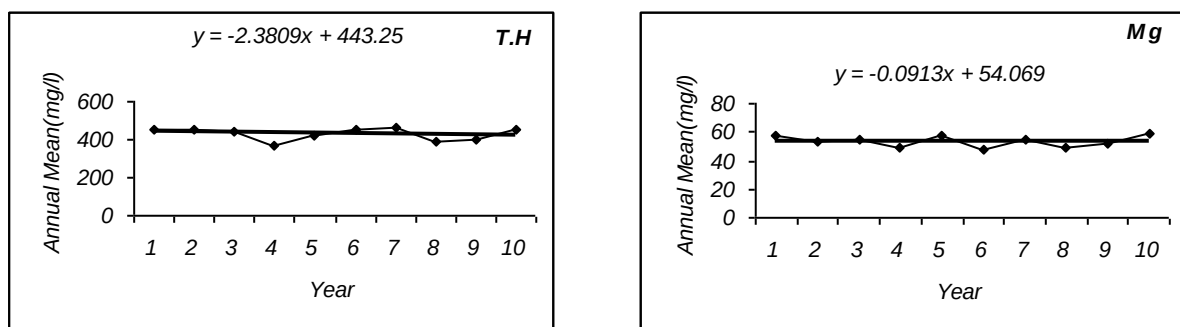


Fig. B.6: Annual of homogenous data of Diwaniya station after removing variation in mean from 1992 to 2001.

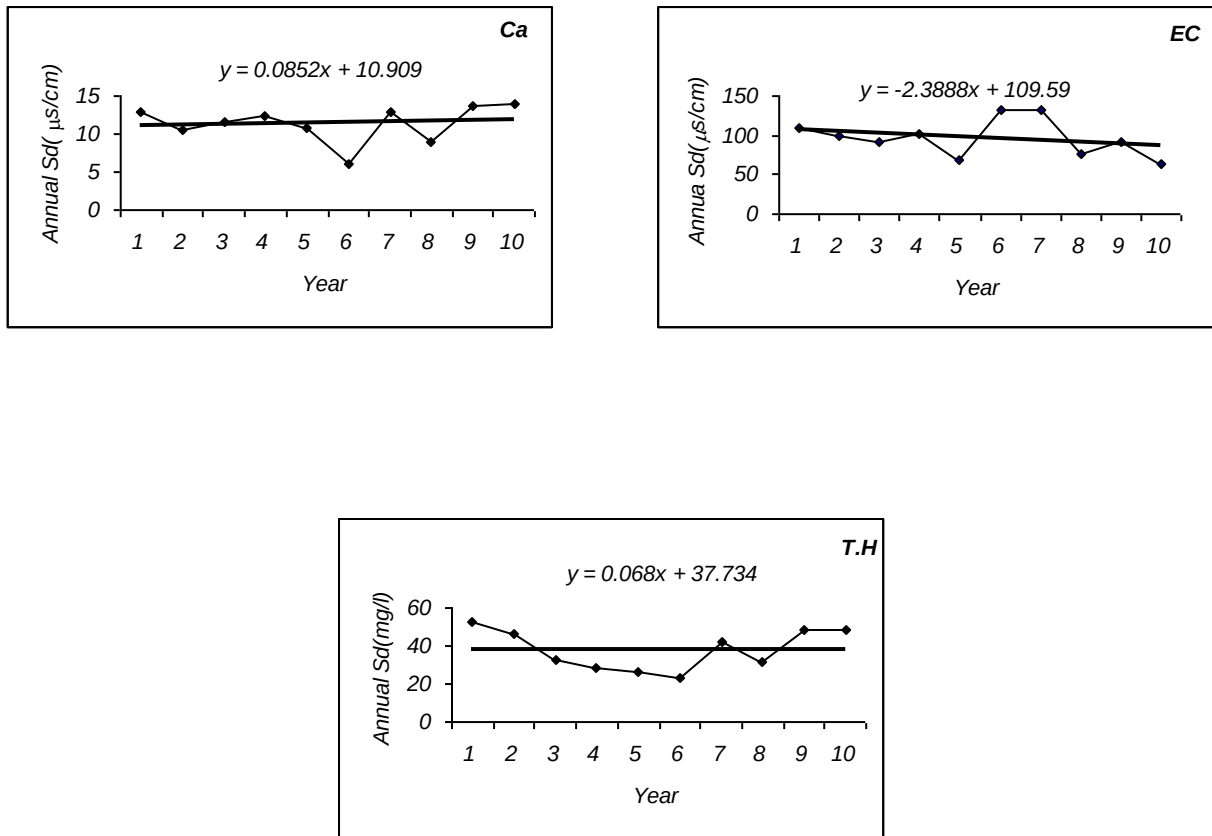


Fig. B.7: Annual standard deviations of homogenous data of Kufa station after removing variation in standard deviation from 1992 to 2001.

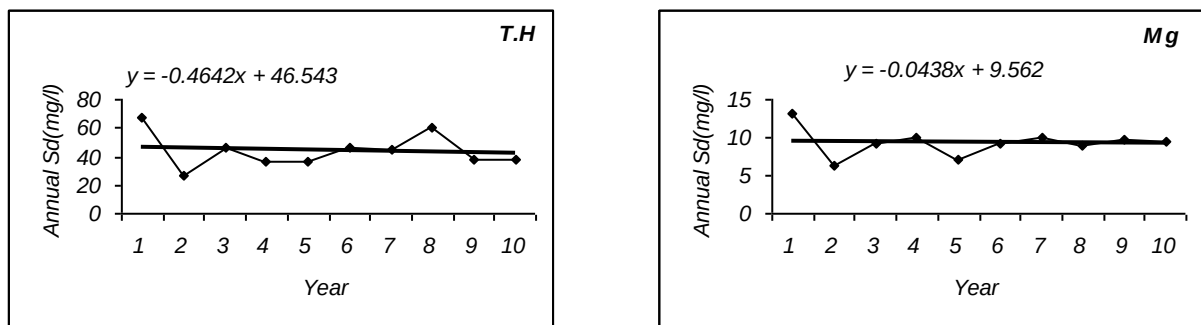


Fig . B.8 :Annual standard deviation of homogenous data of Diwaniya station after removing variation in standard deviation from 1992 to 2001 .

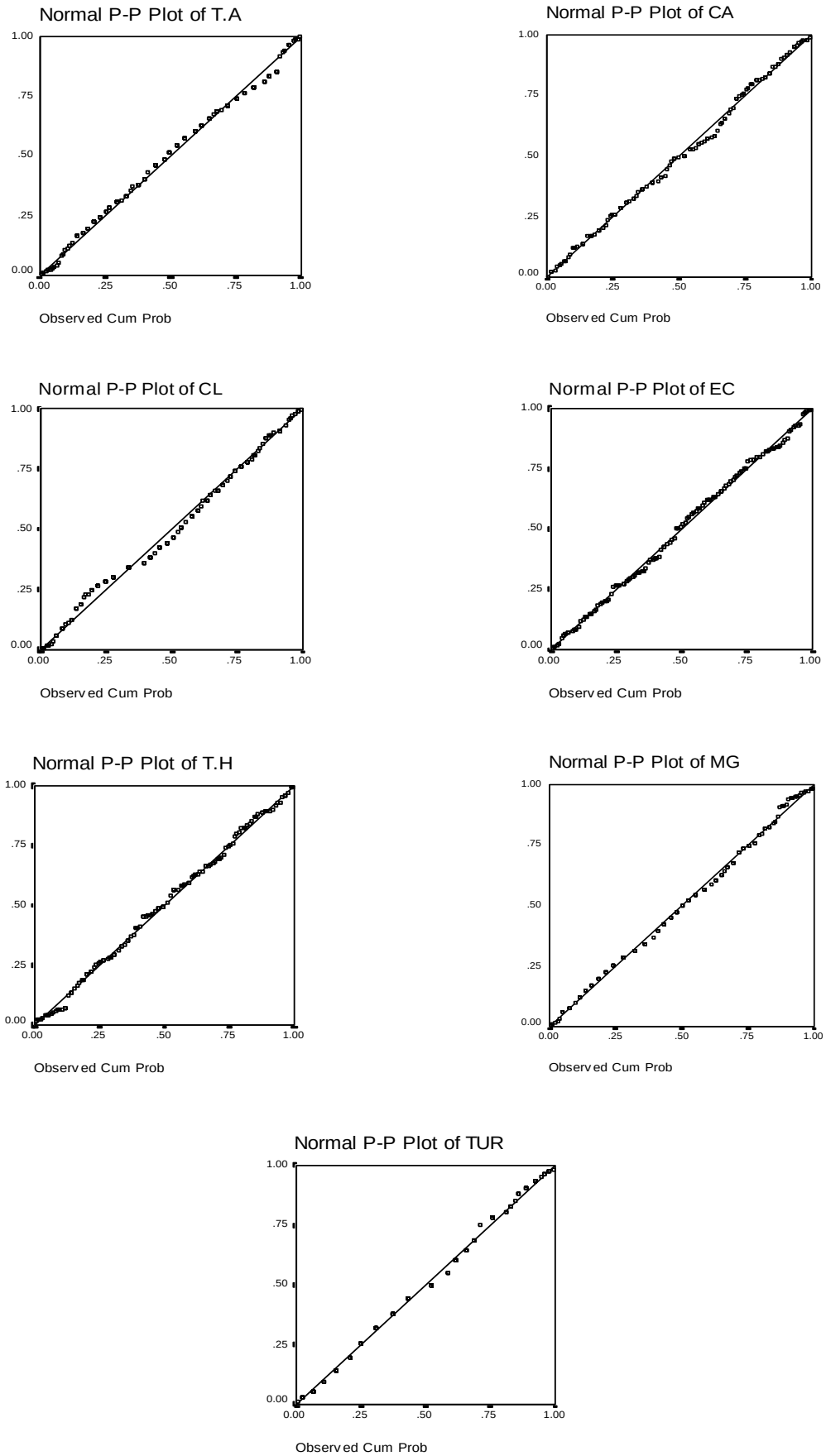


Fig. B.9:Normal distribution test of Kufa station parameters .

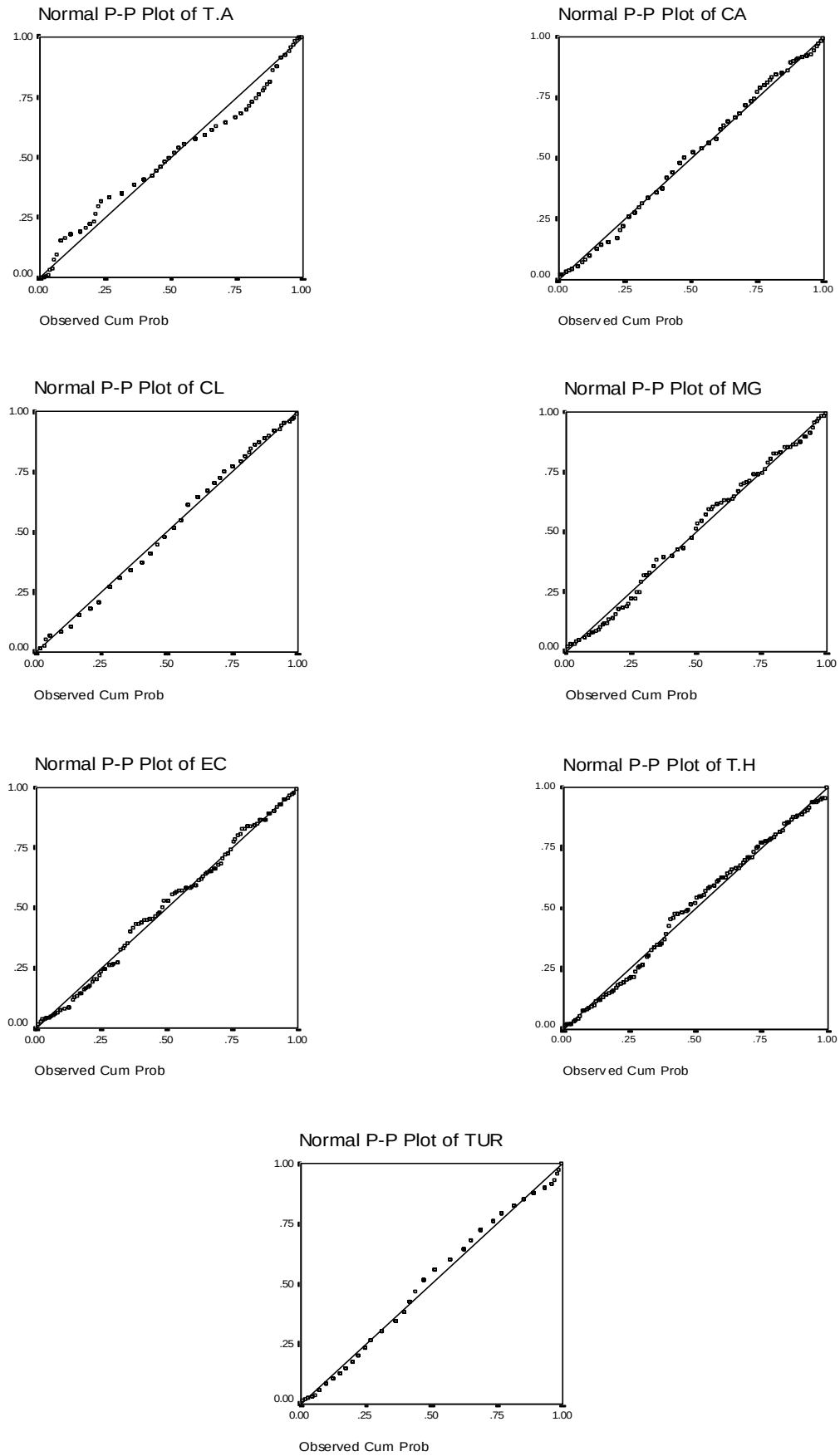


Fig. B.10:Normal distribution test of Diwaniya station parameters .

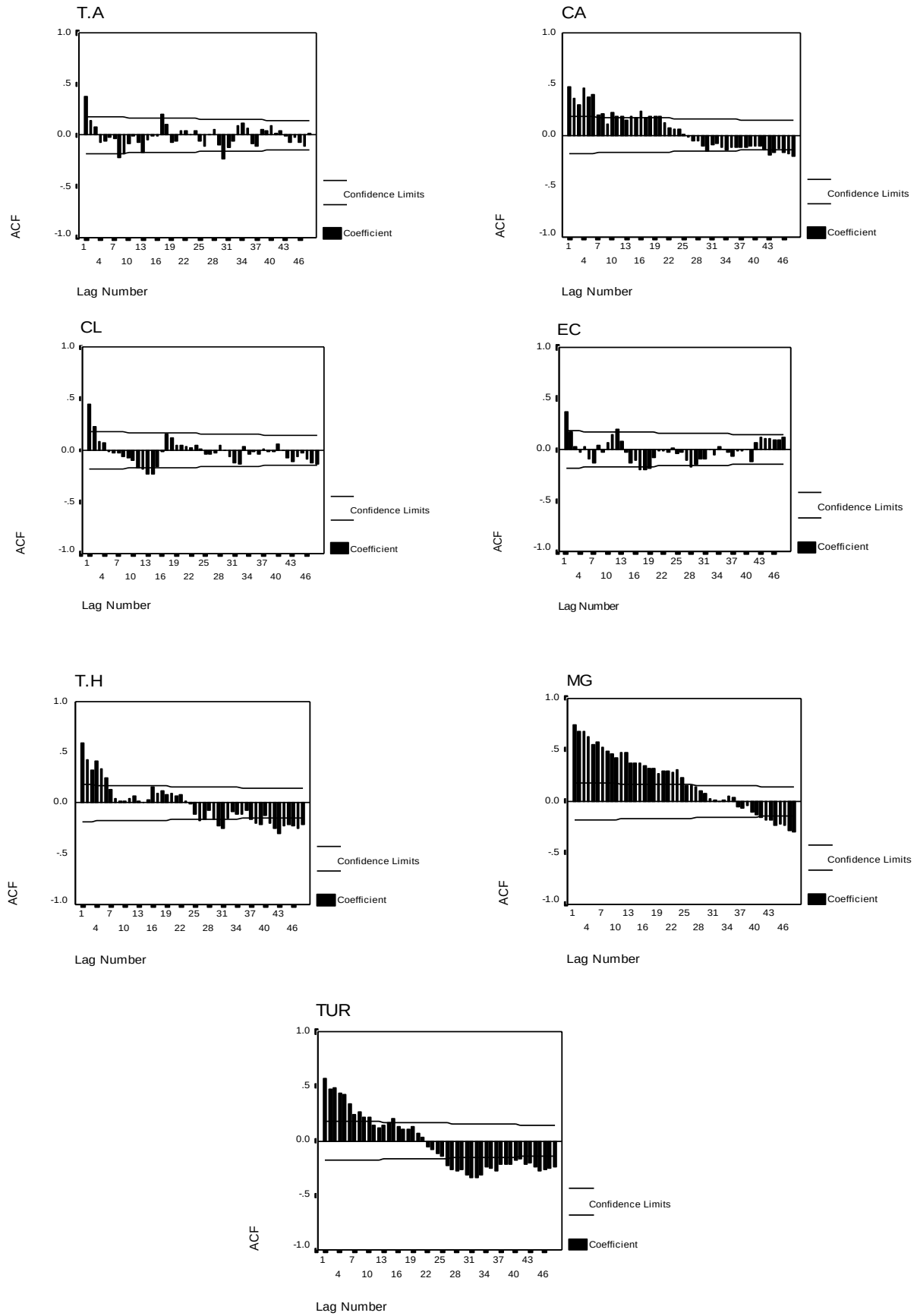


Fig. B.11 : Correlogram of normalized series of Kufa Station.

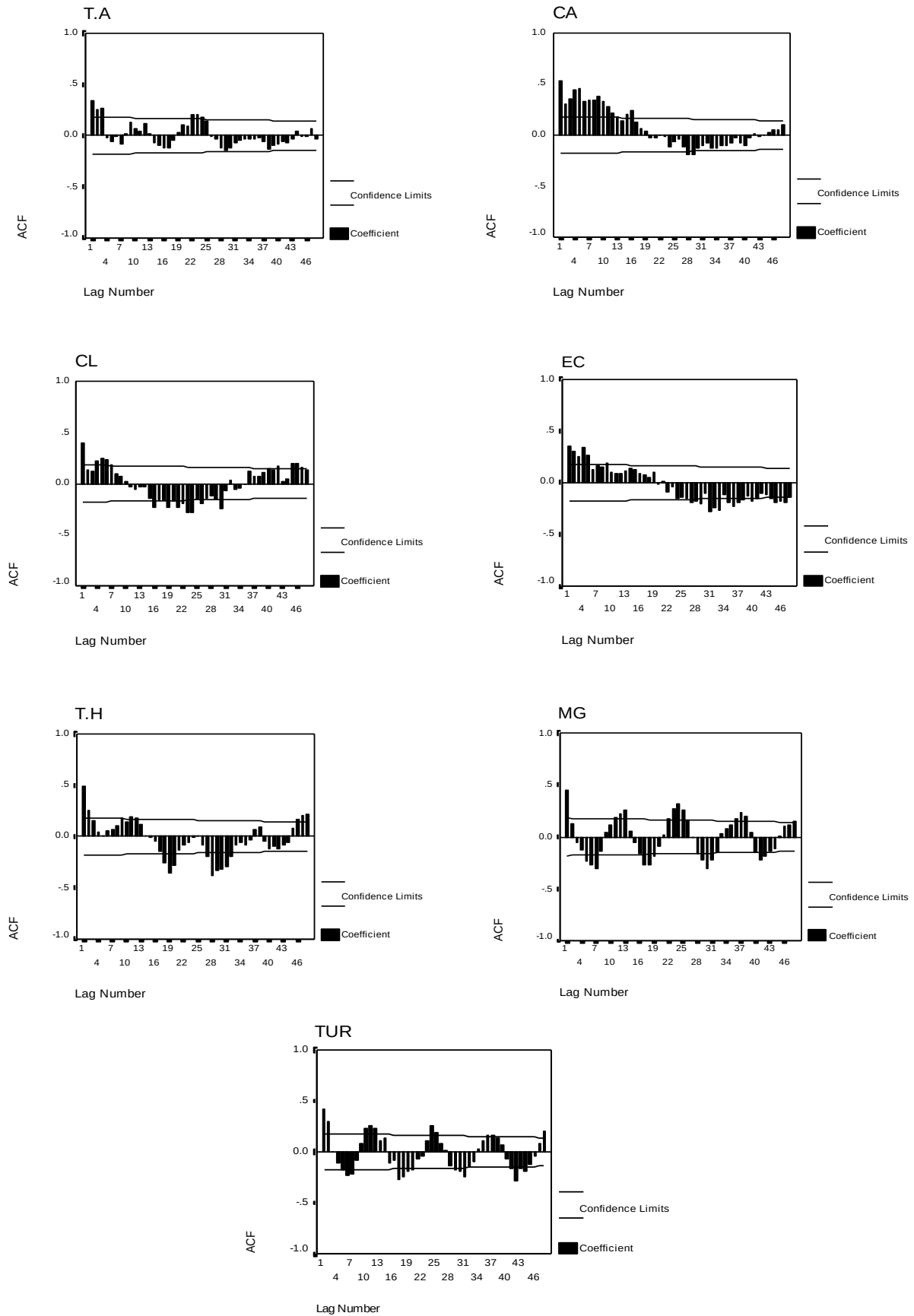


Fig. B.12 : Correlogram of normalized series of Diwaniya Station.

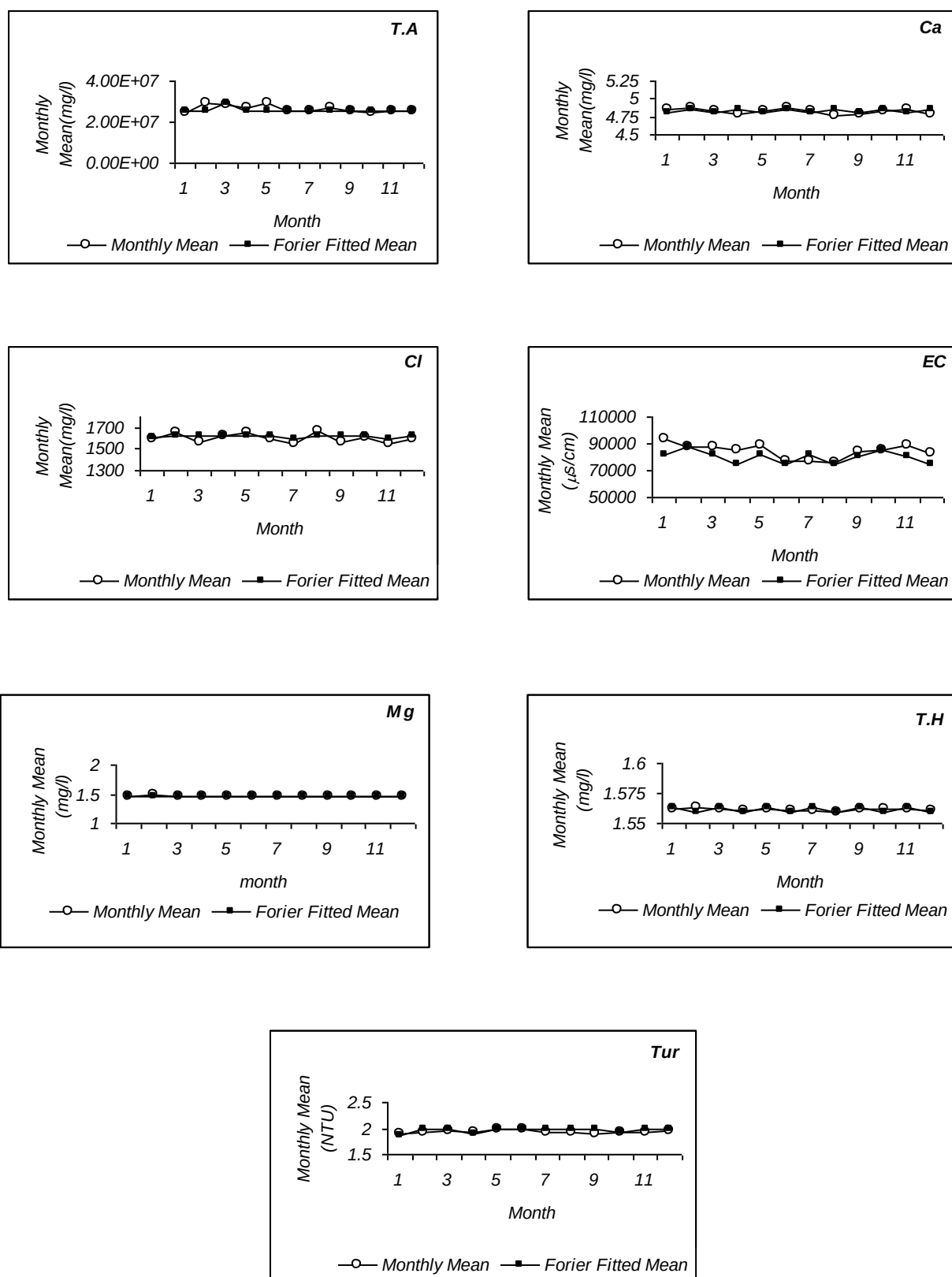


Fig. B.13: Periodic component of monthly mean of Kufa station.

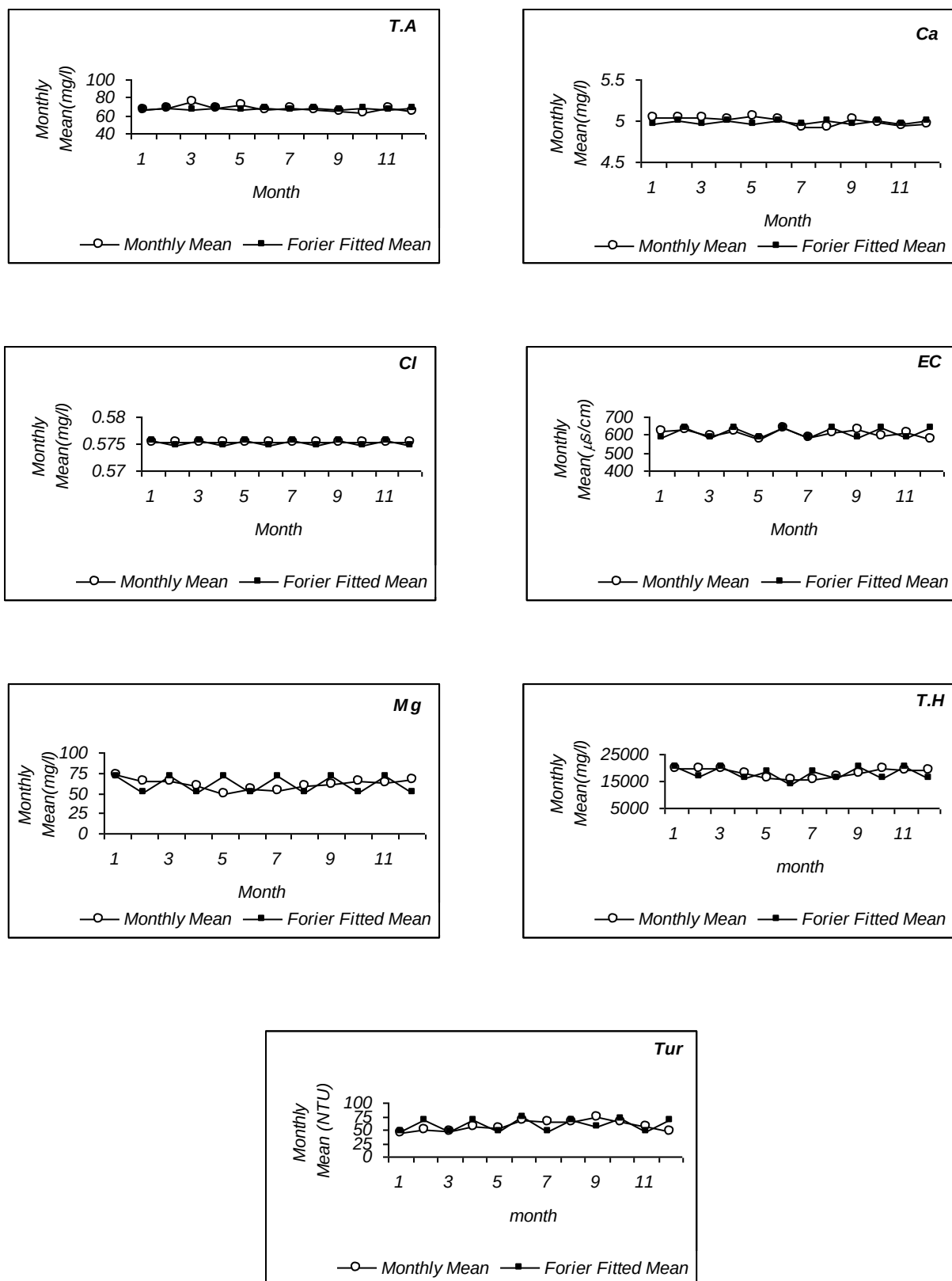


Fig. B.14: Periodic component of monthly mean of Diwaniya station.

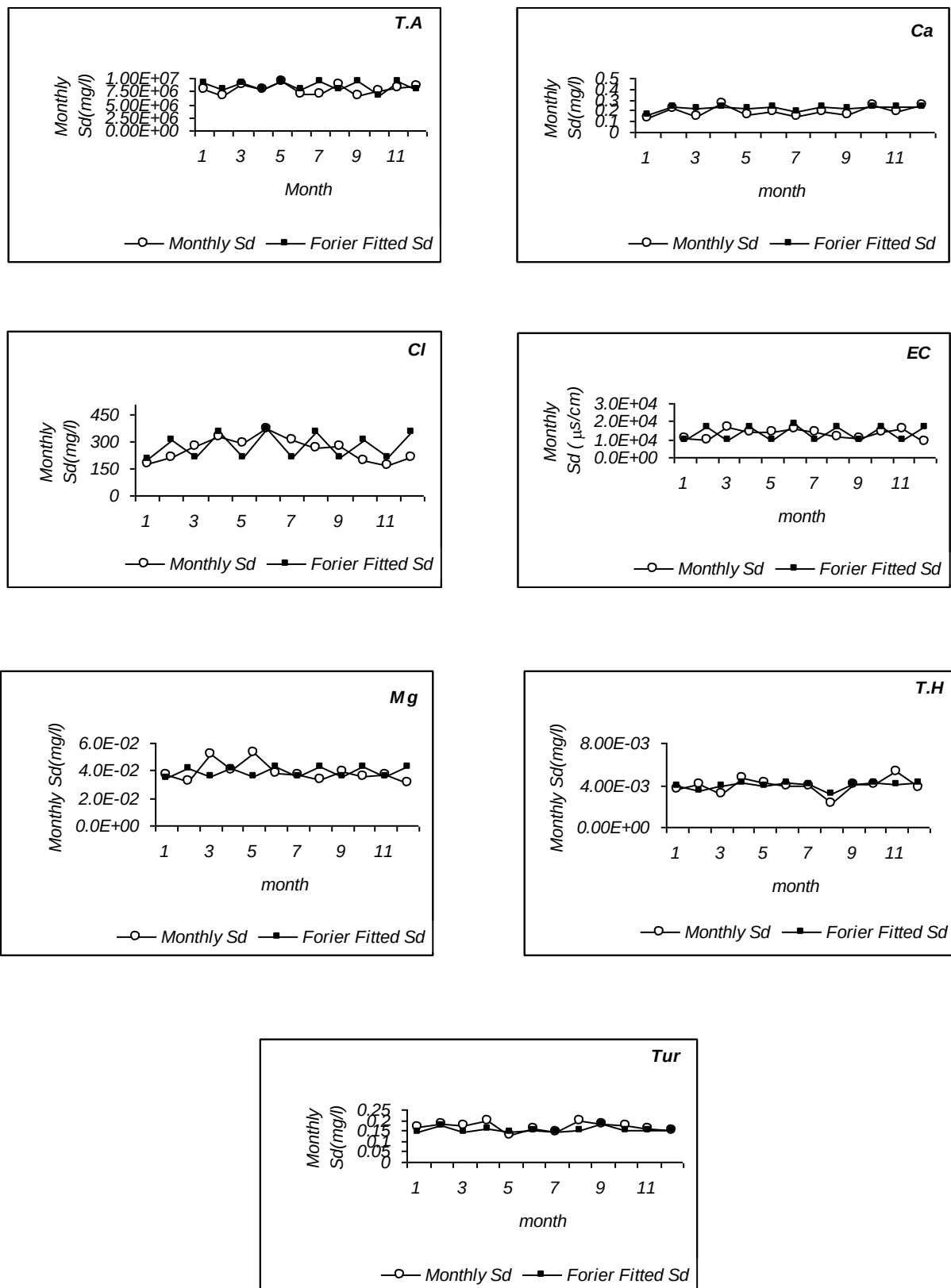


Fig. B.15: Periodic component of monthly standard deviation of Kufa station.

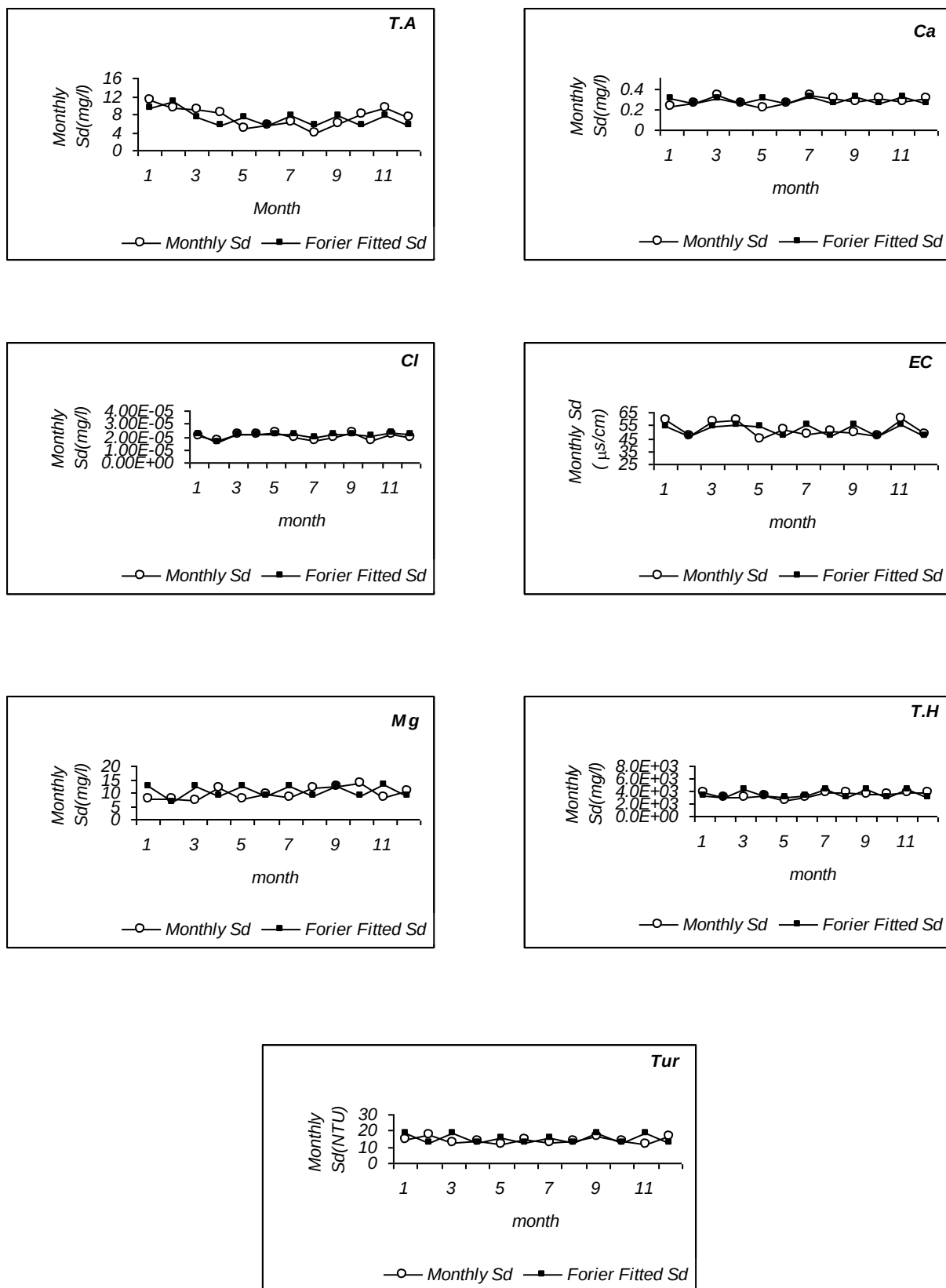


Fig. B.16: Periodic component of monthly standard deviation of Diwaniya station.

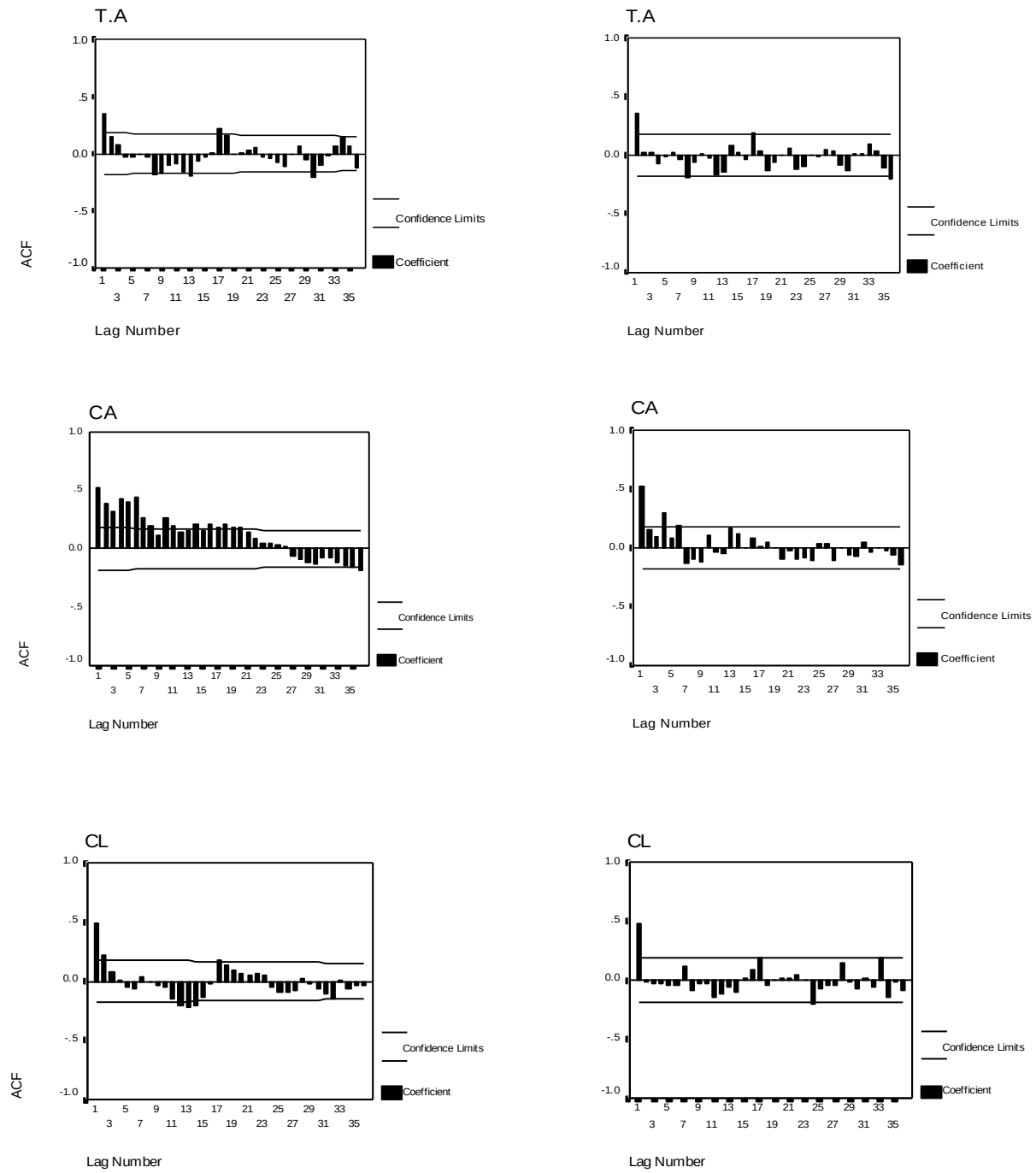


Fig. B.17:ACF&PACF of T.A, Ca and Cl of Kufa station parameters.

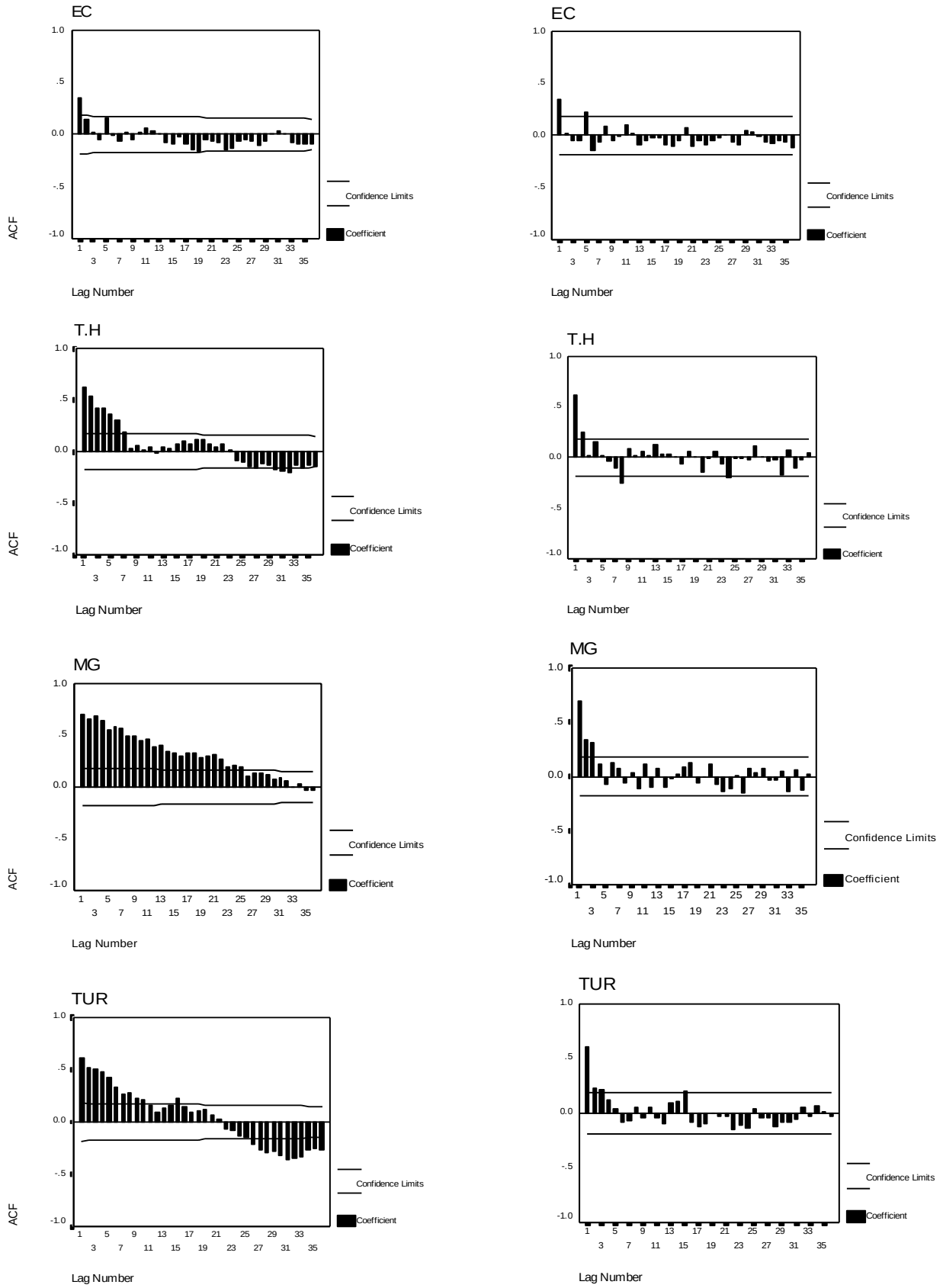


Fig. B.18: ACF&PACF of EC, T.H, Mg and Tur of Kufa station.

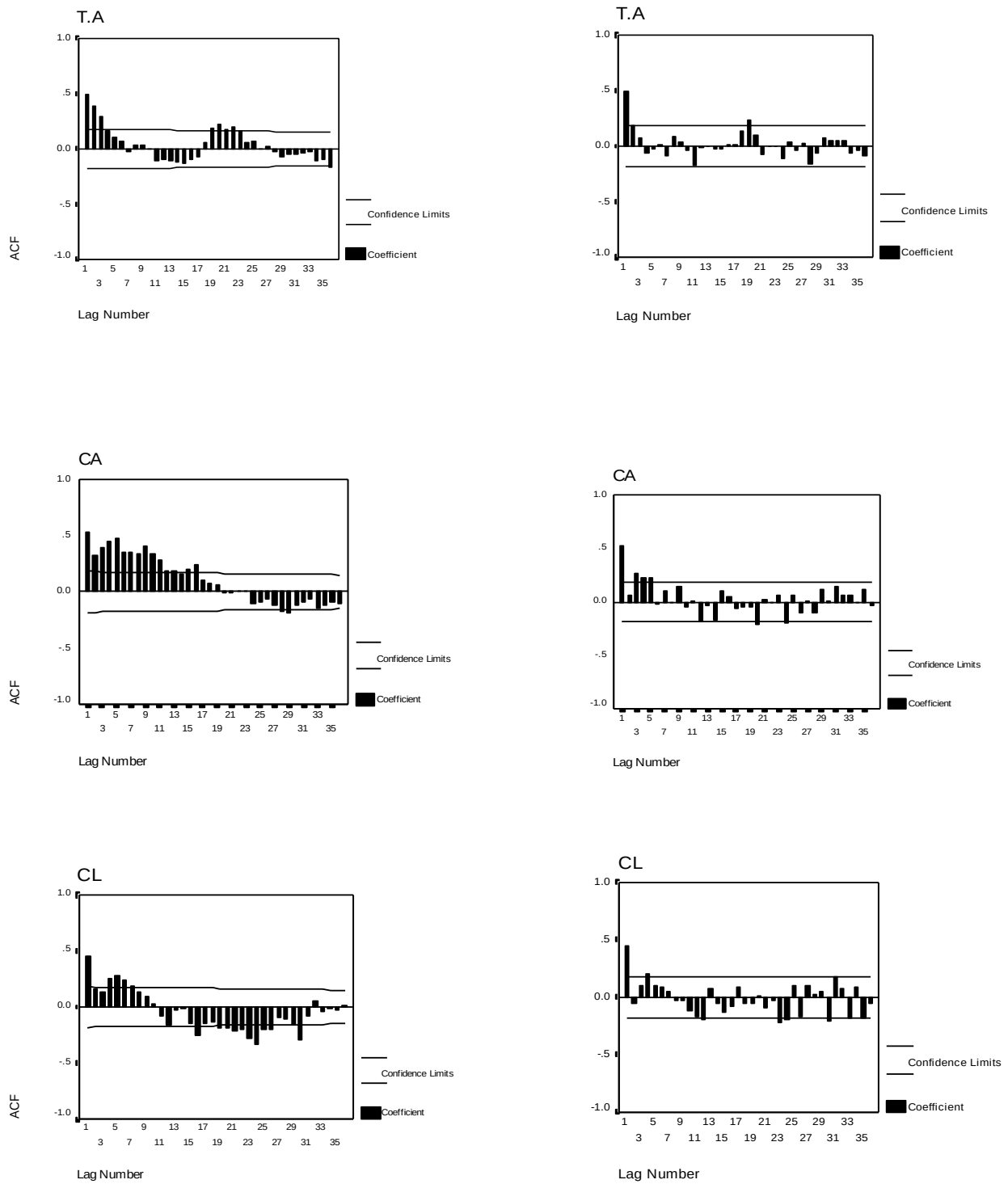


Fig B.19: ACF&PACF of T.A, Ca and Cl of Diwaniya station.

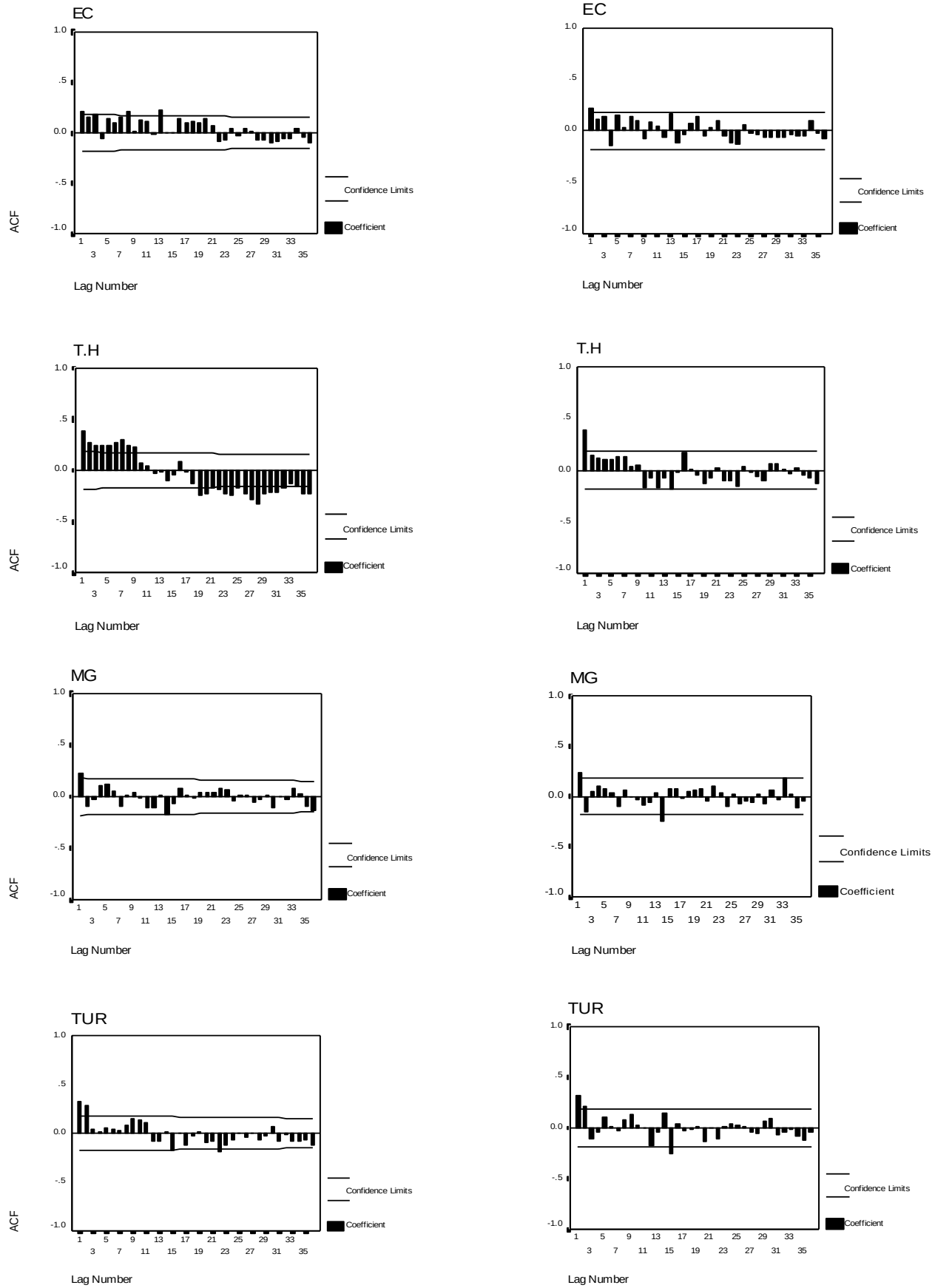


Fig . B.20: ACF&PACF of EC, T.H, Mg , and Tur of Diwaniya station.

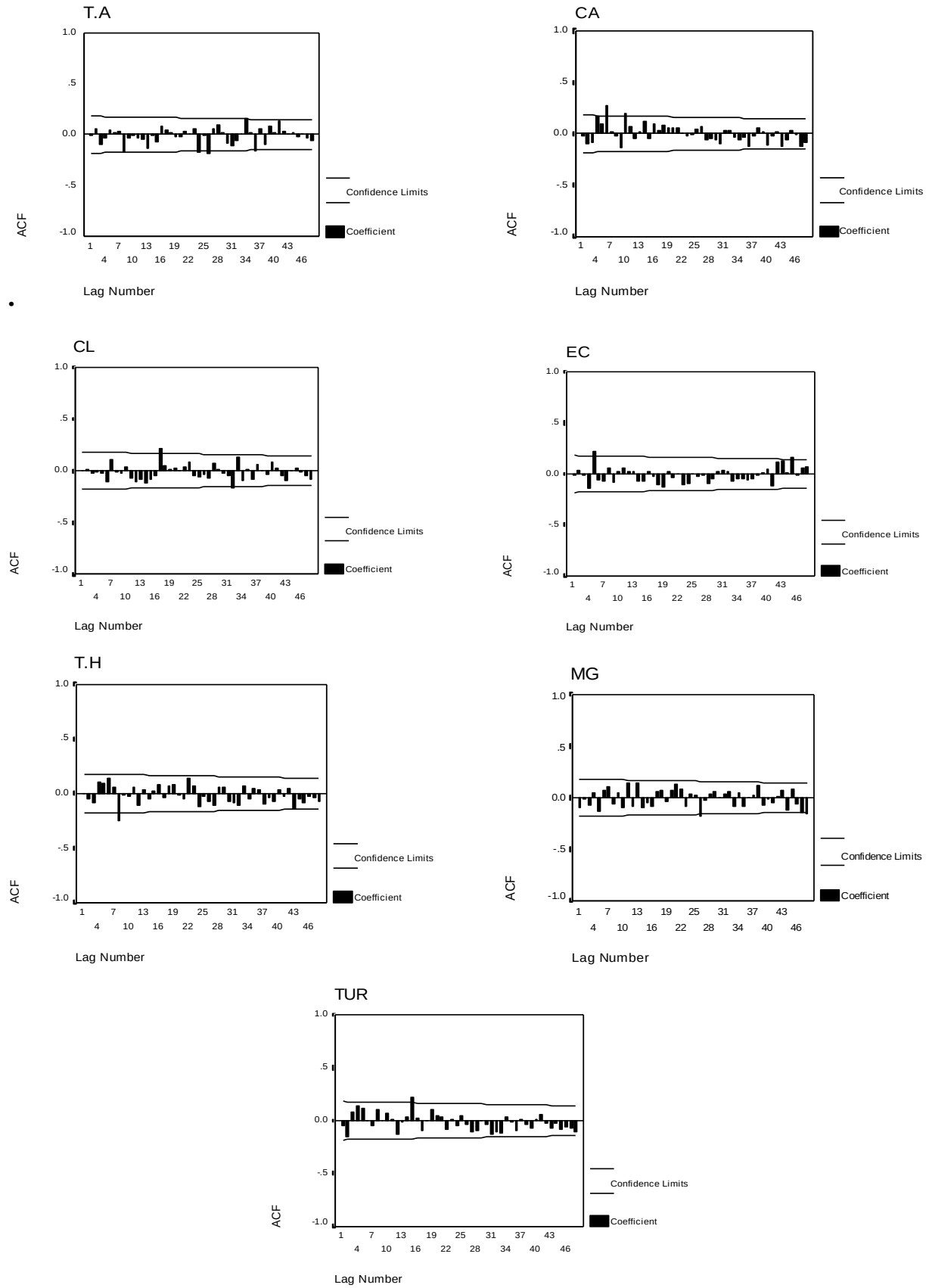


Fig. B.21: Autocorrelogram of residual of Kufa station parameters .

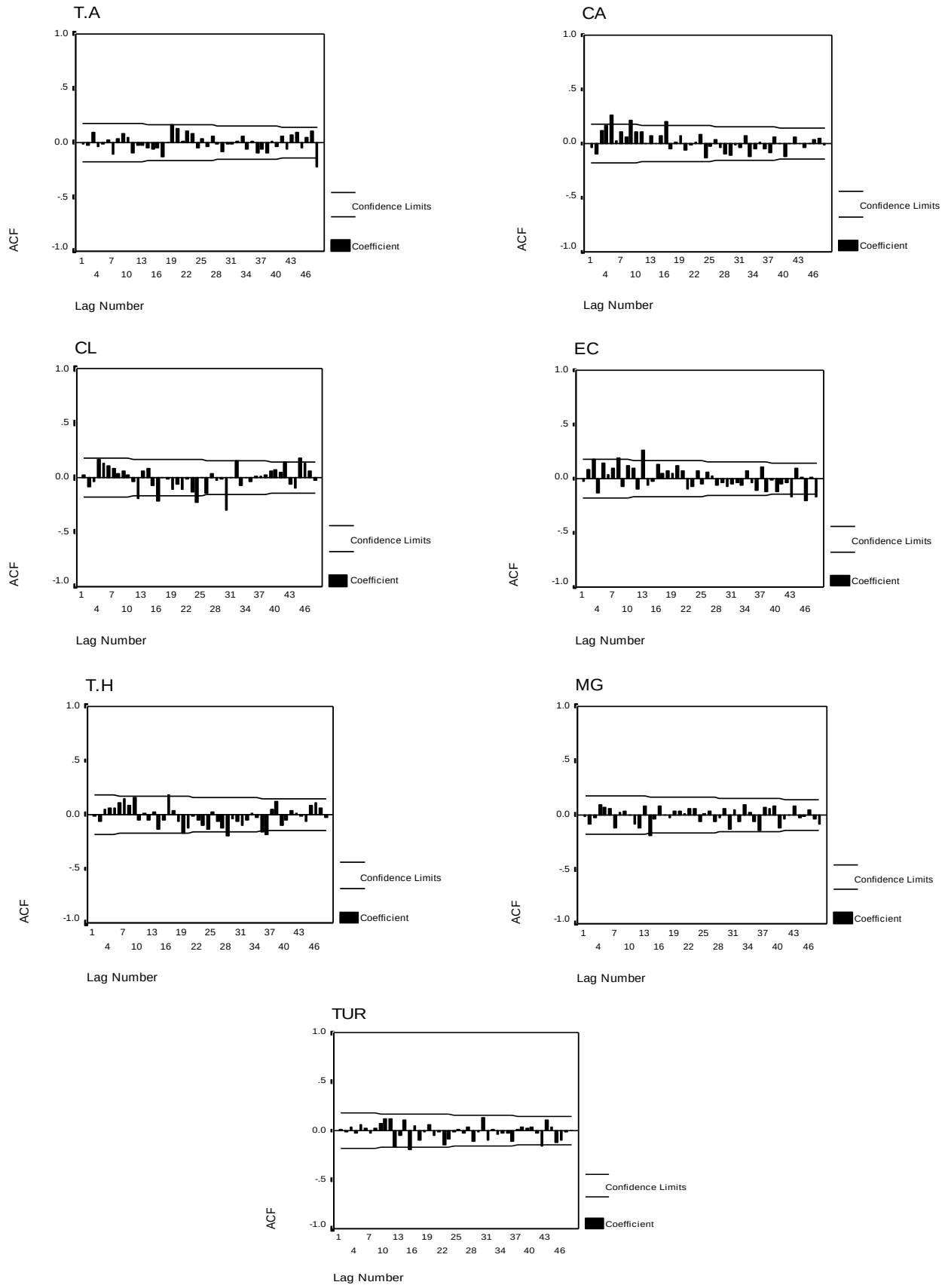


Fig. B.22: Autocorrelogram of residual of Diwaniya station parameters .

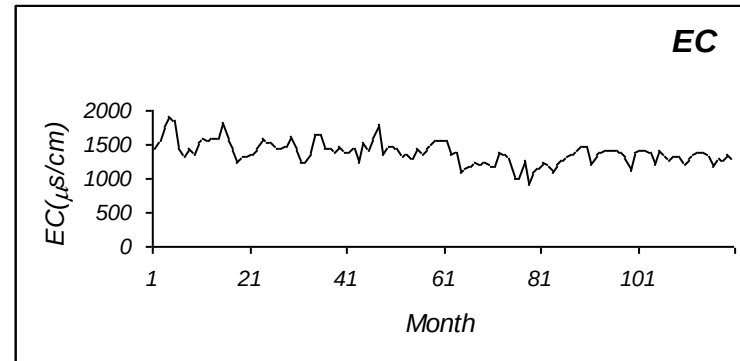
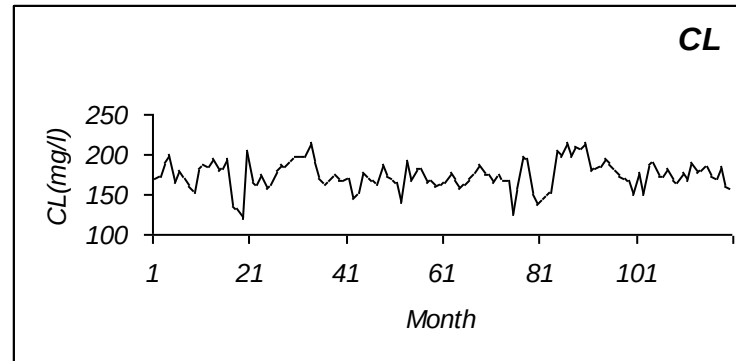
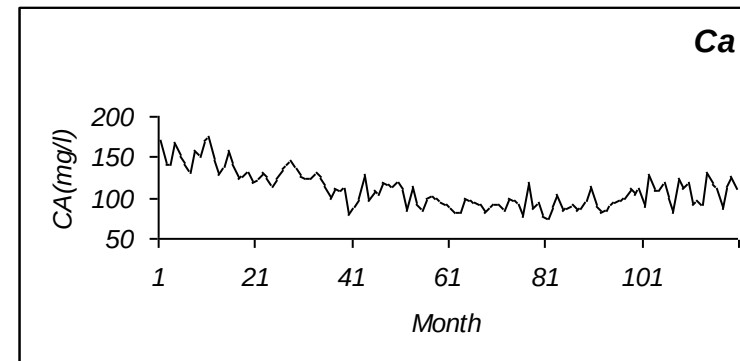
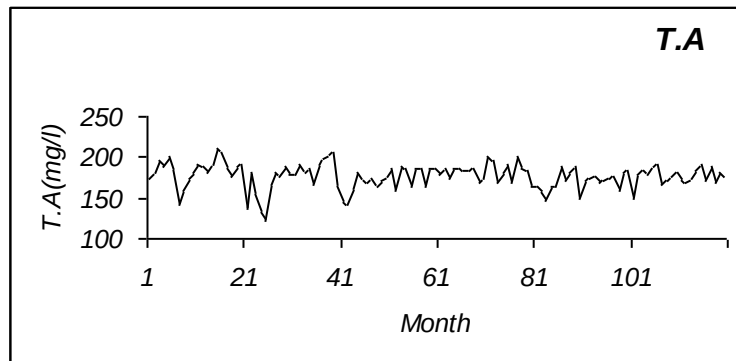


Fig. B.23: Monthly observed values of T.A, Ca ,Cl and EC of Kufa station from 1992 to 2001.

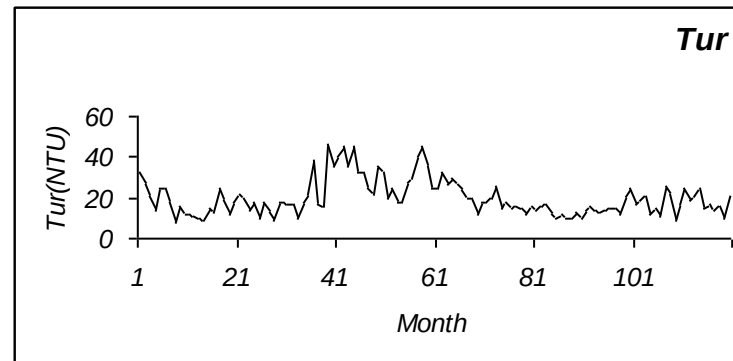
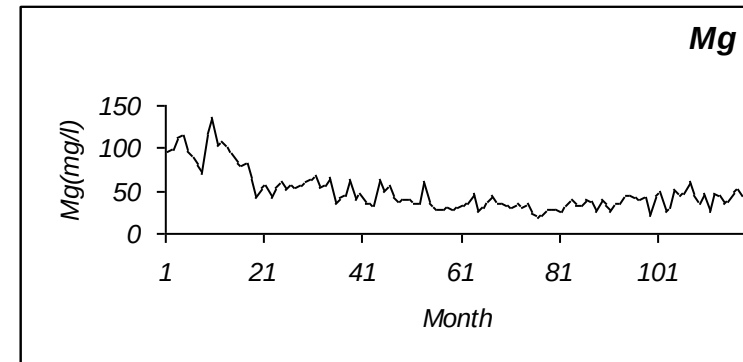
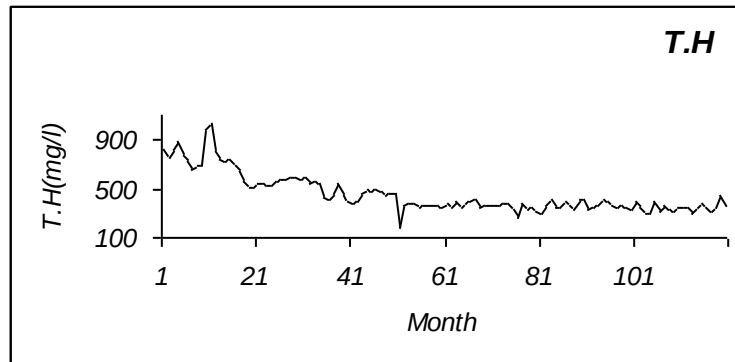


Fig. B.24: Monthly observed values of T.H , Mg and Tur of Kufa station from 1992 to 2001.

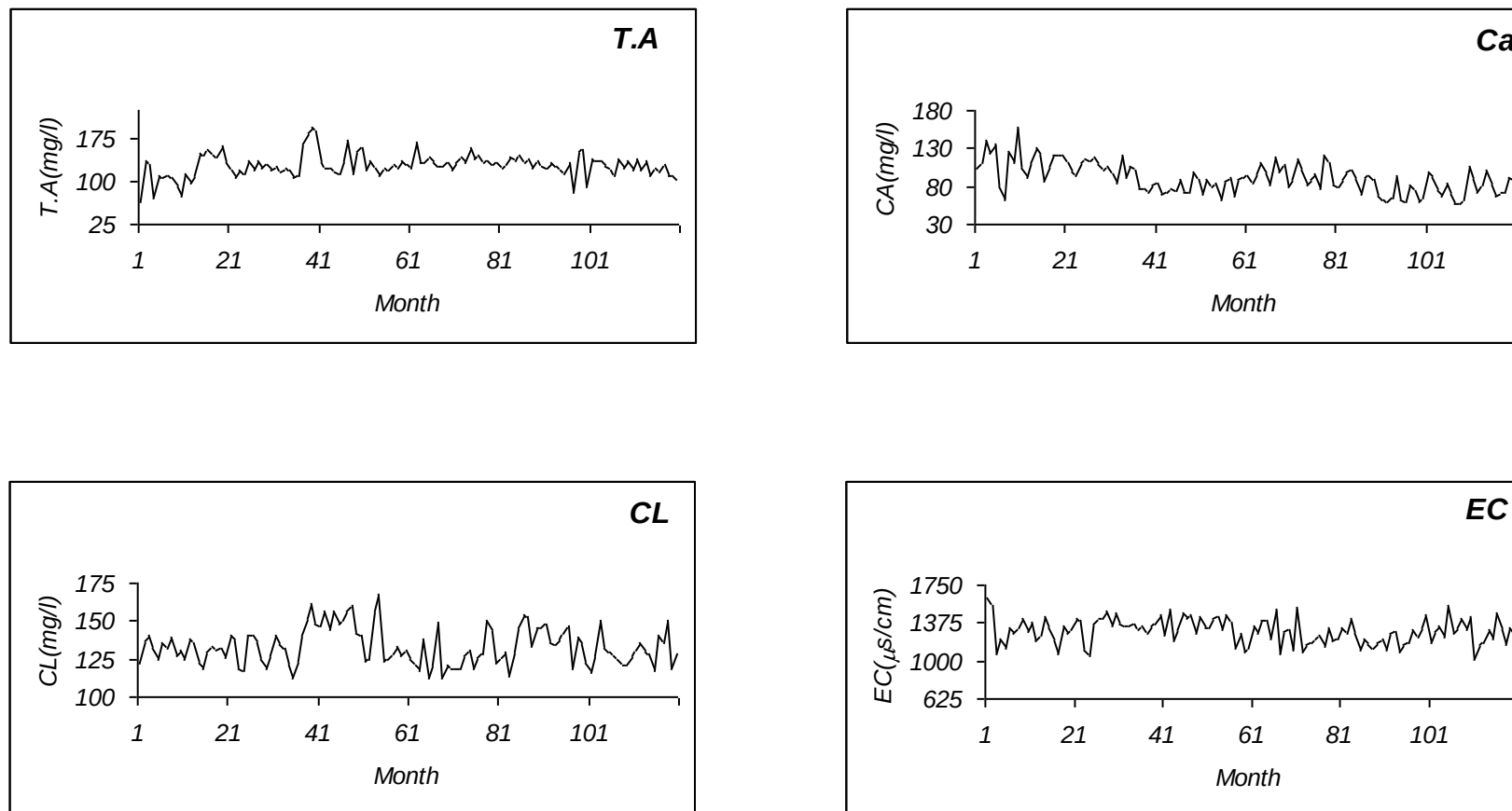


Fig. B.25: Monthly observed values of T.A, Ca ,Cl and EC of Diwaniya station from 1992 to 2001.

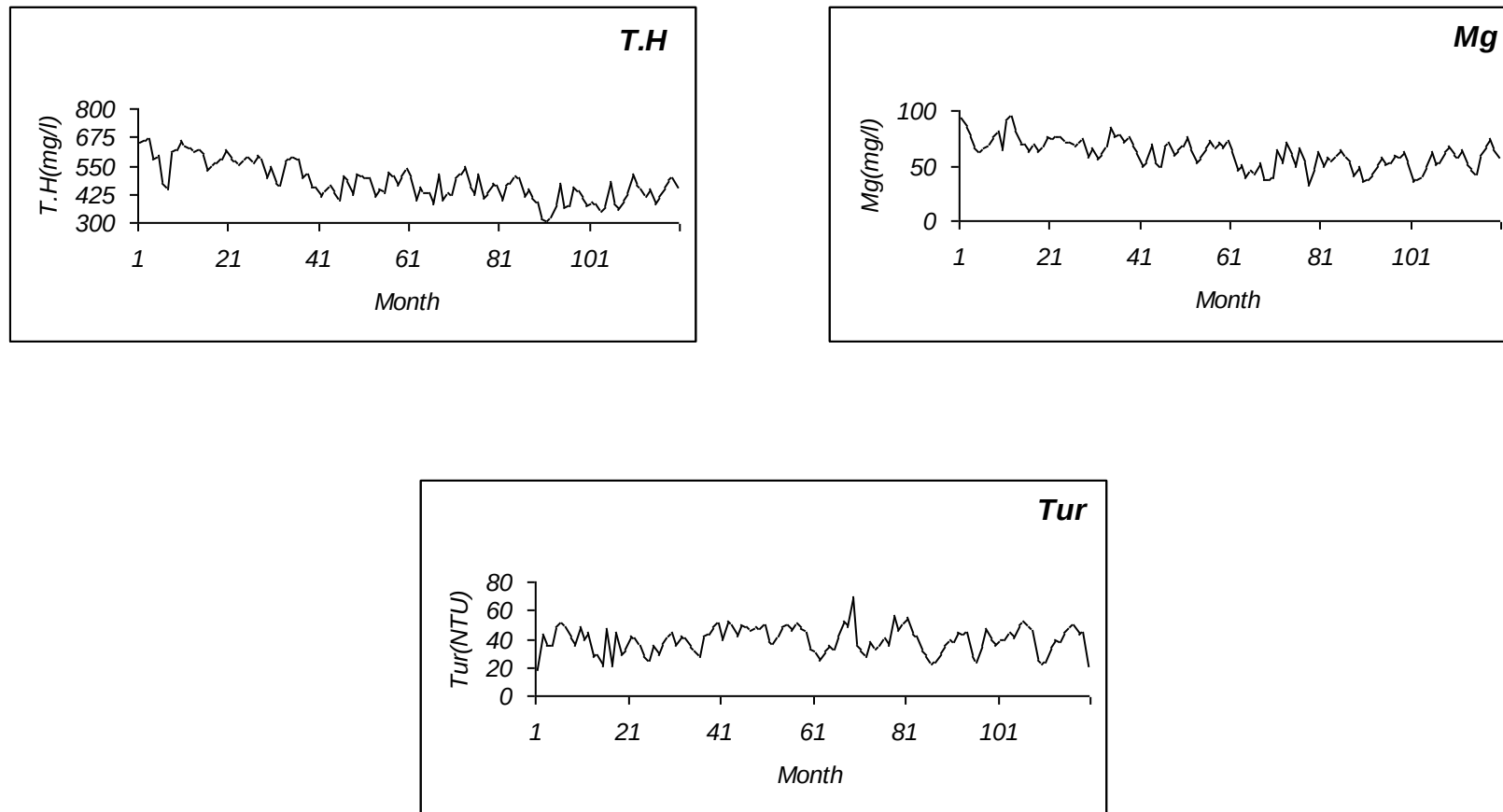


Fig. B.26: Monthly observed values of T.H , Mg and Tur of Diwaniya station from 1992 to 2001.

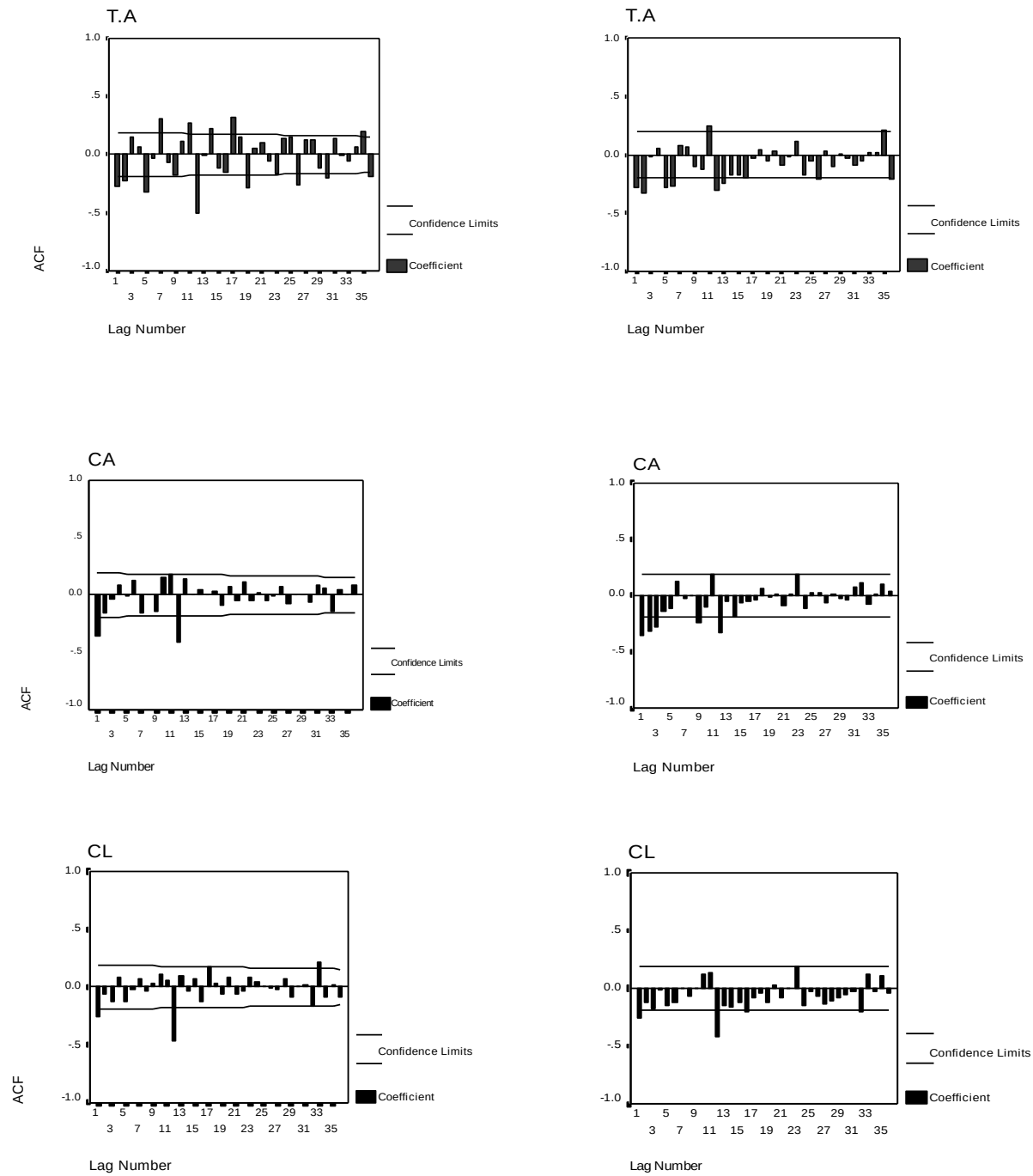


Fig. B.27: ACF&PACF of T.A, Ca and Cl of Kufa station.

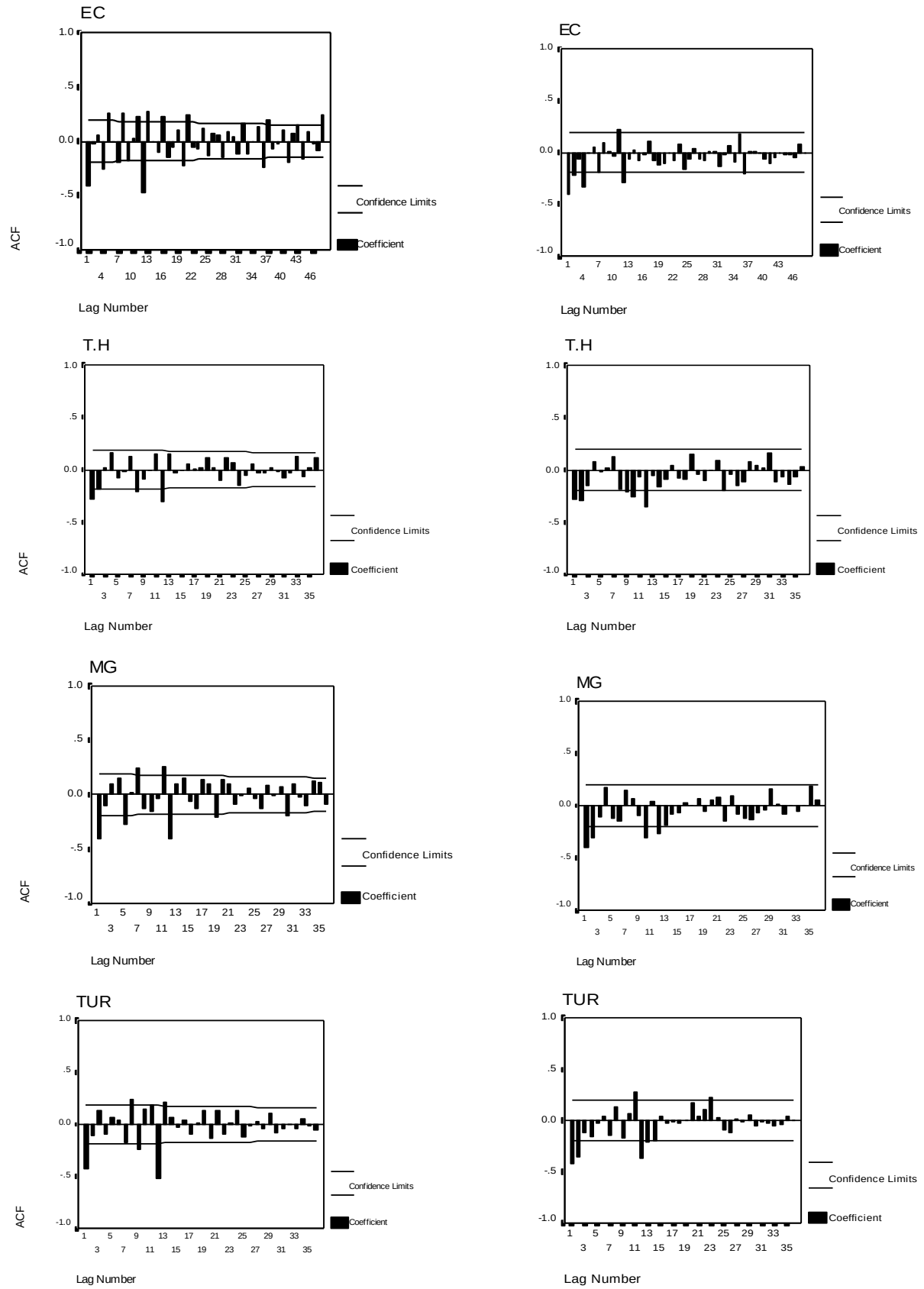


Fig. B.28: ACF&PACF of EC, T.H, Mg and Tur of Kufa station.

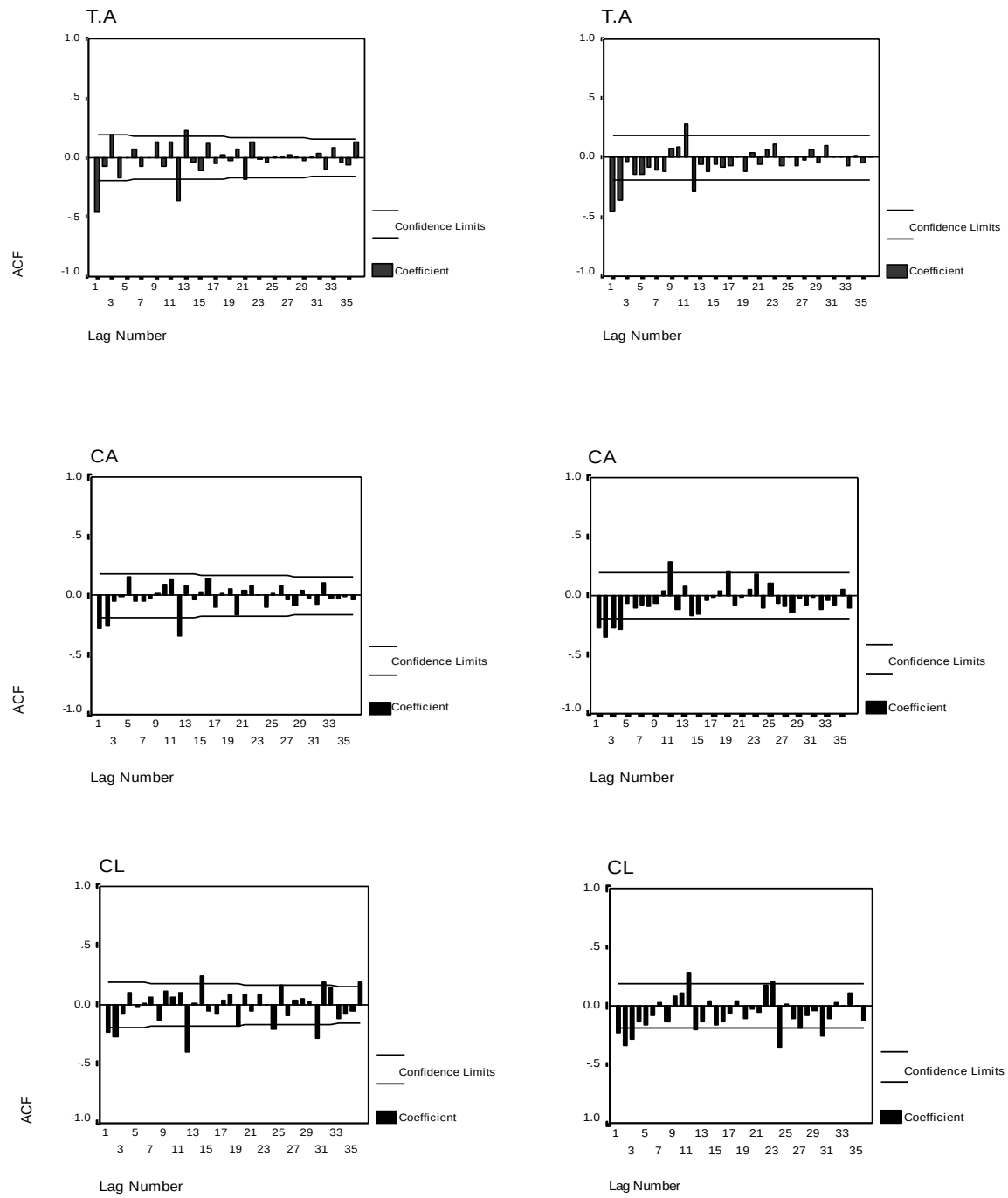


Fig. B.29: ACF&PACF of T.A, Ca and Cl of Diwaniya station.

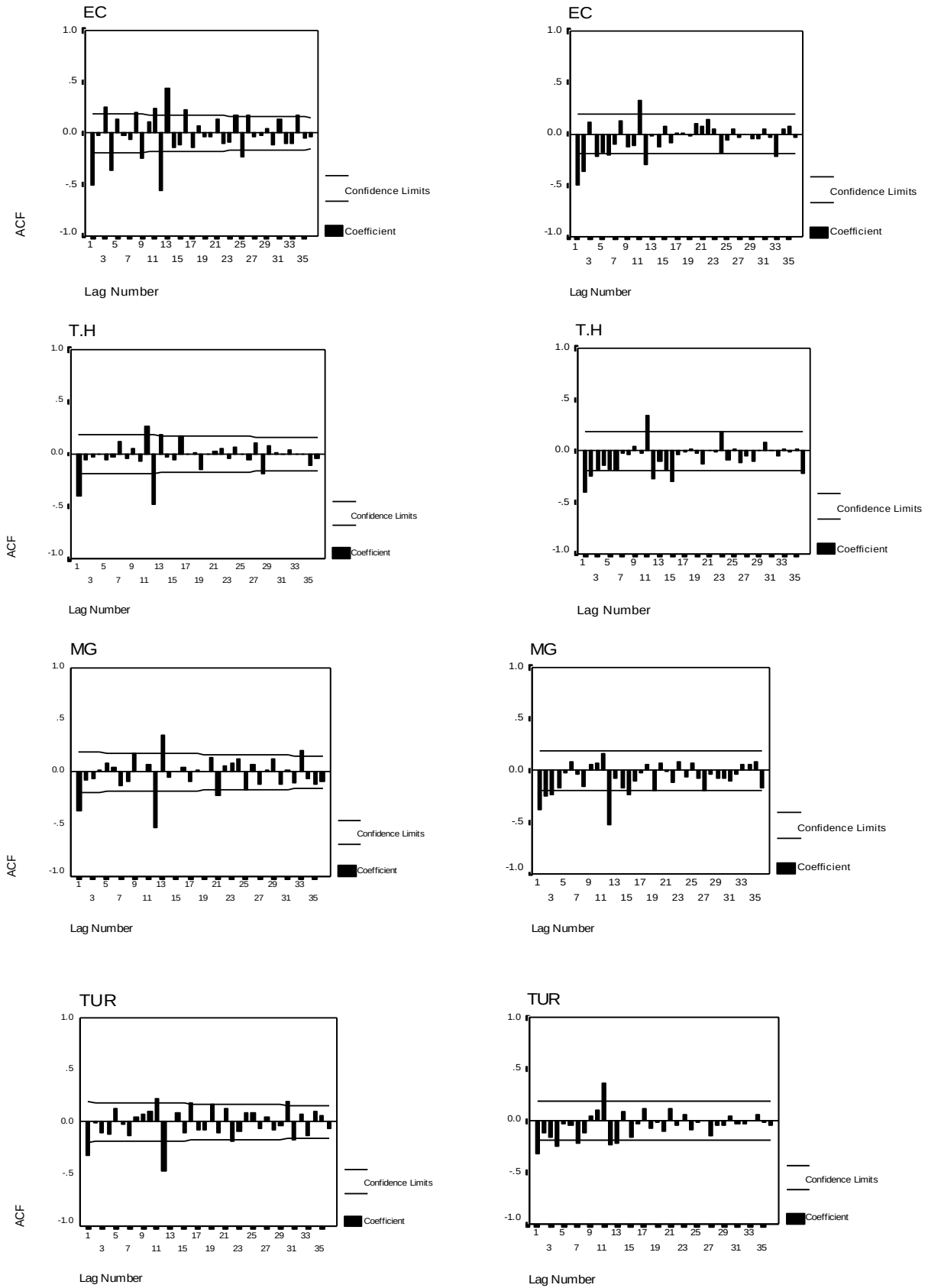
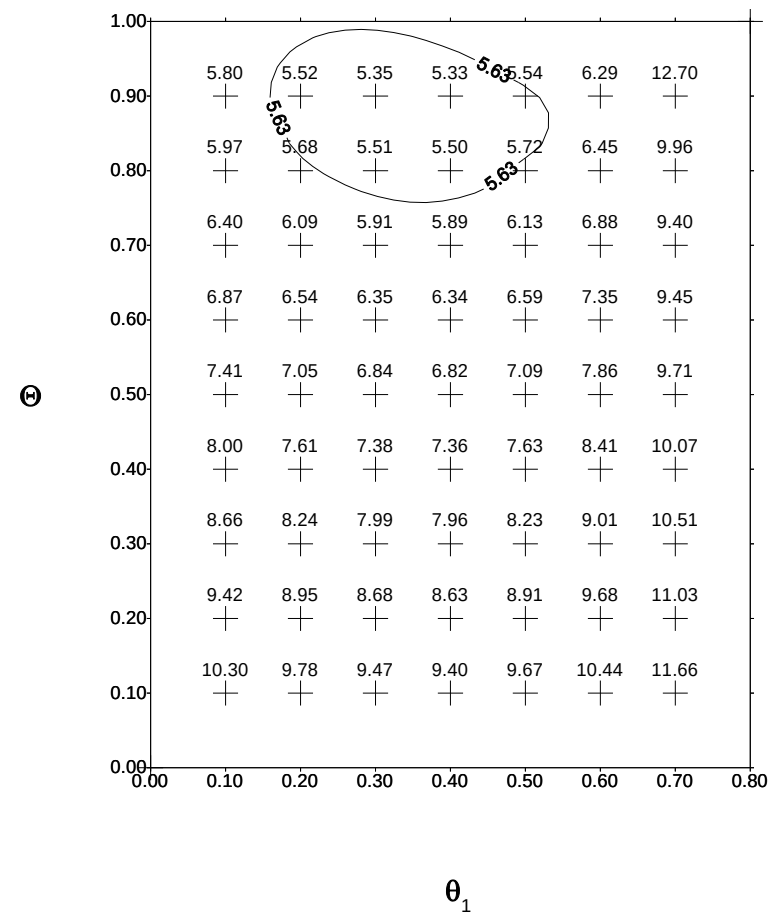
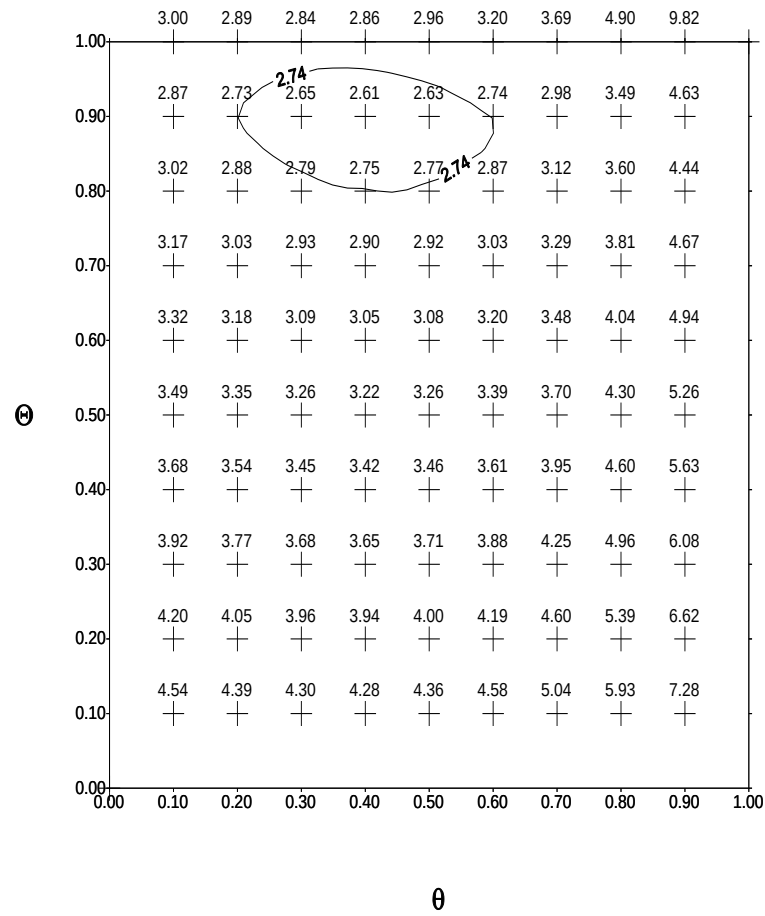
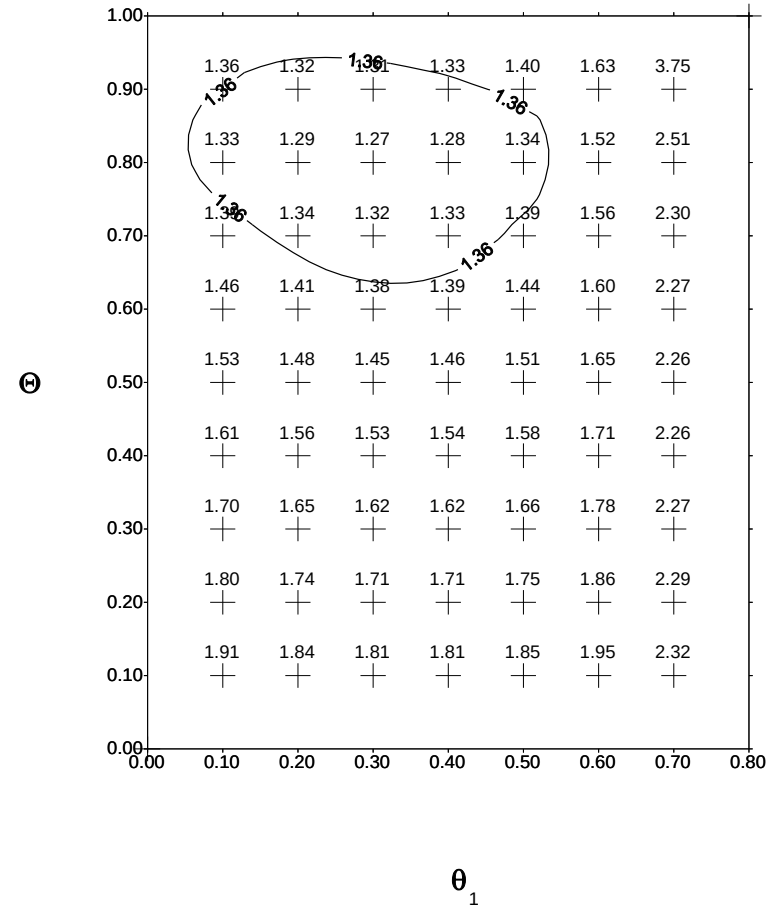
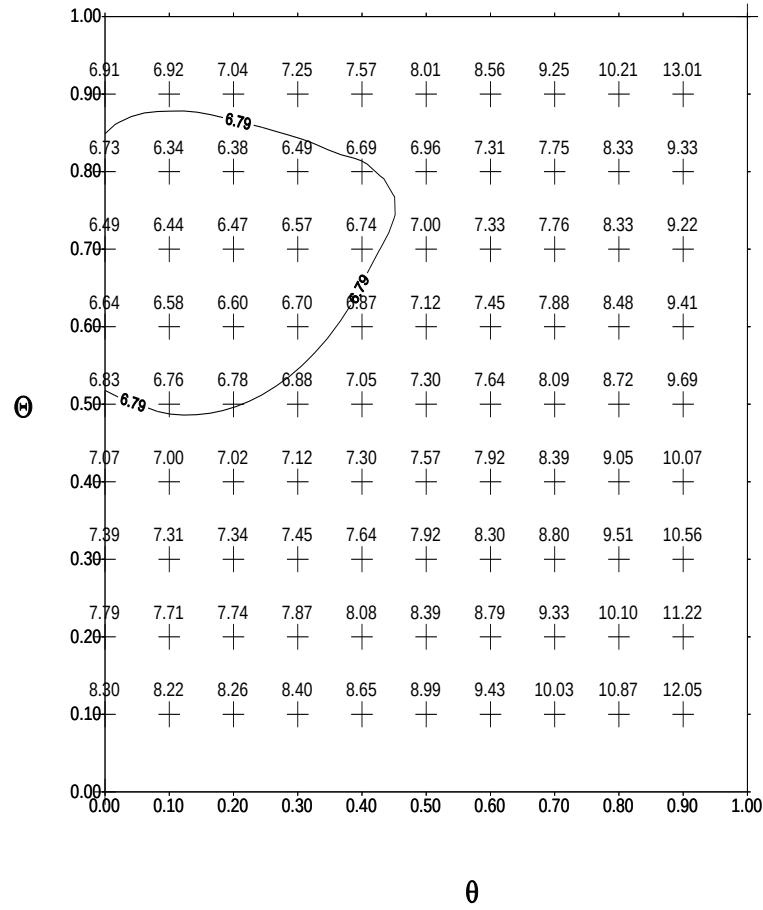


Fig. B.30: ACF&PACF of EC, T.H, Mg and Tur of Diwaniya station.



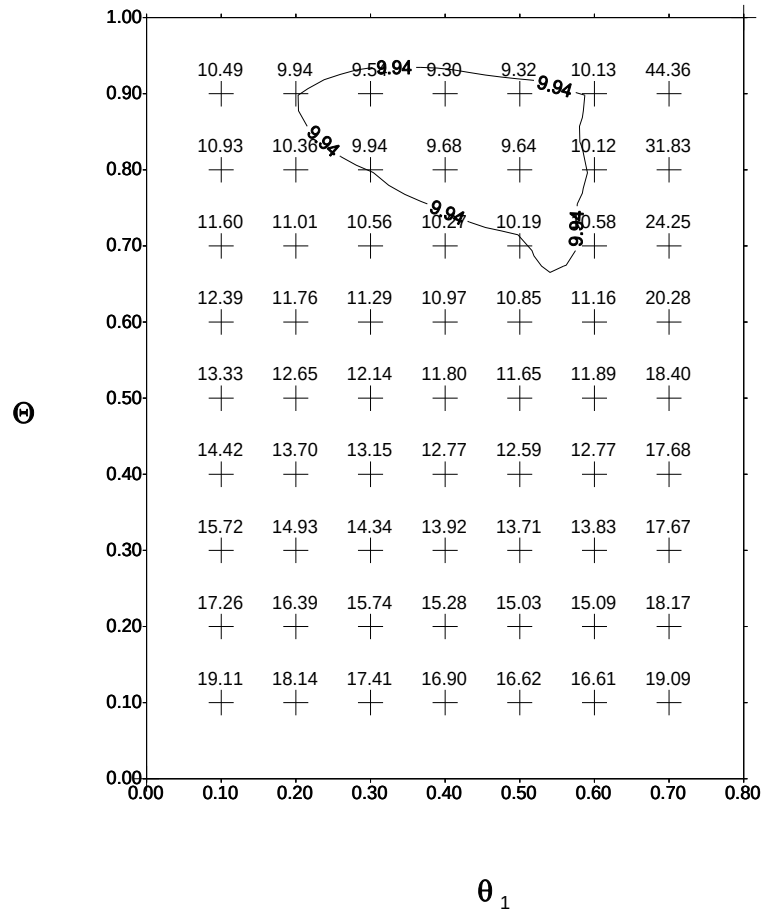
$\theta_2=0.288$

Fig. B.31: Confidence region of T.A of AL-Hindya station. Fig B.32: Confidence region of Ca AL-Hindya station .



$\theta_2=0.261$

Fig.B.33:Confidence region of Cl of AL-Hindya station. Fig B.34: Confidence region of EC of AL-Hindya station.



$\theta_2=0.294$

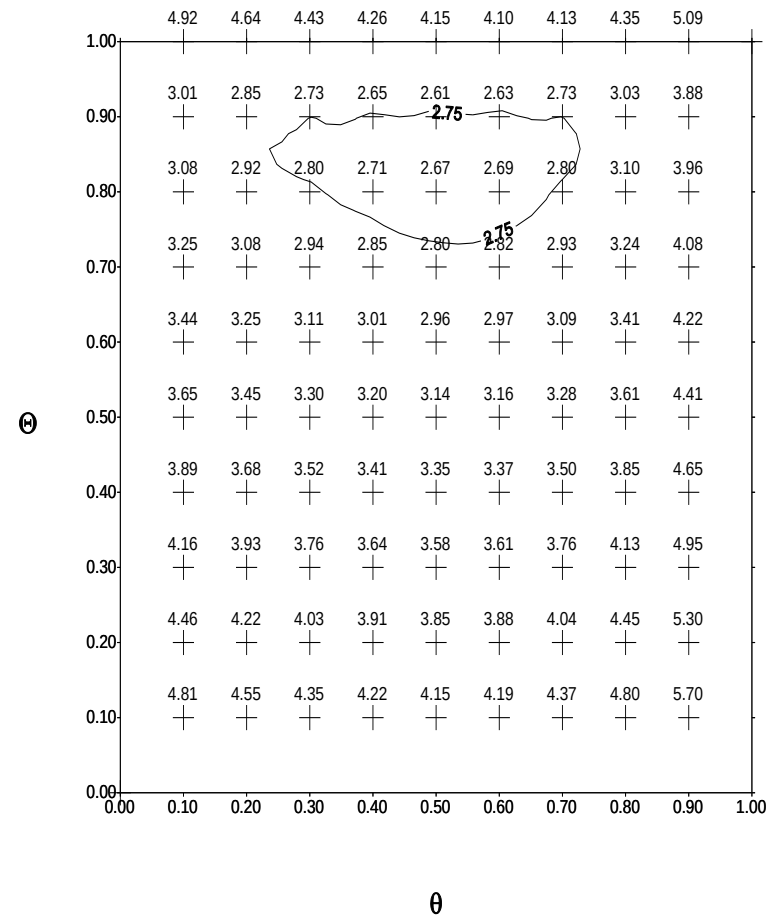


Fig B.35: Confidence region of Mg of AL-Hindya station . Fig B.36: Confidence region of T.H of AL-Hindya station .

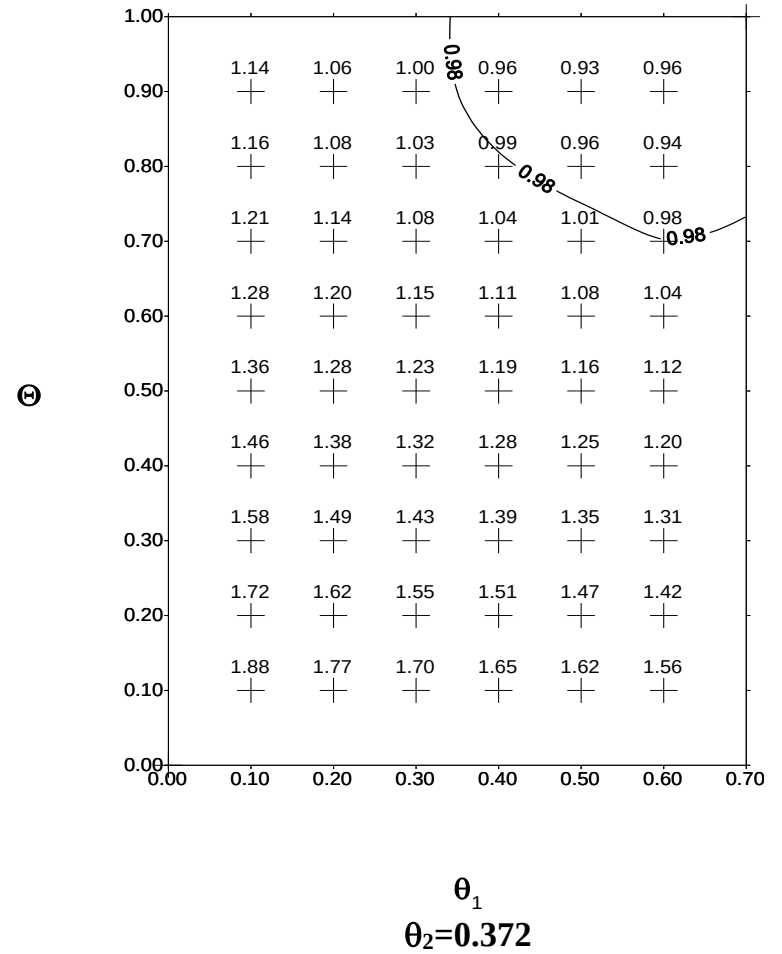
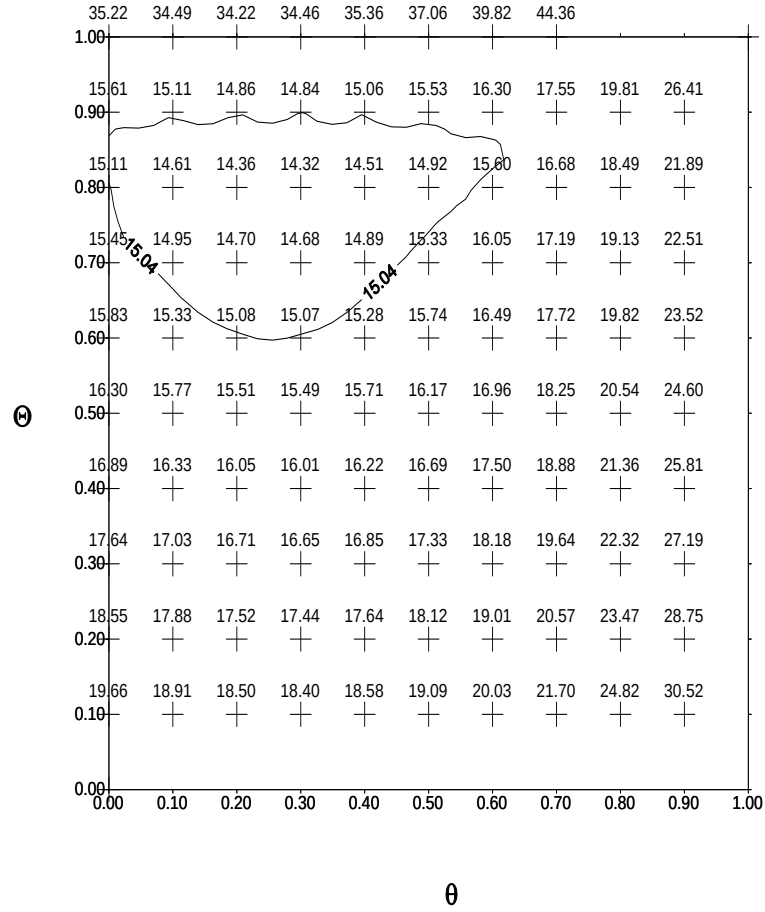


Fig B.37: Confidence region of Tur of AL-Hindya station . Fig B.38: Confidence region of T.A of Kufa station .

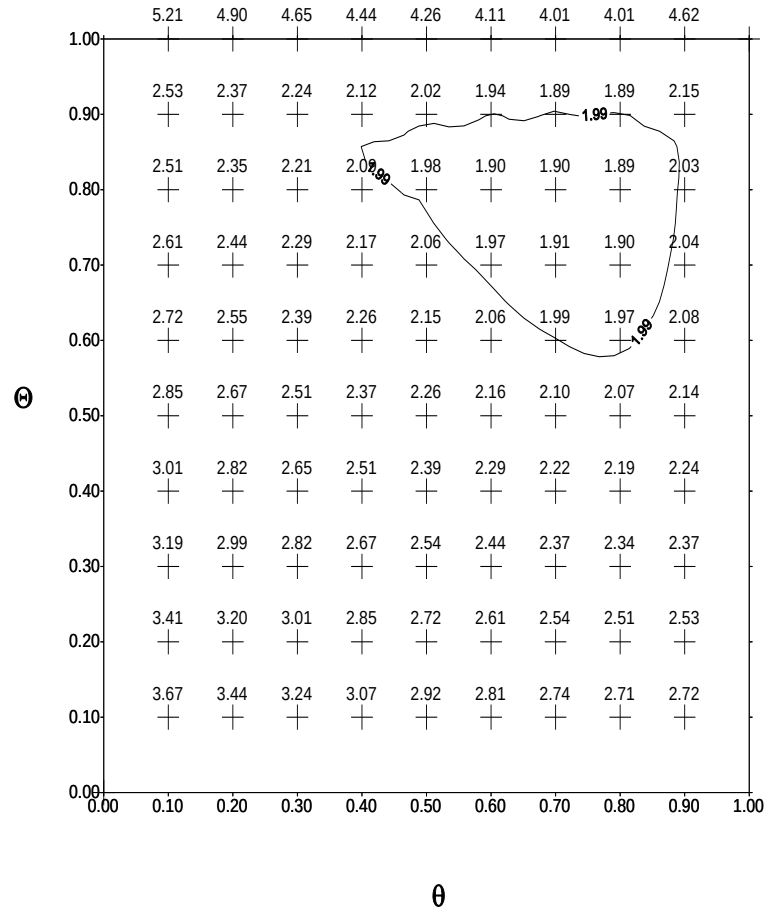


Fig B.39: Confidence region of Ca of Kufa station .

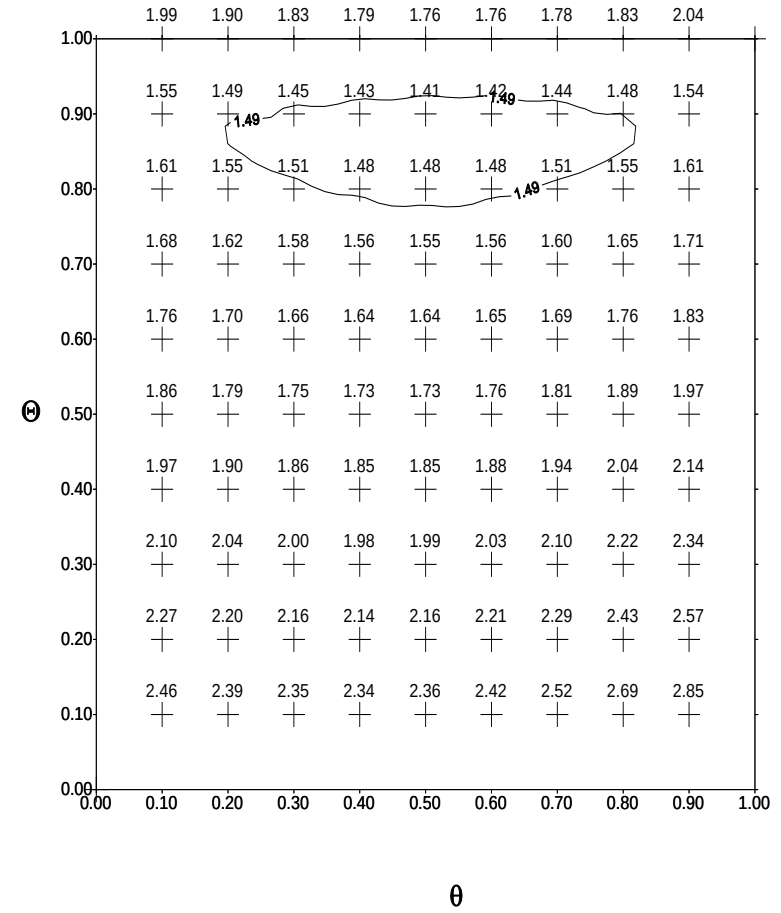


Fig B.40: Confidence region of Cl of Kufa station.

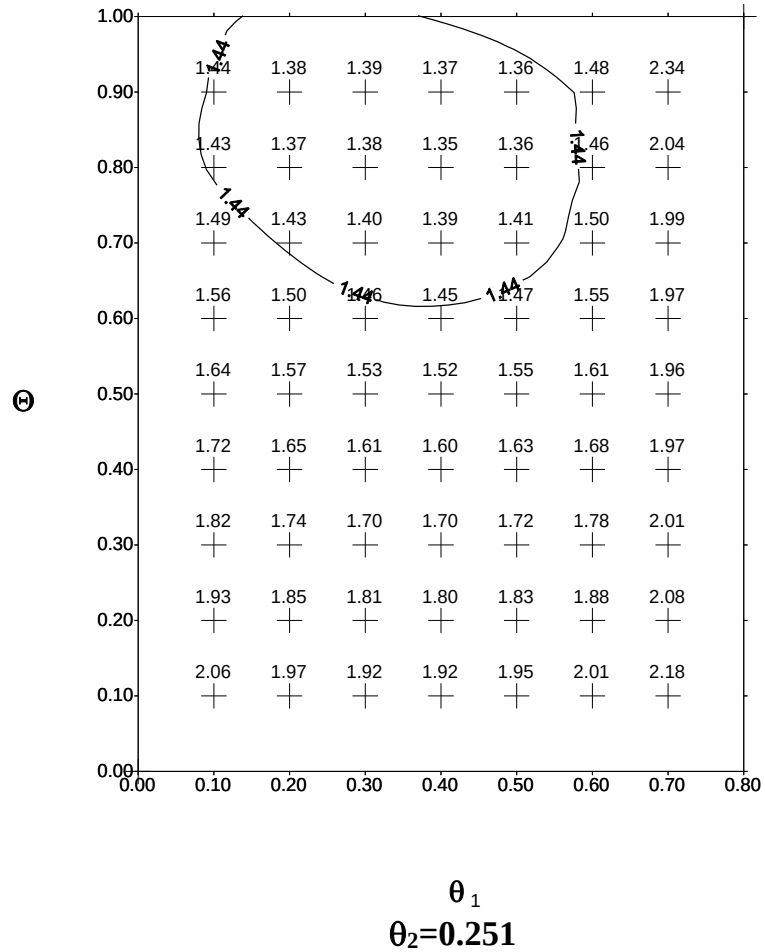


Fig B.41: Confidence region of T.H of Kufa station.

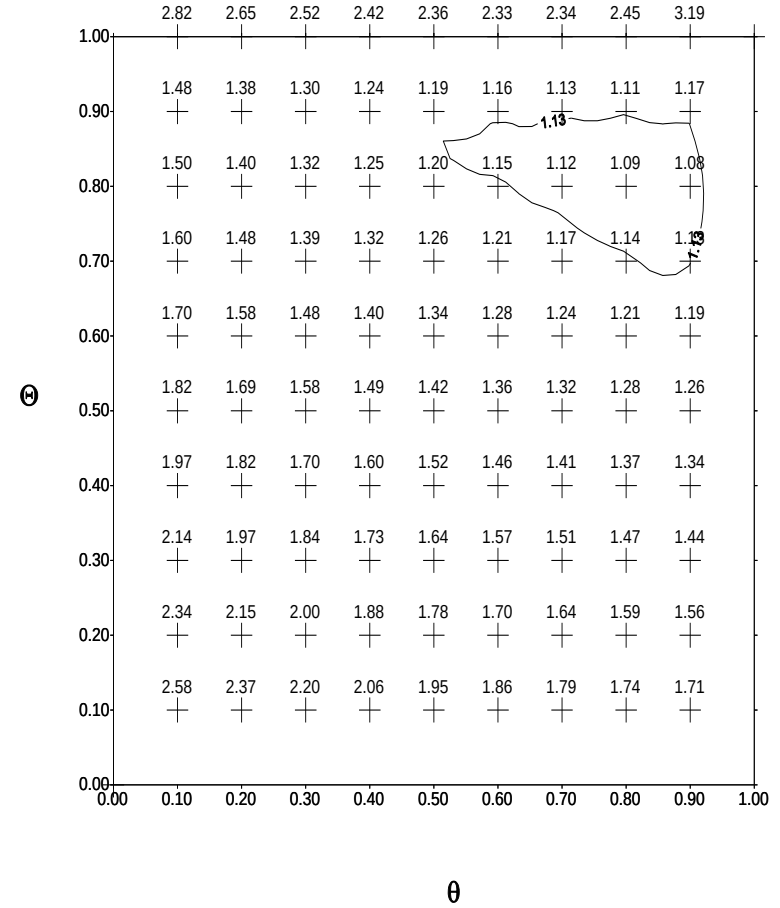


Fig B.42: Confidence region of Mg of Kufa station.

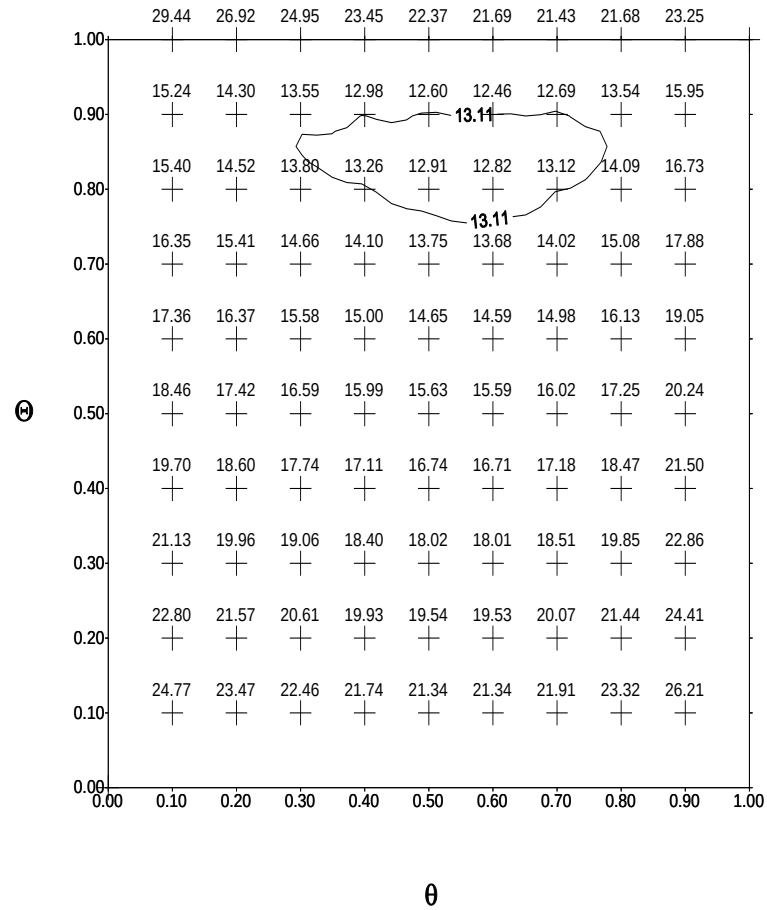


Fig B.43: Confidence region of Tur of Kufa station.

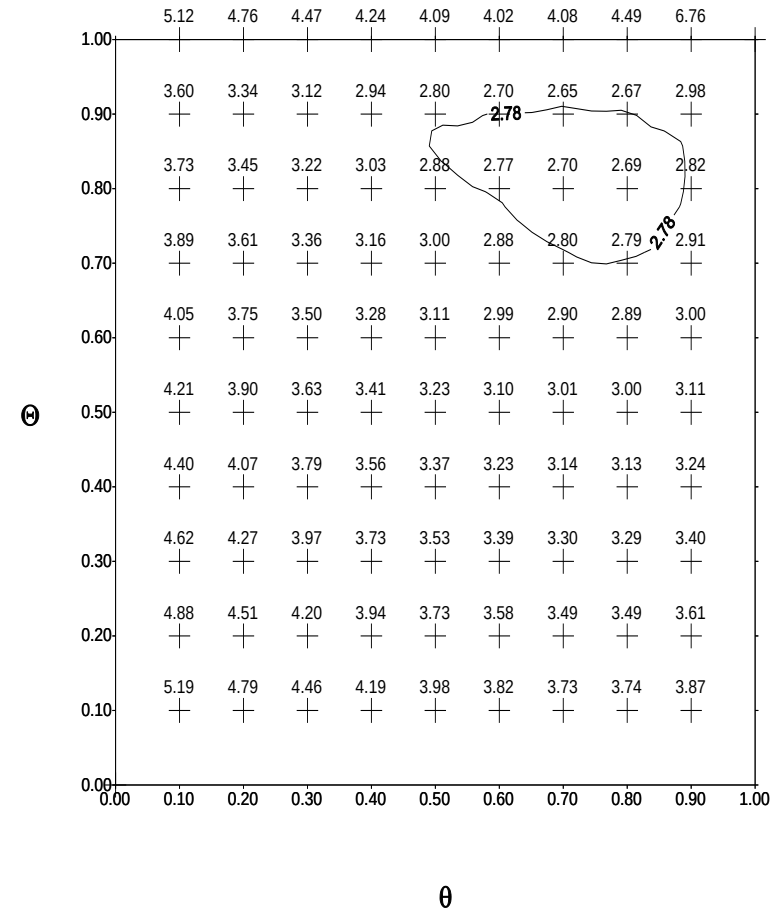


Fig B.44: Confidence region of T.A of Diwaniya station.

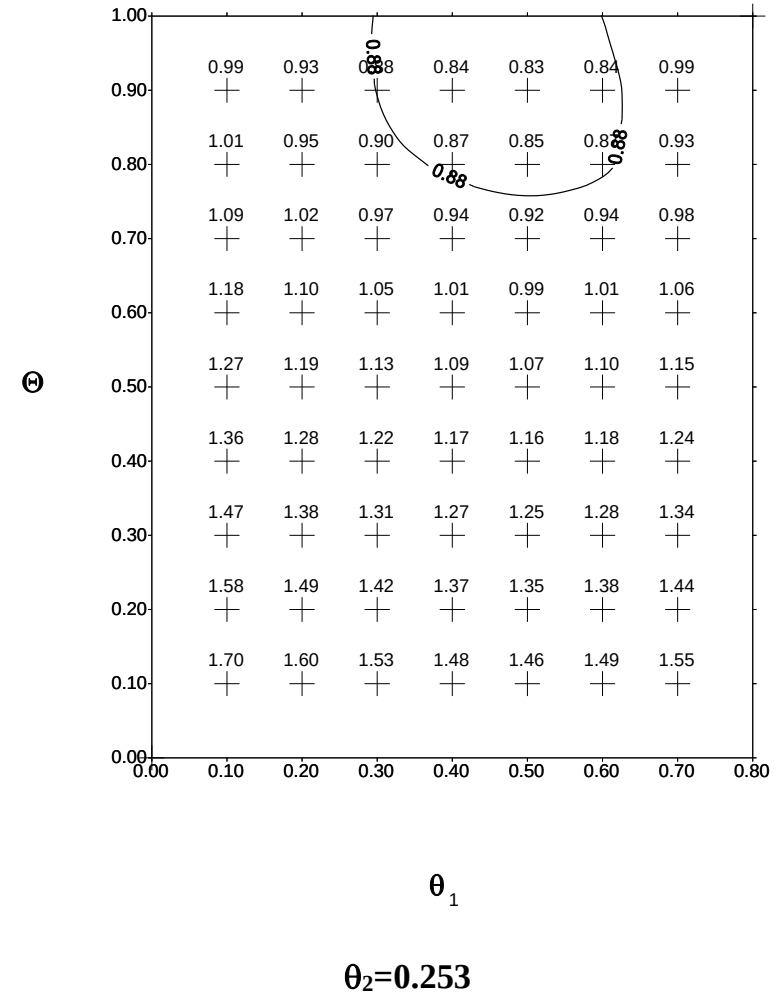
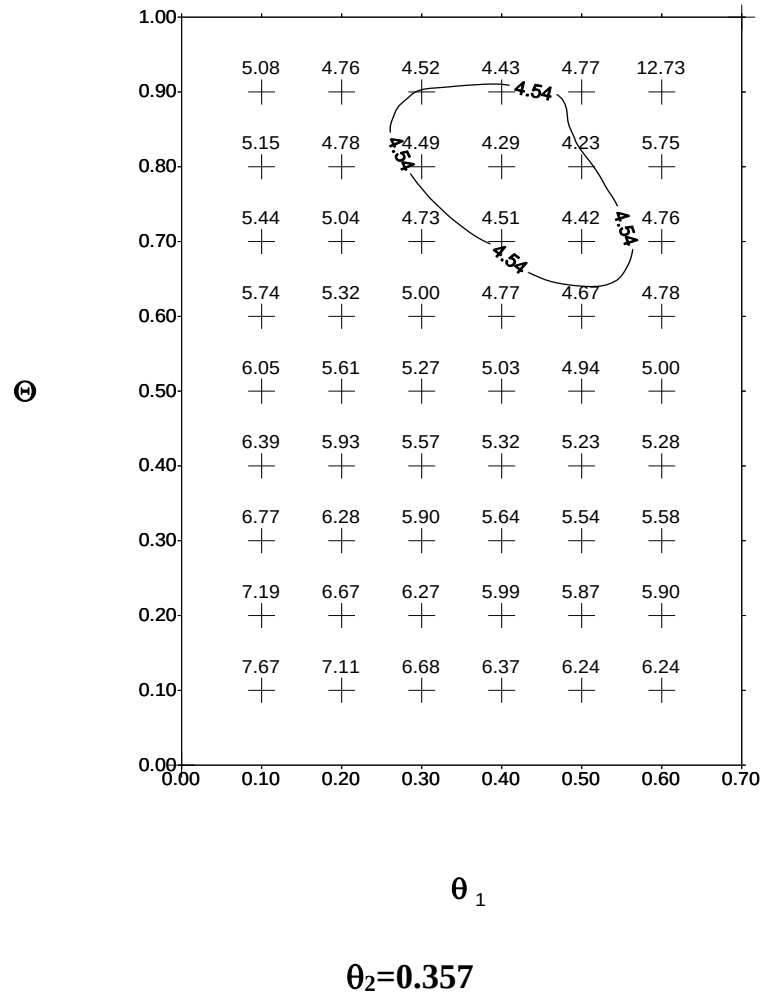


Fig B.45: Confidence region of Ca of Diwaniya station.

Fig B.46: Confidence region of Cl of Diwaniya station .

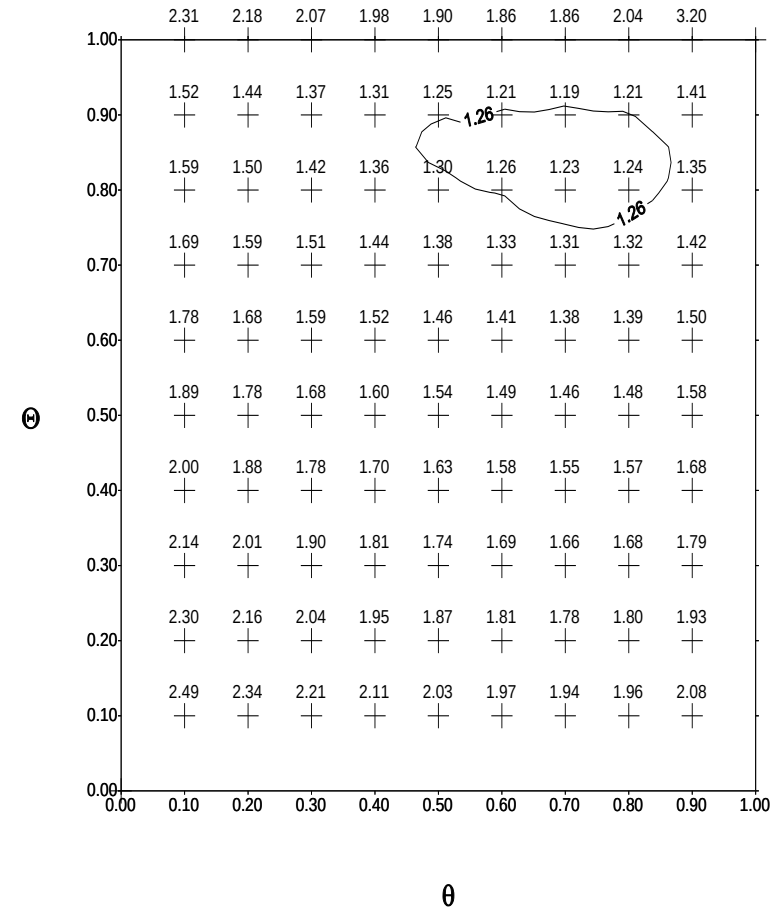
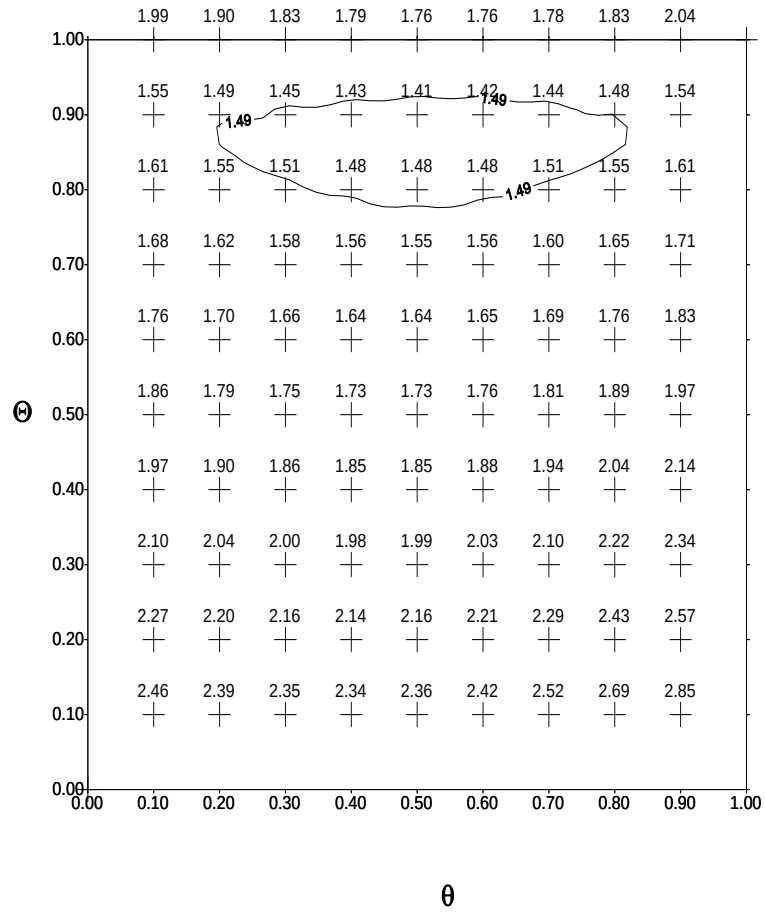


Fig B.47: Confidence region of ECof Diwaniya station . Fig B.48: Confidence region of T.H of Diwaniya station.

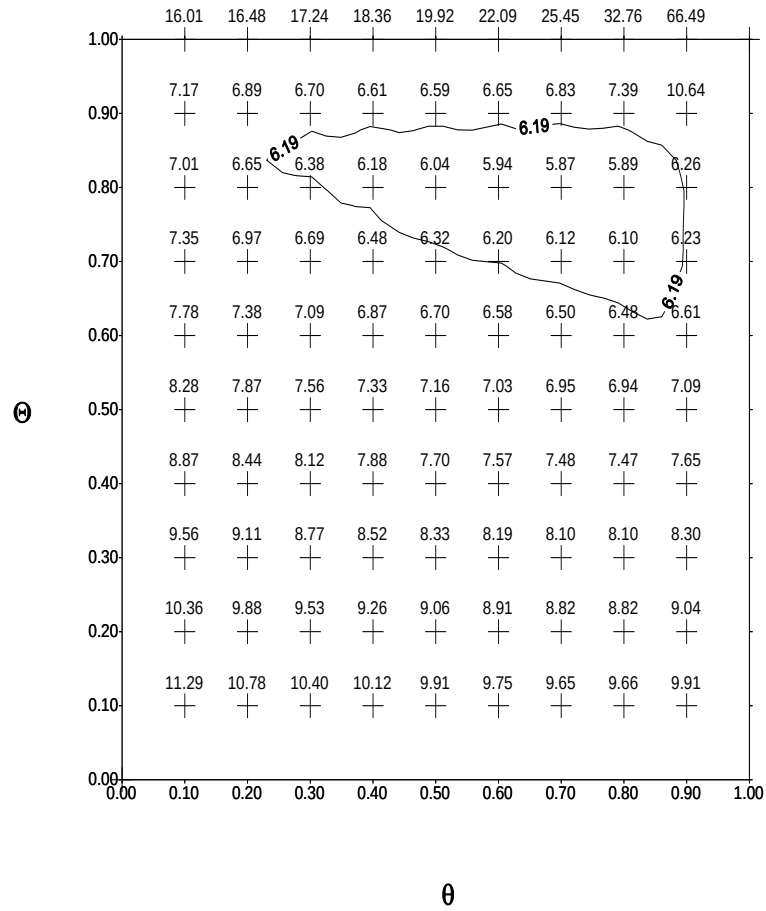


Fig B.49: Confidence region of Mg of Diwaniya station

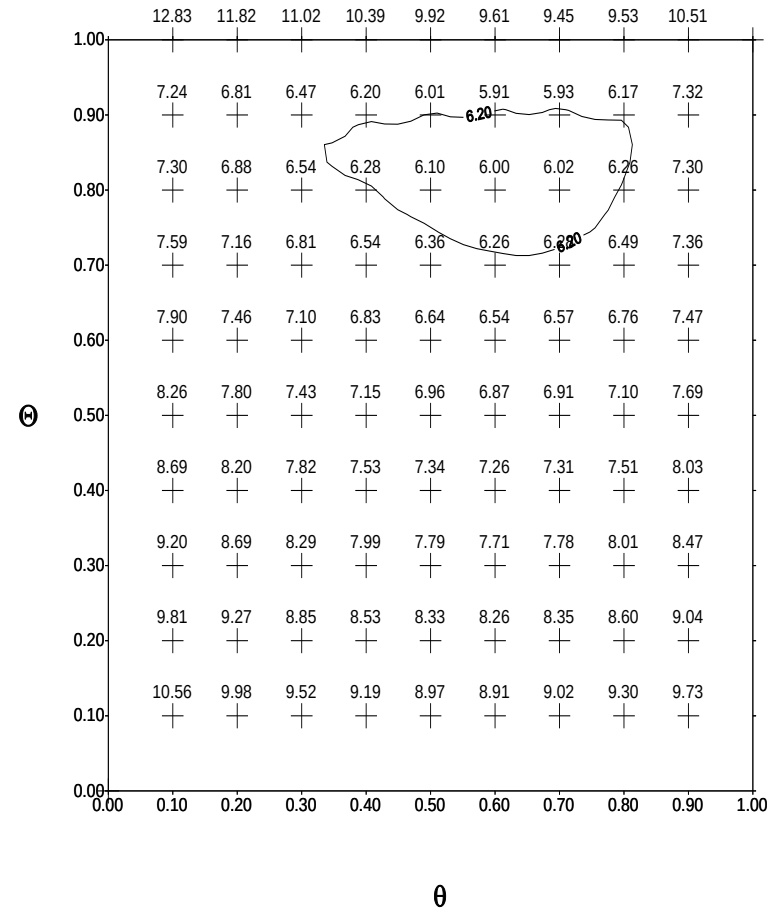


Fig B.50: Confidence region of Tur of Diwaniya station.

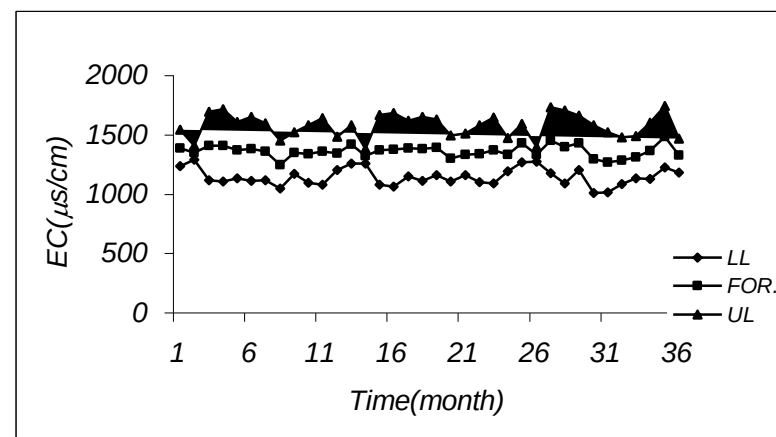
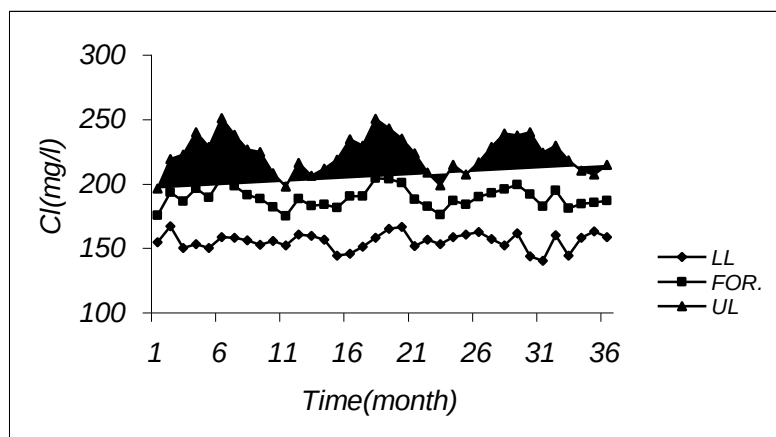
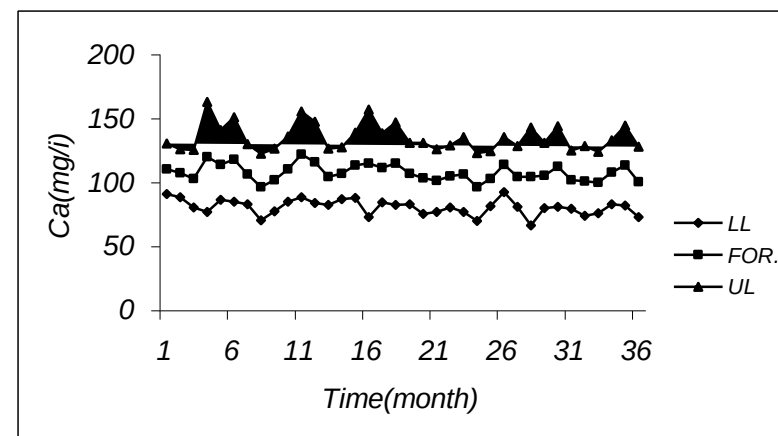
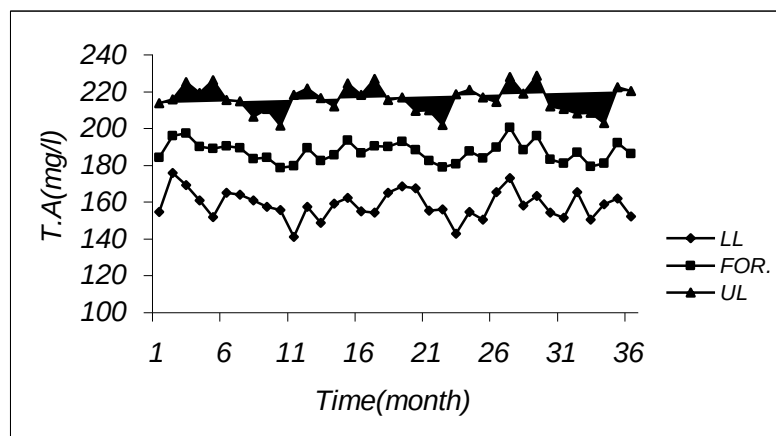


Fig. B.51: Probability limits of forecasted values obtained by ARMA model of T.A ,Ca ,Cl &EC of Kufa station from 2002 to 2004 .

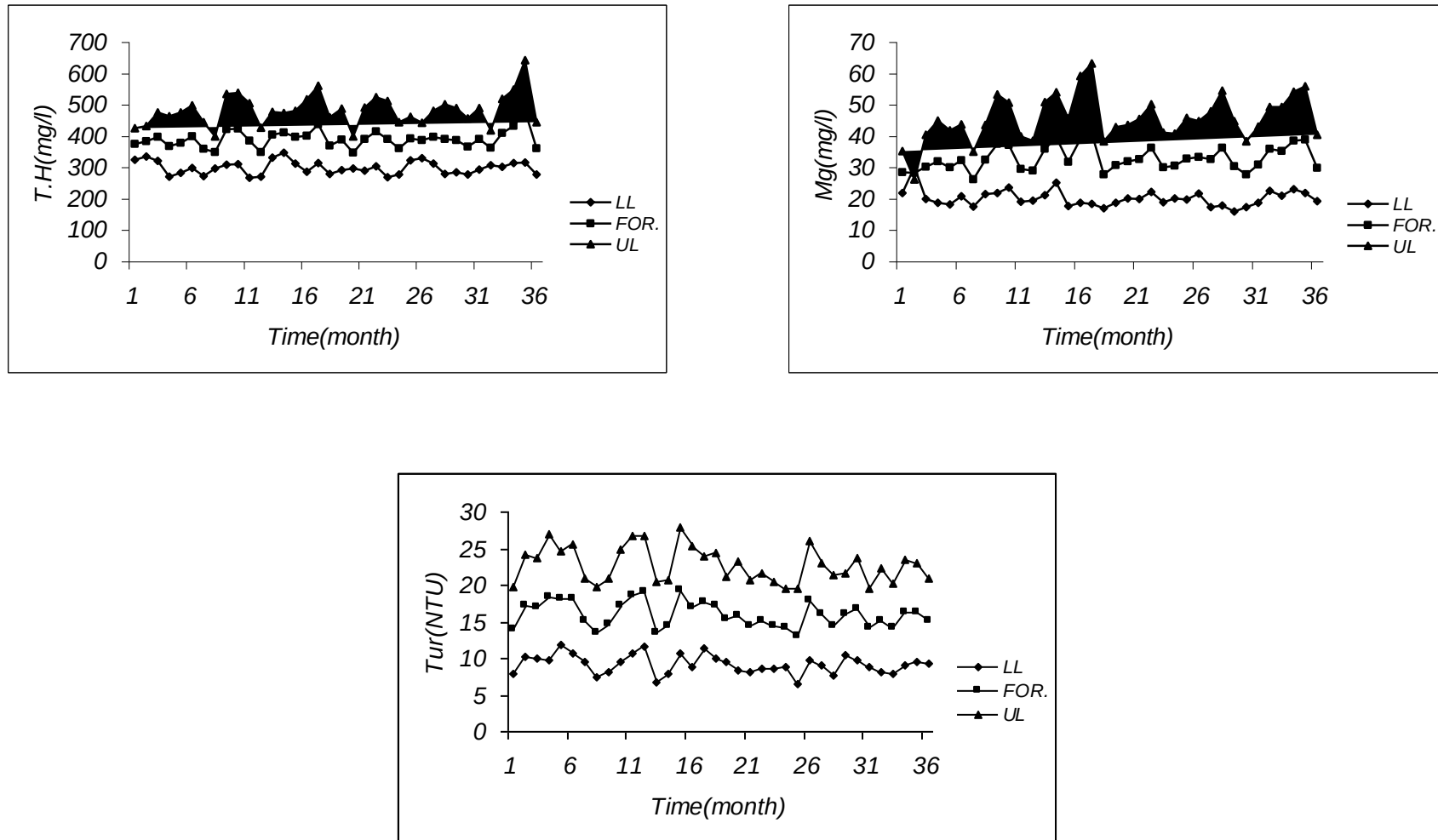


Fig. B.52: Probability limits of forecasted values obtained by ARMA model of T.H, Mg, Tur of Kufa station from 2002 to 2004.

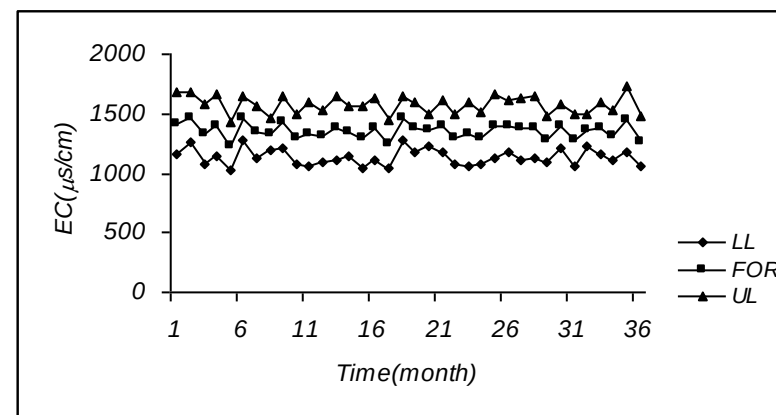
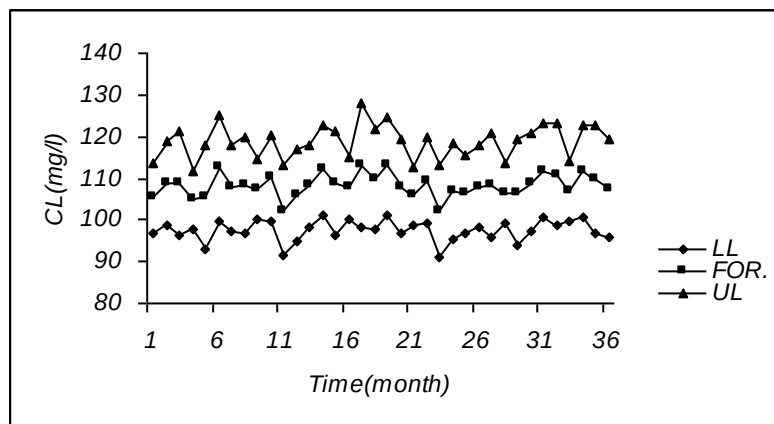
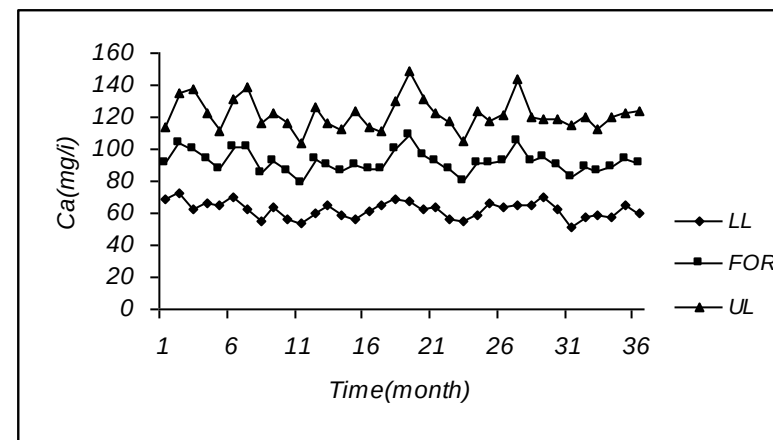
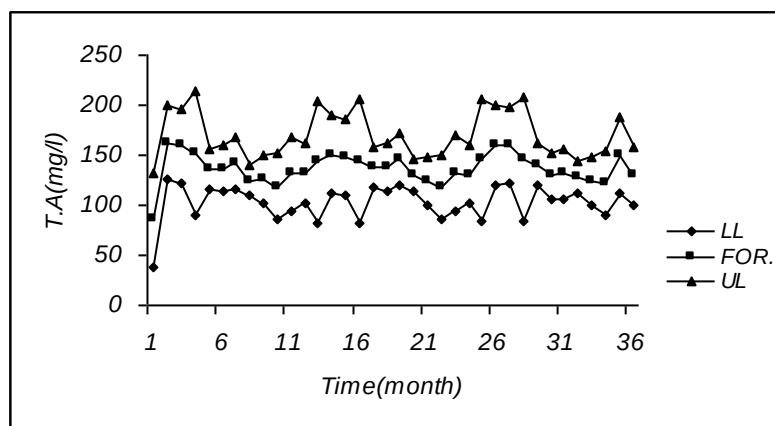


Fig. B.53: Probability limits of forecasted values obtained by ARMA model of T.A ,Ca ,Cl &EC of Diwaniya station from 2002 to 2004.

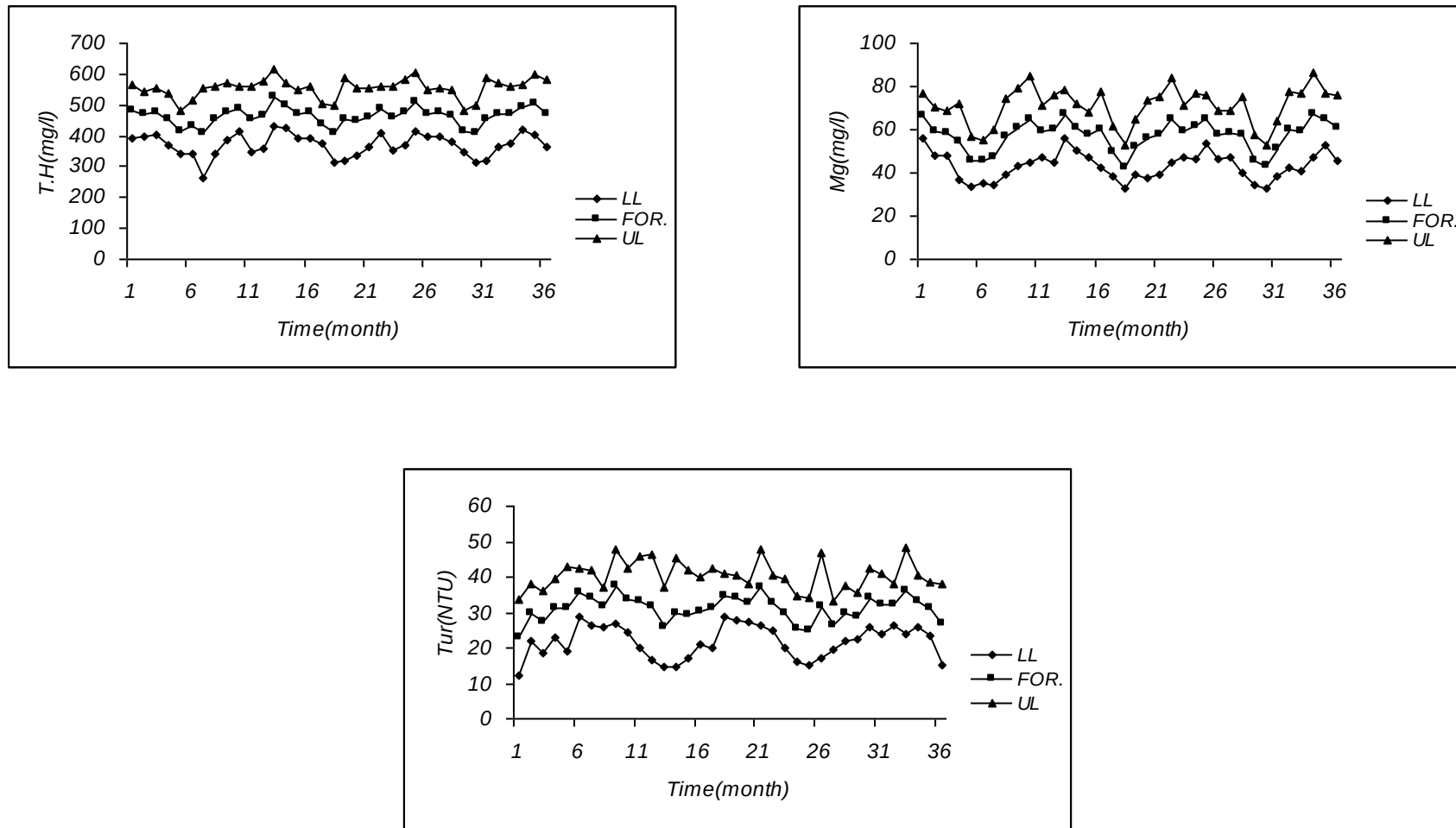


Fig. B.54: Probability limits of forecasted values obtained by ARMA model of T.H, Mg, Tur of Diwaniya station from 2002 to 2004.

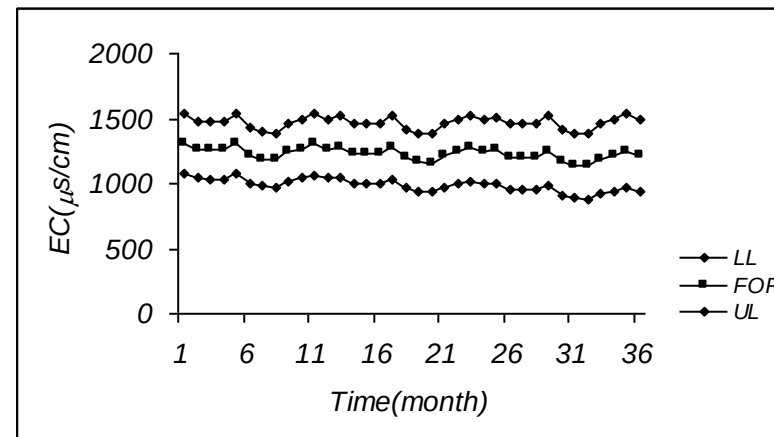
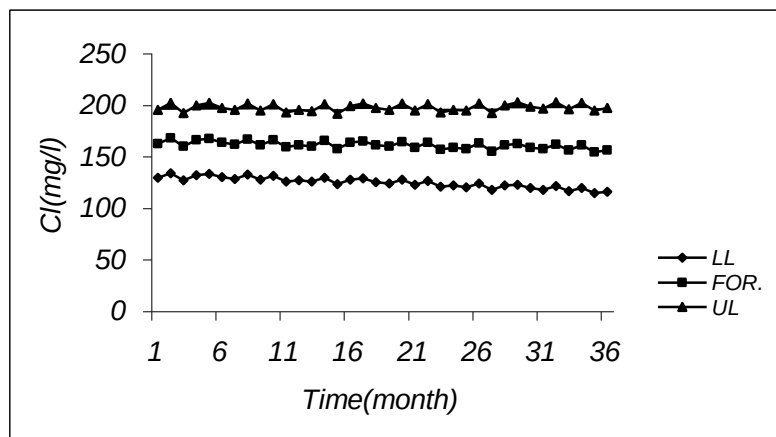
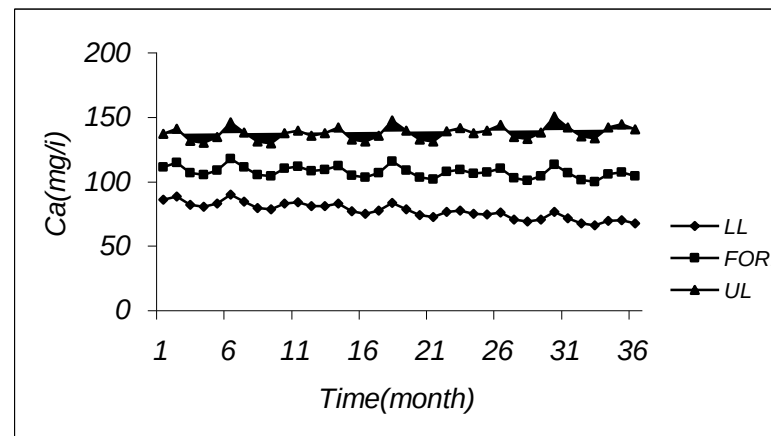
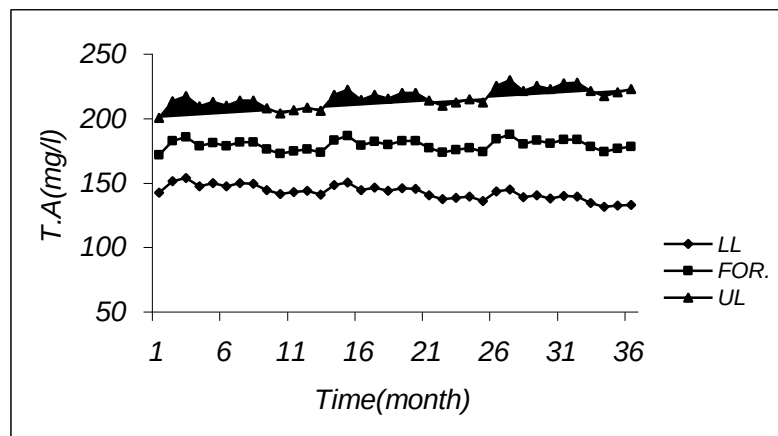


Fig. B.55: Probability limits of forecasted values obtained by Box-Jenkins Seasonal model for T.A , Ca, Cl & EC of Kufa station from 2002 to 2004 .

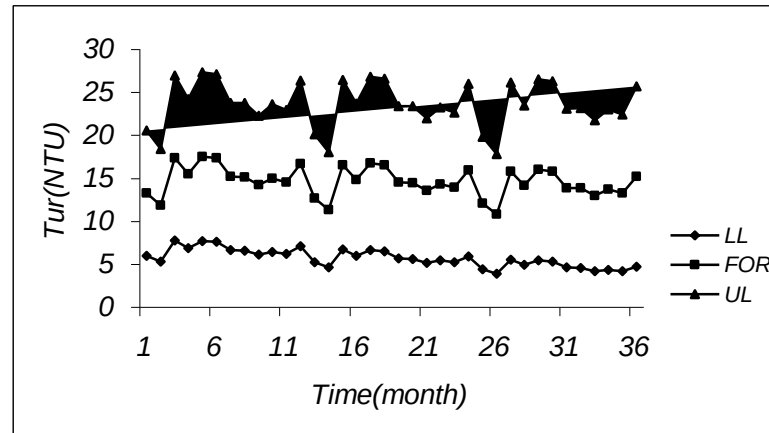
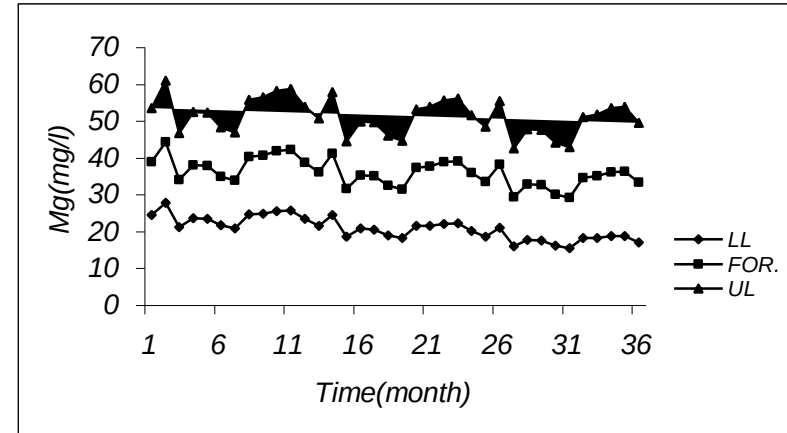
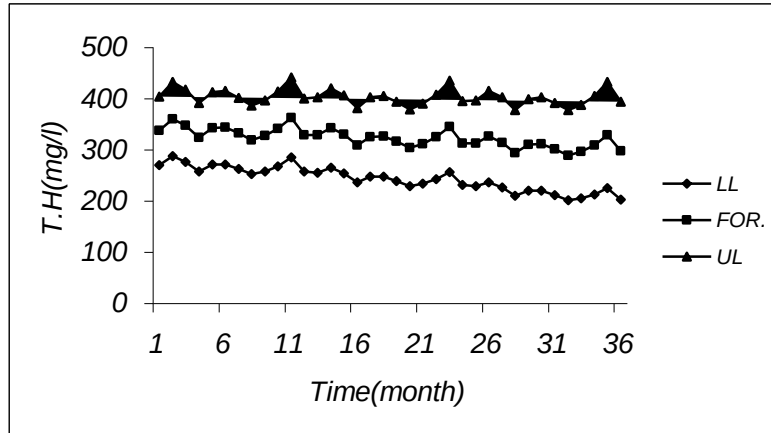


Fig. B.56: Probability limits of forecasted values obtained by Box-Jenkins Seasonal model for T.H, Mg, Tur of Kufa station from 2002 to 2004 .

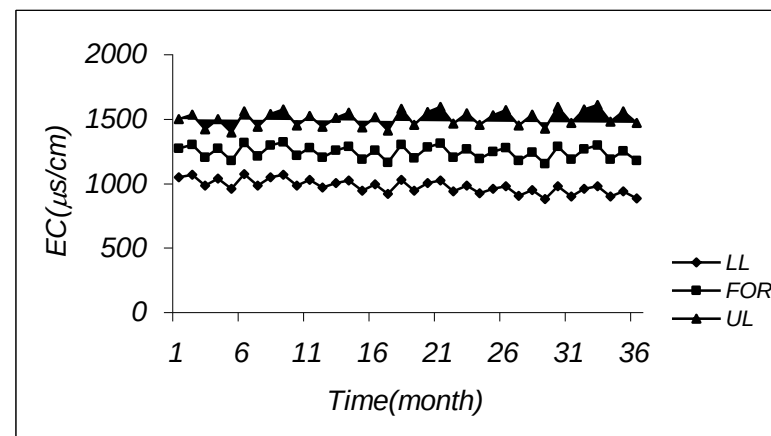
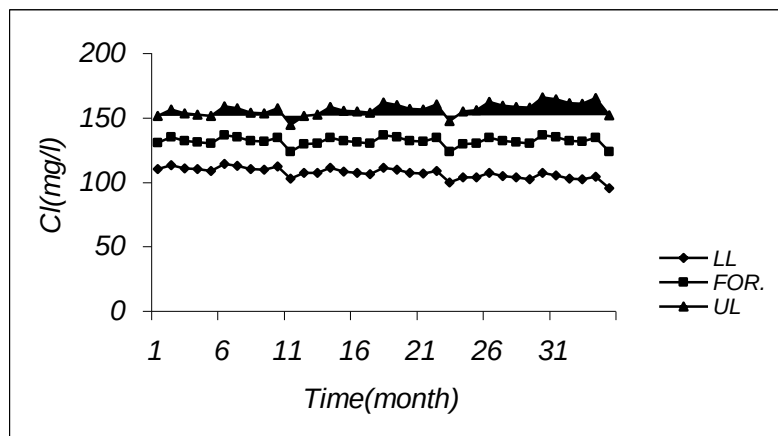
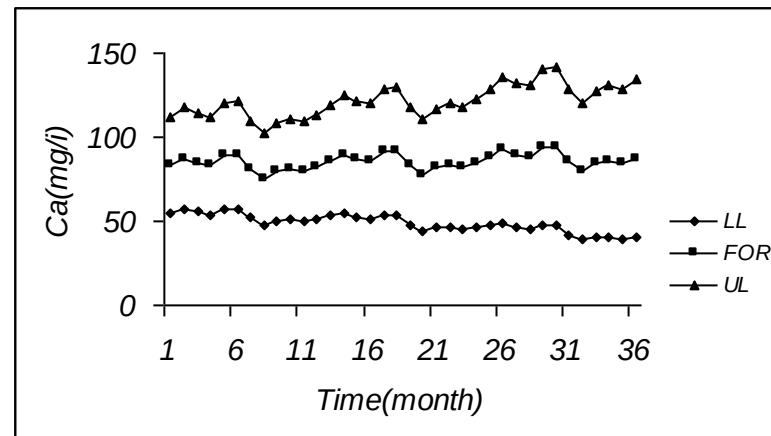
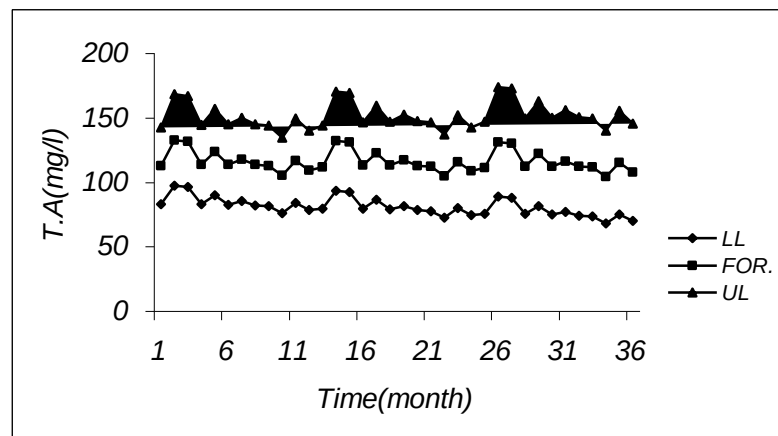


Fig. B.57: Probability limits of forecasted values obtained by Box-Jenkins Seasonal model for T.A, Ca, Cl& EC of Diwaniya station from 2002 to 2004 .

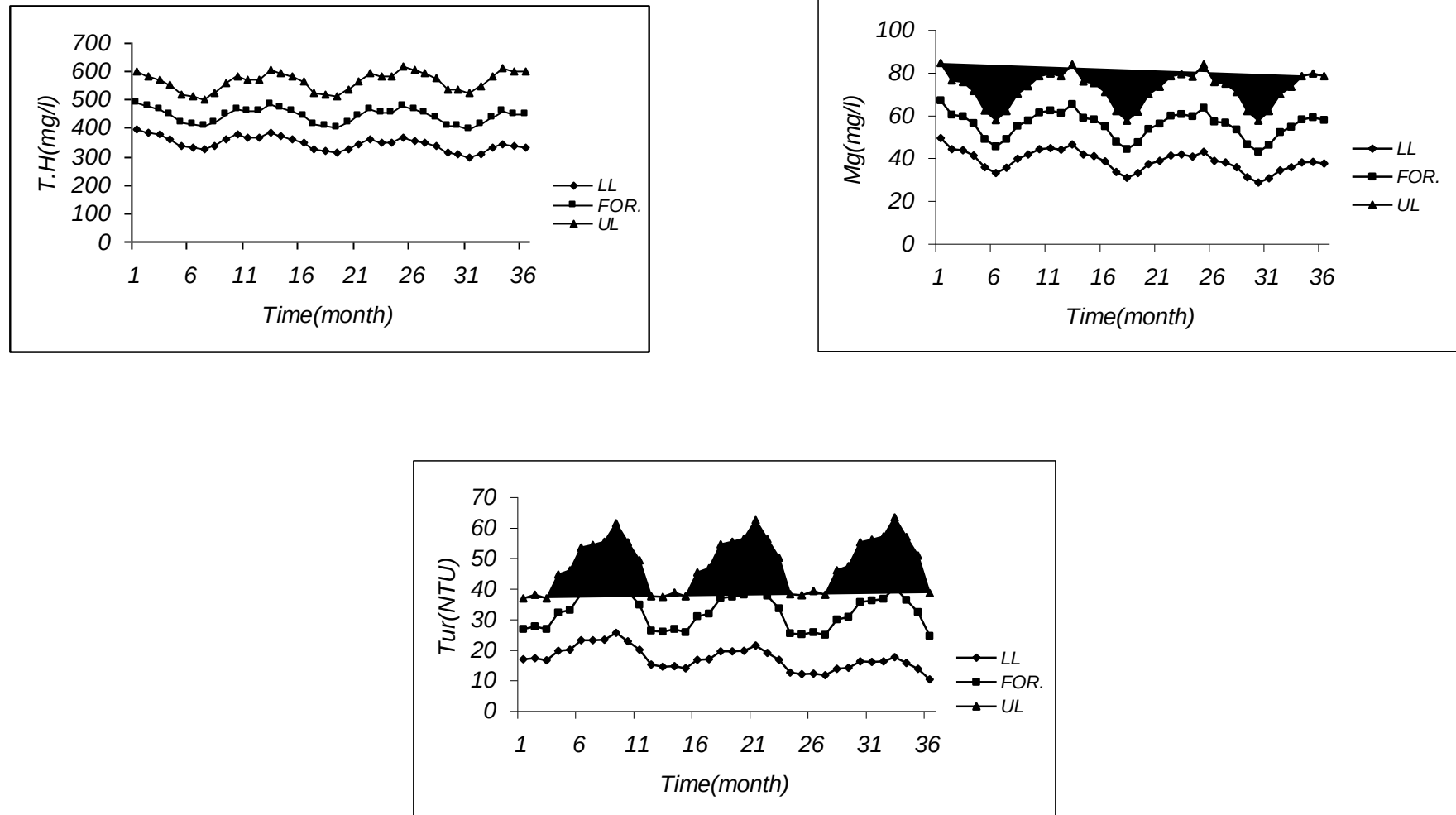


Fig. B.58: Probability limits of forecasted values obtained by Box-Jenkins Seasonal model for T.H, Mg, Tur of Diwaniya station from 2002 to 2004 .

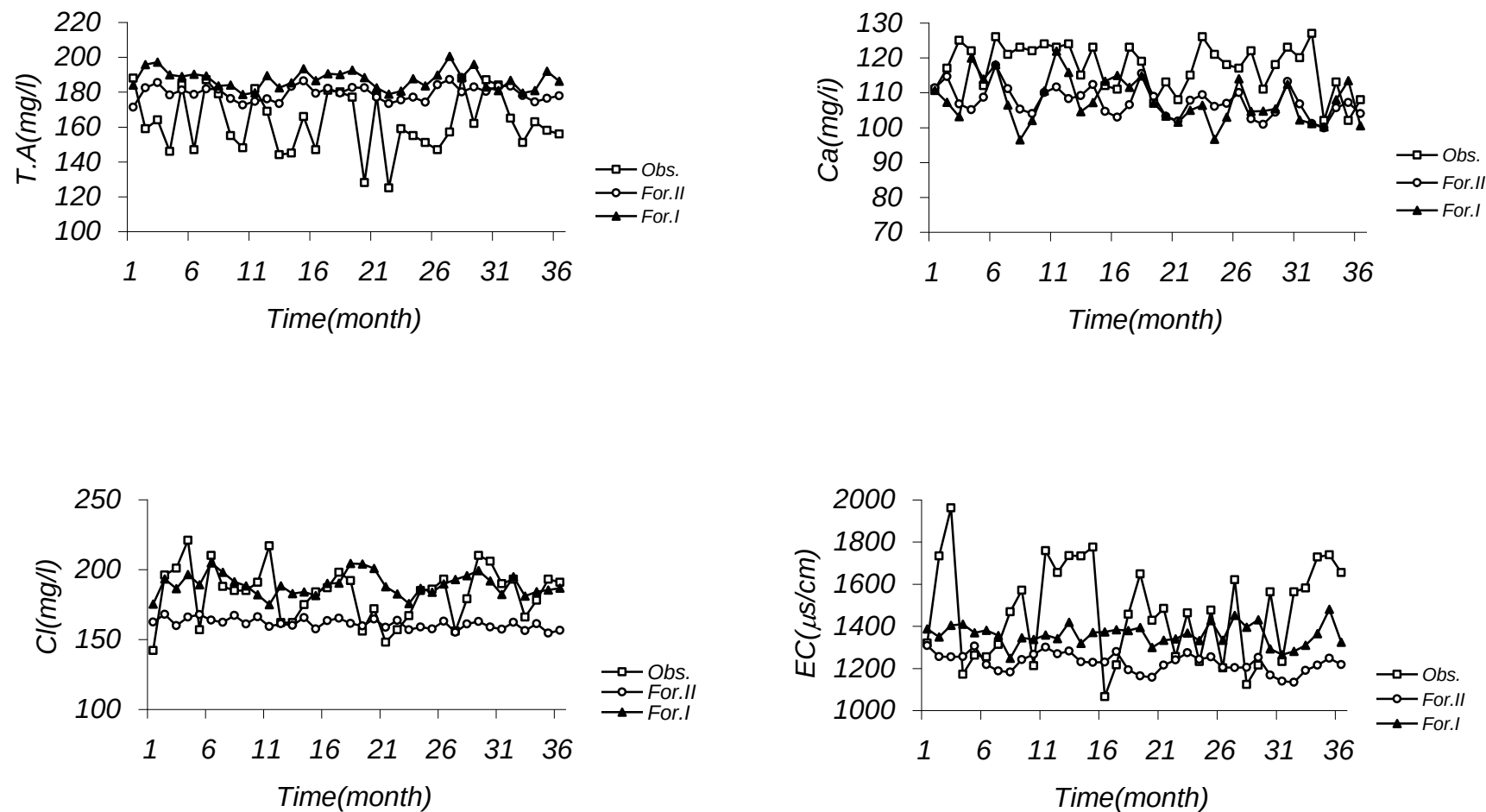


Fig. B.59: Comparison between observed series and the series generated by (I-ARMA ,II-Box-Jenkins) methods for T.A ,Ca ,Cl &EC at Kufa station from 2002 to 2004.

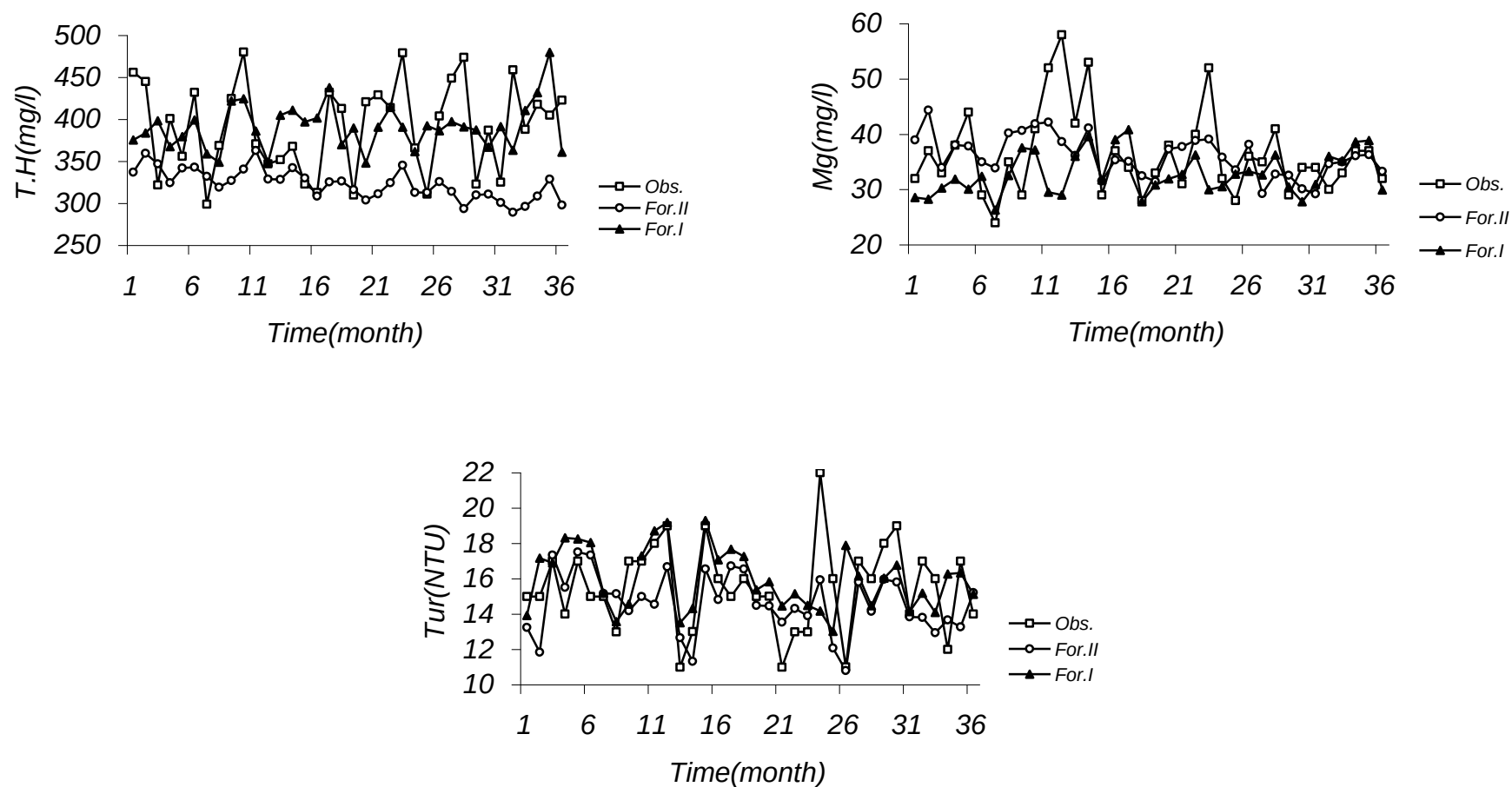


Fig. B.60: Comparison between observed series and the series generated by (I-ARMA ,II-Box-Jenkins) methods for T.H, Mg, Tur at Kufa station from 2002 to 2004.

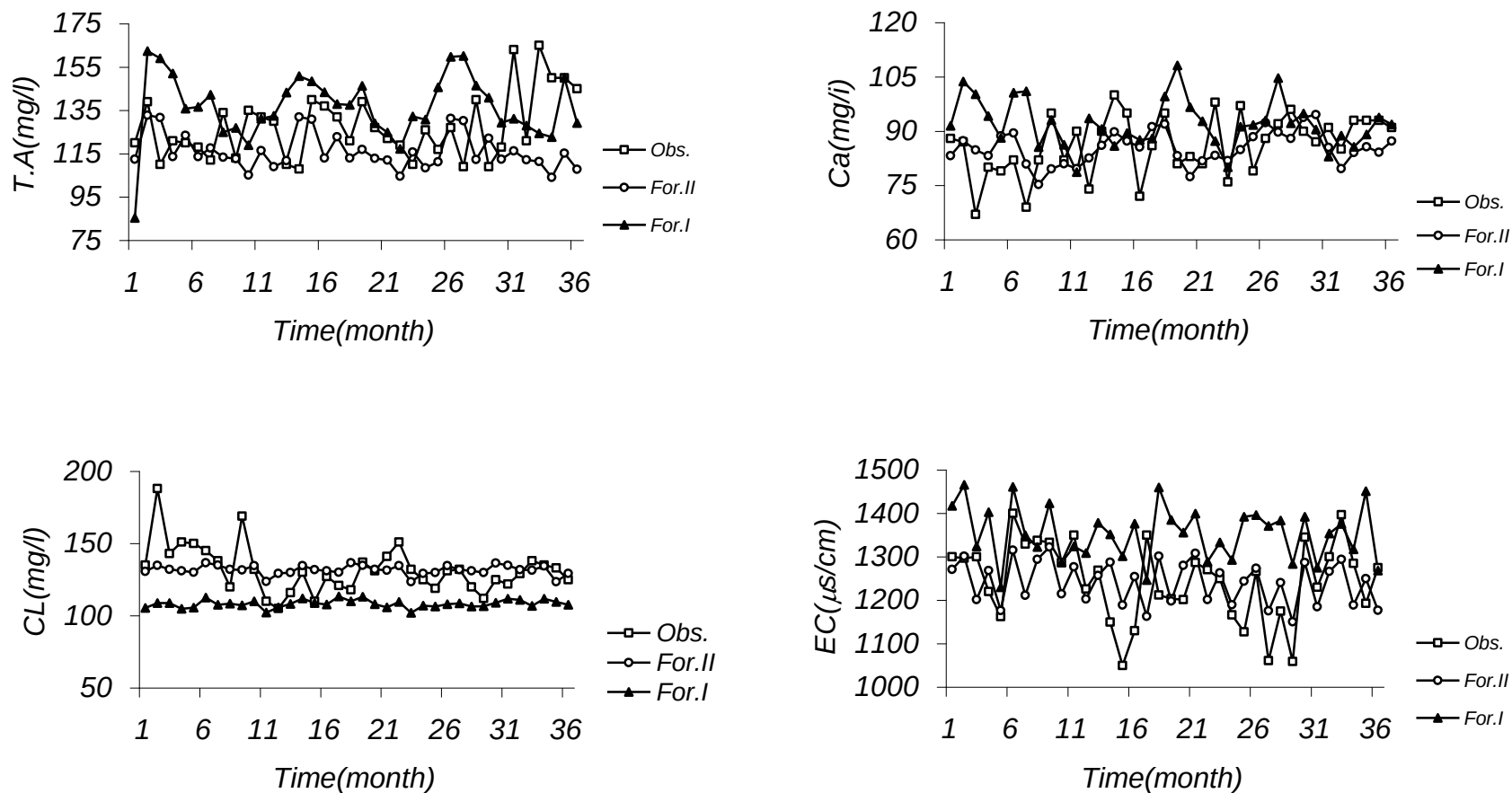


Fig. B.61: Comparison between observed series and the series generated by (I-ARMA ,II-Box-Jenkins) methods for T.A ,Ca ,Cl & EC at Diwaniya station from 2002 to 2004.

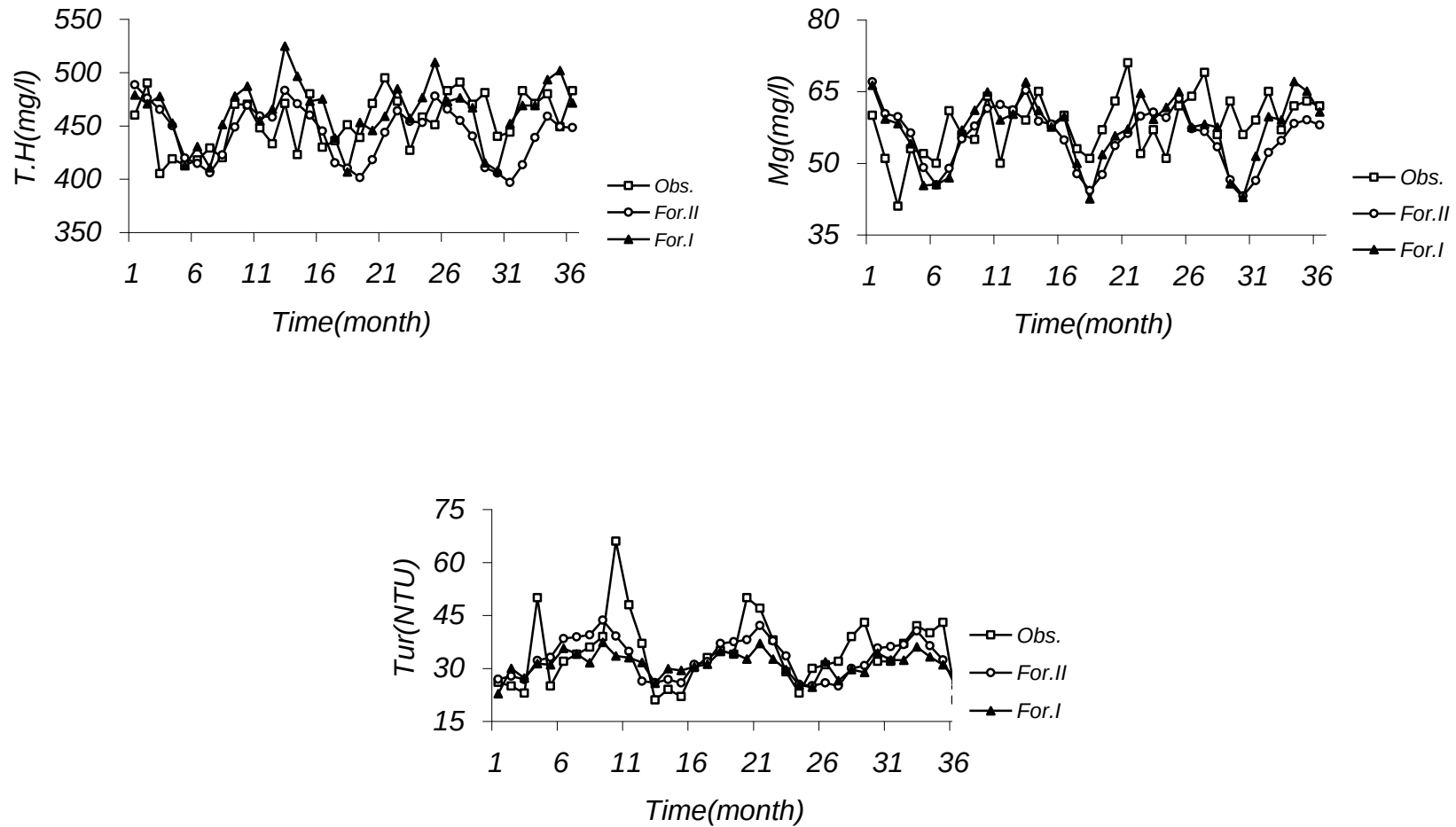


Fig. B.62: Comparison between observed series and the series generated by (I-ARMA ,II-Box-Jenkins) methods for T.H, Mg, Tur at Diwaniya station from 2002 to 2004.

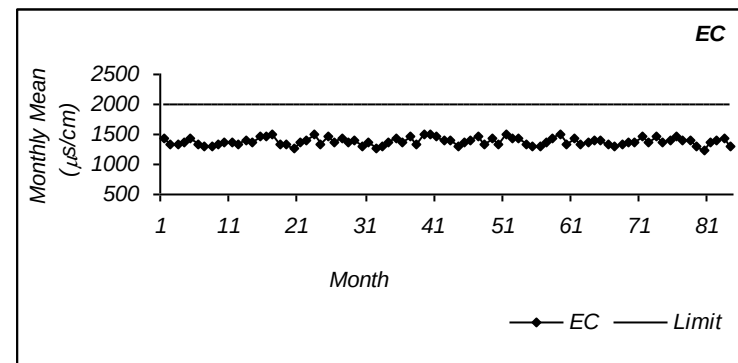
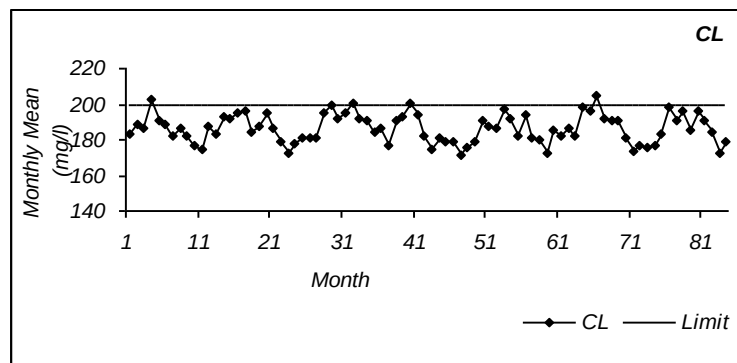
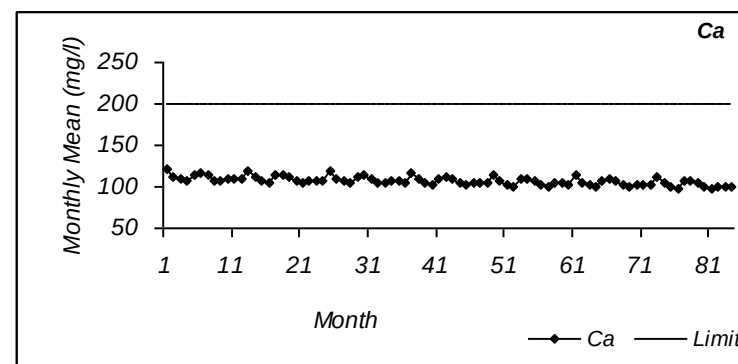
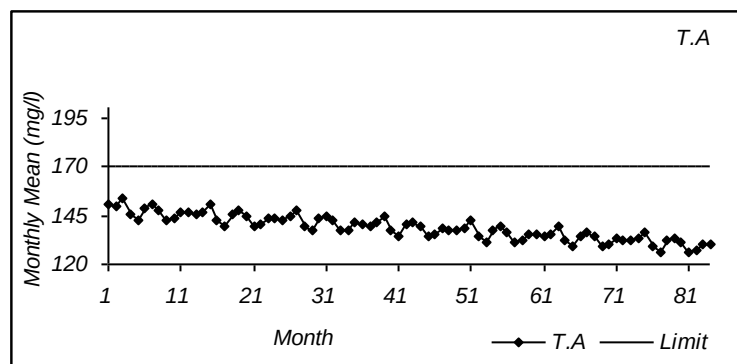


Fig.B.63: Comparison of 7-years(from 2005 to 2011) forecast for T.A ,Ca ,Cl &EC for Kufa station with the river standard .

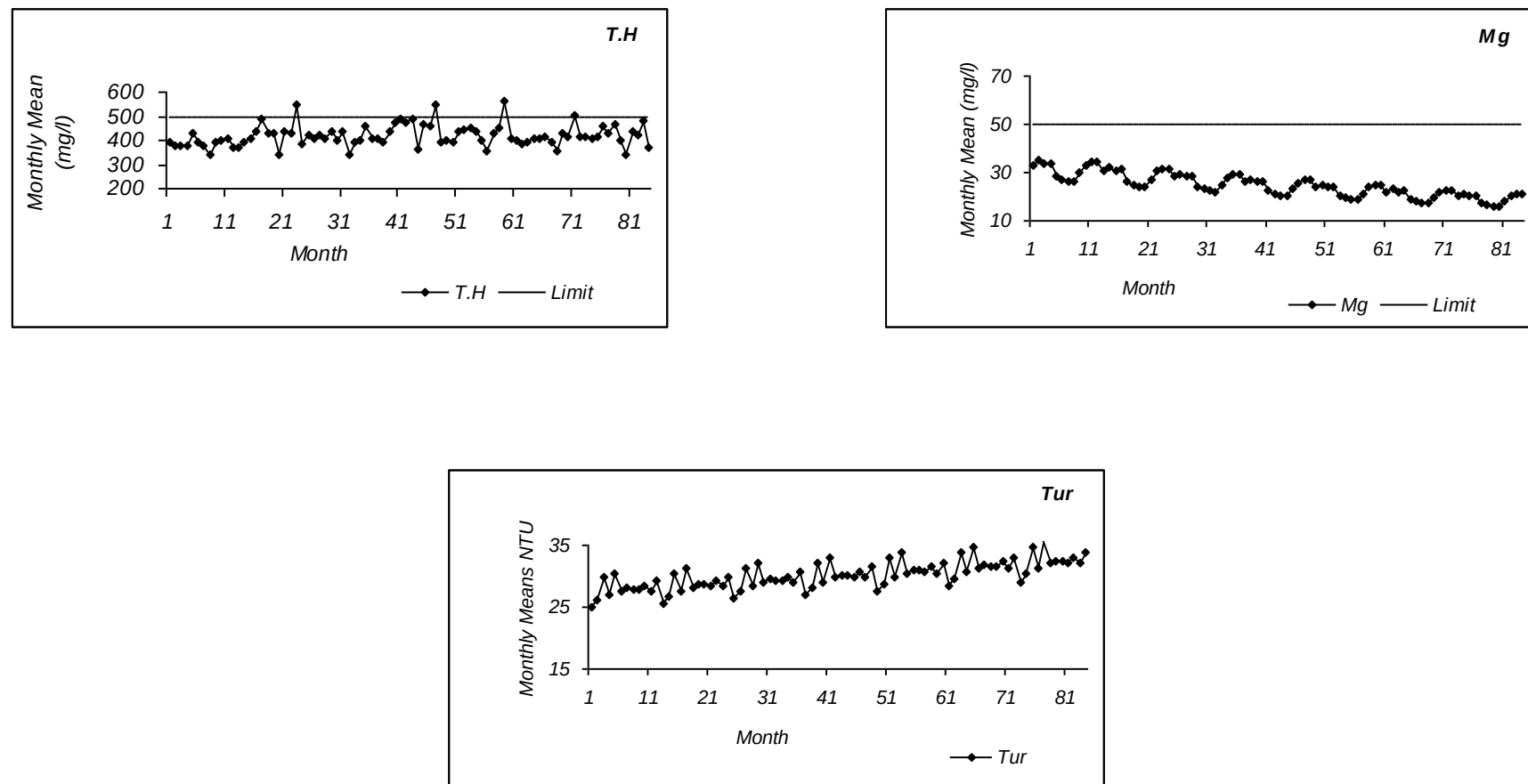


Fig.B.64: Comparison of 7-years(from 2005 to 2011) forecast for T.H ,Mg &Tur for Kufa station with the river standard .

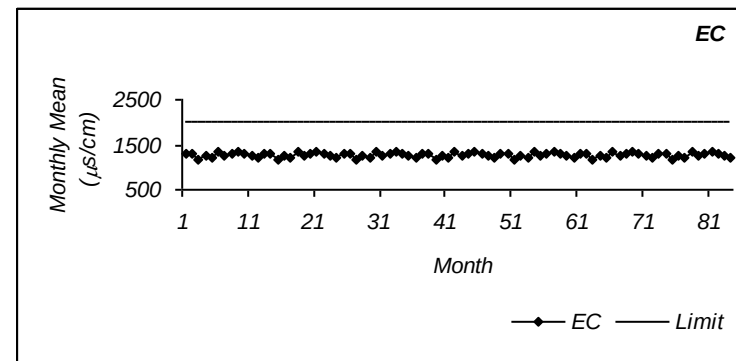
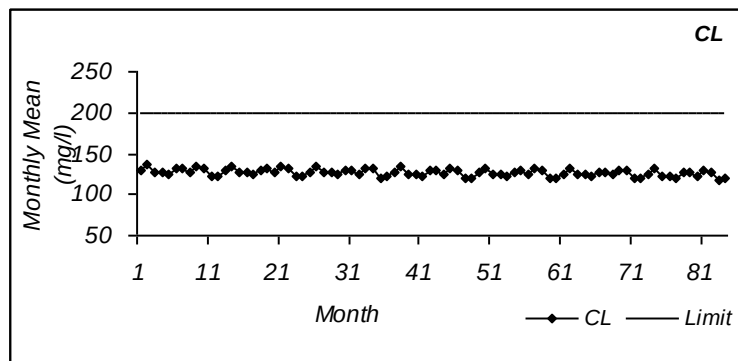
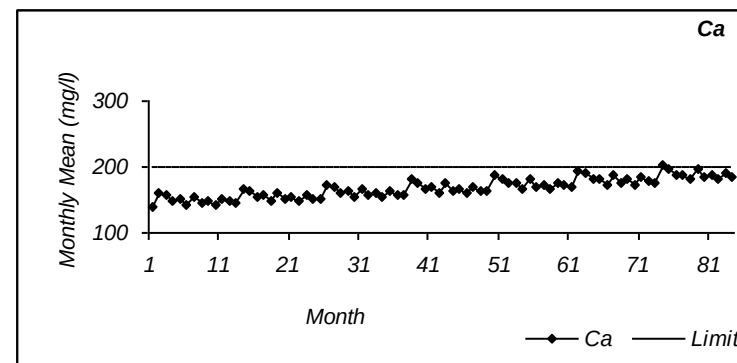
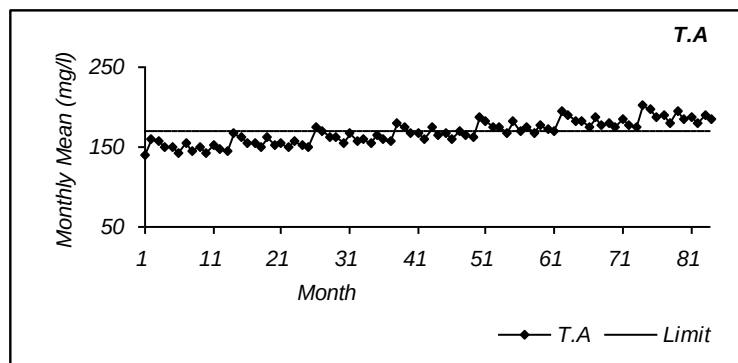


Fig.B.65: Comparison of 7-years(from 2005 to 2011) forecast for T.A ,Ca ,Cl &EC for Diwaniya station with the river standard .

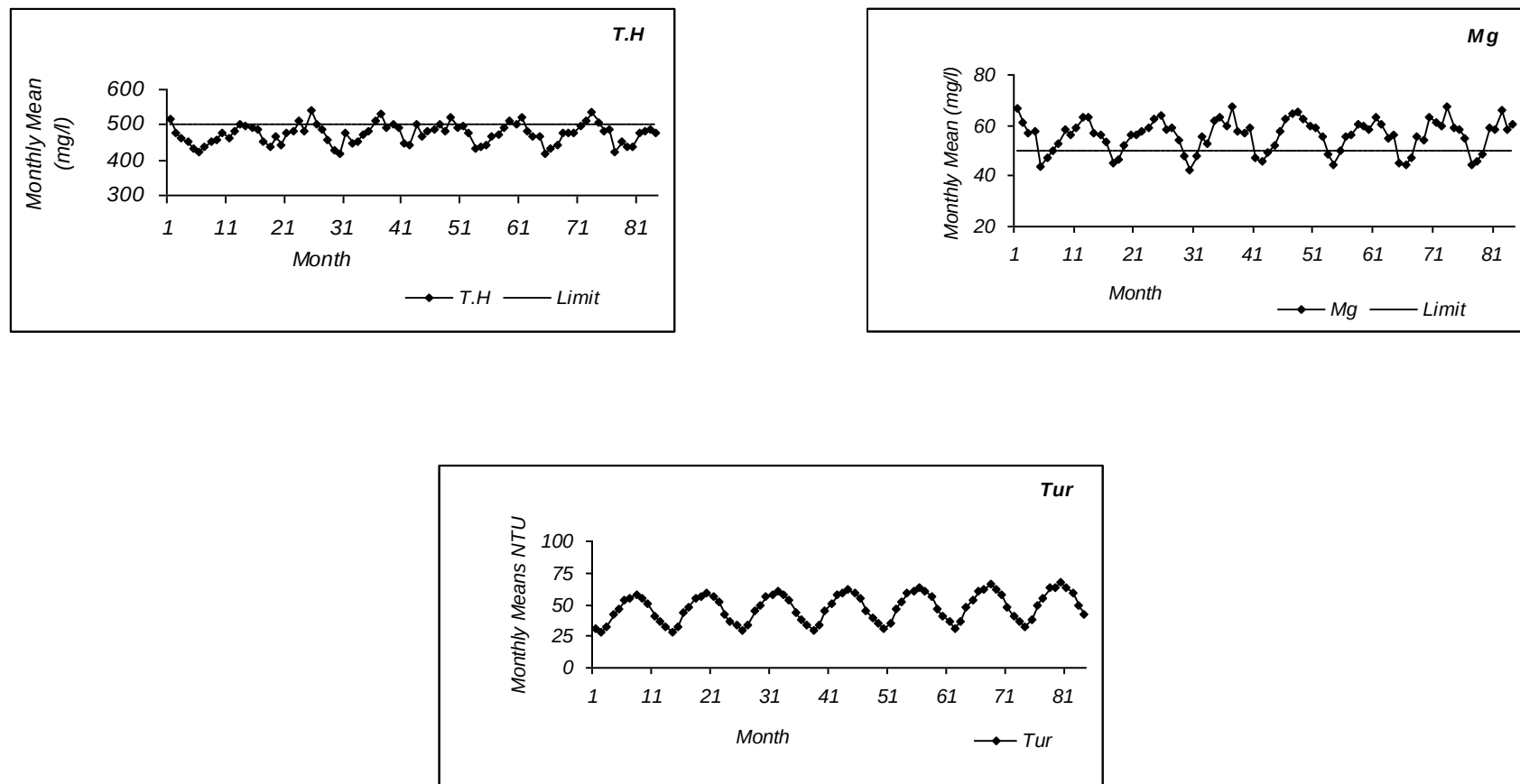


Fig.B.66: Comparison of 7-years(from 2005 to 2011) forecast for T.H ,Mg &Tur for Diwaniya station with the river standard .

Chapter One

Introduction

1.1 Prelude

Among the most important issues facing civilization in the 21st century is the growing scarcity of fresh and clean water. Maintaining and improving the quality and quantity of freshwater have long term economic, health, and ecological implication. The purpose of most water quality studies is to point out the information and knowledge necessary for management of water resources and their use, control and development.

Because of the rapid economic development, industrialization, and population increase, there becomes necessary to analyze and assess water quality parameters and to examine methods for forecasting water quality in water bodies for developing appropriate control strategies in the future. Analysis and examine forecasting methods of water quality parameter have great attention of researchers and communities in the world especially in developing countries. Stochastic models are widely used in forecasting water quality parameters, flow, rainfall and other hydrologic phenomena. Time series analysis approach has been used to develop parametric models for (daily , weekly , monthly, ...) data by many researchers like (Mahloch, 1974), (Stidinger, 1981), (Cooper and Wood, 1982), (Salas and Abedlmohsen, 1993).

A model which describes the probability structure of a sequence of observations is called a "stochastic process ". A time series of N successive observations $z = (z_1, z_2, \dots, z_n)$ is regarded as a sample realization, from an infinite population of such samples, which could have been generated by the process. An important class of stochastic processes is the stationary processes. They are assumed to be in a specific form of statistical equilibrium and in particular vary about a fixed mean. Particular stationary stochastic processes of value in modeling time series are the autoregressive

(AR), moving average (MA), and mixed autoregressive moving average processes (ARMA). Another class of stochastic processes which is non stationary processes like autoregressive integrated moving average (ARIMA) models (Box and Jenkins, 1976).

The water quality variation in a river may be modeled by deterministic approach, in which the cause – effect relationships are represented by differential equations, or by the stochastic approach in which the variation of the magnitude of one or more parameters of water quality is represented as a function of time. The former approach has been used for predicting the steady – state water quality condition along a river and to predict the short – term transient state of water quality parameter. The latter approach enables the identification of trends periodic phenomena presented in natural data series, their decomposition, and subsequent synthesis for data generation and /or forecasting. There is also third approach in which the water quality equations have been transformed from deterministic to stochastic differential equations (Jayawardena & lai, 1989).

The special feature of time series analysis is the fact that successive observations are usually not independent and that the analysis must take into account the time order of the observation. When successive observations are dependent future value may be predicted from past observations. If a time series can be predicted exactly, it is said to be deterministic. But most time series are stochastic in that the future is only partly determined by past values. For stochastic series exact predictions are impossible and must be replaced by the idea that future values have a probability distribution which is conditioned by a knowledge of past values, (Chatfield, 1982).

There are several possible objectives in analysis of a time series. These objectives may be classified as description, explanation, prediction

and control. Univariate stochastic models are widely used and in these models variables are separated from each other. Other models are used in the generation of multivariate synthetic sequences, where the sequences may pertain stations, seasons, or combination of stations and seasons which are called multivariate models (Matalas, 1967). Single site models deal with the variables at a certain point(the correlation exists in time only), while multisite models deal with the correlations in both time and space.

This research covers the study and the analysis of seven monthly water quality parameters-which are listed in table (1.1)- of water of three stations (AL- Hindya, Kufa, and Diwaniya) on the branches of Euphrates river. The data which are used in the analysis span over ten years (1992-2001), whereas three years (2002-2004) are used for comparison between the observed and fitted data .

Table 1.1:List of water quality parameter which are used in the analysis.

Parameter	Symbol	Unit
Total Alkalinity	T.A	mg /l
Calcium	Ca	mg/l
Chloride	Cl	mg/l
Electrical Conductivity	EC	μ s/cm
Total Hardness	T.H	mg/l
Magnesium	Mg	mg/l
Turbidity	Tur	NTU

1.2 Description of Study Area and Data :

Euphrates river has a large importance for the Iraqi environmental researchers because of the detrimental effect of pollutants resulting from treated and untreated domestic waste water, treated and untreated industrial waste water and farming and agricultural pollutants.

The main source of Euphrates river water are rain water, stored water such as lakes and reservoirs. AL-Hilla river is a branch of Euphrates river, it branches at AL-Hindya Barrage town (Fig 1.1), it's length through Babylon governorate is about (104)km ,then AL-Hilla river branches into two main rivers; AL-Dagarah and AL-Diwaniya river. The length of the latter is about (124)km, the center of AL-Diwaniya city lies about (53) km from the river. Kufa river is a branch of Euphrates river ,it branches after AL-Kifil town directly, the length of Kufa river is about (43)km, the center of Kufa city lies about (24) km from the river (Sousa, 1945).

Three stations on Euphrates river branches are used for this research include; AL-Hindya station at AL-Hindya Barrage town, Kufa station at Kufa city and Diwaniya station at Diwaniya city. Monthly mean values of seven water quality parameters for ten years (1992-2001) are used for basic analysis and three years (2002-2004) are used for comparison with the series of data generated using the time series model.

1.3 Objectives of the Study

The main objectives of the present study are:

1. Applying the well-known mixed autoregressive moving average ARMA and Box-Jenkins seasonal models to water quality data from three governorates in the middle of Iraq i.e. Babylon, Najaf, and Diwaniya governorates and comparing the results of the two models .
2. Using the above mentioned models to forecast seven years of future time series compared with water quality standards.

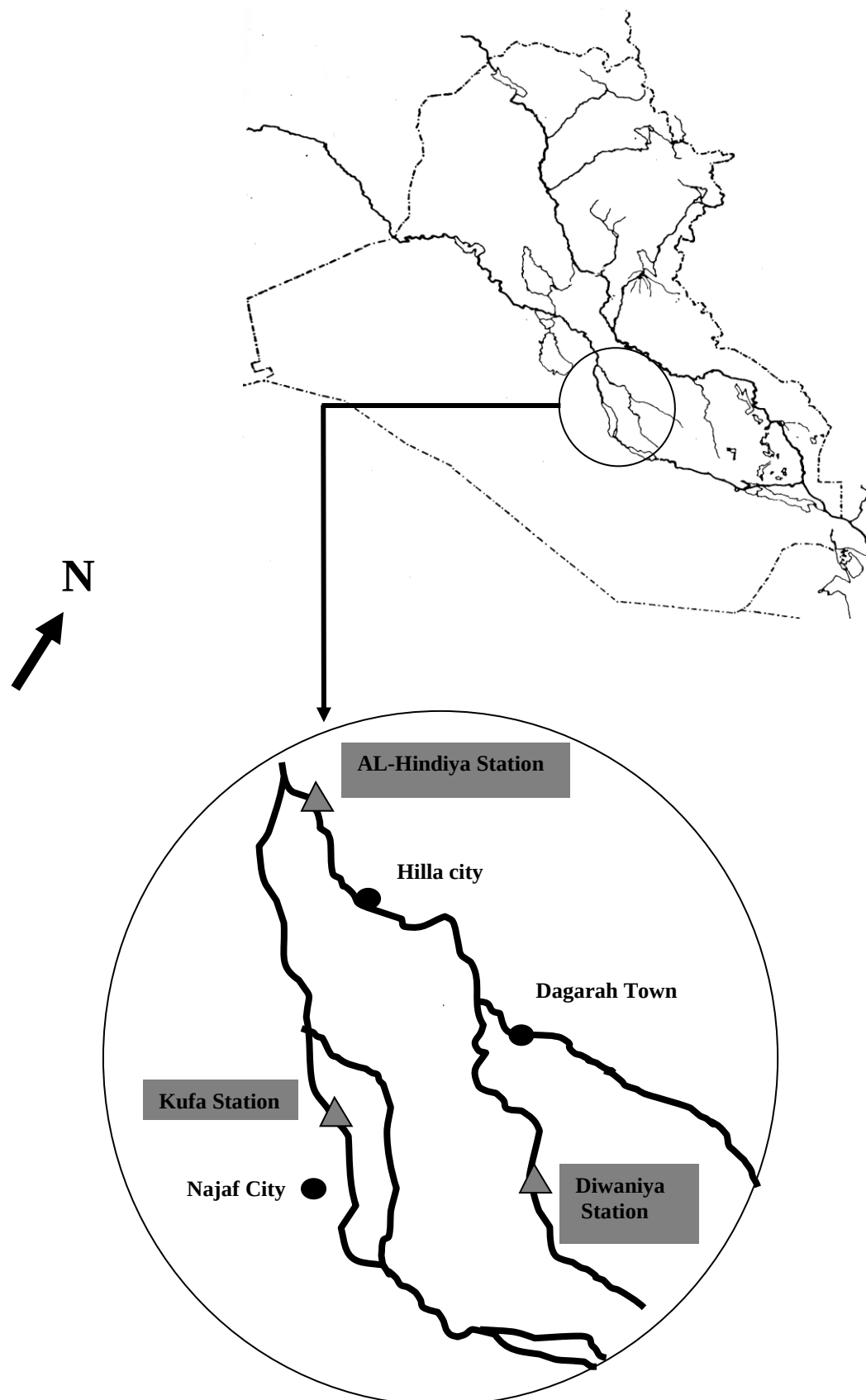


Fig. 1.1:Map of studying area

Chapter Two

Stochastic Methods

2.1 Components of Time Series

A time series is a collection of observations made sequentially in time. A time series is said to be continuous when observations are made continuously in time. The term continuous is used for series of this type even when the measured variable can only take a discrete set of values, a time series is said to be discrete when observations are taken only at specific time usually spaced, the term discrete is used for the series of this type even when the measured variable is a continuous variable (Chatfield, 1982).

The time series may consist of four components depending on the type of variable and the averaging time interval. These components may exist in monthly time series, which may be formulated by (AL-Suhaili, 1986) :

$$X_t = J_t + T_t + P_t + W_t, \dots\dots\dots(2-1)$$

where X_t is the time series observations at time $t = 1, 2, 3, \dots, N$.

J_t is the jump component.

T_t is the trend component.

P_t is the periodic or seasonal component.

W_t is the stochastic component includes depended and independent parts.

N is the No. of observations.

Jump, trend and periodic components represent the deterministic part of the process while the stochastic component represents the non deterministic part, therefore, the first three components should be detected and identified by suitable formulations and decomposed from the stochastic component.

The methodology for the identification and decomposition are explained herein this chapter.

2.2 Stationarity

A stationary time series is a series statistic properties do not change with time. That is there is no synthetic change in mean (no trend), if there is synthetic change in variance and if strictly periodic variance have been removed. The time series is said to be strictly stationary if its properties are unaffected by a change of time origin; that, if the joint probability distribution associated with (n) observations $Z_{t1}, Z_{t2}, \dots, Z_{tn}$ made at any set of time $t_1, t_2, \dots, t_n$ is the same as that associated with (n) observation $Z_{t1+k}, \dots, Z_{tn+k}$, made at time $t_{1+k}, t_{2+k}, \dots, t_{n+k}$. Thus for a discrete process to be strictly stationary, the joint distribution of any set of observations must be unaffected by shifting all the times of observation forward or backward by any integer amount (k) (Mahmood, 2000). This implies that no trend or jump exists in the series. Hence, the classification of a time series as stationary or non stationary is not merely for descriptive purpose. Available statistical methods are designed to analyze stationary time series. Therefore, if the time series is defined to be non stationary, it must be transformed into stationary one before these are applied (Matalas, 1967).

In Box and Jenkins model the differencing technique is used to make the mean of time series as a constant as possible and its degree of the regressive equation equal to zero by taking appropriate number of differences for the original time series X_t which is explained as follow:

$$\begin{aligned} Y_t &= \nabla X_t = X_t - X_{t-1} \\ &= X_t - \beta X_t \\ &= (1-\beta)X_t, \dots\dots\dots(2-2) \end{aligned}$$

where ∇ is the backward difference coefficient .

Y_t is the new time series after taking the difference .

β is the backward shift operator that is $\beta^j X_t = X_{t-j}$

The second –order difference is :

$$\begin{aligned} Y_t &= \nabla^2 X_t = \nabla(\nabla X_t) \\ &= \nabla(X_t - X_{t-1}) = \nabla X_t - \nabla X_{t-1} \\ &= X_t - X_{t-1} - X_{t-1} + X_{t-2} = X_t - 2X_{t-1} + X_{t-2} \\ &= (1 - 2\beta + \beta^2)X_t \\ &= (1 - \beta)^2 X_t \quad \dots\dots\dots(2-3) \end{aligned}$$

In general if (d) differences are taken from the original time series,

$$Y_t = \nabla^d X_t = (1 - \beta)^d X_t \quad , \quad \dots\dots\dots(2-4)$$

where d is the maximum number of differencing necessary to make the time series stationary in the mean (Box and Jenkins, 1976) .

2.3 Homogeneity of Time Series

Homogeneity definition requires at least two conditions:-

1. The data series must not contain any systematic error.
2. All conditions should be constant.

If these conditions are satisfied then the series may be considered as homogenous. These two conditions imply that the homogenous series should be free from trend and jump component respectively, therefore homogeneity should be insured at the beginning of the analysis by detection and removal of trend and jump components (AL-Suhaili, 1986). In the ARMA model the trend and jumps are detected and removed as described in the following sections.

2.4 Trend and Jump Detection

Many authors define the jump component as a trend component, while these two components differ somehow from each other. Some of those authors define the jump component as a man-made trend, it may be

defined accurately as a sudden slippage (either negative or positive) in the parameters of historical data such as means , standard deviations ... etc. Human activities may cause jump, for example; the construction of factories, waste water treatment plants or drains upstream of the observation station which will cause a positive jump in pollution levels while reducing or shifting the industrial waste water or drainage water away from the point of observation may result in negative jump in pollution levels. Trend is often the most important component time series. It may be linear, growing at a constant amount over time or it may be follow any of the nonlinear patterns , for example it may be growing at a constant rate over time (in which case it would be linear on a log scale) (Wonnacott and Wonnacott, 1990) .

Trend component may be defined as a long term change in the mean. The main causes of trend in water quality data are either sequential man – made changes within the catchments area (such drains many heavy metals and ions from factories and other source of pollution to the river), or natural causes (such as climatic changes), or inconsistency (systematic errors) (AL-Husseini, 2000). Applications of statistical tests for trend have generally been univariate in which they have considered only a single water quality variable at a single monitoring location whereas routine monitoring programs typically measure several variables at several locations ,and water quality process are perhaps more logically viewed as multivariate (Loftis et .al., 1991).

Split – sample method is used for the detection of trend and jump which is explained as follows :

Split – Sample Method

This method is carried out by using the split – sample approach which divides the entire sample into two sub- samples, then testing the

difference between the means and standard deviations of these two sub-samples at the (97.5) percent probability level of significance using the (t-test) & (f – test) (Appendix A). This method is the most powerful for testing homogeneity, which is suggested by Yevjevich. This method is used mostly for detection of jump component, but it may also be used for the detection of trend component when the trend is steep enough to produce high differences in mean and standard deviations of the two sub- samples (Yevjevich, 1975, quoted in AL-Tikriti, 2001).

2.5 Removal of Trend and Jump Component

For nonhomogeneity removal Yevjevich (1975) suggests fitting a linear regression equations for both annual means and annual standard deviations as follows (AL-Suhaili, 1986) :

$$Y_{i,j} = \frac{X_{i,j} - \bar{X}_i}{S_i} Sd_2 + M_2 , \quad \dots\dots\dots(2-4)$$

where i,j are the annual and seasonal positions of observations respectively. $Y_{i,j}$ is the homogeneous series.

X is the original or non – homogeneous series .

M_2 , Sd_2 are the mean and standard deviation of the second sub- sample respectively.

$\bar{X}_i$, S_i are obtained from linear regression equations for annual means and standard deviation against years, and expressed as follows:

$$\bar{X}_i = A_i + B \quad , \quad \dots\dots\dots(2-5)$$

$$S_i = A_{1i} + B_1 \quad , \quad \dots\dots\dots(2-6)$$

where $\bar{X}_i$ is the mean at year i.

S_i is the standard deviation at year i.

A , B , A_1 and B_1 are the regression coefficients .

2.6 Detection of Periodic or Seasonal Component

It is said that the time series has seasonal behavior or periodic component if it has irregular pattern that repeats itself every period of time. The period of time is named seasonal cycle (S) which may be a day, a month, a season, or a year (Mahmood, 2000). When the time series has monthly data, seasonal pattern usually repeats itself every year, which means $S=12$. Cyclic behavior may appear in more than statistical parameter such as mean, standard deviation, skewness, kurtosis and serial correlation coefficients.

Detection of periodicity or seasonality of any time series can be made by autocorrelation analysis (Chatfield, 1982). The sample autocorrelation coefficients measure the correlation between observation at different time-lags. These coefficients often provide insight into the probability model which generates the data. If the series is periodic, the plot of autocorrelation coefficient with different time-lags will be periodic also, otherwise it is not.

The autocorrelation coefficient between observations at lag (k) for any random variable X_i is given by

$$r_k = \frac{\sum_{i=1}^{N-K} (X_i - \bar{X})(X_{i+k} - \bar{X})}{\sum_{i=1}^N (X_i - \bar{X})^2} \dots\dots\dots(2-7)$$

where N is the number of observations.

r_k is the correlation coefficient at lag k.

X_i is the observation at time i.

$\bar{X}$ is the over all mean.

$$\bar{X} = \sum_{i=1}^N X_i / N \dots\dots\dots(2-8)$$

The autocorrelation coefficients are usually calculated by computing the series of autocovariance coefficients c_k , which is defined as:

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (X_t - \bar{X})(X_{t+k} - \bar{X}) \quad \dots\dots\dots(2-9)$$

where c_k is the autocovariance coefficient at lag k .

The autocorrelation coefficients may then be computed from

$$r_k = \frac{c_k}{c_0} \quad \dots\dots\dots(2-10)$$

where c_0 is the autocovariance coefficient at lag zero.

2.7 Removal of Periodic or Seasonal Component

Two practical methods may be used to remove the periodic component from a time series as follows :

a-The Nonparametric Method

The removing periodicity in the original series is necessary and can be employed for monthly sequences is to deseasonalize the series by subtracting estimated monthly mean and dividing by the estimated monthly standard deviation for the series as follows :

$$W_{i,j} = \frac{Y_{i,j} - \hat{\mu}_j}{\hat{\sigma}_j} \quad , \quad \dots\dots\dots(2-11)$$

where $W_{i,j}$ is the stochastic component at year i , and month j .

$Y_{i,j}$ is the homogeneous series.

$\hat{\mu}_j, \hat{\sigma}_j$ are the mean and standard deviation for month (j) respectively computed by:-

$$\hat{\mu}_j = \frac{1}{n} \sum_{i=1}^n Y_{i,j} \quad , \quad \dots\dots\dots(2-12)$$

$$\hat{\sigma}_j = \left[\frac{1}{n-1} \sum_{i=1}^n (Y_{i,j} - \hat{\mu}_j)^2 \right]^{1/2} \quad , \quad \dots\dots\dots(2-13)$$

where n is the number of years.

b- The Harmonic Method

If a periodicity exists, it can be described mathematically by harmonic functions and represented by a Fourier series, the periodic component in the mean can be estimated from (Srikanthan & McMahon, 1982):

$$\mu_j = \mu + \sum_{h=1}^{w/2} \left(A_h \cos \frac{2\pi jh}{w} + B_h \sin \frac{2\pi jh}{w} \right), \dots \dots (2-14)$$

in which

$$A_h = \frac{2}{w} \sum_{j=1}^w \bar{Y}_j \cos \frac{2\pi jh}{w} \dots \dots \dots (2-15)$$

$$B_h = \frac{2}{w} \sum_{j=1}^w \bar{Y}_j \sin \frac{2\pi jh}{w} \dots \dots \dots (2-16)$$

$$\bar{Y}_j = \frac{1}{n} \sum_{i=1}^n Y_{i,j} \dots \dots \dots (2-17)$$

$$\mu = \frac{1}{w} \sum_{j=1}^w \bar{Y}_j \dots \dots \dots (2-18)$$

where n is the number of years.

$w=12$ for monthly data (Fernando and Jayawardena, 1994).

$\hat{\mu}_j$ is the harmonically fitted means or variance.

A_h, B_h are the Fourier coefficients .

$Y_{i,j}$ is the homogeneous series at year i & month j .

Similarly, the periodic component in standard deviation can be estimated from

$$\sigma_j = \sigma + \sum_{h=1}^{w/2} \left(A_h \cos \frac{2\pi jh}{w} + B_h \sin \frac{2\pi jh}{w} \right) \dots \dots \dots (2-19)$$

In which A_h, B_h , and σ can be computed from equation (2-15), (2-16) & (2-18) by replacing of $(\bar{Y}_j)$ by (S_j) which is obtained from :

$$S_j = \left[\frac{1}{n-1} \sum_{i=1}^n (Y_{i,j} - \bar{Y}_j)^2 \right]^{1/2} \dots \dots \dots (2-20)$$

The periodic components in the means standard deviation are removed from $(Y_{i,j})$ by

$$Y'_{i,j} = \frac{Y_{i,j} - \mu_j}{\sigma_j} \dots\dots\dots(2-21)$$

in which $Y'_{i,j}$ is stationary in the mean and variance. The mean and standard deviation of $Y'_{i,j}$ will not exactly equal to 0 and 1, respectively, and a further transformation is required :

$$W_{i,j} = \frac{Y'_{i,j} - \bar{Y}'}{S} \quad , \quad \dots\dots\dots(2-22)$$

where $W_{i,j}$ is the stochastic series .

$\bar{Y}'_j, S_j$ is the mean and standard deviation respectively of $Y'_{i,j}$ at month j

For monthly data the maximum number of harmonics is $(\frac{W}{2})$. The method of critical explained variance (ΔP_h) was used to estimate the number of significant harmonic to represent the periodic function.

The ratio (ΔP_h) is calculated by (Resoner and Yevjevich, 1966, quoted in AL-Husseini, 2000).

$$\Delta P_h = \frac{(A_h^2 + B_h^2)}{2 * \text{Var } \mu_j} \dots\dots\dots(2-23)$$

These values are then ranked in descending order sequence and summed to

$$P_h = \sum_{i=1}^h \Delta P_i \quad \text{for } h=1,2,3,\dots,\frac{W}{2} \quad , \quad \dots\dots\dots(2-24)$$

where P_h is the variance explained by h , harmonics.

For most hydrological series the harmonics of any parameter are usually not significant beyond the first six harmonics, so P_h is calculated from equation (2-26) and then.

a) If $P_6 \leq P_{\min}$ then no harmonic is significant.

- b) If $P_{\min} < P_6 \leq P_{\max}$ then all harmonics are significant .
- c) If $P_6 > P_{\max}$ then the first , h harmonics whose P_h value exceeds $P_{\max}$ are significant.

where the critical values $P_{\min}$ and $P_{\max}$ are calculated from

$$P_{\min} = a * \sqrt{\frac{W}{nc}} , \quad \dots\dots\dots(2-25)$$

$$P_{\max} = 1 - P_{\min} , \quad \dots\dots\dots(2-26)$$

respectively ,where c is the order of the highest moment used in estimating the parameter part ,c=1 for mean and c=2 for standard deviation.

w is the number of seasons.

n is the number of years .

a is a constant, which is suggested to be (0.033) by Yevjevich (AL-Suhaili, 1986).

2.8 Stochastic Component

Stochastic component is the remaining part after the separation of periodic component defined by Fourier series from homogenous (free from trend component). It is termed as the dependent stochastic component of the presses and denote as $(W_{i,j})$. The stochastic component contains a dependent and independent part. The dependent part may be represented by an ARMA(p,q)or ARIMA (p, d, q) models, where p and q is the order of the autoregressive and moving average models respectively and d is the degree of differencing. The independent part can only be described by some probability distribution function. These models may be for, singlesite, for multisite, for univariate or for multivariate stochastic process.

2.9 ARMA(p,q) Stochastic Models

These models are applied to a single time series in one point. These models are classified as follows:

(i) Autoregressive Model AR(p) or ARMA(p,0)

In this model the current value of the process is expressed as a finite linear aggregate of previous values of the process and a shock (a_t). By denoting the value of a process at equally spaced times $t, t-1, t-2, \dots$ by $W_t, W_{t-1}, W_{t-2}, \dots$ then

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \dots + \phi_p W_{t-p} + a_t, \dots \quad (2-27)$$

where:

$\phi_1, \phi_2, \dots, \phi_p$ are the autoregressive model parameters .

p is the order of the model.

W_t is the value of stochastic series at time (t) .

a_t is the shock (independent part of stochastic process) .

In equation (2-27) the variable is regressed on the previous value of itself ; hence the model is autoregressive.

If an autoregressive operator of order p is defined by:-

$$\phi(\beta) = 1 - \phi_1\beta - \phi_2\beta^2 - \dots - \phi_p\beta^p \dots \quad (2-28)$$

where (β) is the backward shift operator such that $\beta W_t = W_{t-1}$, then the autoregressive model may be written economically as

$$\phi(\beta)W_t = a_t \dots \quad (2-29)$$

The model contains $(p+2)$ unknown parameters $\mu_1, \phi_1, \phi_2, \dots, \phi_p, \sigma_a^2$, which in practice have to be estimated from the data. Where (μ_1) is the mean of the stochastic series (W_t) and the additional parameter (σ_a^2) is the variance of the white noise process (a_t).

(ii) Moving Average Model MA(q) or ARMA(0,q)

In this model each value of (W_t) is assumed to arise from the effect of the previous value of the independent stochastic component (a_t) rather than the dependent stochastic component (W_t) thus .

$$W_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \dots \dots \dots (2-30)$$

This model is called moving average (MA) of order (q). Where $(\theta_1, \theta_2, \dots, \theta_q)$ are the moving average model parameters.

If the moving average operator of order q can be defined by

$$\theta(\beta) = 1 - \theta_1 \beta - \theta_2 \beta^2 - \dots - \theta_q \beta^q \dots \dots \dots (2-31)$$

Then the moving average model may be written economically as

$$W_t = \theta(\beta) a_t \dots \dots \dots (2-32)$$

It includes (q+2) unknown parameter $(\mu_1, \theta_1, \theta_2, \dots, \theta_q, \sigma_a^2)$, which in practice have to be estimated from the data.

(iii) Mixed Models

A useful class of model for time series is that formed from a combination of AR(p) and MA(q) model. A mixed autoregressive moving average model containing (p) AR term and (q) MA term (abbreviated to an ARMA(p, q) model is said to be of order (p, q) (Chatfeild, 1982). It is given by:

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \dots + \phi_p W_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \dots (2-33)$$

Or using the backward shift operator β , then it may be written in the form

$$\phi(\beta) W_t = \theta(\beta) a_t \dots \dots \dots (2-34)$$

Which includes (p+q+2) unknown parameters $(\mu_1, \phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q, \sigma_a^2)$ which can be estimated from data.

The importance of ARMA models lies in the fact that a stationary time series may often be described by an ARMA model involving fewer parameters than a MA or AR model by itself that the ARMA models may achieve greater flexibility in fitting of actual time series.

If the process (y_t) is non-stationary in the sense explained above, then it can be converted to a stationary process (W_t) by taking the suitable difference using:

$$W_t = \nabla^d y_t, \dots\dots\dots(2-35)$$

where (∇^d) is the (d^{th}) difference of the process (y_t) . The stationary (W_t) series is then mathematically described by equation (2-40) and the model designated as (ARIMA) (p, d, q) , i.e., the (ARMA) (p, q) model is regarded as special case of the (ARIMA) model with $(d=0)$.

2.10 The Box-Jenkins Seasonal Model

Box and Jenkins (1970) have generalized the autoregressive integrated moving average ARIMA (p, d, q) model to deal with seasonality and define a general multiplicative seasonal model in the form (using the notation)

$$\phi_p(\beta)\Phi_P(\beta^s)W_t = \theta_q(\beta)\Theta_Q(\beta^s)a_t, \dots\dots\dots(2-36)$$

where β denotes the backward shift operator, $\phi_p, \Phi_P, \theta_q, \Theta_Q$ are polynomials of order p, P, q, Q respectively, and (a_t) is purely random process with mean zero and variance σ_a^2 . The values W_t are formed from the original series X_t by differencing to remove both trend and seasonality by :

$$W_t = \nabla^d \nabla_{12}^D X_t \dots\dots\dots(2-37)$$

where for example if $d=D=1$ then

$$\begin{aligned} W_t &= \nabla \nabla_{12} X_t = \nabla_{12} X_t - \nabla_{12} X_{t-1} \\ &= (X_t - X_{t-12}) - (X_{t-1} - X_{t-13}) \end{aligned}$$

$$= X_t - X_{t-1} - X_{t-12} + X_{t-13}, \quad \dots\dots\dots(2-38)$$

where d is the degree of simple differencing (for trend removal)

D is the degree of seasonal differencing (for seasonality removal)

Also in equation (2-37) note that the shift operator β^{12} is such that

$$\beta^{12}W_t = W_{t-12} \quad \dots\dots\dots(2-39)$$

2.11 ARMA Stochastic Model Building

The general method of forecasting does not assume any particular pattern for the historical data of the series .It uses an iterative approach of identifying a possible useful model from a general class of models. The chosen model is then checked against the historical data to see if it accurately describes the series. The model is appropriate if the residuals between the forecasting model and the historical data points are small , randomly, distributed, and independent. If the specified model is not satisfactory, the process is repeated until a satisfactory model is found.

Before forecasting future time series, the following four stages of model building are completed .

1- Model Identification

Identification means the use of the data, and of any information on how the series was generated, to suggest a subclass of parsimonious (the smallest possible number of parameters) models worthily being entertained.

Identification of a model includes the use of the autocorrelation function (ACF) and the partial autocorrelation function (PACF). For stationary series, the correlogram can be used to indicate if a $AR(p)$, $MA(q)$, or $ARMA(p, q)$ is appropriate and also to indicate the order of such a process. The correlogram is a plot between autocorrelation function (ACF) or partial autocorrelation function (PACF) and lags (k) .Whereas the (ACF) of an autoregressive process of order (p) tails off, its (PACF) has a

cut- after lag (p) (only the pth lag is significantly different from zero or lies outside the significant limits $(\pm 2/\sqrt{N})$), where N is the number of observations. Conversely, the (ACF) of moving average process of order (q) has a cut- after lag (q), while its (PACF) tails off (only the qth lag is significantly different from zero or lies outside the significant limits $(\pm 2/\sqrt{N})$). If both the (ACF) and (PACF) tail off, a mixed autoregressive moving average model is suggested.

2-Initial Estimation of Model Parameters

This method is used to obtain an initial estimate of the parameters of ARMA (p,q) process. It is based on the first estimated autocorrelations as follows (Box and Jenkins, 1976):

AR(1) Process

$$\hat{\phi}_1 = r_1 \quad (-1 < \phi_1 < 1) \quad \dots\dots\dots(2-40)$$

where r_1 =lag one estimated autocorrelation coefficient.

AR(2) Process

$$\hat{\phi}_1 = \frac{r_1(1-r_2)}{1-r_1^2}, \quad \dots\dots\dots(2-41)$$

$$\hat{\phi}_2 = \frac{r_2 - r_1^2}{1 - r_1^2}, \quad (-1 < \phi_2 < 1, -1 < \phi_2 + \phi_1 < 1, -1 < \phi_2 - \phi_1 < 1) \quad \dots\dots(2-42)$$

where r_2 =lag two estimated autocorrelation coefficient.

AR(3) Process

$$\hat{\phi}_3 = \frac{r_3 - a_1 r_2 - a_2 r_1}{1 - a_1 r_1 - a_2 r_2}, \quad \dots\dots\dots(2-43)$$

$$\hat{\phi}_2 = a_2 - a_1 \hat{\phi}_3, \quad \dots\dots\dots(2-44)$$

$$\hat{\phi}_1 = a_1 - a_2 \hat{\phi}_3, \quad \dots\dots\dots(2-45)$$

$$a_1 = \frac{r_1(1-r_2)}{1-r_1^2}, \quad \dots\dots\dots(2-46)$$

$$a_2 = \frac{r_2 - r_1^2}{1-r_1^2}, \quad \dots\dots\dots(2-47)$$

where r_3 = lag three estimated autocorrelation coefficient.

MA(1) Process

$$\hat{\theta}_1 + \frac{\hat{\theta}_1}{r_1} + 1 = 0 \quad (-1 < \theta_1 < 1) \quad \dots\dots\dots(2-48)$$

MA(2) Process

$$r_1 = \frac{-\hat{\theta}_1(1-\hat{\theta}_2)}{1+\hat{\theta}_1^2+\hat{\theta}_2^2}, \quad \dots\dots\dots(2-49)$$

$$r_2 = \frac{-\hat{\theta}_2}{1+\hat{\theta}_1^2+\hat{\theta}_2^2}, \quad \dots\dots\dots(2-50)$$

$$(-1 < \theta_2 < 1, -1 < \theta_2 + \theta_1 < 1, -1 < \theta_2 - \theta_1 < 1)$$

ARMA(1,1)Process

$$r_1 = \frac{(1-\hat{\phi}_1\hat{\theta}_1)(\hat{\phi}_1-\hat{\theta}_1)}{1+\hat{\theta}_1^2-2\hat{\phi}_1\hat{\theta}_1}, \quad \dots\dots\dots(2-51)$$

$$r_2 = \hat{\phi}_1 r_1 \quad \dots\dots\dots(2-52)$$

3-Final Estimation of Model Parameters

The identification process leads to a tentative formulation for the model and needs to obtain an efficient estimate of the parameters (Box and Jenkins, 1976). The conditional and unconditional sum of squares of independent stochastic components (a_t) are used in iterative procedures in the exact estimation of model parameter, the minimum sum of squares determines the maximum likelihood estimate of model parameters.

1. The conditional likelihood for an ARIMA process:

Supposing that the $N=n+d$ original observations Z form a time series and is denoted by $Z_{d+1}, \dots, Z_0, Z_1, \dots, Z_n$ it is assumed to be generated by ARIMA(p, d, q) model. From these observations it can generate a series W of $n = N-d$ differences $W_1, W_2, \dots, W_n$, where $W_t = \nabla^d Z_t$. Thus, the general problem of fitting the parameters ϕ and θ of the ARIMA model is equivalent to fitting to the W 's, the stationary ARMA(p, q) model which may be written

$$a_t = \tilde{W}_t - \phi_1 \tilde{W}_{t-1} - \phi_2 \tilde{W}_{t-2} - \dots - \phi_p \tilde{W}_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}, \quad \dots (2-53)$$

$$\text{where } W_t = \nabla^d Z_t, \quad \tilde{W}_t = W_t - \mu, \quad \mu = \frac{1}{n} \sum_{t=1}^n W_t \quad \text{and} \quad E(W_t) = \mu$$

For $d > 0$, it would often be appropriate to assume the $\mu = 0$. Suppose that the p values W^* of the W 's and the q values a^* of the a 's prior to the commencement of the W series were given. Then the values of $a_1, a_2, \dots, a_n$, conditional on this choice, could be calculated in turn from (2-53)

Thus, for any given choice of parameters (ϕ, θ) and of starting values (W_*, a_*) a successive set of value $a_t(\phi, \theta | W_*, a_*, W)$, $t=1, 2, \dots, n$ can be calculated.

Given a particular set of data W , the log likelihood associated with parameter values (ϕ, θ, σ_a) , conditional on the choice of (W_*, a_*) would then be

$$l_*(\phi, \theta, \sigma_a) = -n \ln \sigma_a - \frac{S_*(\phi, \theta)}{2\sigma_a^2}, \quad \dots (2-54)$$

where

$$S_*(\phi, \theta) = \sum_{t=1}^n a_t^2(\phi, \theta | W_*, a_*, W) \quad \dots (2-55)$$

In the above, star subscripts, on the likelihood and sum of squares functions, are used to emphasize that they are conditional on the choice of the starting values.

To start up the process ,initial values of (ϕ, θ) are used, a_0 is set equals to zero and a_t is calculated for $(t=1,2,3,\dots,n)$. Adding (0.1) to the values $(\phi \& \theta)$ of the last iteration, the process is repeated with the next iteration until getting the minimum sum of squares $S_*(\phi, \theta)$ (Box & Jenkins, 1976).

2- The unconditional likelihood and sum of squares function:

The log likelihood of an ARIMA model is given by(Box & Jenkins 1976).

$$l(\phi, \theta, \sigma_a) = f(\phi, \theta) - n \ln \sigma_a - \frac{S(\phi, \theta)}{2\sigma_a^2} \quad \dots\dots\dots (2-56)$$

Where $f(\phi, \theta)$ is a function of ϕ and θ . The unconditional sum of squares function is given by:

$$S(\phi, \theta) = \sum_{t=-\infty}^n [a_t | \phi, \theta, W]^2 \quad \dots\dots\dots (2-57)$$

where $n = N - d$

The calculation of the unconditional sum of squares of ARIMA process was explained by Box and Jenkins (1976) who showed that the difference between the conditional and unconditional sum of squares is usually very small. Therefore, we use the first method in this study for the ARMA models and the unconditional sum of squares method for the Box-Jenkins seasonal methods. The computations of conditional sum of squared errors for (Ca) values at AL-Hindya station are shown in table (4.28), in chapter 4 as an example. The same procedure is applied for all the rest of the stochastic series. The unconditional method is illustrated in tables (5.1)&(5.2) in chapter 5.

4- Model Diagnostic Checking

After identifying and estimation of model parameters, Akaike information criteria (AIC) test has been used to select the best ARMA model out of the various models. AIC (p,q) is used to test the parsimonious model (the best fitted model), defined as

$$AIC(p,q) = N \ln(\sigma_a^2) + 2(p+q) \quad \dots\dots\dots(2-58)$$

where σ_a^2 is the maximum likelihood estimate of the residual variance.

$$\sigma_a^2 = \frac{1}{N-p-q} * \sum_{t=1}^N a_t^2 \quad \dots\dots\dots(2-59)$$

The most parsimonious model is the one that give minimum AIC value (Jayawardena and Lai, 1989), where N is the number of observations.

The diagnostic checking are then applied to test the adequacy of the fitted model. For checking the adequacy of the model fitted to data series, the following statistical tests may be used In this study the Q- statistics and residual autocorrelation function (RACF) tests have been used for testing the independency of the residuals.

(i)Port Manteau Lack of Fit Test

It is a test of the residual independency and uses the Q-statistic

defined as $Q = N \sum_{k=1}^M r_k^2(a_t) \quad \dots\dots\dots(2-60)$

where $r_k(a_t)$ is the autocorrelation coefficient of the residual (a_t) at lag k.

M is the maximum lag considered (about N/5).

If the a_t is independent, then the calculated Q, which is approximately chi-squared distributed with (M-p-q) degree of freedom, should be less than χ^2 (M-p-q) degree of freedom (Jayawardena and Lai, 1989).

(ii) Residual Autocorrelation Function (RACF) Test

The test for whiteness of the residuals is carried out to determine whether the residuals a_t are white noise, i.e., ARIMA (0, d, 0), or not. An appropriate procedure is to check the residual, to be sure they are random (white noise series) by checking the autocorrelations of residuals to be sure they are not significantly different from zero or within 95% confidence limits (Mahmood, 2000).

2.12 Forecasting

An observation $Z_{t+\ell}$ may be expressed directly in terms of difference equation by

$$Z_{t+\ell} = \phi_1 Z_{t+\ell-1} + \dots + \phi_p Z_{t+\ell-p} + a_{t+\ell} - \theta_1 a_{t+\ell-1} - \dots - \theta_q a_{t+\ell-q} \quad \dots\dots\dots(2-61)$$

The conditional expectation of $Z_{t+\ell}$ gives knowledge of all Z 's to time t , $E_t[Z_{t+\ell}]$. Then

$$\hat{Z}_t(\ell) = \psi_\ell a_t + \psi_{\ell+1} a_{t-1} + \dots = E_t[Z_{t+\ell}] \quad \dots\dots\dots(2-62)$$

Thus, the minimum mean square error forecast at origin t , for lead time ℓ , is the conditional expectation of $Z_{t+\ell}$ at time t . When $\hat{Z}_t(\ell)$ is as a function of ℓ for fixed t , it will be called the forecast function for origin t .

Taking the conditional expectation at time t and for $\ell > 0$ forecasts for difference equation can be obtained as:

$$[Z_{t+\ell}] = \hat{Z}_t(\ell) = \phi_1 [Z_{t+\ell-1}] + \dots + \phi_p [Z_{t+\ell-p}] - \theta_1 [a_{t+\ell-1}] - \dots - \theta_q [a_{t+\ell-q}] + [a_{t+\ell}] \quad \dots\dots\dots(2-63)$$

To calculate the conditional expectations which occur in the equation above, it is noted that if j is nonnegative integer

$$[Z_{t-j}] = E_t[Z_{t-j}] = Z_{t-j} \quad j = 0, 1, 2, \dots$$

$$[Z_{t+j}] = E_t[Z_{t+j}] = \hat{Z}_t(j) \quad j = 1, 2, \dots$$

$$[a_{t-j}] = E_t[a_{t-j}] = a_{t-j} = Z_{t-j} - \hat{Z}_{t-j-1}(1) \quad j = 0, 1, 2, \dots$$

$$[a_{t+j}] = E_t[a_{t+j}] = 0 \quad j=1, 2, \dots$$

Therefore, to obtain forecast $\hat{Z}_t(\ell)$, are write down

To obtain probability limits for these forecasts at lead time ℓ and also to allow new forecasts to be calculated by a process of updating the old, an approximate $1-\varepsilon$ probability limits $Z_{t+\ell}(-)$ and $Z_{t+\ell}(+)$ for $Z_{t+\ell}$ will be given by (Box-Jenkins, 1976)

$$Z_{t+\ell}(\pm) = \hat{Z}_t(\ell) \pm u_{\varepsilon/2} \left\{ 1 + \sum_{j=1}^{\ell-1} \psi_j^2 \right\}^{1/2} \sigma_a, \quad \dots\dots\dots(2-64)$$

where t is the origin time .

ℓ is the lead time .

$Z_{t+\ell}(\pm)$ is the upper and lower probability limits.

$\hat{Z}_t(\ell)$ is the forecast series at origin time t and lead time ℓ .

σ_a is the residual standard deviation .An estimate S_a^2 may be obtained from the time series and may be substituted by σ_a^2 .

$u_{\varepsilon/2}$ is the deviate exceeded by a proportion $u_{\varepsilon/2} = 0.68$, or 1.96

where the probability limits of the future value lies in interval 50 % or 95% respectively.

The ψ 's weights may be computed from the following equation

$$\left. \begin{aligned} \psi_1 &= \phi_1 - \theta_1 \\ \psi_2 &= \beta_1 \psi_1 + \phi_2 - \theta_1 \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \\ \psi_j &= \phi_1 \psi_{j-1} + \phi_2 \psi_{j-2} + \dots + \phi_p \psi_{j-p} - \theta_j \end{aligned} \right\} \quad \dots\dots(2-65)$$

For seasonal models , suppose there is moving average operator in the form

$$(1 - \theta\beta)(1 - \Theta\beta^{12}) = (\nabla + \lambda\beta)(\nabla_{12} + \Lambda\beta^{12}), \quad \dots\dots\dots(2-66)$$

where $\lambda = 1 - \theta$, $\Lambda = 1 - \Theta$ and $\nabla_{12} = 1 - \beta^{12}$. Hence the model

$$\nabla \nabla_{12} Z_t = (\nabla + \theta \beta)(\nabla_{12} + \Theta \beta^{12}) a_t \dots\dots\dots(2-66)$$

of order $(0,1,1) \times (0,1,1)_{12}$ may be written

$$\nabla \nabla_{12} Z_t = (\nabla + \lambda \beta)(\nabla_{12} + \Lambda \beta^{12}) a_t \dots\dots\dots(2-67)$$

On summing we obtain

$$Z_t = \lambda S a_{t-1} + \Lambda S_{12} a_{t-12} + \lambda \Lambda S S_{12} a_{t-13} + a_t \dots\dots\dots(2-68)$$

where $S_{12} = 1 + \beta^{12} + \beta^{24} \dots\dots$. Thus the ψ weights for this process are

$$\psi_1 = \psi_2 = \dots = \psi_{11} = \lambda \quad \quad \psi_{12} = \lambda + \Lambda \quad \quad \dots\dots\dots(2-69)$$

$$\psi_{13} = \psi_{14} = \dots = \psi_{23} = \lambda(1 + \Lambda) \quad \quad \psi_{24} = \lambda(1 + \Lambda) + \Lambda \quad \quad \dots\dots\dots(2-70)$$

$$\psi_{25} = \psi_{26} = \dots = \psi_{35} = \lambda(1 + 2\Lambda) \quad \quad \psi_{36} = \lambda(1 + 2\Lambda) + \Lambda \quad \quad \dots\dots\dots(2-71)$$

and so on .

Writing ψ_j as $\psi_{r,m}$, where $r=1,2,3,\dots$ and $m=1,2,3,\dots,12$ refer respectively to years and months, then

$$\psi_{r,m} = \lambda(1 + r \Lambda) + \delta \Lambda \quad \quad \dots\dots\dots(2-72)$$

$$\text{where } \delta = \begin{cases} 1, & \text{when } m = 12 \\ 0, & \text{when } m \neq 12 \end{cases}$$

In order to evaluate the accuracy of the forecasting and to compare the forecasting ability of the two modeling procedures, the mean absolute relative error (MARE) was used (Kurunc, et. al., 2005).

$$\text{MARE \%} = \frac{1}{N} \sum_{t=1}^N \left| \frac{X_t - Z_t}{X_t} \right| * 100, \quad \dots\dots\dots(2-73)$$

where X_t is the observed data at time t , Z_t is the forecasted data at time t and N is the number of observation. The best model is the one that gives minimum MARE% value.

2.13 Synthetic Data Generation

In hydrology, annual stream flow is often assumed first order autoregressive series known as Markov model which may be written as (McMahon and Mein, 1986).

$$X_{i+1} = \bar{X} + r_1(X_i - \bar{X}) + t_i s \sqrt{1 - r_1^2} \quad , \quad \dots\dots\dots(2-74)$$

where X_{i+1} and X_i = annual stream flow for $(i+1)^{\text{th}}$ and i^{th} years .

$\bar{X}$ = mean annual historical stream flow .

s = standard deviation of annual stream flow.

r_1 = lag one autocorrelation coefficient of annual stream flow and ,

t_i = normal random variate with a mean of zero and a variance of unity.

To initialize the model operation X_1 is set equals to $\bar{X}$. Hence, the model is used to generate very long synthetic record, for example 10000 years long, which is then used in capacity- yield analysis of a proposed reservoir. However, since in this study we are concerned with forecasting rather than data generation, the forecasting method developed by Box and Jenkins (1976) which is described in section (2.12) is adopted to forecast future values of water quality parameters .

Chapter Three

Literature Review

3.1 Some Studies on Time Series from other Countries

Iturbe et al., (1971) used Markov stochastic models for the modeling of monthly flow of the Arroyo Seco river near Splended, California, Kootenai river and Wennatchee river at Plain, Washington and tested the adequacy of these models with cyclic components of data. Amplitude and phase of the annual stream flow cycle were studied by using demodulation techniques. A case of study performed with several markovian type model showed that the seasonal variation in standard deviation should be incorporated in any model used for design. It was also shown that the type so distribution used for the random component did not have a great influence on stream flows generated for storage design.

Hirsch (1979) considered six different single – site monthly generators of synthetic of hydrology which may be used in the evaluating the consequence of water supply management decisions and the compared between them to find the best one. The six synthetic stream flow generators (SSG's) are applied to the case of the Potomac River at point of Rocks, Maryland where an 81-year historical record was available. The synthetics stream flow record produced by each SSG is 1001 years long. The SSG model was identified by combinations of the following five acronyms: LN, log normal, LT, log transform NT, normalizing transform; AR, first – order autoregressive, and Autoregressive – moving average model ARMA (1,1). The results suggest that the log normal LN – ARMA (1,1) was the best of the six generators.

Miller et. al., (1981) used an autoregressive model of the third order AR(3) for modeling of daily river flow by using not only the past flow record but also rainfall. A univariate time series of daily river flows was compared with a regression model that incorporates precipitation in

nonlinear fashion. The analyzed data was the average daily flow of Menomonoc river during 1975-1976, and the rainfall data of the same area and at which the flow data were measured. Even though the univariate time series model was constructed by using regression techniques it was a transfer function plus noise model. the logarithms transformation was used. Autoregressive model of third order was found to be the best univariate model for the data simulation.

Lohani and Wang (1987) used AR(1) model to represent 12 monthly water quality parameters taken at three stations on Chung Kang river in Taiwan. Five years of data (1976 – 1980) were used for basic analysis and the data in the sixth year (1981) were used for the comparison of the results forecasted from the time series model. Different type of transformation were used such as (square root, logarithmic, and power transformation). Autocorrelation function (acf) and partial autocorrelation function (pacf) were used for model identification, maximum likelihood estimate (MLE) was used for the model parameter estimation diagnostic checks were also applied. Akaik information criteria (AIC) was used to find the suitable model (which have minimum AIC). It was found that the autoregressive model of order one the best fit of data with minimum (AIC).

Gupta and Chauhan (1985) used second order autoregressive model AR(2) to model the stochastic structure of weekly irrigation water requirement (IWR) of paddy crop grown in the Nainitd Tarai region of India. The IWR values were obtained by computing the crop evapotranspiration and effective rainfall during the ten growing periods (1971 -1980) of paddy crops. The turning point test and Kendall's rank correlation test are applied for detecting the trend. The periodicity was detected by using correlogram technique which was analyzed by Fourier series method. Harmonic analysis was done for identifying the number of significant harmonies. The series was tested for stationary and the

dependent part of the stochastic component was found to be well expressed by second order autoregressive model AR(2).

Jayawardena and Lai (1989) AR(1), AR(2) and a higher order model ARMA (5,8) were used for the modeling of 21 years (1965 -1985) of mean monthly water quality data in the Guangzhou reach of the Pearl River in the southern China. Eighteen years of data was used for model development while the model performance was compared with the data for the remaining three years. The turning point test and Kendall's rank correlation test were used for the detection of the trend component and found that it can be removed by fitting a regression equation. The autocorrelation and spectral analysis were used for detection the periodicity which can be represented by Fourier series. The stochastic component of the series is obtained by subtracting the periodic mean and it was standardized by dividing by the periodic standard deviations. The comparison between observed and generated series were found to be satisfactory.

Mattikalli (1996) presented time series of historical water quality data as indexed by total oxidized nitrogen (TON) from 1976 to 1990 in a river catchments in the south – eastern part of England. The trend value can be identified by fitting least squares regression lines to the TON time series data paired with the observation dates. Seasonal indices can be estimated by using an index that compares the seasonal mean with grand mean values or by the ratio to moving average method. Standard time series analyses (autocorrelation and partial autocorrelation functions) were not applicable to irregularly sampled water quality data sets, and there was need for an established method capable of handling such data sets. An approach based on multiple regression analysis had therefore developed to quantify both seasonal and annual trends in the TON water quality data.

The results were demonstrated that this approach was useful for the analysis of historical water quality data sets.

Cun and Vilagines (1997) used graphical and statistical time series techniques to analyze the trends and specified time changes of annual average values of stream flow and concentrations of ammonium , nitrates, chlorides and dissolved oxygen observed at water control station on the seine near Paris over a period of 93 years from 1901 to 1993. The analysis included three steps ;the first step was graphical and numerical studies; the second step was graphical analysis of trend and dispersion by using Cumul curve and Box and Whisker graphs; and third step was statistical data analysis with non parametric test like the Kruskall – Wallis (to test the groups of median homogeneity, Kendall's tau (to test trends); Bartlett (to test groups of dispersion homogeneity) and Mann – Whitney tests (to identity sudden changes in dates) , were used to confirm the type and dates of changes and trends. Kurskall – Wallis test for these water quality parameters showed that the 1931-1960 period presented no significant evolution in the median values but the last 30 years presented marked evolution every 10 years. For Bartlett test, for the dispersion of the result showed no significant differences between the analyzed series standard deviation whatever the tested parameter, except for the 1911- 1960 period of ammonium data. For Mann- Whitney test for all parameters, the heterogeneity showed by the Kurkall – Wallis test for the 1921-1930 and 1980-1993 periods respectively but did not seem to be confirmed by the Mann- Whitney test. The difference between medians is not statistically significant and can not be considered as a sudden increase. For Kendall's tau test of trends, the chloride data presented an increase tendency for all the test periods from 1921 to 1993, nitrate and ammonium average value presented the same increase period between 1961 to 1970 but the nitrate parameter still presented an increasing trend in the 1971 -1980 period of

time and, earlier between 1901 and 1910 the dissolved oxygen parameter presented no monotonic trend period.

Montanari et. al., (1997) considered the fractionally differenced autoregressive integrated moving average (FARIMA) model where the difference order(d) is less than 1(between 0) and 1). For preliminary identification of (FARIMA) model, hurst parameter(h) was found at the first step by using R/S method and the aggregated variance methods to find (d) where ($d = h-0.5$), finding the underlying series by some methods and then estimating the appropriate ARMA of the underlying series. The analysis of daily inflows to lake Maggiore showed that the FARIMA approach provided better representation of the observed data than traditional ARIMA models.

Ahmad et. al., (2001) three different approaches of (i) multiplicative autoregressive integrated moving average (ARIMA) model, (ii) deseasonalized model using Fourier series technique and (iii) set of Thomas –Fiering model for various months have been used to model 10 years (1981-1990) of mean monthly water quality data consisting of six water quality parameters (temperature, pH, chlorides, electrical conductivity, DO and BOD) at ten typical locations along the river Ganges in India. The Box and Whisker plots and Kendall's tau test were used to identity the trends during the study period. For detection the possible intervention in the data, the time series plots and Cusum charts were used. The multiplicative ARIMA model having both nonseasonal and seasonal components were, in general identified as appropriate models. In the deseasonolized modeling approach the lower order ARIMA models were found appropriate for the stochastic components. The set of Thomas – Fiering models were formed for each month for all water quality parameter these models were used to forecast the future values. The error estimates of forecasts – which were computed by root

mean square error (RMSE) and mean absolute percentage error (MAPE) – from the three approaches were compared to identify the most suitable approach for the reliable forecast. The deseasonalized modeling approach was recommended for forecasting of water quality parameters.

Lhmann and Rode (2001) studied the analysis weekly data sampled from the river Elbe at Magdeburg in Germany between 1984 and 1996 to investigate the changes in metabolism and water quality in the river after German reunification in 1990. Autoregressive component models and autoregressive integrated moving average (ARIMA) models were used to reveal the improvement of water quality due to reduction of waste water emissions since 1990 and to determine the long – term and second behavior of important water quality variables. For data modeling the trend component was detected by curve fitting using polynomials and growth curves such as a logistic curve and a Gompertz curve and removed by simple difference. The seasonal variations were removed by seasonal differencing . ARIMA model was identified by (acf) and (pacf) plots of simply and seasonal differenced series. The maximum likelihood parameter estimation procedure was used to estimate the ARIMA model and the diagnostic check was done. Also a new procedure for testing the significance of sample correlation coefficient was used to show the cross – correlation between variables. The cross- correlation analysis was applied to hydrophysical, biological and chemical water quality variables. Univariate modeling of important water quality variables with component, ARIMA models and cross – correlation analysis between these variables show that the trends and seasonal behavior of almost water quality indicator differs significantly in the periods before and after 1990.

Kahya and Kalayci (2004) studied the trends computed for the 31 years period of monthly stream flows obtained from 26 basins over Turkey. Four non – parametric trend test (the Sen's T, the Spearman's Rho,

the Mann – Kendall , and the seasonal Kendall which are known as appropriate tools in detecting linear trends of hydrological time series) were adapted in this study. Moreover, the Van Belle and Hughes basin wide trend test were included in the analysis for the same purpose. Homogeneity of trends in monthly stream flows was also tested using a procedure developed by Van Belle and Hughes. As a result, basins located in western Turkey exhibited downward trend, significant at the 0.05 or lower level, whereas basins located in eastern Turkey showed no trend. In most cases, the first four tests provided the same conclusion about trend existence. Some of basins located in southern Turkey exhibited a global trend, implying the homogeneity of trends in seasons and stations together based on the Van Belle and Hughes basin wide trend test.

Kurunc et al., (2005) evaluated the forecasting performance of two modeling methods ARIMA& Thomas–Fiering method for selected water quality constituents and stream flow of the Yesilirmak River at Durucasu monitoring station in Turkey. For this purpose 18 years (1984 – 2001) of mean monthly water data which include : temperature, pH, electric conductivity, sodium, potassium, calcium + magnesium, chlorides and sulfates and monthly stream flow were used to obtain the best model. 13 (1984- 1996) years were used for analysis and 5(1997-2001) years were used for the comparison of observed to generated data. For ARIMA model unit root test was used to detect the trend in the observed data. The autocorrelation function (acf) and partial autocorrelation function (pacf) were used to identity the model. The closeness of fits was checked by using Ljung – Box portmanteau Q (LB/ Q) statistics. Schwarz Bayesian criterion (SBC) was used to select the best ARIMA model among possible that fulfilled all the diagnostic checks for water quality parameter and stream- flow, the model with the minimum SBC value was selected as the best model. The results of forecast accuracy were measured by root mean

square error (RMSE) and mean absolute error (MAE) for two approaches. Thomas – Fiering model presented more reliable forecasting of water quality constituents and streamflow than ARIMA model.

Tosfaye and Meerschoert (2005) studied the modeling of the periodic hydrological time series in general and river flows in particular. The innovations algorithm was used to obtain estimates of certain model parameters. Asymptotic distribution for the innovations estimates and model parameter were used with general technique for identifying the periodic autoregressive moving average(PARMA) models. Monthly river flow data for Fraser River in British Colombia and the Salt River Arizona were used in a detailed simulation study to test the effectiveness of the innovation estimation procedures and asymptotic as a method for PARMA model identification, to illustrate residual modeling procedure and to prove the ability to generate realistic synthetic river flow.

The asymptotic distribution of the discrete Fourier transform of the innovation estimates and PARMA model parameters was developed to obtain a parsimonious model. The applicability of these methods was demonstrated by using weakly river flow data for the Fraser River.

3.2 Previous Application of Time Series Methods to Data from Iraq

AL-Shuhaili (1986) used singlesite AR(1), ARIMA(1,0,1) and multisite autoregressive model (Matalas model) for four Tigris river flow stations. These models were used for daily stream flow of Tigris river at Mosul, Fatha, Baghdad and Kut stations for the period from (1936- 1982). T-test was used for detecting of nonhomogeneity and Split- sample method was used for removing it (if any).The homogeneous series was normalized using a power transformation. The periodicity was detected by the correlogram technique and removed by harmonic analysis. The

ARIMA(1,0,1) was found to be the most attractive one and the comparison between generated and observed data was found to be satisfactory.

AL-Husseini (2000) used AR(1), MA(1) and ARMA(1,1) as univariate models and first order multivariate model (Matalas model) to fit stochastic component of eight years (1992 -1998) of mean monthly water quality parameter at AL-Hilla station. Nonhomogeneity in mean and standard deviation was detected by t-test and removed by using split sample method. Periodicity was detected by plotting the autocorrelation function (correlogram) and removed by using Fourier analysis Autocorrelation functions (acf) and partial autocorrelation functions (pacf) were used for model identifications. Univariate stochastic model proved to be good in forecasting the concentrations of all parameter except hardness which was represented by Matalas model. The comparison was found to be satisfactory.

Mahmood (2000) applied AR(1), AR(2) and ARIMA(1,1,1) models in analyzing monthly time series of water quality data which included ten parameter on the Euphrates river at kufa city for period (1991 -1998). Eight years were used for basic analysis and the other two years were used for the comparison of the results. Common log was used in the series normalization and autocorrelation function (ACF) and partial autocorrelation function (PACF) were used for model identification. The seasonal component was not existed in these types of observations. The comparison was found to be satisfactory. The related equations are given table (7.3) .

AL- Tikriti (2001) used singlesite AR(1) to model seven parameter of average weekly water quality data and discharges at two station on Euphrates river (AL-Hindiya and AL –Samawa station) for period (1984-1997). Ten years of data (1984 – 1993) were used for building the stochastic model and the last four years (1994 – 1997) were used for the

checking of model acceptability. Nonhomogeneity was removed by using split – sample method. The periodicity was detected by correlogram technique and removed by harmonic analysis. The single site autoregressive model with first order AR(1) model was found convenient for the parameter in the two stations while multisite model was not used because of the weak correlation between the two stations. The comparison of the statistical parameters of these records and with generated series was found to be acceptable. The related equations are given table (7.3) .

AL-Mousawi (2003) applied two stochastic singlesite autoregressive models with first order AR(1) and multisite model with first order AR(1) (matalas model) to model of monthly water quality data (raw water and water supply) of eight hydrochemical parameters with discharges of four stations (AL- Hindya , AL- Hilla , AL-Hssien &AL- Hashimiya) on AL-Hilla River for the period (1987-2001). Twelve years (1987-1998) of data were used for basic analysis and model building and the remaining three years from (1999- 2001) were used to cheek the models acceptability. Nonhomogeneity was detected by suitable statistical tests and removed by split – sample method . Power transformation was used to transform data to normal distribution. Periodicity was detected by correlogram technique and removed by harmonic analysis. Multisile model was used to study the cross correlations exist between the stations, this model was found to be adequate to describe the process. The singlesite first order autoregressive model AR(1) was found to be adequate for all station and the comparison of the statistical parameter of observed and generated series was found to be satisfactory. The related equations are given table (7.3) .

Chapter Four

Application of ARMA Model

4.1 ARMA Model

Before the building of ARMA model, historical annual series must be tested for jump and trend components by statistical tests and these components must be removed by using an appropriate method (split –sample method). Then, a suitable transformation is selected to convert the data to the normal distribution. The periodic component can be detected by plotting the autocorrelation coefficients of the normalized data against the lags and it can be removed by parametric method. After the removing of periodic component, the standardization is applied to the series, and the remaining series is called stochastic series which is used for the building of stochastic models.

The procedure used for data analysis is shown schematically in figure (4-1) and the procedure can be summarized by the following steps (AL-Suhaili,1986) :

1. Test the homogeneity by some statistical testes ,if the nonhomogeneity exists, it will be removed by split –sample method.
2. Use a suitable transformation like (Box-Cox, square root, or log transformation) to normalize the homogeneous data .
3. Plot the correlogram of monthly means and standard deviations of normalized data to detect the periodicity.
4. The removal of periodicity by applying harmonic analysis and determination the number of significant harmonic for monthly means and standard deviations of the normalized series.
5. Standardization of the remaining series by subtracting mean and dividing by standard deviation .
6. Initial estimation of model parameter by plotting the correlograms of the water quality parameter.

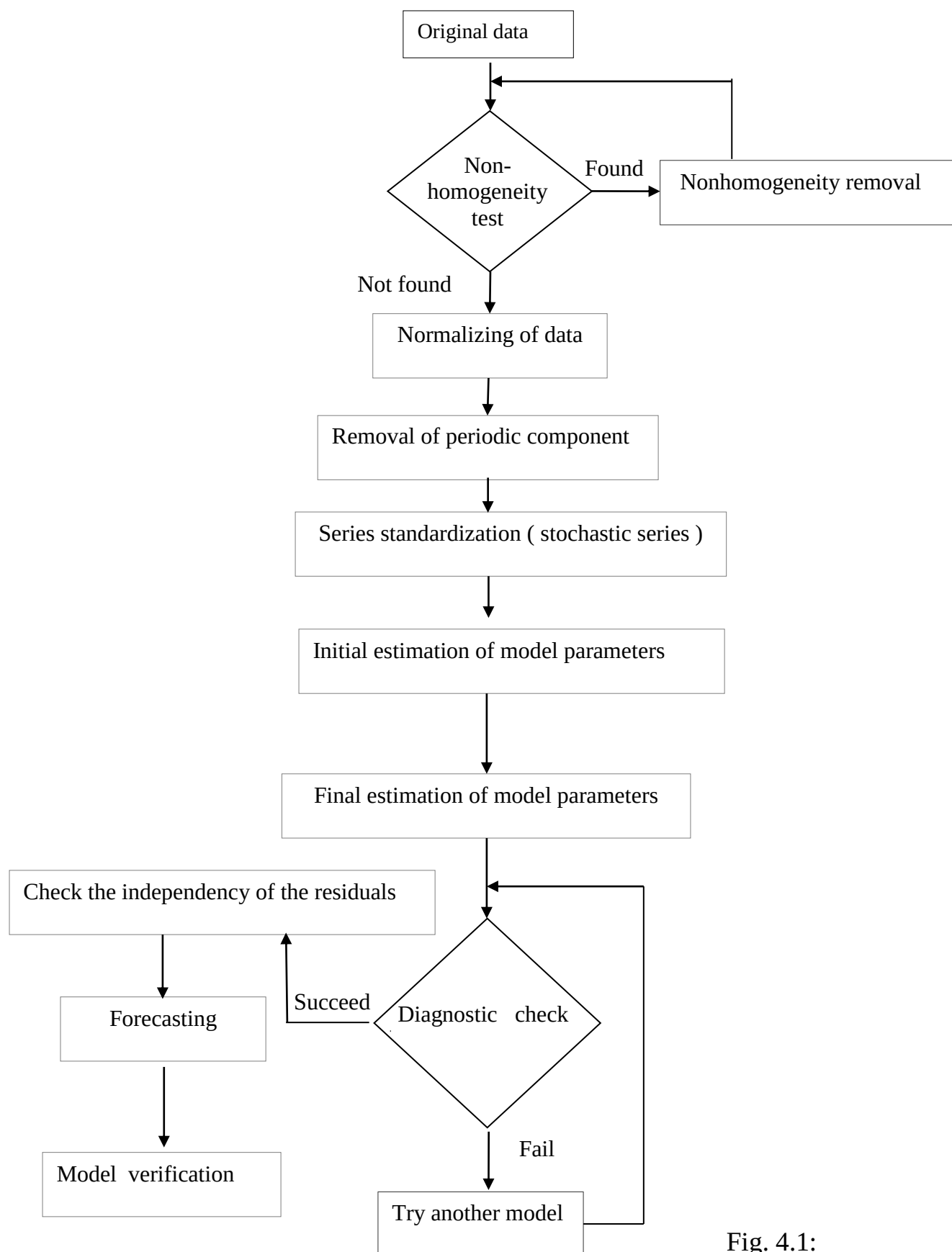


Fig. 4.1:
Steps of ARMA method

- 7.Exact estimation of model parameter by using the general least squares algorithm for the conditional model.
- 8.The diagnostic checks of model by one of some tests and testing the independency of the residuals.
- 9.Forecasting and verifying the forecasted models.

4.2 Test and Removal of Nonhomogeneity

Hydrologic and water quality time series are commonly nonstationary with trends and seasonality (Zou and Yu,1996). This section includes nonhomogeneity test of the mean and standard deviation ,and it's removed when the series is found to be non -homogeneous. Split –sample method described in section (2.4) is used to ascertain whether or not the differences between the means and standard deviations of two sub-samples are significantly different from zero at (97.5) percent confidence limit, therefore, each water quality parameter in the three stations is divided into two sub-samples of years ,the first is five years long (1992- 1996) and the second is five years long also (1997- 2001). Figures (4.2) and (4.3) show the annual mean and annual standard deviation respectively of AL-Hindya station while the remaining figures are given in the appendix (B)(figures from (B.1) to (B.4)) and the solid line with its equation on these figures represents the trend of annual means and annual standard deviations of the water quality parameters respectively. The test is applied using equations (A_1 , A_2 , A_3) and ($n_1=5$, $n_2=5$) (see appendix A). If the calculated value of t and f is more than the critical value of t and f respectively, then the trend is found. The results are summarized in tables (4.1) ,(4.2) & (4.3). For the water quality parameter in the three stations which show significant differences. The split-sample method is used for removing the nonhomogeneity. The annual mean ($\bar{X}_j$) and annual standard deviation (S_j) are obtained by applying equations from (2-4) to (2-6).

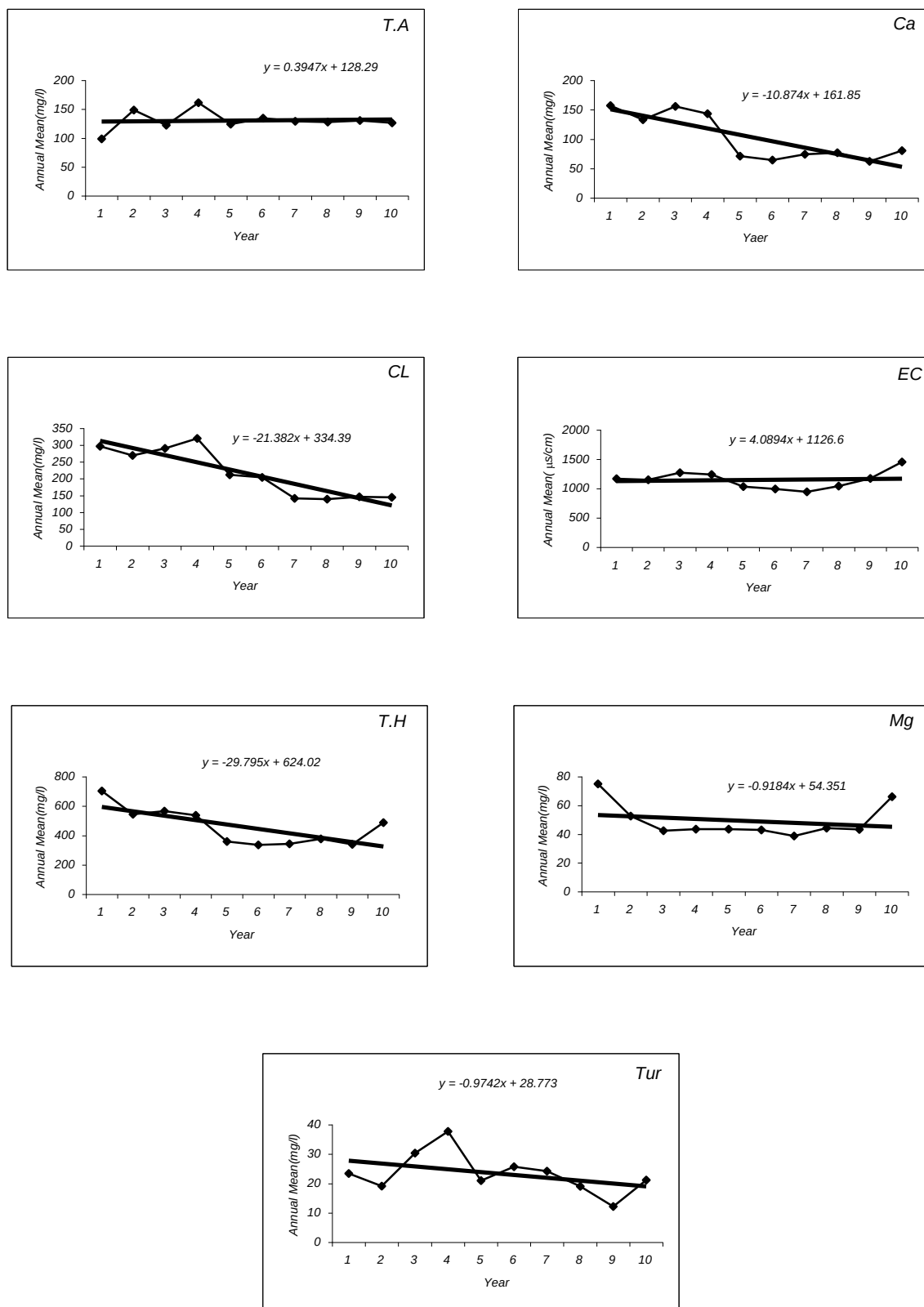


Fig. 4.2:Annual means of observed AL-Hindya data before removing variation in mean form 1992 to 2001 .

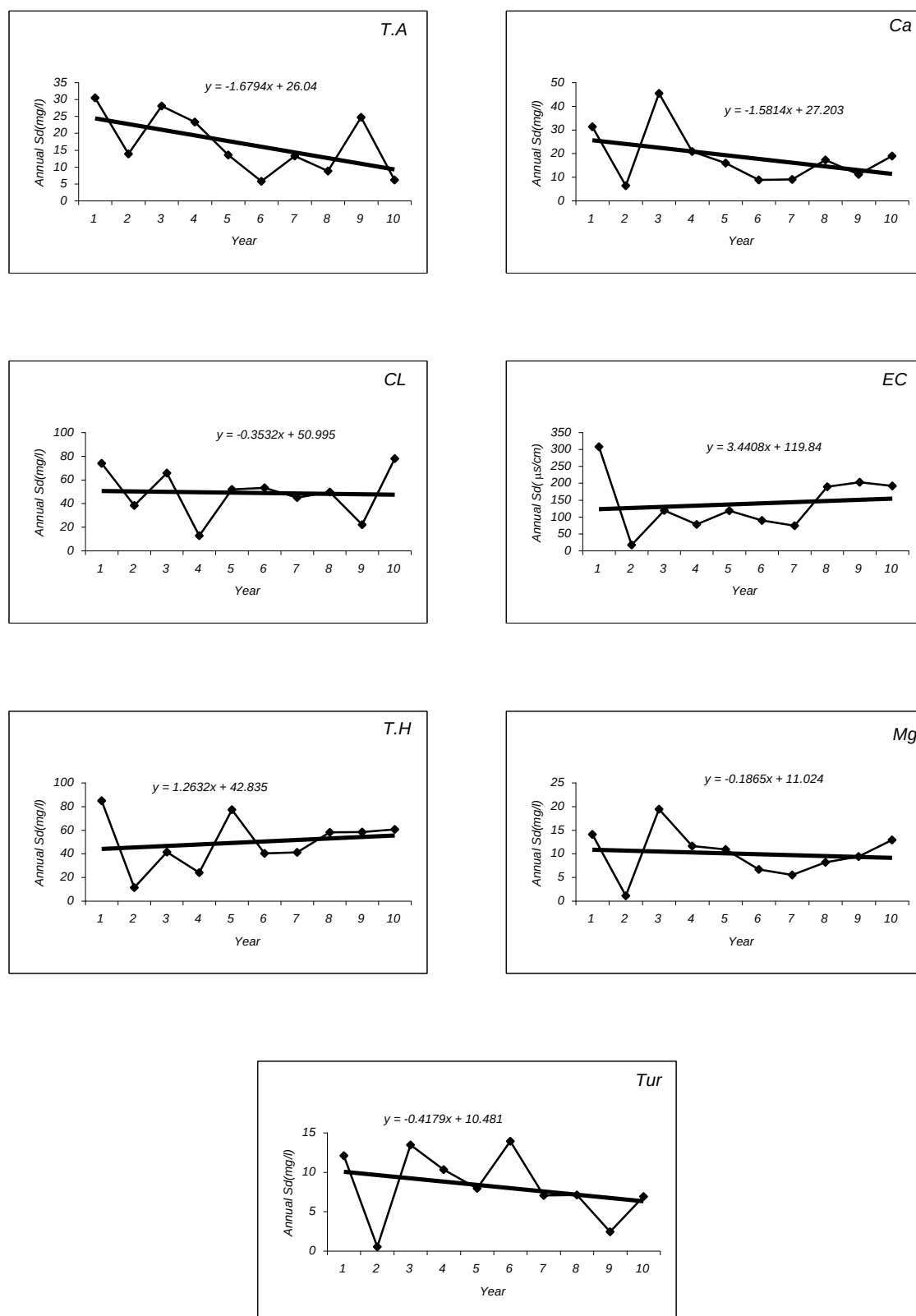


Fig. 4.3: Annual standard deviations of observed AL-Hindya data before removing variation in standard deviation form 1992 to 2001 .

Table 4.1: Result of nonhomogeneity test of AL-Hindya station where critical t and f –value (2.306) and (9.60) respectively.

Parameter	T.A	Ca	Cl	EC	Mg	T.H	Tur
t-calculated	0.120	3.718	5.525	0.666	0.550	2.662	0.950
Change in mean	None	Exist	Exist	None	None	Exist	None
f- calculated	5.281	7.489	1.174	3.042	1.840	2.582	1.602
Change in Sd	None	None	None	None	None	None	None

Table 4.2: Results of nonhomogeneity test of kufa station where critical t and f –value (2.306) and (9.60) respectively .

Parameter	T.A	Ca	Cl	EC	T.H	Mg	Tur
t-calculated	-0.009	2.449	-0.288	4.370	5.029	1.934	1.415
Change in mean	None	Exist	None	Exist	Exist	None	None
f- calculated	2.501	2.673	1.056	1.372	22.639	9.110	3.544
Change in Sd	None	None	None	None	Exist	None	None

Table 4.3: Result of nonhomogeneity test of Diwaniya station where critical t and f –value (2.306) and (9.60) respectively .

Parameter	T.A	Ca	Cl	EC	T.H	Mg	Tur
t-calculated	-0.289	1.481	0.938	1.938	3.329	5.213	0.424
Change in mean	None	None	None	None	Exist	Exist	None
f- calculated	2.926	1.563	1.185	1.103	1.806	1.081	1.974
Change in Sd	None	None	None	None	None	None	None

The regression coefficients of the second sub –samples are calculated and the result are summarized in table (4.4). The nonhomogeneity can be removed by using equation (2-4) for all parameters show nonhomogeneity in order to

obtain the first new sub-samples, so the new series is free from trend component .For **Ca** of AL-Hindya station as an example equation (2-4) becomes:

$$Y_{i,j} = \frac{X_{i,j} - 161.85 + 10.874i}{27.203 - 1.581i} * 13.874 + 71.85 \quad \dots\dots\dots(4-1)$$

Table 4.4: Regression coefficients before removing nonhomogeneity.

Station	Parameter	Reg . coff. of mean		Reg. coff. Sd	
		A	B	A ₁	B ₁
AL-Hindya	Ca	-10.874	161.847	-1.581	27.203
	Cl	-21.382	334.391	-0.354	50.995
	T.H	-29.795	624.021	1.263	42.835
Kufa	Ca	-4.788	136.981	-0.073	12.264
	EC	-29.608	1543.305	-11.294	182.07
	T.H	-43.779	699.188	-6.372	89.595
Diwanya	Mg	-2.666	75.686	0.095	8.410
	T.H	-20.850	597.231	-0.773	48.850

The test is now repeated by using ($Y_{i,j}$) series-which obtained from equation (2-4) -as the new first sub-sample, and the original second sub-sample. Figures (4.4) and (4.5) show the annul mean and annul standard deviation respectively of AL-Hindya station data after removing of nonhomogeneity and the other related figures are given in the appendix (B)(figures from B.5 to B.8). The results which are shown in table (4.5) ensure that the nonhomogeneity is removed from the data .

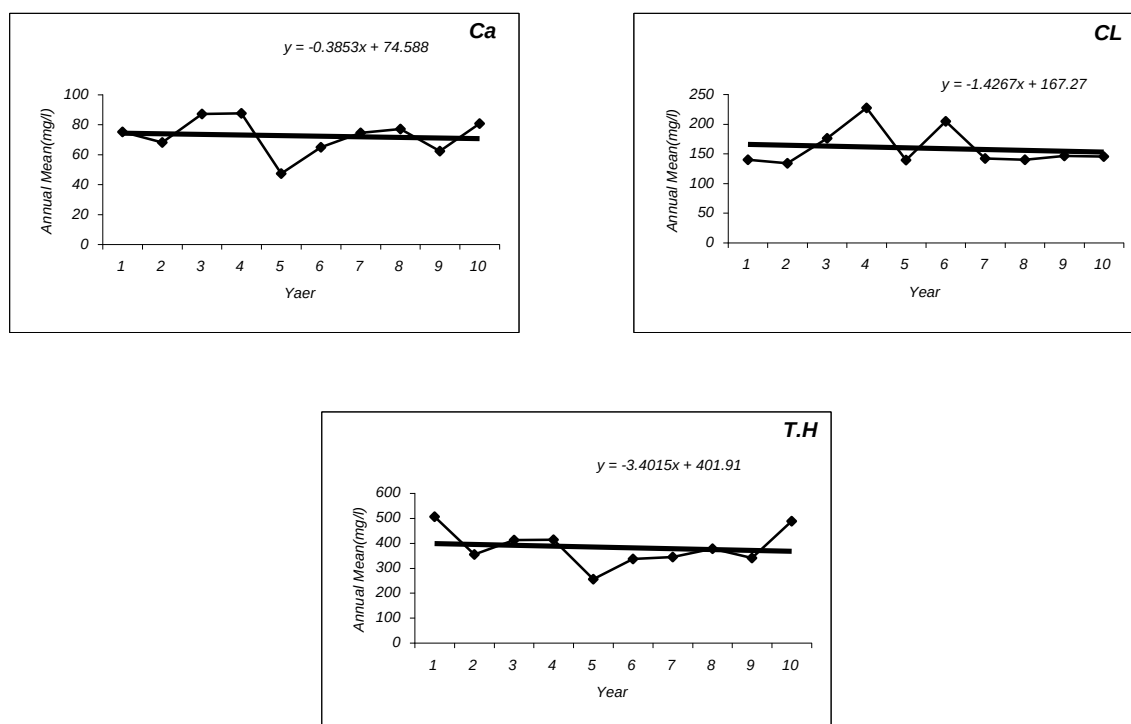


Fig. 4.4: Annual means of homogenous data of AL-Hindya station after removing variation in mean from 1992 to 2001.

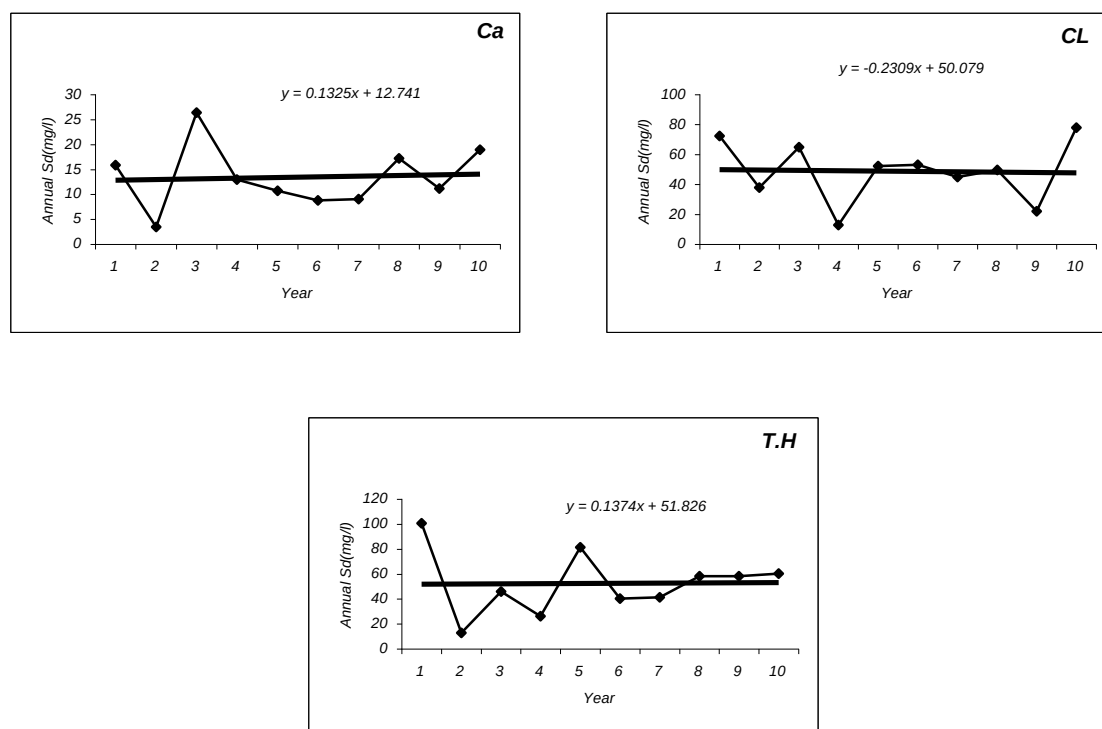


Fig.4.5 :Annual standard deviation of homogenous data of AL-Hindya station after removing variation in standard deviation from 1992 to 2001.

Table 4.5: Result of nonhomogeneity test after removal trend component where critical $t = 2.306$ and critical $f = 9.60$.

Station	AL-Hindya			Kufa			Diwaniya	
Parameter	Ca	Cl	T.H	Ca	EC	T.H	Mg	T.H
t- Calculated	0.151	0.356	0.218	0.073	0.460	-0.404	0.748	-0.016
Change in mean	None	None	None	None	None	None	None	None
f-Calculated	2.012	1.212	1.800	1.349	1.843	2.149	1.116	1.011
Change in Sd	None	None	None	None	None	None	None	None

4.3 Transforming Data to Normal Distribution

The time series observation of a given phenomenon required a certain type of transformation (Hipel et. al., 1977). Before transformation data to normal distribution , the coefficients of skewness (C_s) and Coefficients of kurtosis (C_k) must be determined. Tables (4.6), (4.7) and (4.8) show the values of the four moments of the homogeneous water quality data series of three stations . These tables indicate that the water quality parameters at Diwaniya station are closer to normal distribution from AL-Hindya and Kufa stations, because of the small (C_s) and (C_k) near to 3.

It is often better to transform the data to the normal distribution to utilize its simple properties, its familiarity to most engineers , and to obtain satisfactory fit to data, (AL –Suhaili, 1986).

The other reason for transforming the date includes stabilizing the variance and improving the normality assumption of the white noise series (AL–Husseini, 2000) .

Several transformation may be used to normalize the data but the most common and useful class of transforms for stabilizing the variance is known as the Box - Cox transformation (Chandra et. al, 1978). The transformation as

Table 4.6: The values of the first four moments of homogeneous series of AL-Hindya station.

	T.A	Ca	Cl	EC	T.H	Mg	Tur
Mean	130.420	72.462	159.404	1157.046	383.190	49.300	23.413
Sd	23.845	18.403	59.466	198.183	90.544	15.602	11.042
C_s	-0.160	0.474	0.351	0.577	0.407	0.767	1.187
C_k	3.695	3.034	2.450	3.379	4.918	4.998	4.220

Table 4.7: The values of the first four moments of homogeneous series of Kufa station .

	T.A	Ca	Cl	EC	T.H	Mg	Tur
Mean	176.892	98.163	174.133	1291.196	351.57	48.946	20.029
Sd	15.265	15.033	15.252	110.434	50.057	23.416	8.693
C_s	-0.802	0.384	-0.186	-0.393	0.633	1.537	1.120
C_k	4.172	2.773	3.270	3.705	3.492	5.016	3.810

Table 4.8: The values of the first four moments of homogeneous series of Diwaniya station .

	T.A	Ca	Cl	EC	T.H	Mg	Tur
Mean	128.29	90.958	133.025	1292.283	430.453	53.574	39.400
Sd	20.868	20.188	12.219	125.339	53.801	9.832	9.338
C_s	0.125	0.464	0.498	0.026	-0.209	-0.014	-0.113
C_k	4.821	2.871	2.586	2.436	2.625	2.390	2.887

follows :

$$\begin{bmatrix} F_{i,j} = (Y_{i,j}^\lambda - 1) / \lambda & \lambda \neq 0 \\ F_{i,j} = \log Y_{i,j} & \lambda = 0 \end{bmatrix}, \quad \dots\dots\dots(4-2)$$

where $F_{i,j}$ is the transformed series.

$Y_{i,j}$ is the variates of given series .

λ is the constant of transformation.

The relationships between λ and C_s is being some form of second degree polynomial as

$$\lambda = B_0 + B_1 C_s + B_2 C_s^2 \quad \dots\dots\dots(4-3)$$

The value of (λ) is found by choosing random eight values of (λ) between (-1) and (1) and computing the corresponding (C_s) values for the series after transforming it by equation (4-2) (AL-Suhaili, 1986). Then by fitting equation (4-3) to these eight points, the value of (λ) is found as equal to (B_0) corresponding to $C_s = 0$. Tables (4.9), (4.10) and (4.11) show the effect of (λ) values on the first four moment mean, Sd, C_s & C_k of each water quality parameter in the three stations. For normally distributed data the ($C_s = 0$) and ($C_k \approx 3$). Tables (4.12), (4.13) and (4.14) show the values of (λ) at which the series is normally distributed together with other parameters .

Square root transformations is also applied by using the following equation.

$$F_{i,j} = \sqrt{Y_{i,j}} \quad , \quad \dots\dots\dots(4-4)$$

where $F_{i,j}$ & $Y_{i,j}$ are defined above.

Table 4.9: Effect of (λ) values on the first four moments of each series in AL-Hindya station.

Parameter	λ	Mean	Sd	C_s	C_k
T.A	-0.8	1.224	4.528×10^{-3}	-1.849	8.187
	-0.6	1.575	1.173×10^{-2}	-1.637	7.313
	-0.4	2.141	3.000×10^{-2}	-1.430	6.534
	-0.2	3.104	7.720×10^{-2}	-1.229	5.853
	0.2	8.207	0.516	-0.846	4.772
	0.4	14.969	1.341	-0.665	4.367
	0.6	29.185	3.497	-0.492	4.044
	0.8	60.137	9.149	-0.326	3.797
Ca	-0.8	01.207	9.197×10^{-2}	-0.981	4.495
	-0.6	01.534	2.104×10^{-2}	-0.798	4.052
	-0.4	02.040	4.834×10^{-2}	-0.620	3.687
	-0.2	02.860	0.111	-0.447	3.399
	0.2	06.715	0.602	-0.118	3.037
	0.4	11.260	1.407	3.84×10^{-2}	2.953
	0.6	19.940	3.303	0.189	2.928
	0.8	37.014	7.782	0.334	2.957
Cl	-0.8	01.226	8.177×10^{-3}	-1.062	4.066
	-0.6	01.581	2.134×10^{-2}	-0.871	3.441
	-0.4	02.157	5.612×10^{-2}	-0.693	2.974
	-0.2	03.154	0.149	-0.528	2.637
	0.2	08.627	1.005	-0.221	2.270
	0.4	16.138	2.881	-0.076	2.212
	0.6	32.686	7.842	6.737×10^{-2}	2.226
	0.8	70.209	21.433	0.201	2.305

Table 4.9: Cont.

Paramete	λ	Mean	Sd	C_s	C_k
EC	-0.8	1.245	6.07×10^{-4}	-0.204	2.532
	-0.6	1.642	2.47×10^{-3}	-0.129	2.538
	-0.4	2.350	1.01×10^{-2}	-4.9×10^{-2}	2.562
	-0.2	3.776	4.12×10^{-2}	3.27×10^{-2}	2.606
	0.2	15.449	0.692	0.202	2.757
	0.4	39.365	2.841	0.291	2.868
	0.6	112.740	11.677	0.382	3.006
	0.8	350.880	48.085	0.476	3.171
T.H	-0.8	1.239	3.25×10^{-3}	-5.283	44.598
	-0.6	1.618	9.47×10^{-3}	-4.270	33.712
	-0.4	2.264	0.028	-3.318	24.546
	-0.2	3.467	8.58×10^{-2}	-2.471	17.447
	0.2	11.351	0832	-1.151	8.951
	0.4	24.305	2.648	-0.655	6.831
	0.6	57.06	8.518	-0.244	5.627
	0.8	143.846	27.659	0.103	5.054
Mg	-0.8	1.190	2.06	-3.392	20.815
	-0.6	1.497	0.041	-2.721	15.745
	-0.4	1.958	0.081	-2.097	11.730
	-0.2	2.677	0.165	-1.532	8.743
	0.2	5.813	0.716	-0.585	5.347
	0.4	9.242	1.520	-0.191	4.630
	0.6	15.407	3.271	0.164	4.390
	0.8	26.708	7.109	0.489	4.532
Tur	-0.8	1.134	4.13×10^{-2}	-0.662	2.891
	-0.6	1.389	7.39×10^{-2}	-0.477	2.677
	-0.4	1.751	0.134	-0.283	2.542
	-0.2	2.274	0.244	-0.083	2.497
	0.2	4.245	0.837	0.334	2.689
	0.4	6.116	1.573	0.584	2.933
	0.6	9.128	2.987	0.764	2.274
	0.8	14.089	5.732	0.981	3.707

Table 4.10: Effect of (λ) value on the four moments of each series in Kufa station .

Parameter	λ	Mean	Sd	C_s	C_k
T.A	-0.8	1.230	$1.52*10^{-3}$	-1.490	6.391
	-0.6	1.591	$4.21*10^{-3}$	-1.406	6.051
	-0.4	2.189	$1.17*10^{-3}$	-1.325	5.741
	-0.2	3.223	0.033	-1.245	5.452
	0.2	9.071	0.252	-1.091	4.946
	0.4	17.298	0.703	-1.017	4.724
	0.6	35.492	1.961	-0.944	4.521
	0.8	77.240	5.486	-0.872	4.338
Ca	-0.8	1.218	$3.97*10^{-3}$	-0.349	3.216
	-0.6	1.559	$9.87*10^{-3}$	-0.269	3.071
	-0.4	2.098	$2.46*10^{-2}$	-0.178	2.954
	-0.2	2.996	$6.16*10^{-2}$	$-9.38*10^{-2}$	2.864
	0.2	7.491	0.381	$7.24*10^{-2}$	2.755
	0.4	13.109	0.953	0.153	2.732
	0.6	24.384	2.388	0.231	2.728
	0.8	47.688	5.988	0.368	2.742
CI	-0.8	1.229	$1.18*10^{-3}$	-0.953	3.843
	-0.6	1.589	$3.29*10^{-3}$	-0.911	3.741
	-0.4	2.182	$9.21*10^{-3}$	-0.871	3.645
	-0.2	3.216	$2.58*10^{-2}$	-0.831	3.554
	0.2	9.021	0.202	-0.749	3.392
	0.4	17.165	0.564	-0.712	3.315
	0.6	35.128	1.579	-0.671	3.246
	0.8	76.234	4.424	-0.631	3.182
EC	-0.8	1.246	$2.96*10^{-4}$	-1.090	5.493
	-0.6	1.644	$1.2*10^{-3}$	-0.997	5.194
	-0.4	2.357	$5.09*10^{-3}$	-0.914	4.928
	-0.2	3.806	0.021	-0.834	4.688
	0.2	15.934	0.366	-0.680	4.279
	0.4	41.348	1.525	-0.606	4.106
	0.6	120.811	6.354	-0.533	3.954
	0.8	383.737	26.481	-0.462	3.820

Table 4.10 : Cont.

Parameter	λ	Mean	Sd	C_s	C_k
T.H	-0.8	1.238	$1.34 \cdot 10^{-3}$	$-6.81 \cdot 10^{-2}$	2.554
	-0.6	1.617	$4.31 \cdot 10^{-3}$	$8.47 \cdot 10^{-3}$	2.587
	-0.4	2.259	$1.39 \cdot 10^{-2}$	$7.94 \cdot 10^{-2}$	2.634
	-0.2	3.448	$4.49 \cdot 10^{-2}$	0.152	2.697
	0.2	11.123	0.469	0.304	2.576
	0.4	23.516	1.518	0.383	2.996
	0.6	54.355	4.931	0.464	3.1.7
	0.8	134.578	16.011	0.547	3.202
Mg	-0.8	1.186	0.020	-0.195	2.840
	-0.6	1.489	0.042	0.017	2.732
	-0.4	1.944	0.090	0.229	2.733
	-0.2	2.649	0.195	0.441	2.846
	0.2	5.712	0.942	0.865	3.399
	0.4	9.064	2.099	1.076	3.838
	0.6	15.111	4.721	1.287	4.385
	0.8	26.367	10.724	1.496	5.038
Tur	-0.8	1.122	0.040	-0.422	2.596
	-0.6	1.363	0.070	-0.256	2.467
	-0.4	1.711	0.126	-0.088	2.401
	-0.2	2.201	0.225	0.082	2.399
	0.2	3.987	0.737	0.426	2.599
	0.4	5.630	1.350	0.600	2.802
	0.6	8.207	2.492	0.774	3.073
	0.8	12.332	4.633	0.949	3.412

Table 4.11: Effect of (λ) values the first four moments of each series in Diwaniya station.

Parameter	λ	Mean	Sd	C_s	C_k
T.A	-0.8	1.221	0.004	-1.856	9.455
	-0.6	1.575	0.099	-1.613	8.444
	-0.4	2.139	0.026	-1.375	7.559
	-0.2	3.101	0.062	-1.142	6.809
	0.2	8.172	0.445	-0.695	5.691
	0.4	14.871	1.158	0.480	5.713
	0.6	28.908	3.027	-0.272	5.046
	0.8	59.359	7.935	-0.071	4.885
Ca	-0.8	1.215	6.2×10^{-3}	-0.37	2.413
	-0.6	1.553	0.015	-0.288	2.369
	-0.4	2.083	3.70×10^{-3}	-0.201	2.343
	-0.2	2.959	9.04×10^{-3}	-0.112	2.319
	0.2	7.276	0.543	0.071	2.402
	0.4	12.600	1.337	0.165	2.473
	0.6	23.147	3.296	0.262	2.573
	0.8	44.709	8.145	0.361	2.706
EC	-0.8	1.246	3.2×10^{-4}	-0.343	2.465
	-0.6	1.644	1.3×10^{-3}	-0.302	2.444
	-0.4	2.357	5.6×10^{-3}	-0.262	2.427
	-0.2	3.805	2.34×10^{-2}	-0.222	4.415
	0.2	15.937	0.409	-0.141	2.404
	0.4	41.352	1.709	-0.099	2.405
	0.6	120.844	7.151	-0.058	2.410
	0.8	383.936	29.932	-1.5×10^{-2}	2.421
CI	-0.8	1.225	1.79×10^{-3}	0.163	2.307
	-0.6	1.578	4.76×10^{-3}	0.198	2.326
	-0.4	2.146	0.013	0.234	2.348
	-0.2	3.118	3.37×10^{-3}	0.269	2.375
	0.2	8.288	0.239	0.341	2.438
	0.4	15.162	0.639	0.377	2.476
	0.6	29.646	1.703	0.413	2.518
	0.8	61.222	4.545	0.452	2.564
T.H	-0.8	1.240	-0.001	-0.752	3.126
	-0.6	1.623	-0.004	-0.692	3.027
	-0.4	2.278	-0.012	-0.632	2.94
	-0.2	3.510	0.039	-0.572	2.864
	0.2	11.796	0.430	-0.452	2.745
	0.4	25.728	1.437	-0.392	2.704
	0.6	61.629	4.803	-0.332	2.675
	0.8	158.522	16.068	-0.271	2.658

Table 4.11: Cont.

Parameter	λ	Mean	Sd	C_s	C_k
Mg	-0.8	1.197	0.008	-0.692	2.775
	-0.6	1.511	0.0182	-0.618	2.674
	-0.4	1.986	-0.039	-0.544	2.586
	-0.2	2.735	0.087	-0.469	2.513
	0.2	6.055	0.417	0.319	2.411
	0.4	9.738	0.917	-0.244	2.383
	0.6	16.423	2.019	-0.168	2.370
	0.8	28.872	4.452	-0.091	2.372
Tur	-0.8	1.181	0.015	-0.201	4.147
	-0.6	1.478	0.030	-1.078	3.824
	-0.4	1.916	0.061	-0.956	3.544
	-0.2	2.586	0.123	-0.834	3.308
	0.2	5.384	0.511	-0.591	2.975
	0.4	8.308	1.046	-0.467	2.884
	0.6	13.368	2.144	-0.341	2.848
	0.8	22.330	4.429	-0.213	2.873

Table 4.12: Values of statistical parameters of normally distributed series at AL-Hindya station.

	T.A	Ca	Cl	EC	T.H	Mg	Tur
λ	1.195	0.352	0.505	-0.239	0.711	0.513	-0.127
Mean	282.891	9.903	23.177	3.409	94.396	12.268	2.521
Sd	61.786	1.149	4.860	0.031	16.310	3.340	0.304
C_s	-0.019	0.001	0.000	0.017	-0.046	0.014	-0.008
C_k	3.505	2.967	2.211	2.596	5.247	4.443	2.503

Table 4.13: Values of statistical parameters of normally distributed series at Kufa station .

	T.A	Ca	Cl	EC	T.H	Mg	Tur
λ	3.543	0.023	1.509	2.255	-0.624	-0.616	-0.297
Mean	2.68×10^7	4.827	1602.531	46404668	1.561	1.461	1.939
Sd	7504989	0.169	251.160	871711.8	0.004	-0.039	0.169
C_s	-0.002	0.000	-0.003	0.001	0.000	0.000	-0.001
C_k	3.271	2.792	3.136	3.338	2.582	2.736	2.392

Table 4.14: Values of statistical parameters of normally distributed series at Diwaniya station.

	T.A	Ca	Cl	EC	T.H	Mg	Tur
λ	0.839	0.0471	-1.735	0.877	1.7047	1.047	1.139
Mean	88.719	4.997	0.572	607.18	18228.82	60.869	57.287
Sd	9.589	0.273	1.827×10^{-5}	51.762	3811.273	11.866	15.218
C_s	-0.031	0.000	0.006	0.000	0.010	0.004	0.015
C_k	4.865	2.366	2.265	2.426	2.756	2.397	3.076

Tables (4.15), (4.16) and (4.17) show a comparison among transformation procedures. The best procedure which displays the value of skewness coefficient ($C_s=0$) and kurtosis coefficient ($C_k \approx 3$) is the Box-Cox transformation for most series in the three stations. The normality was tested by plotting observed cumulative probability against expected cumulative probability where

$$p(x) = \frac{m}{N+1}, \quad \dots\dots\dots (4-5)$$

where $p(x)$ is the observed cumulative probability of the value (x) of transformed water quality parameter .

m is the rank of (x) in ascending order.

N is the number of tested data ($N=120$).

The resulting plots for AL-Hindya station are shown in figure (4.6) and the other related figures are given in the appendix (B)(figures B.9 & B.10).All these figures show good agreement .

Table 4.15: The comparison among the three transformation procedures of the data in AL –Hindya station.

	Transformation	T.A	Ca	Cl	EC	T.H	Mg	Tur
C _s	Box-Cox	-0.019	0.001	0.000	0.017	-0.046	0.014	-0.008
	Log	-0.736	-0.058	-0.226	0.448	-1.522	-0.858	0.249
	Square root	-0.578	0.114	-0.004	0.336	-0.441	-0.009	0.656
C _k	Box-Cox	3.505	2.697	2.211	2.596	5.247	4.443	2.503
	Log	4.832	3.098	2.338	2.729	11.756	6.413	2.567
	Square root	4.195	2.934	2.210	2.934	6.134	4.458	3.092

Table 4.16: The comparison among the three transformation procedures of the data in Kufa station.

	Transformation	T.A	Ca	Cl	EC	T.H	Mg	Tur
C _s	Box-Cox	-0.002	0.000	-0.003	0.001	0.000	0.000	-0.001
	Log	-0.494	-0.033	0.359	-0.091	0.615	0.784	0.392
	Square root	-0.980	-0.375	0.192	-0.591	0.428	1.181	0.689
C _k	Box-Cox	3.271	2.792	3.136	3.338	2.582	2.736	2.392
	Log	4.091	3.385	2.797	3.732	2.448	3.184	2.515
	Square root	4.620	3.492	2.729	4.027	3.115	4.098	2.929

Table 4.17: The comparison among the three transformation procedure of the data in Diwaniya station .

	Transformation	T.A	Ca	Cl	EC	T.H	Mg	Tur
C _s	Box-Cox	-0.031	0.000	0.006	0.000	0.010	0.004	0.015
	Log	-0.570	0.325	0.234	0.388	-0.064	-0.092	-0.480
	Square root	-0.375	0.280	0.213	-0.078	-0.362	-0.206	-0.404
C _k	Box-Cox	4.865	2.366	2.265	2.426	2.756	2.397	3.076
	Log	5.707	3.451	2.361	2.313	2.512	2.687	2.899
	Square root	5.166	3.529	2.519	2.407	2.687	2.309	2.859

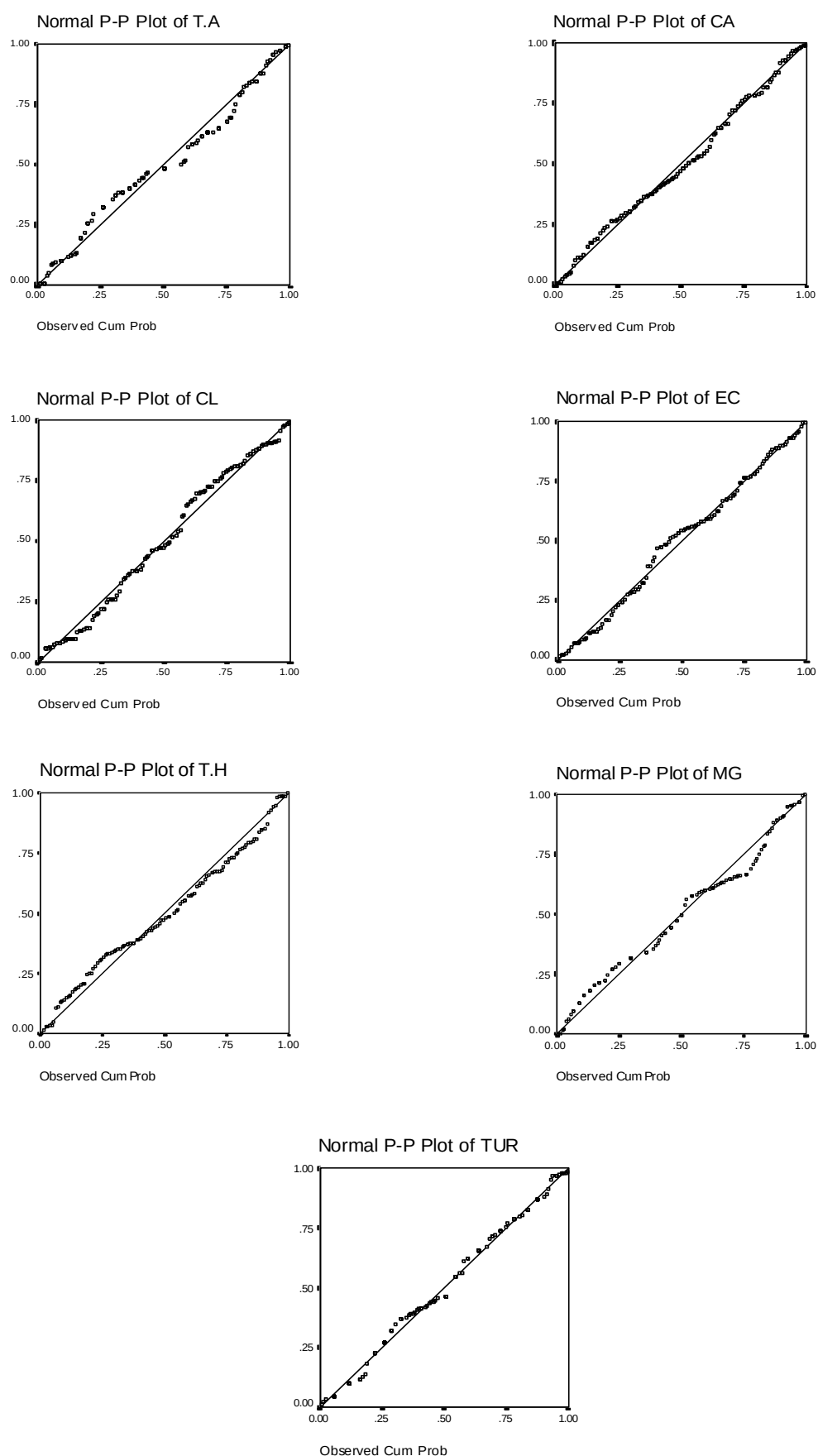


Fig. 4.6: Normal distribution test of AL-Hindya station parameters .

4.4 Detection and Removal of Periodic Components

Generally, water quality constituents show seasonal patterns which may due to the influence of the annual cyclic pattern of the hydrological inputs to the river environment (Kurunc et. al., 2005). If a time series contains a seasonal fluctuation then the correlogram – which is a plot of autocorrelation coefficient (r_k) against the lag (k) – will also exhibit an oscillation at the same frequency (Chatfield, 1982) Then the periodicity can be detected by the correlogram. The correlograms of normalized data for AL –Hindya station parameters are shown in figure (4.7) and the figures of the other two stations are presented in the appendix(B) (figures B.11 & B.12). These figures indicate that the periodicity is found in the series because the correlogram is periodic too.

Removal of periodicity from data is done by harmonic analysis of monthly means and standard deviations using Fourier series .

Harmonic analysis is done by using the following steps :

- 1.Calculation of sample monthly means and monthly standard deviations for each of the twelve months.
- 2.Calculation of Fourier coefficients (A_h, B_h) using equations (2-15)&(2-16).
- 3.Calculation of the number of significant harmonic for means and standard deviations using the empirical method with $n = 10$, $w=12$, $C= 1$ for mean and $C=2$ for standard deviation.
- 4.Appling equations (2-14) and (2-19) to calculate the periodic component of means (μ_j) and standard deviation (σ_j).
- 5.Removal of periodic component from the series by:

$$Y'_{i,j} = \frac{F_{i,j} - \mu_j}{\sigma_j}, \dots\dots\dots(4-6)$$

where $Y'_{i,j}$ is the series free from periodic component at year (i) & month (j).

$F_{i,j}$ is the transformed series .

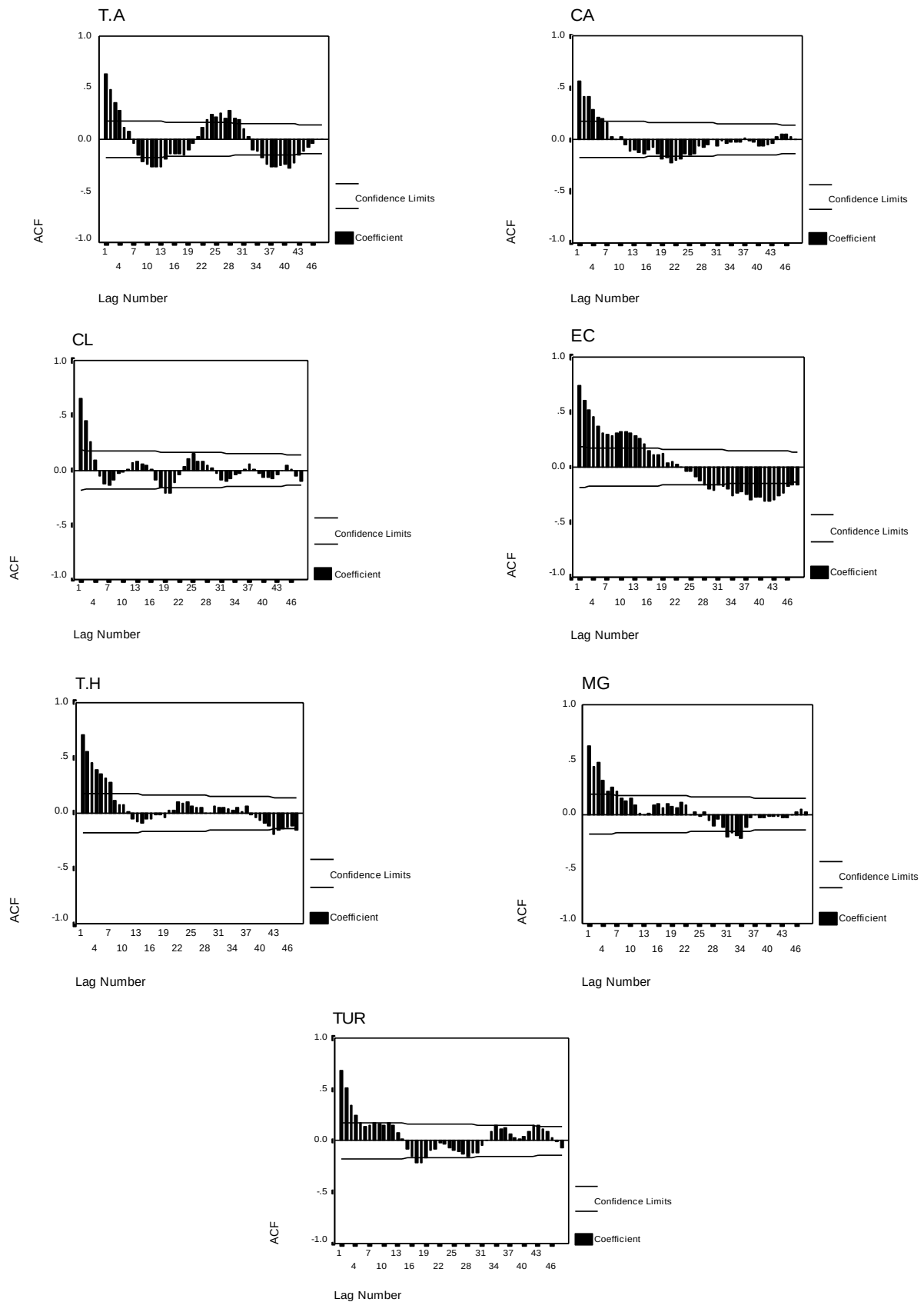


Fig. 4.7 : Correlogram of normalized series of AL-Hindya station.

The results of the application of harmonic analysis are shown in tables from (4.18) to (4.23), which indicate that the significant harmonic of mean and standard deviation are between (1-6) for each water quality parameter in the three stations. Table (4.24) shows the number of significant harmonics of mean and standard deviation of the three stations.

Table 4.18 : Results of harmonic analysis (explained variance) for monthly means of AL-Hindya station with , $w= 12$, $n=10$, $C=1$, $a= 0.033$ $P_{\min} = 0.036$, $P_{\max} =0.964$.

Harmonics h	T.A	Ca	Cl	EC	T.H	Mg	Tur
1	0.486	0.461	0.919	>1.0	0.787	0.204	0.696
2	0.544	0.513	0.932	>1.0	0.831	0.279	0.874
3	0.691	0.536	0.962	>1.0	0.871	0.307	0.954
4	0.753	0.849	0.999	>1.0	>1.0	0.446	0.987
5	0.817	0.954	>1.0	>1.0	>1.0	0.506	>1.0
6	>1.0	>1.0	>1.0	>1.0	>1.0	0.552	>1.0

Table 4.19: Results of harmonic analysis (explained variance) for monthly means of Kufa station with , $w= 12$, $n=10$, $C=1$, $a= 0.033$, $P_{\min}= 0.036$, $P_{\max} =0.964$.

Harmonics h	T.A	Ca	Cl	EC	T.H	Mg	Tur
1	>1.0	0.219	0.089	0.609	>1.0	0.659	0.171
2	>1.0	0.391	0.112	0.689	>1.0	0.703	0.489
3	>1.0	0.911	0.313	0.786	>1.0	0.859	0.509
4	>1.0	0.919	0.595	0.831	>1.0	0.928	0.741
5	>1.0	0.945	0.678	0.839	>1.0	0.931	>1.0
6	>1.0	>1.0	>1.0	>1.0	>1.0	>1.0	>1.0

Table 4.20: Results of harmonic analysis (explained variance) for monthly means of Diwaniya station with , $w= 12$, $n=10$, $C=1$, $a= 0.033$, $P_{\min} = 0.036$, $P_{\max} =0.964$.

Harmonics h	T.A	Ca	Cl	EC	T.H	Mg	Tur
1	0.764	0.497	>1.0	0.009	>1.0	>1.0	0.831
2	0.915	0.531	>1.0	0.145	>1.0	>1.0	0.870
3	>1.0	0.978	>1.0	0.246	>1.0	>1.0	0.916
4	>1.0	>1.0	>1.0	0.256	>1.0	>1.0	0.933
5	>1.0	>1.0	>1.0	0.927	>1.0	>1.0	0.987
6	>1.0	>1.0	>1.0	>1.0	>1.0	>1.0	0.991

Table 4.21: Results of harmonic analysis (explained variance) for monthly standard deviation of AL-Hindya station with , $w= 12$, $n=10$, $C=2$, $a= 0.033$ $P_{\min} = 0.026$, $P_{\max} =0.974$.

Harmonics h	T.A	Ca	Cl	EC	T.H	Mg	Tur
1	0.693	0.002	0.061	0.426	0.184	0.103	0.264
2	0.699	0.567	0.370	0.621	0.753	0.509	0.505
3	0.714	0.728	0.617	0.833	0.754	0.512	0.630
4	0.836	0.805	0.653	0.959	0.829	0.786	0.734
5	0.986	0.973	0.979	>1.0	0.976	0.959	0.926
6	>1.0	>1.0	>1.0	>1.0	>1.0	>1.0	>1.0

Table 4.22: Results of harmonic analysis (explained variance) for monthly standard deviation of Kufa station with , $w= 12$, $C=2$, $a= 0.033$, $P_{\min} = 0.026$, $P_{\max}=0.974$.

Harmonics h	T.A	Ca	Cl	EC	T.H	Mg	Tur
1	0.391	0.069	0.860	0.195	0.68	0.501	0.120
2	0.614	0.336	0.863	0.549	0.779	0.849	0.399
3	0.707	0.650	0.873	0.705	0.814	0.888	0.480
4	0.762	0.970	0.969	0.787	0.843	0.899	0.623
5	0.768	0.853	0.987	0.949	0.945	0.903	0.980
6	>1.0	>1.0	>1.0	>1.0	>1.0	>1.0	>1.0

Table 4.23: Results of harmonic analysis (explained variance) for monthly standard deviation of Diwaniya station with , $w= 12$, $C=2$, $a=0.033$, $P_{\min} = 0.026$, $P_{\max}=0.974$.

Harmonics h	T.A	Ca	Cl	EC	T.H	Mg	Tur
1	0.483	0.100	0.182	0.399	0.224	0.329	0.285
2	0.498	0.235	0.364	0.500	0.558	0.595	0.387
3	0.513	0.724	0.546	0.538	0.865	0.756	0.469
4	0.836	0.846	0.727	0.678	0.936	0.839	0.585
5	>1.0	0.992	0.909	0.783	0.974	0.859	0.956
6	>1.0	>1.0	>1.0	>1.0	>1.0	>1.0	>1.0

Table 4.24: The number of significant harmonics of mean (M) and standard deviation (Sd) of the three stations.

Station	T.A		Ca		Cl		EC		T.H		Mg		Tur	
	M	Sd	M	Sd	M	Sd	M	Sd	M	Sd	M	Sd	M	Sd
AL-Hindya	6	5	6	6	4	5	1	5	4	5	6	6	4	6
Kufa	1	6	6	6	6	5	6	6	1	6	6	6	5	5
Diwaniya	3	5	3	5	1	6	6	6	1	5	1	6	5	6

Since the $Y'_{i,j}$ series obtained from equation (4-6) with a limited number of harmonics, then, this series is further standardized to ascertain a zero mean and unit variance (AL-Suhaili, 1986) thus

$$W_{i,j} = \frac{Y'_{i,j} - \mu}{\sigma}, \quad \dots\dots\dots(4-7)$$

where $W_{i,j}$ is the dependent stochastic components of year i and month j

μ is the mean of $Y'_{i,j}$ series.

σ is the standard deviation of $Y'_{i,j}$ series.

The result series is called stochastic series, which contains a depended part on time which may be represented by AR (p), MA(q) or ARMA(p,q) models, and an independent part (a_t) can be described by some probability distribution function.

Figures (4.8) & (4.9) and figures from (B.13) to (B.16) in the appendix (B) show a suitable fit between periodic mean and Fourier fitted mean and between a periodic standard deviation and Fourier fitted standard deviation for each water quality parameter in the three stations, so the harmonic method is used for the removal of periodic component.

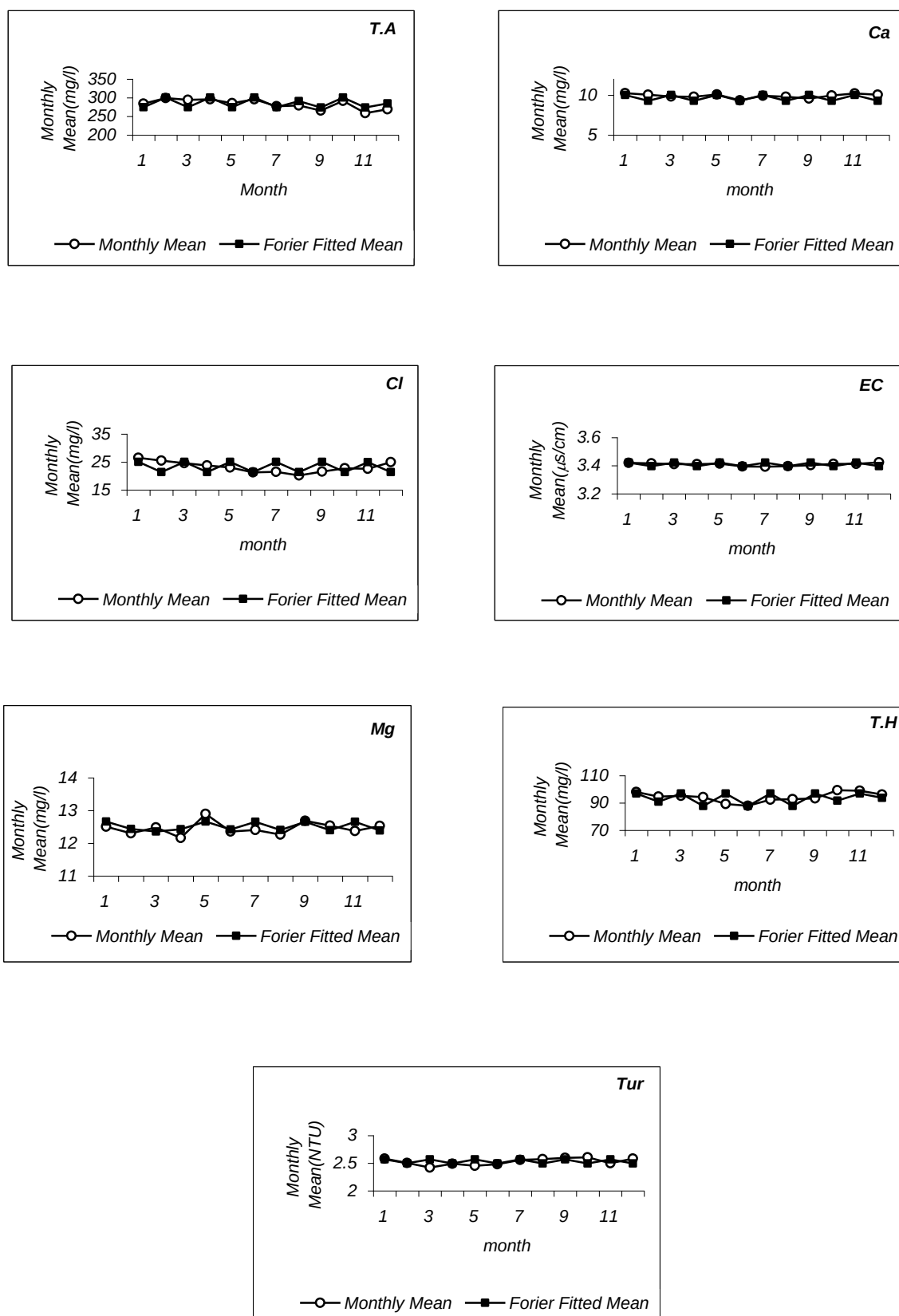


Fig. 4.8 :Periodic component of monthly mean of AL-Hindya station.

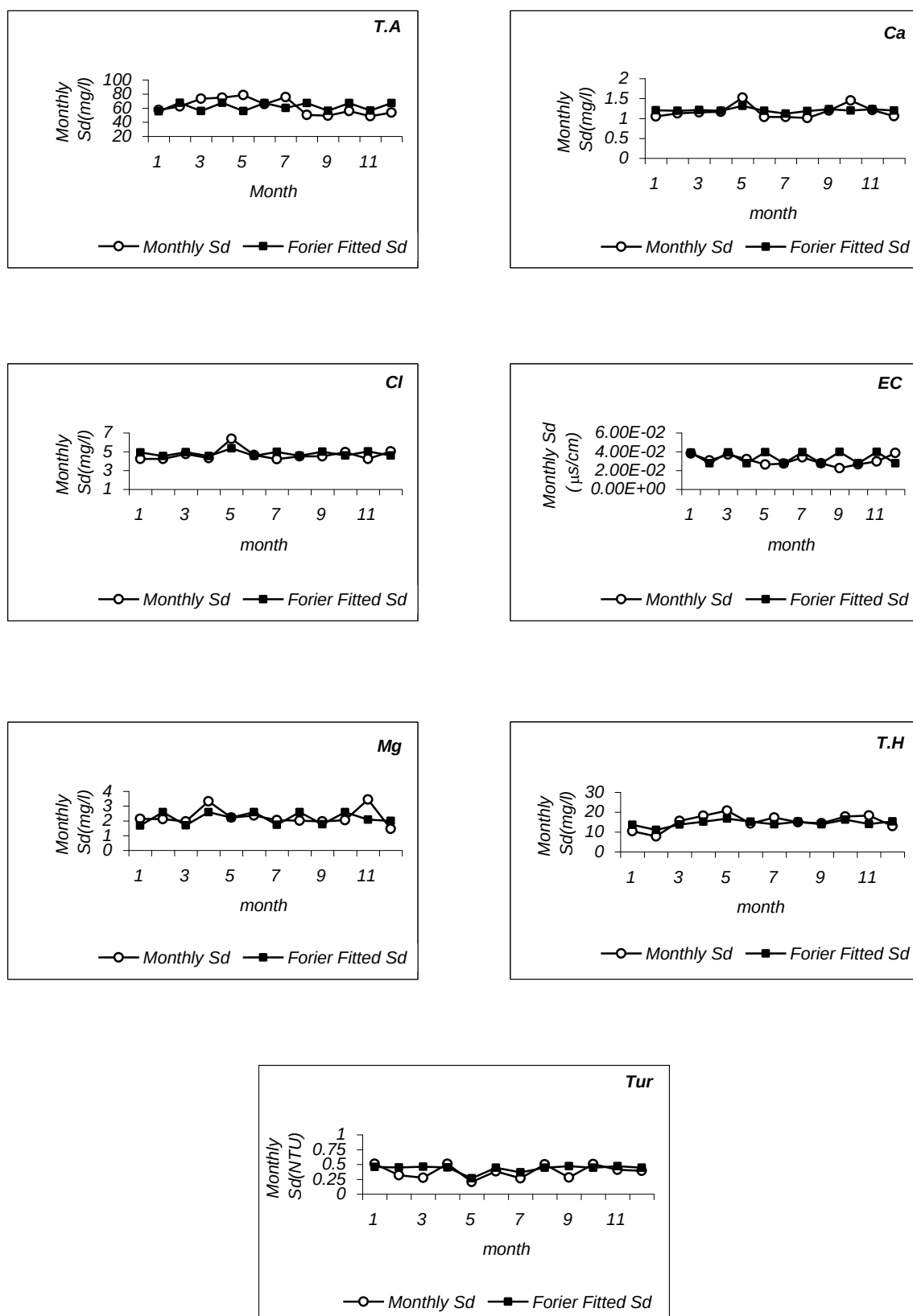


Fig. 4.9: Periodic component of monthly standard deviation of AL-Hindya station.

4.4 Stochastic Model

This section contains three steps of model building, identification, estimation of model parameters and diagnostic checks. These stages are made after obtaining the stochastic series.

4.4.1 Model Identification

As explained in section (2.11) autocorrelation function (ACF) and partial autocorrelation function (PACF) is an important guide to the properties of a time series because they often provide insight into the probability model which generated the data. Figures (4.10) and (4.11) show the behavior of the autocorrelation function (ACF) and partial autocorrelation function (PACF) of dependent stochastic series ($W_{i,j}$) of AL-Hindya water quality parameters with 95% confidence limits and the other figures of the two stations are given in the appendix (B)(figures from (B.17) to (B.20)). For example the suggested model of T.A of AL-Hindya station was AR(1) because the (ACF)) tails off while its (PACF) has a cut off after lag (1) and so on.

4.4.2 Estimation of Model Parameters

The identification process having led to a tentative formulation for the model, then, it is needed to obtain efficient estimates of the model parameters. Estimation of model parameter contains two stages:

- a. Initial estimation of model parameters.
- b. Final estimation of model parameters .

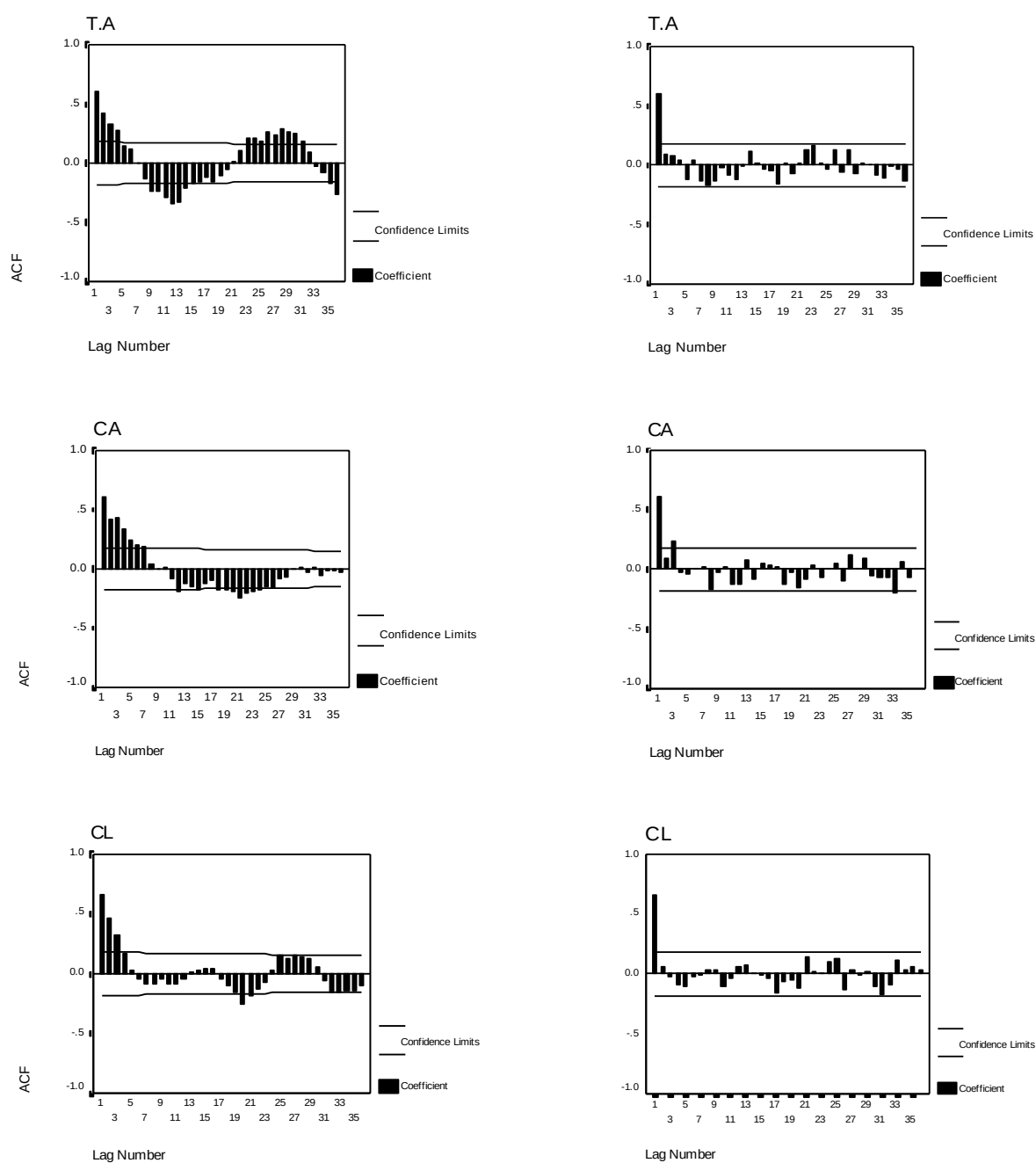


Fig. 4.10 :ACF&PACF of T.A, Ca and Cl of AL-Hindya station.

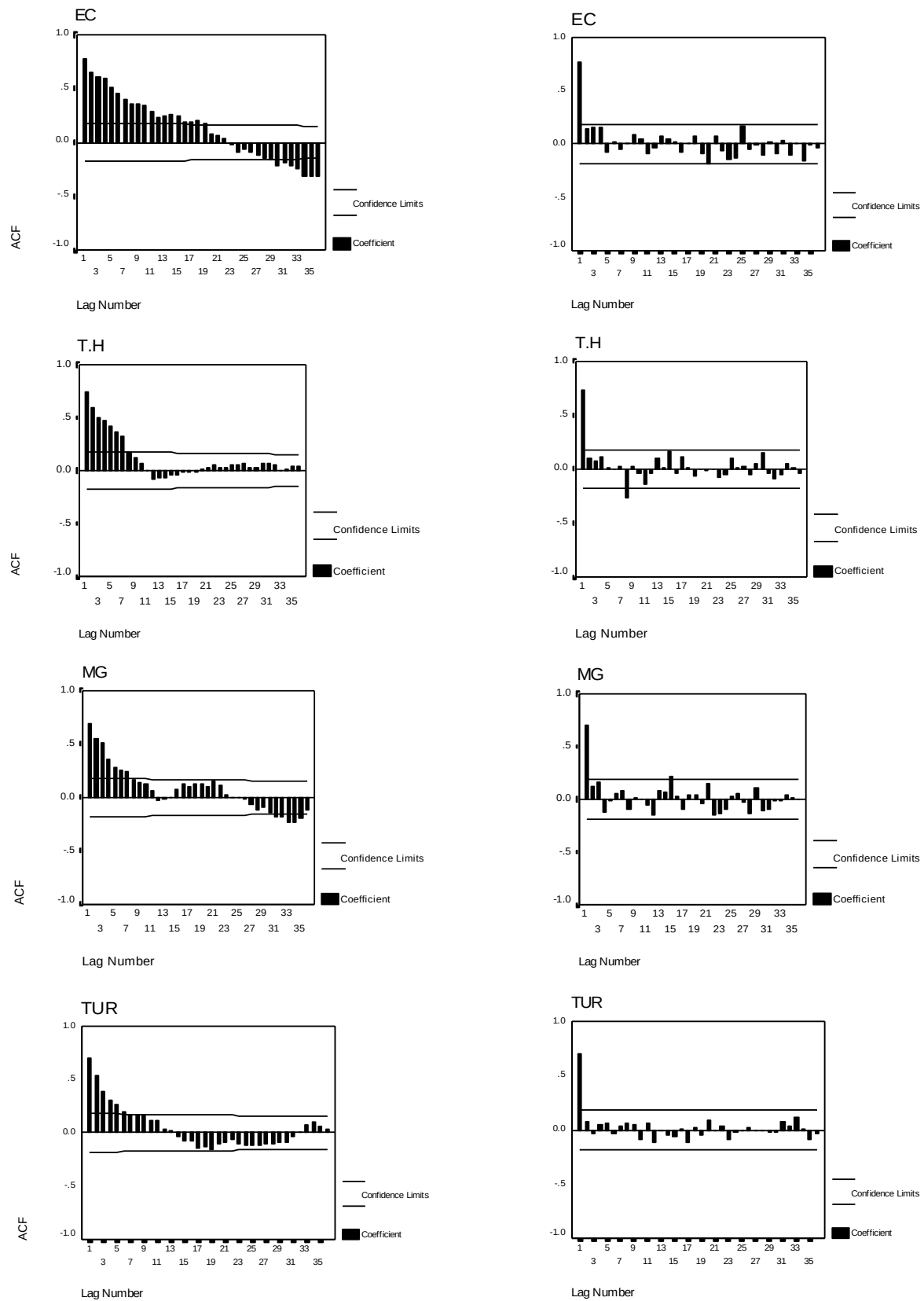


Fig. 4.11: ACF&PACF of EC, T.H, Mg and Tur of AL-Hindya station.

The calculation of the initial estimation begins with determining of the autocorrelation coefficients of the dependent stochastic series which are then used in the estimation of the model's parameters as explained in chapter two. Tables (4.25), (4.26) and (4.27) summarize the initial estimation of model parameters.

Conditional sum of squares method explained in chapter 2 is used in the final estimation of model parameter. As an example, the calculation of sum of squared errors for the Ca parameter at AL-Hindya is given table(4.28).Tables (4.29), (4.30) and (4.31) show the results of final estimation.

Table 4.25: Initial estimation of model parameter of AL-Hindya station.

Parameter	Suggested model	AR parameter		MA parameter
		ϕ_1	ϕ_2	θ_1
T.A	AR(1)	0.601		
Ca	AR(1)	0.604		
Cl	AR(1)	0.658		
EC	AR(1)	0.733		
	AR(2)	0.662	0.144	
	ARMA(1,1)	0.848		0.189
T.H	AR(1)	0.739		
	AR(2)	0.665	0.099	
Mg	AR(1)	0.698		
	AR(2)	0.609	0.127	
Tur	AR(1)	0.702		

Table 4.26: Initial estimation of model parameter of Kufa station.

Parameter	Suggested model	AR parameter		MA parameter	
		ϕ_1	ϕ_2	θ_1	θ_2
T.A	AR(1)	0.352			
	AR(2)	0.243	0.028		
	ARMA(1,1)	0.422			0.08
Ca	AR(1)	0.524			
	AR(2)	0.443	0.154		
Cl	AR (1)	0.482			
	MA(1)				- 0.768
EC	AR(1)	0.346			
	MR(1)				- 0.400
T.H	AR(1)	0.628			
	AR(2)	0.467	0.278		
Mg	AR(1)	0.698			
	AR(2)	0.465	0.333		
	AR(3)	0.363	0.192	0.305	
Tur	AR(1)	0.611			
	AR(2)	0.473	0.227		

Table 4.27: Initial estimation of model parameter of Diwaniya station.

Parameter	Suggested model	AR parameter		MA parameter	
		ϕ_1	ϕ_2	θ_1	θ_2
T.A	AR(1)	0.495			
	AR(2)	0.402	0.188		
Ca	AR(1)	0.524			
Cl	AR (1)	0.449			
	MA(2)				
EC	AR (1)	0.217		- 0.623	
	AR(2)	0.193	0.113		
	MA(1)			- 0.229	
	MA(2)			- 0.199	- 0.165
T.H	AR(1)	0.391			
	AR(2)	0.336	0.141		
Mg	AR(1)	0.231			
	AR(2)				
	ARMA(1,1)				
Tur	AR (1)	0.318			
	AR(2)	0.251	0.207		
	MA(1)			-0.359	
	MA(2)			-0.284	-0.343

Table 4.28: Conditional sum of squares of Ca of AL-Hindya station where the model is AR(1), $\phi=0.609$ and $a_t = W_t - 0.609 W_{t-1}$.

t	W_t	a_t	a_t^2	t	W_t	a_t	a_t^2	t	W_t	a_t	a_t^2
0		0	0	40	0.98659	0.108	0.0117	80	0.38831	0.2818	0.079401
1	-0.97412	-0.974	0.94891	41	0.5651	-0.0357	0.0013	81	0.57309	0.3366	0.113304
2	-0.66659	-0.073	0.00538	42	1.6266	1.2825	1.6447	82	0.25166	-0.0974	0.009477
3	0.865639	1.2716	1.61695	43	0.63285	-0.3578	0.128	83	-0.2958	-0.449	0.201623
4	-0.6056	-1.133	1.28318	44	0.78231	0.3969	0.1575	84	0.2601	0.4402	0.193794
5	-0.35885	0.01	9.92E-05	45	1.13675	0.6603	0.436	85	-0.3808	-0.5392	0.290711
6	-0.15786	0.0607	0.00368	46	0.29734	-0.3949	0.156	86	-0.6475	-0.4156	0.172742
7	6.26E-02	0.1587	0.02519	47	-0.5451	-0.7262	0.5273	87	-0.8271	-0.4328	0.187309
8	-0.12433	-0.162	0.02639	48	-0.1094	0.2225	0.0495	88	-1.162	-0.6583	0.43331
9	0.685145	0.7609	0.57891	49	-0.7066	-0.6399	0.4095	89	1.31869	2.0263	4.106047
10	1.215753	0.7985	0.6376	50	-0.7479	-0.3176	0.1009	90	-0.0343	-0.8374	0.701199
11	1.146328	0.4059	0.16478	51	-0.8284	-0.373	0.1391	91	0.67488	0.6958	0.484091
12	0.613791	-0.084	0.00711	52	-0.6859	-0.1814	0.0329	92	0.99307	0.5821	0.338804
13	-0.17618	-0.55	0.30247	53	-1.9611	-1.5434	2.3821	93	1.02108	0.4163	0.173304
14	-0.0438	0.0635	0.00403	54	-2.1017	-0.9074	0.8234	94	0.5691	-0.0527	0.002781
15	9.01E-02	0.1168	0.01364	55	-1.8863	-0.6064	0.3677	95	0.49323	0.1466	0.021505
16	8.16E-02	0.0267	0.00071	56	-1.8972	-0.7484	0.5601	96	0.69745	0.3971	0.157672
17	-0.17189	-0.222	0.04909	57	-1.5026	-0.3473	0.1206	97	-0.4418	-0.8666	0.750942
18	0.391584	0.4963	0.24628	58	-2.023	-1.1079	1.2274	98	-0.8986	-0.6295	0.396257
19	-0.22491	-0.463	0.21472	59	-1.1311	0.1009	0.0102	99	-1.1575	-0.6102	0.372376
20	-0.16201	-0.025	0.00063	60	-2.3801	-1.6912	2.8603	100	-0.583	0.1219	0.014863
21	-0.01059	0.0881	0.00776	61	-0.756	0.6935	0.4809	101	-0.5957	-0.2406	0.057903
22	-0.32648	-0.32	0.10242	62	-0.7717	-0.3113	0.0969	102	-0.4465	-0.0838	0.00702
23	-0.63295	-0.434	0.18846	63	-1.1575	-0.6875	0.4726	103	-1.2981	-1.0261	1.052943
24	-0.65873	-0.273	0.07467	64	-0.4622	0.2427	0.0589	104	-1.0464	-0.2558	0.06545
25	-0.53447	-0.133	0.01777	65	-0.1124	0.1691	0.0286	105	0.78303	1.4203	2.01714
26	-0.37733	-0.052	0.00269	66	0.28992	0.3583	0.1284	106	-0.1763	-0.6532	0.426673
27	1.083317	1.3131	1.72426	67	0.45079	0.2742	0.0752	107	-1.1502	-1.0428	1.087384
28	2.15934	1.4996	2.2488	68	-0.6135	-0.888	0.7885	108	0.64429	1.3447	1.808342
29	1.554489	0.2395	0.05734	69	-0.2782	0.0954	0.0091	109	1.92662	1.5342	2.353901
30	-0.82538	-1.772	3.14022	70	-0.8505	-0.6811	0.4638	110	1.38039	0.2071	0.04288
31	-0.19164	0.311	0.09673	71	-0.2432	0.2747	0.0755	111	0.88398	0.0433	0.001877
32	1.443232	1.5599	2.43342	72	0.26007	0.4082	0.1666	112	0.6739	0.1356	0.018377
33	-1.10289	-1.982	3.92761	73	0.50974	0.3514	0.1235	113	0.131	-0.2794	0.078069
34	1.401416	2.0731	4.29765	74	1.01251	0.7021	0.4929	114	0.59881	0.519	0.269392
35	1.95857	1.1051	1.22126	75	-0.3954	-1.012	1.0241	115	1.60492	1.2403	1.538221
36	1.126723	-0.066	0.00436	76	-0.4028	-0.162	0.0263	116	0.23641	-0.741	0.549062
37	1.533549	0.8474	0.71804	77	-0.3694	-0.1241	0.0154	117	-1.3048	-1.4487	2.098826
38	1.760463	0.8265	0.68316	78	0.65891	0.8839	0.7813	118	-0.359	0.4356	0.189774
39	1.442734	0.3706	0.13735	79	0.17492	-0.2264	0.0512	119	0.40012	0.6187	0.382825
								120	-0.4542	-0.6979	0.487034
											ss=68.454

Table 4.29: Final estimation of model parameter of AL-Hindya station.

Parameter	Suggested model	AR parameter		MA parameter	σ_a^2
		ϕ_1	ϕ_2	θ_1	
T.A	AR(1)	0.601			0.579
Ca	AR(1)	0.609			0.575
Cl	AR(1)	0.683			0.501
EC	AR(1)	0.793			0.349
	AR(2)	0.687	0.134		0.348
	ARMA(1,1)	0.878		0.254	0.341
T.H	AR(1)	0.741			0.410
	AR(2)	0.655	0.104		0.409
Mg	AR(1)	0.714			0.455
	AR(2)	0.615	0.145		0.449
Tur	AR(1)	0.702			0.459

Table 4.30: Final estimation of model parameter of Kufa station.

Parameter	Suggested model	AR parameter			MA parameter	σ_a^2
		ϕ_1	ϕ_2	ϕ_3	θ_1	
T.A	AR(1)	0.352				0.795
	AR(2)	0.342	0.03			0.801
Ca	AR(1)	0.529				0.657
	AR (2)	0.443	0.164			0.644
Cl	AR(1)	0.488				0.693
	MA (1)				-0.420	0.729
EC	AR(1)	0.351				0.797
	MA(1)				-0.315	0.811
T.H	AR(1)	0.622				0.557
	AR(2)	0.467	0.253			0.527
Mg	AR(1)	0.698				0.465
	AR(2)	0.470	0.333			0.417
	AR (3)	0.365	0.191	0.305		0.373
Tur	AR (1)	0.611				0.568
	AR(2)	0.468	0.232			0.543

Table 4.31 : Final estimation of model parameter of Diwaniya station.

Parameter	Suggested model	AR parameter		MA parameter		σ_a^2
		ϕ_1	ϕ_2	θ_1	θ_2	
T.A	AR(1)	0.500				0.683
	AR(2)	0.402	0.198			0.58
Ca	AR(1)	0.524				0.658
Cl	AR(1)	0.449				0.725
	MA(1)			-0.418		0.726
EC	AR(1)	0.217				0.846
	AR(2)	0.193	0.113			0.854
	MA(1)	-0.189				0.865
	MA(2)	-0.154	-0.130			0.863
T.H	AR(1)	0.391				0.769
	AR(2)	0.336	0.141			0.753
Mg	AR(1)	0.231				0.859
	MA(1)	-0.295				0.869
	ARMA(1,1)	-0.02		-0.303		0.844
Tur	AR(1)	0.318				0.815
	AR(2)	0.256	-0.207			0.780
	MA(1)			-0.219		0.824
	MA(2)			-0.259	-0.348	0.762

4.4.3 Diagnostic Checking

After exact estimation , Akaik information criterion (AIC) as explained in sec.(2.11) may be used to select the best model from a group of suggested models. The best model is the one that gives minimum AIC. AIC values are summarized in table(4.32). However, the adequacy of model fitting is also checked by examining the residuals (a_t) from the fitted process.

Table 4.32: AIC values of selected model of AL-Hindya , Kufa , and Diwaniya Stations

Model Type		AR(1)	AR(2)	AR(3)	MA(1)	MA(2)	ARMA(1,1)
AL-Hindya	EC	-124.701	-122.701				-125.105
	T.H	-104.790	-103.314				
	Mg	-92.495	-91.879				
Kufa	T.A	-25.529	-22.668				
	Ca	-48.501	-48.806				
	Cl	-41.955			-35.946		
	EC	-25.198			-23.065		
	T.H	-68.339	-73.186				
	Mg	-89.960	-101.047	-112.341			
	Tur	-65.927	-69.273				
Diwanya	T.A	-43.967	-46.280				
	Cl	-36.640			-36.504		
	EC	-15.542	-14.998		-15.403	-13.736	
	T.H	-29.590	-30.036				
	Mg	-16.258			-15.750		-16.406
	Tur	-22.551	-24.790		-21.231	-27.541	

Table (4.33) shows the Porte Manteau lack of fit test. The table shows that all models are succeeded because the calculated (Q) by equation (2-60) with $M=48$ is less the χ^2 -table with degree of freedom ($M-p-q$).

Table 4.33: Diagnostic check results with 95% confidence limits and ($M-p-q$) degree of freedom , where $M=48$.

Parameter		T.A	Ca	Cl	EC	T.H	Mg	Tur
ALHindya	Q-calculated	41.001	35.109	31.239	41.893	29.487	39.586	22.502
	χ^2 -table	63.981	63.981	63.981	62.806	63.981	63.881	63.981
Kufa	Q-calculated	38.597	41.455	29.357	31.011	35.976	41.070	35.492
	χ^2 -table	63.981	62.806	63.981	63.981	62.806	61.632	62.806
Diwaniya	Q-calculated	30.398	44.260	60.625	58.98	48.028	28.009	34.808
	χ^2 -table	62.806	63.981	63.981	63.981	62.806	62.806	62.806

The second test is the independency of the resulting (a_t) series, the correlogram of this series are computed for lag ($M=48$) as shown in figure (4.12) and the other related figures for the two stations are shown in the appendix (B) (figures (B.21) & (B.22)). The result shows that the most of computed lags lie inside the tolerance interval ($\pm 2/\sqrt{N}$, at 95% confidence limits). Hence, the selected models can be considered as appropriate models because of its capability of removing the dependency from data. The final ARMA models and model equation—which have minimum AIC and succeeded in the above two tests—are shown in table (4.34).

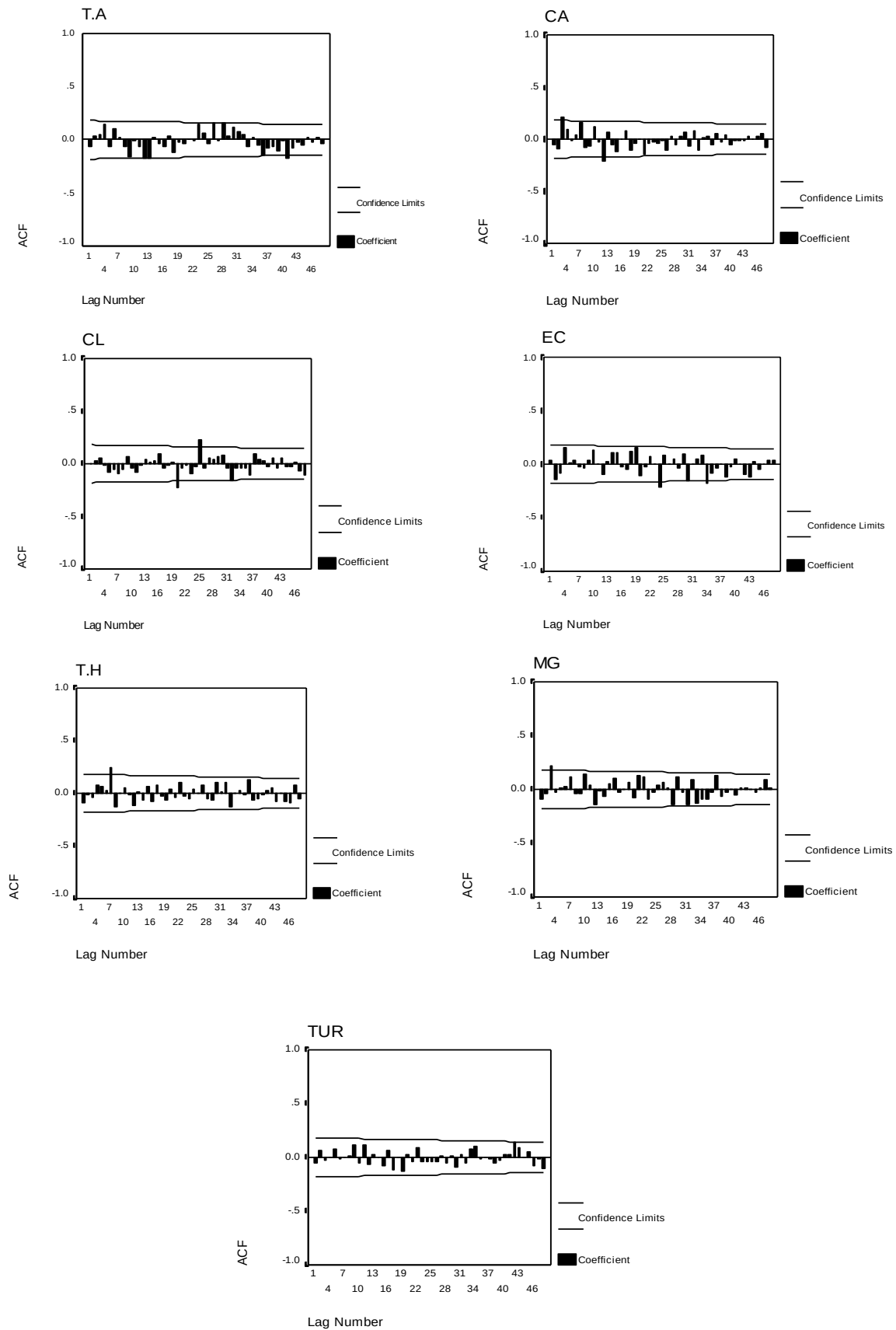


Fig.4.12:Autocorrelogram of residual series of AL-Hindya station parameters.

Table 4.34: Final model types of each water quality parameter of AL-Hindya , Kufa and Diwaniya stations.

Parameter	AL-Hindya		Kufa		Diwaniya	
	Model	Model equation	Model	Model equation	Model	Model equation
T.A	AR(1)	$W_t = 0.601W_{t-1} + a_t$	AR(1)	$W_t = 0.352W_{t-1} + a_t$	AR(2)	$W_t = 0.402W_{t-1} + 0.198W_{t-2} + a_t$
Ca	AR(1)	$W_t = 0.609W_{t-1} + a_t$	AR(2)	$W_t = 0.443W_{t-1} + 0.164W_{t-2} + a_t$	AR(1)	$W_t = 0.524W_{t-1} + a_t$
Cl	AR(1)	$W_t = 0.683W_{t-1} + a_t$	AR(1)	$W_t = 0.488W_{t-1} + a_t$	AR(1)	$W_t = 0.449W_{t-1} + a_t$
EC	ARMA(1,1)	$W_t = 0.878W_{t-1} - 0.254a_{t-1} + a_t$	AR(1)	$W_t = 0.351W_{t-1} + a_t$	AR(1)	$W_t = 0.217W_{t-1} + a_t$
T.H	AR(1)	$W_t = 0.741W_{t-1} + a_t$	AR(2)	$W_t = 0.467W_{t-1} + 0.253W_{t-2} + a_t$	AR(2)	$W_t = 0.336W_{t-1} + 0.141W_{t-2} + a_t$
Mg	AR(1)	$W_t = 0.714W_{t-1} + a_t$	AR(3)	$W_t = 0.365W_{t-1} + 0.191W_{t-2} + 0.305W_{t-3} + a_t$	ARMA(1,1)	$W_t = -0.02W_{t-1} + 0.303a_t + a_t$
Tur	AR(1)	$W_t = 0.702W_{t-1} + a_t$	AR(2)	$W_t = 0.468W_{t-1} + 0.232W_{t-2} + a_t$	MA (2)	$W_t = a_t + 0.259a_{t-1} + 0.348a_{t-2}$

Chapter Five

Fitting Box – Jenkins Seasonal Model

5.1 Identification of Models

For fitting seasonal ARIMA (p,d,q) $\times$ (P,D,Q) model –which is called "seasonal multiplicative autoregressive integrated moving average model "or Box – Jenkins seasonal model – to a time series of water quality data, three stages procedure are adopted. These stages are model identification, estimation of model parameters and diagnostic checking of the estimated parameters.

First of all the observation is plotted against time this will show up important features such as trend, seasonality, discontinuities and outlier.

Figures (5.1) and (5.2), for example show that there is little trend, fluctuation and seasonal variation for each water quality parameter in AL-Hindya station. Other related figures are presented in the appendix (B)(figures from B. 23 to B.26).

The procedure of fitting Box – Jenkins seasonal ARIMA model is summarized by the following steps :-

1. Transform the data by using natural log transformation which was found the most appropriate transformation (Box and Jenkins ,1976) and (Mcleod et. al., 1987).
2. Removing trend component by using the first order differencing.
3. Removing the seasonal variation by using the first order seasonal differencing .
4. Model identification by plotting (ACF) and (PACF) of each water quality parameter in the three stations.

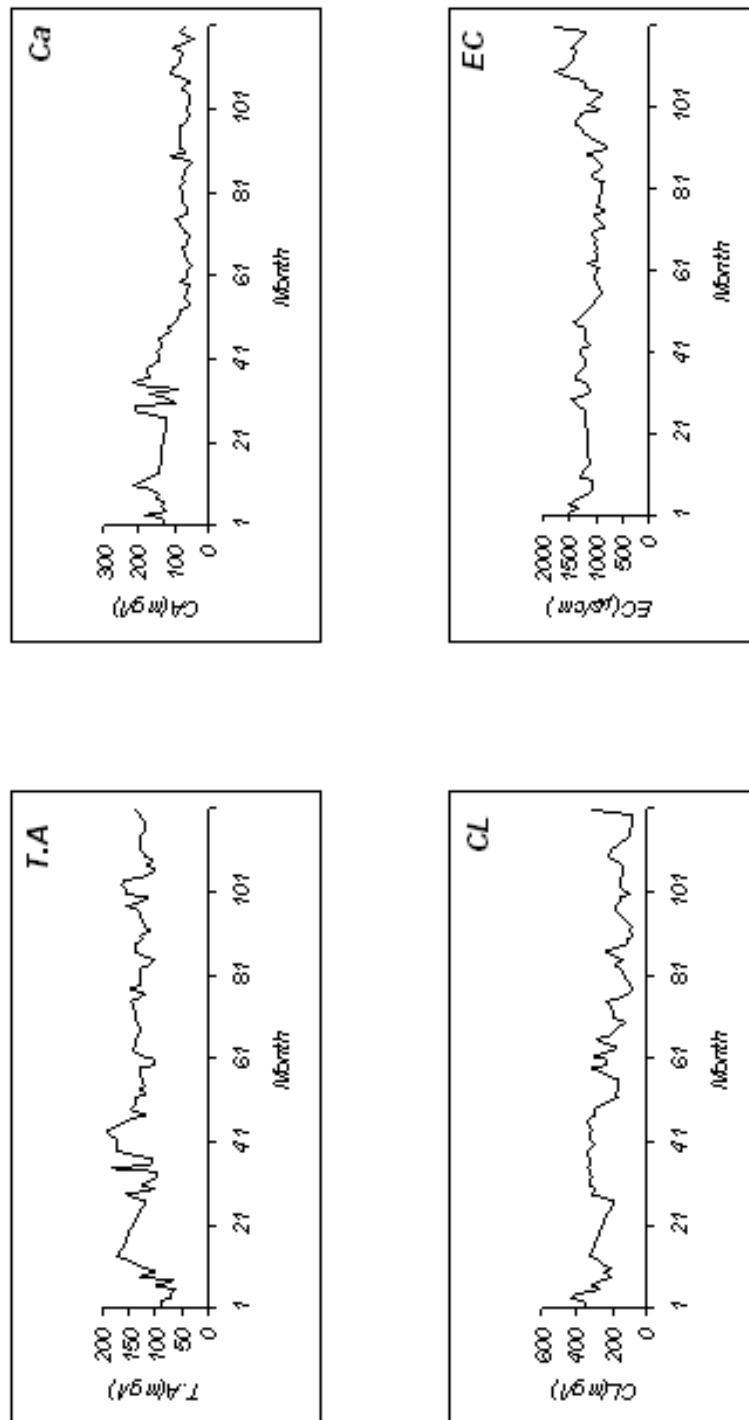


Fig. 5.1: Monthly observed values of T.A, Ca, Cl and EC of AL-Hindya station from 1992 to 2001.

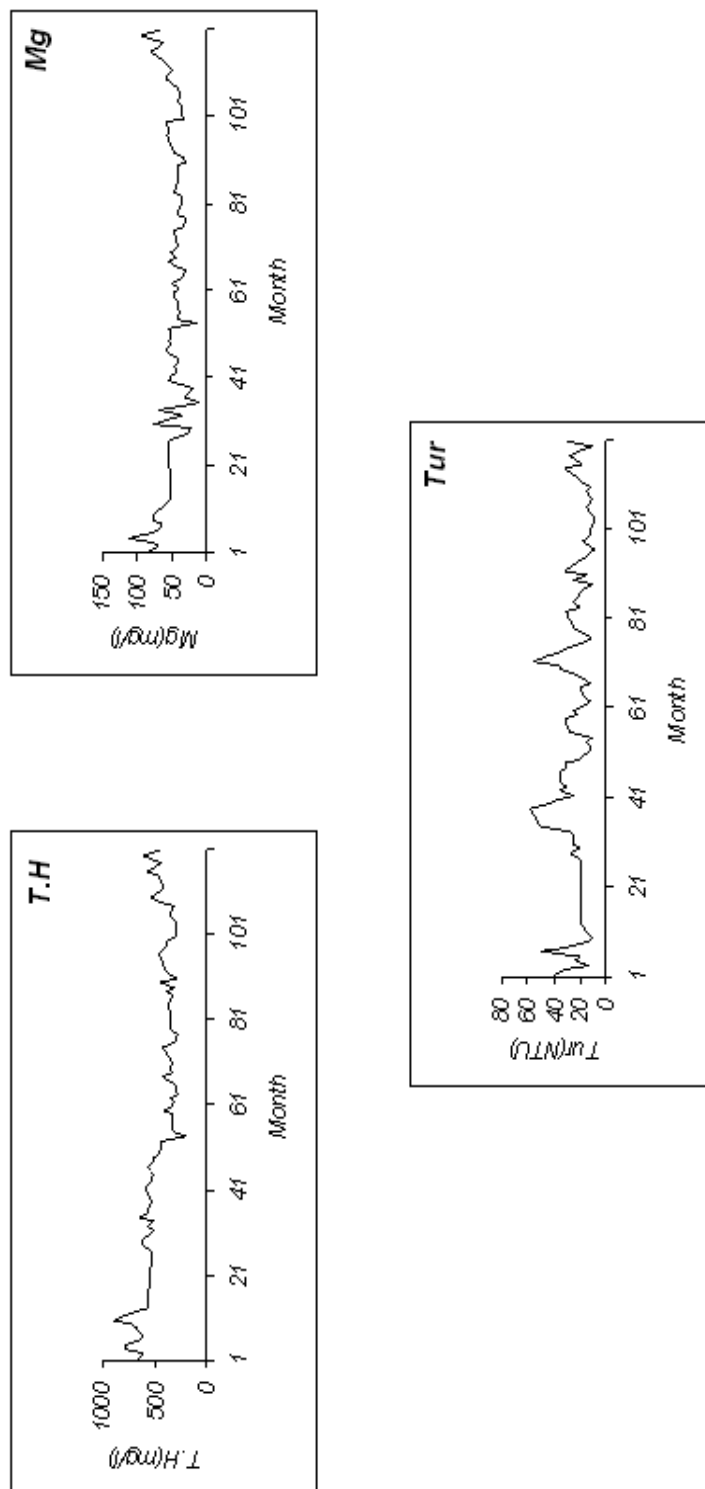


Fig. 5.2: Monthly observed values of T.H., Mg and Tur of AL-Hindya station from 1992 to 2001.

Figures (5.3) &(5.4) show the (ACF) and (PACF) of each water quality parameter in AL-Hindya station and the other related figures are given in the appendix (B)(figures from B.27 to B.30). Chatfield in 1982 explained that the (ACF) of moving average (MA(q)) cuts off after lag q whereas the (PACF) is a mixture of damped exponentials sinusoids and dies out slowly (or attenuates), then, these figures indicate that the model is seasonal multiplicative model(seasonal moving average model) of order $(0,1,1) \times (0,1,1)_{12}$ for all water quality parameter except (Ca, EC and Mg) at AL-Hindya station, (T.A and T.H) at Kufa station and (Ca, Cl) at Diwaniya station. Two types of model $(0,1,1) \times (0,1,1)_{12}$ and $(0,1,2) \times (0,1,1)_{12}$ are suggested, and the selected model is the one which give minimum sum of squares(SS).

5.2 Estimation of Parameters

The unconditional sum of square is used to estimate the model parameters. To illustrate this procedure the calculation of (EC) parameter of Kufa station is given in tables (5.1)& (5.2) .It is assumed that $(\Theta=0.1)$ and $(\theta=0.1)$ as a trial values and the sum of squares(SS) is computed as shown in the table (5.1), i.e. $SS= 1.591$.The procedure is repeated with different values of Θ and θ until obtaining the minimum sum of squares. Now if the model is

$$\nabla \nabla_{12} Z_t = W_t = (1 - \theta\beta)(1 - \Theta\beta^{12})a_t \quad \dots\dots\dots(5-1)$$

and the number of observations is $N=120$, $n=N-d-SD =120-1-12= 107$ which represent the number of W_t values.

where Z_t is the transformed series ($Z_t = \ln X_t$).

X_t is the original series.

d is the degree of simple difference ($d=1$).

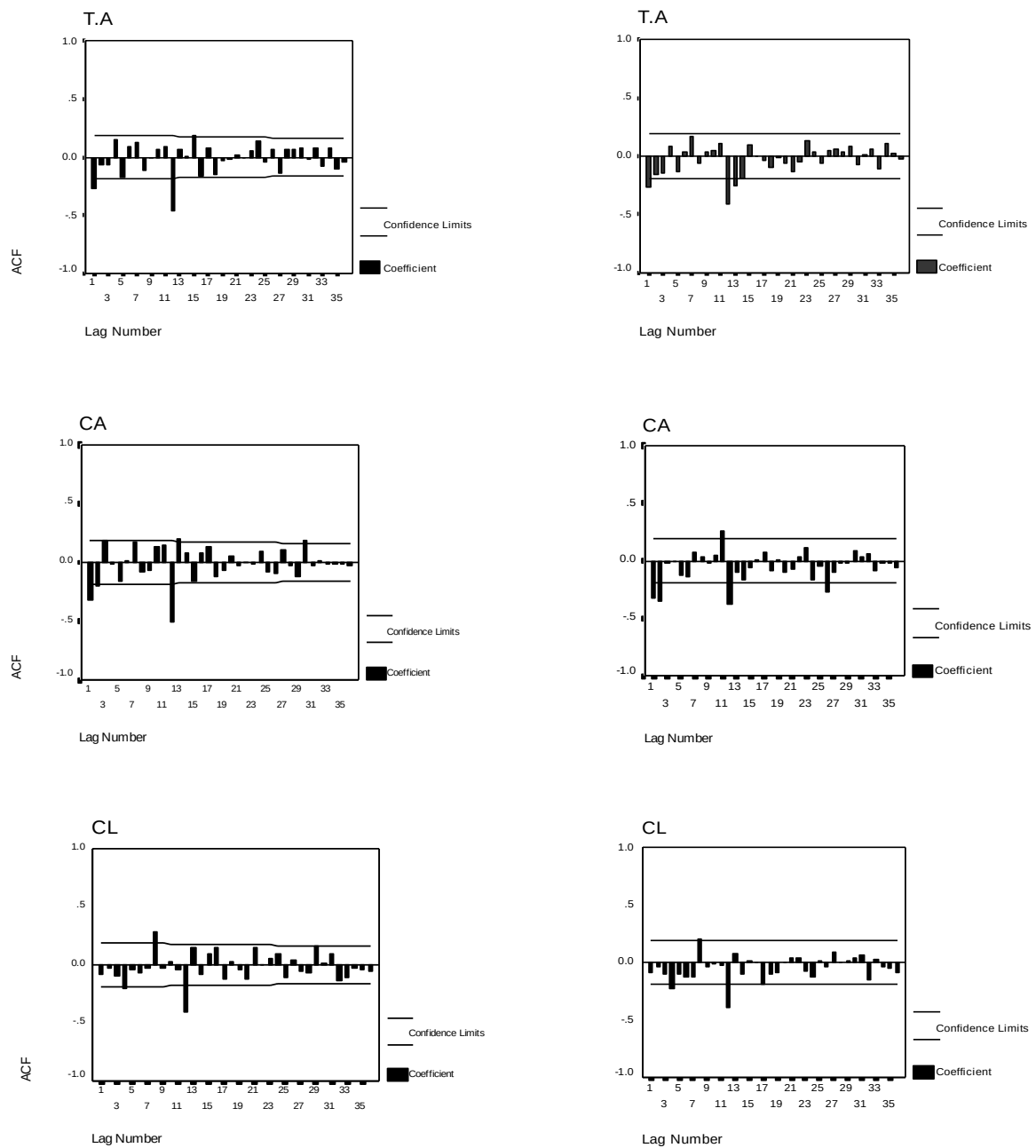


Fig. 5.3:ACF&PACF T.A, Ca and Cl of AL-Hindya station .

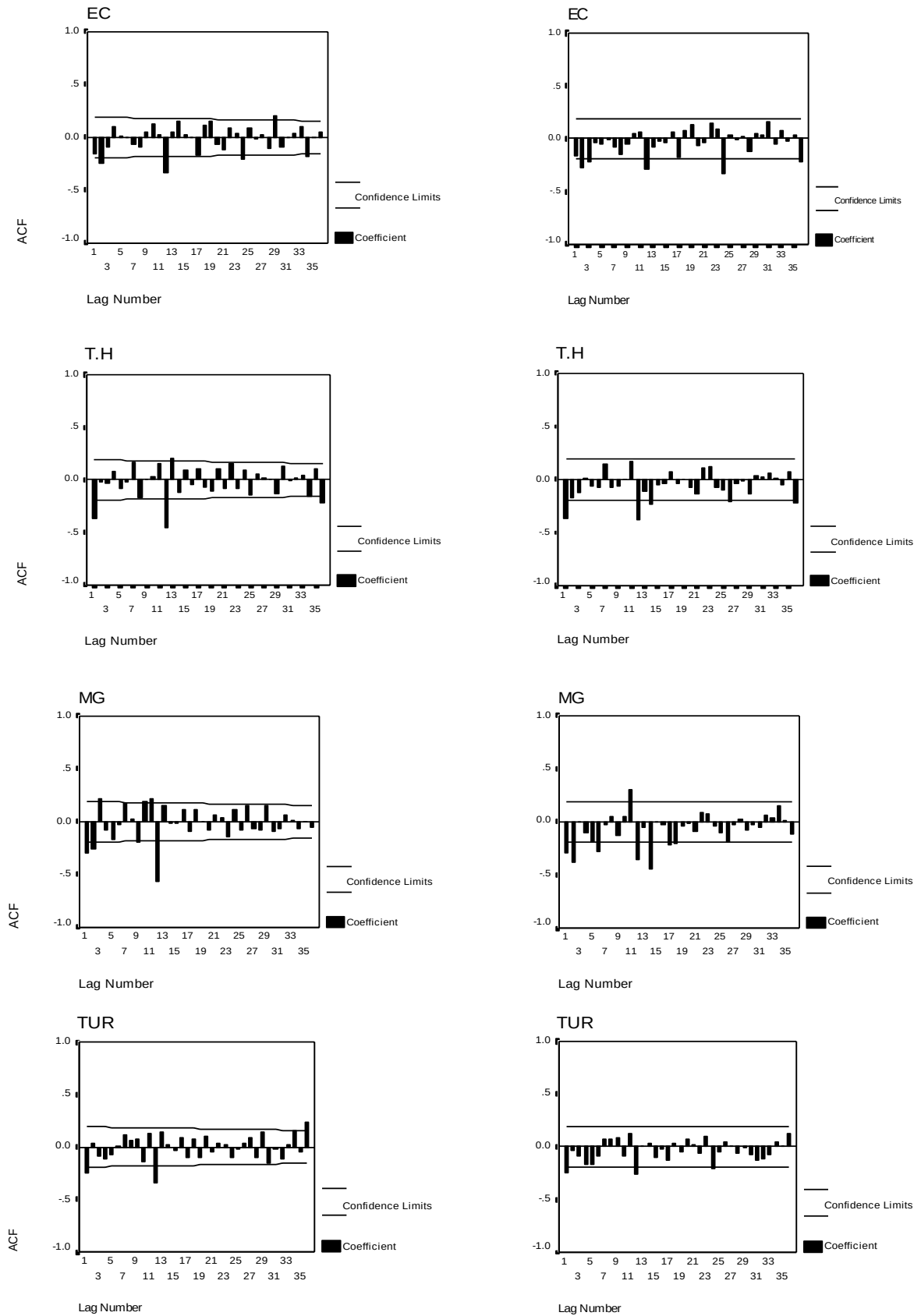


Fig. 5.4:ACF&PACF EC, T.H, Mg and Cl of AL-Hindya station .

D is the degree of seasonal difference (D=1).

S is the length of periodic cycle (S=12).

Then the model may be written in either forward form

$$[a_t] = [W_t] + \theta[a_{t-1}] + \Theta[a_{t-12}] - \theta\Theta[a_{t-13}] \quad \dots\dots\dots(5-2)$$

or backward form

$$[e_t] = [W_t] + \theta[e_{t+1}] + \Theta[e_{t+12}] - \theta\Theta[e_{t+13}] \quad \dots\dots\dots(5-3)$$

and it is convenient to use a numbering system so that the first observation in the X_t series (the second column in table (5.1)) has a subscript (-12), so that the first observation in the W_t series has a subscript (1) and the last has subscript (107). Accordingly, in table (5.1) the observation designated as $X_{-12}, X_{-11}, \dots, X_{107}$. The W 's obtained by differencing, designated as $W_1, W_2, \dots, W_{107}$ (the eighth column in table (5.1)). The beginning of calculation of the unconditional sum of squares (SS) i.e. when $t=107$, backward form (equation (5-3)) is used to compute the $[e_t]$'s (column 12 in table (5.1)) for $t=107, 106, \dots, 1$ by setting $[e_t]=0$ for $t=108, 109, \dots$, then the W_t series for $t=0, -1, -2, \dots, -12$ is computed by equation (5-3) by setting $[e_t]=0$ for $t=0, -1, -2, \dots, -12$. After this a_t 's (in fourth column) is computed by forward form (equation (5-2)) for $t=-12, -11, \dots, -107$ by setting $[a_t]$'s for $t=-13, -14, \dots$. The sum of squares (SS) is computed by

$$SS = \sum_{t=-12}^{t=107} a_t^2 \quad \dots\dots\dots(5-4)$$

Table (5.2) shows that the minimum (SS=0.992) when $\Theta=0.895$ and $\theta=0.550$ and table (5.3) summarizes the estimation of model parameters in the three stations. Table (5.4) shows the model equation in the three stations.

22	1639	7.40184	-0.08288	0.01251	0.00628	-0.00077	-0.10244	-0.00117	0.01063	-0.01144	-0.10207	0.00869
23	1437	7.27031	-0.11445	-0.00829	-0.00096	0.00063	-0.10457	0.00067	-0.01175	0.00263	-0.11436	0.01306
24	1448	7.27794	0.00347	-0.01144	-0.00262	-9.6E-05	0.01743	0.00031	0.00668	0.00248	0.02629	1.2E-05
25	1397	7.24208	0.01271	0.00035	-0.00604	-0.00026	0.01814	-0.00076	0.00305	0.00282	0.02477	0.000162
26	1467	7.28413	0.0301	0.00127	-0.01382	-0.0006	0.04205	-0.00047	-0.00756	-0.00675	0.02821	0.000906
27	1395	7.24065	-0.03578	0.00301	0.01297	-0.00138	-0.05315	0.00015	-0.00468	-0.00949	-0.06747	0.00128
28	1392	7.2385	-0.09891	-0.00358	0.02039	0.0013	-0.11443	-0.00058	0.00147	0.01742	-0.09495	0.009784
29	1450	7.27932	0.17277	-0.00989	0.00304	0.00204	0.18166	0.00252	-0.00583	0.00091	0.17422	0.029651
30	1227	7.11233	-0.03706	0.01728	-0.01744	0.0003	-0.03658	-0.00265	0.02521	0.0178	0.00908	0.001373
31	1538	7.33824	0.21863	-0.00371	-0.0029	-0.00174	0.2235	0.00109	-0.02652	-0.01788	0.178	0.0478
32	1425	7.26193	-0.1538	0.02186	0.00837	-0.00029	-0.18432	-0.00068	0.01094	-0.0061	-0.1788	0.023653
33	1579	7.36455	-0.06925	-0.01538	0.01251	0.00084	-0.06554	-0.00089	-0.00676	0.01063	-0.06099	0.004795
34	1792	7.49109	0.1113	-0.00692	-0.00829	0.00125	0.12776	0.00289	-0.00686	-0.01175	0.10626	0.012387
35	1347	7.20564	-0.15341	0.01113	-0.01144	-0.00083	-0.15392	-0.00086	0.02891	0.00668	-0.11747	0.023535
36	1457	7.28413	0.05702	-0.01534	0.00035	-0.00114	0.07087	-0.00152	-0.0086	0.00305	0.06884	0.003252
37	1483	7.30182	0.06048	0.0057	0.00127	3.5E-05	0.05354	0.00027	-0.01518	-0.00756	0.03053	0.003658
38	1435	7.26892	-0.06802	0.00605	0.00301	0.00013	-0.07495	-0.00131	0.00273	-0.00468	-0.0756	0.004359
39	1327	7.19068	-0.04524	-0.0066	-0.00358	0.0003	-0.03476	0.00047	-0.0131	0.00147	-0.04685	0.002047
40	1346	7.20489	0.00231	-0.00452	-0.00989	-0.00036	0.01637	0.00046	0.00466	-0.00583	0.01474	5.34E-06
41	1283	7.15696	-0.07026	0.00023	0.01728	-0.00099	-0.08876	-0.00068	0.00459	0.02521	-0.05828	0.004937
42	1445	7.27586	0.27344	-0.00703	-0.00371	0.00173	0.2859	0.00048	-0.00676	-0.02652	0.25214	0.07477
43	1367	7.22037	-0.23182	0.02734	0.02186	-0.00037	-0.2814	-0.00041	0.00483	0.01094	-0.26522	0.053742
44	1427	7.26333	0.07852	-0.02318	-0.01538	0.00219	0.11927	-0.00101	-0.00409	-0.00676	0.10943	0.006165
45	1503	7.31522	-0.04827	0.00785	-0.00692	-0.00154	-0.05073	-0.0001	-0.01012	-0.00686	-0.06761	0.00233
46	1551	7.34666	-0.08811	-0.00483	0.01113	-0.00069	-0.0951	0.00141	-0.00103	0.02891	-0.06863	0.007763
47	1548	7.34472	0.25825	-0.00881	-0.01534	0.00111	0.28352	-0.00011	0.01412	-0.0086	0.28914	0.066694
48	1563	7.35436	-0.03579	0.02583	0.0057	-0.00153	-0.06886	0.00094	-0.00106	-0.01518	-0.08604	0.001281
49	1347	7.20564	-0.16452	-0.00358	0.00605	0.00057	-0.16641	-0.00244	0.00945	0.00273	-0.1518	0.027066
50	1394	7.23993	0.04354	-0.01645	-0.0066	0.0006	0.0672	0.00245	-0.02435	-0.0131	0.0273	0.001896
51	1100	7.00307	-0.15813	0.00435	-0.00452	-0.00066	-0.15862	0.00151	0.02451	0.00466	-0.13096	0.025006
52	1142.67	7.04112	0.00871	-0.01581	0.00023	-0.00045	0.02384	-0.00308	0.01509	0.00459	0.04659	7.38E-05
53	1185.33	7.07778	0.07842	0.00087	-0.00703	2.3E-05	0.0846	0.00118	-0.03078	-0.00676	0.04688	0.006149
54	1228	7.11314	-0.04766	0.00784	0.02734	-0.0007	-0.08355	0.00065	0.01179	0.00483	-0.06757	0.002271
55	1217	7.10414	0.01581	-0.00477	-0.02318	0.00273	0.04649	0.00063	0.00654	-0.00409	0.04831	0.00025
56	1224	7.10988	-0.02547	0.00158	0.00785	-0.00232	-0.03722	-0.00014	0.00634	-0.01012	-0.04086	0.000549
57	1167	7.06219	-0.10774	-0.00255	-0.00483	0.00079	-0.09958	-0.00079	-0.00141	-0.00103	-0.10123	0.011607
58	1184	7.07665	-0.03608	-0.01077	-0.00881	-0.00048	-0.01697	-0.00042	-0.00786	0.01412	-0.01029	0.001301

□

59	1369	7.22184	0.17022	-0.00361	0.02583	-0.00088	0.14712	0.00068	-0.00418	-0.00106	0.1412	0.028974
60	1347	7.20564	-0.01498	0.01702	-0.00358	0.00258	-0.02584	0.00097	0.00676	0.00945	-0.01061	0.000225
61	1298	7.16858	0.09408	-0.0015	-0.01645	-0.00036	0.11167	0.00257	0.00974	-0.02435	0.09448	0.008851
62	1002	6.90975	-0.27772	0.00941	0.00435	-0.00165	-0.29312	0.00065	0.02574	-0.02451	-0.24353	0.077127
63	987	6.89467	0.17776	-0.02777	-0.01581	0.00044	0.22178	-0.0017	0.00654	0.01509	0.24512	0.0316
64	1254	7.13409	0.2216	0.01778	0.00087	-0.00158	0.20137	0.00266	-0.01705	-0.03078	0.15089	0.049106
65	917	6.82111	-0.31973	0.02216	0.00794	8.7E-05	-0.34965	-0.00348	0.02661	0.01179	-0.30776	0.102228
66	1100	7.00307	0.10907	-0.03197	-0.00477	0.00078	0.1466	0.00035	-0.03484	0.00654	0.11794	0.011897
67	1152	7.04925	0.06815	0.01091	0.00158	-0.00048	0.05519	-0.00038	0.00351	0.00634	0.06542	0.004646
68	1242	7.12448	0.0736	0.00682	-0.00255	0.00016	0.06949	0.00085	-0.00381	-0.00141	0.06343	0.006417
69	1167.5	7.06262	-0.01733	0.00736	-0.01077	-0.00025	-0.01417	0.00049	0.00845	-0.00786	-0.01406	0.0003
70	1093	6.99668	-0.08466	-0.00173	0.00361	-0.00108	-0.0804	-0.00115	0.00486	-0.00418	-0.07858	0.007168
71	1217	7.10414	-0.0288	-0.00847	0.01702	-0.00036	-0.03772	-0.00062	-0.0115	0.00676	-0.04184	0.00083
72	1275	7.1507	0.05668	-0.00288	-0.0015	0.0017	0.06276	-0.00133	-0.00621	0.00974	0.06762	0.003212
73	1336	7.19744	0.09901	0.00567	0.00941	-0.00015	0.08379	-0.00117	-0.01331	0.02574	0.09739	0.009804
74	1343	7.20266	0.24524	0.0099	-0.02777	0.00094	0.26405	0.00149	-0.01173	0.00654	0.25738	0.060143
75	1415	7.25488	0.11238	0.02452	0.01778	-0.00278	0.06731	-0.0003	0.01488	-0.01705	0.06543	0.01263
76	1481	7.30047	-0.16222	0.01124	0.02216	0.00178	-0.19384	0.00028	-0.00295	0.02661	-0.17046	0.026314
77	1461	7.28688	0.24898	-0.01622	-0.03197	0.00222	0.29939	0.00131	0.00285	-0.03484	0.26609	0.061991
78	1214	7.10168	-0.32816	0.0249	0.01091	-0.0032	-0.36716	-0.00215	0.01306	0.00351	-0.34844	0.107686
79	1352	7.20934	0.03438	-0.03282	0.00682	0.00109	0.06147	0.0011	-0.02146	-0.00381	0.03511	0.001182
80	1375	7.22621	-0.04824	0.00344	0.00736	0.00068	-0.05835	-0.00082	0.011	0.00845	-0.03808	0.002327
81	1411	7.25205	0.08041	-0.00482	-0.00173	0.00074	0.0877	-0.00018	-0.00823	0.00486	0.08451	0.006466
82	1405.67	7.24827	0.0619	0.00804	-0.00847	-0.00017	0.06215	0.00026	-0.00181	-0.0115	0.04858	0.003832
83	1400.33	7.24447	-0.10711	0.00619	-0.00286	-0.00085	-0.11126	0.00016	-0.00259	-0.00621	-0.11504	0.011472
84	1395	7.24065	-0.05513	-0.01071	0.00567	-0.00029	-0.05037	1.5E-05	0.00161	-0.01331	-0.0621	0.003039
85	1297	7.16781	-0.11575	-0.00551	0.0099	0.00057	-0.11957	0.00198	0.00015	-0.01173	-0.13313	0.013399
86	1118	7.0193	-0.14178	-0.01158	0.02452	0.00099	-0.15374	-0.0018	0.01981	0.01488	-0.11725	0.020101
87	1396	7.24137	0.16445	-0.01418	0.01124	0.00245	0.16985	0.00017	-0.01797	-0.00295	0.14876	0.027046
88	1412	7.25276	-0.03509	0.01645	-0.01622	0.00112	-0.03419	-0.00017	0.00167	0.00285	-0.0295	0.001231
89	1417	7.2563	0.04014	-0.00351	0.0249	-0.00162	0.01713	-1E-05	-0.00173	0.01306	0.02847	0.001611
90	1371	7.2233	0.12091	0.00401	-0.03282	0.00249	0.1522	-5.7E-06	-0.0001	-0.02146	0.13064	0.014619
91	1218	7.10497	-0.20718	0.01209	0.00344	-0.000328	-0.22599	-0.00044	-5.7E-05	0.011	-0.21461	0.042925
92	1401	7.24494	0.09722	-0.02072	-0.00482	0.00034	0.12311	0.0005	-0.0044	-0.00623	0.10998	0.009462
93	1321	7.18614	-0.0664	0.00972	0.00804	-0.00048	-0.08461	0.00084	0.00502	-0.00181	-0.06228	0.004409
94	1277	7.15227	-0.03134	-0.00664	0.00619	0.0008	-0.03009	-0.00096	0.00843	0.00259	-0.01811	0.000982
95	1316	7.18235	0.01942	-0.00313	-0.01071	0.00062	0.03368	0	-0.00956	0.00161	0.02593	0.000377

96	1332	7.19444	0.0134	0.00194	-0.00551	-0.00107	0.0159	0	0.00015	0.01605	0.00018
97	1216	7.10332	-0.02796	0.00134	-0.01158	-0.00055	-0.01827	0	0.01981	0.00154	0.000782
98	1301	7.17089	0.20026	-0.0028	-0.01418	-0.00116	0.21608	0	-0.01797	0.19811	0.040105
99	1355	7.21156	-0.14351	0.02003	0.01645	-0.00142	-0.1814	0	0.00167	-0.17973	0.020595
100	1396	7.24137	-0.00109	-0.01435	-0.00351	0.00164	0.01841	0	-0.00173	0.01668	1.19 E-05
101	1377	7.22766	-0.01298	-0.00011	0.00401	-0.00035	-0.01724	0	-0.0001	-0.01734	0.000169
102	1331	7.19369	0.00942	-0.0013	0.01209	0.0004	-0.00098	0	-5.7 E-05	-0.00103	8.87 E-05
103	1187	7.07918	-0.01716	0.00094	-0.02072	0.00121	0.00383	0	-0.0044	-0.00057	0.000294
104	1300	7.17012	-0.03896	-0.00172	0.00972	-0.00207	-0.04904	0	0.00502	-0.04403	0.001518
105	1278	7.15305	0.03022	-0.0039	-0.00664	0.00097	0.04173	0	0.00843	0.05016	0.000913
106	1357	7.21303	0.09441	0.00302	-0.00313	-0.00066	0.09386	0	-0.00956	0.0843	0.000913
107	1271	7.14756	-0.08386	0.00944	0.00194	-0.00031	-0.09556	0	0	-0.09556	0.007032
SS=1.591											

Table 5.2: Calculation of sum of squares of EC of Kufa station with $\theta=0.550, \Theta=0.895$. Calculations start from the bottom of the table.

t	z_t	$\ln z_t$	a_t	∂a_{t-1}	∂a_{t-2}	∂a_{t-3}	w_t	$\partial \partial a_{t-1,2}$	$\partial \partial a_{t-2,2}$	$\partial \partial a_{t-3,2}$	e_t	a_t^2
-12	1453	7.281	-0.09				-0.09	-0.09	0	0	0	0.008028
-11	1562	7.354	0.067	-0.049			0.117	-0.046	-0.163	0	0	0.00452
-10	1768	7.478	0.127	0.037			0.09	0.006	-0.084	0	0	0.016245
-9	1903	7.551	0.146	0.07			0.076	0.087	0.011	0	0	0.021418
-8	1848	7.522	0.023	0.08			-0.058	0.101	0.159	0	0	0.000526
-7	1448	7.278	-0.16	0.013			-0.173	0.011	0.184	0	0	0.025651
-6	1327	7.191	-0.145	-0.088			-0.057	-0.037	0.02	0	0	0.021141
-5	1427	7.263	0.011	-0.08			0.091	0.024	-0.067	0	0	0.000121
-4	1352	7.209	-0.096	0.006			-0.102	-0.059	0.043	0	0	0.009212
-3	1537	7.338	0.046	-0.053			0.099	-0.009	-0.108	0	0	0.002151
-2	1584	7.368	0.014	0.026			-0.012	-0.027	-0.015	0	0	0.000187
-1	1567	7.357	0.042	0.008			0.034	-0.016	-0.05	0	0	0.001724
0	1593	7.373	0.045	0.023	-0.08		0.102	-0.027	-0.029	-0.1	0	0.002015
1	1587	7.37	0.053	0.025	0.06	-0.044	-0.076	0.006	-0.048	-0.052	-0.182	0.002791
2	1813	7.503	0.119	0.029	0.114	0.033	0.009	0.121	0.011	0.007	-0.094	0.014235
3	1550	7.346	-0.096	0.066	0.131	0.063	-0.23	0.075	0.22	0.097	0.012	0.009296
4	1425	7.262	-0.159	-0.053	0.021	0.072	-0.055	0.017	0.136	0.113	0.177	0.025361
5	1238	7.121	-0.139	-0.088	-0.143	0.011	0.103	-0.058	0.032	0.012	0.205	0.019314
6	1317	7.183	0.021	-0.076	-0.13	-0.079	0.149	-0.021	-0.106	-0.041	0.023	0.000456
7	1327	7.191	0.028	0.012	0.01	-0.072	-0.065	-0.002	-0.038	0.026	-0.075	0.000786
8	1367	7.22	0.008	0.015	-0.086	0.005	0.084	-0.033	-0.003	-0.066	0.048	6.07 E-05
9	1427	7.263	0.008	0.004	0.042	-0.047	-0.085	-0.035	-0.06	-0.01	-0.12	5.97 E-05
10	1579	7.365	0.065	0.004	0.012	0.023	0.071	-0.006	-0.063	-0.031	-0.017	0.004193
11	1537	7.338	0.05	0.036	0.037	0.007	-0.016	0.012	-0.01	-0.018	-0.056	0.002489
12	1522	7.328	0.021	0.027	0.04	0.02	-0.026	-0.003	0.021	-0.03	-0.032	0.000438
13	1442	7.274	-0.014	0.012	0.047	0.022	-0.05	0.005	-0.005	0.007	-0.054	0.000183
14	1442	7.274	-0.06	-0.007	0.107	0.026	-0.133	-7E-04	0.009	0.135	0.012	0.003574
15	1456	7.283	-0.011	-0.033	-0.086	0.059	0.166	0.003	-0.001	0.084	0.246	0.000132
16	1629	7.396	0.095	-0.006	-0.143	-0.047	0.196	0.069	0.006	0.019	0.152	0.009015
17	1415	7.255	0.006	0.052	-0.124	-0.078	-2E-04	0.025	0.126	-0.065	0.035	3.63 E-05
18	1242	7.124	-0.101	0.003	0.019	-0.068	-0.192	-0.051	0.046	-0.023	-0.118	0.010295
19	1245	7.127	-0.046	-0.056	0.025	0.011	-0.005	-0.058	-0.093	-0.002	-0.043	0.00215
20	1387	7.235	0.046	-0.026	0.007	0.014	0.078	-0.06	-0.105	-0.037	-0.003	0.002115
21	1641	7.403	0.154	0.025	0.007	0.004	0.125	0.044	-0.11	-0.039	-0.067	0.023587

22	1639	7.402	0.036	0.084	0.058	0.004	-0.102	0.042	0.08	-0.006	-0.071	0.001309
23	1437	7.27	-0.072	0.02	0.045	0.032	-0.105	-0.004	0.076	0.013	-0.011	0.005166
24	1448	7.278	-0.028	-0.04	0.019	0.025	0.017	-0.016	-0.007	-0.003	0.024	0.000779
25	1397	7.242	-0.02	-0.015	-0.012	0.01	0.018	-3E-04	-0.03	0.006	-0.006	0.000384
26	1467	7.284	-0.016	-0.011	-0.054	-0.007	0.042	0.031	-6E-04	-7E-04	0.01	0.000243
27	1395	7.241	-0.043	-0.009	-0.01	-0.029	-0.053	0.007	0.056	0.004	-0.001	0.001813
28	1392	7.238	-0.047	-0.023	0.085	-0.006	-0.114	-0.03	0.013	0.077	0.006	0.00223
29	1460	7.279	0.114	-0.026	0.005	0.047	0.182	0.014	-0.055	0.028	0.141	0.01308
30	1227	7.112	-0.067	0.063	-0.091	0.003	-0.037	-0.12	0.026	-0.057	0.051	0.004551
31	1538	7.338	0.195	-0.037	-0.041	-0.05	0.223	0.046	-0.217	-0.064	-0.104	0.037954
32	1425	7.262	-0.013	0.107	0.041	-0.023	-0.184	-0.051	0.084	-0.068	-0.117	0.000174
33	1579	7.365	0.042	-0.007	0.137	0.023	-0.066	0.013	-0.093	0.049	-0.123	0.001766
34	1792	7.491	0.108	0.023	0.032	0.076	0.128	0.109	0.024	0.047	0.089	0.011599
35	1347	7.206	-0.177	0.059	-0.064	0.018	-0.154	-0.046	0.199	-0.004	0.085	0.031265
36	1457	7.284	-0.016	-0.097	-0.025	-0.035	0.071	-0.021	-0.081	-0.018	-0.008	0.000256
37	1483	7.302	0.041	-0.009	-0.018	-0.014	0.054	0.049	-0.038	-4E-04	-0.033	0.001675
38	1435	7.269	-0.057	0.023	-0.014	-0.01	-0.075	0.048	0.089	0.034	-7E-04	0.00322
39	1327	7.191	-0.096	-0.031	-0.038	-0.008	-0.035	-6E-04	0.088	0.008	0.062	0.009295
40	1346	7.205	-0.058	-0.053	-0.042	-0.021	0.016	-0.033	-0.001	-0.034	0.015	0.00336
41	1283	7.157	0.005	-0.032	0.102	-0.023	-0.089	-0.072	-0.061	0.016	-0.062	2.45E-05
42	1445	7.276	0.172	0.003	-0.06	0.056	0.286	-0.008	-0.131	-0.134	0.029	0.029577
43	1367	7.22	0.021	0.095	0.174	-0.033	-0.281	-9E-04	-0.014	0.051	-0.243	0.00043
44	1427	7.263	0.023	0.011	-0.012	0.096	0.119	-0.033	-0.002	-0.057	0.093	0.000529
45	1503	7.315	0.006	0.013	0.038	-0.006	-0.051	0.008	-0.06	0.015	-0.104	3.64E-05
46	1551	7.347	-0.016	0.003	0.096	0.021	-0.095	0.015	0.015	0.122	0.027	0.000258
47	1548	7.345	0.063	-0.009	-0.158	0.053	0.284	0.039	0.028	-0.05	0.222	0.004024
48	1563	7.354	0.039	0.035	-0.014	-0.087	-0.069	0.07	0.071	-0.023	-0.091	0.001498
49	1347	7.206	-0.101	0.021	0.037	-0.008	-0.166	0.058	0.127	0.054	-0.042	0.010124
50	1394	7.24	-0.059	-0.055	-0.051	0.02	0.067	0.127	0.105	0.054	0.099	0.003489
51	1100	7.003	-0.249	-0.032	-0.086	-0.028	-0.159	-0.026	0.232	-7E-04	0.098	0.052241
52	1142.667	7.041	-0.118	-0.137	-0.052	-0.047	0.024	-0.06	-0.047	-0.037	-0.001	0.013882
53	1185.333	7.078	0.053	-0.065	0.004	-0.029	0.085	-0.036	-0.108	-0.081	-0.068	0.002782
54	1228	7.113	0.097	0.029	0.154	0.002	-0.084	-0.012	-0.066	-0.008	-0.146	0.009399
55	1217	7.104	0.034	0.053	0.019	0.085	0.046	0.039	-0.021	-1E-03	-0.015	0.00114
56	1224	7.11	-0.008	0.019	0.021	0.01	-0.037	-4E-04	0.072	-0.037	-0.002	6.82E-05
57	1167	7.062	-0.11	-0.005	0.005	0.011	-0.1	-0.024	-7E-04	0.009	-0.067	0.012109

58	1184	7.077	-0.095	-0.061	-0.014	0.003	-0.017	-0.061	-0.044	0.017	0.016	0.008997
59	1369	7.222	0.16	-0.052	0.057	-0.008	0.147	0.049	-0.111	0.044	0.031	0.025482
60	1347	7.206	0.065	0.088	0.035	0.031	-0.026	0.062	0.09	0.078	0.08	0.004276
61	1298	7.169	0.039	0.036	-0.09	0.019	0.112	0.147	0.114	0.065	0.142	0.001485
62	1002	6.91	-0.275	0.021	-0.053	-0.05	-0.293	2E-04	0.268	0.142	0.117	0.075784
63	987	6.895	-0.124	-0.151	-0.223	-0.029	0.222	-0.066	3E-04	-0.029	0.259	0.015339
64	1254	7.134	0.151	-0.068	-0.105	-0.123	0.201	0.068	-0.12	-0.067	-0.053	0.022564
65	917	6.821	-0.162	0.083	0.047	-0.058	-0.35	-0.146	0.123	-0.041	-0.121	0.026137
66	1100	7.003	0.118	-0.089	0.087	0.026	0.147	-0.058	-0.266	-0.013	-0.074	0.014041
67	1152	7.049	0.103	0.065	0.03	0.048	0.055	0.018	-0.106	0.044	-0.024	0.010584
68	1242	7.124	0.102	0.057	-0.007	0.017	0.069	0.021	0.032	-5E-04	0.08	0.010418
69	1167.5	7.063	-0.052	0.056	-0.098	-0.004	-0.014	-0.003	0.038	-0.027	-8E-04	0.002752
70	1093	6.997	-0.14	-0.029	-0.085	-0.054	-0.08	-0.104	-0.005	-0.068	-0.05	0.019601
71	1217	7.104	0.075	-0.077	0.143	-0.047	-0.038	-0.048	-0.189	0.055	-0.124	0.005597
72	1275	7.151	0.084	0.041	0.059	0.079	0.063	-0.055	-0.088	0.07	0.1	0.007038
73	1336	7.197	0.132	0.046	0.034	0.032	0.084	0.021	-0.1	0.165	0.127	0.01749
74	1343	7.203	0.071	0.073	-0.246	0.019	0.264	0.003	0.039	2E-04	0.3	0.005104
75	1415	7.255	0.131	0.039	-0.111	-0.135	0.067	-0.001	0.006	-0.074	4E-04	0.017213
76	1481	7.3	0.074	0.072	0.135	-0.061	-0.194	0.014	-0.002	0.076	-0.134	0.005475
77	1461	7.287	0.121	0.041	-0.145	0.074	0.299	0.024	0.025	-0.163	0.138	0.01472
78	1214	7.102	-0.115	0.067	0.106	-0.08	-0.367	-0.092	0.043	-0.065	-0.297	0.013167
79	1352	7.209	0.032	-0.063	0.092	0.058	0.061	0.032	-0.167	0.02	-0.118	0.00103
80	1375	7.226	3E-05	0.018	0.091	0.051	-0.058	-0.013	0.058	0.023	0.036	8.54E-10
81	1411	7.252	-0.009	2E-05	-0.047	0.05	0.088	0.019	-0.024	-0.003	0.042	8.92E-05
82	1405.667	7.248	-0.043	-0.005	-0.125	-0.026	0.062	-0.014	0.034	-0.116	-0.006	0.001809
83	1400.333	7.244	0.001	-0.023	0.067	-0.069	-0.111	0.021	-0.026	-0.054	-0.211	1.4E-05
84	1395	7.241	-0.011	7E-04	0.075	0.037	-0.05	0.023	0.037	-0.062	-0.098	0.000131
85	1297	7.168	-0.049	-0.006	0.118	0.041	-0.12	0.058	0.042	0.024	-0.112	0.00238
86	1118	7.019	-0.182	-0.027	0.064	0.065	-0.154	-0.087	0.106	0.003	0.043	0.033012
87	1396	7.241	0.152	-0.1	0.117	0.035	0.17	0.004	-0.158	-0.001	0.006	0.023161
88	1412	7.253	0.051	0.084	0.066	0.065	-0.034	-0.009	0.007	0.016	-0.002	0.00252
89	1417	7.256	0.117	0.028	0.109	0.036	0.017	-0.001	-0.017	0.027	0.028	0.013798
90	1371	7.223	0.054	0.065	-0.103	0.06	0.152	-0.002	-0.003	-0.103	0.049	0.002952
91	1218	7.105	-0.111	0.03	0.029	-0.056	-0.226	-0.007	-0.003	0.036	-0.187	0.012295
92	1401	7.245	0.046	-0.061	3E-05	0.016	0.123	0.032	-0.012	-0.014	0.065	0.002149
93	1321	7.186	-0.068	0.025	-0.008	1E-05	-0.085	0.02	0.058	0.021	-0.026	0.004572

94	1277	7.152	-0.101	-0.037	-0.038	-0.005	-0.03	-0.047	0.037	-0.016	0.038	0.010141
95	1316	7.182	5E-04	-0.055	0.001	-0.021	0.034	0	-0.086	0.023	-0.029	2.36E-07
96	1332	7.194	0.005	3E-04	-0.01	6E-04	0.016	0	0	0.026	0.042	2.85E-05
97	1216	7.103	-0.053	0.003	-0.044	-0.006	-0.018	0	0	0.065	0.047	0.002848
98	1301	7.171	0.048	-0.029	-0.163	-0.024	0.216	0	0	-0.097	0.119	0.002315
99	1355	7.212	0.071	0.026	0.136	-0.089	-0.181	0	0	0.004	-0.177	0.004993
100	1396	7.241	0.028	0.039	0.046	0.075	0.018	0	0	-0.01	0.008	0.000796
101	1377	7.228	0.078	0.016	0.105	0.025	-0.017	0	0	-0.002	-0.019	0.006119
102	1331	7.194	0.033	0.043	0.049	0.058	-1E-03	0	0	-0.002	-0.003	0.001081
103	1187	7.079	-0.104	0.018	-0.099	0.027	0.004	0	0	-0.007	-0.004	0.010829
104	1300	7.17	-0.01	-0.057	0.041	-0.055	-0.049	0	0	0.035	-0.014	0.000105
105	1278	7.153	-0.047	-0.006	-0.061	0.023	0.042	0	0	0.023	0.064	0.00223
106	1357	7.213	0.011	-0.026	-0.09	-0.033	0.094	0	0	-0.053	0.041	0.000122
107	1271	7.148	-0.04	0.006	4E-04	-0.05	-0.096	0	0	0	-0.096	0.001561
SS=0.992												

Table 5.3: Estimated model parameters and sum of square of the three stations.

Parameter	AL-Hindya				Kufa				Diwaniya			
	θ_1	θ_2	θ	SS	θ_1	θ_2	θ	SS	θ_1	θ_2	θ	SS
T.A	0.401		0.925	2.592	0.631		0.898	1.018	0.748		0.875	2.629
					0.567	0.372	0.866	0.911				
Ca	0.550		0.875	5.393	0.601		0.850	1.884	0.774		0.849	4.344
	0.405	0.288	0.875	5.245					0.407	0.357	0.845	4.229
Cl	0.251		0.798	6.427	0.550		0.903	1.414	0.752		0.873	0.884
									0.512	0.253	0.877	0.821
EC	0.375		0.851	1.351	0.550		0.895	0.992	0.848		0.852	1.067
	0.318	0.261	0.827	1.264								
T.H	0.553		0.851	1.351	0.550		0.860	1.428	0.725		0.875	1.189
					0.405	0.251	0.860	1.343				
Mg	0.651		0.902	9.716	0.675		0.875	5.978	0.726		0.874	2.589
	0.452	0.294	0.885	9.246								
Tur	0.275		0.841	14.24	0.624		0.876	12.42	0.751		0.798	5.864

Table 5.4: Seasonal model equations of the three stations.

Par.	AL-Hindya	Kufa	Diwaniya
T.A	$W_t = (1 - 0.401\beta)(1 - 0.925\beta^{12})a_t$	$W_t = (1 - 0.567\beta - 0.372\beta^{12})(1 - 0.864\beta^{12})a_t$	$W_t = (1 - 0.748\beta)(1 - 0.875\beta^{12})a_t$
C.a	$W_t = (1 - 0.405\beta - 0.288\beta^{12})(1 - 0.875\beta^{12})a_t$	$W_t = (1 - 0.601\beta)(1 - 0.850\beta^{12})a_t$	$W_t = (1 - 0.407\beta - 0.357\beta^{12})(1 - 0.84\beta^{12})a_t$
CI	$W_t = (1 - 0.251\beta)(1 - 0.798\beta^{12})a_t$	$W_t = (1 - 0.550\beta)(1 - 0.903\beta^{12})a_t$	$W_t = (1 - 0.512\beta - 0.253\beta^{12})(1 - 0.87\beta^{12})a_t$
EC	$W_t = (1 - 0.318\beta - 0.261\beta^{12})(1 - 0.827\beta^{12})a_t$	$W_t = (1 - 0.550\beta)(1 - 0.895\beta^{12})a_t$	$W_t = (1 - 0.848\beta)(1 - 0.852\beta^{12})a_t$
T.H	$W_t = (1 - 0.553\beta)(1 - 0.851\beta^{12})a_t$	$W_t = (1 - 0.405\beta - 0.251\beta^{12})(1 - 0.84\beta^{12})a_t$	$W_t = (1 - 0.725\beta)(1 - 0.875\beta^{12})a_t$
Mg	$W_t = (1 - 0.452\beta - 0.294\beta^{12})(1 - 0.885\beta^{12})a_t$	$W_t = (1 - 0.675\beta)(1 - 0.875\beta^{12})a_t$	$W_t = (1 - 0.726\beta)(1 - 0.874\beta^{12})a_t$

5.3 Diagnostic Check

After estimating the model parameters, the diagnostic checking is applied to see if the model is adequate or not. Porte Manteau lack of fit test(sec.2.10) is used for this purpose. The results of this test are summarized in table (5.4) which indicate that all models in the three stations are adequate because the calculated Q less the χ^2 - table with degree of freedom (M-P-p-Q-q) where M=48.

Table 5.5: Diagnostic check results of the three station with K=48 & degree of freedom =(M-P-p-Q-q)at 95% confidence limits.

Parameter	AL-Hindya		Kufa		Diwaniya	
	Q-Calculated	χ^2 -Table	Q-Calculated	χ^2 -Table	Q-Calculated	χ^2 -Table
T.A	27.841	62.806	40.835	61.632	39.975	62.806
Ca	20.590	61.632	36.704	62.806	31.063	61.632
Cl	53.105	62.806	35.216	62.806	44.799	61.632
EC	30.698	61.632	28.846	62.806	39.734	62.806
T.H	25.789	62.806	49.499	61.632	35.125	62.806
Mg	32.542	61.632	56.801	62.806	38.602	62.806
Tur	29.837	62.806	23.670	62.806	44.963	62.806

5.4 Confidence Region

When several parameters have been considered simultaneously it needs some means of judging the precision of the estimate jointly. One of means of doing this is to determine the confidence region . The approximate

$1-\xi$ confidence region bounded by contour on the sum of squares surface obtained by (Box and Jenkins, 1976)

$$S(\beta) = S(\hat{\beta}) \left\{ 1 + \frac{\chi_{\xi}^2(k)}{n} \right\}, \dots \dots \dots (5-4)$$

where $S(\beta)$ is the quadratic surface .

$S(\hat{\beta})$ is the minimum sum of squares.

$\chi_{\xi}^2(k)$ is the significant point expected by a proportion ξ of the χ^2 distribution having k degree of freedom.

K is the no. of model parameters.

$n=107$.

It may be shown (Box-Jenkins, 1976) that a $1-\xi$ confidence region has the property that if repeated samples of size n are imagined to be drawn from the same population and a confidence region constructed from each such sample, then the proportion $1-\xi$ of these regions will include the true parameter point.

Figure(5.5) shows the 95% confidence region bounded by the contour line with $SS=1.05$ with $SS(\theta, \Theta)$ grids for EC of Kufa station and the other related figures are listed in the appendix (B)(figures from B.31 to B.50).

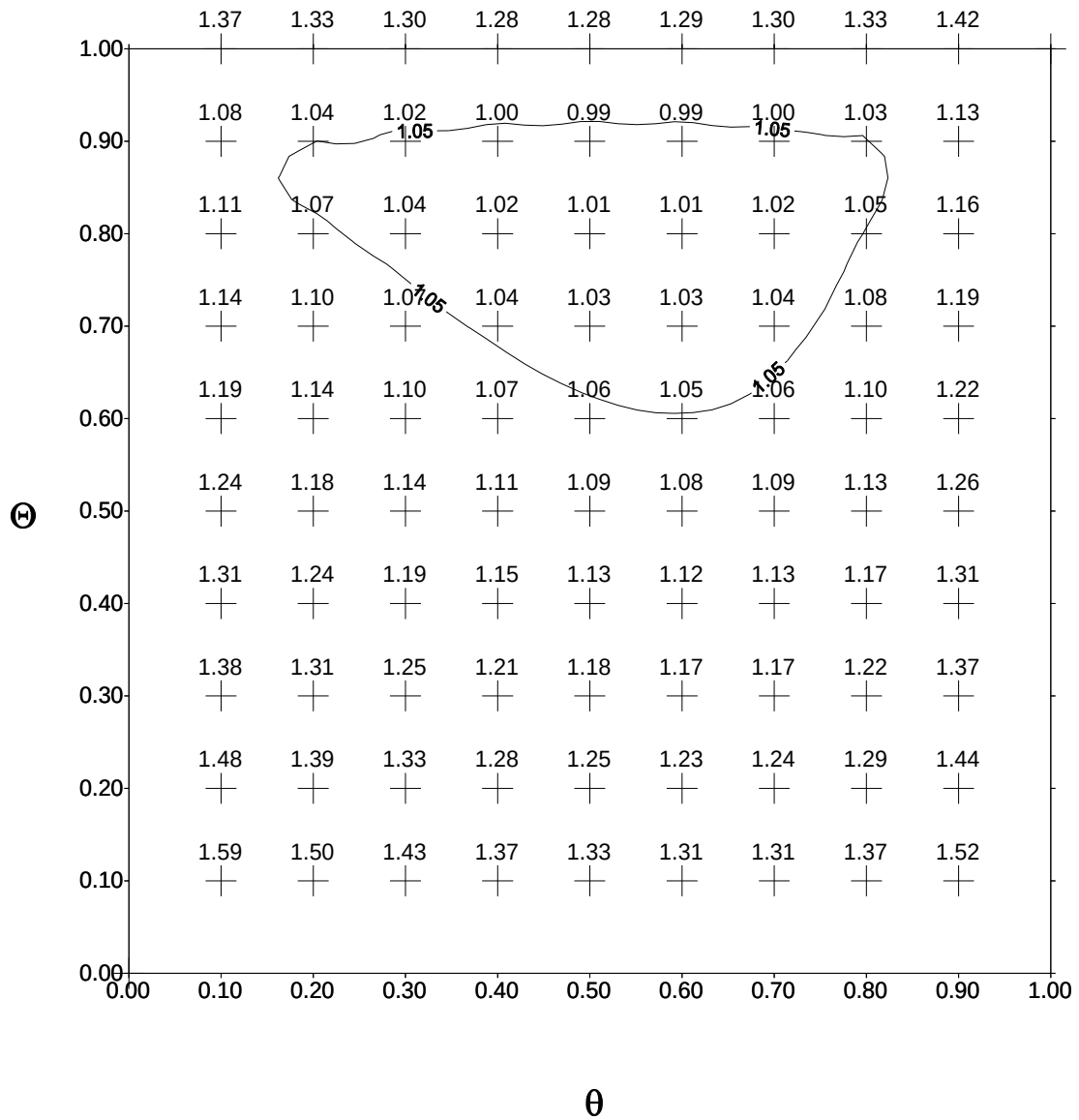


**Fig. 5.5: Confidence region of model parameters of EC of Kufa station
when $\theta=0.550$ & $\Phi=0.895$.**

Chapter Six

Forecasting

6.1 Forecasting Using ARMA Models

Forecasted monthly data are computed for the period from 2002 to 2004 by applying equation (2-63) and the rules given in section (2.12). Hence, the forecasting equations for all the stations, and water quality parameters are as given in tables from (6.1) to (6.3).

Table 6.1: ARMA models equations for AL-Hindya station parameters .

Parameter	Model Type	Model Equation
T.A	AR(1)	$\hat{Z}_t(1) = 0.601 Z_t$ $\hat{Z}_t(\ell) = 0.601 \hat{Z}_t(\ell-1) \text{ for } \ell = 2, 3, \dots, 36$
Ca	AR(1)	$\hat{Z}_t(1) = 0.609 Z_t$ $\hat{Z}_t(\ell) = 0.609 \hat{Z}_t(\ell-1) \text{ for } \ell = 2, 3, \dots, 36$
Cl	AR(1)	$\hat{Z}_t(1) = 0.683 Z_t$ $\hat{Z}_t(\ell) = 0.683 \hat{Z}_t(\ell-1) \text{ for } \ell = 2, 3, \dots, 36$
EC	ARMA(1,1)	$\hat{Z}_t(1) = 0.878 Z_t - 0.254 a_t$ $\hat{Z}_t(\ell) = 0.878 \hat{Z}_t(\ell-1) \text{ for } \ell = 2, 3, \dots, 36$
T.H	AR(1)	$\hat{Z}_t(1) = 0.741 Z_t$ $\hat{Z}_t(\ell) = 0.741 \hat{Z}_t(\ell-1) \text{ for } \ell = 2, 3, \dots, 36$
Mg	AR(1)	$\hat{Z}_t(1) = 0.714 Z_t$ $\hat{Z}_t(\ell) = 0.714 \hat{Z}_t(\ell-1) \text{ for } \ell = 2, 3, \dots, 36$
Tur	AR(1)	$\hat{Z}_t(1) = 0.702 Z_t$ $\hat{Z}_t(\ell) = 0.702 \hat{Z}_t(\ell-1) \text{ for } \ell = 2, 3, \dots, 36$

Table 6.2: ARMA model equations for Kufa Station parameters.

Parameter	Model Type	Model Equation
T.A	AR(1)	$\hat{Z}_t(1) = 0.352 Z_t$ $\hat{Z}_t(\ell) = 0.352 \hat{Z}_t(\ell-1)$ for $\ell = 2, 3, \dots, 36$
Ca	AR(2)	$\hat{Z}_t(1) = 0.443Z_t + 0.164Z_{t-1}$ $\hat{Z}_t(2) = 0.443\hat{Z}_t + 0.164Z_t$ $\hat{Z}_t(\ell) = 0.443\hat{Z}_t(\ell-1) + 0.164\hat{Z}_t(\ell-2)$ for $\ell = 2, 3, \dots, 36$
Cl	AR(1)	$\hat{Z}_t(1) = 0.488 Z_t$ $\hat{Z}_t(\ell) = 0.488 \hat{Z}_t(\ell-1)$ for $\ell = 2, 3, \dots, 36$
EC	AR(1)	$\hat{Z}_t(1) = 0.351 Z_t$ $\hat{Z}_t(\ell) = 0.351 \hat{Z}_t(\ell-1)$ for $\ell = 2, 3, \dots, 36$
T.H	AR(2)	$\hat{Z}_t(1) = 0.467Z_t + 0.253Z_{t-1}$ $\hat{Z}_t(2) = 0.467\hat{Z}_t + 0.253Z_t$ $\hat{Z}_t(\ell) = 0.467\hat{Z}_t(\ell-1) + 0.253\hat{Z}_t(\ell-2)$ for $\ell = 2, 3, \dots, 36$
Mg	AR(3)	$\hat{Z}_t(1) = 0.365 Z_t + 0.191Z_{t-1} + 0.305 Z_{t-2}$ $\hat{Z}_t(2) = 0.365 \hat{Z}_t(1) + 0.191Z_t + 0.305 Z_{t-1}$ $\hat{Z}_t(3) = 0.365 \hat{Z}_t(2) + 0.191\hat{Z}_t(1) + 0.305 Z_t$ $\hat{Z}_t(\ell) = 0.365\hat{Z}_t(\ell-1) + 0.191\hat{Z}_t(\ell-2) + 0.305\hat{Z}_t(\ell-3)$ for $\ell = 2, 3, \dots, 36$
Tur	AR(2)	$\hat{Z}_t(1) = 0.468Z_t + 0.232Z_{t-1}$ $\hat{Z}_t(2) = 0.468\hat{Z}_t + 0.232Z_t$ $\hat{Z}_t(\ell) = 0.468\hat{Z}_t(\ell-1) + 0.232\hat{Z}_t(\ell-2)$ for $\ell = 2, 3, \dots, 36$

Table 6.3 : ARMA model equations for Diwaniya station parameters.

Parameter	Model Type	Model equation
T.A	AR(2)	$\hat{Z}_t(1) = 0.402Z_t + 0.198Z_{t-1}$ $\hat{Z}_t(2) = 0.402\hat{Z}_t + 0.198Z_t$ $\hat{Z}_t(\ell) = 0.402\hat{Z}_t(\ell-1) + 0.198\hat{Z}_t(\ell-2)$ for $\ell = 2, 3, \dots, 36$
Ca	AR(1)	$\hat{Z}_t(1) = 0.524 Z_t$ $\hat{Z}_t(\ell) = 0.524 \hat{Z}_t(\ell-1)$ for $\ell = 2, 3, \dots, 36$
Cl	AR(1)	$\hat{Z}_t(1) = 0.449 Z_t$ $\hat{Z}_t(\ell) = 0.449 \hat{Z}_t(\ell-1)$ for $\ell = 2, 3, \dots, 36$
EC	AR(1)	$\hat{Z}_t(1) = 0.217 Z_t$ $\hat{Z}_t(\ell) = 0.217 \hat{Z}_t(\ell-1)$ for $\ell = 2, 3, \dots, 36$
T.H	AR(2)	$\hat{Z}_t(1) = 0.336Z_t + 0.141Z_{t-1}$ $\hat{Z}_t(2) = 0.336\hat{Z}_t + 0.141Z_t$ $\hat{Z}_t(\ell) = 0.336\hat{Z}_t(\ell-1) + 0.141\hat{Z}_t(\ell-2)$ for $\ell = 2, 3, \dots, 36$
Mg	ARMA(1,1)	$\hat{Z}_t(1) = -0.02Z_t + 0.303a_t$ $\hat{Z}_t(\ell) = -0.20\hat{Z}_t(\ell-1)$ for $\ell = 2, 3, \dots, 36$
Tur	MA(2)	$\hat{Z}_t(1) = 0.259 a_t + 0.348 a_{t-1}$ $\hat{Z}_t(2) = 0.348 a_t$ $\hat{Z}_t(\ell) = 0$ for $\ell = 2, 3, \dots, 36$

To illustrate the forecasting procedure, the calculation of (Ca) parameter of AL-Hindya station is given in table (6.4). In this table Z_t (col. 2) is calculated from model forecasting equation for (Ca) which is given in table (6.1) when the origin time is ($t = 120$) and lead time ($\ell = 1, 2, \dots, 36$). In col.3 Y'_t is calculated by reversing the standardization presses (reversing equation 4-7), then $Y'_t = Z_t * \sigma + \mu$ where $\mu = -0.0197$ and $\sigma = 0.057$ (where μ & σ are the mean and standard

deviation of Z_t for $t=121$ to $t=156$). In col. 6 F_t is calculated by reversing equation (4-6) (multiply $\hat{Y}_t$ by seasonal Sd (col. 5) and adding seasonal mean(col. 4) which are calculated by harmonic analysis in section 4.5). Forecasted series (col. 7) is calculated by reversing Box-Cox transformation with $\lambda=0.352$.

Table 6.4: Caculation of forecasting values of Ca of AL-Hindya station by using ARMA model.

Time (t) (months)	Z_t	Y'_t	Seasonal Means	Seasonal Sd	F_t	Forecasted data
120	-0.45421					
121	-0.276614	-0.03542	10.1738	1.202704	10.1312	74.7706
122	-0.168458	-0.02925	9.70454	1.188038	9.66979	67.4599
123	-0.102591	-0.0255	10.01626	1.208882	9.98544	72.40833
124	-0.062478	-0.02321	9.27175	1.190184	9.24412	61.14158
125	-0.038049	-0.02182	10.019	1.314964	9.99031	72.48651
126	-0.023172	-0.02097	9.3157	1.192336	9.2907	61.81333
127	-0.014112	-0.02045	10.161	1.120949	10.1381	74.88317
128	-0.008594	-0.02014	9.9676	1.184495	9.94374	71.7417
129	-0.005234	-0.01995	9.7132	1.226836	9.68873	67.75046
130	-0.003187	-0.01983	9.276306	1.196659	9.25257	61.26311
131	-0.001941	-0.01976	10.3228	1.232625	10.2984	77.54178
132	-0.001182	-0.01972	9.0178047	1.190828	8.99432	57.61914
133	-0.00072	-0.01969	10.1738	1.202704	10.1501	75.08078
134	-0.000438	-0.01967	9.70454	1.188038	9.68117	67.63437
135	-0.000267	-0.01967	10.01626	1.208882	9.99249	72.52148
136	-0.000163	-0.01966	9.27175	1.190184	9.24835	61.20237
137	-9.9E-05	-0.01966	10.019	1.314964	9.99315	72.53218
138	-6.03E-05	-0.01965	9.3157	1.192336	9.29227	61.83607
139	-3.67E-05	-0.01965	10.161	1.120949	10.139	74.89791
140	-2.24E-05	-0.01965	9.9676	1.184495	9.94432	71.75093
141	-1.36E-05	-0.01965	9.7132	1.226836	9.68909	67.75607
142	-8.3E-06	-0.01965	9.276306	1.196659	9.25279	61.26624
143	-5.05E-06	-0.01965	10.3228	1.232625	10.2986	77.54407
144	-3.08E-06	-0.01965	9.0178047	1.190828	8.9944	57.62025
145	-1.87E-06	-0.01965	10.1738	1.202704	10.1502	75.08159
146	-1.14E-06	-0.01965	9.70454	1.188038	9.68119	67.63483
147	-6.95E-07	-0.01965	10.01626	1.208882	9.99251	72.52177
148	-4.23E-07	-0.01965	9.27175	1.190184	9.24836	61.20253
149	-2.58E-07	-0.01965	10.019	1.314964	9.99316	72.5323
150	-1.57E-07	-0.01965	9.3157	1.192336	9.29227	61.83613
151	-9.56E-08	-0.01965	10.161	1.120949	10.139	74.89795
152	-5.82E-08	-0.01965	9.9676	1.184495	9.94432	71.75095
153	-3.55E-08	-0.01965	9.7132	1.226836	9.68909	67.75608
154	-2.16E-08	-0.01965	9.276306	1.196659	9.25279	61.26624
155	-1.31E-08	-0.01965	10.3228	1.232625	10.2986	77.54407
156	-8.01E-09	-0.01965	9.0178047	1.190828	8.9944	57.62025

The accuracy of forecast may be provided by determining the upper and lower probability limits by using equation (2-64). Figures (6.1) & (6.2) show the forecast series for the water quality parameters in AL-Hindya station by using ARMA model and upper and lower probability limits (at 95% confidence limits) and other related figures are given in the appendix (B)(figures from B.51 to B.54) .

6.2 Forecasting Using Box-Jenkins Seasonal Model

Suppose that the model is seasonal moving average model of order $(0,1,1) \times (0,1,1)_{12}$, forecasted monthly data are computed for the period from 2002 to 2004 by applying equation (5-1) and the rules given in section (2.12) . Then , the forecasting equation becomes

$$\hat{Z}_t(\ell) = Z_{t+\ell-1} + Z_{t+\ell-12} - Z_{t+\ell-13} + a_{t+\ell} - \theta a_{t+\ell-1} - \Theta a_{t+\ell-12} + \theta\Theta a_{t+\ell-13}, \dots (6-1)$$

where t is the origin time ($t=120$) and ℓ is the lead time ($\ell=1,2,\dots,36$). After obtaining the forecast series (Z_t for $t=121,122,\dots,156$), then the final series (X_t) is determined by reversing ($\ln$) transformation .

Figures (6.3) & (6.4) show the forecast series for the water quality parameters in AL-Hindya station by using Box-Jenkins seasonal model- which are given in table (5.3) -and upper and lower probability limits (at 95% confidence limits) while other related figures are given in the appendix (B)(figures from B.55 to B.58) .

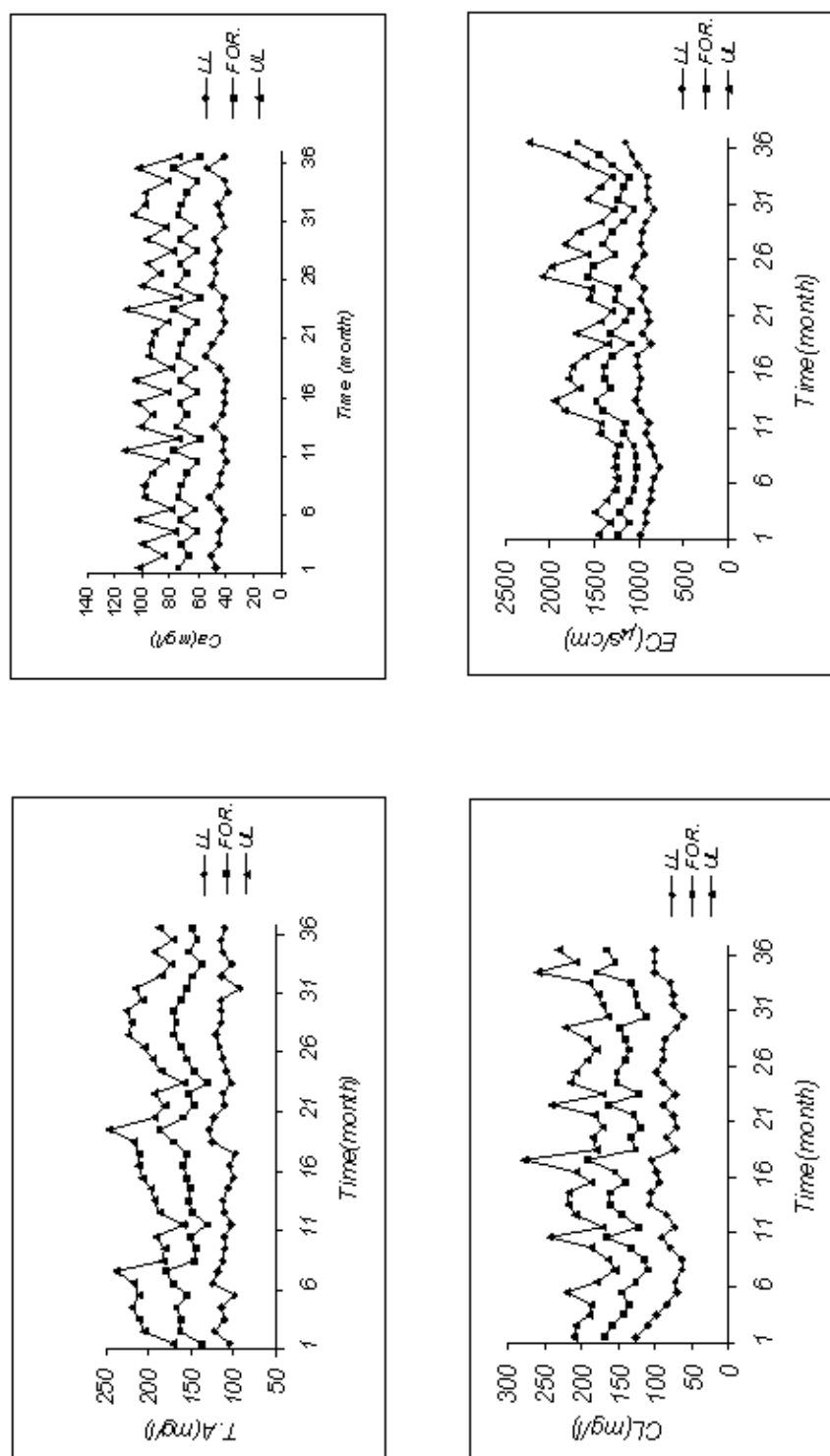


Fig 6.1: Probability limits of forecasted values obtained by ARMA model of T.A, Ca, Cl & EC of AL-Hindya station from 2002 to 2004.

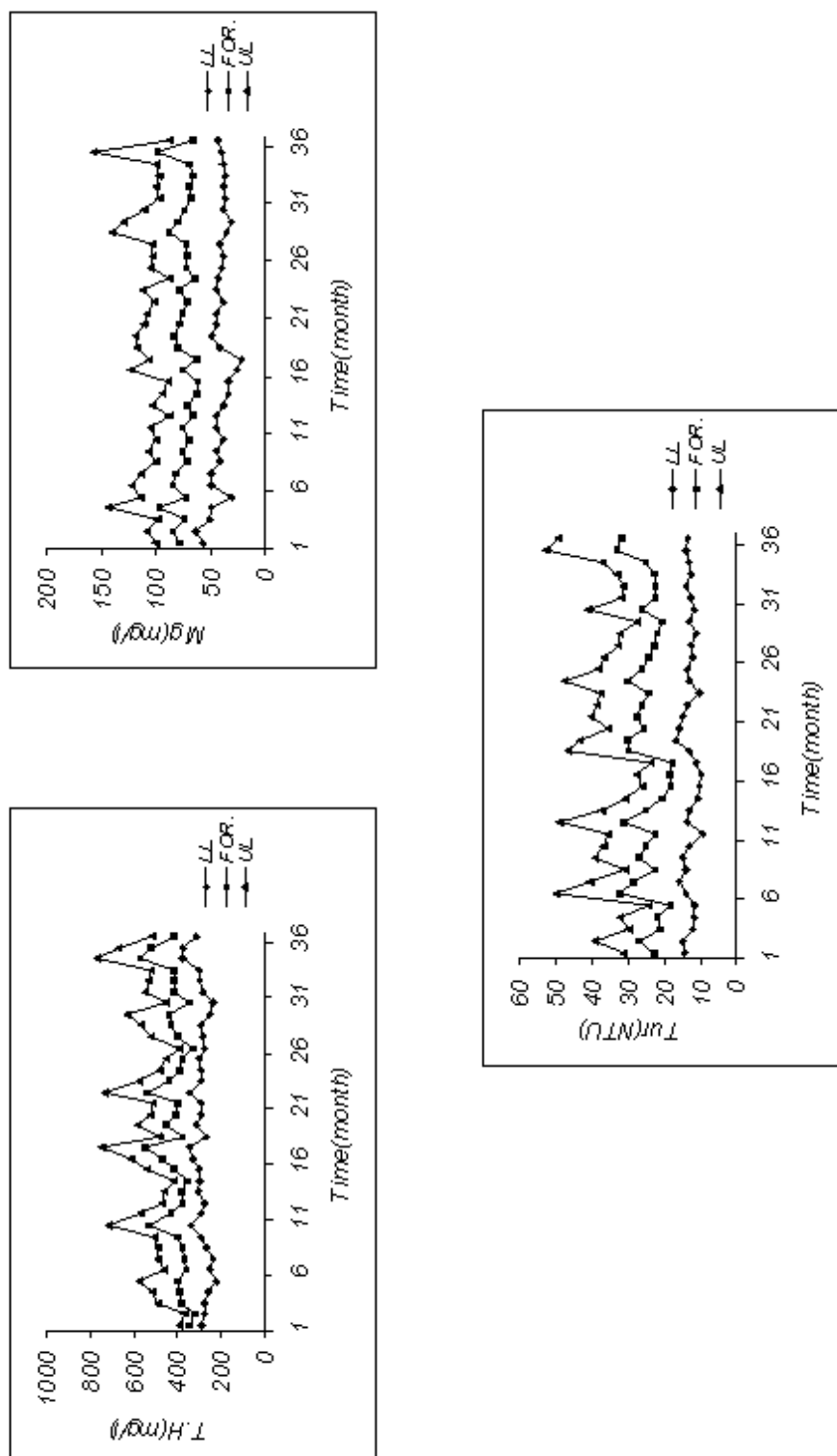


Fig 6.2: Probability limits of forecasted values obtained by ARMA model of T.H, Mg, Tur of AL-Hindya station from 2002 to 2004.

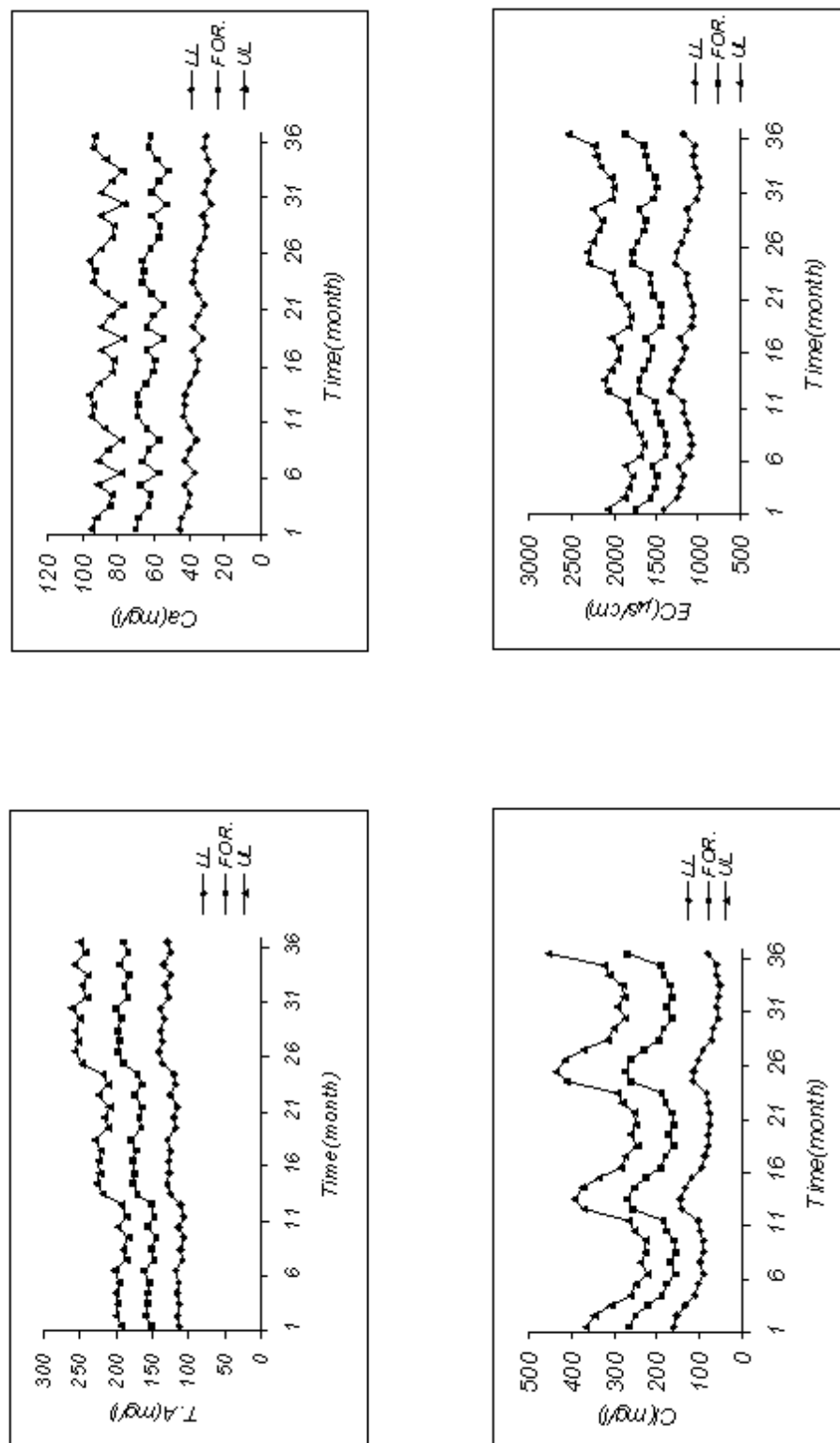


Fig. 6.3: Probability limits of forecasted values obtained by Box-Jenkins seasonal model of T.A, Ca, Cl & EC of AL-Hindya station from 2002 to 2004.

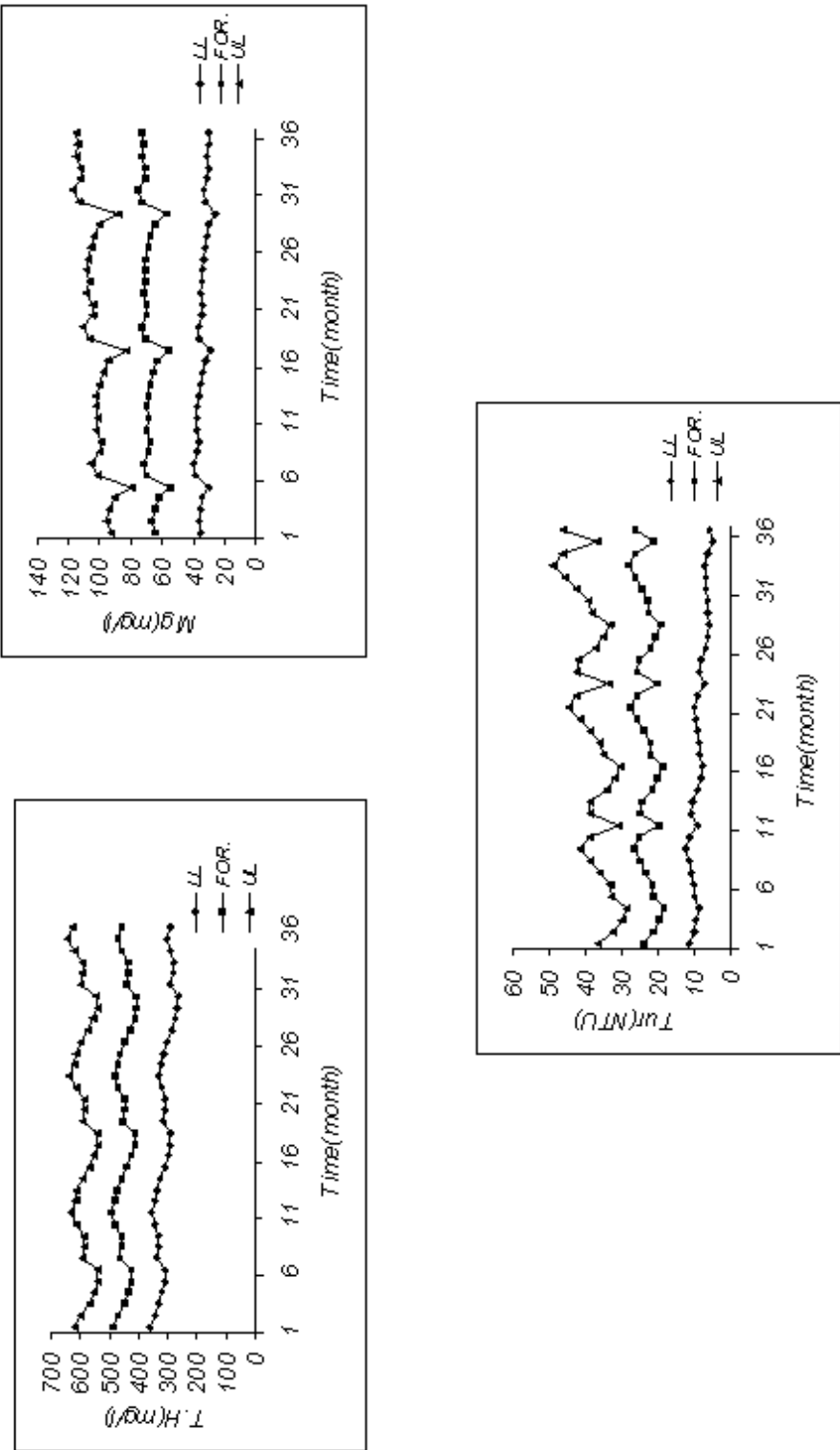


Fig. 6.4:Probability limits of forecasted values obtained by Box-Jenkins seasonal model of T.H, Mg, Tur of AL-Hindya station from 2002 to 2004.

6.3 Mean Absolute Relative Errors

In order to evaluate the accuracy of the forecasting and to compare the forecasting ability of the two modeling methods, the mean absolute relative error (MARE%) obtained by equation (2-73) was used. The best method is the one that gives the minimum (MARE%) value. As an example, the calculation of (MARE%) for the Ca parameter at Kufa station is given in table (6.5). Results of application of this measure are summarized in table (6.6) to select the proper model for each water quality parameter for each station.

Table 6.5: The calculation of MARE% for the Ca parameter at Kufa station.

t	Obs.	Box-Jenkins	ARMA	MARE% (Box-Jenkins)	MARE% (ARMA)	t	Obs.	Box-Jenkins	ARMA	MARE% (Box-Jenkins)	MARE% (ARMA)
1	111	111.4	110.7	0.33117	0.28731	19	107	108.9	107	1.79084	0.031205
2	117	114.6	107.3	2.05624	9.06894	20	113	103.2	103.3	8.68982	9.426354
3	125	106.7	103.1	14.6054	21.2448	21	108	101.9	101.5	5.6488	6.365167
4	122	105.1	119.9	13.8584	1.71863	22	115	107.8	104.9	6.27009	9.595203
5	112	108.7	113.9	2.96116	1.68108	23	126	109.4	106.4	13.1871	18.42528
6	126	117.9	118	6.4646	6.75478	24	121	106.1	116.6	12.3166	3.737099
7	121	111.1	106.5	8.1514	13.621	25	118	107	103	9.35449	14.57846
8	123	105.3	113.5	14.4031	8.38644	26	117	110.1	114	5.93103	2.620597
9	122	104	102	14.7732	19.5825	27	122	102.5	104.7	15.9669	16.57595
10	124	110	110.5	11.3008	12.2339	28	111	100.9	107.4	9.06739	3.354939
11	123	111.6	121.9	9.25667	0.89104	29	118	104.4	105.3	11.5391	12.01461
12	124	108.3	115.8	12.6935	7.10189	30	123	113.2	112.4	7.9739	9.437128
13	115	109.1	104.6	5.09357	9.98238	31	120	106.7	102.2	11.05	17.41476
14	123	112.3	107.2	8.69553	14.745	32	127	101.1	101.2	20.3787	25.51925
15	112	104.6	113.3	6.59768	1.13396	33	102	99.86	100.2	2.09483	1.824353
16	111	103	115	7.21351	3.43922	34	113	105.6	107.9	6.51699	4.700108
17	123	106.5	111.5	13.405	10.3155	35	102	107.2	113.4	5.09676	10.07908
18	119	115.5	114.8	2.94134	3.67949	36	108	104	100.5	3.725	7.421321
										sum=8.650	sum=8.862

6.4 Comparison Between ARMA and Box-Jenkins Models

Table (6.7) shows the preferable models (the model which give minimum MARE% according to table (6.6)). Figures (6.5) & (6.6) show the comparison between the forecasted values obtained from the two models for AL-Hindya station and the observed value for the three years period from 2002 to 2004. Other figures are given in the appendix (B)(see

figures from B.59 to B.62) .It is found that the Box-Jenkins seasonal models represent the observed values more closely than the ARMA models .

Table 6.6: MARE % result of the three stations.

Parameter	AL-Hindya		Kufa		Diwaniya	
	ARMA	Box-Jenkins	ARMA	Box-Jenkins	ARMA	Box-Jenkins
T.A	13.964	17.205	16.551	12.734	15.516	12.610
Ca	13.332	9.413	8.862	8.650	12.057	8.45
Cl	18.721	17.601	8.817	12.683	16.858	8.600
EC	13.565	16.182	14.608	16.909	10.061	5.456
T.H	16.123	11.054	12.496	19.718	5.764	6.019
Mg	15.419	9.884	14.687	13.805	11.236	12.298
Tur	16.796	10.725	12.964	12.104	16.877	15.961

Table 6.7: Methods and models used for 7-years forecasts.

Parameter	AL-Hindya		Kufa		Diwaniya	
	Model	Type	Model	Type	Model	Type
T.A	ARMA	AR(1)	Box-Jenkins	(0,1,1)×(0,1,2)	Box-Jenkins	(0,1,1)×(0,1,1)
Ca	Box-Jenkins	(0,1,1)×(0,1,2)	Box-Jenkins	(0,1,1)×(0,1,1)	Box-Jenkins	(0,1,1)×(0,1,2)
Cl	Box-Jenkins	(0,1,1)×(0,1,1)	ARMA	AR(1)	Box-Jenkins	(0,1,1)×(0,1,2)
EC	ARMA	ARMA(1,1)	ARMA	AR(1)	Box-Jenkins	(0,1,1)×(0,1,1)
T.H	Box-Jenkins	(0,1,1)×(0,1,1)	ARMA	AR(2)	ARMA	AR(2)
Mg	Box-Jenkins	(0,1,1)×(0,1,2)	Box-Jenkins	(0,1,1)×(0,1,1)	ARMA	ARMA(1,1)
Tur	Box-Jenkins	(0,1,1)×(0,1,1)	Box-Jenkins	(0,1,1)×(0,1,1)	Box-Jenkins	(0,1,1)×(0,1,1)

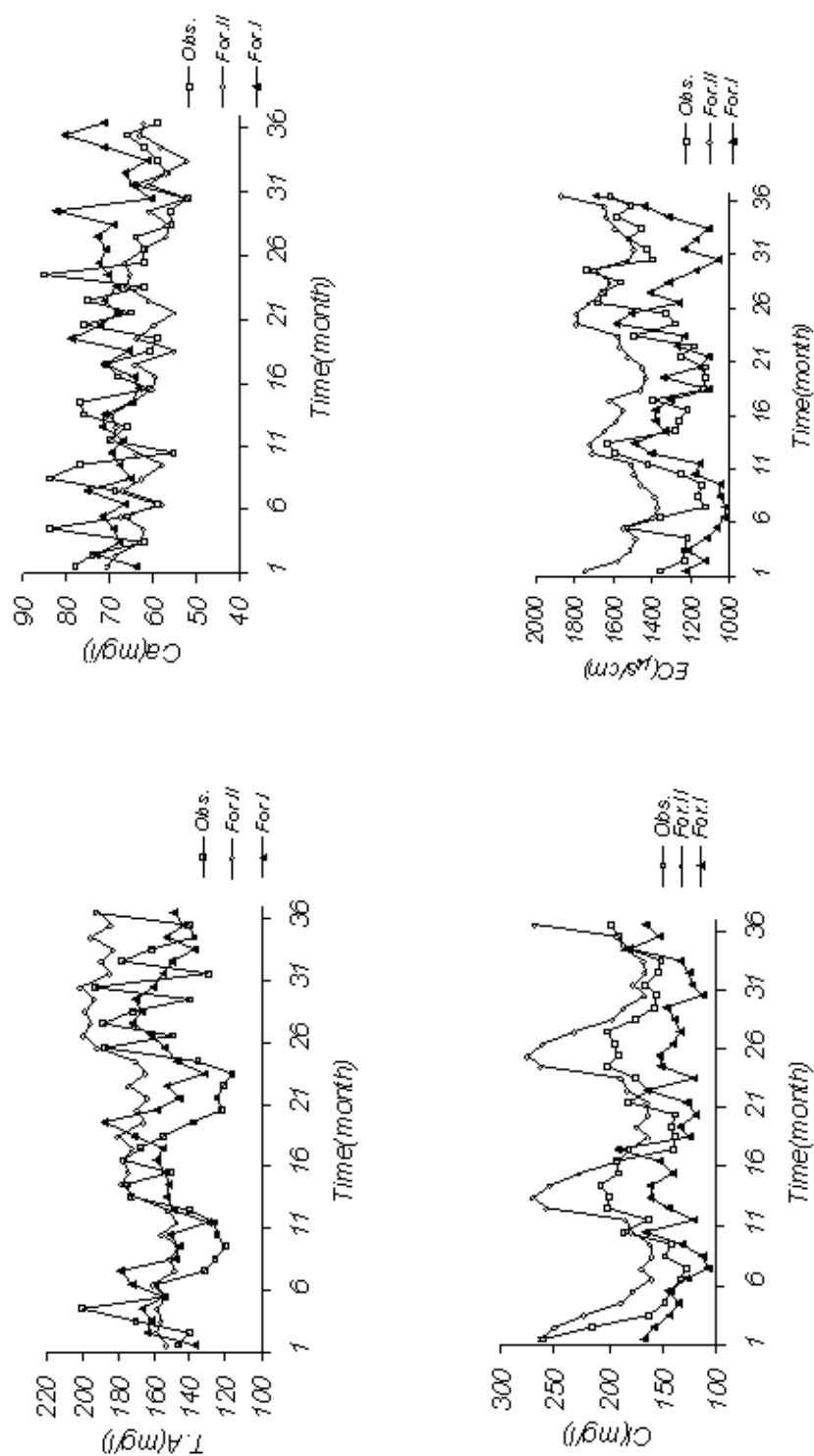


Fig. 6.5: Comparison between observed series and the series generated by (I-ARMA, II-Box-Jenkins) models for T.A, Ca, Cl & EC at AL-Hindya station.

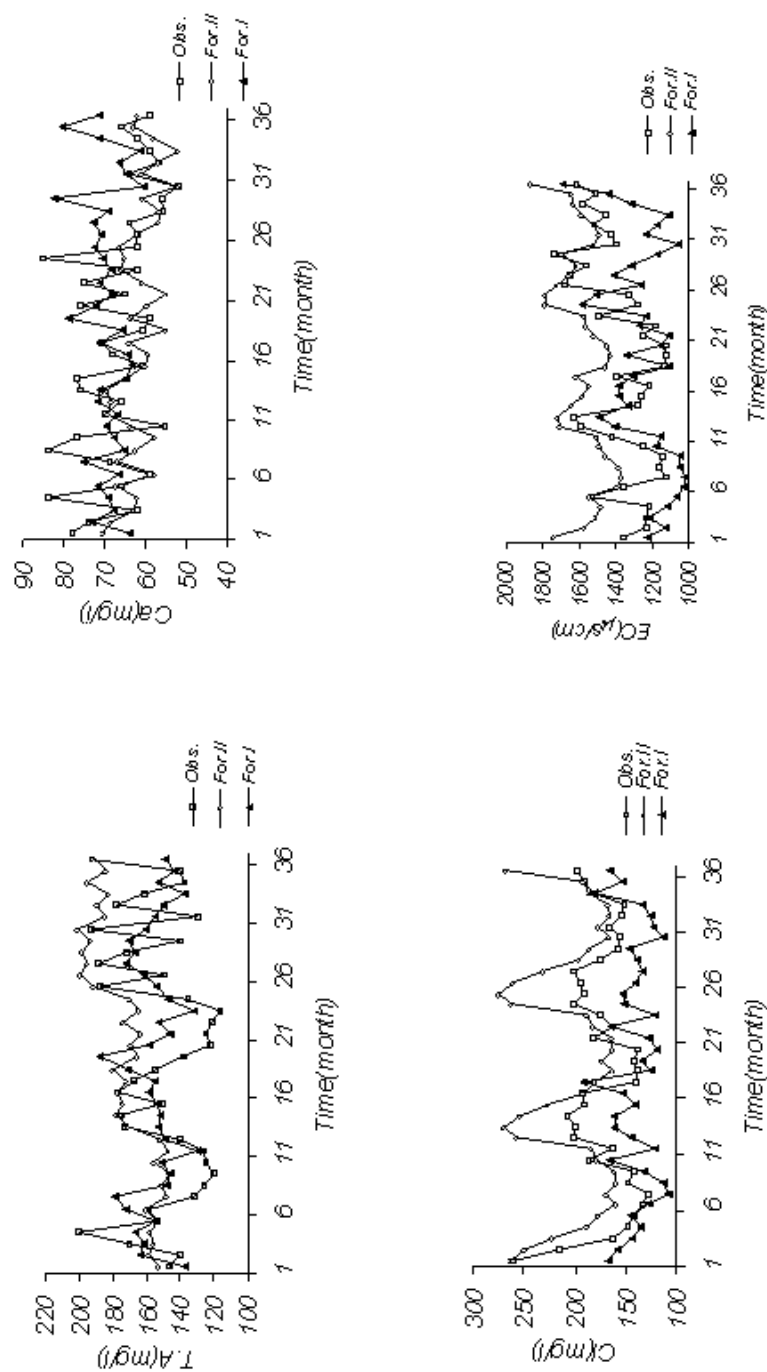


Fig. 6.5: Comparison between observed series and the series generated by (I-ARMA, II-Box-Jenkins) models for T.A, C'a, C'I & EC at AL-Hindya station.

6.5 Comparison of Forecasted Data with the River Standards

According to the models given in table (6.7), this section includes making forecasting for seven years extend from 2005 to 2011 and comparing the forecasting results with the river standard –which are given in table(6.8) -to evaluate the river suitability as a drinking water source. Figures (6.7)and (6.8) show the comparison between the forecasted values and river standards for AL-Hindya station while other related figures are given in the appendix (B)(figures from B.63 to B.66). Some of forecasted values exceed the specifications in the present and future time like Mg at AL-Hindya and Diwaniya stations, some of these values exceed the specifications after a period of time like T.A at Diwaniya station, some of these values decreased with time like Cl &T.H at AL-Hindya station and some of these values do not change with time like Cl & EC at Diwaniya station. As an example, the calculation of future values of T.A parameter of AL-Hindya station from 2005 to 2011 is given in table(6.9).

Table 6.8: Allowable limits of water quality parameters in surface water body used as drinking water source according to Iraqi specification. (Abbawi & Hassan, 1990).

Parameter	Allowable Limits
T.A	170 mg/l
EC	2000 μ s/cm
Ca	200 mg/l
Cl	200 mg/l
T.H	500 mg/l
Mg	50 mg/l
Tur	Non

Table 6.9: Future values of T.A of AL-Hindya station from 2005 to 2011.

t month	Future values	t month	Future values	t month	Future values	t month	Future values
1	165.547	22	152.753	43	172.865	64	161.655
2	181.714	23	131.743	44	147.362	65	159.01
3	173.922	24	147.27	45	145.335	66	161.536
4	174.439	25	154.179	46	163.012	67	154.193
5	159.616	26	161.467	47	147.17	68	147.78
6	174.174	27	172.163	48	163.695	69	139.902
7	180.489	28	167.67	49	156.405	70	158.043
8	148.589	29	170.54	50	165.909	71	139.615
9	146.323	30	160.5	51	180.711	72	143.802
10	151.332	31	154.348	52	175.777	73	163.236
11	130.243	32	149.899	53	176.908	74	172.86
12	149.13	33	137.846	54	170.906	75	174.645
13	152.672	34	153.123	55	172.88	76	170.96
14	151.977	35	142.867	56	154.266	77	157.42
15	154.153	36	148.642	57	145.559	78	167.884
16	158.191	37	163.02	58	157.169	79	168.824
17	154.957	38	166.431	59	135.222	80	158.265
18	170.636	39	164.246	60	138.081	81	152.432
19	187.154	40	179.064	61	144.785	82	170.569
20	158.292	41	177.66	62	168.439	83	137.293
21	146.559	42	174.303	63	159.458	84	142.973

6.6 Comparison Between the Present Study and Previous Studies in Iraq

The previous studies which deal with AL-Hindya station i.e. AL-Tikrity in 2001 and AL-Mousawi in 2003 show that the best models are as given in table (6.10) compared with models suggested in this study and this table shows that a clear difference according to the author. This difference because the difference of time studying period and suggesting more than one model during the initial model identification stage and using AIC to select the best model. In Kufa station the selected models by Mahmood in 2000 are different from the present study because the difference of time studying period and method of analysis. It should be noted that most previous studies are restricted to ARMA models only. In this study we use both ARMA and Box-Jenkins seasonal models.

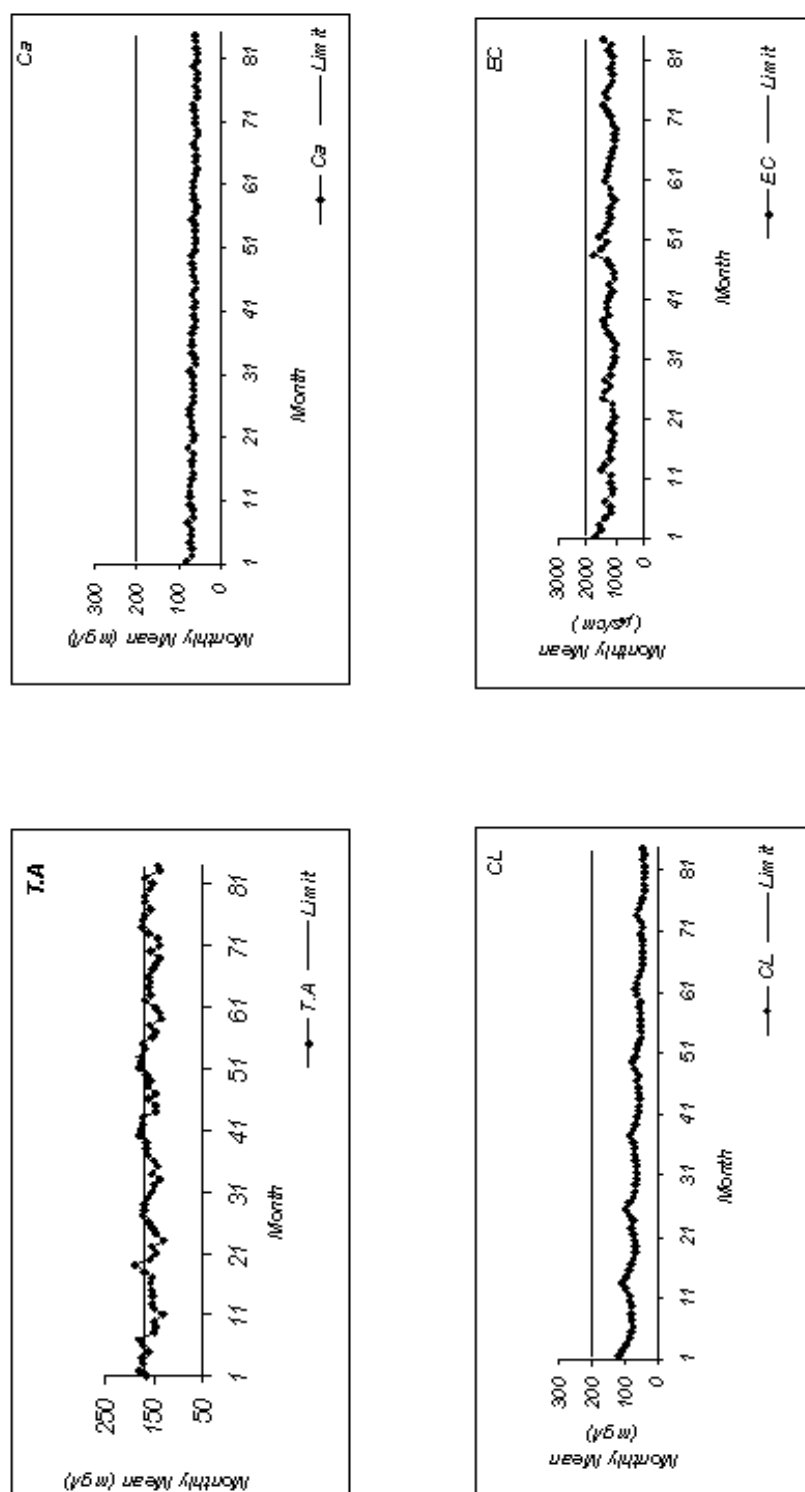


Fig. 6.7 : Comparison of 7-years(2005-2011)forecast of T.A, Ca, Cl & EC for AL-Hindya station with river standard .

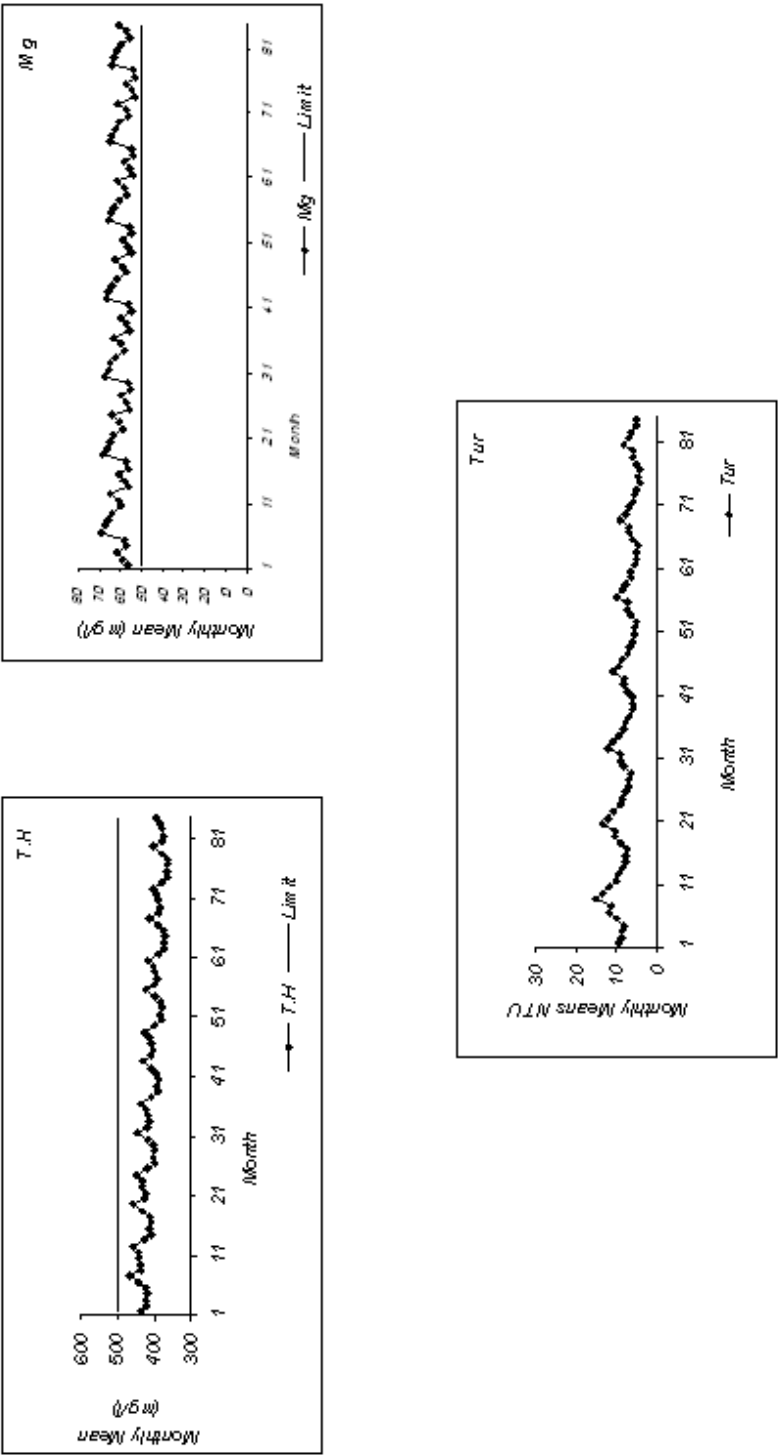


Fig.6.8 : Comparison of 7-years(2005-2011)forecast of T.H ,Mg & Tur for AL-Hindya station with river standard .

Table 6.10: Comparison between the present study and previous studies.

Author Sta. ¶.	AL-Tikrity		AL-Mousawi		Mahmood		Present author	
	Model	Model equation	Model	Model equation	Model	Model equation	Model	Model equation
AL-Hindya	T.A		AR(1)	$W_t = 0.414W_{t-1} + a_t$			AR(1)	$W_t = 0.601W_{t-1} + a_t$
	Ca	$W_t = 0.890W_{t-1} + a_t$	AR(1)	$W_t = 0.632W_{t-1} + a_t$			AR(1)	$W_t = 0.609W_{t-1} + a_t$
	Cl	$W_t = 0.920W_{t-1} + a_t$	AR(1)	$W_t = 0.679W_{t-1} + a_t$			AR(1)	$W_t = 0.683W_{t-1} + a_t$
	EC	$W_t = 0.920W_{t-1} + a_t$	AR(1)	$W_t = 0.639W_{t-1} + a_t$			ARMA (1,1)	$W_t = 0.878W_{t-1} - 0.254a_{t-1} + a_t$
	Mg	$W_t = 0.750W_{t-1} + a_t$	AR(1)	$W_t = 0.241W_{t-1} + a_t$			AR(1)	$W_t = 0.741W_{t-1} + a_t$
	T.H	$W_t = 0.920W_{t-1} + a_t$	AR(1)	$W_t = 0.709W_{t-1} + a_t$			AR(1)	$W_t = 0.714W_{t-1} + a_t$
	Tur		AR(1)	$W_t = 0.189W_{t-1} + a_t$			AR(1)	$W_t = 0.702W_{t-1} + a_t$
Kufa	T.A				AR(1)	$W_t = 2.25 + 0.41W_{t-1} + a_t$	AR(1)	$W_t = 0.352W_{t-1} + a_t$
	Ca				AR(2)	$W_t = 2.06 + 0.57W_{t-1} + 0.31W_{t-2} + a_t$	AR(2)	$W_t = 0.443W_{t-1} + 0.164W_{t-2} + a_t$
	Cl				AR(1)*	$W_t = 172.19 + 0.32W_{t-1} + a_t$	AR(1)	$W_t = 0.488W_{t-1} + a_t$
	EC				AR(1)	$W_t = 3.15 + 0.7W_{t-1} + a_t$	AR(1)	$W_t = 0.351W_{t-1} + a_t$
	Mg				ARIMA (1,1,1)	$Z_t = 0.56Z_{t-1} + a_t - 0.79a_{t-1}$	AR(2)	$W_t = 0.467W_{t-1} + 0.253W_{t-2} + a_t$
	T.H				AR(2)	$W_t = 1.7 + 0.62W_{t-1} + 0.31W_{t-2} + a_t$	AR(3)	$W_t = 0.365W_{t-1} + 0.191W_{t-2} + 0.305W_{t-3} + a_t$
	Tur				AR(1)	$W_t = 1.29 + 0.64W_{t-1} + a_t$	AR(2)	$W_t = 0.468W_{t-1} + 0.232W_{t-2} + a_t$

*(The analysis not included transformation)

Chapter Seven

Conclusions and Recommendations

7.1 Conclusions

From this study the following Conclusions are obtained :

1. Some of water quality parameters were found to be nonhomogeneous in mean, in standard deviation or both this means that the river receives various concentrations of various pollutants.
2. All water quality parameters show seasonal pattern, these patterns may be due to the influence of the annual cyclic pattern of the hydrological inputs to the river water environment.
3. According to ARMA method ,most of the monthly water quality parameters (18 from 21 cases) in the three stations give the autoregressive (AR) model with order (1, 2 or 3) which states that the data are dependent on the data of the last (1,2 or 3) months respectively, and some parameters like (Tur) at kufa station is represented by moving average model (MA(2)) which indicates that the correlation between the consecutive observations is not significant.
4. According to Box-Jenkins seasonal model method ,some of water quality parameters in the three stations (14 from 21 cases) give the ARIMA (0,1,1) ($\times$ 0,1,1) model and the others give the ARIMA (0,1,2) $\times$ (0,1,1) model.
- 5 . MARE% of forecast of the two methods were compared and it is found that Box-Jenkins seasonal model provides a satisfactory forecast for 14 cases from 21 cases whereas ARMA model provides a satisfactory forecast for other cases of the three stations.
- 6 . The obtained seven years forecast values by using both ARMA and Box – Jenkins models show that ,some of these values exceed the Iraqi Standard limits i.e. Mg at AL-Hindya station and T.A &Mg at Diwaniya station, whereas some values stay within the standard limits i.e. (Ca, Cl, EC & T.H at AL-Hindya station), (T.A, Ca, Cl, EC, T.H & Mg at Kufa station) and

(Ca, Cl & EC at Diwaniya station). Some of these values increase with time like Tur at Kufa station .

7.2 Recommendations

The following suggestions are put forward for further research :

1. Using more measuring or sampling stations of water quality parameters in the future studies to draw complete future picture of river pollution.
2. Using the Box-Jenkins seasonal models as a preferable modeling method because the Box-Jenkins seasonal models more simple and closely representing the observed values than the ARMA models .
3. Multisite and multivariate using the Box-Jenkins seasonal models are recommended as future studies to account the correlations between the stations and the parameters respectively .

1. الدكتور احمد سوسة, 1945. (ري العراق:نهر الفرات), مطبعة الحكومة, بغداد.
2. الدكتورة سعاد عبد عباوي و محمد سليمان حسن, 1990. (الهندسة العملية للبيئة وفحوصات الماء) جامعة الموصل.
3. مديرية ماء بابل, مديرية ماء النجف و مديرية ماء الديوانية-وزارة البلديات والاشغال.
4. Ahmad , S., Khan ,I. H . and Parida , B. P., 2001." Performance of Stochastic Approaches for Forecasting River Water Quality ", Journal of Water Resource Research, Vol.35, No.18, Pages 4261-4266.
5. AL-Husseini ,A. H., 2000." Hydrochemical Forecasting Model of Shutt AL-Hilla ", M. Sc. Thesis ,College of Engineering , University of Babylon.
6. AL-Mousawi ,E. M., 2003." Multisite Stochastic Model of Water Quality Properties at Selected Regions ", M. Sc. Thesis ,College of Engineering , University of Babylon.
7. AL-Suhaili ,R. H., 1986." Stochastic Analysis of Daily Streamflow of Tigris River ", M. Sc. Thesis, College of Engineering, University of Baghdad.
8. AL-Tikriti ,H. A., 2001. " Forecasting of Pollution Levels in Accordance with Discharge Reduction in selected Area on Euphrates River " M. Sc. Thesis, College of Engineering, University of Baghdad.
9. Box, G. E. P. and Jenkins, G. M., 1976. " Time Series Analysis; Forecasting and Control ", Holden Day, San Francisco, California.
10. Chandra, S., Spaila, S.K. and Kumar, A., 1978." Flood Frequency Analysis by Power Transformation ", Journal of Hydraulic Division, ASCE, Vol. 104, No. HY11, Pages 1945-1505.
11. Chatfield ,C.,1982. " The Analysis of Time Series : An Introduction ", Chapman and Hill, 2nd Edition.
12. Cooper, D. M., Wood, 1982. " Parameter Estimation of Multiple Input – Output Time Series Models: Application to Rainfall –Runoff Process ", Journal of Water Resource Research, Vol.18, No.5 , Pages 1352-1364.

- 13.Cun, C. and Vilagines, R., 1997. " Time Series Analysis on Chloride ,Ammonium and Dissolved Oxygen Concentrations in the Seine River Near Paris ", Journal of the Science of the Total Environment, 208, Pages 59-69.
- 14.Fernando, A. K. and Jayawardena, A. W., 1994. " Generation and Forecasting of Monsoon Rainfall Data " , Affordable Water Supply and Sanitation, 20th WEDC Conference, Colombo, Sri Lanka, (Internet).
- 15.Gupta, R. K. and Chanhana, H. S .,1986. " Stochastic Modeling of Irrigation Requirements ", Journal of Irrigation and Drainage Engineering, ASCE, Vol . 112 .No.6.
- 16.Hipel ,K. W., McLeod, A.I. and Lennox, W.C., 1977. "Advance in Box-Jenkins Modeling :1-Model Construction ", Journal of Water Resource Research , Vol.13, No.3.
- 17.Hirsch ,R. M., 1979. " Synthetic Hydrology and Water Supply Reliability ", Journal of Water Resource Research, Vol.15,No.6,Pages 1603-1615.
- 18.Iturbe ,I. R, Dawdy , D. R and Garcia ,L. E., 1971." Adequacy of Markovian with Cyclic Components for Stochastic streamflow Simulation ", Journal of Water Resource Research , Vol.2, No.5.
- 19.Jayawardena, A. W., and Lai, F., 1989. " Time Series of Water Quality Data in Pearl River –China ".Journal of Environmental Engineering, ASCE, Vol. 115. No.3.
- 20.Kahya, E. and Kalayci, S., 2004. " Trend Analysis of Streamflow in Turkey " , Journal of Hydrology, 289, Pages 128-144.
- 21.Kurunc, A., Yureki, k . and Cevik, O., 2005. " Performance of Two Stochastic Approaches for Forecasting Water Quality and Streamflow Data From Yesilirmak River, Turkey ", Journal of Environmental Modeling and Software, 20, Pages 1195-1200 .

22. Lehmann, A. and Rode, M., 2001. " Long Term Behavior and Cross-Correlation Water Quality Analysis of the River Elbe ", Journal of Water Resource Research , Vol.35, No.9, Pages 2153-2160.
23. Loftis, J.C, Taylor, C. H. and Chapman, P. L., , 1991. " Multivariate Tests for Trend in Water Quality ", Journal of Water Resource Research, Vol.27, No.7, Pages 1419-1427.
24. Lohani, B. N. and Wang, M. M., 1987. " Water Quality Data Analysis in Chang Kang River ", Journal of Environmental Engineering, ASCE, Vol. 113. No.1.
25. Mahloch, J. L., 1974. " Multivariate Techniques for Water Quality Analysis ", Journal of Environmental Engineering Division, Vol.100, No. EE5.
26. Mahmood, H. S., 2000. "Time Series Analysis of Hydrochemical Pollutant : Application of Box-Jenkins Models on Euphrates River at Kufa ". M. Sc. Thesis, College of Engineering, University of Babylon.
27. Matalas, N. C., 1967. " Mathematical Assessment of Synthetic Hydrology", Journal of Water Resource Research, Vol. 3, No.4.
28. Matalas, N.C., 1967. " Time Series Analysis ", Journal of Water Resource Research , Vol. 3, No.3, Pages 817-829.
29. Mattikalli, N. M., 1996. " Time Series of Historical Surface Water Quality Data of the River Glen Catchment, U.K.", Journal of Environmental Management 46, Pages 149-172.
30. McLeod, A. I., Noakes, D. J., Hipel, K. W. and Thompson, R. M., 1987. "Combining Hydrologic Forecasts ", Journal of Water Resource Planning and Management, Vol. 113, No.1, Pages 29-39.
31. McMahon, T.A, and Mein, R.G, 1986. "River and Reservoir Yield ", Water Resource Publications , U.S.A.

32. Miller, R. B., Bell, W., Ferreiro, O. and Wang, R. Y., 1981. " Modeling of Daily River Flows with Precipitation Input " , Journal of Water Resource Research, Vol.17, No.1, Pages 209-215.
33. Montanary, A., Rosso, R. and Taqqu, M. S., 1997. " Fractionally Differenced ARIMA Models Applied to Hydrologic Time Series: Identification, Estimation, and Simulation ", Journal of Water Resource Research, Vol.33, No.5, Pages 1035-1044.
34. Salas, J. D. and Abdulmohsen, M. W., 1993. " Initializing for Generation Single-Site and Multisite Low- Order Periodic Autoregressive and Moving Average Process ", Journal of Water Resource Research, Vol.29, No.6, Pages 1771-1776.
35. Srikanthan, R. and McMahon, T.A., 1982. " Stochastic Generation of Streamflows ", Journal of Hydraulic Division, ASCE, Vol. 108, No. HY3.
36. Stedinger, J. R., 1981. " Estimation Correlations in Multivariate Stream Flow Models ", Journal of Water Resource Research, Vol.17, No.1.
37. Tesfaye, Y. G., 2005. " Seasonal Time Series Model and Their Application to the Modeling of River Flows ", Ph.D. Thesis, University of Nevada, Reno.
38. Wonnacott, T.H, and Wonnacott, R. J., 1990. " Introductory Statistic for Business and Economic ", 4th Edition, John Wiley & Sons .
39. Zou, S., and Yu, Y. S., 1996. " A Dynamic Factor Model for Multivariate Water Quality Time Series with Trend ", Journal of Hydrology ,178.