FORCASTING ANALYSIS OF TOTAL DISSOLVED SOLIDS AND CHLORIDE CONCENTRATIONS IN EUPHRATES RIVER IN BABYLON PROVINCE-HILLA CITY

Nassrin J. AL-Mansori
Nassrin20052001@yahoo.com
Environmental Engineering Department/ Collage of Engineering/University of Babylon

ABSTRACT

The monthly time series of the Total Dissolved Solids (TDS) and Chloride(CL) concentrations in Euphrates River at Babylon were analyzed as a time series.

The data used for the analysis was the monthly series during (2000-2013). The series was tested for non-homogeneity and found to be nonhomogeneous. A significant positive jump was observed at 2002 for Total Dissolved Solids (TDS) and at 2006 for Chloride (CL). This non homogeneity was removed using a method suggested by Yevichevich (7). The homogeneous series was then normalized using Box and Cox (2) transformation. The periodic component of the series was fitted using harmonic analyses, and removed from the series to obtain the dependent stochastic component. This component was then modeled using first order autoregressive model (Markovian chain). The above analysis was conducted using the data for the period (2001-2011), the remaining 2-years (2012-2013) of the observed data was left for the verification of the model. The observed model was used to generate future series. Those series were compared with the observed series using t-test. The comparison indicates the capability of the model to produce acceptable future data.

Key words: Total dissolved solids, Chloride, nonhomogeneous, periodic component, and dependent stochastic

. التنبؤ بالبيانات الشهرية لتراكيز المواد المذابة الكلية وتراكيز الكلور لنهر الفرات في مدينة بابل- الحلة

نسرين جاسم حسين المنصوري

قسم هندسة البيئة /كلية الهندسة / جامعة بابل

الخلاصــــــة:

تم في هذا البحث تحليل البيانات الشهرية لتركيز المواد المذابة الكلية وتراكيز الكلور لنهر الفرات في محافظة بابل (مدينة الحلة) ، حيث تم اخذ المعدل الشهري للبيانات المتوفرة من سنة (2000-2013).(دائرة بيئة بابل، بيانات).

في البداية تم اختبار البيانات لمعرفة فيما اذا كانت متجانسة ام غير متجانسة ووجد انها غير متجانسة عند سنة 2002 للمواد المذابة الكلية وعند سنة 2006 لتركيز الكلوريد باستخدام طريقة Yevichevich تم ازالة عدم التجانس في البيانات . ثم تم توزيع البيانات توزيعا طبيعيا باستخدام طريقة العالمين Cox,Box .

بعد ذلك تم اخذ القيم المعدلة لإزالة المركبة الدورية عنها وذلك للحصول على الدالة المستقلة التي تم نمذجتها بموديل من نوع سلسلة Markovian.

ان التحليل اعلاه تم تطبيقه على البيانات من سنة 2001-2013 حيث تركت قيم التراكيز للسنتين المتبقيتين (2012-2013) لاستخدامها في التحقق من نتائج الموديل.

ومن خلال النموذج الذي تم التوصل اليه تم التنبؤ بقيم لسنتين مستقبليه لمقارنتها مع القيم المقاسة وباستخدام اختبار t-test وقد بينت النتائج امكانية الاعتماد على النموذج لا عطاء نتائج مستقبلية مقبولة.

SYMBOLS

AR : Autoregressive.
ARIMA: Autoregressive Integrated Moving Average.
BOD: Biological Oxidation Demand
DO: Dissolved Oxygen
RMSE: Root Mean Square Error
MAE: Mean Absolute Error.

INTRODUCTION

Water quality express the suitability of water to sustain various uses or processes, such as drinking water, irrigation water and nature conservation.

Solids can be dispersed in water in both suspended and dissolved forms. Total Dissolved solids may be present in suspension and /or in solution and may be divided into organic matter and inorganic matter and these material may be objectionable in water for several reasons. It is aesthetically displeasing and provides adsorption sites for chemical and biological agents. Biologically active suspended solids may include disease causing organisms as well as organisms such as producing strain of algae.

Chlorides are usually present in water in the form of sodium chloride. These impart a salty taste to water. When present in concentration more than 200mg/l, the taste maybe unacceptable by some consumers.

During the last 25 years, Time Series Analysis had become one of the most important and widely used branches of Mathematical Statistics. The technique of time-series analysis uses estimated statistical parameters to build a mathematical model. This model is capable of describing the evolution of possible sequences of events in time, at the site of observations, which have the same statistical properties as the historical sample. In this research, the data of TDS and CL in Euphrates River will be utilized to build the mathematical model to predict the future concentration. Numbers of projects were by now under construction in the southern of Iraq after March 2003, such as water treatment plants and projects related to restoration of marshes and irrigation projects. The modeling of TDS and CL could be used to predict future values, that are useful in the operation of such project. Moreover these data are useful also in planning, design of new projects.

AL-Suhaili R.H., (1985) used daily stream flow data of Tigris river to build stochastic model. The data were collected from four measurement station (Mosul, Fatha, Baghdad, Kut) located downstream of a river. Two different single site model were used to detect the changes in the stream flow data ,(AR and ARIMA) models and a multisite model (MATLABS).

Mustafa et al.(2012)., used modelling technique (Artificial neural network) represent the non-liner relationship among the input and output variables of a water resources system by using a grouped neurons or nodes in layers

Zhiyong Huang and Hiroshi Morimoto**.** (2006) designed a structure of input/mixing/output model (called it the three-step model) and represented each process of material input, mixing and output by a stochastic numerical methods. Data of BOD and DO concentrations of river

Cam used to build the model. The stochastic numerical methods applied in this paper are discrete time approximation methods. Additional equations representing mixing processes and biochemical reactions also used in model. The researcher concluded that the theory of stochastic differential equations is a beneficial tool for studying water pollution.

Kadri Yurkel and Ahmet Kurunc. (2006) had analyzed the daily discharge data of each month from three gauge stations on Cekerek Stream for forecasting using stochastic approaches. Initially non-parametric test (Mann-Kendall) was used to identify the trend during the study period. The two approaches of stochastic modeling, ARIMA and Thomas-Fiering models were used to simulate the monthly-minimum daily discharge data of each month. The error estimates (RMSE and MAE) forecasts from both approaches were compared to identify the most suitable approach for reliable forecast. The two error estimates calculated for two approaches indicate that ARIMA model appear slightly better than Thomas-Fiering model. However, both approaches were identified as an appropriate method for simulating the monthly-minimum daily discharge data of each month from three gauge stations on Cekerek Stream.

The use of time series analysis (stochastic analysis) in generating possible future dissolved solids and sulfate concentrations assumes that those concentrations are extracted from a common statistical population, and that the recorded historical data forms a sample from this population.

Generally, a hydrologic time series may consist of four components depending on the type of variable and the averaging time interval. Each of seasonal TDS,CL concentration series four components may exist and can be considered to arise from a combination of those components, which are termed the jump component (Jt), trend component (Tt), periodic or cyclic component (Pt) and stochastic or random component (εt).These components may be formulated by:

$$TDSt= Jt +Tt +Pt +\varepsilon t \qquad (1)$$

$$CLt= Jt +Tt +Pt +\varepsilon t \qquad (2)$$

The first three components represents the deterministic part of the process while the fourth component represents the non-deterministic part, therefore those three components should be detected and identified by suitable formulations and decomposed from the stochastic component.

METHEDOLOGY

The procedure used for data analysis may be summarized by the following steps:

1- Test and Removal of Non-Homogeneity

The modeling process required a set data to be homogenous. Hence, the first step before starting the analysis is to test the homogeneity of the data series. If the test indicates non-homogeneity, then this non-homogeneity should be removed. This was achieved by plotting the average monthly data and computing the annual mean and standard deviation for each year then using the spilt-sample approach which divides the entire sample into two sub-samples. Then testing the differences between the means and standard deviations of these two sub-samples at the 95 percent probability level of significance using the t-test method.

The data were tested for non-homogeneity and found to be non-homogeneous at year 2002,2006 respectively for both parameters, see **Fig. (1 and 2)**. The calculated t-value is greater than the tabulated t-value.

For non-homogeneity removal, Yevjevich(1972) suggest fitting linear regression equations for both annual means and annual standard deviations, then applying the following equation:-

$$YJ,t = \frac{XJ,t - XJ}{SJ}\ Sd2 + Av2 \qquad (3)$$

where:
j, t: the annual and seasonal positions of the observations, respectively.
Y: transformed series(homogeneous)
X : historical non-homogeneous series.
Av_2,Sd_2: the average and standard deviation of the second sub –sample respectively, and
Xj,Sj : linear regression equations for annual means and standard deviations against years.

Thus the data are divided into two sub-samples, For (TDS), the first (3) years long (2000-2002) and the second (9) years long (2003-2011) and for (CL), the first (7) years long (2000-2006) and the second (5) years long (2007-2011) The first sub-sample will be transformed according to the above equation. The two sub-samples were then tested again using t-test to check the homogeneity.

The result of applying the above equation are shown in **Fig. (1)**, **Fig.(2)** and **Table (1)** below, hence the data are homogeneous. Since the calculated t-values is less than tabulated t-value.

2- Transformation to Normally Distributed Data

It is of common practice in time- series analysis to transform the data to the normal distribution. This means, to remove the skewness in the data and try to make it nearly zero. For the normalization process several transformations may be used to normalize the data, but the most common one is the power transformation. The power transformation used in this research is the one suggested by Box and Cox (1976) (see equation (4) below. The application of this transformation begins with the estimation of the transformation coefficient value (λ). This coefficient has a value between (-2) and (2) and is strongly related to the skewness coefficient (Cs).**Table (2)** shows values of Cs computed for each transformed series, using different λ -values.

The values above were found to best fitted by a second polynomial equations.

$$\lambda_{TDS} = 0.0079Cs^2+0.4712Cs+ 0.042 \qquad (4)$$

$$\lambda_{CL} = 0.0396Cs^2+0.523Cs+ 0.023 \qquad (5)$$

In order to find the λ-value that will normalize the data, the skewness coefficient Cs in the above equations are substituted by zero, which gives a λ value as 0.042 for TDS and 0.023 for CL. This value will be used to get the transformed series according to the Box and Cox transformation as follows:

$$Y = \frac{(X-1)^{\lambda}}{\lambda} \qquad (6)$$

where: Y: The transformed Series, X : The original Series Data. **Table (3)** shows the monthly means and standard deviations for the original and transformed series.

3- Determination of the Independent Stochastic Component

The series obtained after the removal of non-homogeneity and non-stationary (periodic component of mean and standard deviation) is termed as the dependent stochastic component of the process and denoted as ($\varepsilon p,t$). The values of monthly mean and standard deviations were used to find the value of the independent stochastic component by the following equation:-

$$\varepsilon_{p,t} = \frac{Y_{P,t} - \mu_t}{\sigma_t} \tag{7}$$

where: $\varepsilon p,t$ = is the dependent stochastic component. μt =is the mean value of yp,t data at position p(month). σ_t is the standard deviation value of $y_{p,t}$ data at position p(month). The values of $\varepsilon_{p,t}$ may be fitted by a suitable model whose parameters will depend directly or indirectly on the amount of existing correlation represented by the lag rk serial correlation coefficient model (rk), **Fig.(3a,3b)** for TDS and CL respectively. One of the most familiar models, are the autoregressive model. It is preferable to try the first degree model, and then check its adequacy to remove the dependency of the $\varepsilon p,t$ series in the first degree model fails to remove the dependency, the second degree model will be used, and so on. The first degree autoregressive model required the calculation of lag-one (r_1) serial correlation coefficient, which was found to be (r_1=0.45626, r_1=0.5736) for TDS and CL respectively.

The model is represented as the relation between the dependent stochastic component ($\varepsilon_{p,t}$) and the independent stochastic component ($\zeta_{p,t}$). The independent stochastic series ($\zeta_{p,t}$) is a series of random numbers usually with zero mean and unit variance.

As mentioned above one of the most used models is the first order autoregressive model (Markov model). This model express the relationship between the $\varepsilon_{p,t}$ and $\zeta_{p,t}$ as follows:

$$\varepsilon_{p,t} = a * \varepsilon_{p,t-1} + \sqrt{1-a^2}\ \zeta_{p,t} \tag{8}$$

$$\varepsilon_{p,t} = b * \varepsilon_{p,t-1} + \sqrt{1-b^2}\ \zeta_{p,t} \tag{9}$$

where: a= r_1 for TDS, b=r1 for CL Substituting the value of a =(0.45626) in the above equation (8) and b=(0.5736) in the above equation (9), the independent stochastic component $\zeta_{p,t}$, for both parameters could be found using:

$$\zeta_{p,t} = \frac{\varepsilon_{p,t} - 0.45626 * \varepsilon_{p,t-1}}{0.8898} \quad for\ TDS\ data \tag{10}$$

$$\zeta_{p,t} = \frac{\varepsilon_{p,t} - 0.5736 * \varepsilon_{p,t-1}}{0.8191} \quad for\ CL\ data \tag{11}$$

In order to test the validity of the proposed first order autoregressive model, the correlogram of the $\zeta_{p,t}$ component should be found and tested. This correlogram is shown below which show that the first order autoregressive model, is suitable since the values of the serial correlation coefficient are fluctuated around the zero-values. Hence the proposed model was capable of removing the dependency between the values of $\varepsilon_{p,t}$, Figure (4a and 4b) for both parameters.

4- Model Verification

Upon the completion of the first four steps above, the model parameters were found. As mentioned before the observed data series of TDS and CL were divided into two parts (2000-2011), was used for the analysis (i.e., models parameters estimation), the other part (2012-2013), will be used now for model verification.

Usually in practice, the model is used to generate future values (series). The model validity will be decided upon the comparison between the statistical properties of the generated series with those of the observed one that was not used in the estimation of the parameters of the model.

The Microsoft Excel program was used for generating future series for the TDS-values and CL-values. Two series were generated as shown in **Table (4)** and **Table(5)**. The generation process begins by generating a standardized normally distributed random series (i.e., with zero mean and unit variance) then, using those as $\zeta p,t$ values to generate the $\varepsilon p,t$ values using the first autoregressive model. The TDS,CL values were found using a reverse process of the analysis conducted in steps (2-4).

Table (6) and **Table (7)** show the generated monthly TDS and CL values respectively using the two generated randomized series, proposed to be for years 2012 and 2013. The observed TDS and CL concentrations for these 2-years are shown in **Table (8)**.

Table (9) shows the values of the mean, standard deviation, skewness coefficient and kurtoses coefficient of observed data and generated one. **Table (10)** shows the results of the t-test for monthly means of observed and generated TDS and CL Series.

CONCLUSIONS

1) The series of monthly TDS and CL concentrations in Euphrates River at Babylon Province is non-homogeneous. The non-homogeneity can be attributed to the disposal of effluent wastewater from the constructed treatment plants.

2) The suitable value of the power transformation parameter λ that can be used to transform data to the normal distribution was found to be 0.042 for TDS and 0.023 for CL.

3) The correlogram of the observed independent stochastic component indicate the capability of the first order autoregressive model to model to time-dependency of the dependent stochastic component.

4) The T-test result shows that the obtained model can presence future forecasted values for the monthly TSS values.

Table(1) Mean and standard deviation of each sub-samples before and after applying the procedure of removal of non-homogeneity.

TDS Data	Before Removal		After Removal	
	Mean	Standard deviation	Mean	Standard deviation
Set 1	759.0278	70.51142	856.6753	120.7332
Set 2	858.25	120.6797	858.25	120.6797
CL	Before Removal		After Removal	
	Mean	Standard deviation	Mean	Standard deviation
Set 1	149.9167	22.24545	150.8985	7.65487
Set 2	150.95	22.67004	138.4667	19.03058

Table(2) Variation of Skewness Coefficient with Box and Cox transformations Coefficient for TDS and CL

λ and Cs for TDS								
λ	-0.6	-0.4	-0.2	0.2	0.6	0.8	1	1.2
Cs	-1.73	-1.4	-1.43	-0.73	-0.38	-0.23	-0.09	0.03
λ and Cs for CL								
λ	-0.8	-0.6	-0.4	0.4	0.6	0.8	1.1	1.2
Cs	-0.49	-0.3	-0.10	-0.09	0.043	0.12	0.370	0.54

Table (3) Monthly means and standard deviations for the original homogeneous series and normalized series for TDS and CL in a period from 2000 to 2011.

Mon.	Original homogeneous Series(x)				Normalized series(y)			
	TDS		CL		TDS		CL	
	Mean	St.d	Mean	St.d	Mean	St.d	Mean	St.d
Jan.	807.	131	1444	20.3	32.6	0.24	55.2	0.17
Feb.	798.	121	146.	19.2	32.6	0.19	55.2	0.17
Mar.	805.	134	150.	20.4	32.6	0.21	55.2	0.16
Ap.	759.	150	145.	12.40	32.5	0.32	55.2	0.10
May	855.	189	139.	20.73	32.7	0.28	55.1	0.17
Jun.	843.	109	135.	25.94	32.7	0.17	55.1	0.22
Jul.	874.	99.	142.	21.41	32.7	0.15	55.1	0.18
Aug	958.	145	155.	20.07	32.8	0.20	55.2	0.14
Sep.	908.	113	146.	20.42	32.8	0.16	55.2	0.16
Oct.	920.	104	143.	21.91	32.8	0.15	55.1	0.16
Nov	925.	123	146.	13.13	32.8	0.17	55.2	0.10
Dec.	836.	148	151.	18.40	32.7	0.24	55.2	0.14

Table(4) TDS Values of generated randomized numbers ($\zeta_{p,t}$) for 2012,2013 years.

months	Rand1		Rand2		Rand3	
Jun.	0.399	0.220	0.178	0.161	0.625	0.225
Feb.	0.015	0.864	0.093	0.574	0.437	0.594
Mar.	0.569	0.213	0.038	0.338	0.104	0.587
Apr.	0.792	0.772	0.833	0.872	0.129	0.297
May.	0.957	0.935	0.011	0.690	0.677	0.178
Jun.	0.013	0.969	0.384	0.664	0.095	0.270
July.	0.198	0.600	0.170	0.882	0.343	0.613
Aug.	0.133	0.120	0.963	0.273	0.839	0.056
Sep.	0.015	0.500	0.674	0.563	0.704	0.564
Oct.	0.985	0.887	0.558	0.833	0.898	0.562
Nov.	0.982	0.209	0.945	0.165	0.171	0.362
Dec.	0.128	0.384	0.060	0.478	0.115	0.275

Table(5) CL Values of generated randomized numbers ($\zeta_{p,t}$) for 2012,2013 years.

months	Rand1		Rand2		Rand3	
Jun.	0.674	0.192	0.180	0.723	0.010	0.192
Feb.	0.713	0.830	0.566	0.209	0.329	0.229
Mar.	0.773	0.410	0.000	0.479	0.9682	0.173
Apr.	0.532	0.163	0.572	0.032	0.178	0.751
May.	0.096	0.134	0.363	0.203	0.969	0.764
Jun.	0.387	0.720	0.406	0.647	0.023	0.229
July.	0.354	0.521	0.426	0.072	0.098	0.548
Aug.	0.802	0.041	0.698	0.505	0.457	0.830
Sep.	0.130	0.980	0.657	0.088	0.633	0.520
Oct.	0.360	0.621	0.933	0.822	0.397	0.433
Nov.	0.767	0.472	0.160	0.024	0.911	0.887
Dec.	0.790	0.839	0.776	0.910	0.994	0.889

Table (6) generated monthly TDS concentrations (mg/l) for 2012,2013 years

months	TDS con. Rand1		TDS con. Rand2		TDS con. Rand3	
	2012	2013	2012	2013	2012	2013
Jun.	794.9	897.8	794.9	887.8	794.9	858.6
Feb.	832.8	933.5	808.2	893.4	858.7	883.3
Mar.	818.0	889.0	815.3	886.0	881.6	913.0
Apr.	856.7	949.3	760.6	967.9	815.4	871.8
May.	1031.	1103.	992.0	1061.	891.9	924.4
Jun.	998.4	1019.	880.1	971.8	925.1	889.7
July.	933.1	1000.	923.0	1011.	913.2	949.6
Aug.	1017	1049.	1007.	1078.	1024.	1008.
Sep.	940.3	991.6	1028.	1008.	1021.	983.8
Oct.	931.9	1041.	1034.	1042.	1034.	1003.
Nov.	1044.2	1009.7	1049.	1005.	1092	1008.
Dec.	1055.5	934.3	1052.	946.5	948.3	916.6

Table (7) generated monthly CL concentrations (mg/l) for 2012,2013 years.

months	TDS con. Rand1		TDS con. Rand2		TDS con. Rand3	
	2012	2012	2012	2012	2012	2013
Jun.	163.1	173.9	163.1	172.1	163.1	181.7
Feb.	170.0	179.7	159.7	165.1	156.1	170.7
Mar.	177.8	176.7	168.6	169.9	161.6	166.7
Apr.	163.4	155.9	151.6	152.3	160.4	159.6
May.	166.5	150.5	155.0	148.4	156.0	167.5
Jun.	156.2	161.2	155.0	157.5	172.6	160.5
July.	159.6	164.9	159.5	153.6	159.1	165.2
Aug.	168.9	165.8	170.1	168.7	164.1	180.4
Sep.	170.0	171.6	168.7	155.8	160.0	171.5
Oct.	158.0	168.9	167.9	163.5	162.3	164.9
Nov.	156.6	161.8	167.5	154.4	158.6	165.8
Dec.	172.3	178.1	170.2	173.0	176.7	182.4

Table (8) observed TDS and CL concentrations (mg/l) for 2012,2013 years.

months	TDS		CL	
	2012	2013	2012	2013
Jun.	942	849	158	141
Feb.	918	889	133	146
Mar.	857	768	106	131
Apr.	603	801	130	118
May.	688	707	116	131
Jun.	623	752	116	130
July.	771	757	130	125
Aug.	757	798	144	117
Sep.	748	767	133	123
Oct.	723	811	127	118
Nov.	760	876	128	152
Dec.	848	619	153	118

Table (9) statistical prosperities of observed data and generated series for TDS & CL

	TDS			
	Observed data	Rand1	Rand2	Rand3
Mean	776.33	961.46	954.4939	933.9309
Standard deviation	89.855	82.4558	93.63728	74.85573
skewness	-0.1821	-0.43764	-0.65001	0.206083
kurtoses	-0.2509	-0.6019	-0.72835	-0.51551
CL				
	Observed data	Rand1	Rand2	Rand3
Mean	130.16	166.3603	162.1797	166.190463
Standard deviation	13.26213	7.95224767	7.55870447	7.85893117
skewness	0.483443	-0.0530578	-0.2132529	0.84967267
kurtoses	-0.27098	-0.7890046	-1.4118451	-0.2194869

Table(10) results of t-test for monthly TDS, CL means, t tabulated= 2.65

months	T Rand1	T Rand2	T Rand3
Jun.	0. 348	0. 368	0. 406
Feb.	0. 024	0. 013	0. 845
Mar.	0. 855	0. 021	0. 665
Apr.	0. 615	0. 938	0. 317
May.	0. 613	0. 695	0. 552
Jun.	0.674	0. 736	0. 684
July.	0. 869	0. 962	0. 656
Aug.	0.0773	0. 043	0. 034
Sep.	0. 742	0. 615	0. 714
Oct.	0. 201	0.7 59	0. 2905
Nov.	0. 755	0. 724	0. 363
Dec.	0. 487	0. 335	0. 744

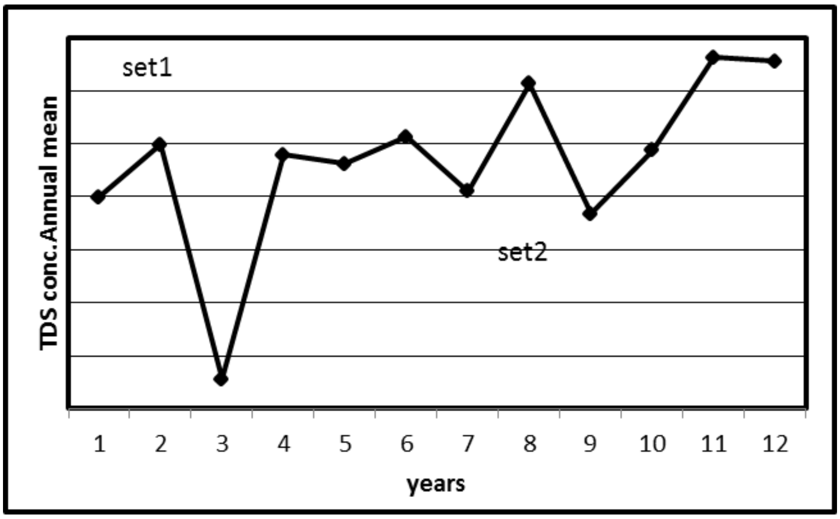

Fig.1 Split sample test of the original historical Data for TDS.

Fig.2 Split sample test of the original historical Data for CL.

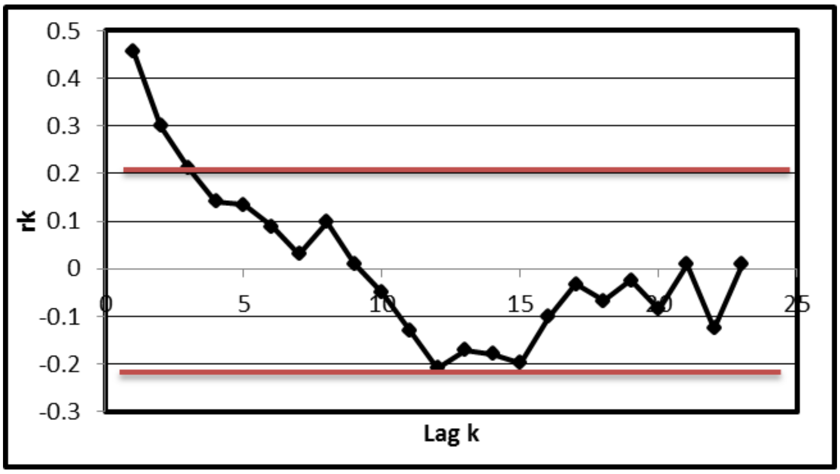

Fig.3.a. corrologram of the dependent stochastic component ($\varepsilon_{p,t}$), for TDS data.

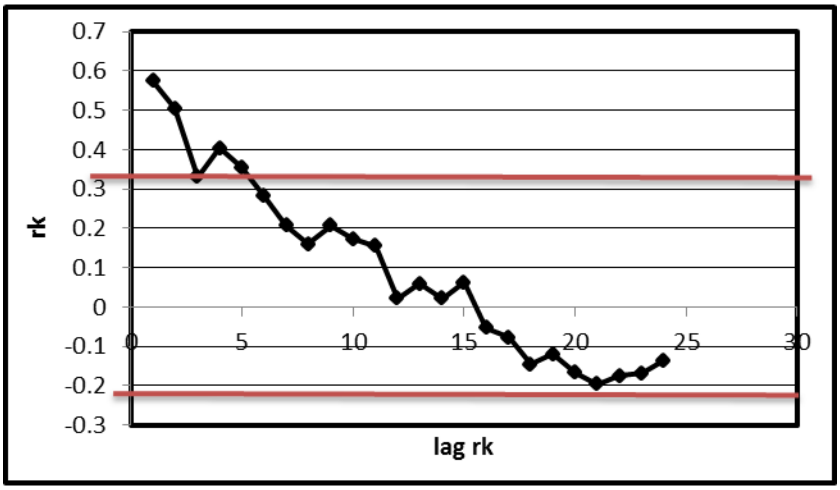

Fig.3.b. corrologram of the dependent stochastic component ($\varepsilon_{p,t}$), for CL data.

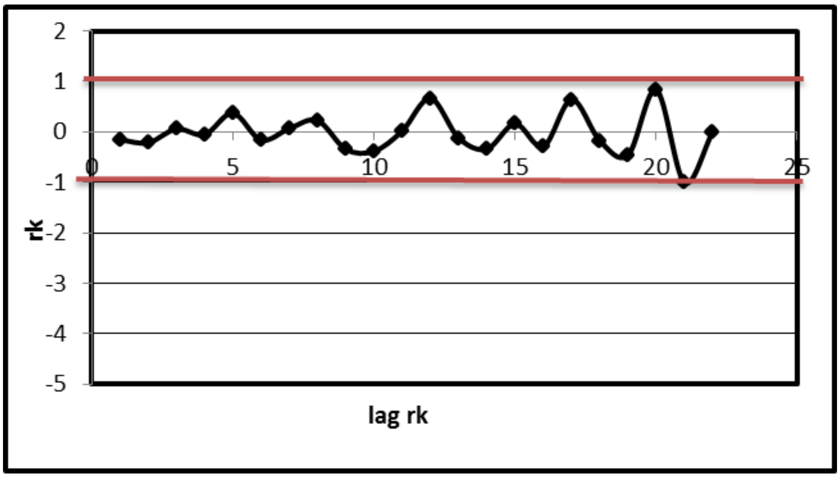

Fig.4.a. corrologram of the independent stochastic component ($\zeta_{p,t}$), for TDS data.

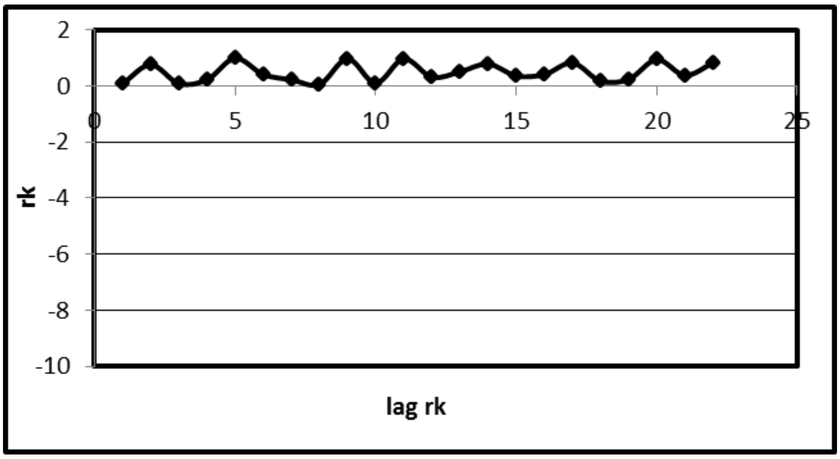

Fig.4.b. corrologram of the independent stochastic component ($\zeta_{p,t}$), for CL data.

REFRENCES

Al-Suhaili R.H., (1985) "Stochastic Analysis of Daily Stream flow of Tigris River", M.Sc. Thesis, College of Engineering, University of Baghdad.

Box, G.E.P., and, Jenkins, G.M., (1976)"Time Series Analysis, Forecasting and Control". Holdened Day, San Fransisco,. P.575.

Kadri Yurkel and Ahmet Kurunc.,(2006)" Performances of Stochastic Approaches in Generating Low Stream flow Data for Drought Analysis". Journal of Spatial Hydrology Vol. 5 no. 1. Internet.

Mustafa, M.R., Isa, M.H. and Rezaur, R.B. (2012) "Artificial neural networks modeling in water resources engineering: infrastructure and applications", World Academy of Science, Engineering and Technology, Vol. 6, No. 2, pp.317–325.

Yevjevich,W.,M.,(1972), "Structural Analysis of Hydrology time Series", Hydrology Paper No.56,Fort Collins, Colorado, Nov.

Zhiyong Hung and Hiroshi Morimoto, (2006)" Water Pollution Models based on Stochastic Differential Equations". Department of Earth and Environmental Sciences Graduate School of Environmental Studies, Nagoya University, PDF, Internet.

دائرة بيئة بابل ، "التحاليل الكيمياويه لنهر الفرات في مدينة الحله" ، بيانات (2000-2013) .