Analysis and Projection of The Influent Rate and Quality Parameters of Water in King Talal Dam

By
Ahmad Abdel Razzak Bani Hani

Supervisor
Dr. Ahmad I. Al Jamrah

Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Civil Engineering

Faculty of Graduate Studies
University of Jordan

August - 2002

This thesis was successfully defended and approved on:…………………………………

Examination Committee	Signature
Dr. Ahmad Jamrah, Chairman Associate Professor, Environmental Engineering Department of Civil Engineering University of Jordan	…………………………...
Dr. Sulieman Tarawneh, Member Associate Professor, Environmental Engineering Department of Civil Engineering University of Jordan	…………………………...
Dr. Tawfik Samaneh, Member Assistant Professor, Environmental Engineering Department of Civil Engineering University of Jordan	…………………………...
Dr. Munjed Al-Sharif, Member Assistant Professor, Environmental Engineering Department of Civil Engineering Jordan University of Science and Technology	..……………………….….

DIDICATION

To my family

To whom I am proud to belong

With Best Regards

Eng. A. Bani Hani

ACKNOWLEDGEMENTS

To all those who

gave me the full assistance and support to

prepare and complete this thesis, especially

Dr. A. Al Jamrah,

Dr. Suleiman Tarawneh

and the head of Irrigation Department in Jordan

Valley Authority

My full thanks and gratitude

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

P : The estimated probability of an X value

M : The Rank

n : The number of data

S : Standard deviation

$\bar{X}$: The mean

C_k : Coefficient of Kurtosis.

K : The fourth moment about the mean.

X_i : Observed data at time i.

Wi : Shapiro-Wilk statistics

U_t : Random process with mean equals to zero and variance equals to one.

β_r : Unknown parameter with $\beta_0 = 1$ and $|\beta_r| < 1$.

Y_i : Observation (data) at time equals to i.

λ : The lag operator

α_q : Unknown parameter, $\alpha_0 = 1$ and $|\alpha_q| < 1$.

v_t : Is a sequence of independent random variables with $\xi\, v_t = 0$ and $\xi\, v_t^2 = \sigma^2$.

$W_t = \nabla$: Differentiation of Yi

P : Unknown parameter with $\rho_0 = 1$ and $|\rho_r| < 1$.

a : The regression constant

b_y : The regression coefficient

θ : Smoothing constant , $0 < \theta < 1$

ρ_{xy} : The correlation coefficient

p,d,q : ARIMA coefficients

Analysis and Projection of The Influent Rate and Quality Parameters of Water in King Talal Dam

By
Ahmad Abdel Razzak Bani Hani

Supervisor
Dr. Ahmad I. Al Jamrah

Abstract

King Talal Reservoir (KTR) is one of the most important projects in Jordan. This project is used to irrigate wide regions in the Jordan Valley. Studying the water quality and quantity in King Talal Reservoir (KTR), and the trends of influent rate and quality should be a priority in Jordan.

This study concentrates on studying and projection quantity and quality variables in Zarka River (main branch and supplier of KTR). The quantity variable used was the flow, and, the quality variables were (TSS, BOD5, COD, T-P, and T-N) in Zarka River. The data collected for each variable was containing 156 months from the year 1988 till the end of the year 2000. The procedure used in analyzing the six variables in Zarka River is through auto, cross, and vertical distance correlation between a point in Zarka River and another point, which was a point in the Samra's effluent. Deterministic and stochastic (ARIMA model) forecasting of the six variables were used in finding the best model to be used in projection.

The study results indicate that the ARIMA model is a good model in predicting most of the six variables. In forecasting the BOD5 variable non of the modeling satisfied the 10 percentage of mean error but ARIMA model gave the best model and the least percentage of mean error. However, ARIMA model did not give the best modeling in

the COD variable. The least percentage of mean error of all variables by ARIMA modeling was equal to 4.8% in the T-P variable. The cross and distance correlation gave information about the variables, such as: the relation of the variables together, the form of the variables in Zarka River, the source of the variables, and other information.

1. INTRODUCTION

1.1 Introduction

Jordan is one of the countries of the Middle East. Jordan has an area of 90 thousand Km^2 and population of about 4.9 million in the year 1999 according to (General Statistical Department, 2000). The population growth rate is about 3.6% per year. One of the problems facing Jordan is water; Jordan's climate is classified as a semi-arid one. The climate is mainly characterized by low precipitation and humidity. Approximately 80% of Jordan's area receives an average precipitation of less than 100 mm/year. The evaporation rate in some parts of Jordan ranges from 5 to 80 times the average amount of precipitation (Salameh, 1996).

Water consumption categories in Jordan are mainly divided into three types. These include domestic, industrial, and agricultural water consumption. These categories constitute 25%, 5%, and 70% of the total water consumption respectively. It should be noted that the annual agricultural water consumption in 1995 was about 639.7 MCM (RSS, 1998).

Many projects were done to increase the usage of water. One of the most important projects was King Talal Reservoir (KTR). (KTR) is one of the most important projects in Jordan and is used to irrigate wide regions in the Jordan Valley, which is estimated to be about 100 Km^2. The capacity of (KTR) is estimated to be about 86 million m^3 with a depth of 108 m and a catchment area of 3175 Km^2. The inlet of water to King Talal Reservoir is mainly from Zarka Stream and Ramemen Wadi. The influent water to the reservoir is a mixture of rainfall and springs water mixed with domestic and industrial waste (WAJ, 1998).

The difference in quality and quantity of water and wastewater in (KTR) mandates the implementation of a continuous monitoring program of quantity and quality. Monitoring activities should aim at investigating whether the effluent of (KTR) is within the limits of the irrigation standards. Forecasting the water quantity and quality should help in establishing some precautional measures, which should assist in resolving the anticipated problems.

The term water quality refers to the suitability of water to a particular purpose. Any physical, chemical, or biological property that influences the use of water is called water quality variable. Water quality standards have been developed to serve as guidelines for selecting water supplies for various activities (Boyd, 2000).

The water quality variables that will be forecasted in this thesis are TSS, BOD, COD, T-P, and T-N. Total suspended solids (TSS) is the most important physical characteristics of wastewater, which is composed of floating matter, settleable matter, colloidal matter, and matter in solution. Analytically, the total suspended solids content is defined as all matter that remains as residue upon filtration and then evaporation at 103 to 105 ^{o}C. Biochemical oxygen demand (BOD) is a chemical characteristic. The most widely used parameter of organic pollution applied to both wastewater and surface water is the 5-days BOD (BOD_5). This determination involves the measurements of the dissolved oxygen used by microorganisms in the biochemical oxidation of organic matter. Chemical oxygen demand (COD) is a chemical characteristic used to measure the organic and inorganic pollutants in wastewater. Finally the total Nitrogen and

Phosphorus are chemical characteristics and essential to the growth of the protista and plants. Inspite of the fact that they are chemical characteristics, they are necessary to evaluate the treatability of wastewater by biological processes (Metcalf & Eddy, 1991).

These five water quality variables will be analyzed and forecasted using statistical methods. Any statistical analysis of data has to be based upon some assumed probability model for those data (Green & Margerison, 1977). If a series has shown some trend or persistent pattern in its variations for long period of time in the past, it will be sensible to assume that such patterns or regularities will continue in the future (Chao, 1974). But one should take into consideration that in decision-making, forecasts are usually wrong. The magnitude of the forecasting errors experienced will depend upon the forecasting system used (Montogomery & Johnson 1976).

The main objectives of forecasting, for environmental quality can be summarized as; firstly, the data is frequently expensive to accumulate so that using forecasting can minimize it. Secondly, correlation between the constituents may help the filling of missing data or the identification of outlier data. Finally, it is useful in detection of any deterioration in environmental quality, because early detection may provide the opportunity for controlling the problem at a lower cost before the problem magnifies (Bean & Rover 1998).

The aim of this research is to study the water quality and quantity in (KTR). Trends of influent rate and quality will be established and analyzed as a function of time. A time series model will then be derived to represent the trends and help in making future projections.

1.2 Literature Review

Many researchers have investigated the subject of forecasting in the field of water quantity and quality. Some of the topics investigated include water quality of surface and ground water, water level, water flow, floods and rainfall, design and operation of biological wastewater treatments, wastewater treatment plant performance, long term effluent quality, and other papers in forecasting field .

Miao-Hsiang PENG and Jin-King LIU (2000) have studied the groundwater level forecasting with time series analysis. This study investigates the application of time series analysis methods for forecasting groundwater levels. The study site is located in western Taiwan where serious land subsidence has occurred. A series of monthly groundwater level observations made during the period 1993 and 1999 is used for the experiments. Univariate time series models including ARIMA models and the time series decomposition method are applied and the resulting accuracy is compared. Empirical results indicate that groundwater level data series in this study are cyclical. ARIMA models generate more accurate forecasts than the decomposition model. The forecasting of ARIMA models presents the characteristics of trend and seasonal variation.

A. Lehmann and M. Rode (1999) have studied long-term behavior and cross-correlation analysis of water quality parameters of the Elbe River at Magdeburg, Germany. This study analyses weekly data samples from the Elbe River at Magdeburg between 1984 and 1996 to investigate the changes in metabolism and water quality in the Elbe River since the German reunification in 1990. Modeling water quality variables by Univariate time series models such as autoregressive component models and

ARIMA models reveals the improvement of water quality due to the reduction of wastewater emissions since 1990. The models are used to determine the long-term and seasonal behavior of important water quality parameters. A new procedure for testing the significance of a sample correlation coefficient is discussed. The cross-correlation analysis is applied to hydro-physical, biological, and chemical water quality variables of the Elbe River since 1990. Special emphasis is laid on the detection of spurious sample correlation coefficients.

D. P. Solomatine, C. J. Rojas, S. Velickov and J. C. Wüst (2000) have studied Chaos theory in predicting surge water levels in the North Sea. The problem of predicting surge water levels is important for ship guidance and navigation. The data collected in the coastal waters of the Netherlands (Hook of Holland) is analyzed with an objective of making such prediction. It was found that the correlation between data on surge, temperature, air pressure and wind is not sufficient to rely only on the input-output (connectionist) models like neural networks. It appeared that the surge time series in itself has enough information to make predictions. The applied linear prediction methods including autocorrelation and ARIMA models could not provide sufficient accuracy. Features of chaotic behavior were identified in surge, and methods of chaos theory were applied. The predictions are quite accurate (RMSE is 3.6 *cm* for 1 hour, and 6.1 c*m* for 3 hours). Possible techniques allowing for increase of the prediction accuracy and horizon (wavelet analysis, data mining techniques) were also identified.

Rodel (1997) studied the time series analysis of rising underground salt-water from the abandoned Werra potash-mining district in Germany. Time series analysis has provided detailed information about the rates of discharge of salt-water disposed of under the Werra potash-mining district.

Hammett (1988) had studied the water use, land use, stream flow, and water quality characteristics of the Charlotte Harbor inflow area, USA-Florida. Florida is being subjected to increasing environmental stress by rapid population growth and development. So increase in freshwater demand is required. The three major rivers exist there flow into the harbor. Time series analysis was used in one of the rivers stations to forecast a decreasing trend after a long-term analysis. The increased population will require an additional 70 million gal/ day.

Berthouex and Box (1996) had studied time series models for forecasting wastewater treatment plant performance. The time series model has the form of an exponentially weighted moving average (EWMA). The interpretation of the model is that response of the system can be predicted by deviations from the EWMA smoothed values of the predictor variables.

Stegmann, Ehring, and Liem (1978) studied the application of time series analysis in water quality management and the prediction of the changes in water quality of the Oker River in Germany, on the basis of trends in dissolved oxygen and chloride levels. The environmental impacts of external factors were qualified by measuring the difference between measured water quality indicators and the extrapolated values.

Ellis, Ge, and Grasso (1990) studied the time series analysis of wastewater quality. In the treatment facilities, accurate time series forecasts must be available as input data. The wastewater influent variables were analyzed using the application of ARMA and multivariate ARMA.

Paul A. Conrads (1998) has studied the effects of model output time averaging on the determination of the assimilative capacity of the Waccamaw River and Atlantic Intracoastal Waterway near Myrtle Beach, South Carolina. A branched Lagrangian transport model was calibrated and validated for the tidally influenced portions of the Waccamaw and Pee Dee Rivers, Bull Creek, and the Atlantic Intracoastal Waterway near Myrtle Beach, South Carolina. In determining the assimilative capacity of the Atlantic Intracoastal Waterway, 1-hour, 24-hour, 14-day, and 30-day averaging intervals were used. For each averaging interval, point-source loadings in the model were increased until the state dissolved-oxygen standard was violated. Results of the averaging intervals and point-source loadings for two locations were evaluated by comparing time series of dissolved-oxygen concentration at critical locations and longitudinal profiles of average dissolved-oxygen concentration for particular reaches of the system. The concentrations of the oxygen-consuming constituents that can be assimilated vary by 180 percent, depending upon the averaging interval used for interpreting the simulation model output

M. Stein and J. Lloret (1999) have studied forecasting of air and water temperatures using Fishery Statistical Methods to describe and forecast monthly mean air and bottom water temperatures from 3 sites in the Northwest Atlantic region, up to one year in advance. ARIMA (Auto-Regressive-Integrated-Moving- Average) models were developed that accounted for 92% of the total variability in the long-term time series of monthly means of air temperature and 80% for bottom water temperatures.

These models were then used to forecast conditions in 1999 with results showing good agreement between the predicted and observed values of both air and bottom water temperatures. Intervention analysis that models events as step-like features was also carried out. While this provided a better model fit to the observed data series, such events cannot be predicted. Since nearly all fitted interventions appeared during winter (December–March), prediction of temperatures during these months must be viewed with caution. Results showed that the use of ARIMA models yields better forecasts for highly variable time series than simple models based upon averages of previous monthly averages alone.

Hare, S.R. and R.C. Francis (1994) have studied the climate change and salmon production in the Northeast Pacific Ocean. Alaskan salmon stocks have exhibited enormous fluctuations in production during the century. They investigate their hypothesis that large-scale salmon-production variation was driven by the climatic processes in the Northeast Pacific Ocean. Using a time-series analytical techniques known as intervention analysis, they demonstrate that Alaskan salmonids alternate between high and low production regimes. The transition from high (low) regime to low (high) regime is called an intervention. To test for intervention, they first fitted the salmon time series to univariate autoregressive integrated moving average (ARIMA) models. On the basis of tentatively identified climatic regression, potential interventions were then identified and incorporated into the models, and the result was compared with the non-intervention models. A highly significant positive step intervention in the 1970s and a significant negative step intervention in the late 1940s were identified in the Alaska salmon stocks analyzed. We review the evidence for synchronous regime shift in the 1940s and late 1970s that coincide with the shifts in salmon production. Potential mechanism in North Pacific climate processes to salmon production is identified.

Lu Guanghua, Tang Jie, Yuan Xing, and Zhao Yuanhui (2001) have studied Correlation for the structure and biodegradability of substituted benzenes in the Songhua river water. The biodegradability of 47 substituted benzenes was determined by BOD technique. The molecular weight (M_w), the total surface area (*TSA*), the energy of the highest occupied molecular orbital (E_{HOMO}), the heat of formation (H_f), and the moment of dipole of 47 studied compounds were calculated by the quantum chemical method MOPAC6.0-AM1. The ionization constant (pK_a) and n-octanol/water partition coefficient (*logP*) were obtained from Qsar software and Biobyte software, respectively. The quantitative structure-biodegradability relationship studies were developed by linear regression analysis. The correct prediction rate of obtained model is up to 85% for the testing set. It has been shown that the biodegradability of substituted benzenes in natural river water is mainly related to electronic parameter E_{HOMO}, H_f and steric parameter M_w.

P.D. LaValle, V.C. Lakhan, and A.S. Trenhaile (2000) have studied the short term fluctuations of Lake Erie water levels and the El Nio/Southern Oscillation. This study assesses the relationship between short term fluctuations of Lake Erie water levels and the El Niõ/Southern Oscillation (ENSO) using data collected from May 1978 to May 1997. After standardizing the collected data, graphical and Box-Jenkins time series techniques are utilized to assess the temporal interrelationship of the Southern Oscillation Index and Lake Erie water level variables. The statistical results demonstrate that a first-order auto regressive model AR (1) provides the best fit for the data sets of the analyzed variables. Both the graphical and statistical results suggest that short term Lake Erie water levels are fluctuating in response to the two ENSO phases, El Nio and La Niő. Negative values of the Southern Oscillation Index are related to higher lake levels while positive values are associated with lower lake levels.

N. Koning and JC Roos (1999) have studied the continued influence of organic pollution on the water quality of the turbid Modder River. The Modder River is a relatively small river, which drains an area of 7,960 km2, in the central region of the Free State Province, South Africa, and has a mean annual runoff of 184 x 106 m3. Botshabelo is a city that was developed in the catchment area of the river and treated sewage is released into the Klein Modder River. This study determined seasonal and spatial patterns in the system as well as the continued influence that Botshabelo's treated sewage outflow has on the water quality of the river. Box-Jenkins plots were used in every sample point in this study, as well as, plots of the relations between different sample points. It was determined that the Modder River and Klein Modder River follow distinctive seasonal patterns in terms of algal growth and physical factors. There were periods when the waters of the Modder River and Klein Modder River are of acceptable quality. However, outflows from Botshabelo have a detrimental effect on the water quality in terms of nutrient concentrations and algal biomass. The inflow of the Klein Modder River into the Modder River caused an average of 112% increase in phosphate-phosphorus (PO4-P), 171% increase in nitrate-nitrogen (NO3-N) and a 50% increase in chlorophyll-*a* concentration in the Modder River. The long-term detrimental effect of Botshabelo on the system can clearly be seen by comparing predicted nutrient increases with measured values.

2. KING TALAL DAM AND DATA COLLECTION

This chapter is divided into four parts. The first part is about the philosophy of King Talal Dam construction. The second part is about the water supplier of King Talal Dam. The third part is about water characteristics, and the last part is about data collection.

2.1 Philosophy of King Talal Dam Construction.

Since water is the main factor in agricultural production, and the main factor for extension in lands reclamation and since water resources are very limited in Jordan because of very low average of rainfall, it was necessary to think about ways of collecting surface water by dams to be constructed downstream rivers and valleys to be distributed to farms when required. In 1977 it was agreed to construct King Talal Dam as the first trial on Zarka River with a capacity of 56 MCM, which was increased to 86 MCM in 1988, 78 MCM of which were used for irrigation purposes (Salameh, 1996). The height of the dam was 108 m, the catchment area was 3157 Km^2 with a population of 2.439 million Capita (1996). This dam is considered the most important water project in Jordan. The water stored in it is used for irrigation of wide areas in the Ghor, the area irrigated by water coming from the dam mixed with water coming by King Abdullah channel is estimated by 100,000 donums. (RSS, 1988-2001).

Figure (1) shows the site of King Talal Dam, the borders of the catchment area and the branches that supply the lake behind the dam. The main branches

Figure (1): King Talal's catchment area and branches. (RSS reports.1988 till 2001) are Al Zarka River (main supplier), Al Dhlale Valley that pours in Al Zarka River, and Rumaimeen Valley that pours directly in the lake. The percent of water drained from Al Zarka River to the lake of the river is about 90%, while for Rumaimeen Valley it is about 10%. Population of the catchment area is about half of the kingdom population, which is distributed in Amman, Zarka, Sweileh, Rusaifeh, Baka'a and Jerash with several villages around these cities. (RSS, 1988-2001).

2.2 Water Suppliers for King Talal Reservoir:

The only two water suppliers for King Talal reservoir are Zarka River and Wadi Rumaimeen.

2.2.1 Zarka River

It is a river coming from the east directly to the lake of the dam (see figure 2). Zarka River is the main water branch of King Talal Reservoir. The water in Zarka River consists of the water in Wadi Dhuliel, the effluent from Al Samra, and Jerash Wastewater Treatment plant (see Figure 2). In addition to the treated and untreated effluents, which are drained from industrial factories and farms located on the river's banks. The underground water and rainfall affect water quality of this River.

2.2.2 Wadi Rumaimeen

It is a wadi coming from the south directly to the lake of the dam (see figure 2). This wadi was connected in 1988 to increase the capacity of the lake and to study the effect of AL Baka'a waste water plant on the quality of the

Figure (2): King Talal's reservoir inlets and outlets. (RSS reports.1988 till 2001) water in this valley and finally on the water in the lake and the springs water extending around this wadi that supply it with extra water.

2.3 Water Characteristics

There are three characteristics of water, these characteristics are: physical, chemical, and biological. According to these three characteristics, one can determine the usage of water. Water usage is generally divided into four categories: (1) domestic (water used for sanitary and general purposes, (2) industrial (no domestic purposes), (3) public service (water used for irrigation, fire fighting, and industrial system), (4) unaccounted for system losses and leakage . In this section, the water characteristics will be discussed for each of the five properties.

2.3.1 Total suspended solids (TSS):

TSS has physical characteristics; the term of the total suspended solids refers to non-filterable residue that is retained on a glass–fiber disk after filtration of a sample of water or wastewater. A measured portion of a sample is drown through a glass-fiber filter, retained in a funnel, applying a vacuum to the section flask under the filter with dump suspended solids adhering to the surface is transferred from the filtration apparatus to an aluminum or stainless steel planked as a support. After drying at 103º - 105º C in an oven the filtered dried suspended solids in milligrams divided by the volume of the sample by liters gives the total suspended solids expressed by milligrams per liters. (Viessman,1985).

The total suspended solids fraction consists of the particulate matter with an approximate size range from 0.001 to 1 µm. The dissolved solids consist of both organic and inorganic molecules and ions that are present in true solution in water (Metcalf and eddy, 1991).

2.3.2 Biochemical oxygen demand (BOD):

BOD has chemical characteristics. Biochemical oxygen demand is the quantity of oxygen used in the aerobic stabilization of wastewater and polluted waters. The standard 5-days BOD value is commonly used to measure the amount of pollution in wastewater, to evaluate the efficiency of treatment by measuring oxygen demand remaining in the effluent, and to determine the amount of organic pollutant in surface waters (Viessman, 1985).

Laboratory analyses of wastewaters and polluted waters are considered using 300 ml of BOD bottle incubated at a room temperature of 20º C. a measured portion is placed in the BOD bottle, then the bottle is filled with aerated dilution water containing phosphate buffer and inorganic nutrients. The sample is then diluted with distilled water, when the sample contains a large population of microorganisms (untreated wastewater, for example) seeding is not necessary. If required, the dilution water is seeded with a bacterial culture that has been acclimated with a bacterial matter or other materials that may be present in the wastewater. Readings are taken for five days or more, the biochemical oxygen demand exerted by a diluted wastewater progress approximately by first-order kinetics. Within 5 days, the oxidation of the carbonaceous organic matter is about 60-70 percent completion. (Metcalf and eddy, 1991).

It should be noted that the initial depletion of dissolved oxygen is the result of carbonaceous oxygen demand resulting from organic matter degradation. If presented in sufficient numbers, nitrifying bacteria exerts a secondary oxygen demand by lags several days behind the start of carbonaceous oxygen demand. (Viessman, 1985).

2.3.3 Chemical oxygen demand (COD):

COD has chemical characteristics, the chemical oxygen demand test is used to measure the content of organic matter of both wastewater and natural matter. The oxygen equivalent to the organic matter that can be oxidized is measured by using a strong chemical oxidized agent in an acidic medium. The test must be performed at an elevated temperature. A catalyst (silver sulfate) is required to aid the oxidation of certain classes of organic compounds. The COD test is also used to measure the organic matter in industrial and municipal wastes that contain compounds that are toxic to biological life. The value of COD is higher than BOD and that is because there are more compounds that can be chemically oxidized than biologically. For many types of wastes, it can be correlated COD with BOD, and this is useful because the value of COD can be determined in 3 hours, where as the value of BOD can be determined in 5 days. (Metcalf and Eddy, 1991).

2.3.4 Total phosphorus (T-P):

T-P has chemical characteristics. It is essential for the growth of algae and other biological organisms and that is because of noxious algal blooms that occur in surface waters, so there is presently much interest in controlling the amount of phosphorus compounds that enter surface waters in domestic and industrial waste discharges of natural run off. (Metcalf and Eddy, 1991)

2.3.5 Total nitrogen (T-N):

T-N has chemical characteristics, and it is essential for the growth of the biological characteristics; such as: Protista and plants. This happens because nitrogen is an essential factor in building the synthesis of protein.

In nature the nitrogen is presented in several ways, in water and wastewater, it is combined in proteinacous matter and urea. Nitrogen data will be required to evaluate the treatability of wastewater by biological processes, also the amount of nutrients should be controlled because of the algal growth. (Metcalf and eddy, 1991).

2.4 Data Collection:

The data collected in this thesis was collected from reports made by Jordan Valley Authority in cooperation with Royal Scientific Society. These reports contained monthly information covering the period from January 1988 till December 2000 (156 months) for two sights: Zarka River and Samra Waste water treatment plant. The information includes the monthly analysis of quality of the flow entering the dam through AL Zarka River and Samra Waste Water Treatment Plant effluent. The data contains the concentrations of total suspended solids in (mg/l), biochemical oxygen demand in (mg/l), chemical oxygen demand (mg/l), total phosphorus (mg/l), total nitrogen (mg/l), and finally the rate of flow (MCM/month).

Table 1: Data of water entering King Talal Reservoir from Zarka River

YEAR	MONTH	Zarka River Flow MCM/month	TSS mg/l	BOD5 mg/l	COD mg/l	T-P mg/l	T-N mg/l
1988	January	10.202	227.00	47.00	118.00	3.00	52.70
	February	31.634	42.00	10.00	37.00	2.20	21.50
	March	10.411	34.00	6.00	16.00	2.70	27.30
	April	5.033	41.00	34.00	96.00	4.50	20.20
	May	5.306	72.00	36.00	119.00	4.90	23.70
	June	4.43	83.00	33.00	128.00	5.20	20.20
	July	3.962	103.00	35.00	123.00	5.90	13.90
	August	3.325	60.00	17.00	79.00	6.10	11.00
	September	3.2	55.00	36.00	103.00	7.10	19.90
	October	3.306	47.00	30.00	94.00	6.50	27.10
	November	3.664	61.00	26.00	155.02	5.80	42.90
	December	16.348	72.00	36.00	103.00	6.70	34.60
	Monthly Average	8.402	72.34	24.40	76.91	4.15	27.64
	Total	100.824	72.34	24.40	76.91	4.15	27.64
1989	January	7.247	83.00	12.00	51.00	3.10	23.21
	February	5.101	27.00	11.50	63.99	5.05	30.02
	March	6.825	30.00	22.00	58.50	5.70	35.75
	April	4.829	36.50	14.50	64.50	5.67	36.93
	May	3.87	59.00	37.00	127.51	6.50	33.57
	June	2.878	55.00	25.00	107.50	5.89	20.21
	July	3.298	60.49	25.00	103.98	8.10	27.10
	August	3.15	54.01	18.50	97.51	6.54	22.13
	September	2.967	52.00	16.00	72.00	11.80	28.24
	October	2.946	64.99	16.00	87.49	8.36	32.96
	November	3.555	74.99	18.00	82.49	8.97	45.46
	December	4.953	45.50	21.00	91.50	8.60	51.94
	Monthly Average	4.302	52.58	19.07	78.99	6.55	32.75
	Total	51.619	52.58	19.07	78.99	6.55	32.75
1990	January	7.539	35.00	13.00	64.00	8.10	47.40
	February	7.559	43.00	19.00	68.00	8.10	29.00
	March	7.901	40.00	15.00	56.00	5.70	30.70
	April	5.073	36.00	22.00	80.00	6.54	35.20
	May	3.859	54.00	27.00	74.01	8.10	32.50
	June	3.176	69.99	36.00	112.99	10.40	31.40
	July	3.607	50.00	33.00	99.99	8.50	25.90
	August	3.568	91.99	25.00	100.99	8.00	22.80
	September	3.2	93.99	30.00	112.98	9.90	19.10
	October	3.528	125.00	42.00	116.00	10.20	32.50
	November	3.789	92.00	36.00	119.99	10.10	40.80
	December	4.269	82.00	39.00	157.00	11.50	43.30
	Monthly Average	4.756	60.75	25.33	89.46	8.39	33.51
	Total	57.066	60.76	25.33	89.47	8.39	33.52
1991	January	8.355	79.00	47.00	132.00	10.50	39.40
	February	6.941	20.00	29.00	125.00	8.13	36.10
	March	8.773	47.00	11.00	54.00	3.70	35.00
	April	4.061	51.00	28.00	78.00	7.45	29.00

Cont. Table 1:Data of water entering King Talal Reservoir from Zarka River.

YEAR	MONTH	Zarka River Flow MCM/month	TSS mg/l	BOD5 mg/l	COD mg/l	T-P mg/l	T-N mg/l
1991	May	3.871	36.00	30.00	88.01	8.70	27.70
	June	3.244	38.00	26.00	95.00	5.90	21.30
	July	2.86	57.99	21.00	88.98	7.78	19.50
	August	2.901	56.01	21.00	116.01	7.58	28.00
	September	2.841	72.50	25.17	126.87	8.51	27.19
	October	3.22	85.01	44.01	92.01	9.25	31.00
	November	3.935	104.99	40.00	115.99	10.80	41.00
	December	33.86	28.00	11.00	35.00	4.30	39.50
	Monthly Average	7.073	45.36	22.06	74.23	6.30	35.27
	Total	84.882	45.36	22.02	74.21	6.30	35.27
1992	January	24.598	59.00	23.00	51.00	3.40	17.00
	February	68.544	65.00	19.00	59.00	2.05	19.00
	March	29.245	71.00	15.00	67.00	0.70	21.00
	April	15.179	59.00	21.00	60.00	3.70	28.80
	May	10.266	89.00	25.00	73.00	4.60	27.00
	June	8.693	118.00	28.00	85.00	5.50	25.20
	July	6.909	102.99	33.00	85.99	6.40	21.25
	August	6.596	86.98	36.99	86.98	7.30	17.30
	September	6.187	90.00	24.00	93.00	6.84	26.20
	October	6.537	92.99	11.00	98.99	6.37	35.00
	November	8.206	75.00	13.00	77.00	6.29	26.80
	December	14.78	56.00	15.00	54.00	6.20	18.50
	Monthly Average	17.145	71.50	20.18	65.50	3.48	21.45
	Total	205.737	71.50	20.15	65.51	3.48	21.45
1993	January	14.694	56.00	15.00	54.00	6.20	18.50
	February	11.175	54.00	22.00	84.00	5.04	27.51
	March	9.936	57.00	33.00	91.00	5.67	27.70
	April	8.196	59.00	43.00	97.00	6.30	27.88
	May	8.179	65.00	38.00	116.01	7.77	30.55
	June	8.385	70.00	33.00	135.01	9.24	33.22
	July	5.015	74.00	44.00	116.00	9.62	35.76
	August	5.136	76.99	54.00	96.99	10.00	38.29
	September	4.887	86.00	47.00	129.01	10.91	40.71
	October	5.318	93.99	40.00	160.99	11.82	43.12
	November	7.513	81.00	45.00	122.00	9.96	41.90
	December	6.992	67.00	50.00	83.00	8.10	40.68
	Monthly Average	7.952	66.64	35.21	100.35	7.79	31.59
	Total	95.427	66.63	35.23	100.35	7.79	31.59
1994	January	9.827	47.00	45.00	86.00	7.16	38.60
	February	8.346	26.00	39.00	88.00	6.21	36.52
	March	8.184	60.00	28.00	90.00	6.10	40.62
	April	5.201	37.00	58.00	127.00	6.81	37.51
	May	4.561	63.99	42.00	127.99	12.61	49.55
	June	4.017	67.00	52.00	123.00	13.96	48.20
	July	3.344	82.99	60.99	120.98	13.37	48.81
	August	3.803	80.00	63.00	149.01	12.99	47.46
	September	3.395	93.99	47.00	115.99	14.71	50.20

Cont. Table 1:Data of water entering King Talal Reservoir from Zarka River

YEAR	MONTH	Zarka River Flow MCM/month	TSS mg/l	BOD5 mg/l	COD mg/l	T-P mg/l	T-N mg/l
1995	October	4.844	79.00	28.00	166.00	12.90	53.74
	November	26.245	22.00	45.00	133.00	5.89	40.27
	December	18.312	37.00	15.00	39.00	5.98	36.11
	Monthly Average	8.34	44.58	38.90	104.20	7.91	41.21
	Total	100.077	44.58	38.90	104.20	7.91	41.21
1996	January	5.989	62.95	92.00	101.02	8.01	52.10
	February	8.115	60.01	32.04	83.06	7.02	37.09
	March	7.693	50.96	41.99	88.00	8.71	47.06
	April	6.653	41.94	55.01	94.99	8.27	47.35
	May	6.568	33.95	61.97	103.99	10.66	50.24
	June	5.197	72.93	50.03	113.91	11.55	47.91
	July	5.07	111.05	54.04	140.04	13.41	69.23
	August	5.117	86.97	45.93	154.97	14.27	39.48
	September	4.785	84.01	51.62	174.92	3.34	50.37
	October	4.766	71.97	45.95	147.08	12.38	58.54
	November	5.165	70.09	18.97	143.08	11.62	49.76
	December	6.413	79.06	36.96	104.01	10.92	55.98
	Monthly Average	5.961	66.60	48.65	116.09	9.73	49.82
	Total	71.532	66.54	48.64	116.17	9.80	49.75
1996	January	11.057	97.04	38.98	114.05	5.06	46.40
	February	6.058	32.02	47.05	117.04	8.58	44.24
	March	9.268	33.99	43.05	114.05	10.36	49.74
	April	5.624	54.94	65.08	91.93	9.42	55.30
	May	5.566	169.96	63.06	127.02	10.96	71.69
	June	4.516	81.93	54.03	110.05	11.07	65.54
	July	4.072	52.06	33.89	107.07	8.10	54.52
	August	3.807	112.95	69.87	127.92	11.56	79.33
	September	4.242	151.11	74.02	181.05	13.44	68.84
	October	3.688	75.11	33.08	120.93	11.12	70.23
	November	7.934	136.00	95.03	275.02	15.00	76.76
	December	6.859	73.04	80.04	188.07	10.93	63.27
	Monthly Average	6.058	87.32	57.94	142.46	10.07	60.09
	Total	72.691	87.30	58.03	142.45	10.14	60.06
1997	January	18.864	127.01	125.00	231.98	12.72	66.05
	February	11.551	70.04	41.04	72.03	8.57	43.98
	March	8.071	187.96	85.00	223.02	8.67	55.88
	April	5.275	73.93	68.06	115.07	6.26	37.54
	May	5.45	55.96	44.04	79.08	7.89	48.26
	June	4.826	98.01	77.08	156.03	9.12	63.41
	July	4.863	92.95	56.96	104.05	9.05	49.15
	August	4.839	75.02	36.99	93.20	11.57	61.38
	September	4.674	65.90	40.01	90.07	12.84	57.55
	October	4.609	70.95	31.89	116.08	14.10	62.05
	November	6.854	106.07	46.98	147.07	10.50	60.84
	December	10.711	43.97	135.00	302.96	12.32	52.56
	Monthly Average	7.549	94.32	77.76	166.38	10.60	55.64
	Total	90.587	94.26	77.78	166.39	10.56	55.65

Cont. Table 1:Data of water entering King Talal Reservoir from Zarka River

YEAR	MONTH	Zarka River Flow MCM/month	TSS mg/l	BOD5 mg/l	COD mg/l	T-P mg/l	T-N mg/l
1998	January	10.245	52.03	29.97	68.03	7.81	53.29
	February	6.073	53.02	30.96	126.96	9.22	45.94
	March	9.734	66.98	36.98	99.96	8.32	51.78
	April	5.943	37.02	25.07	88.00	11.44	47.96
	May	3.661	39.06	31.96	93.96	8.47	57.36
	June	3.45	6.96	57.97	184.93	9.86	53.62
	July	3.777	96.90	45.01	88.96	12.71	54.28
	August	4.079	71.10	57.12	103.95	10.05	46.58
	September	4.394	68.96	38.92	94.90	9.33	53.03
	October	4.518	71.93	59.10	106.91	10.18	57.55
	November	4.428	63.01	51.94	111.11	9.49	65.72
	December	4.5	48.00	47.11	152.00	11.78	76.89
	Monthly Average	5.4	56.67	40.19	104.63	9.63	54.63
	Total	64.802	56.68	40.18	104.67	9.58	54.54
1999	January	7.318	40.04	40.99	203.06	12.16	66.41
	February	9.728	39.99	24.98	86.04	8.53	66.51
	March	5.202	33.06	62.09	104.00	8.84	68.63
	April	4.201	74.98	64.03	194.00	12.62	77.60
	May	4.798	67.94	61.90	255.94	12.71	66.07
	June	3.961	89.88	77.00	268.11	14.64	62.11
	July	3.14	80.89	42.04	230.89	11.46	50.32
	August	3.317	106.12	56.07	177.87	9.04	61.80
	September	3.529	79.91	30.04	207.14	14.45	54.41
	October	4.37	97.03	30.89	102.97	13.73	59.50
	November	4.68	64.10	35.90	176.07	16.03	72.22
	December	5.551	--	43.06	206.99	21.80	79.45
	Monthly Average	4.983	63.21	45.15	174.59	12.84	66.43
	Total	59.795	63.18	45.20	174.55	12.76	66.43
2000	January	12.228	44.00	35.98	180.98	12.43	63.13
	February	6.552	58.00	39.07	219.02	8.55	67.00
	March	7.031	33.00	22.05	164.98	9.53	62.58
	April	4.31	54.06	57.08	229.93	12.76	77.49
	May	4.226	78.09	57.97	136.06	12.07	79.51
	June	3.971	67.99	45.08	184.08	11.08	57.92
	July	3.651	87.92	30.13	205.97	19.99	64.64
	August	3.275	105.95	37.86	258.93	15.57	59.85
	September	3.355	85.84	34.87	298.96	12.22	64.68
	October	7.113	85.06	58.06	162.94	13.07	73.67
	November	4.384	57.03	18.93	135.95	14.14	69.34
	December	8.792	68.02	40.04	226.00	20.02	79.85
	Monthly Average	5.741	63.75	39.54	195.26	13.41	68.63
	Total	68.888	63.76	39.47	195.26	13.36	68.66

Table 2: Data of Samra Wastewater Treatment Plant Effluent

YEAR	MONTH	Samra WWTP flow MCM/month	TSS mg/l	BOD5 mg/l	COD mg/l	T-P mg/l	T-N mg/l
1988	January	1.88	184.0	127.0	338.9	13.0	123.2
	February	2.30	168.0	144.0	391.1	13.0	99.2
	March	2.18	157.0	126.0	343.0	13.0	84.2
	April	1.79	180.0	142.0	366.0	18.0	76.2
	May	1.68	204.1	173.0	351.1	18.5	99.2
	June	1.87	210.0	114.0	372.0	8.0	95.2
	July	2.32	190.0	112.0	348.9	19.0	68.2
	August	2.07	209.0	126.0	339.1	15.0	79.2
	September	2.15	195.0	157.0	355.9	17.9	68.2
	October	2.17	191.0	162.0	358.0	16.5	89.3
	November	1.99	193.0	132.0	298.0	17.4	79.2
	December	2.50	195.0	135.0	354.9	17.0	88.2
	Monthly Average	2.07	189.2	137.1	351.8	15.6	86.9
	Total	24.88	2270.3	1644.8	4221.1	186.8	1043.4
1989	January	2.27	126.0	102.0	344.0	17.3	95.2
	February	1.96	105.0	124.0	420.1	17.3	102.2
	March	2.33	145.0	145.0	455.0	17.6	106.2
	April	2.23	172.0	118.0	340.0	17.1	98.2
	May	2.20	196.0	62.0	324.0	18.4	98.2
	June	2.06	137.0	31.0	275.0	18.9	90.5
	July	2.26	130.0	69.0	276.0	16.0	93.9
	August	2.54	144.0	64.0	263.0	13.8	95.5
	September	2.08	170.0	93.1	293.0	16.0	91.6
	October	2.19	203.0	64.0	350.0	13.0	103.1
	November	2.25	221.0	107.0	335.0	21.3	109.3
	December	2.40	172.0	95.0	332.0	8.6	112.0
	Monthly Average	2.23	160.4	89.5	333.2	16.2	99.8
	Total	26.76	1924.9	1074.4	3998.5	194.2	1197.4
1990	January	2.45	137.0	123.0	346.0	15.0	108.0
	February	2.29	109.0	154.0	375.0	12.0	100.0
	March	2.39	124.0	134.0	361.0	15.8	89.0
	April	2.81	160.0	110.0	345.0	17.2	94.0
	May	2.46	189.0	52.0	230.0	19.2	92.0
	June	2.41	173.0	98.0	289.0	19.0	93.5
	July	2.74	147.0	74.0	282.0	16.4	102.0
	August	2.46	203.0	88.0	348.0	18.5	86.8
	September	2.41	198.0	94.0	256.0	22.6	98.8
	October	2.40	211.0	112.0	255.0	22.8	97.1
	November	2.47	242.0	111.0	352.0	20.6	102.0
	December	2.36	232.0	112.0	416.0	24.0	111.0
	Monthly Average	2.47	176.9	104.6	320.7	18.6	97.8
	Total	29.64	2122.4	1254.8	3848.7	222.8	1173.7
1991	January	2.67	246.8	146.8	412.3	32.8	123.1
	February	2.46	56.5	81.9	353.2	23.0	102.0
	March	2.70	152.8	35.8	175.5	12.0	113.8

Cont. Table 2: Data of Samra Wastewater Treatment Plant Effluent

YEAR	MONTH	Samra WWTP flow MCM/month	TSS mg/l	BOD5 mg/l	COD mg/l	T-P mg/l	T-N mg/l
1991	April	2.36	87.8	48.2	134.3	12.8	49.9
	May	2.50	55.7	46.4	136.0	13.4	42.8
	June	2.64	46.8	32.0	116.9	7.3	26.2
	July	2.41	68.7	24.9	105.5	9.2	23.1
	August	2.53	64.3	24.1	133.2	8.7	32.2
	September	2.43	84.7	29.4	148.2	9.9	31.8
	October	2.40	114.2	59.1	123.6	12.4	41.7
	November	2.39	172.6	65.8	190.7	17.8	67.4
	December	2.85	333.2	130.9	416.4	51.2	470.0
	Monthly Average	2.53	126.9	61.6	207.7	18.0	98.7
	Total	30.34	1523.1	739.2	2491.8	216.0	1184.3
1992	January	3.24	70.0	134.0	323.0	14.9	70.0
	February	3.15	78.0	149.0	328.0	13.0	82.1
	March	3.29	73.0	177.0	289.0	13.9	51.0
	April	3.26	144.0	117.0	367.0	19.6	84.4
	May	3.35	218.0	123.0	424.0	21.0	87.2
	June	3.21	221.0	106.0	391.0	21.3	76.8
	July	3.36	205.0	99.0	338.0	20.9	73.1
	August	3.35	169.0	83.0	291.0	17.9	77.1
	September	3.25	146.0	84.0	250.0	19.0	75.2
	October	3.40	177.0	87.0	303.0	18.6	81.5
	November	3.16	167.0	135.0	320.0	17.8	78.0
	December	3.36	135.0	112.0	372.0	16.9	88.5
	Monthly Average	3.28	150.2	117.2	333.0	17.9	77.1
	Total	39.38	1803.0	1406.0	3996.0	214.8	924.9
1993	January	3.45	98.0	169.0	404.0	17.1	80.0
	February	3.10	78.0	200.0	448.0	15.0	79.0
	March	3.41	91.0	248.0	501.0	18.6	87.0
	April	3.26	140.0	188.0	500.0	21.3	93.6
	May	3.47	187.0	134.0	440.0	23.0	93.4
	June	3.23	182.0	92.0	339.0	22.6	80.6
	July	3.34	183.0	118.0	325.0	21.0	75.4
	August	3.36	167.0	117.0	341.0	23.7	75.6
	September	3.26	267.0	106.0	288.0	19.6	73.4
	October	3.38	174.0	96.0	324.0	19.9	75.0
	November	3.32	146.0	133.0	313.0	18.8	79.0
	December	3.46	107.0	134.0	398.0	20.0	90.0
	Monthly Average	3.34	151.7	144.6	385.1	20.0	81.8
	Total	40.03	1820.0	1735.0	4621.0	240.6	982.0
1994	January	3.39	79.0	177.0	428.0	20.7	94.4
	February	2.93	66.0	162.0	423.0	17.8	90.0
	March	3.19	111.0	102.0	398.0	18.8	76.4
	April	3.07	213.0	117.0	390.0	19.7	78.1
	May	3.18	164.0	85.0	398.0	20.5	77.2
	June	3.07	193.0	90.0	331.0	21.3	80.6
	July	3.42	169.0	41.0	311.0	20.7	79.1

Cont. Table 2: Data of Samra Wastewater Treatment Plant Effluent

YEAR	MONTH	Samra WWTP flow MCM/month	TSS mg/l	BOD5 mg/l	COD mg/l	T-P mg/l	T-N mg/l
1994	August	3.41	144.0	40.0	324.0	20.0	71.1
	September	3.50	143.0	68.0	289.0	19.8	76.5
	October	3.73	149.0	39.0	379.0	19.6	76.8
	November	3.69	134.0	93.0	410.0	17.6	76.4
	December	4.02	103.0	117.0	292.0	15.5	76.1
	Monthly Average	3.38	139.0	94.2	364.4	19.3	79.4
	Total	40.61	1668.0	1131.0	4373.0	232.0	952.7
1995	January	4.11	62.1	134.9	322.9	15.6	81.1
	February	3.70	68.9	104.9	410.1	15.7	93.0
	March	4.13	98.1	194.9	409.9	17.9	90.1
	April	3.83	136.0	120.1	402.9	19.6	89.0
	May	3.67	157.0	107.9	340.1	19.1	85.0
	June	3.49	158.1	56.0	447.1	18.4	39.3
	July	3.43	153.1	44.9	313.1	19.2	79.0
	August	3.44	167.0	63.9	340.0	20.0	78.1
	September	3.47	150.0	45.9	242.9	19.6	72.1
	October	3.45	145.9	67.0	291.0	18.9	82.1
	November	3.58	139.0	55.9	360.0	17.9	86.4
	December	3.58	75.1	137.9	415.1	17.0	89.1
	Monthly Average	3.66	125.9	94.5	357.9	18.2	80.4
	Total	43.86	1510.3	1134.2	4295.2	218.9	964.4
1996	January	4.27	70.1	150.0	462.0	15.5	99.1
	February	3.98	72.1	193.1	423.0	13.8	92.9
	March	4.29	74.1	161.9	468.1	16.3	98.1
	April	3.99	79.0	193.1	522.9	18.6	103.1
	May	3.83	109.9	216.1	539.9	18.8	108.0
	June	3.64	120.0	132.9	414.1	18.9	112.0
	July	3.59	145.0	54.0	341.0	18.4	106.1
	August	3.61	140.9	140.0	424.0	17.5	100.1
	September	3.62	164.0	138.0	401.2	18.8	104.1
	October	4.29	194.1	103.9	364.0	20.0	107.9
	November	3.73	106.9	266.1	694.0	20.6	110.9
	December	3.75	102.9	330.9	790.1	21.1	113.1
	Monthly Average	3.88	114.9	173.3	487.0	18.2	104.6
	Total	46.60	1379.0	2080.1	5844.5	218.2	1255.4
1997	January	3.72	243.1	182.0	583.8	19.1	109.0
	February	3.38	197.0	365.9	621.9	16.0	104.0
	March	3.74	96.0	249.9	604.9	16.0	102.1
	April	3.14	114.0	308.1	584.9	17.8	107.0
	May	3.87	129.9	205.0	492.9	18.1	112.1
	June	3.79	185.1	156.1	452.1	19.0	101.9
	July	4.77	139.0	88.0	250.1	17.0	70.0
	August	4.28	140.0	114.1	225.1	18.0	76.0
	September	4.02	153.0	98.0	227.1	19.4	97.0
	October	4.27	144.9	82.1	257.9	20.4	80.9
	November	4.17	94.1	79.9	271.0	20.6	71.1

Cont. Table 2: Data of Samra Wastewater Treatment Plant Effluent

YEAR	MONTH	Samra WWTP flow MCM/month	TSS mg/l	BOD5 mg/l	COD mg/l	T-P mg/l	T-N mg/l
1997	December	4.70	64.0	186.1	265.0	18.9	90.0
	Monthly Average	3.99	141.7	176.3	403.1	18.4	93.4
	Total	47.83	1700.1	2115.2	4836.7	220.4	1121.1
1998	January	4.87	61.0	229.0	324.9	17.9	94.7
	February	3.83	77.1	176.9	355.9	18.6	98.0
	March	4.72	111.9	177.0	382.0	19.3	109.4
	April	4.30	78.0	149.9	274.0	18.4	93.4
	May	4.58	105.9	161.0	318.0	19.0	93.5
	June	4.46	121.0	138.0	285.9	22.9	99.0
	July	4.63	128.1	94.0	252.9	21.0	93.1
	August	4.36	145.0	89.0	208.1	17.2	128.9
	September	4.92	134.1	64.0	329.9	27.0	130.0
	October	4.80	140.9	89.1	305.0	14.4	95.5
	November	4.40	119.0	67.9	265.9	14.1	100.1
	December	4.65	94.9	111.9	384.9	19.8	107.0
	Monthly Average	4.54	109.7	129.0	307.3	19.1	103.6
	Total	54.51	1316.9	1547.8	3687.5	229.4	1242.6
1999	January	4.61	47.9	164.9	399.0	21.3	108.9
	February	4.07	76.1	107.0	318.9	18.4	105.1
	March	4.56	56.0	126.4	478.1	14.9	106.9
	April	4.24	96.0	91.1	521.1	25.0	118.0
	May	4.35	111.1	80.0	327.0	20.0	106.3
	June	4.17	124.9	82.0	247.9	22.1	99.3
	July	4.44	105.0	114.9	454.9	24.1	93.5
	August	4.36	142.0	75.0	253.0	20.6	97.9
	September	4.33	112.0	70.9	382.1	14.6	97.0
	October	4.66	129.0	86.1	258.9	20.0	100.5
	November	4.45	115.1	102.9	351.1	23.6	115.1
	December	4.70	75.9	209.9	364.7	20.0	110.8
	Monthly Average	4.41	99.3	109.3	363.1	20.4	104.9
	Total	52.93	1191.1	1311.2	4356.8	244.5	1259.2
2000	January	4.99	67.0	121.9	427.1	16.2	103.1
	February	4.22	54.9	114.1	339.0	18.0	91.1
	March	4.55	49.0	152.1	386.1	--	--
	April	4.66	117.1	110.0	349.9	21.1	116.2
	May	4.71	111.1	139.1	499.0	--	--
	June	4.40	138.1	--	303.9	20.9	102.7
	July	4.29	151.1	97.0	354.9	--	--
	August	4.33	184.9	54.0	256.9	21.9	102.3
	September	4.53	144.0	81.0	336.9	--	--
	October	4.59	132.9	62.1	269.1	18.3	112.2
	November	4.50	115.9	108.0	416.9	--	--
	December	4.98	109.0	137.1	453.0	22.1	81.3
	Monthly Average	4.56	114.6	106.9	366.1	19.8	101.3
	Total	54.75	1375.0	1283.3	4393.0	237.4	1215.2

3. STATISTICAL ANALYSIS

This Chapter is divided into five sections. The first section about exploration of data, the second section about equations needed for forecasting, the third section about the statistical softwares, the fourth section about the methodology, and the fifth section is about the analysis of data for the variables in king Talal Reservoir and it's relation with AL Samra WWTP.

3.1. Exploration of Data:

The first step of time series analysis is to plot the data in a scatter diagram. Graphic presentation of statistical data gives a pictorial effect. The collected data will generally be complex; it will be very difficult to understand the importance of collected data. Graphic presentation makes the data easy to be understood and grasped. Also it shows if there is any trend that maybe present and the direction in which the trend may change. (Pillai and Bagavalthi, 1997).

The plot of time series contains fluctuations, which means that it is hard to bring order into seemingly hazard movement of the data through time. Nevertheless, in making simplifying assumptions, we can identify, explain and measure fluctuations that appear in the time series. The fluctuations that appear in the plot are due to four basic types of variations: secular trend, seasonal, cyclic and irregular variations (A. Sakakini, 2001). The variation in the data will lead to outlier. In general, an outlier is an observation that is far from the rest of the data.

3.2. Equations Needed in Forecasting:

Through applying mathematical models to the past data, one can analyze the data through mathematical models so that future forecasting of the data can be estimated. The mathematical models used in this thesis were:

3.2.1 Normality of data

3.2.1.1. Weibull distribution:

Is one of the most efficient formulas for computing plotting positions for unspecified distributions, the mathematical formula of Weibull Distribution is:

$$P = \frac{m}{n+1} \qquad \text{.................. (1)}$$

Where:

P: The estimated probability of an X value

m: The Rank

n: The Number of years of record

The technique in all cases is to arrange the data in increasing or decreasing order of magnitude and to assign order number m to the rank values. (Viessman and Lewis, 1996).

3.2.1.2. coefficient of variation process (COV)

The coefficient of variation is often used to describe the relative amount of variation in a population. The sample estimate for the coefficient of variation is defined as:

$$COV = \frac{S}{\bar{X}} \qquad \text{.................. (2)}$$

Where:

COV: The coefficient of variation (dimensionless)

S : Standard deviation

$\bar{X}$: The mean

The coefficient of variable (COV) test is the simplest test available to determine whether the data can be characterized by normal distribution or not (McBean and Rovers, 1998).

3.2.1.3. Kurtosis process:

The fourth moment about the mean is the kurtosis, which is a measure of the peakness of the distribution. The sample estimate of the Kurtosis is obtained from the following equation:

$$K = \frac{n^2 \sum_{i=1}^{n} (X_i - \bar{X})^4}{(n-1)(n-2)(n-3)} \qquad \text{..................(3)}$$

$$C_k = \frac{K}{S^4} - 3 \qquad \text{.................. (4)}$$

Where:

C_k: Coefficient of Kurtosis.

K : The fourth moment about the mean.

n: The number of data.

X_i: Observed data.

$\bar{X}$: Mean.

S : Standard deviation. (McBean and Rovers,1998)

Kurtosis concerns the relative concentration of values in the center of the distribution as compared to the tails. In term of this property, three types of distributions can be defined: leptokurtic, mesokurtic and platykutric. Leptokurtic distribution is characterized by a prominent peak and by a relatively large proportion of values falling in the tail. A mesokurtic distribution is one, which the values are mainly located in the center of the distribution. A platykutric distribution is characterized that the peak is relatively flat and very few values appear in the tails (Levine, Ramsey and Smidt, 2001)

3.2.1.4. Shapiro-Wilk test

The Shapiro-Wilk “W’ test is another statistical goodness of fit test that performs for a small sample (3-50). The steps for calculating W are:

A- Order the sample data

B- Compute a weighted sum (b) of the differences between the most extreme observations

$$b = \sum_{i=1}^{n} a_{n-i+1}(X_{n-i+1} - X_i) \qquad \text{.................. (5)}$$

C- Divide the weighted sum by a multiple of the standard deviation, and square the results to get the Shapiro-Wilk statistics Wi.

$$W = \left\{ \frac{b^2}{S\,(n-1)^{0.5}} \right\} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

D- Compare the computed value of W with five percent critical value for a specific sample size. (McBean and Rovers, 1998).

3.2.2 Order of autoregressive

The value of autoregressive is indicated how the joint distribution of some P consecutive y_t's and the identical distribution of the (independently distributed) u_t's together with:

$$U_t = \sum_{r=0}^{p} \beta_r \, y_{t-r} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (7)$$

Where:

U_t : Random process with mean equals to zero and variance equals to one.

β_r : Unknown parameter with $\beta_0 = 1$ and $|\beta_r| < 1$.

y_{t-r}: observation (data) at time equals to t-r.

For first order case, where p = 1

$$U_t = \beta_0 \, y_t + \beta_1 \, y_{t-1}$$

For $\beta_0 = 1$ and $\rho = \beta_r$ the equation will be

$$y_t = \rho \, y_{t-1} + U_t \qquad \ldots\ldots\ldots\ldots\ldots\ldots (8)$$

For second order case, where P=2, replacing y_{t-1} in equation (7) and replacing t with t-1 (that is, $y_{t-1} = \rho\, y_{t-2} + u_{t-1}$) we obtain:

$$y_t = U_t + \rho\, U_{t-1} + \rho^2\, y_{t-2} \qquad \dots\dots\dots\dots\dots\dots (9)$$

So in general:

$$y_t = U_t + \rho\, U_{t-1} + \dots + \rho^{p-1}\, U_{(t-(p-1))} + \rho^p\, y_{t-p} \qquad \dots\dots\dots\dots\dots\dots (10)$$

Let λ be the lag operator, so the relation between the observation and the lag will be :

$$\lambda\ y_t = y_{t-1} \qquad \dots\dots\dots\dots\dots\dots (11)$$

So equation (7) can be written as:

$$U_t = \sum_{r=0}^{p} \beta_{rr}\, \lambda^r\, y_{t-r} \qquad \dots\dots\dots (12) \qquad \text{(Anderson,1971)}$$

3.2.3 The moving average model

The general formula for the moving average model is given by the equation:

$$y_t = \sum_{j=0}^{q} \alpha_q\, v_{t-j} \qquad \dots\dots\dots\dots\dots (13)$$

Where:

y_t : Current observation.

α_q : unknown parameter, $\alpha_0 = 1$ and $|\alpha_q| < 1$.

v_t: is a sequence of independent random variables with $\xi\, v_t = 0$ and $\xi\, v_t^2 = \sigma^2$.

When adding the lag parameter (λ) into equation (13), so the equation can be written as:

$$y_t = \sum_{j=0}^{q} \alpha_q\, \lambda^j\, v_{t-j} \qquad \dots\dots\dots\dots (14) \qquad \text{(Anderson,1971)}$$

3.2.4 Autoregressive process with moving average residuals: -

This model is summarized by the following equation

$$\sum_{j=0}^{q} \alpha_q \, v_{t-j} = \sum_{r=0}^{p} \beta_r \, y_{t-r} \qquad \ldots\ldots\ldots\ldots\ldots\ldots \ (15)$$

Equation (14) is valued when lag parameter is not used, when using lag parameter the equation will be

$$\sum_{j=0}^{q} \alpha_q \lambda^j v_{t-j} = \sum_{r=0}^{p} \beta_r \lambda^r y_{t-r} \qquad \ldots\ldots\ldots\ldots \ (16) \quad \text{(Anderson,1971)}$$

After analysis of equation (15), the final equation that is used in the ARMA modeling is:

$$Y_t = \rho_1 Y_{t-1} + \ldots + \rho_p Y_{t-p} + U_t + \alpha_q U_t + \ldots + \alpha_q U_{t-q} \qquad \ldots\ldots\ldots\ldots\ldots\ldots \ (17)$$

The residual sum of squares can be calculated at every point on a suitable grid of parameter values, and the values that give the minimum sum of squares may then be assessed. (Chatfield, 1984)

3.2.5 Integrated models:

In practice, most time series are non-stationary. In order to fit a stationary model, it is necessary to remove non-stationary sources of variation. If the observed time series is non-stationary in mean then we can difference the series. This is done by replacing Y_t in equation (17) by $\nabla^d X_t$. And representing or replacing $\nabla^d X_t$ with W_t. So the general Autoregressive integrated moving average process (abbreviated ARIMA process) is of the form

$$W_t = \rho_1 W_{t-1} + \ldots + \rho_p W_{t-p} + U_t + \alpha_q U_t + \ldots + \alpha_q U_{t-q} \qquad \ldots\ldots\ldots\ldots\ldots \ (18)$$

The above ARIMA process describing the d^{th} differences of the data is said to be of order (p, d, q). (Chatfield, 1984)

3.2.6 Linear regression

The mechanics of using the regression line to predict Y from X are simple enough once the line has been fitted through the scatter plot. The equation that defines the regression line is shown in the following equation:

$$Y' = a + b_y X \quad \text{……………..} \ (19)$$

Where

Y' : The predicted value of Y

a : The regression constant

b_y : The regression coefficient

X : Observed value

In words, the predicted value of the variable (Y') for any value of the variable X is computed by multiplying X by the regression coefficient (b_y) and adding the regression constant (a). The equation that estimate the regression constant (a) and coefficient (b_y) is called the least square equation. (Diekhoff, 1996).

3.2.7 Quadratic regression

If the relationship was clearly non-linear. In practice you may find it necessary to fit several types of curve to the data and choose the one, which gives the best fit, so if a quadratic relationship is thought to be the appropriate, the regression curve is given by:

$$Y = a_0 + a_1 X + a_2 x^2 \quad \text{……………….} \ (20)$$

Then the quantities a_0, a_1 and a_2 must be estimated from the data. A general method of estimating the parameter is by the method of least squares. (Chatfield, 1978).

3.2.8 Exponential growth regression

If the relationship was clearly non-linear. In practice you may find it necessary to fit several types of curve to the data and choose the one, which gives the best fit, so if an exponential growth relationship is thought to be the appropriate, the regression curve is given by:

$$Y = a_0 e^x \qquad \text{………………. (21)}$$

Then the quantity a_0 must be estimated from the data. A general method of estimating the parameter is by the method of least squares. (Chatfield, 1978).

3.2.9 Single exponential smoothing

The model should only be used in its basic form for non-seasonal time series showing no trend, so if the data showed trend or seasonality then the first step to make single exponential smoothing regression is to eliminate this trend. The single exponential smoothing equation is written as :

$$\hat{Y}_{(N+1)} = \theta\, \hat{Y}_{(N)} + \theta\,(1-\theta)\, \hat{Y}_{(N-1)} + \theta\,(1-\theta^2)\, \hat{Y}_{(N-2)} + \ldots \qquad \text{………..(22)}$$

Where

$\hat{Y}_{(N+1)}$: Estimated value

$\hat{Y}_{(N)}\ \hat{Y}_{(N-1)}\ \hat{Y}_{(N-2)}$:Past values

θ : Smoothing constant , $0< \theta <1$

The value of the smoothing constant, θ, depends on the properties of the given time series. Values between 0.1 and 0.3 is commonly used (Chatfield, 1984).

3.2.10 Testing differences between means

The mechanism of testing the difference between two means is assigning randomly the data to two samples, the first sample is the experimental group, which is affected with the treatment and the second one is the controlled group, which get nothing special. So the difference between the mean is the difference between the average of the first and second sample, mathematically:

$$\text{Difference in mean} = \bar{X}_1 - \bar{X}_2 \qquad \text{.................... (23)}$$

Where

$\bar{X}1$ = The mean of the experimental group.

$\bar{X}_2$ = The mean of the controlled group.

The difference in mean is a good test to see if one must accept or reject the null theory. (Chase, 1976).

3.2.11 Cross and distance correlation

For cross and distance correlation, the population correlation coefficient ρ_{xy} between the two random variables x and y is equal to:

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \ \sigma_y} \qquad \text{.................. (24)}$$

Where

σ: is the variance

3.3 Statistical software:

Many software were established to satisfy statistical purposes, some of these software are: Minitab13, S-Plus 2000, SPSS, Statistica, NCSS 2001, Matlab, and many other software. In this thesis two packages were used: Minitab13 and S-Plus 2000.

Minitab statistical software is designed to help in solving statistical problems. Minitab is a general-purpose statistical system that can be run in either interactive or batch mode, and on mainframe computers, minicomputers or microcomputers. Minitab is easy to learn and easy to use. It is an ideal companion for the student or the worker faced with the need to analyze data and to formulate problems involving random problems. Minitab is a package used for simulation, regression, statistics identifications, fitting, checking, forecasting using ARIMA models, and other applications. (Miller, 1988).

The S-Plus programming provides an interactive computing environment for programming, graphics, and statistical analysis. A variety of data can be stored, manipulated, plotted, and analyzed using build up functions. S plus can easily accommodate and manipulate data stored in ways suitable for direct manipulation as vectors and matrices, as well as the "observations and variables" model used in most statistical packages. This flexibility allows performing a wide range of tasks. S-Plus can be used in statistical analysis, graphics tools, matrices and lists. (Spector, 1994).

3.4. Methodology:

3.4.1 Outliers and missing data

3.4.1.1 The missing data is determined and calculated by the average of the data of the same month.

3.4.1.2 A scatter diagram was plotted using Excel Charts; the diagram was plotted between the variable vs. the time.

3.4.1.3 A box diagram was plotted for the original data by seasonal period and Residuals by seasonal period using the software "Minitab 13"; the outliers are shown in those two diagrams as points.

3.4.1.4 The outliers ware determined in the two cases and checked if they were real outliers or not. The real outliers were not changed; the others were adjusted to the average monthly value.

3.4.1.5 A new scatter diagram was plotted for the new adjusted data using Excel software.

3.4.2 Normality: -

To determine if the data was normal or not in this thesis, four methods were used.

3.4.2.1. Weibull distribution model histogram:

A) The average monthly value for the variable was calculated.

B) Weibull Distribution Model Histogram was drawn using "Minitab 13".

C) The histogram is determined if skewed to left or right or not.

3.4.2.2. coefficient of variance (COV):

A) The data is divided into four quarters.

B) The coefficient of variable issues found for each quarter Excel sheet.

C) The value of the coefficient of variable is compared with [1], if it is less than $|1|$ then the data is not skewed; otherwise it is skewed to left or right.

3.4.2.3. Kurtosis coefficient:

A) The data were divided into four quarters.

B) The Kurtosis Coefficient was found for each quarter using Excel sheet.

C) The value of the Kurtosis Coefficient was compared with $|1|$, if it is less than $|1|$ then the data is called Mesokurtic, if not, the data is called Leptokurtic for the positive sign and Platykutric for the negative sign.

3.4.2.4. Shapiro-Wilk test:

A) The data were divided into four quarters.

B) The Shapiro-Wilk Test was found for each quarter using Excel sheet.

C) The value of the Shapiro-Wilk Test was compared with the value in the Appendix (3), if it is greater than it, the data is normally distributed, if it is less than it, the data is skewed.

3.4.3 Order of (AR): -

The value of AR (p) is determined through drawing the Autocorrelation function for the variable, taking into consideration that the value of (p) should not exceed (1) in surface water forecasting, because the small rivers (as Zarka River) the water characteristics do not need more than few days to dilute, so the correlation does not exceed one month (Viessman and Lewis, 1996). The figures of autocorrelation function is drawn using "Minitab 13" software.

3.4.4 Order of moving average: -

3.4.4.1 The moving average graph for the variables is drawn using Minitab 13 software using different values of (q).

3.4.4.2 The value of (q) is determined when the graph indicates that the trend was minimized and following graph shows no difference with the previous one.

3.4.5 Order of "I": -

3.4.5.1 Four figures containing the original data, detrended, seasonally adjusted data and seasonally adjusted and detrended data using "Minitab 13" software.

3.4.5.2 If the figures show big difference then it needs differentiations, and if they do not show big difference, then there is no need for differentiation.

3.4.5.3 A diagnostic model diagrams for the ARIMA model is drawn using "Minitab 13" software with I = 0 , 2 , because the seasonality in Jordan is affected by just the summer and winter seasons.

3.4.5.4 The ARIMA that has less residual between the two diagnostic models is considered to be the model, which we will use.

3.4.6 Forecasting future values: -

The method of forecasting was divided into two parts; the deterministic and the stochastic forecasting:

3.4.6.1 deterministic forecasting:

Four different methods were used in this section, these methods are: linear regression, quadratic regression, exponential growth regression and single exponential smoothing models. The acceptance of a model is determined if the model has an error less than 10% of the real data.

3.4.6.2 stochastic forecasting:

Three forecasting methods were used in this section, these methods are: autoregressive model, moving average model and the ARIMA model. If the error in the stochastic forecasting is less than 10%, then the model will be the best one even if there is lower error in the deterministic model.

3.4.7 Results of forecasting: -

The results of the error in each model are shown in the percentage of error table, the calculations in this table were based on the difference in mean between the real data and the forecasted one, the stochastic forecasting is determined to be the best model if it satisfies the maximum acceptance error, which equals to 10%, otherwise the lowest error in the in the deterministic model will be the best model.

3.4.8 Cross and distance correlation: -

3.4.8.2 The cross and distance correlation was drawn using Minitab 13 software

3.4.8.3 Analyses were made for each figure.

3.5. Analysis

In this section, the data that was collected will be analyzed; six variables will be analyzed in King Talal Reservoir. These variables are: Total suspended solids (TSS), biochemical oxygen demand (BOD_5), chemical oxygen demand (COD), Total Phosphorus (T-P), and finally the total Nitrogen (T-N). The method that each variable will be analyzed will be the same as discussed in the methodology. The best forecasting method would be determined in the end of each analysis.

3.5.1. Total suspended solids (TSS) variable:

The consequences that were used to analyze the TSS variable were as follows:

3.5.1.1. detection of missing data and outliers:

From the table (1) it is observed that the contains one missing data in December 1999, the observation will be estimated to be the average of the observations of the same month, which is December in this case. The new calculated value is 58.30 mg/l. To detect the outliers, data should be drawn in a scatter diagram (Figure 3) so that outliers will be clearly observed. These data, which contains 156 observations from January 1988 till December 2000, have approximately five outliers and they are in the following months: January 1988, May 1996, September 1996, March 1997, and finally, June 1998. It was observed that the rainfall in 1996 was high, and it is known that when the rainfall is high then the TSS will get higher (Appendix (1)). So the real data are on May and September 1996, the other three data were assumed to be outliers due to human error. They should be adjusted to a new value since they may greatly influence any statistical calculations and yield biased result. The way that outliers were adjusted was the same as the missing data treated and it was equals to the average monthly value.

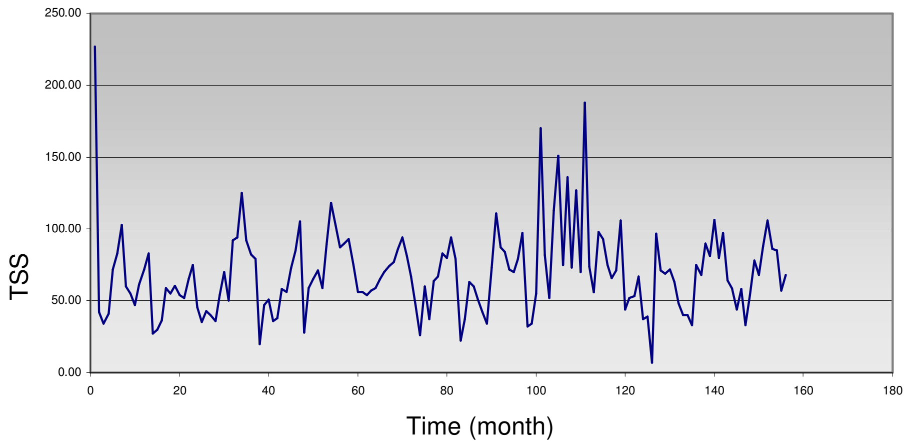


Figure (3): Original Data of TSS

Figure (4) shows the outliers for the seasonal trends for the original and the residual data, one can conclude from the charts that there are five outliers in both the original data and the residual data in the seasonal condition and that they are the same outliers in the scatter diagram of the original data. Also figure (4) shows the variation in the data for the same month, it can be observe the variation was the highest on January, and was the lowest on April.

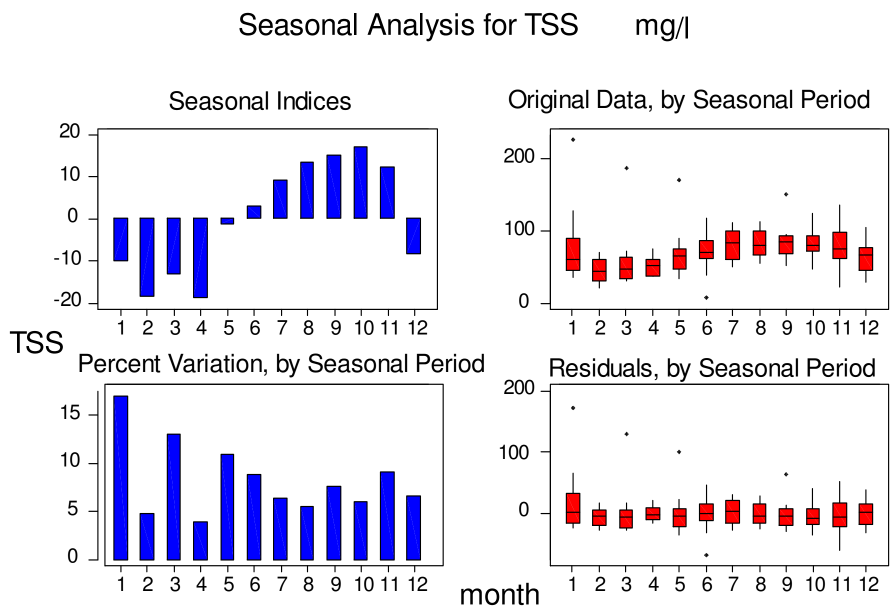


Figure (4) Outliers for Seasonal Analysis for TSS Variable

After adjustment the outliers, the new adjusted data are plotted in Figure (4), the figure shows that there are still outliers but these outliers cannot be omitted because they are real data so it can influence the statistics results. While comparing the old data (Figure 3) with the new adjusted data (Figure 5) it can be observed that two figures are quite the same and they have the same trend, so the effect of the outliers on the data was so little.

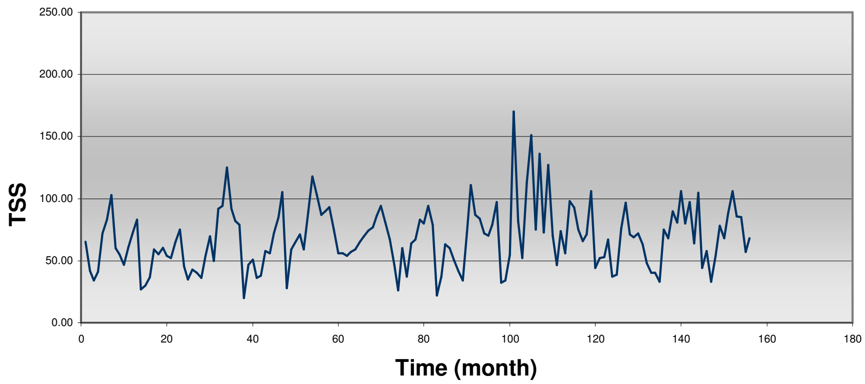


Figure (5): The New Adjusted Data for TSS mg/l

3.5.1.2. normality of data

In this section, normality of data will be checked through four procedures; first one is by drawing a histogram for Weibull's distribution model, second one is through calculating the coefficient of variance, the third one is through calculating the Kurtosis coefficient, and the fourth one is through calculating the Shapiro-Wilk test. From these four procedures, if the data was not normal then a lognormal transformation to the data will be made.

A- Weibull's distribution model histogram:

Data will be transformed to the average monthly value for the TSS variable; the calculated values were as follows

Month	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
TSS mg/l	65.2	45.4	46.4	50.5	68.0	76.0	81.0	81.9	83.0	81.5	77.5	61.9

The Weibull's distribution histogram is drawn for these twelve data. It can be observed from figure (6) that the data of TSS is quite normal and there is a little skewness to the left and bulked to the right, but in general the graph gives an indication that the data is normal.

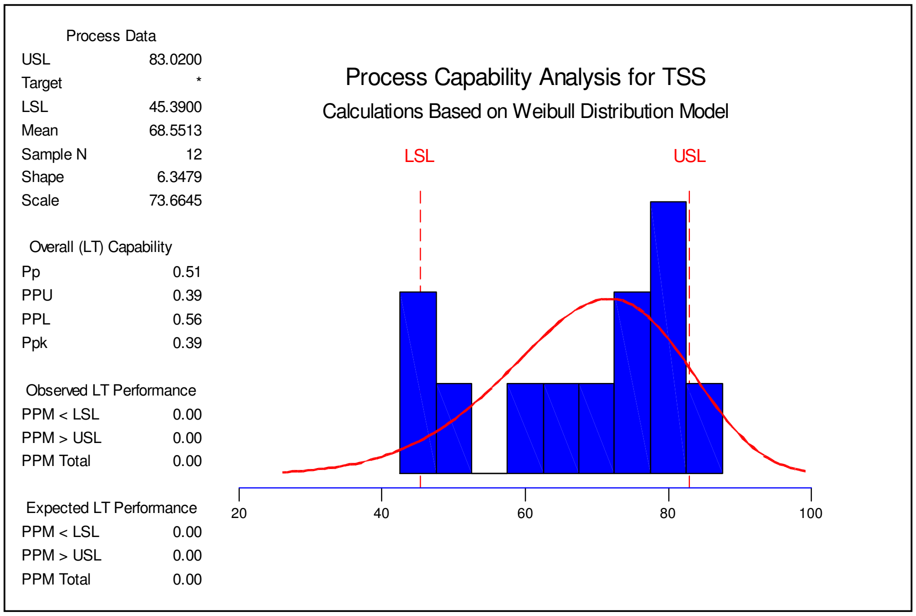


Figure (6) : Weibull Distribution Model Histogram

B- coefficient of variation (COV), preliminary test:

The data were divided into four quarters; each quarter consists of 39 data. Table (3) provides the value of the mean, variance, standard deviation, and the coefficient of variation for the TSS variable.

Table (3): The coefficient of variable for TSS

	MEAN	VARIANCE	ST. DE. (S)	C.O.V.
TSS (1st Quarter)	64.1	1244.5	35.3	0.55
TSS (2nd Quarter)	67.6	442.7	21.0	0.31
TSS (3rd Quarter)	80.5	1332.0	36.5	0.45
TSS (4th Quarter)	65.7	535.5	23.1	0.35

It can be shown from table (3) that the value of the coefficient of variation for each quarter is less than 1, which means that each quarter of the data has a little skewed (either to right or left), so the total data of the TSS variable has less skewness than each of the four TSS quarters, it can be concluded that the TSS variable does not have skewness.

C- Kurtosis coefficient (peakedness), vertical test:

To find the Kurtosis coefficient, one should find the value of K, which depends on the fourth moment about the mean and the number of samples, so that the Kurtosis can be calculated. The Kurtosis will give a good indication if the distribution is leptokurtic or platykutric. The data was divided into four quarters, Table (4) provides the values of the Kurtosis coefficient for each quarter and it provides also the calculations needed to calculate the Kurtosis coefficient, which they are: the mean, the variance or standard error, the value of K, and the Kurtosis coefficient.

Table (4): The Kurtosis Coefficient for TSS

	MEAN	VARIANCE	ST. DE. (S)	K	Kurtosis Coeff. C'_K
TSS (1st Quarter)	64.063	1244.499	35.277	22072292.8	11.25
TSS (2^{nd} Quarter)	67.575	442.716	21.041	607566.2	0.10
TSS (3rd Quarter)	80.456	1331.995	36.497	8230945.2	1.64
TSS (4th Quarter)	65.654	535.502	23.141	833447.0	-0.09

From table (4) one can observe that the data in the first quarter was leptokurtic, in the second quarter it was normally distributed (mesokurtic), in the third quarter it was fairly leptokurtic and in the fourth quarter is was normally distributed.

The total data of the TSS variable can be assumed to be as fairly normally distributed (mesokurtic).

D- Shapiro-Wilk test

This is another test to show that the data we have is normal or not. Data that have been collected were divided into equal quarters, the value of (a_{n-1+I}), was taken for 20 data since the value of n-1+i was equal to 20, the value of (a_{n-1+I}) was taken from Appendix (2). The Shapiro-Wilk value was compared with the five percent critical value for sample size 20 in Appendix (3), if the value of the Shapiro-Wilk test was greater than it then the data will not show evidence of nonnormality.

Table (5): Shapiro-Wilk Test for the Data of TSS's 1st quarter

No	TSS mg/l	Ordering TSS (1)	Inverse order TSS (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
1	65.17	20.00	125.00	105.00	0.3989	41.88
2	42.00	27.00	103.00	76.00	0.2755	20.94
3	34.00	30.00	93.99	63.98	0.2380	15.23
4	41.00	34.00	92.00	58.00	0.2104	12.20
5	72.00	35.00	91.99	56.99	0.1880	10.71
6	83.00	36.00	83.00	47.00	0.1689	7.94
7	103.00	36.50	83.00	46.50	0.1520	7.07
8	60.00	40.00	82.00	42.00	0.1366	5.74
9	55.00	41.00	79.00	38.00	0.1225	4.65
10	47.00	42.00	74.99	32.99	0.1092	3.60
11	61.00	43.00	72.00	29.00	0.0967	2.80
12	72.00	45.50	72.00	26.50	0.0848	2.25
13	83.00	47.00	69.99	23.00	0.0733	1.69
14	27.00	47.00	65.00	18.00	0.0622	1.12
15	30.00	50.00	64.99	15.00	0.0515	0.77
16	36.50	52.00	61.00	9.00	0.0409	0.37
17	59.00	54.00	60.49	6.49	0.0305	0.20
18	55.00	54.01	60.00	5.99	0.0203	0.12
19	60.49	55.00	59.00	4.00	0.0101	0.04
20	54.01	55.00	55.00	0.00		b=139.32
21	52.00	59.00	55.00	-4.00		S=35.277
22	64.99	60.00	54.01	-5.99		
23	74.99	60.49	54.00	-6.49		
24	45.50	61.00	52.00	-9.00		W=0.410< 0.939
25	35.00	64.99	50.00	-15.00		
26	43.00	65.17	47.00	-18.17		Didn't Satisfied
27	40.00	69.99	47.00	-23.00		
28	36.00	72.00	45.50	-26.50		
29	54.00	72.00	43.00	-29.00		
30	69.99	74.99	42.00	-32.99		
31	50.00	79.00	41.00	-38.00		
32	91.99	82.00	40.00	-42.00		
33	93.99	83.00	36.50	-46.50		
34	125.00	83.00	36.00	-47.00		
35	92.00	91.99	35.00	-56.99		
36	82.00	92.00	34.00	-58.00		
37	79.00	93.99	30.00	-63.98		
38	20.00	103.00	27.00	-76.00		
39	47.00	125.00	20.00	-105.00		

Table (6): Shapiro-Wilk Test for the Data of TSS's 2nd quarter

No	TSS mg/l	Ordering TSS (1)	Inverse order TSS (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
40	51.00	26.00	118.00	92.00	0.3989	36.70
41	36.00	28.00	104.99	76.99	0.2755	21.21
42	38.00	36.00	102.99	66.99	0.2380	15.94
43	57.99	37.00	93.99	56.99	0.2104	11.99
44	56.01	38.00	92.99	54.99	0.1880	10.34
45	72.50	47.00	90.00	43.00	0.1689	7.26
46	85.01	51.00	89.00	38.00	0.1520	5.78
47	104.99	54.00	86.98	32.98	0.1366	4.50
48	28.00	56.00	86.00	30.00	0.1225	3.68
49	59.00	56.00	85.01	29.01	0.1092	3.17
50	65.00	56.01	81.00	24.99	0.0967	2.42
51	71.00	57.00	76.99	20.00	0.0848	1.70
52	59.00	57.99	75.00	17.01	0.0733	1.25
53	89.00	59.00	74.00	15.00	0.0622	0.93
54	118.00	59.00	72.50	13.50	0.0515	0.70
55	102.99	59.00	71.00	12.00	0.0409	0.49
56	86.98	60.00	70.00	10.00	0.0305	0.31
57	90.00	63.99	67.00	3.01	0.0203	0.06
58	92.99	65.00	67.00	2.00	0.0101	0.02
59	75.00	65.00	65.00	0.00		b= 128.43
60	56.00	67.00	65.00	-2.00		S= 21.041
61	56.00	67.00	63.99	-3.01		
62	54.00	70.00	60.00	-10.00		
63	57.00	71.00	59.00	-12.00		
64	59.00	72.50	59.00	-13.50		
65	65.00	74.00	59.00	-15.00		
66	70.00	75.00	57.99	-17.01		
67	74.00	76.99	57.00	-20.00		
68	76.99	81.00	56.01	-24.99		
69	86.00	85.01	56.00	-29.01		
70	93.99	86.00	56.00	-30.00		
71	81.00	86.98	54.00	-32.98		
72	67.00	89.00	51.00	-38.00		
73	47.00	90.00	47.00	-43.00		
74	26.00	92.99	38.00	-54.99		
75	60.00	93.99	37.00	-56.99		
76	37.00	102.99	36.00	-66.99		
77	63.99	104.99	28.00	-76.99		
78	67.00	118.00	26.00	-92.00		

W=0.980< 0.939

→ Satisfied

Table (7): Shapiro-Wilk Test for the Data of TSS's 3rd quarter

No	TSS mg/l	Ordering TSS (1)	Inverse order TSS (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
79	82.99	22.00	187.96	165.96	0.3989	66.20
80	80.00	32.02	169.96	137.94	0.2755	38.00
81	93.99	33.95	151.11	117.16	0.2380	27.88
82	79.00	33.99	136.00	102.01	0.2104	21.46
83	22.00	37.00	127.01	90.02	0.1880	16.92
84	37.00	41.94	112.95	71.01	0.1689	11.99
85	62.95	50.96	111.05	60.09	0.1520	9.13
86	60.01	52.06	98.01	45.95	0.1366	6.28
87	50.96	54.94	97.04	42.10	0.1225	5.16
88	41.94	55.96	93.99	38.03	0.1092	4.15
89	33.95	60.01	92.95	32.93	0.0967	3.18
90	72.93	62.95	86.97	24.02	0.0848	2.04
91	111.05	65.90	84.01	18.12	0.0733	1.33
92	86.97	70.04	82.99	12.95	0.0622	0.81
93	84.01	70.09	81.93	11.84	0.0515	0.61
94	71.97	71.97	80.00	8.04	0.0409	0.33
95	70.09	72.93	79.06	6.13	0.0305	0.19
96	79.06	73.04	79.00	5.96	0.0203	0.12
97	97.04	73.93	75.11	1.17	0.0101	0.01
98	32.02	75.02	75.02	0.00		b=215.80
99	33.99	75.11	73.93	-1.17		S=36.50
100	54.94	79.00	73.04	-5.96		
101	169.96	79.06	72.93	-6.13		
102	81.93	80.00	71.97	-8.04		
103	52.06	81.93	70.09	-11.84		
104	112.95	82.99	70.04	-12.95		
105	151.11	84.01	65.90	-18.12		
106	75.11	86.97	62.95	-24.02		
107	136.00	92.95	60.01	-32.93		
108	73.04	93.99	55.96	-38.03		
109	127.01	97.04	54.94	-42.10		
110	70.04	98.01	52.06	-45.95		
111	187.96	111.05	50.96	-60.09		
112	73.93	112.95	41.94	-71.01		
113	55.96	127.01	37.00	-90.02		
114	98.01	136.00	33.99	-102.01		
115	92.95	151.11	33.95	-117.16		
116	75.02	169.96	32.02	-137.94		
117	65.90	187.96	22.00	-165.96		

$W=0.920<0.939$

Didn't Satisfied

Table (8): Shapiro-Wilk Test for the Data of TSS's 4th quarter

No	TSS mg/l	Ordering TSS (1)	Inverse order TSS (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
118	70.95	6.96	106.12	99.16	0.3989	39.56
119	106.07	33.00	106.07	73.07	0.2755	20.13
120	43.97	33.06	105.95	72.89	0.2380	17.35
121	52.03	37.02	104.59	67.57	0.2104	14.22
122	53.02	39.06	97.03	57.96	0.1880	10.90
123	66.98	39.99	96.90	56.91	0.1689	9.61
124	37.02	40.04	89.88	49.84	0.1520	7.58
125	39.06	43.97	87.92	43.95	0.1366	6.00
126	6.96	44.00	85.84	41.84	0.1225	5.13
127	96.90	48.00	85.06	37.06	0.1092	4.05
128	71.10	52.03	80.89	28.87	0.0967	2.79
129	68.96	53.02	79.91	26.89	0.0848	2.28
130	71.93	54.06	78.09	24.03	0.0733	1.76
131	63.01	57.03	74.98	17.96	0.0622	1.12
132	48.00	58.00	71.93	13.94	0.0515	0.72
133	40.04	63.01	71.10	8.09	0.0409	0.33
134	39.99	64.10	70.95	6.85	0.0305	0.21
135	33.06	66.98	68.96	1.98	0.0203	0.04
136	74.98	67.94	68.02	0.07	0.0101	0.00
137	67.94	67.99	67.99	0.00		b=143.76
138	89.88	68.02	67.94	-0.07		S=23.14
139	80.89	68.96	66.98	-1.98		
140	106.12	70.95	64.10	-6.85		
141	79.91	71.10	63.01	-8.09		
142	97.03	71.93	58.00	-13.94		
143	64.10	74.98	57.03	-17.96		
144	104.59	78.09	54.06	-24.03		
145	44.00	79.91	53.02	-26.89		
146	58.00	80.89	52.03	-28.87		
147	33.00	85.06	48.00	-37.06		
148	54.06	85.84	44.00	-41.84		
149	78.09	87.92	43.97	-43.95		
150	67.99	89.88	40.04	-49.84		
151	87.92	96.90	39.99	-56.91		
152	105.95	97.03	39.06	-57.96		
153	85.84	104.59	37.02	-67.57		
154	85.06	105.95	33.06	-72.89		
155	57.03	106.07	33.00	-73.07		
156	68.02	106.12	6.96	-99.16		

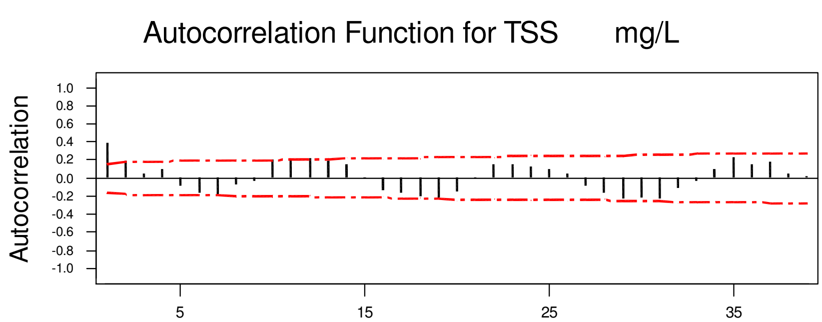
W=1.016< 0.939

Satisfied

From the Tables shown previously, it has been shown that the data in the first quarter is nonnormal, the second quarter has a normal data, the third has a quite nonnormal one, and the fourth one has a normal distribution. It can be safely say that TSS variable is normally distributed.

3.5.1.3. order of (AR)

For water quality like King Talal Dam, the value of AR, which is expressed by the item (p) shall not be more than 1 since the autocorrelation for a particle of TSS does not need more than 1 month till it sediments (Viessman and Lewis, 1996). From Figure (7) it can be seen that the value of AR is equal to 1, so the value of p will be 1 for the TSS variable.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.39	4.86	24.10	13	0.19	1.79	72.55	25	0.10	0.78	119.58	37	0.18	1.33	184.83
2	0.19	2.10	30.02	14	0.15	1.37	76.46	26	0.04	0.35	119.93	38	0.04	0.30	185.20
3	0.05	0.50	30.37	15	-0.01	-0.08	76.48	27	-0.08	-0.64	121.13	39	0.02	0.13	185.27
4	0.10	1.05	31.95	16	-0.14	-1.27	80.01	28	-0.16	-1.31	126.22				
5	-0.09	-0.95	33.27	17	-0.17	-1.48	84.93	29	-0.24	-1.88	137.13				
6	-0.17	-1.78	37.97	18	-0.21	-1.83	92.65	30	-0.22	-1.69	146.34				
7	-0.18	-1.85	43.31	19	-0.22	-1.93	101.67	31	-0.23	-1.79	157.18				
8	-0.07	-0.69	44.09	20	-0.16	-1.32	106.12	32	-0.12	-0.87	159.87				
9	-0.04	-0.37	44.32	21	-0.01	-0.11	106.15	33	-0.03	-0.20	160.02				
10	0.20	2.04	51.32	22	0.15	1.25	110.28	34	0.10	0.73	161.99				
11	0.20	1.95	58.11	23	0.15	1.27	114.70	35	0.23	1.72	172.89				
12	0.22	2.07	66.15	24	0.13	1.05	117.82	36	0.15	1.13	177.81				

Figure (7) Autocorrelation Function for TSS Variable

3.5.1.4. order of moving average (MA)

After finding the value of AR, which was 1, the following procedure is to determine the value of MA, which is expressed by the item (q). Figure (8) shows the change between the real data of the variable TSS and it's moving average with different length of p.

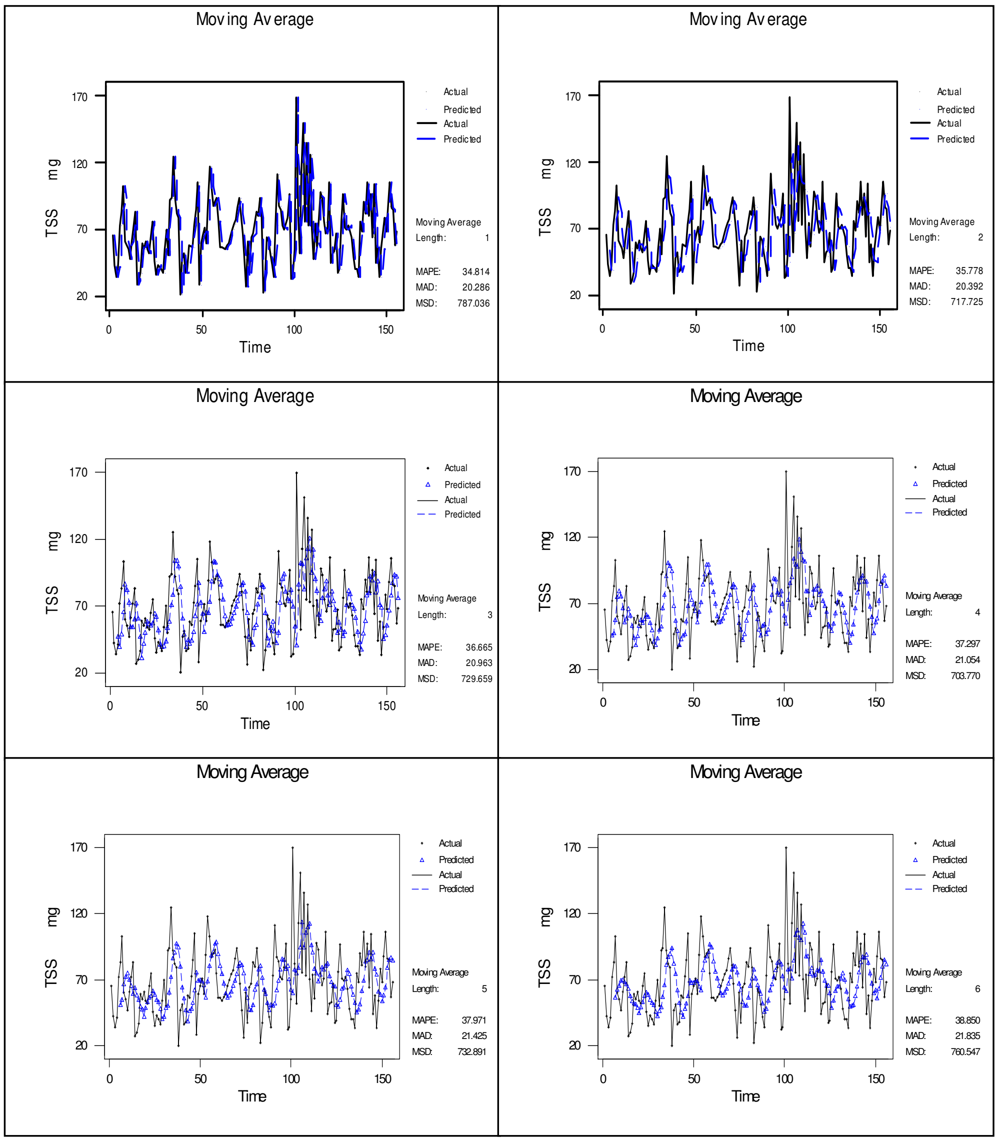


Figure (8) Moving Average of TSS with Different Values of (p)

The moving average can be determined from Figure (8) when the difference between the previous length of p and the followed one have a small difference and that occurred when the value of p was 4 (as shown in Figure (8)), so the TSS variable has a value of MA(4).

3.5.1.5. order of (I)

The last coefficient of ARIMA's parameters is the integrated model (I), which expressed by the item (d). The data should be differenced when there is trend or shift or seasonality in the data, otherwise there is no need to make differentiation for the data. Figure (9) consists of four graphs, which provides a good idea if there is a difference between the original, detrended, seasonally adjusted, and seasonally adjusted and detrended data. It is shown from these four graphs that they are almost the same, which means that the detrended and seasonally effects are almost negligible.

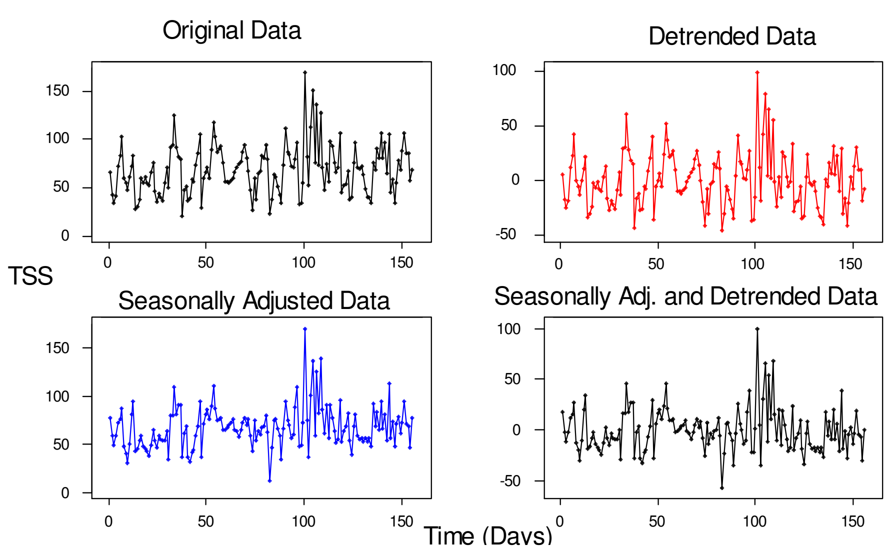


Figure (9): Component Analysis for TSS mg/l

Two season; summer and winter can affect seasonality in Jordan, so if the data has no seasonality effect, then the value of d=0 and if we have seasonality effect then the value of d=2. Figures (10) and (11) provides ARIMA model diagnostics for ARIMA = (1,0,4) and (1,2,4). It is seen from the two graphs that the residual in Figure (10) is less than Figure (11) so the coefficients of ARIMA that will be used are (1,0,4)

Comment: Figure (8)

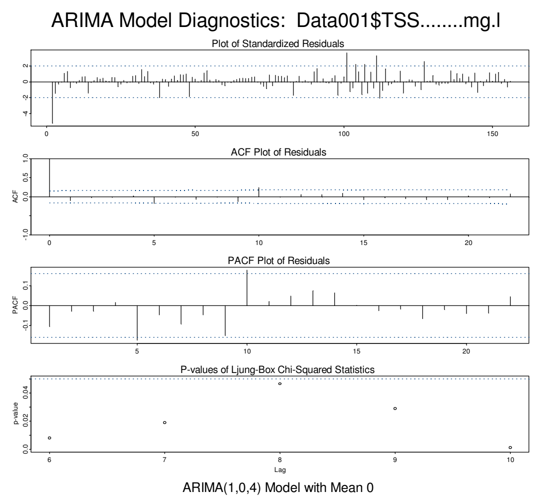


Figure (9): ARIMA (1,0,4) Model Diagnostics for TSS

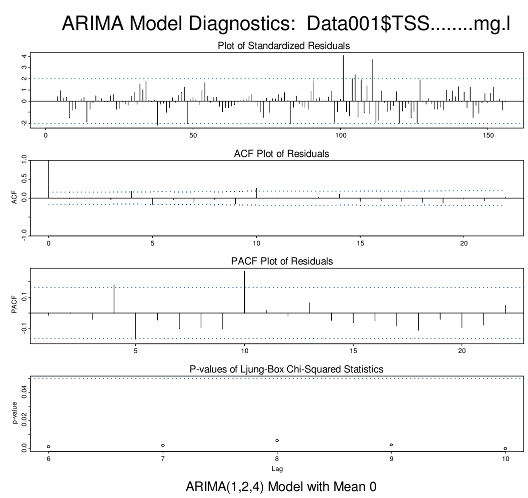


Figure (10): ARIMA (1,2,4) Model Diagnostics for TSS

3.5.1.6. forecasting future values

The following procedure will be used in the forecasting: The values of the data collected will be divided into two parts, the first part consists of 90% of the real data, and this data will be analyzed and predicted. And the second part consists of the last 10% of the real data, and this part will be compared with the predicted values in the mean. The best model is the one that gives the least error in mean.

A- deterministic forecasting

A1- linear regression model

The regression of the additive linear trend is shown in Figure (12).

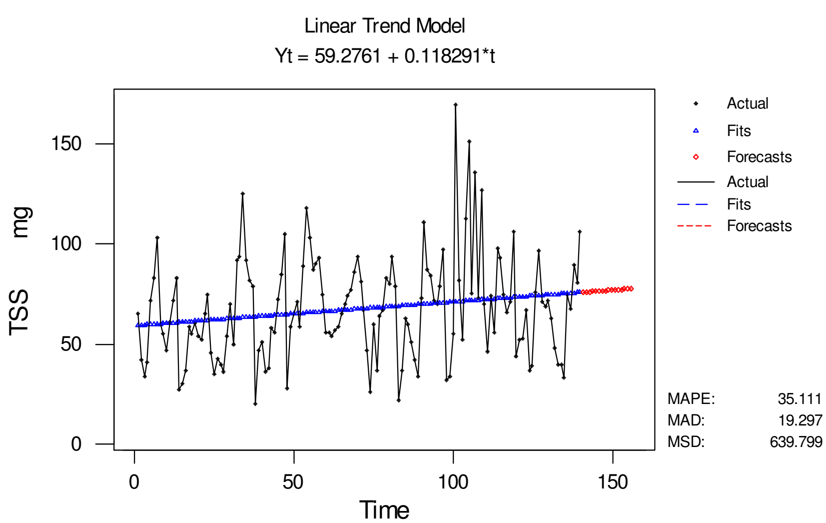


Figure (12): Trend Analysis for TSS mg/l

It can be observed from the above figure and equation of the linear trend that the data is increasing slowly. Table (9) shows the linear prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (9): The values of the predicted and actual data by linear regression for TSS variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	75.96	79.91
2	142	76.07	97.03
3	143	76.19	64.10
4	144	76.31	104.59
5	145	76.43	44.00
6	146	76.55	58.00
7	147	76.66	33.00
8	148	76.78	54.06
9	149	76.90	78.09
10	150	77.02	67.99
11	151	77.14	87.92
12	152	77.26	105.95
13	153	77.37	85.84
14	154	77.49	85.06
15	155	77.61	57.03
16	156	77.73	68.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 5.0%, that the linear trend model has satisfied the forecasting for the TSS variable.

B3- quadratic regression model

The regression of the additive quadratic trend is shown in Figure (13).

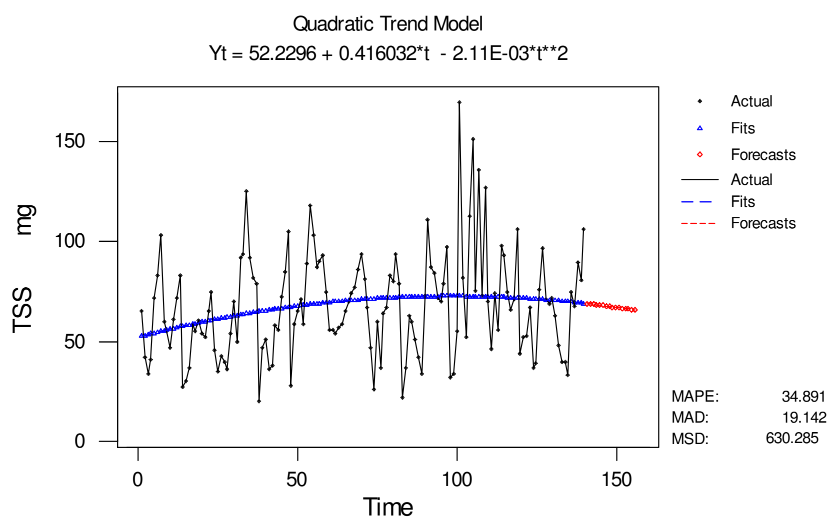


Figure (13): Trend Analysis for TSS mg/l

It can be observed from the above figure and the equation of the quadratic trend that the data is concaved down or in another way that there is an increase in the beginning and then a decrease in the end. Table (10) shows the quadratic prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (10): The values of the predicted and actual data by quadratic regression for TSS variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	68.91	79.91
2	142	68.73	97.03
3	143	68.54	64.10
4	144	68.35	104.59
5	145	68.16	44.00
6	146	67.96	58.00
7	147	67.76	33.00
8	148	67.55	54.06
9	149	67.34	78.09
10	150	67.12	67.99
11	151	66.90	87.92
12	152	66.68	105.95
13	153	66.45	85.84
14	154	66.22	85.06
15	155	65.98	57.03
16	156	65.74	68.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 8.6%, that the quadratic trend model has satisfied the forecasting for the TSS variable.

B3- exponential growth regression model

The regression of the additive exponential growth trend model is shown in Figure (14).

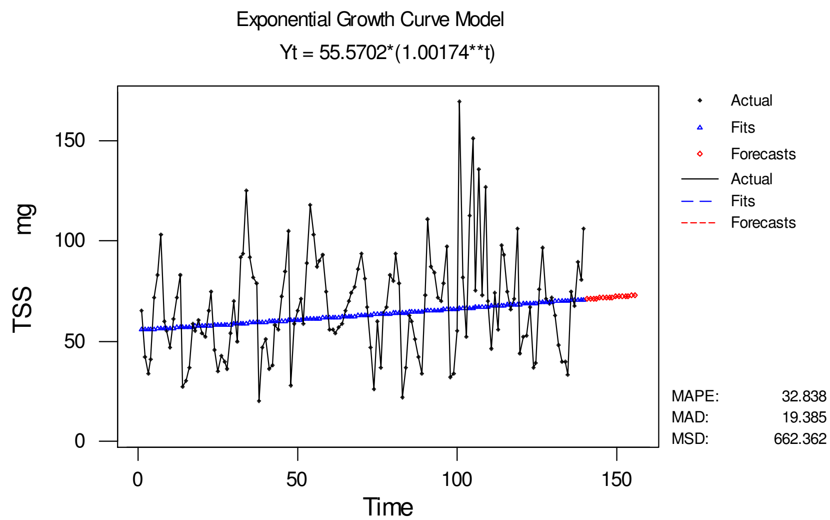


Figure (14): Trend Analysis for TSS mg/l

It can be observed from the above figure and equation of the exponential growth trend that the data has an increasing trend. Table (11) shows the exponential growth prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (11): The values of the predicted and actual data by exponential growth regression for TSS variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	71.01	79.91
2	142	71.14	97.03
3	143	71.26	64.10
4	144	71.38	104.59
5	145	71.51	44.00
6	146	71.63	58.00
7	147	71.76	33.00
8	148	71.88	54.06
9	149	72.01	78.09
10	150	72.13	67.99
11	151	72.26	87.92
12	152	72.38	105.95
13	153	72.51	85.84
14	154	72.64	85.06
15	155	72.76	57.03
16	156	72.89	68.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 1.7%, that the exponential growth trend model has satisfied the forecasting for the TSS variable.

B3- single exponential smoothing model

The regression of the additive single exponential smoothing trend model is shown in Figure (15).

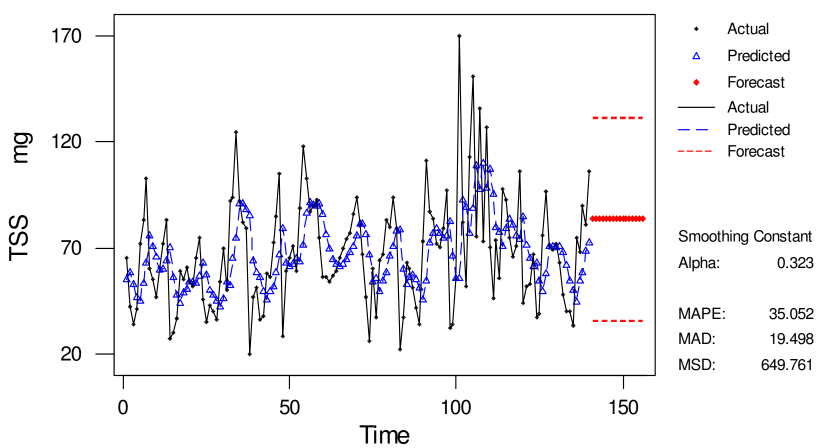


Figure (15): Single Exponential Smoothing for TSS mg/l

Table (12) shows the single exponential smoothing prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (12) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (12): Forecasted, lower, upper and actual values for single exponential smoothing for TSS variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	83.49	35.72	131.26	79.91
2	142	83.49	35.72	131.26	97.03
3	143	83.49	35.72	131.26	64.10
4	144	83.49	35.72	131.26	104.59
5	145	83.49	35.72	131.26	44.00
6	146	83.49	35.72	131.26	58.00
7	147	83.49	35.72	131.26	33.00
8	148	83.49	35.72	131.26	54.06
9	149	83.49	35.72	131.26	78.09
10	150	83.49	35.72	131.26	67.99
11	151	83.49	35.72	131.26	87.92
12	152	83.49	35.72	131.26	105.95
13	153	83.49	35.72	131.26	85.84
14	154	83.49	35.72	131.26	85.06
15	155	83.49	35.72	131.26	57.03
16	156	83.49	35.72	131.26	68.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 12.4%, that the simple exponential smoothing trend model has not satisfied the forecasting for the TSS variable.

B- stochastic forecasting

B3- auto regression model

Table (13) shows the AR(1) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (13) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (13): Forecasted, lower, upper and actual values for AR(1) for TSS variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	83.20	36.59	129.81	79.91
2	142	73.99	23.75	124.22	97.03
3	143	70.28	19.48	121.08	64.10
4	144	68.79	17.90	119.68	104.59
5	145	68.19	17.29	119.10	44.00
6	146	67.95	17.04	118.86	58.00
7	147	67.85	16.95	118.76	33.00
8	148	67.81	16.91	118.72	54.06
9	149	67.80	16.89	118.71	78.09
10	150	67.79	16.88	118.70	67.99
11	151	67.79	16.88	118.70	87.92
12	152	67.79	16.88	118.70	105.95
13	153	67.79	16.88	118.70	85.84
14	154	67.79	16.88	118.70	85.06
15	155	67.79	16.88	118.70	57.03
16	156	67.79	16.88	118.70	68.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 5.4%, that the AR(1) trend model has satisfied the forecasting for the TSS variable.

B3- moving average regression model

The regression of the additive MA(4) trend model is shown in Figure(16).

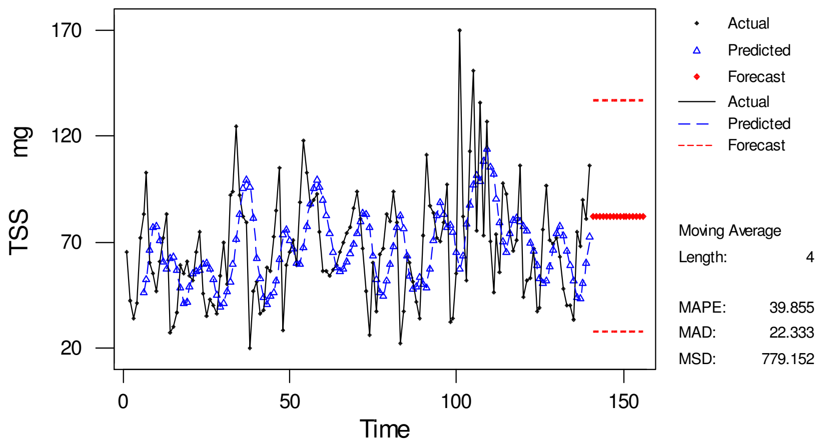


Figure (16): Moving Average Trend for TSS mg/l

Table (14) shows the MA(4) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (14) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (14): Forecasted, lower, upper and actual values for MA(4) for TSS variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	82.32	27.60	137.03	79.91
2	142	82.32	27.60	137.03	97.03
3	143	82.32	27.60	137.03	64.10
4	144	82.32	27.60	137.03	104.59
5	145	82.32	27.60	137.03	44.00
6	146	82.32	27.60	137.03	58.00
7	147	82.32	27.60	137.03	33.00
8	148	82.32	27.60	137.03	54.06
9	149	82.32	27.60	137.03	78.09
10	150	82.32	27.60	137.03	67.99
11	151	82.32	27.60	137.03	87.92
12	152	82.32	27.60	137.03	105.95
13	153	82.32	27.60	137.03	85.84
14	154	82.32	27.60	137.03	85.06
15	155	82.32	27.60	137.03	57.03
16	156	82.32	27.60	137.03	68.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 11.1%, that the MA(5) trend model has not satisfied the forecasting for the TSS variable.

B3- ARMA modeling

Table (15) shows the ARMA(1,4) prediction values for the next 10% of the predicted and the real data, which equals to 16 observations. In Table (15) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (15): Forecasted, lower, upper and actual values for ARIMA(1,0,4) for TSS variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	67.77	66.95	68.58	79.91
2	142	67.82	66.33	69.31	97.03
3	143	67.72	65.76	69.69	64.10
4	144	67.61	65.46	69.75	104.59
5	145	67.60	65.46	69.75	44.00
6	146	67.60	65.46	69.75	58.00
7	147	67.60	65.46	69.75	33.00
8	148	67.60	65.46	69.75	54.06
9	149	67.60	65.46	69.75	78.09
10	150	67.60	65.46	69.75	67.99
11	151	67.60	65.46	69.75	87.92
12	152	67.60	65.46	69.75	105.95
13	153	67.60	65.46	69.75	85.84
14	154	67.60	65.46	69.75	85.06
15	155	67.60	65.46	69.75	57.03
16	156	67.60	65.46	69.75	68.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 8.2%, that the ARIMA(1,0,4) trend model has satisfied the forecasting for the TSS variable.

3.5.1.7 Results of Prediction

The results of error are summarized in the following Table (16), which provides a summary of the models name used in the prediction and also it provides the percentage error.

Table (16) : Percentage of error of each model for TSS variable

Model	Percentage of Mean Error
Linear Method	5.0 %
Quadratic Method	8.6 %
Exponential Growth Method	1.7 %
Simple Exponential Smoothing	12.4 %
Auto Regression, AR(1)	5.4 %
Moving Average, MA(4)	11.1 %
ARIMA (1,0,4)	8.2 %

The previous Table (16) shows that all the methods have satisfied the 10% acceptable prediction limits, except in MA(4) and in simple exponential smoothing. When finding the best model that gave the least error it will be the exponential growth method, this is for deterministic forecasting. But one should take into consideration that we deal with stochastic method, so AR (1) will be the best method for forecasting.

3.5.2 Biochemical oxygen demand (BOD_5) variable:

The consequences that were used to analyze the BOD_5 variable were as follows:

3.5.2.1 detection of missing data and outliers:

From the table (1) it is observed that the data do not contain any missing data, so the second step is to find the outliers, data should be drawn in a scatter diagram (Figure 17) so that outliers will be clearly observed. These data, which contains 156 observations from January 1988 till December 2000, have approximately three outliers and they are in the following months: January 1995, January 1997, and December 1997. It was observed that the rainfall in January 1995 was low, and it is known that when the rainfall is low then the BOD_5 will get high, in January and December 1997 the rainfall was high so the BOD_5 should be low (Appendix (1)). So the real data is on January 1995, the other three data were assumed to be outliers due to human error, and they should be adjusted to a new value since they may greatly influence any statistical calculations and yield biased result. The way that outliers were adjusted was the same as the missing data treated and it was equals to the average monthly value.

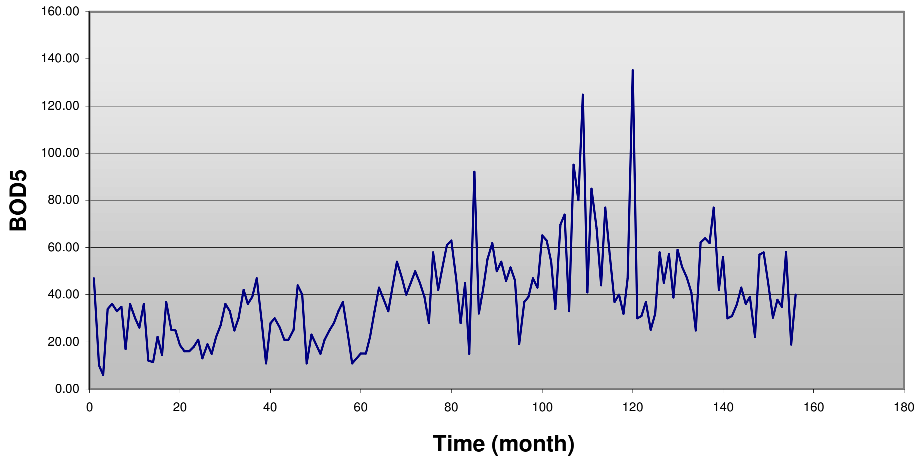


Figure (17): Original Data of BOD_5 mg/l

Figure (18) shows the outliers for the seasonal trends for the original and the residual data, one can conclude from the charts that there are five outliers in both the original data and the residual data in the seasonal condition. Also figure (18) shows the variation in the data for the same month, it can be observe that the variation was the highest on December, and was the lowest on February. Another two outliers were found in the seasonal drawings, they are in November 1996, and March 1997. There was high rainfall in March 1997 so it should be adjusted (Appendix (1)), while in November 1996 it was a real data.

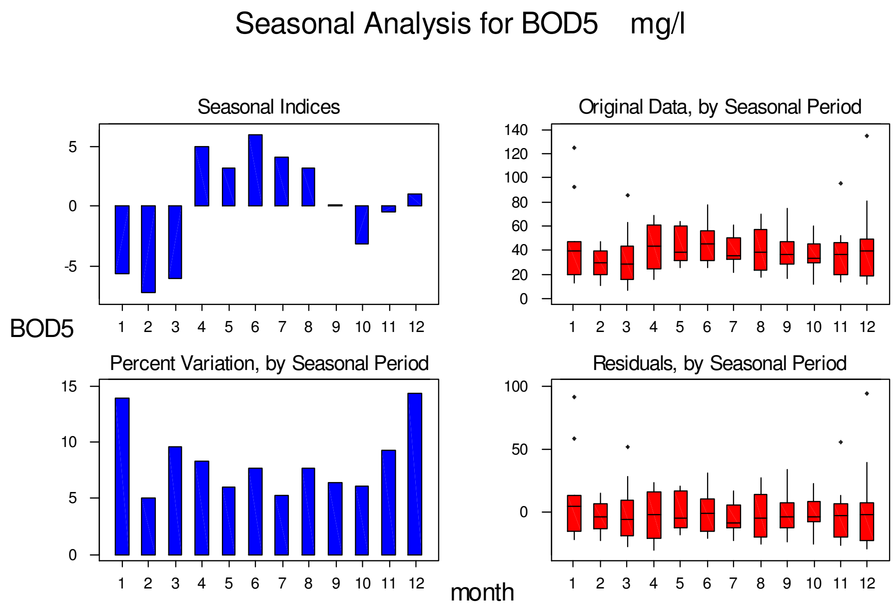


Figure (18) Outliers for Seasonal Analysis for BOD5 Variable

After adjustment the outliers, the new adjusted data are plotted in Figure (17), the figure shows that their still outliers but these outliers cannot be omitted because they are real data so it can influence the statistics results. While comparing the old data (Figure 17) with the new adjusted data (Figure 19) it can be observed that two figures are quite the same and they have the same trend, so the effect of the outliers on the data was so little.

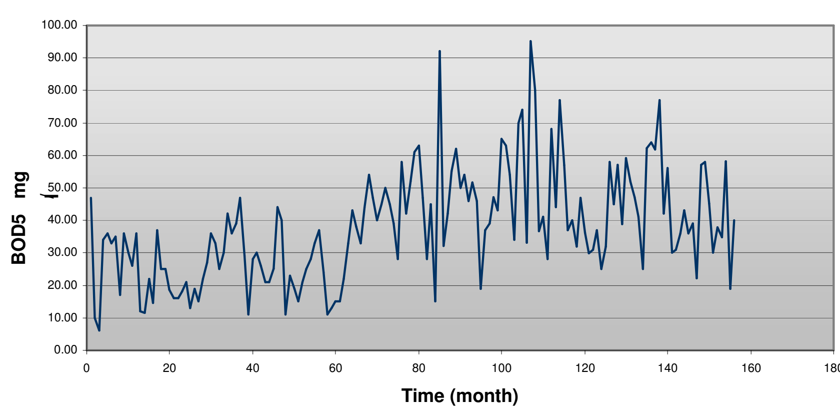


Figure (19): The New Adjusted Data of BOD5 mg/l

3.5.2.2 normality of data

In this section, normality of data will be checked through four procedures; first one is by drawing a histogram for Weibull's distribution model, second one is through calculating the coefficient of variance, the third one is through calculating the Kurtosis coefficient, and the fourth one is through calculating the Shapiro-Wilk test. From these four procedures, if the data was not normal then a lognormal transformation to the data will be made.

A- Weibull's distribution model histogram:

Data will be transformed to the average monthly value for the BOD_5 variable; the calculated values were as follows

Month	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
BOD5	36.7	28.1	28.0	42.7	42.8	45.7	39.5	41.5	38.1	36.2	37.8	36.2

The Weibull's distribution histogram is drawn for these twelve data. It can be observed from figure (20) that the data of BOD_5 is quite normal and there is a little skewness to the left and bulked to the right, but in general the graph gives an indication that the data is normal.

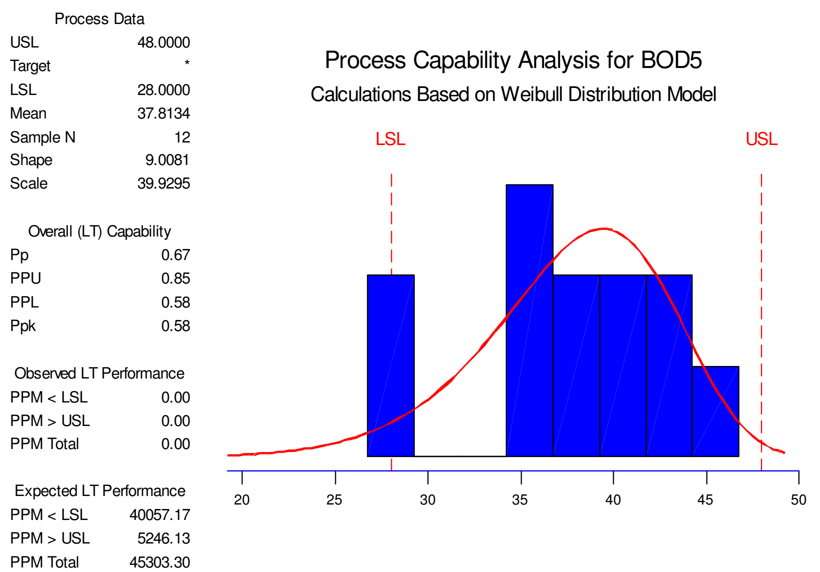


Figure (20) : Weibull Distribution Model Histogram

B- coefficient of variation (COV), preliminary test:

The data were divided into four quarters; each quarter consists of 39 data. Table (17) provides the value of the mean, variance, standard deviation, and the coefficient of variation for the BOD_5 variable.

Table (17): The coefficient of variable for BOD_5

	MEAN	VARIANCE	ST. DE. (S)	C.O.V.
BOD_5 (1st Quarter)	25.8	116.6	10.8	0.4
BOD_5 (2nd Quarter)	31.7	164.9	12.8	0.4
BOD_5 (3rd Quarter)	54.3	479.1	21.9	0.4
BOD_5 (4th Quarter)	45.4	398.8	20.0	0.4

It can be shown from the table that the value of the coefficient of variation for each quarter is less than 1, which means that each quarter of the data has a little skewed (either to right or left), so the total data of the BOD_5 variable has less skewness than each of the four BOD_5 quarters, it can be concluded that the BOD_5 variable does not have skewness.

C- Kurtosis coefficient (peakedness), vertical test:

To find the Kurtosis coefficient, one should find the value of K, which depends on the fourth moment about the mean and the number of samples, so that the Kurtosis can be calculated. The Kurtosis will give a good indication if the distribution is leptokurtic or platykutric. The data was divided into four quarters, Table (18) provides the values of the Kurtosis coefficient for each quarter and it provides also the calculations needed to calculate the Kurtosis coefficient, which they are: the mean, the variance or standard error, the value of K, and the Kurtosis coefficient.

Table (18): The Kurtosis Coefficient for BOD_5

	MEAN	VARIANCE	ST. DE. (S)	K	Kurtosis Coeff. C'_K
BOD5 (1st Quarter)	25.8	116.6	10.8	30758.3	-0.7
BOD5 (2nd Quarter)	31.7	164.9	12.8	60771.7	-0.8
BOD5 (3rd Quarter)	54.3	479.1	21.9	1114502.6	1.9
BOD5 (4th Quarter)	45.4	398.8	20.0	2021622.8	9.7

From table (18) one can observe that the data in the first and second quarters were normally distributed (mesokurtic), in the third quarter it was fairly leptokurtic and in the fourth quarter is was leptokurtic. The total data of the BOD5 variable can be assumed to be as fairly normally distributed (mesokurtic).

D- Shapiro-Wilk test

This is another test to show that the data we have is normal or not. Data that have been collected were divided into equal quarters, the value of (a_{n-1+I}), was taken for 20 data since the value of n-1+i was equal to 20, the value of (a_{n-1+I}) was taken from appendix (2). The Shapiro-Wilk value was compared with the five percent critical value for sample size 20 in Appendix (3), if the value of the Shapiro-Wilk test was greater than it then the data will not show evidence of nonnormality.

From the Tables (19), (20), (21), and (22) it has been shown that the data in the first quarter is nonnormal, the second quarter has a normal data, the third has a quite not normal one, and the fourth one has a normal distribution. It can be safely say that BOD_5 variable is normally distributed.

Table (19): Shapiro-Wilk Test for the Data of BOD_5's 1st quarter

No	BOD_5 mg/l	Ordering BOD_5 (1)	Inverse order BOD_5 (2)	2-1	A(n-1+i)	(2-1) x $a_{(n-1+i)}$
1	47.00	6.00	47.00	41.00	0.3989	16.35
2	10.00	10.00	47.00	37.00	0.2755	10.19
3	6.00	11.00	42.00	31.00	0.2380	7.38
4	34.00	11.50	39.00	27.50	0.2104	5.79
5	36.00	12.00	37.00	25.00	0.1880	4.70
6	33.00	13.00	36.00	23.00	0.1689	3.88
7	35.00	14.50	36.00	21.50	0.1520	3.27
8	17.00	15.00	36.00	21.00	0.1366	2.87
9	36.00	16.00	36.00	20.00	0.1225	2.45
10	30.00	16.00	36.00	20.00	0.1092	2.18
11	26.00	17.00	35.00	18.00	0.0967	1.74
12	36.00	18.00	34.00	16.00	0.0848	1.36
13	12.00	18.50	33.00	14.50	0.0733	1.06
14	11.50	19.00	33.00	14.00	0.0622	0.87
15	22.00	21.00	30.00	9.00	0.0515	0.46
16	14.50	22.00	30.00	7.99	0.0409	0.33
17	37.00	22.00	29.00	7.00	0.0305	0.21
18	25.00	25.00	27.00	2.01	0.0203	0.04
19	25.00	25.00	26.00	1.00	0.0101	0.01
20	18.50	25.00	25.00	0.00		b=65.15
21	16.00	26.00	25.00	-1.00		S=10.80
22	16.00	27.00	25.00	-2.01		
23	18.00	29.00	22.00	-7.00		
24	21.00	30.00	22.00	-7.99		
25	13.00	30.00	21.00	-9.00		
26	19.00	33.00	19.00	-14.00		
27	15.00	33.00	18.50	-14.50		
28	22.00	34.00	18.00	-16.00		
29	27.00	35.00	17.00	-18.00		
30	36.00	36.00	16.00	-20.00		
31	33.00	36.00	16.00	-20.00		
32	25.00	36.00	15.00	-21.00		
33	30.00	36.00	14.50	-21.50		
34	42.00	36.00	13.00	-23.00		
35	36.00	37.00	12.00	-25.00		
36	39.00	39.00	11.50	-27.50		
37	47.00	42.00	11.00	-31.00		
38	29.00	47.00	10.00	-37.00		
39	11.00	47.00	6.00	-41.00		

$W=0.958 > 0.939$

Satisfied

Table (20): Shapiro-Wilk Test for the Data of BOD_5's 2nd quarter

No	BOD_5 mg/l	Ordering BOD_5 (1)	Inverse order BOD_5 (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
40	28.00	11.00	58.00	47.00	0.3989	18.75
41	30.00	11.00	54.00	43.00	0.2755	11.85
42	26.00	13.00	52.00	39.00	0.2380	9.28
43	21.00	15.00	50.00	35.00	0.2104	7.36
44	21.00	15.00	47.00	32.00	0.1880	6.02
45	25.17	15.00	45.00	30.00	0.1689	5.07
46	44.01	19.00	45.00	26.00	0.1520	3.95
47	40.00	21.00	44.01	23.01	0.1366	3.14
48	11.00	21.00	44.00	23.00	0.1225	2.82
49	23.00	21.00	43.00	22.00	0.1092	2.40
50	19.00	22.00	42.00	19.99	0.0967	1.93
51	15.00	23.00	40.00	17.00	0.0848	1.44
52	21.00	24.00	40.00	16.00	0.0733	1.17
53	25.00	25.00	39.00	14.00	0.0622	0.87
54	28.00	25.17	38.00	12.83	0.0515	0.66
55	33.00	26.00	36.99	10.99	0.0409	0.45
56	36.99	28.00	33.00	5.00	0.0305	0.15
57	24.00	28.00	33.00	5.00	0.0203	0.10
58	11.00	28.00	33.00	5.00	0.0101	0.05
59	13.00	30.00	30.00	0.00		b= 77.47
60	15.00	33.00	28.00	-5.00		S= 12.84
61	15.00	33.00	28.00	-5.00		
62	22.00	33.00	28.00	-5.00		
63	33.00	36.99	26.00	-10.99		
64	43.00	38.00	25.17	-12.83		
65	38.00	39.00	25.00	-14.00		
66	33.00	40.00	24.00	-16.00		
67	44.00	40.00	23.00	-17.00		
68	54.00	42.00	22.00	-19.99		
69	47.00	43.00	21.00	-22.00		
70	40.00	44.00	21.00	-23.00		
71	45.00	44.01	21.00	-23.01		
72	50.00	45.00	19.00	-26.00		
73	45.00	45.00	15.00	-30.00		
74	39.00	47.00	15.00	-32.00		
75	28.00	50.00	15.00	-35.00		
76	58.00	52.00	13.00	-39.00		
77	42.00	54.00	11.00	-43.00		
78	52.00	58.00	11.00	-47.00		

W=0.958> 0.939

→ Satisfied

Table (21): Shapiro-Wilk Test for the Data of BOD_5's 3rd quarter

No	BOD_5 mg/l	Ordering BOD_5 (1)	Inverse order BOD_5 (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
79	60.99	15.00	125.00	110.00	0.3989	43.88
80	63.00	18.97	95.03	76.06	0.2755	20.95
81	47.00	28.00	92.00	64.00	0.2380	15.23
82	28.00	32.04	85.00	52.96	0.2104	11.14
83	45.00	33.08	80.04	46.96	0.1880	8.83
84	15.00	33.89	77.08	43.19	0.1689	7.30
85	92.00	36.96	74.02	37.07	0.1520	5.63
86	32.04	36.99	69.87	32.88	0.1366	4.49
87	41.99	38.98	68.06	29.08	0.1225	3.56
88	55.01	40.01	65.08	25.07	0.1092	2.74
89	61.97	41.04	63.06	22.03	0.0967	2.13
90	50.03	41.99	63.00	21.02	0.0848	1.78
91	54.04	43.05	61.97	18.92	0.0733	1.39
92	45.93	44.04	60.99	16.95	0.0622	1.05
93	51.62	45.00	56.96	11.96	0.0515	0.62
94	45.95	45.93	55.01	9.09	0.0409	0.37
95	18.97	45.95	54.04	8.09	0.0305	0.25
96	36.96	47.00	54.03	7.03	0.0203	0.14
97	38.98	47.05	51.62	4.57	0.0101	0.05
98	47.05	50.03	50.03	0.00		b= 131.53
99	43.05	51.62	47.05	-4.57		S= 21.89
100	65.08	54.03	47.00	-7.03		
101	63.06	54.04	45.95	-8.09		
102	54.03	55.01	45.93	-9.09		W=0.950> 0.939
103	33.89	56.96	45.00	-11.96		
104	69.87	60.99	44.04	-16.95		Satisfied
105	74.02	61.97	43.05	-18.92		
106	33.08	63.00	41.99	-21.02		
107	95.03	63.06	41.04	-22.03		
108	80.04	65.08	40.01	-25.07		
109	125.00	68.06	38.98	-29.08		
110	41.04	69.87	36.99	-32.88		
111	85.00	74.02	36.96	-37.07		
112	68.06	77.08	33.89	-43.19		
113	44.04	80.04	33.08	-46.96		
114	77.08	85.00	32.04	-52.96		
115	56.96	92.00	28.00	-64.00		
116	36.99	95.03	18.97	-76.06		
117	40.01	125.00	15.00	-110.00		

Table (22): Shapiro-Wilk Test for the Data of BOD_5's 4th quarter

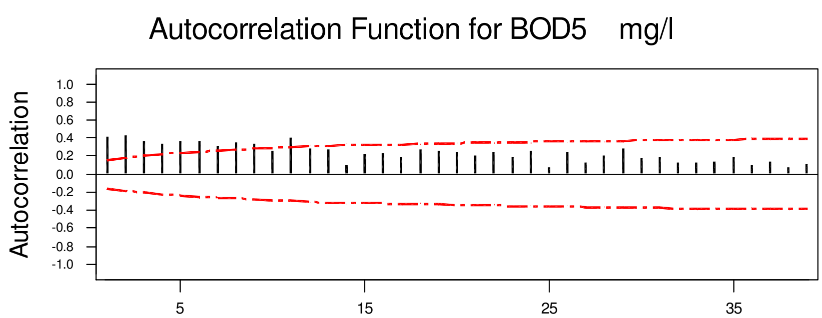
No	BOD_5 mg/l	Ordering BOD_5 (1)	Inverse order BOD_5 (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
118	31.89	18.93	135.00	116.07	0.3989	46.30
119	46.98	22.05	77.00	54.96	0.2755	15.14
120	135.00	24.98	64.03	39.05	0.2380	9.29
121	29.97	25.07	62.09	37.02	0.2104	7.79
122	30.96	29.97	61.90	31.93	0.1880	6.00
123	36.98	30.04	59.10	29.06	0.1689	4.91
124	25.07	30.13	58.06	27.93	0.1520	4.25
125	31.96	30.89	57.97	27.08	0.1366	3.70
126	57.97	30.96	57.97	27.01	0.1225	3.31
127	45.01	31.89	57.12	25.23	0.1092	2.76
128	57.12	31.96	57.08	25.12	0.0967	2.43
129	38.92	34.87	56.07	21.20	0.0848	1.80
130	59.10	35.90	51.94	16.04	0.0733	1.18
131	51.94	35.98	47.11	11.13	0.0622	0.69
132	47.11	36.98	46.98	10.00	0.0515	0.51
133	40.99	37.86	45.08	7.21	0.0409	0.30
134	24.98	38.92	45.01	6.09	0.0305	0.19
135	62.09	39.07	43.06	3.98	0.0203	0.08
136	64.03	40.04	42.04	2.00	0.0101	0.02
137	61.90	40.99	40.99	0.00		b=110.64
138	77.00	42.04	40.04	-2.00		S=19.97
139	42.04	43.06	39.07	-3.98		
140	56.07	45.01	38.92	-6.09		
141	30.04	45.08	37.86	-7.21		
142	30.89	46.98	36.98	-10.00		
143	35.90	47.11	35.98	-11.13		
144	43.06	51.94	35.90	-16.04		
145	35.98	56.07	34.87	-21.20		
146	39.07	57.08	31.96	-25.12		
147	22.05	57.12	31.89	-25.23		
148	57.08	57.97	30.96	-27.01		
149	57.97	57.97	30.89	-27.08		
150	45.08	58.06	30.13	-27.93		
151	30.13	59.10	30.04	-29.06		
152	37.86	61.90	29.97	-31.93		
153	34.87	62.09	25.07	-37.02		
154	58.06	64.03	24.98	-39.05		
155	18.93	77.00	22.05	-54.96		
156	40.04	135.00	18.93	-116.07		

$W=0.808 < 0.939$

→ Did not Satisfied

3.5.2.3 order of (AR)

For water quality like King Talal Dam, the value of AR, which is expressed by the item (p) shall not be more than 1 since the autocorrelation for a particle of BOD_5 does not need more than 1 month till it analyze (Viessman and Lewis, 1996). From Figure (21) it can be seen that the value of AR is more than 1, but the value of p that will be used is 1 for the BOD_5 variable.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.41	5.14	26.97	13	0.26	1.66	256.79	25	0.07	0.39	356.09	37	0.14	0.69	430.48
2	0.42	4.56	55.55	14	0.10	0.64	258.64	26	0.24	1.34	367.44	38	0.08	0.40	431.73
3	0.36	3.48	76.74	15	0.21	1.32	266.69	27	0.13	0.68	370.50	39	0.11	0.54	434.08
4	0.34	3.00	95.06	16	0.24	1.44	276.59	28	0.21	1.11	378.68				
5	0.36	3.03	116.03	17	0.20	1.18	283.45	29	0.28	1.52	394.42				
6	0.36	2.86	137.09	18	0.26	1.57	295.95	30	0.17	0.92	400.35				
7	0.31	2.36	153.08	19	0.25	1.48	307.52	31	0.19	1.02	407.81				
8	0.35	2.59	173.85	20	0.24	1.39	318.06	32	0.12	0.63	410.75				
9	0.33	2.36	192.63	21	0.21	1.19	325.97	33	0.12	0.63	413.67				
10	0.26	1.75	203.79	22	0.24	1.36	336.64	34	0.13	0.69	417.21				
11	0.40	2.67	230.98	23	0.19	1.07	343.44	35	0.19	0.98	424.60				
12	0.28	1.81	244.74	24	0.25	1.39	355.15	36	0.10	0.52	426.68				

Figure (21) Autocorrelation Function for BOD_5 Variable

3.5.2.4 order of moving average (MA)

After finding the value of AR, which was 1, the following procedure is to determine the value of MA, which is expressed by the item (q). Figure (22) shows the change between the real data of the variable BOD_5 and it's moving average with different lengths of p.

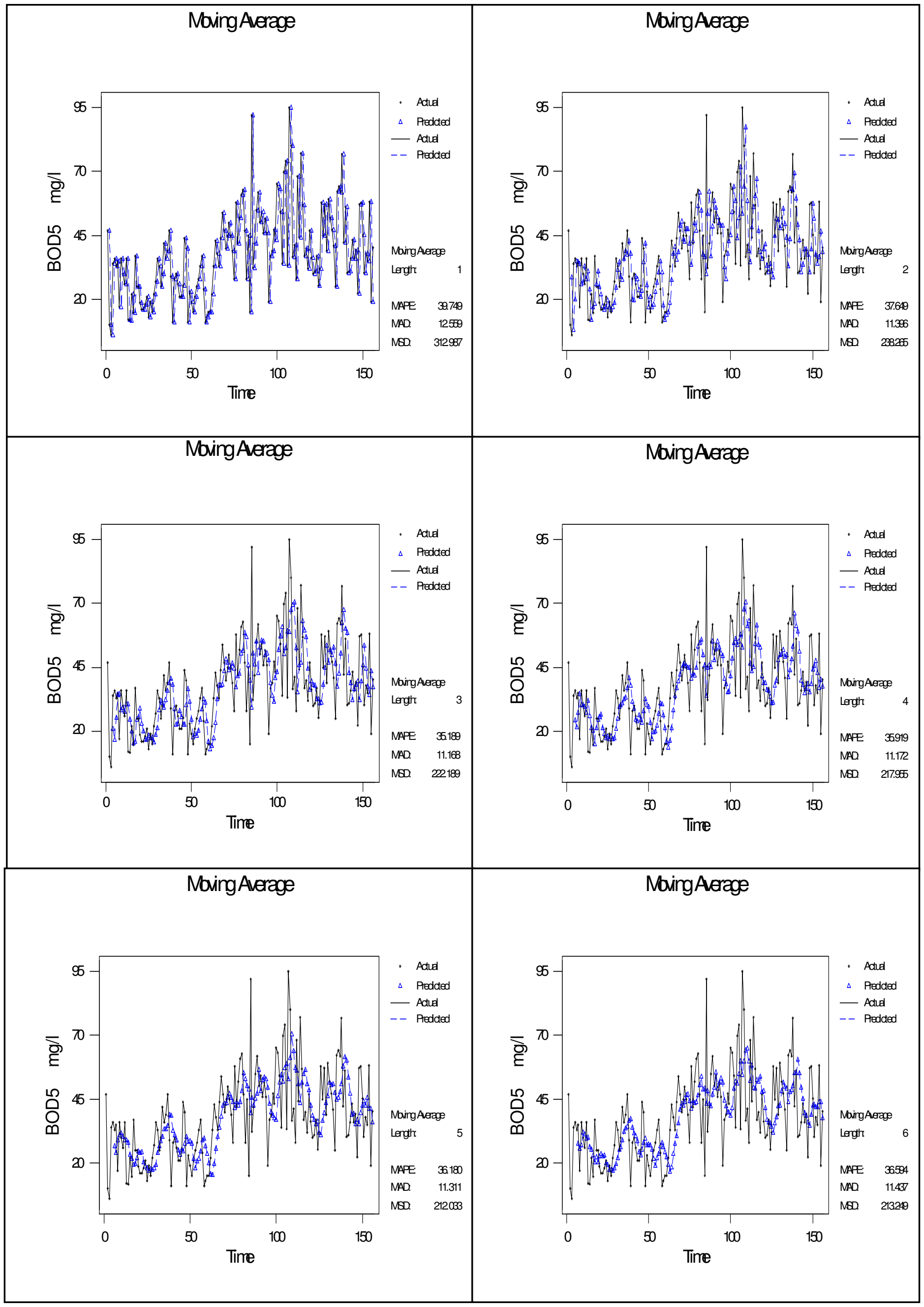


Figure (22) Moving Average of BOD_5 with Different Values of (p)

The moving average can be determined from Figure (22) when the difference between the previous length of p and the followed one have a small difference and that occurred when the value of p was 3 (as shown in Figure (22)), so the BOD5 variable has a value of MA(3).

3.5.2.5 order of (I)

The last coefficient of ARIMA's parameters is the integrated model (I), which expressed by the item (d). The data should be differenced when there is trend or shift or seasonality in the data, otherwise there is no need to make differentiation for the data. Figure (23) consists of four graphs, which provides a good idea if there is a difference between the original, detrended, seasonally adjusted, and seasonally adjusted and detrended data. It is shown from these four graphs that they are almost the same, which means that the detrended and seasonally effects are almost negligible.

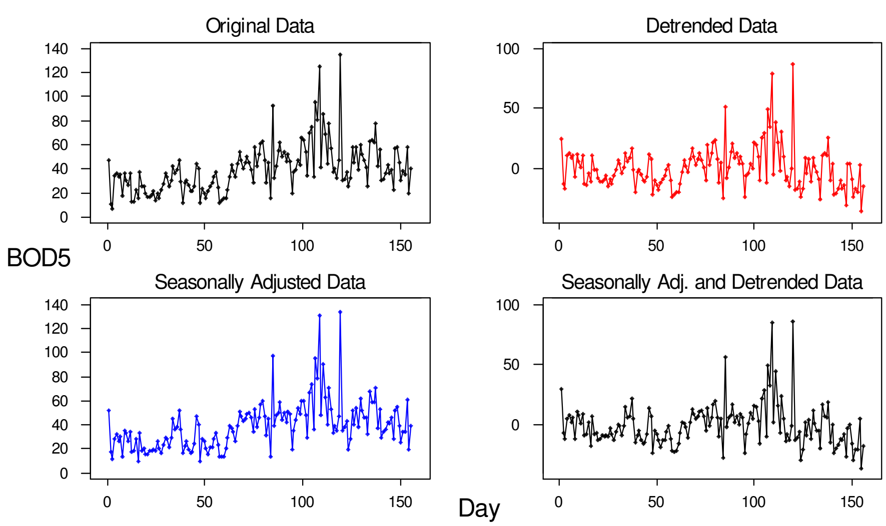


Figure (23): Component Analysis for BOD5 mg/l

Two season; summer and winter can affect seasonality in Jordan, so if the data has no seasonality effect, then the value of d=0 and if we have seasonality effect then the value of d=2. Figure (24) provides ARIMA model diagnostics for ARIMA = (1,0,3), but for ARIMA= (1,2,3) the program could not draw it since it gives singular matrix. It can be concluded that ARIMA (1,0,3) gives the best regression.

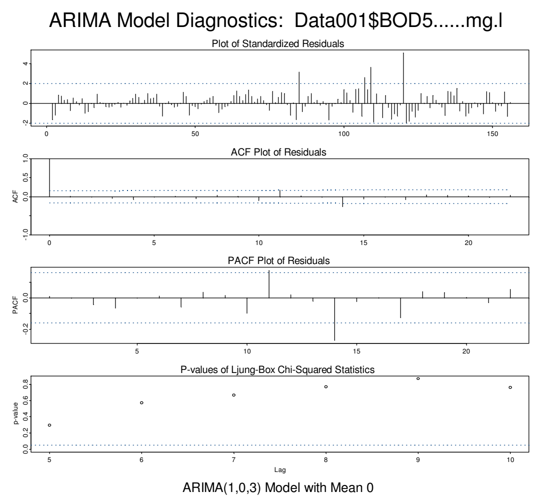


Figure (24): ARIMA (1,0,3) Model Diagnostics for BOD_5

Comment: Figure (8)

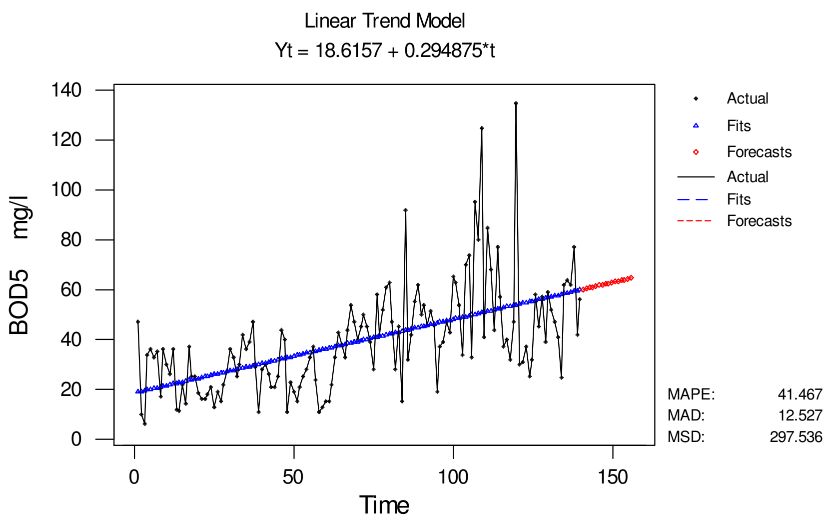
3.5.2.6 forecasting future values

The following procedure will be used in the forecasting: The values of the data collected will be divided into two parts, the first part consists of 90% of the real data, and this data will be analyzed and predicted. And the second part consists of the last 10% of the real data, and this part will be compared with the predicted values in the mean. The best model is the one that gives the least error in mean.

A- deterministic forecasting

A1- linear regression model

The regression of the additive linear trend is shown in Figure (25).



Figure(25): Trend Analysis for BOD5 mg/l

It can be observed from the above figure and equation of the linear trend that the data is increasing slowly. Table (23) shows the linear prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (23): The values of the predicted and actual data by linear regression for BOD_5 variable

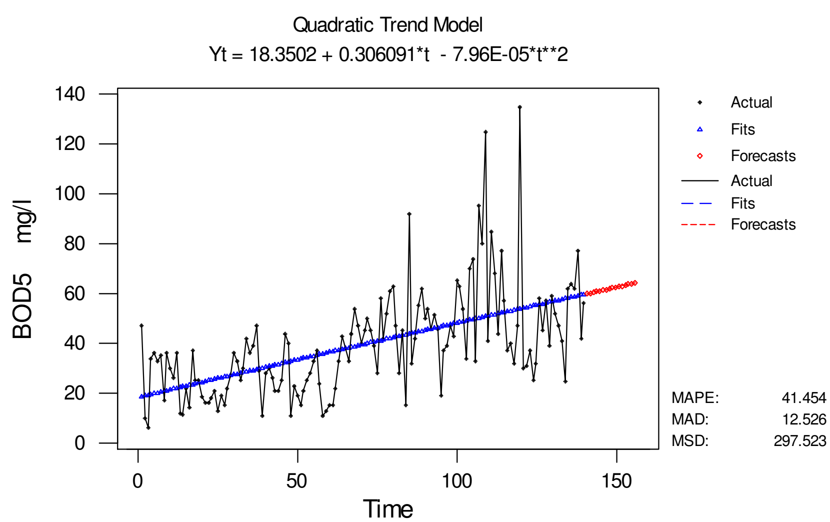
Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	60.19	30.04
2	142	60.49	30.89
3	143	60.78	35.90
4	144	61.08	43.06
5	145	61.37	35.98
6	146	61.67	39.07
7	147	61.96	22.05
8	148	62.26	57.08
9	149	62.55	57.97
10	150	62.85	45.08
11	151	63.14	30.13
12	152	63.44	37.86

13	153	63.73	34.87
14	154	64.03	58.06
15	155	64.32	18.93
16	156	64.62	40.04

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 38.2%, that the linear trend model did not satisfy the forecasting for the BOD5 variable.

A2- quadratic regression model

The regression of the additive quadratic trend is shown in Figure (26).



Figure(26): Trend Analysis for BOD5 mg/l

It can be observed from the above figure and the equation of the quadratic trend that the data is increasing upward slowly. Table (24) shows the quadratic prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (24): The values of the predicted and actual data by quadratic regression for BOD_5 variable

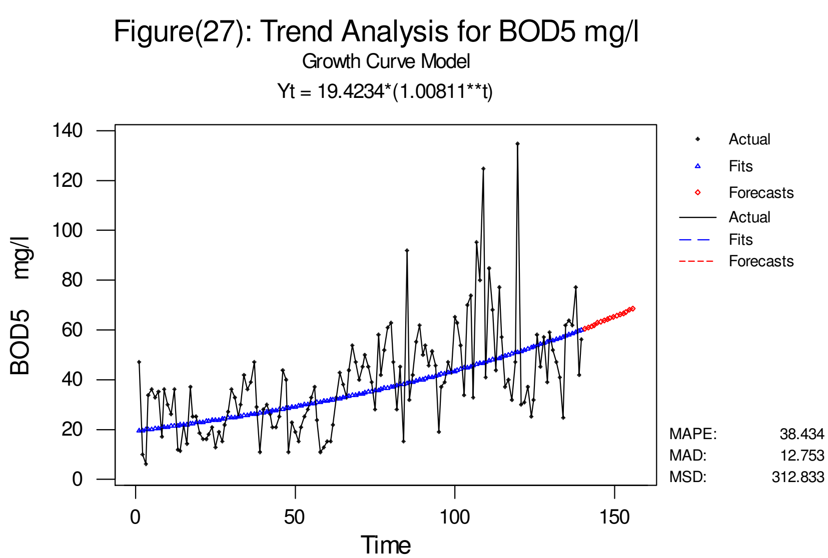
Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	59.93	30.04
2	142	60.21	30.89
3	143	60.49	35.90
4	144	60.78	43.06

5	145	61.06	35.98
6	146	61.34	39.07
7	147	61.63	22.05
8	148	61.91	57.08
9	149	62.19	57.97
10	150	62.47	45.08
11	151	62.76	30.13
12	152	63.04	37.86
13	153	63.32	34.87
14	154	63.60	58.06
15	155	63.88	18.93
16	156	64.16	40.04

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 37.9%, that the quadratic trend model did not satisfy the forecasting for the BOD5 variable.

A3- exponential growth regression model

The regression of the additive exponential growth trend model is shown in Figure (27).



It can be observed from the above figure and equation of the exponential growth trend that the data has an increasing trend. Table (25) shows the exponential growth prediction of the next 10% of the predicted and real data, which equals to 16 observations.

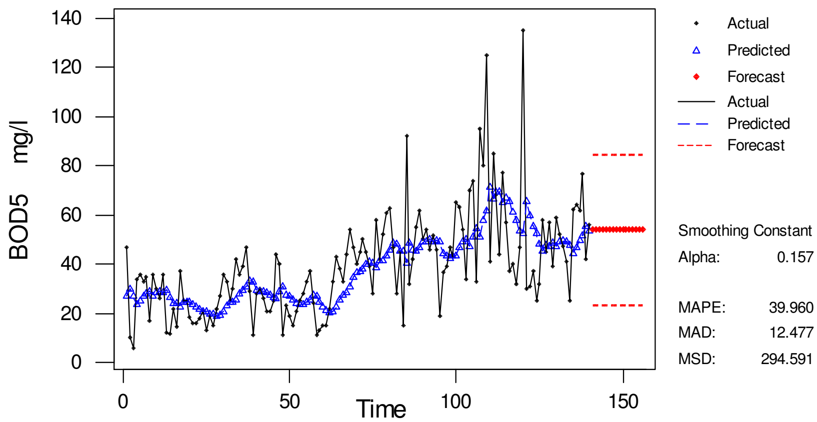
Table (25): The values of the predicted and actual data by exponential growth regression for BOD_5 variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	71.01	30.04
2	142	71.14	30.89
3	143	71.26	35.90
4	144	71.38	43.06
5	145	71.51	35.98
6	146	71.63	39.07
7	147	71.76	22.05
8	148	71.88	57.08
9	149	72.01	57.97
10	150	72.13	45.08
11	151	72.26	30.13
12	152	72.38	37.86
13	153	72.51	34.87
14	154	72.64	58.06
15	155	72.76	18.93
16	156	72.89	40.04

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 46.4%, that the exponential growth trend model did not satisfy the forecasting for the BOD5 variable.

A4- single exponential smoothing model

The regression of the additive single exponential smoothing trend model is shown in Figure (28).



Figure(28): Single Exponential Smoothing for BOD5 mg/l

Table (26) shows the single exponential smoothing prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (26) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (26): Forecasted, lower, upper and actual values by single exponential smoothing for BOD_5 variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	53.91	23.34	84.48	30.04
2	142	53.91	23.34	84.48	30.89
3	143	53.91	23.34	84.48	35.90
4	144	53.91	23.34	84.48	43.06
5	145	53.91	23.34	84.48	35.98
6	146	53.91	23.34	84.48	39.07
7	147	53.91	23.34	84.48	22.05
8	148	53.91	23.34	84.48	57.08
9	149	53.91	23.34	84.48	57.97
10	150	53.91	23.34	84.48	45.08
11	151	53.91	23.34	84.48	30.13
12	152	53.91	23.34	84.48	37.86
13	153	53.91	23.34	84.48	34.87
14	154	53.91	23.34	84.48	58.06

15	155	53.91	23.34	84.48	18.93
16	156	53.91	23.34	84.48	40.04

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 28.5%, that the simple exponential smoothing trend model has not satisfied the forecasting for the BOD5 variable.

B- stochastic forecasting

B1- auto regression model

Table (27) shows the AR(1) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (27) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (27): Forecasted, lower, upper and actual values by AR(1) for BOD_5 variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	82.95	36.96	128.93	30.04
2	142	73.92	24.57	123.28	30.89
3	143	70.41	20.57	120.25	35.90
4	144	69.04	19.13	118.96	43.06
5	145	68.51	18.59	118.44	35.98
6	146	68.30	18.38	118.23	39.07
7	147	68.22	18.30	118.15	22.05
8	148	68.19	18.26	118.12	57.08
9	149	68.18	18.25	118.11	57.97
10	150	68.18	18.25	118.10	45.08
11	151	68.17	18.25	118.10	30.13
12	152	68.17	18.25	118.10	37.86
13	153	68.17	18.24	118.10	34.87
14	154	68.17	18.24	118.10	58.06
15	155	68.17	18.24	118.10	18.93
16	156	68.17	18.24	118.10	40.04

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 44.7%, that the AR(1) trend model has not satisfied the forecasting for the BOD5 variable.

B2- moving average regression model

The regression of the additive MA (3) trend model is shown in Figure (29).

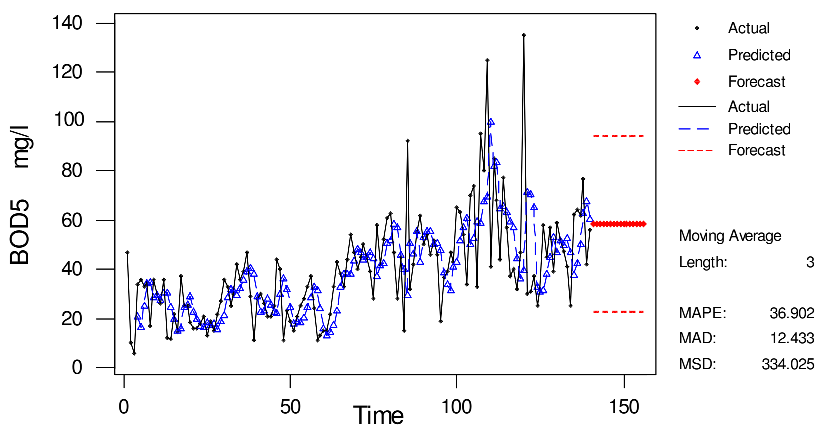


Figure (29): Moving Average Trend for BOD5 mg/l

Table (28) shows the MA(3) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (28) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (28): Forecasted, lower, upper and actual values by MA(3) for BOD_5 variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	58.37	22.55	94.19	30.04
2	142	58.37	22.55	94.19	30.89
3	143	58.37	22.55	94.19	35.90
4	144	58.37	22.55	94.19	43.06
5	145	58.37	22.55	94.19	35.98
6	146	58.37	22.55	94.19	39.07
7	147	58.37	22.55	94.19	22.05

8	148	58.37	22.55	94.19	57.08
9	149	58.37	22.55	94.19	57.97
10	150	58.37	22.55	94.19	45.08
11	151	58.37	22.55	94.19	30.13
12	152	58.37	22.55	94.19	37.86
13	153	58.37	22.55	94.19	34.87
14	154	58.37	22.55	94.19	58.06
15	155	58.37	22.55	94.19	18.93
16	156	58.37	22.55	94.19	40.04

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 33.9%, that the MA(3) trend model has not satisfied the forecasting for the BOD5 variable.

B3- ARMA modeling

Table (29) shows the ARMA(1,3) prediction values for the next 10% of the predicted and the real data, which equals to 16 observations. In Table (29) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (29): Forecasted, lower, upper and actual values by ARMA(1,3) for BOD_5 variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	50.48	17.07	83.89	30.04
2	142	51.34	17.49	85.18	30.89
3	143	50.82	16.23	85.41	35.90
4	144	50.56	15.67	85.45	43.06
5	145	50.30	15.13	85.47	35.98
6	146	50.05	14.61	85.48	39.07
7	147	49.80	14.10	85.49	22.05
8	148	49.55	13.61	85.49	57.08
9	149	49.31	13.14	85.49	57.97
10	150	49.08	12.68	85.48	45.08
11	151	48.85	12.24	85.46	30.13
12	152	48.62	11.81	85.44	37.86
13	153	48.40	11.39	85.41	34.87
14	154	48.19	10.99	85.39	58.06
15	155	47.97	10.59	85.35	18.93
16	156	47.76	10.21	85.31	40.04

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 16.1%, that the ARMA(1,3) trend model has satisfied the forecasting for the BOD5 variable.

3.5.2.7 results of prediction

The results of error are summarized in the following Table (30), which provides a summary of the models name used in the prediction and also it provides the percentage error.

Table (30) : Percentage of error of each model for BOD_5 variable

Model	Percentage of Mean Error
Linear Method	38.2 %
Quadratic Method	37.9 %
Exponential Growth Method	46.4 %
Simple Exponential Smoothing	28.5 %
Auto Regression, AR(1)	44.7 %
Moving Average, MA(3)	33.9 %
ARMA (1, 3)	16.1 %

The previous Table (30) shows that the methods have not satisfied the 10% acceptable prediction limits. When finding the best model that gave the least error it will be ARMA (1,3).

3.5.3 Chemical oxygen demand (COD) variable:

The consequences that were used to analyze the COD variable were as follows:

3.5.3.1 detection of missing data and outliers:

From the table(1) it is observed that the data do not contain any missing data, so the second step is to find the outliers, data should be drawn in a scatter diagram (Figure 30) so that outliers will be clearly observed. These data, which contains 156 observations from January 1988 till December 2000, have approximately five outliers and they are in the following months: March 1988, November 1996, May 1999, June 1999, and August 2000. It was observed that the rainfall in December was high, and it is known that when the rainfall is high then the COD will get low, in November 1996 the rainfall was high, May and June 1999 the rainfall was low so the COD should be high and finally in August 2000 the rainfall was low so the COD should be high (Appendix (1)). So the real data are in March 1988, May 1999, June 1999, and August 2000, the other one data (November 1996) was assumed to be outliers due to human error, and it should be adjusted to a new value since it may greatly influence any statistical calculations and yield biased results. The way that outliers were adjusted was the same as the missing data treated and it was equals to the average monthly value.

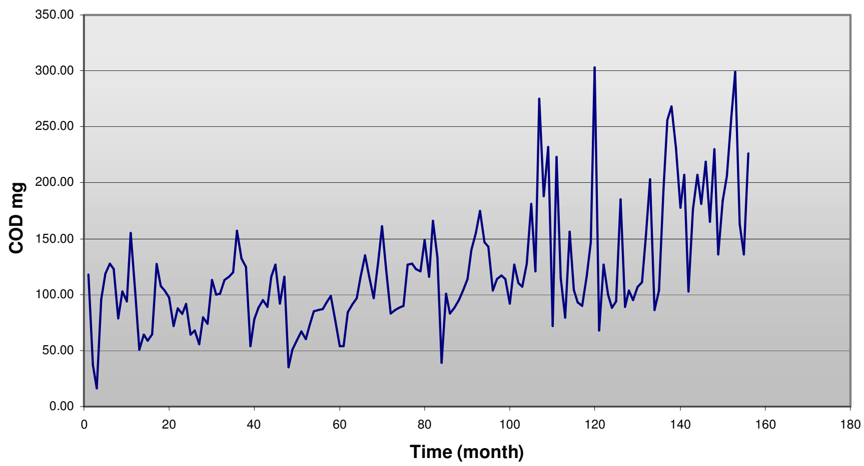


Figure (30) Original Data of COD (mg/l)

Figure (31) shows the outliers for the seasonal trends for the original and the residual data, one can conclude from the charts that there are ten outliers in both the original data and the residual data in the seasonal condition. Also Figure (31) shows the variation in the data for the same month, it can be observe that the variation was the highest on December, and was the lowest on October. Another six outliers were found in the seasonal drawings, they are in March 1997, April 1999, July 1999, February 2000, April 2000, and July 2000. There was high rainfall in March 1997 and February 2000 (Appendix (1)) so they should be adjusted, while in the other months they were real data.

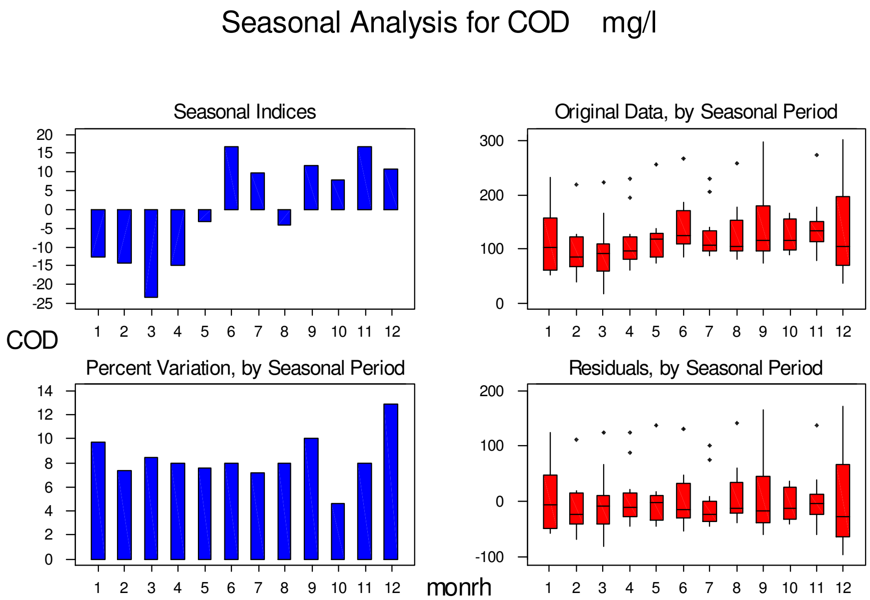


Figure (31) Outliers Seasonal Analysis for COD Variable

After adjustment the outliers, the new adjusted data are plotted in Figure (32), the figure shows that their still outliers but these outliers cannot be omitted because they are real data so it can influence the statistics results. While comparing the old data (Figure 30) with the new adjusted data (Figure 32) it can be observed that two figures are quite the same and they have the same trend, so the effect of the outliers on the data was so little.

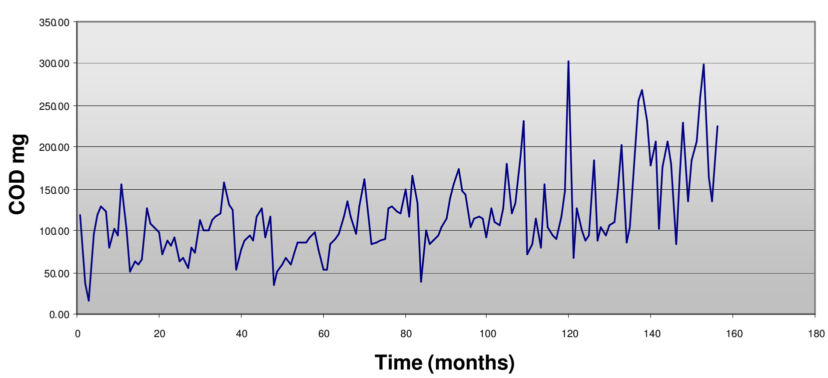


Figure (32): The New Adjusted Data for COD(mg/l)

3.5.3.2 normality of data

In this section, normality of data will be checked through four procedures; first one is by drawing a histogram for Weibull's distribution model, second one is through calculating the coefficient of variance, the third one is through calculating the Kurtosis coefficient, and the fourth one is through calculating the Shapiro-Wilk test. From these four procedures, if the data was not normal then a lognormal transformation to the data will be made.

A- Weibull's distribution model histogram:

Data will be transformed to the average monthly value for the COD variable; the calculated values were as follows

Month	Jan.	Feb.	Mar.	Apr.	May	Jun.
COD mg/l	111.9	84.2	83.6	109.0	117.0	138.7

Month	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
COD mg/l	124.3	126.4	138.5	121.0	127.1	134.0

The Weibull's distribution histogram is drawn for these twelve data. It can be observed from figure (33) that the data of COD is quite normal and there is a little skewness to the left and bulked to the right, but in general the graph gives an indication that the data is normal.

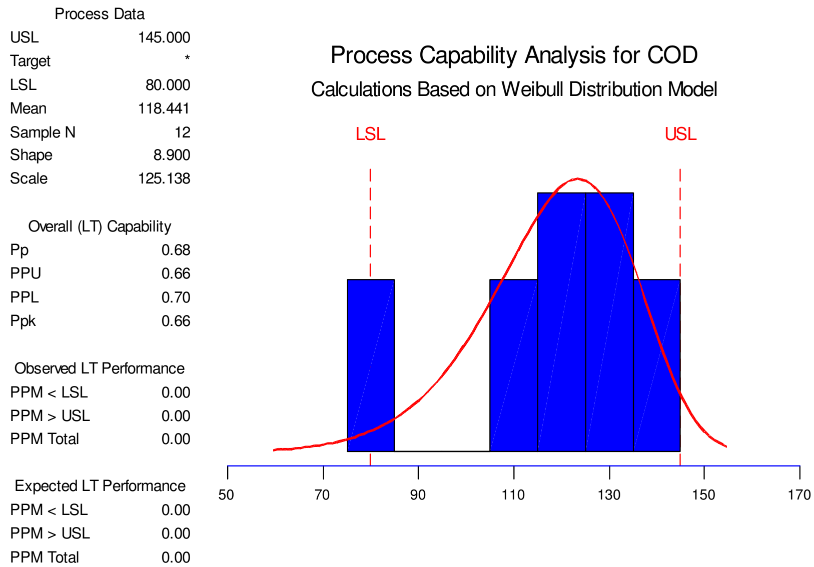


Figure (33) : Weibull Distribution Model Histogram

B- coefficient of variation (COV), preliminary est:

The data were divided into four quarters; each quarter consists of 39 data. Table (31) provides the value of the mean, variance, standard deviation, and the coefficient of variation for the COD variable.

Table (31): The coefficient of variable for COD

	MEAN	VARIANCE	ST. DE. (S)	C.O.V.
COD (1st Quarter)	93.6	977.8	31.3	0.3
COD (2nd Quarter)	93.7	725.1	26.9	0.3
COD (3rd Quarter)	128.5	2122.9	46.1	0.4
COD (4th Quarter)	166.7	4045.4	63.6	0.4

It can be shown from the table that the value of the coefficient of variation for each quarter is less than 1, which means that each quarter of the data has a little skewed (either to right or left), so the total data of the COD variable has less skewness than each of the four COD quarters, it can be concluded that the COD variable does not have skewness.

C- Kurtosis coefficient (peakedness), vertical test:

To find the Kurtosis coefficient, one should find the value of K, which depends on the fourth moment about the mean and the number of samples, so that the Kurtosis can be calculated. The Kurtosis will give a good indication if the distribution is leptokurtic or platykutric. The data was divided into four quarters, Table (32) provides the values of the Kurtosis coefficient for each quarter and it provides also the calculations needed to calculate the Kurtosis coefficient, which they are: the mean, the variance or standard error, the value of K, and the Kurtosis coefficient.

Table (32): The Kurtosis Coefficient for COD

	MEAN	VARIANCE	ST. DE. (S)	K	Kurtosis Coeff. C'_K
COD (1st Quarter)	93.6	977.8	31.3	2939949.6	0.1
COD (2nd Quarter)	93.7	725.1	26.9	45856.6	-2.9
COD (3rd Quarter)	128.5	2122.9	46.1	23428503.4	2.2
COD (4th Quarter)	166.7	4045.4	63.6	39728843.7	-0.6

From table (32) one can observe that the data in the first and fourth quarters were normally distributed (mesokurtic), in the second quarter it was fairly platykutric and in the third quarter is was fairly leptokurtic. The total data of the COD variable can be assumed to be as fairly normally distributed (mesokurtic).

D- Shapiro-Wilk Test

This is another test to show that the data we have is normal or not. Data that have been collected were divided into equal quarters, the value of (a_{n-1+I}), was taken for 20 data since the value of n-1+i was equal to 20, the value of (a_{n-1+I}) was taken from appendix (2). The Shapiro-Wilk value was compared with the five percent critical value for sample size 20 in Appendix (3), if the value of the Shapiro-Wilk test was greater than it then the data will not show evidence of nonnormality.

From the Tables (33), (34), (35), and (36) it has been shown that the data in the first, the second, and the fourth quarters are normal, the third has a quite not normal one. It can be safely say that COD variable is normally distributed.

Table (33): Shapiro-Wilk Test for the Data of COD's 1st quarter

No	COD mg/l	Ordering COD (1)	Inverse order COD (2)	2-1	a(n-1+I)	I*J	
1	118.00	16.00	157.00	141.00	0.3989	56.24	
2	37.00	37.00	155.02	118.02	0.2755	32.52	
3	16.00	51.00	132.00	81.00	0.2380	19.28	
4	96.00	54.00	128.00	74.00	0.2104	15.57	
5	119.00	56.00	127.51	71.51	0.1880	13.44	
6	128.00	58.50	125.00	66.50	0.1689	11.23	
7	123.00	63.99	123.00	59.01	0.1520	8.97	
8	79.00	64.00	119.99	56.00	0.1366	7.65	
9	103.00	64.50	119.00	54.50	0.1225	6.68	
10	94.00	68.00	118.00	50.00	0.1092	5.46	
11	155.02	72.00	116.00	44.00	0.0967	4.25	
12	103.00	74.01	112.99	38.98	0.0848	3.31	
13	51.00	79.00	112.98	33.98	0.0733	2.49	
14	63.99	80.00	107.50	27.50	0.0622	1.71	
15	58.50	82.49	103.98	21.49	0.0515	1.11	
16	64.50	87.49	103.00	15.51	0.0409	0.63	
17	127.51	91.50	103.00	11.50	0.0305	0.35	
18	107.50	94.00	100.99	6.99	0.0203	0.14	
19	103.98	96.00	99.99	3.99	0.0101	0.04	
20	97.51	97.51	97.51	0.00		b=	191.07
21	72.00	99.99	96.00	-3.99		S=	31.27
22	87.49	100.99	94.00	-6.99			
23	82.49	103.00	91.50	-11.50			
24	91.50	103.00	87.49	-15.51			
25	64.00	103.98	82.49	-21.49			
26	68.00	107.50	80.00	-27.50			
27	56.00	112.98	79.00	-33.98			
28	80.00	112.99	74.01	-38.98			
29	74.01	116.00	72.00	-44.00			
30	112.99	118.00	68.00	-50.00			
31	99.99	119.00	64.50	-54.50			
32	100.99	119.99	64.00	-56.00			
33	112.98	123.00	63.99	-59.01			
34	116.00	125.00	58.50	-66.50			
35	119.99	127.51	56.00	-71.51			
36	157.00	128.00	54.00	-74.00			
37	132.00	132.00	51.00	-81.00			
38	125.00	155.02	37.00	-118.02			
39	54.00	157.00	16.00	-141.00			

W= 0.983 > 0.939

Satisfied

Table (34): Shapiro-Wilk Test for the Data of COD's 2nd quarter

No	COD mg/l	Ordering COD (1)	Inverse Order COD (2)	2-1	a(n-1+I)	I*J	
40	78.00	35.00	160.99	125.99	0.3989	50.26	
41	88.01	51.00	135.01	84.00	0.2755	23.14	
42	95.00	54.00	129.01	75.00	0.2380	17.85	
43	88.98	54.00	127.99	73.99	0.2104	15.57	
44	116.01	59.00	127.00	68.00	0.1880	12.78	
45	126.87	60.00	126.87	66.87	0.1689	11.29	
46	92.01	67.00	123.00	56.00	0.1520	8.51	
47	115.99	73.00	122.00	49.00	0.1366	6.69	
48	35.00	77.00	116.01	39.02	0.1225	4.78	
49	51.00	78.00	116.01	38.01	0.1092	4.15	
50	59.00	83.00	116.00	33.00	0.0967	3.19	
51	67.00	84.00	115.99	31.98	0.0848	2.71	
52	60.00	85.00	98.99	13.99	0.0733	1.03	
53	73.00	85.99	97.00	11.00	0.0622	0.68	
54	85.00	86.00	96.99	11.00	0.0515	0.57	
55	85.99	86.98	95.00	8.02	0.0409	0.33	
56	86.98	88.00	93.00	4.99	0.0305	0.15	
57	93.00	88.01	92.01	4.00	0.0203	0.08	
58	98.99	88.98	91.00	2.01	0.0101	0.02	
59	77.00	90.00	90.00	0.00		b =	163.79
60	54.00	91.00	88.98	-2.01		S =	26.93
61	54.00	92.01	88.01	-4.00			
62	84.00	93.00	88.00	-4.99			
63	91.00	95.00	86.98	-8.02			
64	97.00	96.99	86.00	-11.00			
65	116.01	97.00	85.99	-11.00			
66	135.01	98.99	85.00	-13.99			
67	116.00	115.99	84.00	-31.98			
68	96.99	116.00	83.00	-33.00			
69	129.01	116.01	78.00	-38.01			
70	160.99	116.01	77.00	-39.02			
71	122.00	122.00	73.00	-49.00			
72	83.00	123.00	67.00	-56.00			
73	86.00	126.87	60.00	-66.87			
74	88.00	127.00	59.00	-68.00			
75	90.00	127.99	54.00	-73.99			
76	127.00	129.01	54.00	-75.00			
77	127.99	135.01	51.00	-84.00			
78	123.00	160.99	35.00	-125.99			

W=0.974 > 0.939

→ Satisfied

Table (35): Shapiro-Wilk Test for the Data of COD's 3rd quarter

No	COD mg/l	Ordering COD (1)	Inverse Order COD (2)	2-1	a(n-1+I)	I*J	
79	120.98	39.00	275.02	236.02	0.3989	94.15	
80	149.01	72.03	231.98	159.95	0.2755	44.07	
81	115.99	79.08	223.02	143.94	0.2380	34.26	
82	166.00	83.06	188.07	105.02	0.2104	22.10	
83	133.00	88.00	181.05	93.04	0.1880	17.49	
84	39.00	90.07	174.92	84.85	0.1689	14.33	
85	101.02	91.93	166.00	74.07	0.1520	11.26	
86	83.06	93.20	156.03	62.83	0.1366	8.58	
87	88.00	94.99	154.97	59.98	0.1225	7.35	
88	94.99	101.02	149.01	47.99	0.1092	5.24	
89	103.99	103.99	147.08	43.09	0.0967	4.17	
90	113.91	104.01	143.08	39.07	0.0848	3.31	
91	140.04	104.05	140.04	35.99	0.0733	2.64	
92	154.97	107.07	133.00	25.93	0.0622	1.61	
93	174.92	110.05	127.92	17.87	0.0515	0.92	
94	147.08	113.91	127.02	13.11	0.0409	0.54	
95	143.08	114.05	120.98	6.94	0.0305	0.21	
96	104.01	114.05	120.93	6.88	0.0203	0.14	
97	114.05	115.07	117.04	1.96	0.0101	0.02	
98	117.04	115.99	115.99	0.00		b=	272.38
99	114.05	117.04	115.07	-1.96		S=	46.07
100	91.93	120.93	114.05	-6.88			
101	127.02	120.98	114.05	-6.94			
102	110.05	127.02	113.91	-13.11			
103	107.07	127.92	110.05	-17.87			
104	127.92	133.00	107.07	-25.93			
105	181.05	140.04	104.05	-35.99			
106	120.93	143.08	104.01	-39.07			
107	275.02	147.08	103.99	-43.09			
108	188.07	149.01	101.02	-47.99			
109	231.98	154.97	94.99	-59.98			
110	72.03	156.03	93.20	-62.83			
111	223.02	166.00	91.93	-74.07			
112	115.07	174.92	90.07	-84.85			
113	79.08	181.05	88.00	-93.04			
114	156.03	188.07	83.06	-105.02			
115	104.05	223.02	79.08	-143.94			
116	93.20	231.98	72.03	-159.95			
117	90.07	275.02	39.00	-236.02			

$W = 0.920 < 0.939$

didn't Satisfied

Table (36): Shapiro-Wilk Test for the Data of COD's 4th quarter

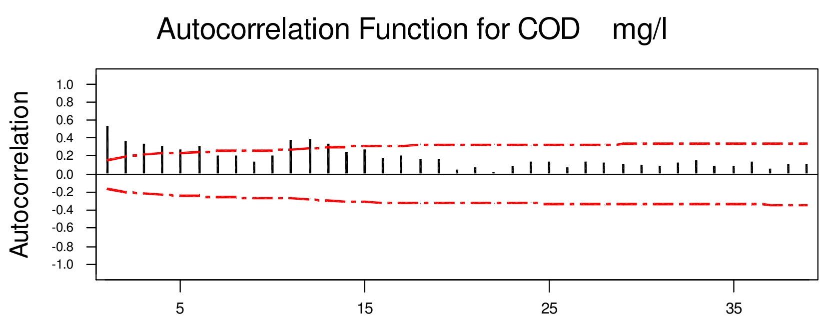
No	COD mg/l	Ordering COD (1)	Inverse order COD (2)	2-1	a(n-1+I)	I*J	
118	116.08	68.03	302.96	234.93	0.3989	93.71	
119	147.07	86.04	298.96	212.92	0.2755	58.66	
120	302.96	88.00	268.11	180.11	0.2380	42.87	
121	68.03	88.96	258.93	169.97	0.2104	35.76	
122	126.96	93.96	255.94	161.98	0.1880	30.45	
123	99.96	94.90	230.89	135.99	0.1689	22.97	
124	88.00	99.96	229.93	129.97	0.1520	19.76	
125	93.96	102.97	226.00	123.03	0.1366	16.81	
126	184.93	103.95	219.02	115.07	0.1225	14.10	
127	88.96	104.00	207.14	103.14	0.1092	11.26	
128	103.95	106.91	206.99	100.08	0.0967	9.68	
129	94.90	111.11	205.97	94.86	0.0848	8.04	
130	106.91	116.08	203.06	86.98	0.0733	6.38	
131	111.11	126.96	194.00	67.05	0.0622	4.17	
132	152.00	135.95	184.93	48.98	0.0515	2.52	
133	203.06	136.06	184.08	48.02	0.0409	1.96	
134	86.04	147.07	180.98	33.91	0.0305	1.03	
135	104.00	152.00	177.87	25.87	0.0203	0.53	
136	194.00	162.94	176.07	13.13	0.0101	0.13	
137	255.94	164.98	164.98	0.00		b=	380.79
138	268.11	176.07	162.94	-13.13		S=	63.60
139	230.89	177.87	152.00	-25.87			
140	177.87	180.98	147.07	-33.91			
141	207.14	184.08	136.06	-48.02			
142	102.97	184.93	135.95	-48.98			
143	176.07	194.00	126.96	-67.05			
144	206.99	203.06	116.08	-86.98			
145	180.98	205.97	111.11	-94.86			
146	219.02	206.99	106.91	-100.08			
147	164.98	207.14	104.00	-103.14			
148	229.93	219.02	103.95	-115.07			
149	136.06	226.00	102.97	-123.03			
150	184.08	229.93	99.96	-129.97			
151	205.97	230.89	94.90	-135.99			
152	258.93	255.94	93.96	-161.98			
153	298.96	258.93	88.96	-169.97			
154	162.94	268.11	88.00	-180.11			
155	135.95	298.96	86.04	-212.92			
156	226.00	302.96	68.03	-234.93			

W=0.943> 0.939

Satisfied

3.5.3.3 order of (AR)

For water quality like King Talal Dam, the value of AR, which is expressed by the item (p) shall not be more than 1 since the autocorrelation for a particle of COD does not need more than 1 month till it analyze (Viessman and Lewis, 1996). From Figure (34) it can be seen that the value of AR is 1, the value of p that will be used is 1 for the COD variable.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.54	6.73	46.22	13	0.33	2.23	220.51	25	0.14	0.86	275.25	37	0.06	0.35	304.30
2	0.37	3.65	67.84	14	0.24	1.55	230.30	26	0.08	0.47	276.39	38	0.11	0.63	306.72
3	0.34	3.12	86.42	15	0.27	1.71	242.76	27	0.14	0.82	279.95	39	0.11	0.65	309.37
4	0.31	2.66	101.68	16	0.18	1.12	248.30	28	0.13	0.78	283.21				
5	0.27	2.23	113.53	17	0.20	1.24	255.34	29	0.12	0.71	285.93				
6	0.31	2.48	129.23	18	0.17	1.04	260.37	30	0.10	0.58	287.81				
7	0.20	1.57	136.06	19	0.16	0.97	264.89	31	0.09	0.53	289.38				
8	0.20	1.53	142.78	20	0.04	0.27	265.23	32	0.13	0.77	292.71				
9	0.14	1.05	146.07	21	0.08	0.46	266.27	33	0.15	0.86	296.95				
10	0.20	1.53	153.14	22	0.02	0.13	266.36	34	0.09	0.52	298.53				
11	0.37	2.72	176.45	23	0.09	0.56	267.93	35	0.08	0.49	299.98				
12	0.38	2.68	201.49	24	0.14	0.83	271.42	36	0.13	0.77	303.57				

Figure (34) Autocorrelation Function for COD Variable

3.5.3.4 order of moving average (MA)

After finding the value of AR, which was 1, the following procedure is to determine the value of MA, which is expressed by the item (q). Figure (35) shows the change between the real data of the variable COD and it's moving average with different lengths of p.

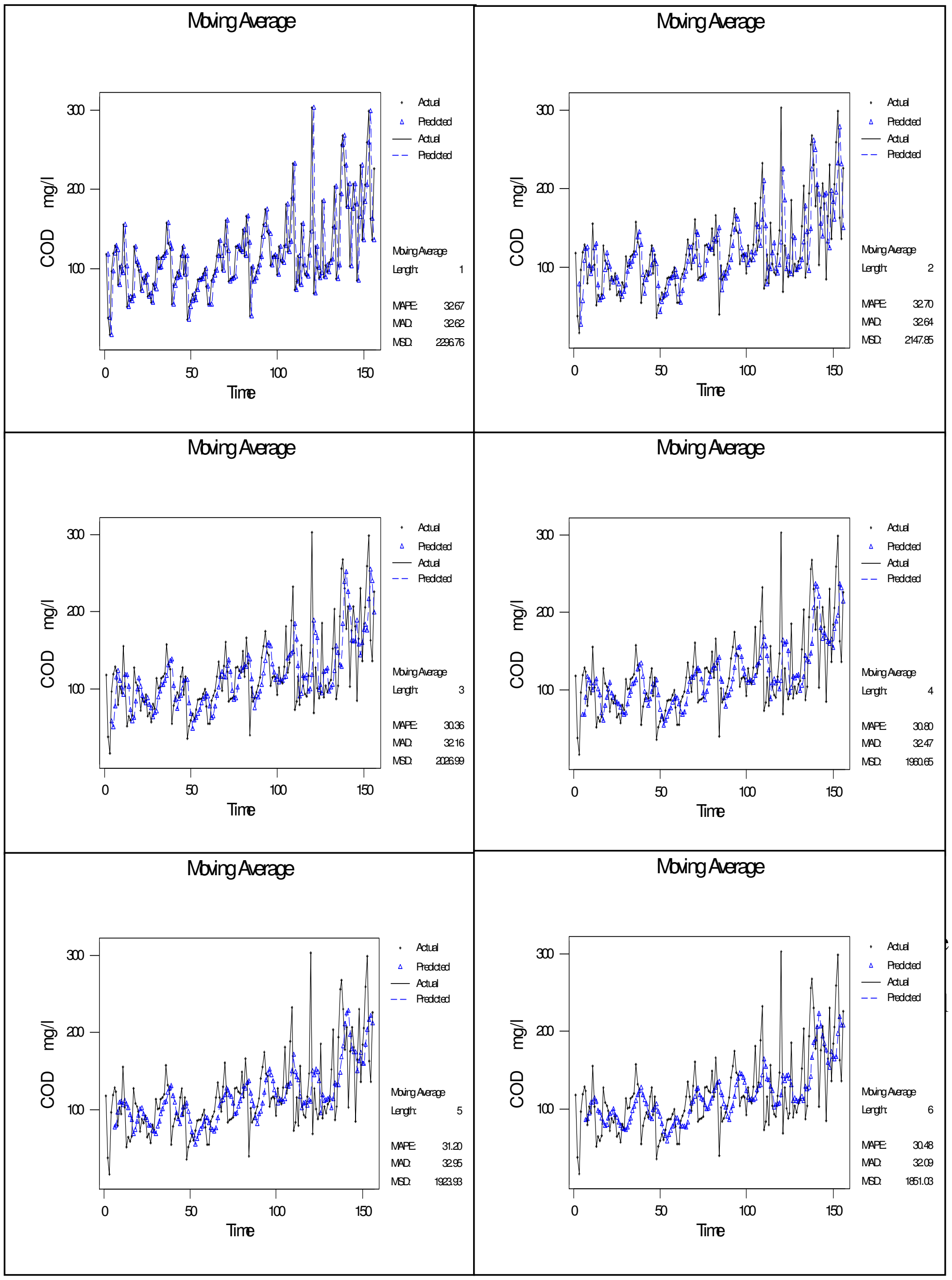


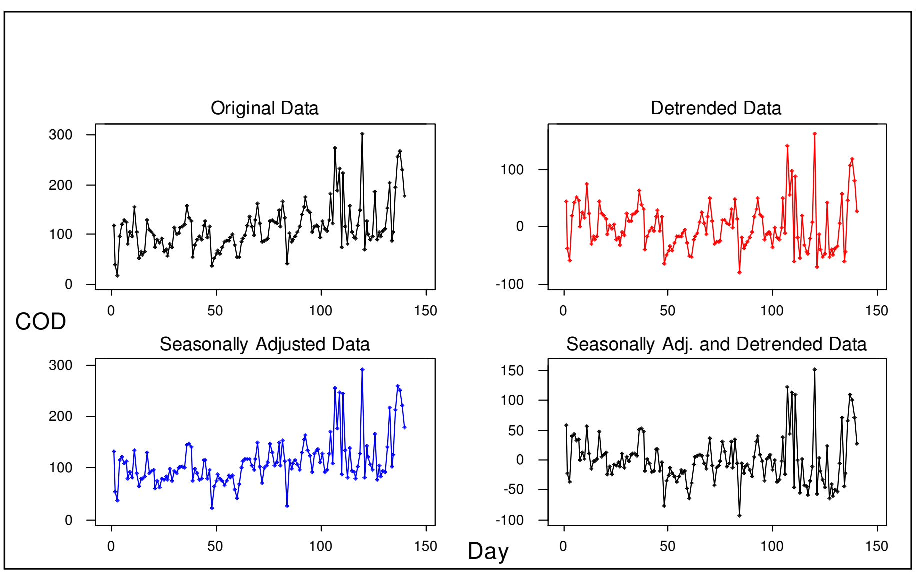
Figure (35) Moving Average of COD with Different Values of (p)

The moving average can be determined from Figure (35) when the difference between the previous length of p and the followed one have a small difference and that occurred when the value of p was 4 (as shown in Figure (35)), so the COD variable has a value of MA(4).

3.5.3.5 order of (I)

The last coefficient of ARIMA's parameters is the integrated model (I), which expressed by the item (d). The data should be differenced when there is trend or shift or seasonality in the data, otherwise there is no need to make differentiation for the data. Figure (36) consists of four graphs, which provides a good idea if there is a difference between the original, detrended, seasonally adjusted, and seasonally adjusted and detrended data. It is shown from these four graphs that there are a difference between the original figure and the detrended one but in the seasonal case they are almost the same, which means that the detrended effect could take into consideration.

Figure (36): Component Analysis for COD mg/l



Two season; summer and winter can affect seasonality in Jordan, so if the data has no trend effect, then the value of d=0 and if we have trend effect then the value of d=2. Figures (37) and (38) provides ARIMA model diagnostics for ARIMA = (1,0,4) and for ARIMA= (1,2,4). The two figures indicate the same results. It can be concluded that the data has trend effect so ARIMA (1,2,4) should be used.

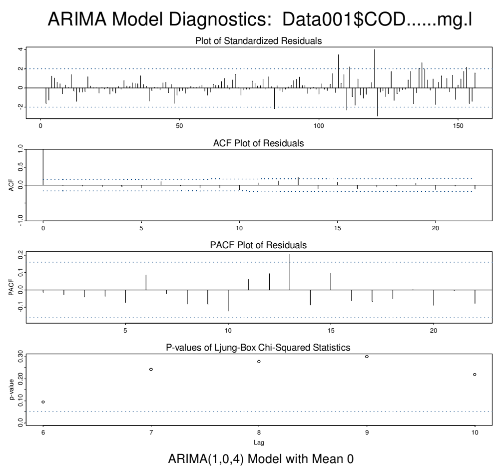


Figure (37): ARIMA (1,0,4) Diagnostics for COD

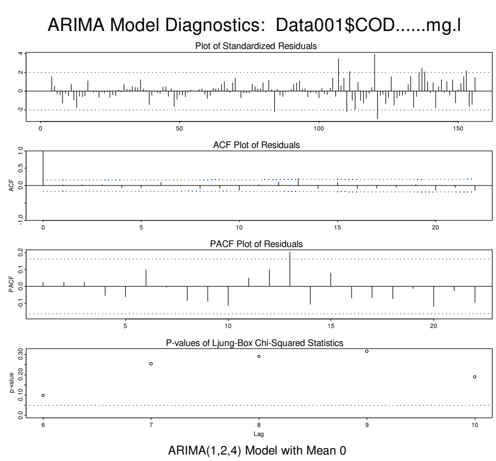


Figure (38): ARIMA (1,2,4) Diagnostics for COD

3.5.3.6 forecasting future values

The following procedure will be used in the forecasting: The values of the data collected will be divided into two parts, the first part consists of 90% of the real data, and this data will be analyzed and predicted. And the second part consists of the last 10% of the real data, and this part will be compared with the predicted values in the mean. The best model is the one that gives the least error in mean.

A- deterministic forecasting

A1- linear regression model

The regression of the additive linear trend is shown in Figure (39).

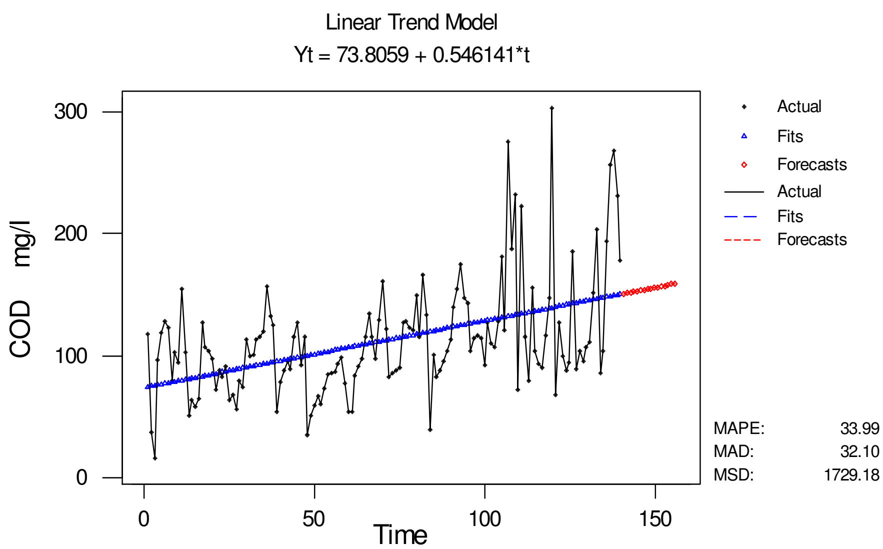


Figure (39): Trend Analysis for COD mg/l

It can be observed from the above figure and equation of the linear trend that the data is increasing. Table (37) shows the linear prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (37): The values of the predicted and actual data by linear regression for COD variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	150.81	207.14
2	142	151.36	102.97
3	143	151.90	176.07
4	144	152.45	206.99
5	145	153.00	180.98
6	146	153.54	84.18
7	147	154.09	164.98
8	148	154.64	229.93
9	149	155.18	136.06
10	150	155.73	184.08
11	151	156.27	205.97
12	152	156.82	258.93
13	153	157.37	298.96
14	154	157.91	162.94
15	155	158.46	135.95
16	156	159.00	226.00

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 19.5%, that the linear trend model did not satisfy the forecasting for the COD variable.

A2- quadratic regression model

The regression of the additive quadratic trend is shown in Figure (40).

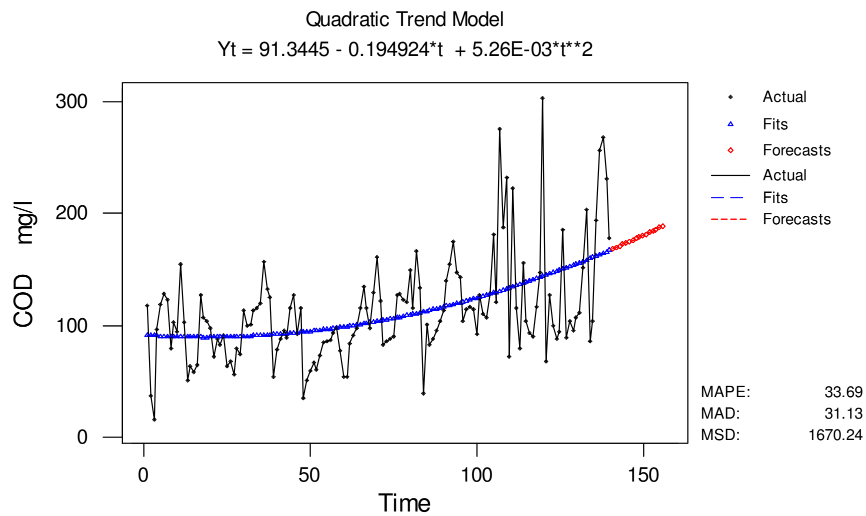


Figure (40):Trend Analysis for COD mg/l

It can be observed from the above figure and the equation of the quadratic trend that the data is increasing upward. Table (38) shows the quadratic prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (38): The values of the predicted and actual data by quadratic regression for COD variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	168.35	207.14
2	142	169.64	102.97
3	143	170.95	176.07
4	144	172.26	206.99
5	145	173.58	180.98
6	146	174.92	84.18
7	147	176.26	164.98
8	148	177.62	229.93
9	149	178.98	136.06
10	150	180.36	184.08
11	151	181.75	205.97
12	152	183.15	258.93
13	153	184.55	298.96
14	154	185.97	162.94
15	155	187.40	135.95
16	156	188.84	226.00

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 3.8%, that the quadratic trend model has satisfied the forecasting for the COD variable.

A3- exponential growth regression model

The regression of the additive exponential growth trend model is shown in Figure (41).

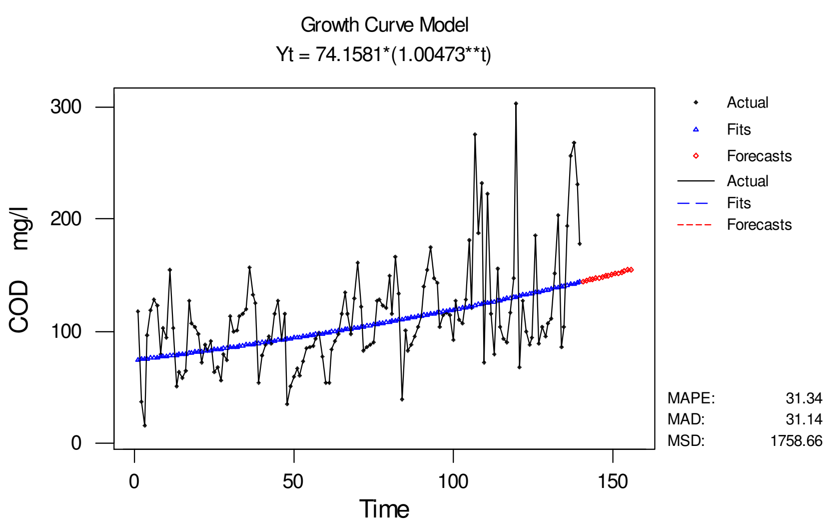


Figure (41): Trend Analysis for COD mg/l

It can be observed from the above figure and equation of the exponential growth trend that the data has an increasing trend. Table (39) shows the exponential growth prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (39): The values of the predicted and actual data by exponential growth regression for COD variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	71.01	207.14
2	142	71.14	102.97
3	143	71.26	176.07

4	144	71.38	206.99
5	145	71.51	180.98
6	146	71.63	84.18
7	147	71.76	164.98
8	148	71.88	229.93
9	149	72.01	136.06
10	150	72.13	184.08
11	151	72.26	205.97
12	152	72.38	258.93
13	153	72.51	298.96
14	154	72.64	162.94
15	155	72.76	135.95
16	156	72.89	226.00

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 23.8%, that the exponential growth trend model did not satisfy the forecasting for the COD variable.

A4- single exponential smoothing model

The regression of the additive single exponential smoothing trend model is shown in Figure (42).

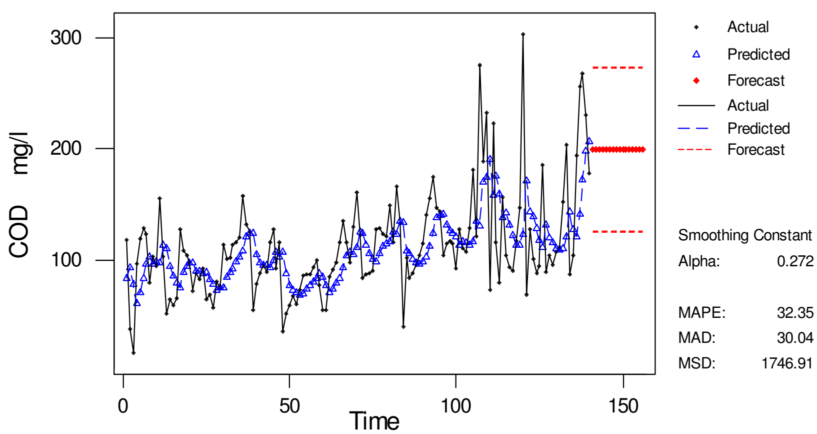


Figure (42): Single Exponential Smoothing for COD mg/l

Table (40) shows the single exponential smoothing prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (40) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (40): Forecasted, lower, upper and actual values by single exponential smoothing for COD variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	199.22	125.62	272.82	207.14
2	142	199.22	125.62	272.82	102.97
3	143	199.22	125.62	272.82	176.07
4	144	199.22	125.62	272.82	206.99
5	145	199.22	125.62	272.82	180.98
6	146	199.22	125.62	272.82	84.18
7	147	199.22	125.62	272.82	164.98
8	148	199.22	125.62	272.82	229.93
9	149	199.22	125.62	272.82	136.06
10	150	199.22	125.62	272.82	184.08
11	151	199.22	125.62	272.82	205.97
12	152	199.22	125.62	272.82	258.93
13	153	199.22	125.62	272.82	298.96
14	154	199.22	125.62	272.82	162.94
15	155	199.22	125.62	272.82	135.95
16	156	199.22	125.62	272.82	226.00

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 7.1%, that the simple exponential smoothing trend model has satisfied the forecasting for the COD variable.

B- stochastic forecasting

B1- auto regression model

Table (41) shows the AR(1) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (41) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (41): Forecasted, lower, upper and actual values by AR(1) for COD variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	151.95	63.25	240.66	207.14
2	142	137.93	37.08	238.78	102.97
3	143	130.35	26.21	234.49	176.07
4	144	126.40	21.16	231.33	206.99
5	145	124.02	18.67	229.38	180.98
6	146	122.82	17.39	228.26	84.18
7	147	122.17	16.71	227.63	164.98
8	148	121.82	16.35	227.29	229.93
9	149	121.63	16.16	227.10	136.06
10	150	121.53	16.06	227.00	184.08
11	151	121.47	16.00	226.94	205.97
12	152	121.44	15.97	226.91	258.93
13	153	121.43	15.96	226.90	298.96
14	154	121.42	15.95	226.89	162.94
15	155	121.41	15.94	226.88	135.95
16	156	121.41	15.94	226.88	226.00

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 47.4%, that the AR(1) trend model has not satisfied the forecasting for the COD variable.

B2- moving average regression model

The regression of the additive MA(4) trend model is shown in Figure(43).

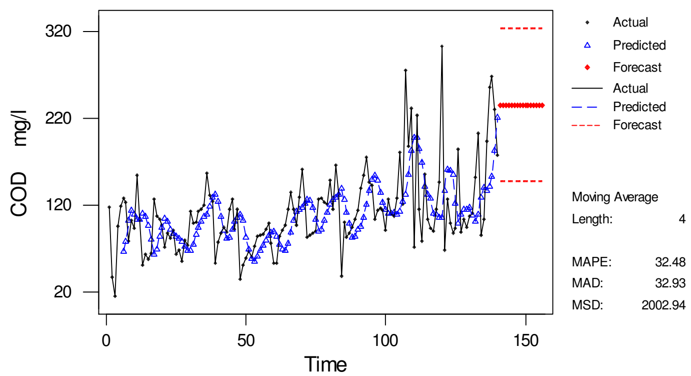


Figure (43): Moving Average Trend for COD mg/l

Table (42) shows the MA(4) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (42) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (42): Forecasted, lower, upper and actual values by MA(4) for COD variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	235.22	147.50	322.94	207.14
2	142	235.22	147.50	322.94	102.97
3	143	235.22	147.50	322.94	176.07
4	144	235.22	147.50	322.94	206.99
5	145	235.22	147.50	322.94	180.98
6	146	235.22	147.50	322.94	84.18
7	147	235.22	147.50	322.94	164.98
8	148	235.22	147.50	322.94	229.93
9	149	235.22	147.50	322.94	136.06
10	150	235.22	147.50	322.94	184.08
11	151	235.22	147.50	322.94	205.97
12	152	235.22	147.50	322.94	258.93
13	153	235.22	147.50	322.94	298.96

14	154	235.22	147.50	322.94	162.94
15	155	235.22	147.50	322.94	135.95
16	156	235.22	147.50	322.94	226.00

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 21.3%, that the MA(4) trend model did not satisfied the forecasting for the COD variable.

B3- ARIMA modeling

Table (43) shows the ARIMA(1,2,4) prediction values for the next 10% of the predicted and the real data, which equals to 16 observations. In Table (43) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values

Table (43): Forecasted, lower, upper and actual values by ARIMA(1,2,4) for COD variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	202.53	116.79	288.26	207.14
2	142	206.83	107.97	305.69	102.97
3	143	217.84	103.66	332.03	176.07
4	144	215.72	94.28	337.16	206.99
5	145	223.00	89.87	356.13	180.98
6	146	223.65	82.43	364.88	84.18
7	147	229.08	77.75	380.42	164.98
8	148	231.17	71.41	390.93	229.93
9	149	235.70	66.59	404.80	136.06
10	150	238.55	60.91	416.19	184.08
11	151	242.65	56.05	429.25	205.97
12	152	245.92	50.77	441.08	258.93
13	153	249.85	45.93	453.76	298.96
14	154	253.37	40.89	465.84	162.94
15	155	257.23	36.11	478.36	135.95
16	156	260.92	31.23	490.60	226.00

Comparing the actual values with the predicted ones, one can conclude, after calcula ting the predication error, which equals to 20.7%, that the ARIMA(1,2,4) trend model did not satisfy the forecasting for the COD variable.

3.5.3.7 results of prediction

The results of error are summarized in the following Table (44), which provides a summary of the models name used in the prediction and also it provides the percentage error.

Table (44) : Percentage of error of each model for COD variable

Model	Percentage of Mean Error
Linear Method	19.5 %
Quadratic Method	3.8 %
Exponential Growth Method	23.8 %
Simple Exponential Smoothing	7.1 %
Auto Regression, AR(1)	47.4 %
Moving Average, MA(4)	21.3 %
ARIMA (1,2, 4)	20.7 %

The previous Table (44) shows that the methods, which have satisfied the 10% acceptable prediction limits, are the quadratic and simple exponential smoothing methods. The best model that gave the least error is the quadratic method.

3.5.4 Total phosphorus (T-P) variable:

The consequences that were used to analyze the T-P variable were as follows:

3.5.4.1 detection of missing data and outliers:

From the table (1) it is observed that the data do not contain any missing data, so the second step is to find the outliers, data should be drawn in a scatter diagram (Figure 44) so that outliers will be clearly observed. These data, which contains 156 observations from January 1988 till December 2000, have approximately two outliers and they are in the following months: March 1992, and December 1999. It was observed that the rainfall in March 1992 was high, and it is known that when the rainfall is high then the T-P will get low, in December 1999 the rainfall was high so the T-P should be low (Appendix (1)). So the real data is on March 1992, the other data was assumed to be an outlier due to human error, and it should be adjusted to a new value since it may greatly influence any statistical calculations and yield biased results. The way that the outlier was adjusted was the same as the missing data treated and it was equal to the average monthly value.

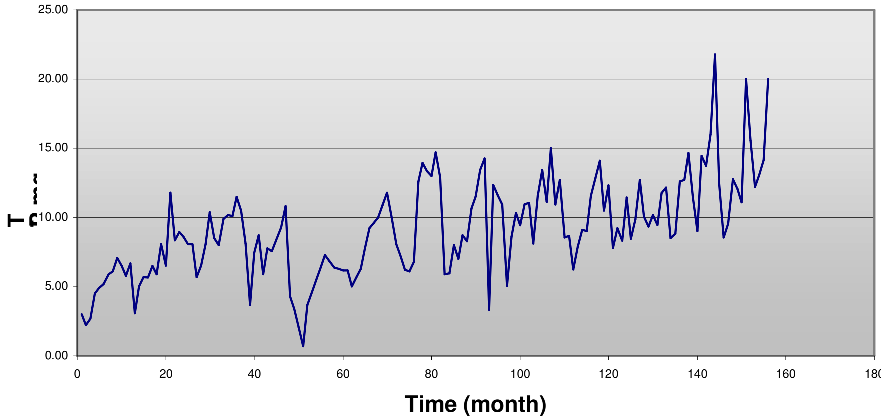


Figure (44):Original Data of T-P mg/l

Figure (45) shows the outliers for the seasonal trends for the original and the residual data, one can conclude from the charts that there are two outliers in both the original data and the residual data in the seasonal condition. Also figure (45) shows the variation in the data for the same month, it can be observe that the variation was the highest on December, and was the lowest on October. Another two outliers were found in the seasonal drawings, they are March 1992, and December 1999, these two outliers were observed in the original data, so no adjustment will be made.

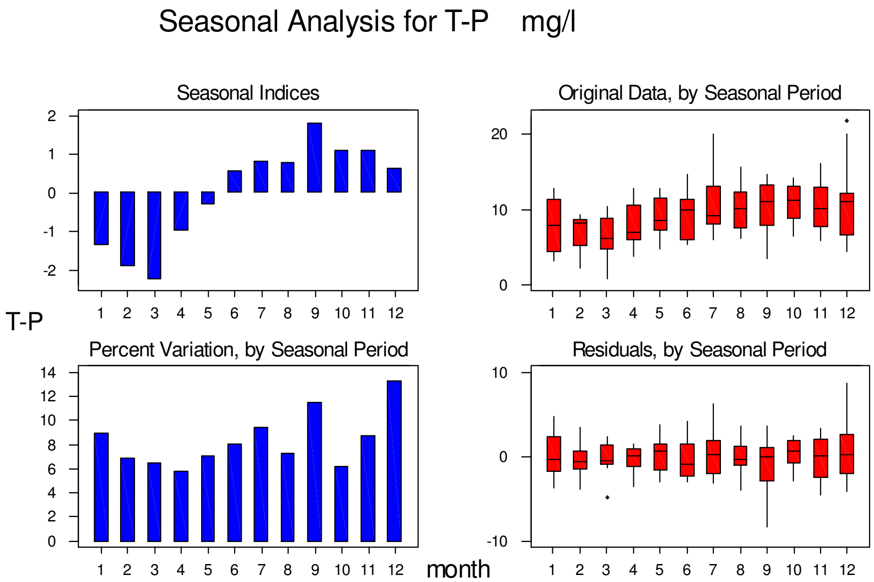


Figure (45) Outliers of Seasonal Analysis for T-P Variable

After adjustment the outliers, the new adjusted data are plotted in Figure (46), the figure shows that there are still outliers but these outliers cannot be omitted because they are real data so it can influence the statistics results. While comparing the old data (Figure 44) with the new adjusted data (Figure 46) it can be observed that two figures are quite the same and they have the same trend, so the effect of the outliers on the data was so little.

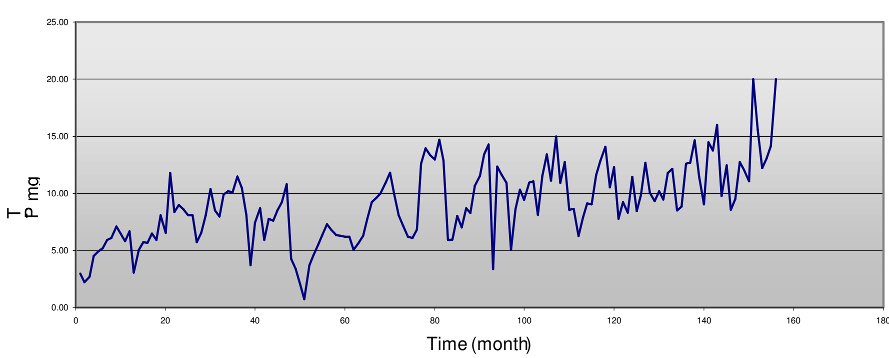


Figure (46): The New Adjusted Data of T-P mg/l

3.5.4.2 normality of data

In this section, normality of data will be checked through four procedures; first one is by drawing a histogram for Weibull's distribution model, second one is through calculating the coefficient of variance, the third one is through calculating the Kurtosis coefficient, and the fourth one is through calculating the Shapiro-Wilk test. From these four procedures, if the data was not normal then a lognormal transformation to the data will be made.

A- Weibull's distribution model histogram:

Data will be transformed to the average monthly value for the T-P variable; the calculated values were as follows

Month	Jan.	Feb.	Mar.	Apr.	May	Jun.
T-P mg/l	7.67	6.71	6.52	7.83	8.92	9.49

Month	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
T-P mg/l	10.34	10.04	10.4	10.8	10.4	9.78

The Weibull's distribution histogram is drawn for these twelve data. It can be observed from figure (47) that the data of T-P is quite normal and there is a little skewness to the left and bulked to the right, but in general the graph gives an indication that the data is normal.

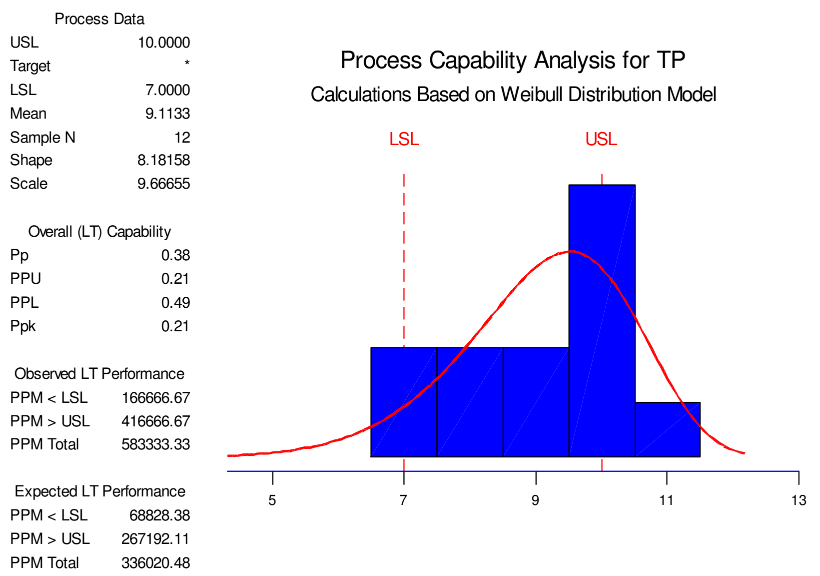


Figure (47) : Weibull Distribution Model Histogram for T-P Variable.

B- Coefficient of Variation (COV), Preliminary Test:

The data were divided into four quarters; each quarter consists of 39 data. Table (45) provides the value of the mean, variance, standard deviation, and the coefficient of variation for the T-P variable.

Table (45): The coefficient of variable for T-P

	MEAN	VARIANCE	ST. DE. (S)	C.O.V.
T-P (1st Quarter)	7.0	5.9	2.4	0.3
T-P (2nd Quarter)	7.3	7.6	2.8	0.4
T-P (3rd Quarter)	10.2	7.9	2.8	0.3
T-P (4th Quarter)	12.1	10.8	3.3	0.3

It can be shown from the table that the value of the coefficient of variation for each quarter is less than 1, which means that each quarter of the data has a little skewed (either to right or left), so the total data of the T-P variable has less skewness than each of the four T-P quarters, it can be concluded that the T-P variable does not have skewness.

C- Kurtosis coefficient (peakedness), vertical test:

To find the Kurtosis coefficient, one should find the value of K, which depends on the fourth moment about the mean and the number of samples, so that the Kurtosis can be calculated. The Kurtosis will give a good indication if the distribution is leptokurtic or platykutric. The data was divided into four quarters, Table (46) provides the values of the Kurtosis coefficient for each quarter and it provides also the calculations needed to calculate the Kurtosis coefficient, which they are: the mean, the variance or standard error, the value of K, and the Kurtosis coefficient.

Table (46): The Kurtosis Coefficient for T-P

	MEAN	VARIANCE	ST. DE. (S)	K	Kurtosis Coeff. C'_K
T-P (1st Quarter)	7.0	5.9	2.4	93.3	-0.3
T-P (2nd Quarter)	7.3	7.6	2.8	208.2	0.6
T-P (3rd Quarter)	10.2	7.9	2.8	174.7	-0.2
T-P (4th Quarter)	12.1	10.8	3.3	555.7	1.8

From table (46) one can observe that the data in the first, second, and third quarters were normally distributed (mesokurtic), the fourth quarter was fairly leptokurtic. The total data of the T-P variable can be assumed to be normally distributed (mesokurtic).

D- Shapiro-Wilk test

This is another test to show that the data we have is normal or not. Data that have been collected were divided into equal quarters, the value of (a_{n-1+I}), was taken for 20 data since the value of n-1+i was equal to 20, the value of (a_{n-1+I}) was taken from Appendix (2). The Shapiro-Wilk value was compared with the five percent critical value for sample size 20 in Appendix (3), if the value of the Shapiro-Wilk test was greater than it then the data will not show evidence of nonnormality.

From the Tables (47), (48), (49), and (50) it has been shown that the data in the first quarter is nonnormal, the second quarter has a normal data, the third has a quite not normal one, and the fourth one has a normal distribution. It can be assumed that the whole data has a tendency to be normal distribution, the T-P variable is assumed to have a normal distribution.

Table (47): Shapiro-Wilk Test for the Data of T-P's 1st quarter

No	T-P mg/l	Ordering T-P (1)	Inverse order T-P (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
1	3.00	2.20	11.80	9.60	0.3989	3.83
2	2.20	2.70	11.50	8.80	0.2755	2.42
3	2.70	3.00	10.50	7.50	0.2380	1.78
4	4.50	3.10	10.40	7.30	0.2104	1.54
5	4.90	3.70	10.20	6.50	0.1880	1.22
6	5.20	4.50	10.10	5.60	0.1689	0.95
7	5.90	4.90	9.90	5.00	0.1520	0.76
8	6.10	5.05	8.97	3.92	0.1366	0.54
9	7.10	5.20	8.60	3.40	0.1225	0.42
10	6.50	5.67	8.50	2.83	0.1092	0.31
11	5.80	5.70	8.36	2.66	0.0967	0.26
12	6.70	5.70	8.13	2.43	0.0848	0.21
13	3.10	5.80	8.10	2.30	0.0733	0.17
14	5.05	5.89	8.10	2.21	0.0622	0.14
15	5.70	5.90	8.10	2.20	0.0515	0.11
16	5.67	6.10	8.10	2.00	0.0409	0.08
17	6.50	6.50	8.00	1.50	0.0305	0.05
18	5.89	6.50	7.10	0.60	0.0203	0.01
19	8.10	6.54	6.70	0.16	0.0101	0.00
20	6.54	6.54	6.54	0.00		b= 14.78
21	11.80	6.70	6.54	-0.16		S= 2.43
22	8.36	7.10	6.50	-0.60		
23	8.97	8.00	6.50	-1.50		
24	8.60	8.10	6.10	-2.00		
25	8.10	8.10	5.90	-2.20		
26	8.10	8.10	5.89	-2.21		
27	5.70	8.10	5.80	-2.30		
28	6.54	8.13	5.70	-2.43		
29	8.10	8.36	5.70	-2.66		
30	10.40	8.50	5.67	-2.83		
31	8.50	8.60	5.20	-3.40		
32	8.00	8.97	5.05	-3.92		
33	9.90	9.90	4.90	-5.00		
34	10.20	10.10	4.50	-5.60		
35	10.10	10.20	3.70	-6.50		
36	11.50	10.40	3.10	-7.30		
37	10.50	10.50	3.00	-7.50		
38	8.13	11.50	2.70	-8.80		
39	3.70	11.80	2.20	-9.60		

$W=0.971 > 0.939$

Satisfied

Table (48): Shapiro-Wilk Test for the Data of T-P's 2nd quarter

No	T-P mg/l	Ordering T-P (1)	Inverse order T-P (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
40	7.45	0.70	13.96	13.26	0.3989	5.29
41	8.70	2.05	12.61	10.56	0.2755	2.91
42	5.90	3.40	11.82	8.42	0.2380	2.00
43	7.78	3.70	10.91	7.21	0.2104	1.52
44	7.58	4.30	10.80	6.50	0.1880	1.22
45	8.51	4.60	10.00	5.40	0.1689	0.91
46	9.25	5.04	9.96	4.92	0.1520	0.75
47	10.80	5.50	9.62	4.12	0.1366	0.56
48	4.30	5.67	9.25	3.58	0.1225	0.44
49	3.40	5.90	9.24	3.34	0.1092	0.36
50	2.05	6.10	8.70	2.60	0.0967	0.25
51	0.70	6.20	8.51	2.31	0.0848	0.20
52	3.70	6.20	8.10	1.90	0.0733	0.14
53	4.60	6.21	7.78	1.57	0.0622	0.10
54	5.50	6.29	7.77	1.48	0.0515	0.08
55	6.40	6.30	7.58	1.28	0.0409	0.05
56	7.30	6.37	7.45	1.08	0.0305	0.03
57	6.84	6.40	7.30	0.90	0.0203	0.02
58	6.37	6.81	7.16	0.35	0.0101	0.00
59	6.29	6.84	6.84	0.00		
60	6.20	7.16	6.81	-0.35		
61	6.20	7.30	6.40	-0.90		
62	5.04	7.45	6.37	-1.08		
63	5.67	7.58	6.30	-1.28		
64	6.30	7.77	6.29	-1.48		
65	7.77	7.78	6.21	-1.57		
66	9.24	8.10	6.20	-1.90		
67	9.62	8.51	6.20	-2.31		
68	10.00	8.70	6.10	-2.60		
69	10.91	9.24	5.90	-3.34		
70	11.82	9.25	5.67	-3.58		
71	9.96	9.62	5.50	-4.12		
72	8.10	9.96	5.04	-4.92		
73	7.16	10.00	4.60	-5.40		
74	6.21	10.80	4.30	-6.50		
75	6.10	10.91	3.70	-7.21		
76	6.81	11.82	3.40	-8.42		
77	12.61	12.61	2.05	-10.56		
78	13.96	13.96	0.70	-13.26		

b= 16.83

S= 2.75

W=0.984> 0.939

Satisfied

Table (49): Shapiro-Wilk Test for the Data of T-P's 3rd quarter

No	T-P mg/l	Ordering T-P (1)	Inverse order T-P (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
79	13.37	3.34	15.00	11.65	0.3989	4.65
80	12.99	5.06	14.71	9.64	0.2755	2.66
81	14.71	5.89	14.27	8.38	0.2380	1.99
82	12.90	5.98	13.44	7.46	0.2104	1.57
83	5.89	6.26	13.41	7.16	0.1880	1.35
84	5.98	7.02	13.37	6.34	0.1689	1.07
85	8.01	7.89	12.99	5.10	0.1520	0.78
86	7.02	8.01	12.90	4.89	0.1366	0.67
87	8.71	8.10	12.84	4.73	0.1225	0.58
88	8.27	8.27	12.72	4.46	0.1092	0.49
89	10.66	8.57	12.38	3.81	0.0967	0.37
90	11.55	8.58	11.62	3.03	0.0848	0.26
91	13.41	8.67	11.57	2.90	0.0733	0.21
92	14.27	8.71	11.56	2.85	0.0622	0.18
93	3.34	9.05	11.55	2.50	0.0515	0.13
94	12.38	9.12	11.12	2.00	0.0409	0.08
95	11.62	9.42	11.07	1.65	0.0305	0.05
96	10.92	10.36	10.96	0.60	0.0203	0.01
97	5.06	10.66	10.93	0.28	0.0101	0.00
98	8.58	10.92	10.92	0.00		b= 17.09
99	10.36	10.93	10.66	-0.28		S= 2.82
100	9.42	10.96	10.36	-0.60		
101	10.96	11.07	9.42	-1.65		
102	11.07	11.12	9.12	-2.00		W=0.968> 0.939
103	8.10	11.55	9.05	-2.50		
104	11.56	11.56	8.71	-2.85		Satisfied
105	13.44	11.57	8.67	-2.90		
106	11.12	11.62	8.58	-3.03		
107	15.00	12.38	8.57	-3.81		
108	10.93	12.72	8.27	-4.46		
109	12.72	12.84	8.10	-4.73		
110	8.57	12.90	8.01	-4.89		
111	8.67	12.99	7.89	-5.10		
112	6.26	13.37	7.02	-6.34		
113	7.89	13.41	6.26	-7.16		
114	9.12	13.44	5.98	-7.46		
115	9.05	14.27	5.89	-8.38		
116	11.57	14.71	5.06	-9.64		
117	12.84	15.00	3.34	-11.65		

Table (50): Shapiro-Wilk Test for the Data of T-P's 4th quarter

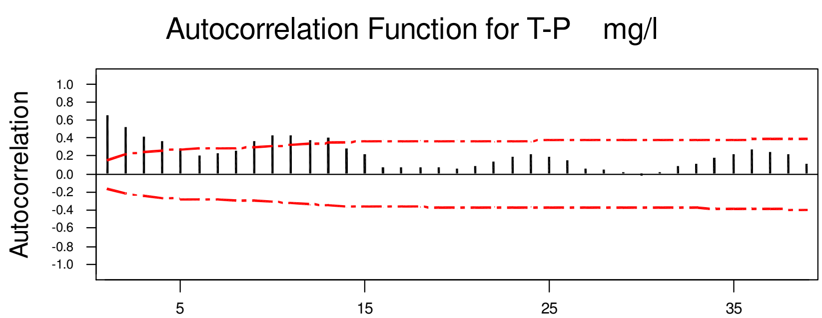
No	T-P mg/l	Ordering T-P (1)	Inverse order T-P (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
118	14.10	7.81	21.80	13.99	0.3989	5.58
119	10.50	8.32	20.02	11.70	0.2755	3.22
120	12.32	8.47	19.99	11.53	0.2380	2.74
121	7.81	8.53	16.03	7.49	0.2104	1.58
122	9.22	8.55	15.57	7.03	0.1880	1.32
123	8.32	8.84	14.64	5.80	0.1689	0.98
124	11.44	9.04	14.45	5.41	0.1520	0.82
125	8.47	9.22	14.14	4.92	0.1366	0.67
126	9.86	9.33	14.10	4.77	0.1225	0.58
127	12.71	9.49	13.73	4.24	0.1092	0.46
128	10.05	9.53	13.07	3.55	0.0967	0.34
129	9.33	9.86	12.76	2.91	0.0848	0.25
130	10.18	10.05	12.71	2.66	0.0733	0.20
131	9.49	10.18	12.71	2.53	0.0622	0.16
132	11.78	10.50	12.62	2.11	0.0515	0.11
133	12.16	11.08	12.43	1.35	0.0409	0.06
134	8.53	11.44	12.32	0.88	0.0305	0.03
135	8.84	11.46	12.22	0.76	0.0203	0.02
136	12.62	11.78	12.16	0.38	0.0101	0.00
137	12.71	12.07	12.07	0.00		b=19.12
138	14.64	12.16	11.78	-0.38		S= 3.28
139	11.46	12.22	11.46	-0.76		
140	9.04	12.32	11.44	-0.88		
141	14.45	12.43	11.08	-1.35		
142	13.73	12.62	10.50	-2.11		
143	16.03	12.71	10.18	-2.53		
144	21.80	12.71	10.05	-2.66		
145	12.43	12.76	9.86	-2.91		
146	8.55	13.07	9.53	-3.55		
147	9.53	13.73	9.49	-4.24		
148	12.76	14.10	9.33	-4.77		
149	12.07	14.14	9.22	-4.92		
150	11.08	14.45	9.04	-5.41		
151	19.99	14.64	8.84	-5.80		
152	15.57	15.57	8.55	-7.03		
153	12.22	16.03	8.53	-7.49		
154	13.07	19.99	8.47	-11.53		
155	14.14	20.02	8.32	-11.70		
156	20.02	21.80	7.81	-13.99		

W=0.891< 0.939

Did not Satisfied

3.5.4.3 order of (AR)

For water quality like King Talal Dam, the value of AR, which is expressed by the item (p) shall not be more than 1 since the autocorrelation for a particle of T-P does not need more than 1 month till it analyze (Viessman and Lewis, 1996). From Figure (48) it can be seen that the value of AR is about 4, but the value of p that will be used is 1 for the T-P variable.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.65	8.13	67.40	13	0.40	2.28	333.06	25	0.19	1.02	385.62	37	0.24	1.23	437.44
2	0.52	4.81	111.18	14	0.28	1.54	346.38	26	0.15	0.78	389.78	38	0.22	1.10	447.39
3	0.41	3.32	138.49	15	0.21	1.17	354.31	27	0.06	0.31	390.43	39	0.11	0.54	449.89
4	0.36	2.72	159.52	16	0.07	0.40	355.27	28	0.05	0.27	390.93				
5	0.28	2.00	172.05	17	0.08	0.42	356.35	29	0.02	0.09	390.98				
6	0.20	1.41	178.67	18	0.07	0.37	357.20	30	-0.01	-0.08	391.03				
7	0.23	1.58	187.25	19	0.08	0.42	358.26	31	0.02	0.13	391.15				
8	0.25	1.73	197.90	20	0.05	0.30	358.81	32	0.09	0.48	392.83				
9	0.37	2.47	220.49	21	0.09	0.49	360.29	33	0.11	0.57	395.17				
10	0.43	2.79	251.93	22	0.14	0.76	363.93	34	0.17	0.90	401.24				
11	0.42	2.61	282.31	23	0.19	1.00	370.33	35	0.22	1.13	410.90				
12	0.37	2.20	305.97	24	0.21	1.13	378.66	36	0.26	1.36	425.32				

Figure (48) Autocorrelation Function for T-P Variable

3.5.4.4 order of moving average (MA)

After finding the value of AR, which was 1, the following procedure is to determine the value of MA, which is expressed by the item (q). Figure (49) shows the change between the real data of the variable T-P and it's moving average with different lengths of p.

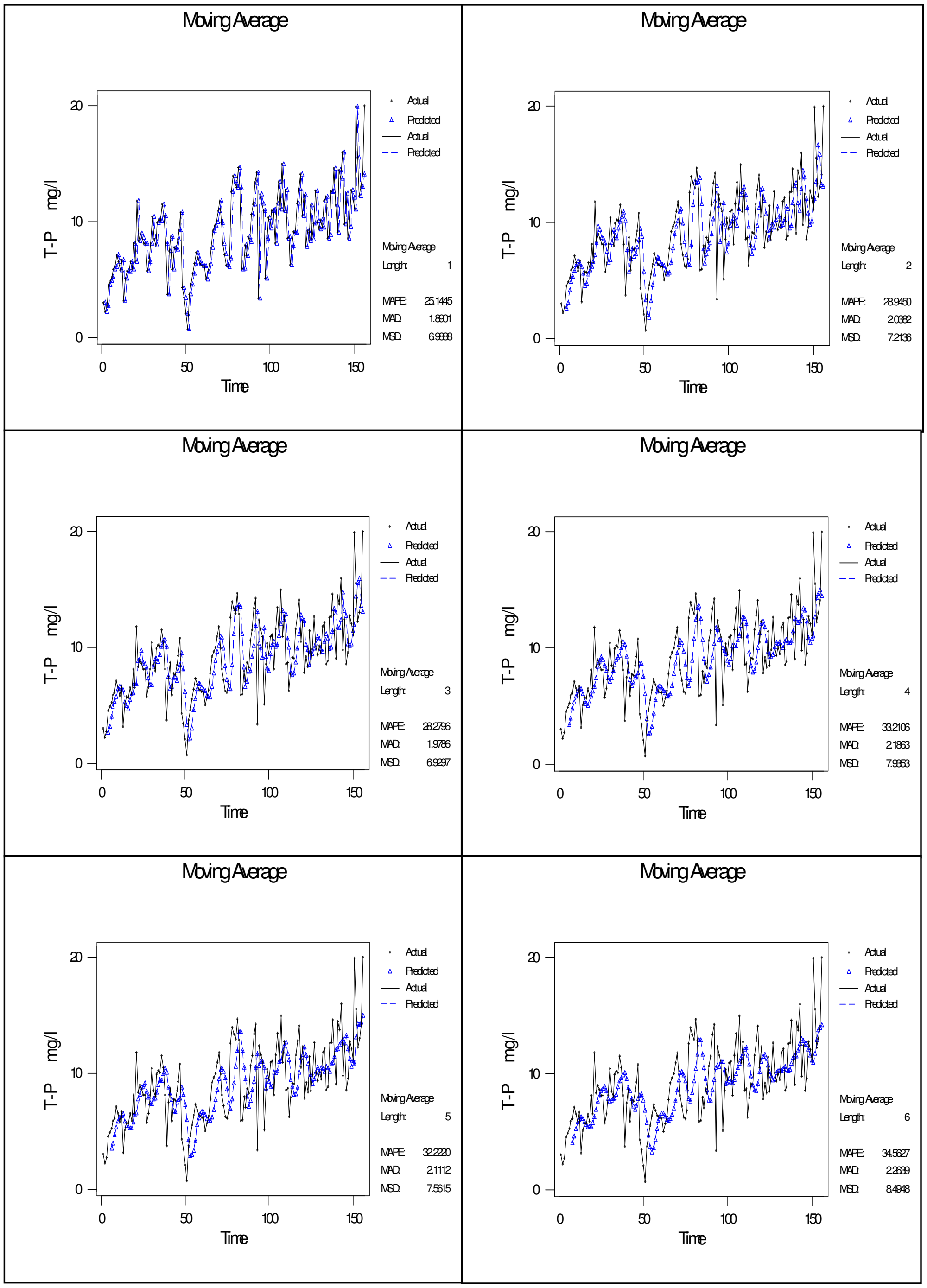


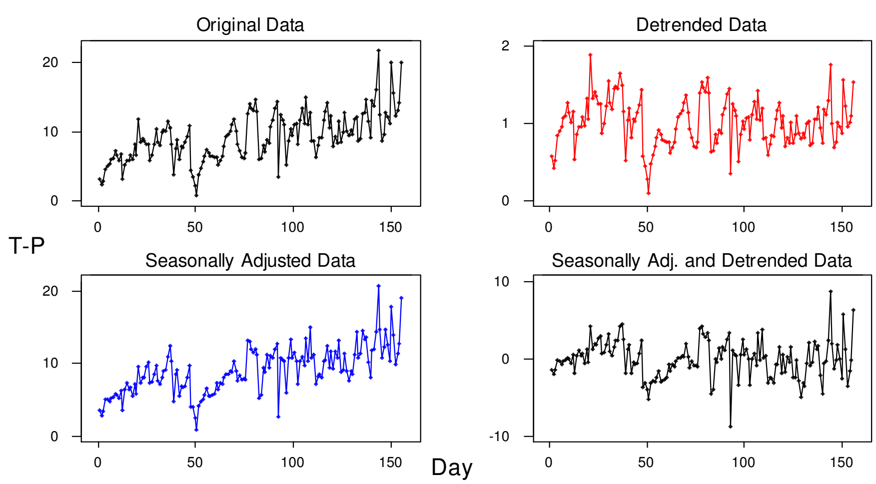
Figure (49) Moving Average of T-P with Different Values of (p)

The moving average can be determined from Figure (49) when the difference between the previous length of p and the followed one have a small difference and that occurred when the value of p was 4 (as shown in Figure (49)), so the T-P variable has a value of MA(4).

3.5.4.5 order of (I)

The last coefficient of ARIMA's parameters is the integrated model (I), which expressed by the item (d). The data should be differenced when there is trend or shift or seasonality in the data, otherwise there is no need to make differentiation for the data. Figure (50) consists of four graphs, which provides a good idea if there is a difference between the original, detrended, seasonally adjusted, and seasonally adjusted and detrended data. It is shown from these four graphs that there are a difference between the original figure and the detrended one but in the seasonal case they are almost the same, which means that the detrended effect could take into consideration

Figure (50): Component Analysis for T-P mg/l



Two season; summer and winter can affect seasonality in Jordan, so if the data has no seasonality effect, then the value of d=0 and if we have seasonality effect then the value of d=2. Figures (51), and (52) provide ARIMA model diagnostics for ARIMA = (1,0,4) and (1,2,4). It is seen from the two graphs that the residual in Figure (52) is less than Figure (51) so the coefficients of ARIMA that will be used are (1,2,4)

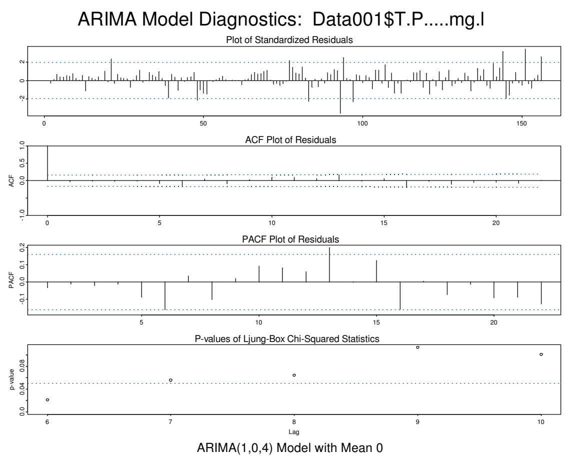


Figure (51): ARIMA (1,0,4) Model Diagnostic for T-P

Comment: Figure (8)

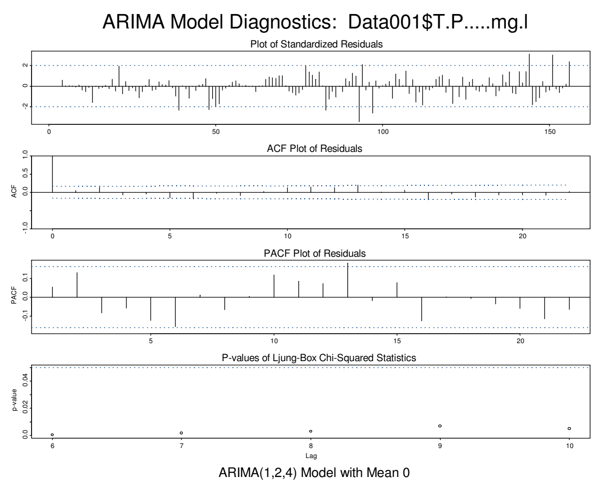


Figure (52): ARIMA (1,2,4) Model Diagnostic for T-P

Comment: Figure (8)

3.5.4.6 forecasting future values

The following procedure will be used in the forecasting: The values of the data collected will be divided into two parts, the first part consists of 90% of the real data, and this data will be analyzed and predicted. And the second part consists of the last 10% of the real data, and this part will be compared with the predicted values in the mean. The best model is the one that gives the least error in mean.

A- deterministic forecasting

A1- linear regression model

The regression of the additive linear trend is shown in Figure (53).

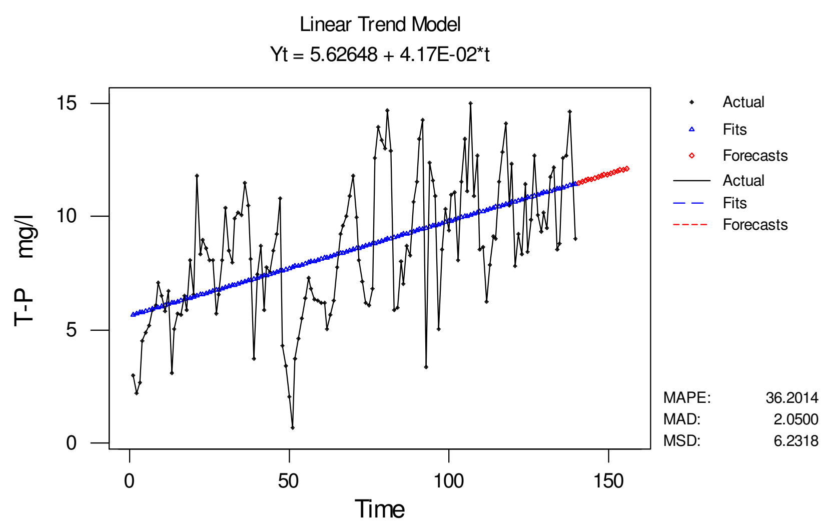


Figure (53): Trend Analysis for T-P mg/l

It can be observed from the above figure and equation of the linear trend that the data is increasing. Table (51) shows the linear prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (51): The values of the predicted and actual data by linear regression for T-P variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	11.51	14.45
2	142	11.55	13.73
3	143	11.59	16.03
4	144	11.63	9.78
5	145	11.67	12.43
6	146	11.72	8.55
7	147	11.76	9.53
8	148	11.80	12.76
9	149	11.84	12.07
10	150	11.88	11.08
11	151	11.92	19.99
12	152	11.97	15.57
13	153	12.01	12.22
14	154	12.05	13.07
15	155	12.09	14.14
16	156	12.13	20.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 13.9%, that the linear trend model did not satisfy the forecasting for the T-P variable.

A2- quadratic regression model

The regression of the additive quadratic trend is shown in Figure (54).

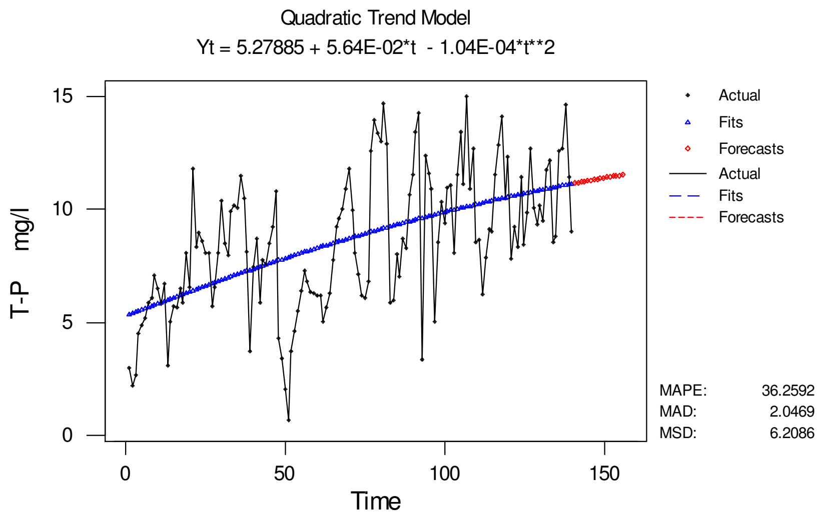


Figure (54):Trend Analysis for T-P mg/l

It can be observed from the above figure and the equation of the quadratic trend that the data is increasing upward. Table (52) shows the quadratic prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (52): The values of the predicted and actual data by quadratic regression for T-P variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	11.16	14.45
2	142	11.19	13.73
3	143	11.21	16.03
4	144	11.24	9.78
5	145	11.27	12.43
6	146	11.29	8.55
7	147	11.32	9.53
8	148	11.34	12.76
9	149	11.37	12.07
10	150	11.39	11.08
11	151	11.42	19.99
12	152	11.44	15.57
13	153	11.47	12.22
14	154	11.49	13.07
15	155	11.52	14.14
16	156	11.54	20.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 18.6%, that the quadratic trend model did not satisfy the forecasting for the T-P variable.

A3- Exponential Growth Regression Model

The regression of the additive exponential growth trend model is shown in Figure (55).

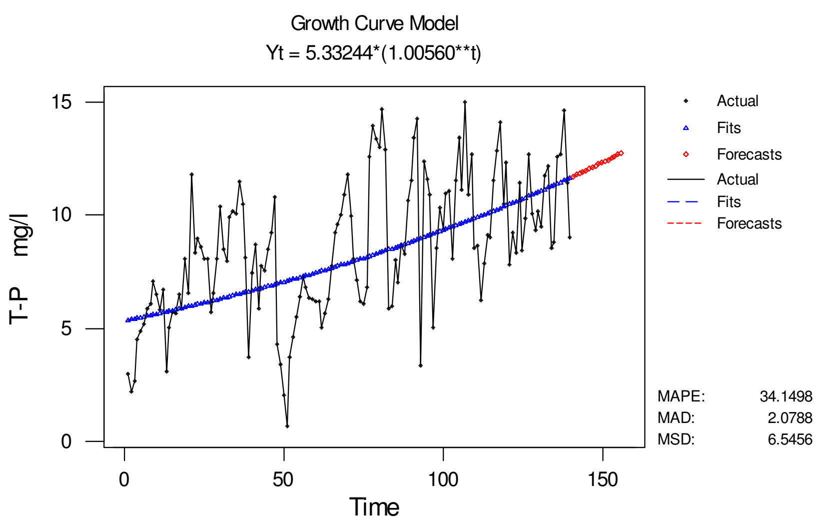


Figure (55): Trend Analysis for T-P mg/l

It can be observed from the above figure and equation of the exponential growth trend that the data has an increasing trend. Table (53) shows the exponential growth prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (53): The values of the predicted and actual data by exponential growth regression for T-P variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	11.72	14.45
2	142	11.79	13.73
3	143	11.85	16.03
4	144	11.92	9.78

5	145	11.99	12.43
6	146	12.05	8.55
7	147	12.12	9.53
8	148	12.19	12.76
9	149	12.26	12.07
10	150	12.33	11.08
11	151	12.39	19.99
12	152	12.46	15.57
13	153	12.53	12.22
14	154	12.60	13.07
15	155	12.67	14.14
16	156	12.75	20.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 10.1%, that the exponential growth trend model did not satisfy the forecasting for the T-P variable.

A4- Single Exponential Smoothing Model

The regression of the additive single exponential smoothing trend model is shown in Figure (56).

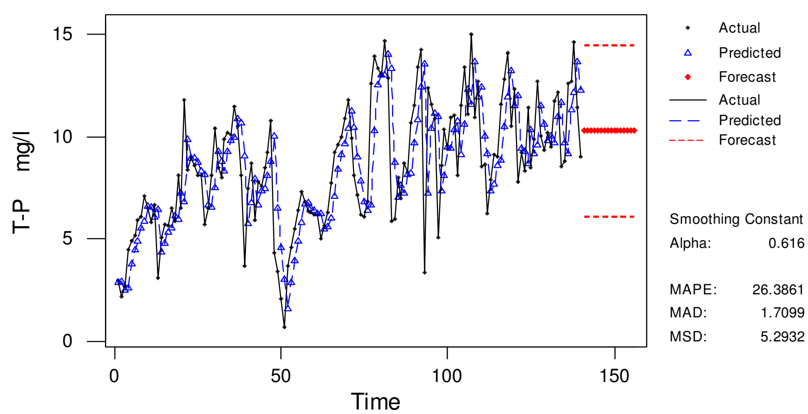


Figure (56): Single Exponential Smoothing for T-P mg/l

Table (54) shows the single exponential smoothing prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (54) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (54): Forecasted, lower, upper and actual values by single exponential smoothing for T-P variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	10.30	6.11	14.49	14.45
2	142	10.30	6.11	14.49	13.73
3	143	10.30	6.11	14.49	16.03
4	144	10.30	6.11	14.49	9.78
5	145	10.30	6.11	14.49	12.43
6	146	10.30	6.11	14.49	8.55
7	147	10.30	6.11	14.49	9.53
8	148	10.30	6.11	14.49	12.76
9	149	10.30	6.11	14.49	12.07
10	150	10.30	6.11	14.49	11.08
11	151	10.30	6.11	14.49	19.99
12	152	10.30	6.11	14.49	15.57
13	153	10.30	6.11	14.49	12.22
14	154	10.30	6.11	14.49	13.07
15	155	10.30	6.11	14.49	14.14
16	156	10.30	6.11	14.49	20.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 30.7%, that the simple exponential smoothing trend model has not satisfied the forecasting for the T-P variable.

B- stochastic forecasting

B1- auto regression model

Table (55) shows the AR(1) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (55) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (55): Forecasted, lower, upper and actual values by AR(1) for T-P variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	8.86	4.47	12.25	14.45
2	142	8.74	3.44	14.05	13.73
3	143	8.66	2.99	14.34	16.03
4	144	8.61	2.77	14.45	9.78
5	145	8.57	2.66	14.49	12.43
6	146	8.55	2.60	14.49	8.55
7	147	8.53	2.56	14.49	9.53
8	148	8.52	2.55	14.49	12.76
9	149	8.51	2.53	14.48	12.07
10	150	8.50	2.53	14.48	11.08
11	151	8.50	2.52	14.48	19.99
12	152	8.50	2.52	14.47	15.57
13	153	8.50	2.52	14.47	12.22
14	154	8.49	2.52	14.47	13.07
15	155	8.49	2.52	14.47	14.14
16	156	8.49	2.52	14.47	20.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 55.6%, that the AR(1) trend model has not satisfied the forecasting for the T-P variable.

B2- Moving Average Regression Model

The regression of the additive MA (4) trend model is shown in Figure (57).

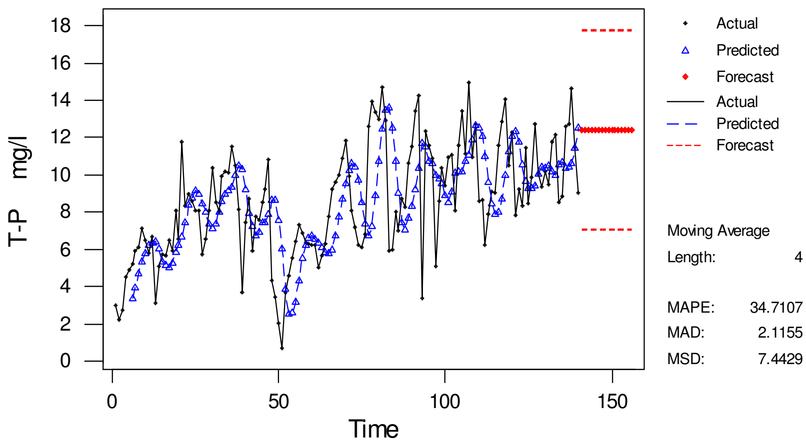


Figure (57): Moving Average for T-P mg/l

Table (56) shows the MA(4) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (56) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (56): Forecasted, lower, upper and actual values by MA(4) for T-P variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	12.41	7.06	17.76	14.45
2	142	12.41	7.06	17.76	13.73
3	143	12.41	7.06	17.76	16.03
4	144	12.41	7.06	17.76	9.78
5	145	12.41	7.06	17.76	12.43
6	146	12.41	7.06	17.76	8.55
7	147	12.41	7.06	17.76	9.53
8	148	12.41	7.06	17.76	12.76
9	149	12.41	7.06	17.76	12.07
10	150	12.41	7.06	17.76	11.08
11	151	12.41	7.06	17.76	19.99
12	152	12.41	7.06	17.76	15.57
13	153	12.41	7.06	17.76	12.22
14	154	12.41	7.06	17.76	13.07
15	155	12.41	7.06	17.76	14.14
16	156	12.41	7.06	17.76	20.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 8.5%, that the MA(4) trend model has satisfied the forecasting for the T-P variable.

B3- ARIMA modeling

Table (57) shows the ARIMA(1,2,4) prediction values for the next 10% of the predicted and the real data, which equals to 16 observations. In Table (57) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (57): Forecasted, lower, upper and actual values by ARIMA(1,2,4) for T-P variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	9.45	4.78	14.13	14.45
2	142	9.85	3.98	15.71	13.73
3	143	10.05	3.05	17.05	16.03
4	144	9.98	2.26	17.69	9.78
5	145	10.06	1.51	18.60	12.43
6	146	10.05	0.81	19.30	8.55
7	147	10.10	0.13	20.06	9.53
8	148	10.38	0.00	20.75	12.76
9	149	10.72	0.00	21.44	12.07
10	150	11.06	0.00	22.11	11.08
11	151	11.39	0.00	22.77	19.99
12	152	11.71	0.00	23.41	15.57
13	153	12.03	0.00	24.05	12.22
14	154	12.34	0.00	24.68	13.07
15	155	12.66	0.00	25.31	14.14
16	156	12.96	0.00	25.92	20.02

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 23.2%, that the ARIMA(1,2,4) trend model did not satisfy the forecasting for the T-P variable.

3.5.4.7 results of prediction

The results of error are summarized in the following Table (58), which provides a summary of the models name used in the prediction and also it provides the percentage error.

Table (58) : Percentage of error of each model for T-P variable

Model	Percentage of Mean Error
Linear Method	13.9 %
Quadratic Method	18.6 %
Exponential Growth Method	10.1 %
Simple Exponential Smoothing	30.7 %
Auto Regression, AR(1)	55.6 %
Moving Average, MA(4)	8.50 %
ARIMA (1,2,4)	23.2 %

The previous Table (58) shows that the method, which has satisfied the 10% acceptable prediction limits, is the Moving Average MA(4) model. The best model that gave the least error is the Moving Average (4) method.

3.5.5 Total nitrogen (T-N) variable:

The consequences that were used to analyze the T-N variable were as follows:

3.5.5.1 detection of missing data and outliers:

From the table (1) it is observed that the data do not contain any missing data, so the second step is to find the outliers, data should be drawn in a scatter diagram (see Figure (58)) so that outliers will be clearly observed. These data, which contains 156 observations from January 1988 till December 2000, have a clearly trend after the month July 1993. So the data will be separated into two parts. The first one is before July 1993, and the second one after it. In the first part there is just one low data that could be an outlier and this data is in August 1988. In the second part it can be seen that the data approximately do not have any outlier. In August 1988 the amount of rainfall was somehow high so the data is assumed to be a real data and not an outlier (Appendix (1)). So the real data is on August 1988, and there is not any outlier according to the data.

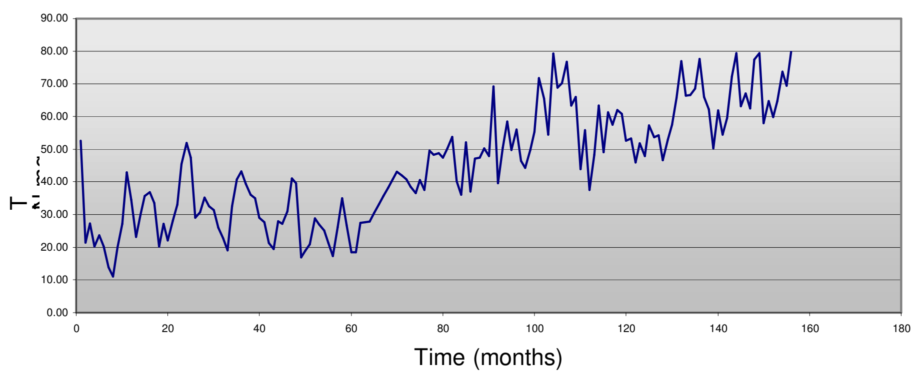


Figure (58): Original Data of T-N mg/l

Figure (59) shows the outliers for the seasonal trends for the original and the residual data, one can conclude from the charts that there are three outliers in the residual data in the seasonal condition. Also figure (59) shows the variation in the data for the same month, it can be observe that the variation was the highest on January, and was the lowest on October. Another three outliers were found in the seasonal graph, they are July 1995, August 1996, and January 1999, these three outliers were not observed in the original data. In July 1995, and August 1996 the amount of rainfall was low, these two months were treated as real data. In January 1999 the amount of rainfall was high (Appendix (1)), the data was treated as an outlier, the adjustment was made according to the average monthly way.

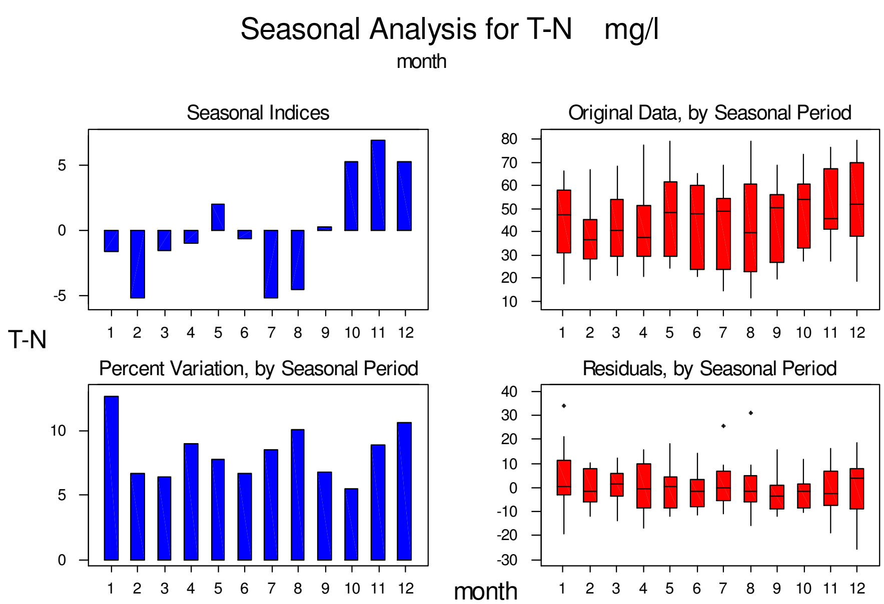


Figure (59) Outliers of Seasonal Analysis for T-N Variable

After adjustment the outliers, the new adjusted data are plotted in Figure (60), the figure shows that their still outliers but these outliers cannot be omitted because they are real data so it can influence the statistics results. While comparing the old data

(Figure 58) with the new adjusted data (Figure 60) it can be observed that two figures are quite the same and they have the same trend, so the effect of the outliers on the data was so little.

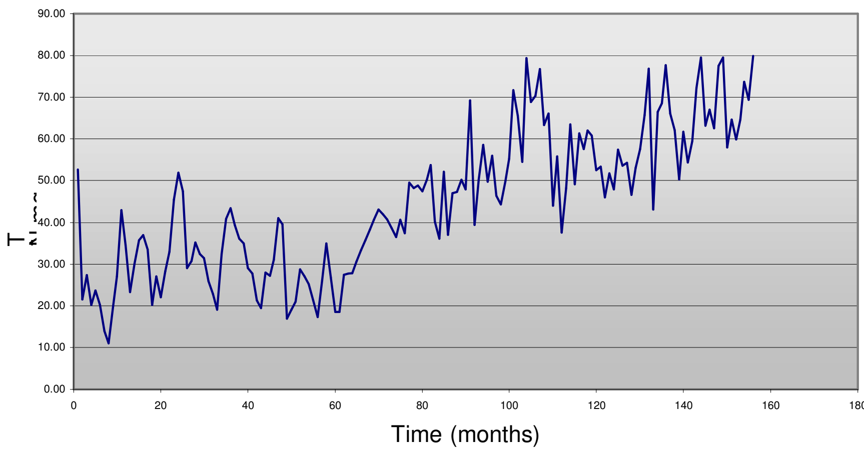


Figure (60): The New Adjusted Data of T-N mg/l

3.5.5.2 normality of data

In this section, normality of data will be checked through four procedures; first one is by drawing a histogram for Weibull's distribution model, second one is through calculating the coefficient of variance, the third one is through calculating the Kurtosis coefficient, and the fourth one is through calculating the Shapiro-Wilk test. From these four procedures, if the data was not normal then a lognormal transformation to the data will be made.

A- Weibull's distribution model histogram:

Data will be transformed to the average monthly value for the T-N variable; the calculated values were as follows

Month	Jan.	Feb.	Mar	Apr.	May	Jun.	Jul.	Aug	Sep.	Oct.	Nov	Dec.
T-N mg/l	43.2	38.8	42.6	43.0	46.0	42.3	41.1	41.2	43.1	49.0	51.8	51.7

The Weibull's distribution histogram is drawn for these twelve data. It can be observed from Figure (61) that the data of T-N is quite normal and there is a little skewness to the left and bulked to the right, but in general the graph gives an indication that the data is normal.

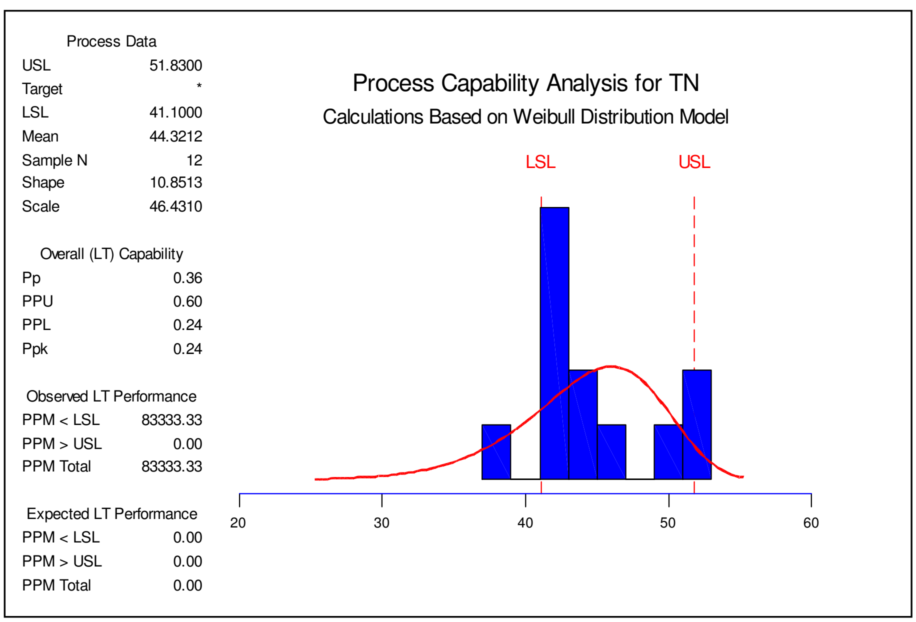


Figure (61) : Weibull Distribution Model Histogram

B- coefficient of variation (COV), preliminary test:

The data were divided into four quarters; each quarter consists of 39 data. Table (59) provides the value of the mean, variance, standard deviation, and the coefficient of variation for the T-N variable.

Table (59): The coefficient of variable for T-N

	MEAN	VARIANCE	ST. DE. (S)	C.O.V.
T-N (1st Quarter)	30.9	98.2	9.9	0.3
T-N (2nd Quarter)	30.9	79.3	8.9	0.3
T-N (3rd Quarter)	54.1	121.2	11.0	0.2
T-N (4th Quarter)	62.7	92.2	9.6	0.2

It can be shown from the table that the value of the coefficient of variation for each quarter is less than 1, which means that each quarter of the data has a little skewed (either to right or left), so the total data of the T-N variable has less skewness than each of the four T-N quarters, it can be concluded that the T-N variable does not have skewness.

C- Kurtosis coefficient (peakedness), vertical test:

To find the Kurtosis coefficient, one should find the value of K, which depends on the fourth moment about the mean and the number of samples, so that the Kurtosis can be calculated. The Kurtosis will give a good indication if the distribution is leptokurtic or platykutric. The data was divided into four quarters, Table (60) provides the values of the Kurtosis coefficient for each quarter and it provides also the calculations needed to calculate the Kurtosis coefficient, which they are: the mean, the variance or standard error, the value of K, and the Kurtosis coefficient.

Table (60): The Kurtosis Coefficient for T-N

	MEAN	VARIANCE	ST. DE. (S)	K	Kurtosis Coeff. C'_K
T-N (1st Quarter)	30.9	98.2	9.9	28379.0	-0.1
T-N (2^{nd} Quarter)	30.9	79.3	8.9	14628.0	-0.7
T-N (3^{rd} Quarter)	54.1	121.2	11.0	41056.8	-0.2
T-N (4^{th} Quarter)	62.7	92.2	9.6	20660.6	-0.6

From table (60) one can observe that the data in each quarter normally distributed (mesokurtic). The total data of the T-N variable can be assumed to be normally distributed (mesokurtic).

D- Shapiro-Wilk test

This is another test to show that the data we have is normal or not. Data that have been collected were divided into equal quarters, the value of (a_{n-1+I}), was taken for 20 data since the value of n-1+i was equal to 20, the value of (a_{n-1+I}) was taken from appendix (2). The Shapiro-Wilk value was compared with the five percent critical value for sample size 20 in Appendix (3), if the value of the Shapiro-Wilk test was greater than it then the data will not show evidence of nonnormality.

From the Tables (61), (62), (63), and (64) it has been shown that the data in each quarter was normal. It can be say that the whole data has a tendency to be normal distribution. It can be safely say that TSS variable is normally distributed.

Table (61): Shapiro-Wilk Test for the Data of T-N's 1st quarter

No	T-N mg/l	Ordering T-N (1)	Inverse order T-N (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
1	52.70	11.00	52.70	41.70	0.3989	16.63
2	21.50	13.90	51.94	38.04	0.2755	10.48
3	27.30	19.10	47.40	28.30	0.2380	6.74
4	20.20	19.90	45.46	25.56	0.2104	5.38
5	23.70	20.20	43.30	23.10	0.1880	4.34
6	20.20	20.20	42.90	22.70	0.1689	3.83
7	13.90	20.21	40.80	20.59	0.1520	3.13
8	11.00	21.50	39.40	17.90	0.1366	2.45
9	19.90	22.13	36.93	14.79	0.1225	1.81
10	27.10	22.80	36.10	13.30	0.1092	1.45
11	42.90	23.21	35.75	12.55	0.0967	1.21
12	34.60	23.70	35.20	11.50	0.0848	0.98
13	23.21	25.90	35.00	9.10	0.0733	0.67
14	30.02	27.10	34.60	7.50	0.0622	0.47
15	35.75	27.10	33.57	6.47	0.0515	0.33
16	36.93	27.30	32.96	5.66	0.0409	0.23
17	33.57	28.24	32.50	4.26	0.0305	0.13
18	20.21	29.00	32.50	3.50	0.0203	0.07
19	27.10	30.02	31.40	1.37	0.0101	0.01
20	22.13	30.70	30.70	0.00		b= 60.35
21	28.24	31.40	30.02	-1.37		S= 9.91
22	32.96	32.50	29.00	-3.50		
23	45.46	32.50	28.24	-4.26		
24	51.94	32.96	27.30	-5.66		W=0.976 > 0.939
25	47.40	33.57	27.10	-6.47		
26	29.00	34.60	27.10	-7.50		Satisfied
27	30.70	35.00	25.90	-9.10		
28	35.20	35.20	23.70	-11.50		
29	32.50	35.75	23.21	-12.55		
30	31.40	36.10	22.80	-13.30		
31	25.90	36.93	22.13	-14.79		
32	22.80	39.40	21.50	-17.90		
33	19.10	40.80	20.21	-20.59		
34	32.50	42.90	20.20	-22.70		
35	40.80	43.30	20.20	-23.10		
36	43.30	45.46	19.90	-25.56		
37	39.40	47.40	19.10	-28.30		
38	36.10	51.94	13.90	-38.04		
39	35.00	52.70	11.00	-41.70		

Table (62): Shapiro-Wilk Test for the Data of T-N's 2nd quarter

No	T-N mg/l	Ordering T-N (1)	Inverse order T-N (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
40	29.00	17.00	49.55	32.55	0.3989	12.98
41	27.70	17.30	48.20	30.91	0.2755	8.51
42	21.30	18.50	43.12	24.62	0.2380	5.86
43	19.50	18.50	41.90	23.40	0.2104	4.92
44	28.00	19.00	41.00	22.00	0.1880	4.14
45	27.19	19.50	40.71	21.21	0.1689	3.58
46	31.00	21.00	40.68	19.68	0.1520	2.99
47	41.00	21.25	40.62	19.37	0.1366	2.65
48	39.50	21.30	39.50	18.20	0.1225	2.23
49	17.00	25.20	38.60	13.40	0.1092	1.46
50	19.00	26.20	38.29	12.09	0.0967	1.17
51	21.00	26.80	37.51	10.71	0.0848	0.91
52	28.80	27.00	36.52	9.52	0.0733	0.70
53	27.00	27.19	35.76	8.57	0.0622	0.53
54	25.20	27.51	35.00	7.49	0.0515	0.39
55	21.25	27.70	33.22	5.52	0.0409	0.23
56	17.30	27.70	31.00	3.30	0.0305	0.10
57	26.20	27.88	30.55	2.67	0.0203	0.05
58	35.00	28.00	29.00	1.00	0.0101	0.01
59	26.80	28.80	28.80	0.00		b= 53.41
60	18.50	29.00	28.00	-1.00		S= 8.90
61	18.50	30.55	27.88	-2.67		
62	27.51	31.00	27.70	-3.30		
63	27.70	33.22	27.70	-5.52		
64	27.88	35.00	27.51	-7.49		
65	30.55	35.76	27.19	-8.57		
66	33.22	36.52	27.00	-9.52		
67	35.76	37.51	26.80	-10.71		
68	38.29	38.29	26.20	-12.09		
69	40.71	38.60	25.20	-13.40		
70	43.12	39.50	21.30	-18.20		
71	41.90	40.62	21.25	-19.37		
72	40.68	40.68	21.00	-19.68		
73	38.60	40.71	19.50	-21.21		
74	36.52	41.00	19.00	-22.00		
75	40.62	41.90	18.50	-23.40		
76	37.51	43.12	18.50	-24.62		
77	49.55	48.20	17.30	-30.91		
78	48.20	49.55	17.00	-32.55		

$W=0.947 > 0.939$

Satisfied

Table (63): Shapiro-Wilk Test for the Data of T-N's 3rd quarter

No	T-N mg/l	Ordering T-N (1)	Inverse order T-N (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
79	48.81	36.11	79.33	43.22	0.3989	17.24
80	47.46	37.09	76.76	39.67	0.2755	10.93
81	50.20	37.54	71.69	34.15	0.2380	8.13
82	53.74	39.48	70.23	30.75	0.2104	6.47
83	40.27	40.27	69.23	28.96	0.1880	5.44
84	36.11	43.98	68.84	24.86	0.1689	4.20
85	52.10	44.24	66.05	21.81	0.1520	3.32
86	37.09	46.40	65.54	19.15	0.1366	2.62
87	47.06	47.06	63.41	16.35	0.1225	2.00
88	47.35	47.35	63.27	15.93	0.1092	1.74
89	50.24	47.46	61.38	13.91	0.0967	1.35
90	47.91	47.91	58.54	10.63	0.0848	0.90
91	69.23	48.26	57.55	9.30	0.0733	0.68
92	39.48	48.81	55.98	7.17	0.0622	0.45
93	50.37	49.15	55.88	6.73	0.0515	0.35
94	58.54	49.74	55.30	5.56	0.0409	0.23
95	49.76	49.76	54.52	4.76	0.0305	0.15
96	55.98	50.20	53.74	3.54	0.0203	0.07
97	46.40	50.24	52.10	1.85	0.0101	0.02
98	44.24	50.37	50.37	0.00		b= 66.27
99	49.74	52.10	50.24	-1.85		S= 11.01
100	55.30	53.74	50.20	-3.54		
101	71.69	54.52	49.76	-4.76		
102	65.54	55.30	49.74	-5.56		W=0.954 > 0.939
103	54.52	55.88	49.15	-6.73		
104	79.33	55.98	48.81	-7.17		Satisfied
105	68.84	57.55	48.26	-9.30		
106	70.23	58.54	47.91	-10.63		
107	76.76	61.38	47.46	-13.91		
108	63.27	63.27	47.35	-15.93		
109	66.05	63.41	47.06	-16.35		
110	43.98	65.54	46.40	-19.15		
111	55.88	66.05	44.24	-21.81		
112	37.54	68.84	43.98	-24.86		
113	48.26	69.23	40.27	-28.96		
114	63.41	70.23	39.48	-30.75		
115	49.15	71.69	37.54	-34.15		
116	61.38	76.76	37.09	-39.67		
117	57.55	79.33	36.11	-43.22		

Table (64): Shapiro-Wilk Test for the Data of T-N's 4th quarter

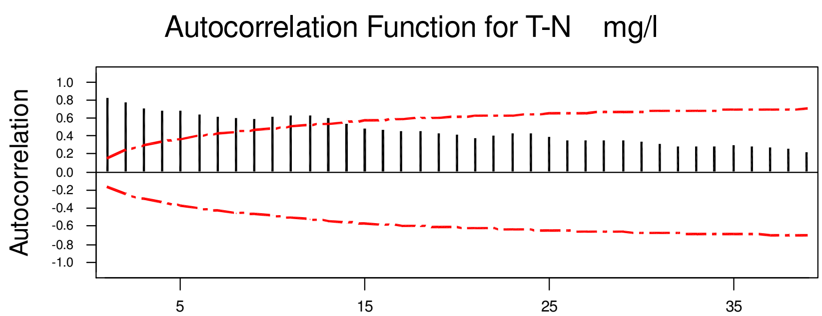
No	T-N mg/l	Ordering T-N (1)	Inverse order T-N (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
118	62.05	45.94	79.85	33.90	0.3989	13.52
119	60.84	46.58	79.51	32.93	0.2755	9.07
120	52.56	47.96	79.45	31.49	0.2380	7.49
121	53.29	50.32	77.60	27.28	0.2104	5.74
122	45.94	51.78	77.49	25.72	0.1880	4.83
123	51.78	52.56	76.89	24.33	0.1689	4.11
124	47.96	53.03	73.67	20.64	0.1520	3.14
125	57.36	53.29	72.22	18.93	0.1366	2.59
126	53.62	53.62	69.34	15.72	0.1225	1.93
127	54.28	54.28	68.63	14.35	0.1092	1.57
128	46.58	54.41	67.00	12.60	0.0967	1.22
129	53.03	57.36	66.51	9.15	0.0848	0.78
130	57.55	57.55	66.41	8.86	0.0733	0.65
131	65.72	57.92	66.07	8.15	0.0622	0.51
132	76.89	59.50	65.72	6.22	0.0515	0.32
133	66.41	59.85	64.68	4.83	0.0409	0.20
134	66.51	60.84	64.64	3.80	0.0305	0.12
135	68.63	61.80	63.13	1.33	0.0203	0.03
136	77.60	62.05	62.58	0.53	0.0101	0.01
137	66.07	62.11	62.11	0.00		b= 57.81
138	62.11	62.58	62.05	-0.53		S= 9.60
139	50.32	63.13	61.80	-1.33		
140	61.80	64.64	60.84	-3.80		
141	54.41	64.68	59.85	-4.83		
142	59.50	65.72	59.50	-6.22		
143	72.22	66.07	57.92	-8.15		
144	79.45	66.41	57.55	-8.86		
145	63.13	66.51	57.36	-9.15		
146	67.00	67.00	54.41	-12.60		
147	62.58	68.63	54.28	-14.35		
148	77.49	69.34	53.62	-15.72		
149	79.51	72.22	53.29	-18.93		
150	57.92	73.67	53.03	-20.64		
151	64.64	76.89	52.56	-24.33		
152	59.85	77.49	51.78	-25.72		
153	64.68	77.60	50.32	-27.28		
154	73.67	79.45	47.96	-31.49		
155	69.34	79.51	46.58	-32.93		
156	79.85	79.85	45.94	-33.90		

$W=0.954 > 0.939$

Satisfied

3.5.5.3 order of (AR)

For water quality like King Talal Dam, the value of AR, which is expressed by the item (p) shall not be more than 1 since the autocorrelation for a particle of T-N does not need more than 1 month till it analyze (Viessman and Lewis, 1996). From Figure (62) it can be seen that the autocorrelation between data is high and that means that the data has random distributions, but the value of p that will be used is 1 for the T-N variable.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.82	10.30	108.15	13	0.60	2.19	937.04	25	0.39	1.20	1354.88	37	0.27	0.77	1589.19
2	0.77	6.25	202.83	14	0.54	1.90	986.97	26	0.35	1.07	1378.61	38	0.25	0.71	1602.30
3	0.70	4.67	282.74	15	0.48	1.67	1027.24	27	0.36	1.07	1402.73	39	0.22	0.63	1612.78
4	0.68	3.96	356.63	16	0.47	1.61	1066.46	28	0.34	1.02	1425.39				
5	0.68	3.64	432.35	17	0.46	1.53	1103.34	29	0.35	1.04	1449.46				
6	0.64	3.18	500.37	18	0.45	1.50	1140.20	30	0.34	1.00	1472.22				
7	0.62	2.87	563.08	19	0.42	1.38	1172.63	31	0.31	0.91	1491.42				
8	0.59	2.63	621.67	20	0.41	1.32	1203.25	32	0.28	0.82	1507.22				
9	0.59	2.48	679.14	21	0.38	1.19	1228.95	33	0.28	0.81	1523.08				
10	0.61	2.51	742.78	22	0.40	1.26	1258.05	34	0.29	0.82	1539.52				
11	0.62	2.45	808.84	23	0.43	1.33	1291.77	35	0.29	0.84	1557.09				
12	0.62	2.36	875.51	24	0.43	1.32	1325.84	36	0.29	0.82	1574.19				

Figure (62) Autocorrelation Function for T-N Variable

3.5.5.4 order of moving average (MA)

After finding the value of AR, which was 1, the following procedure is to determine the value of MA, which is expressed by the item (q). Figure (63) shows the change between the real data of the variable T-N and it's moving average with different lengths of p.

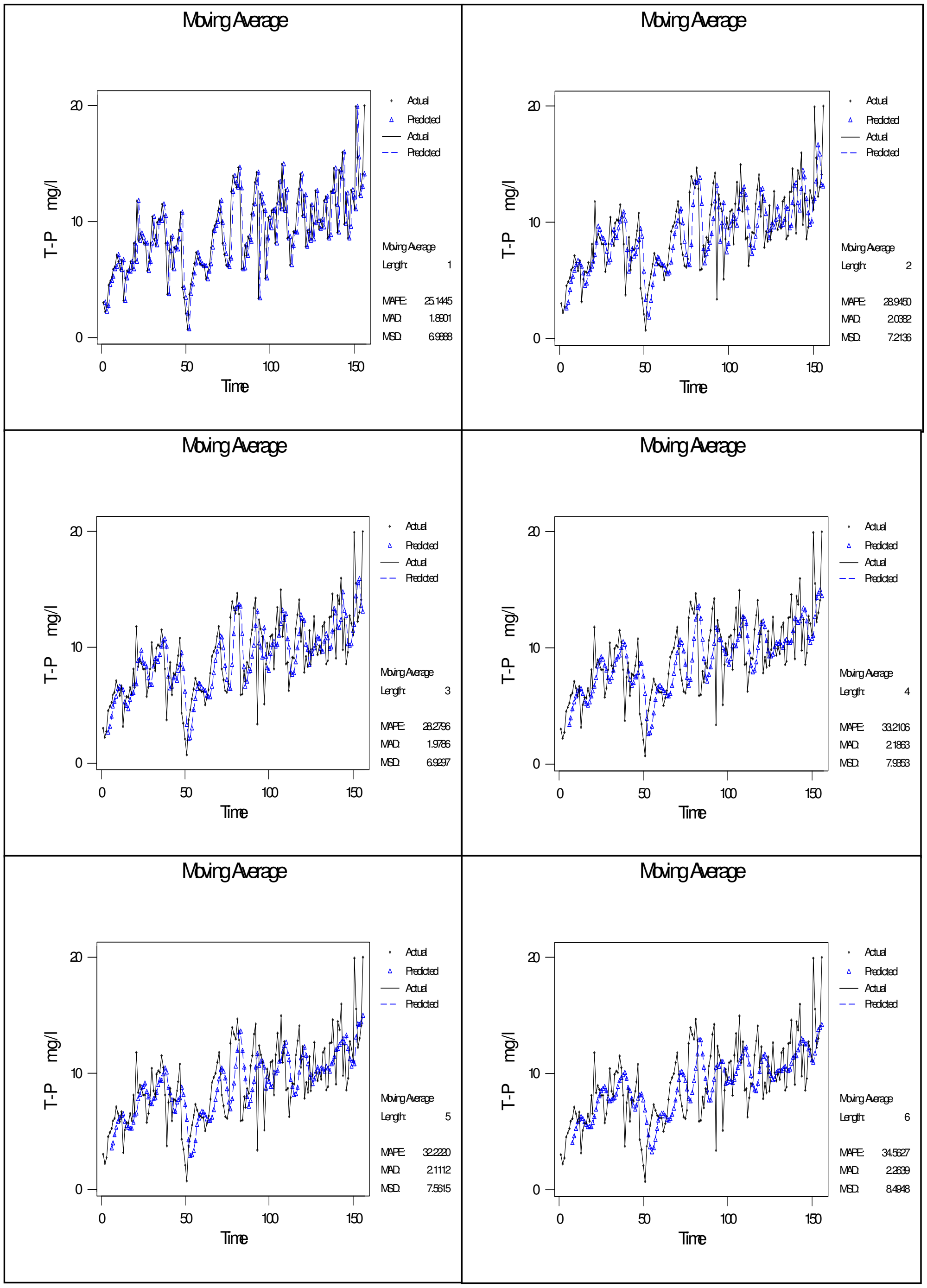


Figure (63) Moving Average of T-N with Different Values of (p)

The moving average can be determined from Figure (63) when the difference between the previous length of p and the followed one have a small difference and that occurred when the value of p was 5 (as shown in Figure (63)), so the T-N variable has a value of MA (5).

3.5.5.5 order of (I)

The last coefficient of ARIMA's parameters is the integrated model (I), which expressed by the item (d). The data should be differenced when there is trend or shift or seasonality in the data, otherwise there is no need to make differentiation for the data. Figure (64) consists of four graphs, which provides a good idea if there is a difference between the original, detrended, seasonally adjusted, and seasonally adjusted and detrended data. It is shown from these four graphs that there are a difference between the original figure and the detrended one but in the seasonal case they are almost the same, which means that the detrended effect should be taken into consideration

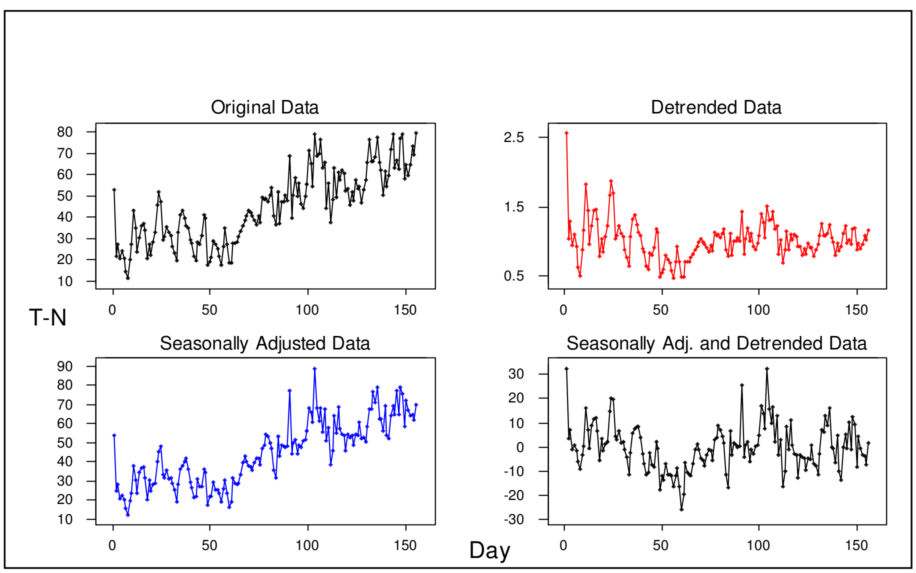


Figure (64): Component Analysis for T-N mg/l

Two season; summer and winter can affect seasonality in Jordan, so if the data has no seasonality effect, then the value of d=0 and if we have seasonality effect then the value of d=2. Figures (65), and (66) provide ARIMA model diagnostics for ARIMA = (1,0,5) and (1,2,5). It is seen from the two graphs that the residual in Figure (65) is quite the same as in Figure (66) so the coefficients of ARIMA that will be used are (1,2,5)

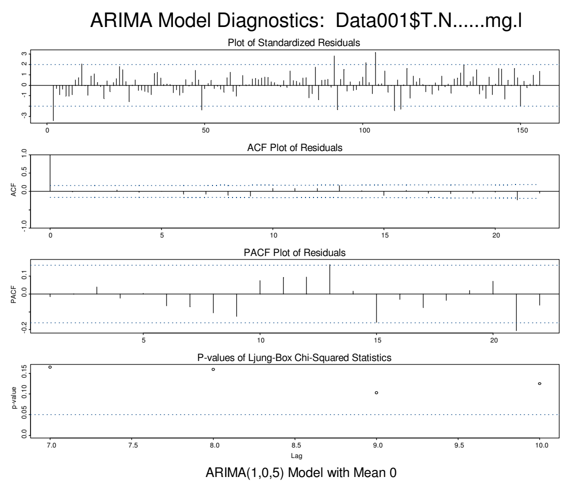


Figure (65): ARIMA (1,0,5) Model Diagnostics for T-N

Comment: Figure (8)

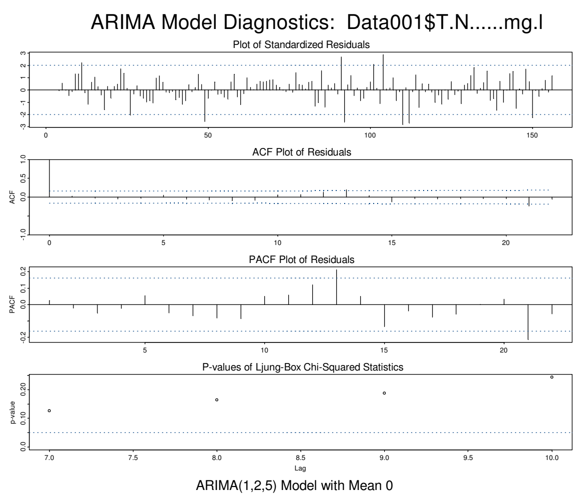


Figure (66): ARIMA (1,2,5) Model Diagnostics for T-N

Comment: Figure (8)

3.5.5.6 forecasting future values

The following procedure will be used in the forecasting: The values of the data collected will be divided into two parts, the first part consists of 90% of the real data, and this data will be analyzed and predicted. And the second part consists of the last 10% of the real data, and this part will be compared with the predicted values in the mean. The best model is the one that gives the least error in mean.

A- deterministic forecasting

A1- linear regression model

The regression of the additive linear trend is shown in Figure (67).

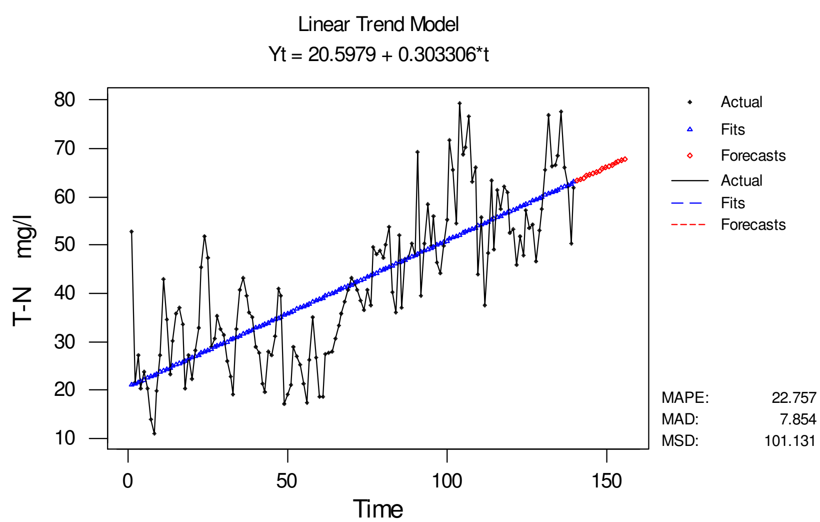


Figure (67): Trend Analysis for T-N mg/l

It can be observed from the above figure and equation of the linear trend that the data is increasing. Table (65) shows the linear prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (65): The values of the predicted and actual data by linear regression for T-N variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	63.36	54.41
2	142	63.67	59.50
3	143	63.97	72.22
4	144	64.27	79.45
5	145	64.58	63.13
6	146	64.88	67.00
7	147	65.18	62.58
8	148	65.49	77.49
9	149	65.79	79.51
10	150	66.09	57.92
11	151	66.40	64.64
12	152	66.70	59.85
13	153	67.00	64.68
14	154	67.31	73.67
15	155	67.61	69.34
16	156	67.91	79.85

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 3.3%, that the linear trend model has satisfied the forecasting for the T-N variable.

A2- quadratic regression model

The regression of the additive quadratic trend is shown in Figure (68).

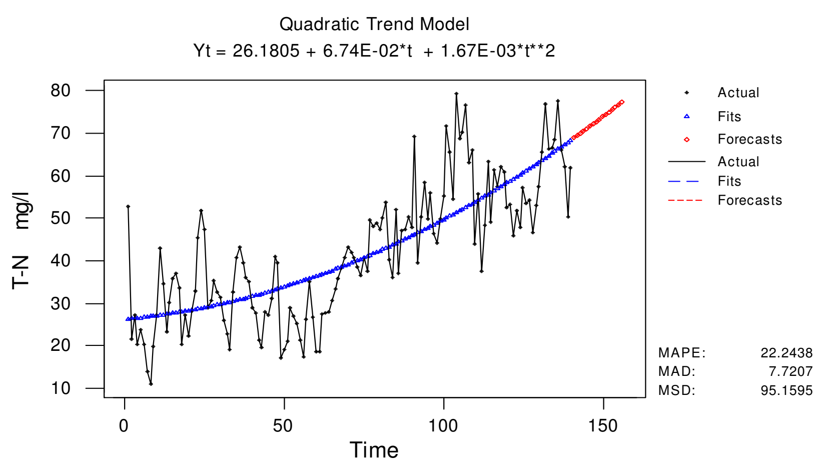


Figure (68): Trend Analysis for T-N mg/l

It can be observed from Figure (68) and the equation of the quadratic trend that the data is increasing upward. Table (66) shows the quadratic prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (66): The values of the predicted and actual data by quadratic regression for T-N variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	68.95	54.41
2	142	69.49	59.50
3	143	70.03	72.22
4	144	70.58	79.45
5	145	71.13	63.13
6	146	71.68	67.00
7	147	72.24	62.58
8	148	72.80	77.49
9	149	73.37	79.51
10	150	73.93	57.92
11	151	74.51	64.64
12	152	75.08	59.85
13	153	75.66	64.68
14	154	76.24	73.67
15	155	76.82	69.34
16	156	77.41	79.85

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 7.2%, that the quadratic trend model has satisfied the forecasting for the T-N variable.

A3- exponential growth regression model

The regression of the additive exponential growth trend model is shown in Figure (69).

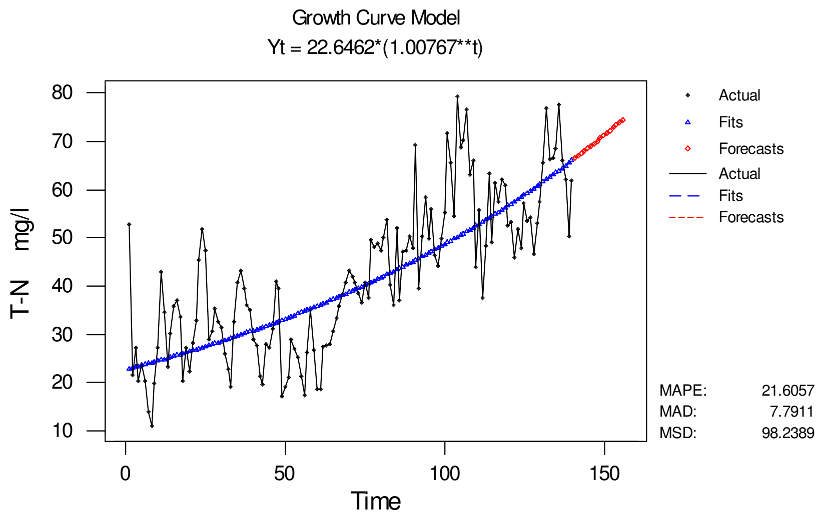


Figure (69): Trend Analysis for T-N mg/l

It can be observed from the above figure and equation of the exponential growth trend that the data has an increasing trend. Table (67) shows the exponential growth prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (67): The values of the predicted and actual data by exponential growth regression for T-N variable

Row	Period (months)	Forecasted (mg/l)	Actual (mg/l)
1	141	66.50	54.41
2	142	67.01	59.50
3	143	67.52	72.22
4	144	68.04	79.45
5	145	68.56	63.13
6	146	69.09	67.00
7	147	69.62	62.58
8	148	70.15	77.49
9	149	70.69	79.51
10	150	71.23	57.92
11	151	71.78	64.64
12	152	72.33	59.85
13	153	72.88	64.68
14	154	73.44	73.67
15	155	74.01	69.34
16	156	74.57	79.85

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 3.7%, that the exponential growth trend model has satisfied the forecasting for the T-N variable.

A4- single exponential smoothing model

The regression of the additive single exponential smoothing trend model is shown in Figure (70).

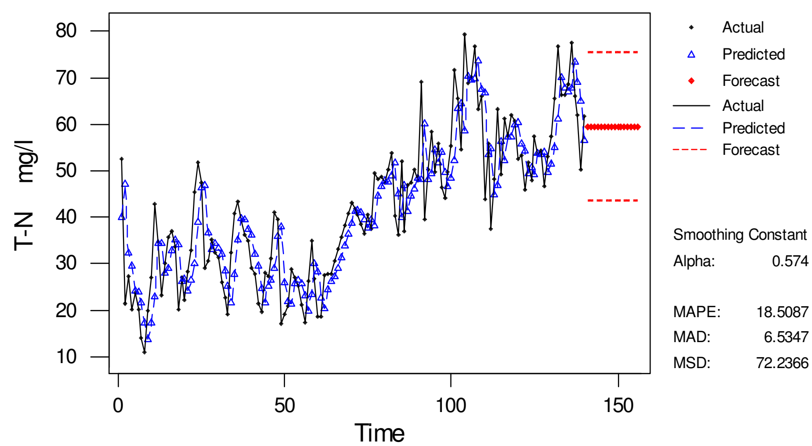


Figure (70): Single Exponential Smoothing for T-N mg/l

Table (68) shows the single exponential smoothing prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (68) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (68): Forecasted, lower, upper and actual values by single exponential smoothing for T-N variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	59.60	43.59	75.61	54.41
2	142	59.60	43.59	75.61	59.50
3	143	59.60	43.59	75.61	72.22

4	144	59.60	43.59	75.61	79.45
5	145	59.60	43.59	75.61	63.13
6	146	59.60	43.59	75.61	67.00
7	147	59.60	43.59	75.61	62.58
8	148	59.60	43.59	75.61	77.49
9	149	59.60	43.59	75.61	79.51
10	150	59.60	43.59	75.61	57.92
11	151	59.60	43.59	75.61	64.64
12	152	59.60	43.59	75.61	59.85
13	153	59.60	43.59	75.61	64.68
14	154	59.60	43.59	75.61	73.67
15	155	59.60	43.59	75.61	69.34
16	156	59.60	43.59	75.61	79.85

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 13.8%, that the simple exponential smoothing trend model did not satisfy the forecasting for the T-N variable.

B- stochastic forecasting

B1- auto regression model

Table (69) shows the AR(1) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (69) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (69): Forecasted, lower, upper and actual values by AR(1) for T-N variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	59.77	42.30	77.23	54.41
2	142	58.00	34.87	81.14	59.50
3	143	56.47	29.84	83.10	72.22
4	144	55.14	26.15	84.12	79.45
5	145	53.98	23.34	84.63	63.13
6	146	52.98	21.14	84.82	67.00
7	147	52.10	19.39	84.82	62.58
8	148	51.34	17.99	84.70	77.49
9	149	50.68	16.85	84.52	79.51

10	150	50.11	15.92	84.30	57.92
11	151	49.62	15.16	84.07	64.64
12	152	49.18	14.53	83.84	59.85
13	153	48.81	14.00	83.61	64.68
14	154	48.48	13.56	83.40	73.67
15	155	48.20	13.19	83.20	69.34
16	156	47.95	12.88	83.02	79.85

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 30.3%, that the AR(1) trend model did not satisfy the forecasting for the T-N variable.

B2- moving average regression model

The regression of the additive MA (5) trend model is shown in Figure (71).

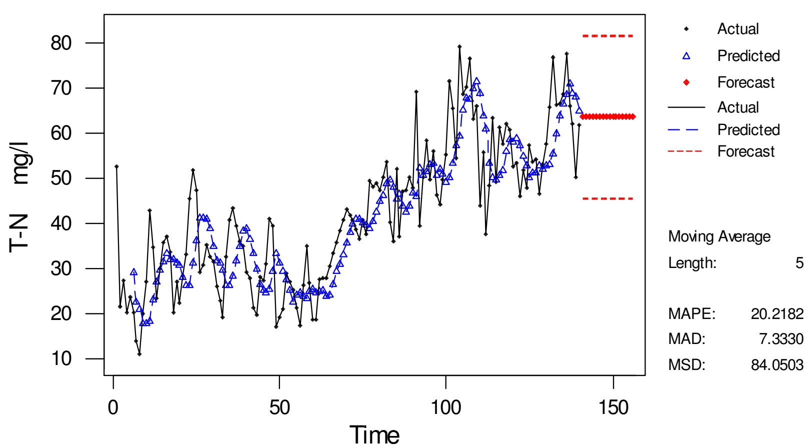


Figure (71): Moving Average for TSS mg/l

Table (70) shows the MA(5) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (70) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (70): Forecasted, lower, upper and actual values by MA(5) for T-N variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	63.58	45.61	81.55	54.41
2	142	63.58	45.61	81.55	59.50
3	143	63.58	45.61	81.55	72.22
4	144	63.58	45.61	81.55	79.45
5	145	63.58	45.61	81.55	63.13
6	146	63.58	45.61	81.55	67.00
7	147	63.58	45.61	81.55	62.58
8	148	63.58	45.61	81.55	77.49
9	149	63.58	45.61	81.55	79.51
10	150	63.58	45.61	81.55	57.92
11	151	63.58	45.61	81.55	64.64
12	152	63.58	45.61	81.55	59.85
13	153	63.58	45.61	81.55	64.68
14	154	63.58	45.61	81.55	73.67
15	155	63.58	45.61	81.55	69.34
16	156	63.58	45.61	81.55	79.85

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 6.7%, that the MA(5) trend model has satisfied the forecasting for the T-N variable.

B3- ARIMA modeling

Table (71) shows the ARIMA(1,2,5) prediction values for the next 10% of the predicted and the real data, which equals to 16 observations. In Table (71) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (71): Forecasted, lower, upper and actual values by ARIMA(1,2,5) for T-N variable

Row	Period (month)	Forecast mg/l	Lower mg/l	Upper mg/l	Actual mg/l
1	141	62.15	45.68	78.62	54.41
2	142	65.63	46.14	85.12	59.50
3	143	66.33	43.22	89.44	72.22
4	144	66.43	41.71	91.16	79.45
5	145	68.21	41.82	94.60	63.13
6	146	68.33	40.43	96.23	67.00

7	147	70.13	40.67	99.60	62.58
8	148	70.27	39.37	101.17	77.49
9	149	72.10	39.70	104.50	79.51
10	150	72.26	38.47	106.05	57.92
11	151	74.11	38.88	109.35	64.64
12	152	74.29	37.71	110.88	59.85
13	153	76.17	38.18	114.17	64.68
14	154	76.37	37.06	115.69	73.67
15	155	78.28	37.58	118.98	69.34
16	156	78.50	36.50	120.50	79.85

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 4.8%, that the ARIMA(1,2,5) trend model has satisfied the forecasting for the T-N variable.

3.5.5.7 results of prediction

The results of error are summarized in the following Table (72), which provides a summary of the models name used in the prediction and also it provides the percentage error.

Table (72) : Percentage of error of each model for T-N variable

Model	Percentage of Mean Error
Linear Method	3.3 %
Quadratic Method	7.2 %
Exponential Growth Method	3.7 %
Simple Exponential Smoothing	13.8 %
Auto Regression, AR(1)	30.3 %
Moving Average, MA(5)	6.7 %
ARIMA (1,2,5)	4.8 %

The previous Table (72) shows that the methods, which have satisfied the 10% acceptable prediction limits, are linear, quadratic, exponential growth, moving average MA(5), ARIMA (1,2,5). The best model is ARIMA(1,2,5) model.

3.5.6 Flow rate (Q) variable:

The consequences that were used to analyze the Q variable were as follows:

3.5.6.1 detection of missing data and outliers:

From table (1) it is observed that the data do not contain any missing data, so the second step is to find the outliers, data should be drawn in a scatter diagram (see Figure (72)) so that outliers will be clearly observed. These data, which contains 156 observations from January 1988 till December 2000, have approximately one outlier and it is in February 1992. It was observed that the rainfall in February 1992 was so high, and it is known that when the rainfall is high (Appendix (1)), the amount of flow rate will be high. So this data is assumed to be a real data and no adjustment will be made on it.

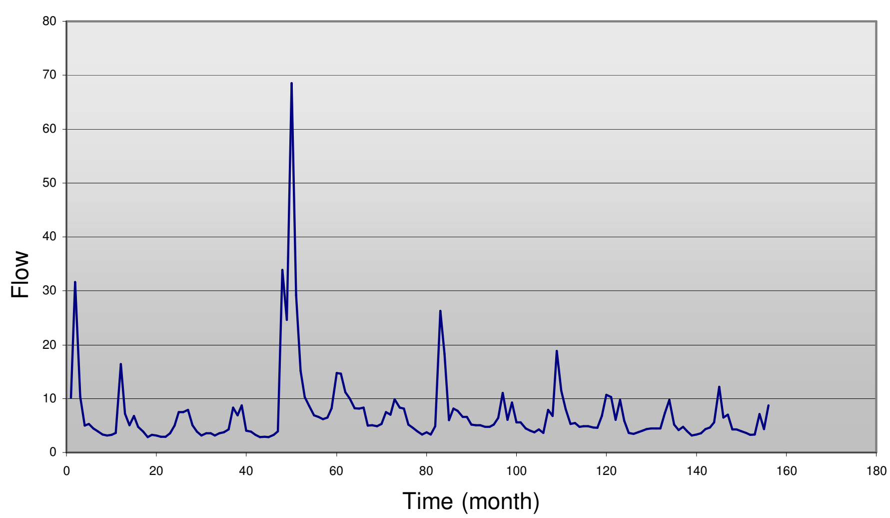


Figure (72): Origin Data of Zarqa River Flow (MCM/month)

.

Figure (73) shows the outliers for the seasonal trends for the original and the residual data, one can conclude from the charts that there are three outliers in the residual data in the seasonal condition. Also Figure (73) shows the variation in the data for the same month, it can be observe that the variation was the highest on February, and was the lowest on September. Nine outliers were found in the seasonal graph, they are in February 1988, December 1991, January 1992, February 1992, March 1992, April 1992, May 1992, June 1992, and November 1994. The amount of rainfall on each of the nine months was high (Appendix (1)), the data was treated as a real data, no adjustment was made on the data. The new adjusted data will be the same as the original data.

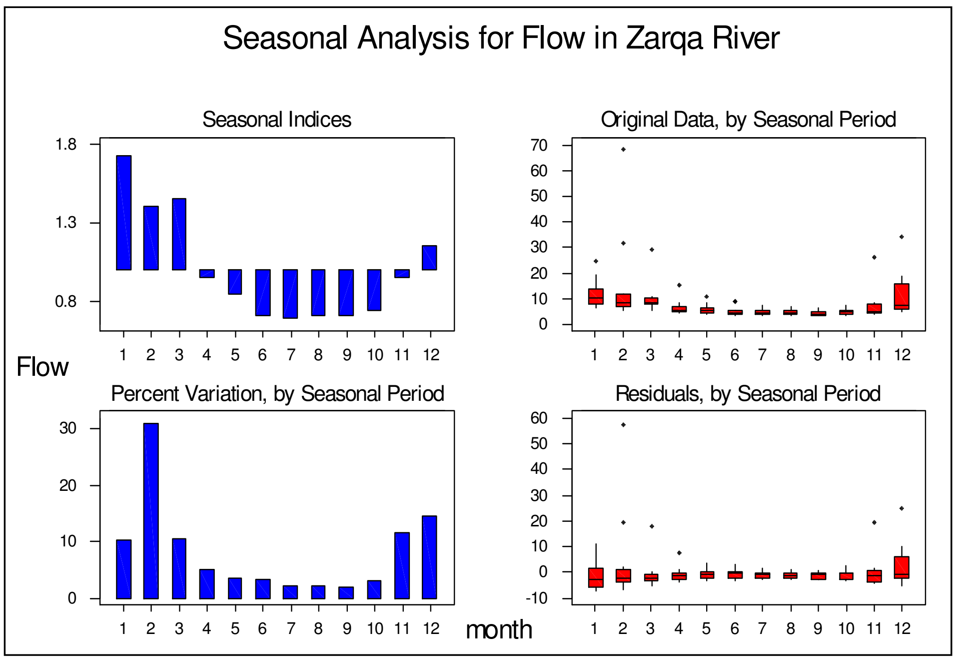


Figure (73) Outliers of Seasonal Analysis for Q Variable

3.5.6.2 normality of data

In this section, normality of data will be checked through four procedures; first one is by drawing a histogram for Weibull's distribution model, second one is through calculating the coefficient of variance, the third one is through calculating the Kurtosis coefficient, and the fourth one is through calculating the Shapiro-Wilk test. From these four procedures, if the data was not normal then a lognormal transformation to the data will be made.

A- Weibull's distribution model histogram:

Data will be transformed to the average monthly value for the Q variable; the calculated values were as follows

Month	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Flow	11.4	14.4	9.9	6.1	5.4	4.7	4.1	4.1	4.0	4.5	7.0	10.9

The Weibull's distribution histogram is drawn for these twelve data. It can be observed from Figure (74) that the data of Q has skewness to the right and bulked to the left, so the data did not satisfy Weibull's Distribution Model for normality. A lognormal test should be made to the data.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul.	Aug	Sep	Oct	Nov	Dec
Log (flow)	1.06	1.16	1.00	0.79	0.73	0.67	0.61	0.61	0.60	0.65	0.85	1.04

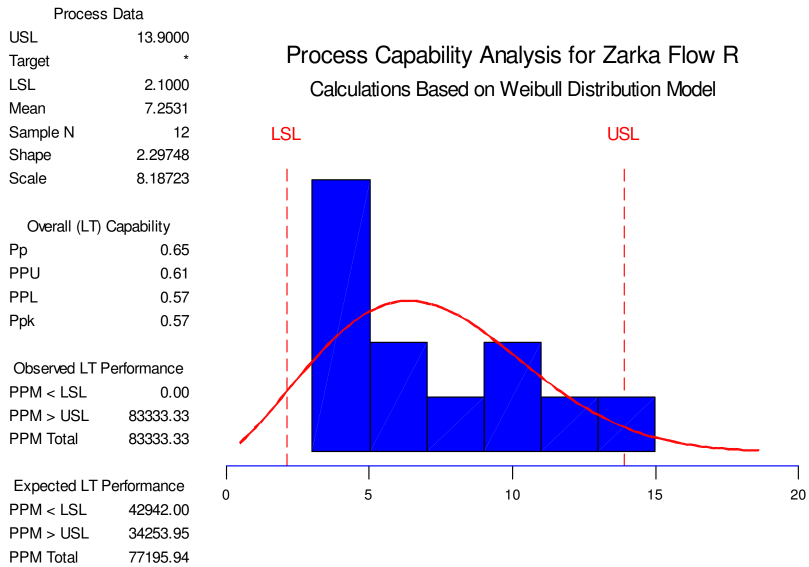


Figure (74) : Weibull Distribution Model Histogram

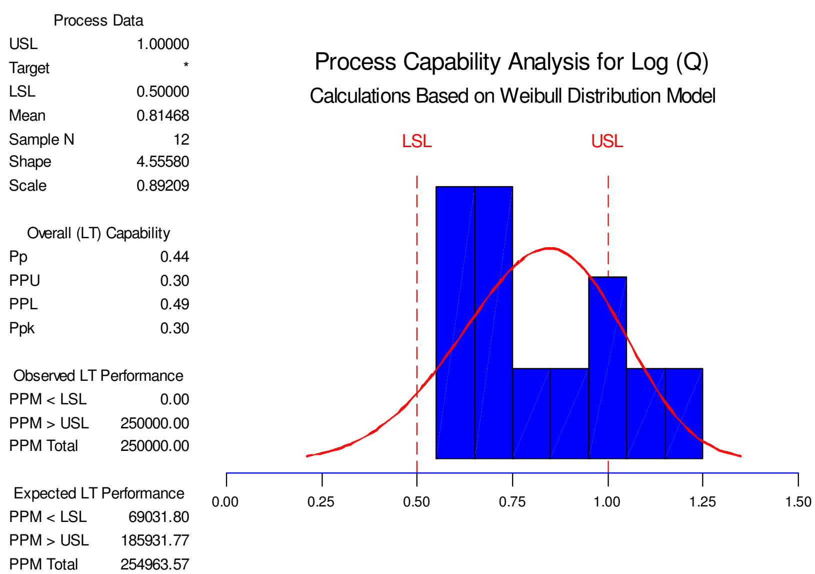


Figure (75) : Weibull Distribution Model Histogram

The Weibull's distribution histogram for the lognormal of the flow is drawn for these twelve data. It can be observed from Figure (75) that the data of log (Q) has not skewness, so the data is normal.

B- coefficient of variation (COV), preliminary test:

The data were divided into four quarters; each quarter consists of 39 data. Table (73) provides the value of the mean, variance, standard deviation, and the coefficient of variation for the Log (Q) variable.

Table (73): The coefficient of variable for Log (Q)

	MEAN	VARIANCE	ST. DE. (S)	C.O.V.
Log (Q) (1^{st} Quarter)	0.702	0.051	0.226	0.32
Log (Q) (2^{nd} Quarter)	0.877	0.095	0.309	0.35
Log (Q) (3^{rd} Quarter)	0.788	0.041	0.202	0.26
Log (Q) (4^{th} Quarter)	0.711	0.026	0.162	0.23

It can be shown from the table that the value of the coefficient of variation for all quarters were less than 1,it can be concluded that the variable Log (Q) does not have skewness.

C- Kurtosis coefficient (peakedness), vertical test:

To find the Kurtosis coefficient, one should find the value of K, which depends on the fourth moment about the mean and the number of samples, so that the Kurtosis can be calculated. The Kurtosis will give a good indication if the distribution is leptokurtic or platykutric. The data was divided into four quarters, Table (74) provides the values of the Kurtosis coefficient for each quarter and it provides also the calculations needed to calculate the Kurtosis coefficient, which they are: the mean, the variance or standard error, the value of K, and the Kurtosis coefficient.

Table (74): The Kurtosis Coefficient for Log (Q)

	MEAN	VARIANCE	ST. DE. (S)	K	Kurtosis Coeff. C'_K
Log (Q) (1st Quarter)	0.702	0.051	0.226	0.016	3.1
Log (Q) (2^{nd} Quarter)	0.877	0.095	0.309	0.042	1.6
Log (Q) (3^{rd} Quarter)	0.788	0.041	0.202	0.009	2.4
Log (Q) (4^{th} Quarter)	0.711	0.026	0.162	0.002	-0.1

From table (74) one can observe that the data in quarter one, two, and three are quietly normally distributed, the fourth quarter is normally distributed. The total data of the log(Q) variable can be assumed to be normally distributed (mesokurtic).

D- Shapiro-Wilk test

This is another test to show that the data we have is normal or not. Data that have been collected were divided into equal quarters, the value of (a_{n-1+I}), was taken for 20 data since the value of n-1+i was equal to 20, the value of (a_{n-1+I}) was taken from appendix (2). The Shapiro-Wilk value was compared with the five percent critical value for sample size 20 in Appendix (3), if the value of the Shapiro-Wilk test was greater than it then the data will not show evidence of nonnormality.

Table (75): Shapiro-Wilk Test for the Data of log (Q's) 1st quarter

No	Log (Q) MCM/ MONTH	Ordering Log (Q) (1)	Inverse order Log (Q) (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
1	1.01	0.46	1.50	1.04	0.40	0.42
2	1.50	0.47	1.21	0.74	0.28	0.21
3	1.02	0.47	1.02	0.55	0.24	0.13
4	0.70	0.50	1.01	0.51	0.21	0.11
5	0.72	0.50	0.94	0.44	0.19	0.08
6	0.65	0.51	0.92	0.42	0.17	0.07
7	0.60	0.51	0.90	0.39	0.15	0.06
8	0.52	0.52	0.88	0.36	0.14	0.05
9	0.51	0.52	0.88	0.36	0.12	0.04
10	0.52	0.52	0.86	0.34	0.11	0.04
11	0.56	0.55	0.84	0.29	0.10	0.03
12	1.21	0.55	0.83	0.28	0.08	0.02
13	0.86	0.55	0.72	0.17	0.07	0.01
14	0.71	0.56	0.71	0.15	0.06	0.01
15	0.83	0.56	0.71	0.14	0.05	0.01
16	0.68	0.58	0.70	0.12	0.04	0.01
17	0.59	0.59	0.69	0.11	0.03	0.00
18	0.46	0.59	0.68	0.10	0.02	0.00
19	0.52	0.60	0.65	0.05	0.01	0.00
20	0.50	0.63	0.63	0.00		b= 1.29
21	0.47	0.65	0.60	-0.05		S= 0.23
22	0.47	0.68	0.59	-0.10		
23	0.55	0.69	0.59	-0.11		
24	0.69	0.70	0.58	-0.12		W= 0.861 < 0.939
25	0.88	0.71	0.56	-0.14		
26	0.88	0.71	0.56	-0.15		Did not satisfy
27	0.90	0.72	0.55	-0.17		
28	0.71	0.83	0.55	-0.28		
29	0.59	0.84	0.55	-0.29		
30	0.50	0.86	0.52	-0.34		
31	0.56	0.88	0.52	-0.36		
32	0.55	0.88	0.52	-0.36		
33	0.51	0.90	0.51	-0.39		
34	0.55	0.92	0.51	-0.42		
35	0.58	0.94	0.50	-0.44		
36	0.63	1.01	0.50	-0.51		
37	0.92	1.02	0.47	-0.55		
38	0.84	1.21	0.47	-0.74		
39	0.94	1.50	0.46	-1.04		

Table (76): Shapiro-Wilk Test for the Data of log (Q's) 2nd quarter

No	Log (Q) MCM/ MONTH	Ordering Log (Q) (1)	Inverse order Log (Q) (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
40	0.61	0.45	1.84	1.38	0.40	0.55
41	0.59	0.46	1.53	1.07	0.28	0.30
42	0.51	0.46	1.47	1.00	0.24	0.24
43	0.46	0.51	1.39	0.88	0.21	0.19
44	0.46	0.51	1.18	0.67	0.19	0.13
45	0.45	0.59	1.17	0.58	0.17	0.10
46	0.51	0.59	1.17	0.57	0.15	0.09
47	0.59	0.60	1.05	0.44	0.14	0.06
48	1.53	0.61	1.01	0.40	0.12	0.05
49	1.39	0.66	1.00	0.34	0.11	0.04
50	1.84	0.69	0.99	0.30	0.10	0.03
51	1.47	0.70	0.94	0.24	0.08	0.02
52	1.18	0.71	0.92	0.21	0.07	0.02
53	1.01	0.72	0.92	0.21	0.06	0.01
54	0.94	0.73	0.91	0.19	0.05	0.01
55	0.84	0.79	0.91	0.12	0.04	0.00
56	0.82	0.82	0.91	0.10	0.03	0.00
57	0.79	0.82	0.91	0.09	0.02	0.00
58	0.82	0.84	0.88	0.04	0.01	0.00
59	0.91	0.84	0.84	0.00		
60	1.17	0.88	0.84	-0.04		
61	1.17	0.91	0.82	-0.09		
62	1.05	0.91	0.82	-0.10		
63	1.00	0.91	0.79	-0.12		
64	0.91	0.91	0.73	-0.19		
65	0.91	0.92	0.72	-0.21		
66	0.92	0.92	0.71	-0.21		
67	0.70	0.94	0.70	-0.24		
68	0.71	0.99	0.69	-0.30		
69	0.69	1.00	0.66	-0.34		
70	0.73	1.01	0.61	-0.40		
71	0.88	1.05	0.60	-0.44		
72	0.84	1.17	0.59	-0.57		
73	0.99	1.17	0.59	-0.58		
74	0.92	1.18	0.51	-0.67		
75	0.91	1.39	0.51	-0.88		
76	0.72	1.47	0.46	-1.00		
77	0.66	1.53	0.46	-1.07		
78	0.60	1.84	0.45	-1.38		

b= 1.83

S= 0.31

W=0.921 < 0.939

Did not satisfy

Table (77): Shapiro-Wilk Test for the Data of log (Q's) 3rd quarter

No	Log (Q) MCM/ MONTH	Ordering Log (Q) (1)	Inverse order Log (Q) (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
79	0.52	0.52	1.42	0.89	0.40	0.36
80	0.58	0.53	1.28	0.74	0.28	0.21
81	0.53	0.57	1.26	0.70	0.24	0.17
82	0.69	0.58	1.06	0.48	0.21	0.10
83	1.42	0.58	1.04	0.46	0.19	0.09
84	1.26	0.61	0.97	0.36	0.17	0.06
85	0.78	0.63	0.91	0.28	0.15	0.04
86	0.91	0.65	0.91	0.25	0.14	0.03
87	0.89	0.67	0.90	0.23	0.12	0.03
88	0.82	0.68	0.89	0.21	0.11	0.02
89	0.82	0.68	0.84	0.16	0.10	0.02
90	0.72	0.68	0.82	0.14	0.08	0.01
91	0.71	0.68	0.82	0.13	0.07	0.01
92	0.71	0.69	0.81	0.12	0.06	0.01
93	0.68	0.69	0.78	0.10	0.05	0.00
94	0.68	0.71	0.78	0.07	0.04	0.00
95	0.71	0.71	0.75	0.04	0.03	0.00
96	0.81	0.71	0.75	0.03	0.02	0.00
97	1.04	0.72	0.74	0.02	0.01	0.00
98	0.78	0.72	0.72	0.00		b= 1.16
99	0.97	0.74	0.72	-0.02		S= 0.2
100	0.75	0.75	0.71	-0.03		
101	0.75	0.75	0.71	-0.04		
102	0.65	0.78	0.71	-0.07		W=0.866 < 0.939
103	0.61	0.78	0.69	-0.10		
104	0.58	0.81	0.69	-0.12		Did not satisfy
105	0.63	0.82	0.68	-0.13		
106	0.57	0.82	0.68	-0.14		
107	0.90	0.84	0.68	-0.16		
108	0.84	0.89	0.68	-0.21		
109	1.28	0.90	0.67	-0.23		
110	1.06	0.91	0.65	-0.25		
111	0.91	0.91	0.63	-0.28		
112	0.72	0.97	0.61	-0.36		
113	0.74	1.04	0.58	-0.46		
114	0.68	1.06	0.58	-0.48		
115	0.69	1.26	0.57	-0.70		
116	0.68	1.28	0.53	-0.74		
117	0.67	1.42	0.52	-0.89		

Table (78): Shapiro-Wilk Test for the Data of log (Q's) 4th quarter

No	Log (Q) MCM/ MONTH	Ordering Log (Q) (1)	Inverse order Log (Q) (2)	2-1	a(n-1+i)	(2-1) x $a_{(n-1+i)}$
118	0.66	0.50	1.09	0.59	0.40	0.24
119	0.84	0.52	1.03	0.51	0.28	0.14
120	1.03	0.52	1.01	0.49	0.24	0.12
121	1.01	0.53	0.99	0.46	0.21	0.10
122	0.78	0.54	0.99	0.45	0.19	0.08
123	0.99	0.55	0.94	0.40	0.17	0.07
124	0.77	0.56	0.86	0.30	0.15	0.05
125	0.56	0.56	0.85	0.29	0.14	0.04
126	0.54	0.58	0.85	0.27	0.12	0.03
127	0.58	0.60	0.84	0.24	0.11	0.03
128	0.61	0.60	0.82	0.22	0.10	0.02
129	0.64	0.61	0.78	0.17	0.08	0.01
130	0.65	0.62	0.77	0.15	0.07	0.01
131	0.65	0.63	0.74	0.12	0.06	0.01
132	0.65	0.63	0.72	0.08	0.05	0.00
133	0.86	0.64	0.68	0.04	0.04	0.00
134	0.99	0.64	0.67	0.03	0.03	0.00
135	0.72	0.64	0.66	0.02	0.02	0.00
136	0.62	0.65	0.65	0.01	0.01	0.00
137	0.68	0.65	0.65	0.00		b= 0.95
138	0.60	0.65	0.65	-0.01		S= 0.16
139	0.50	0.66	0.64	-0.02		
140	0.52	0.67	0.64	-0.03		
141	0.55	0.68	0.64	-0.04		
142	0.64	0.72	0.63	-0.08		
143	0.67	0.74	0.63	-0.12		
144	0.74	0.77	0.62	-0.15		
145	1.09	0.78	0.61	-0.17		
146	0.82	0.82	0.60	-0.22		
147	0.85	0.84	0.60	-0.24		
148	0.63	0.85	0.58	-0.27		
149	0.63	0.85	0.56	-0.29		
150	0.60	0.86	0.56	-0.30		
151	0.56	0.94	0.55	-0.40		
152	0.52	0.99	0.54	-0.45		
153	0.53	0.99	0.53	-0.46		
154	0.85	1.01	0.52	-0.49		
155	0.64	1.03	0.52	-0.51		
156	0.94	1.09	0.50	-0.59		

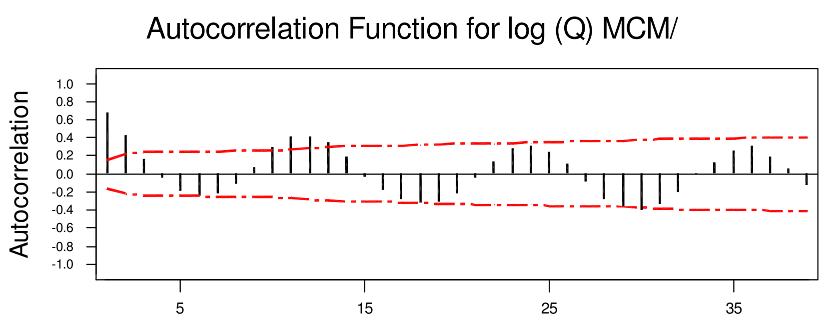
$W=0.902 < 0.939$

Did not satisfy

From the Tables (75), (76), (77), and (78) it has been shown that the data in each quarter did not satisfy Shapiro-Wilk test for normality. So this test did not give any indication about the whole data if it is normal or not.

3.5.6.3 order of (AR)

For water quality like King Talal Dam, the value of AR, which is expressed by the item (p) shall not be more than 1 since the autocorrelation for a particle of log (Q) does not need more than 1 month till it analyze (Viessman and Lewis, 1996). From Figure (76) it can be seen that there is a high correlation for the first autocorrelation, the value of p that will be used is 1 for the Q variable.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.68	8.48	73.22	13	0.35	2.34	230.14	25	0.25	1.39	350.16	37	0.19	0.93	501.40
2	0.43	3.88	102.87	14	0.19	1.20	236.21	26	0.11	0.59	352.32	38	0.07	0.32	502.31
3	0.16	1.34	107.15	15	-0.03	-0.17	236.33	27	-0.09	-0.48	353.74	39	-0.12	-0.60	505.49
4	-0.04	-0.33	107.41	16	-0.18	-1.16	242.18	28	-0.29	-1.57	369.40				
5	-0.19	-1.53	113.14	17	-0.28	-1.78	256.33	29	-0.37	-2.01	396.07				
6	-0.24	-1.97	122.99	18	-0.32	-1.99	274.95	30	-0.40	-2.10	426.98				
7	-0.22	-1.73	131.03	19	-0.32	-1.90	292.90	31	-0.34	-1.75	449.79				
8	-0.12	-0.89	133.26	20	-0.22	-1.27	301.39	32	-0.20	-1.02	457.94				
9	0.08	0.59	134.27	21	-0.05	-0.29	301.85	33	-0.01	-0.06	457.97				
10	0.30	2.30	149.58	22	0.14	0.79	305.25	34	0.13	0.65	461.37				
11	0.41	3.06	178.71	23	0.28	1.63	319.93	35	0.25	1.26	474.31				
12	0.42	2.92	208.61	24	0.32	1.80	338.51	36	0.31	1.53	493.93				

Figure (76) Autocorrelation Function for Q Variable

3.5.6.4 order of moving average (MA)

After finding the value of AR, which was 1, the following procedure is to determine the value of MA, which is expressed by the item (q). Figure (77) shows the change between the real data of the variable log (Q) and it's moving average with different lengths of p.

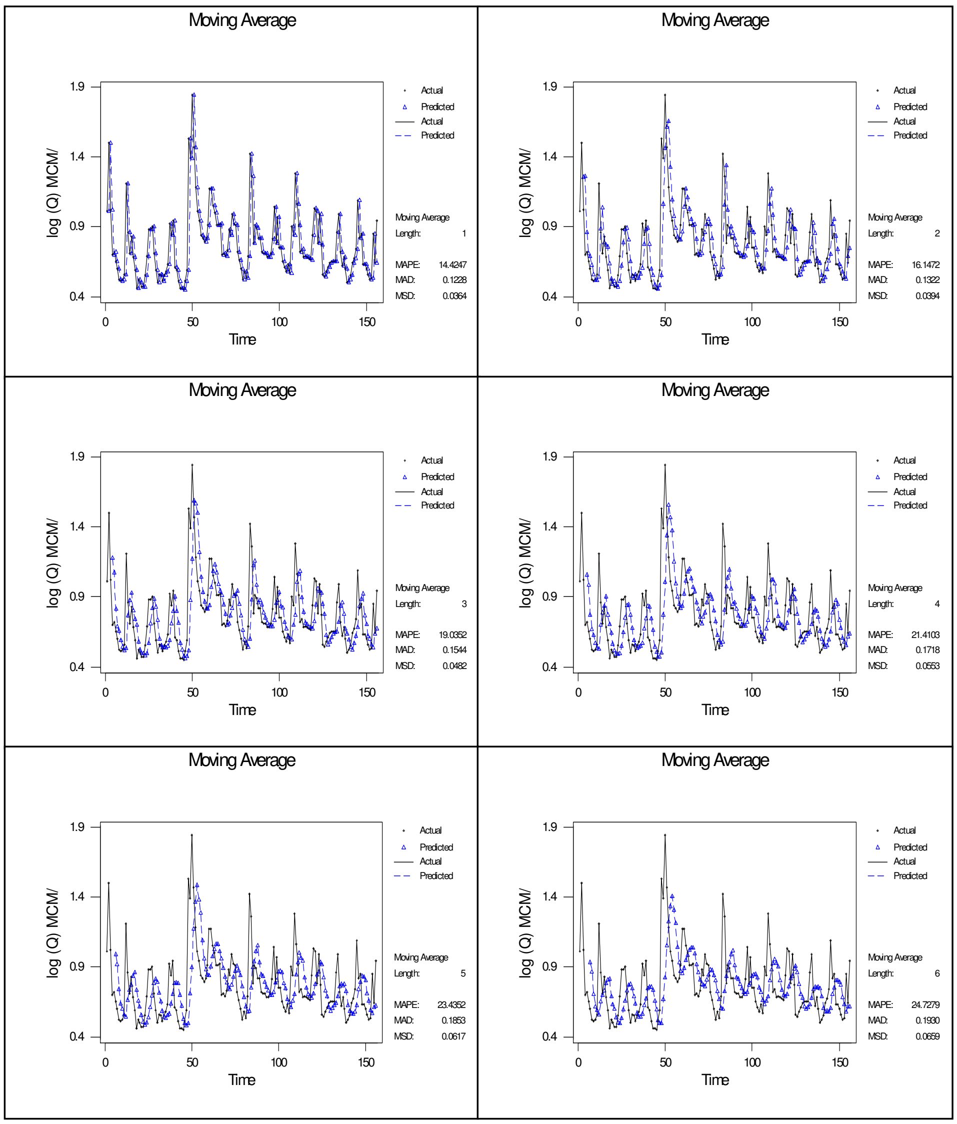
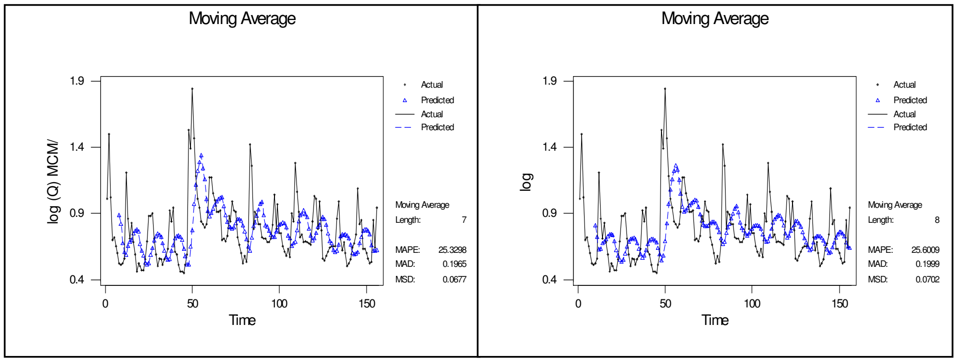


Figure (77) Moving Average of Log (Q) with Different Values of (p)



Cont. Figure (77) Moving Average of log (Q) with Different Values of (p)

The moving average can be determined from Figure (77) when the difference between the previous length of p and the followed one have a small difference and that occurred when the value of p was 6 (as shown in Figure (77)), so the log (Q) variable has a value of MA(6).

3.5.6.5 order of (I)

The last coefficient of ARIMA's parameters is the integrated model (I), which expressed by the item (d). The data should be differenced when there is trend or shift or seasonality in the data, otherwise there is no need to make differentiation for the data. Figure (78) consists of four graphs, which provides a good idea if there is a difference between the original, detrended, seasonally adjusted, and seasonally adjusted and detrended data. It is shown from these four graphs that there are a difference between the original figure and the seasonally one but in the detrended case they are almost the same, which means that the seasonally effect should be taken into consideration.

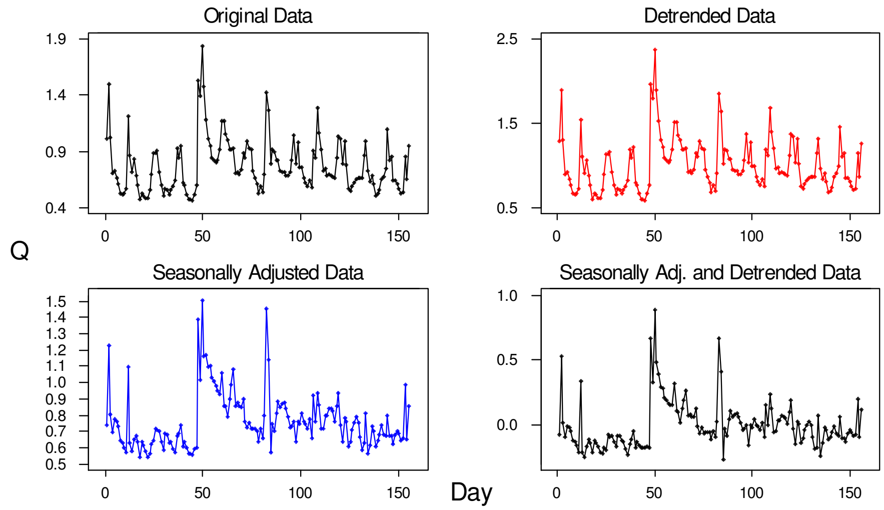


Figure (78): Component Analysis for log (Q) MCM/month

Two season; summer and winter can affect seasonality in Jordan, so if the data has no seasonality effect, then the value of d=0 and if we have seasonality effect then the value of d=2. Figures (79), and (80) provide ARIMA model diagnostics for ARIMA = (1,0,6) and (1,2,6). It is seen from the two graphs that the residual in Figure (79) is quite the same as in Figure (80) so the coefficients of ARIMA that will be used are (1,2,6).

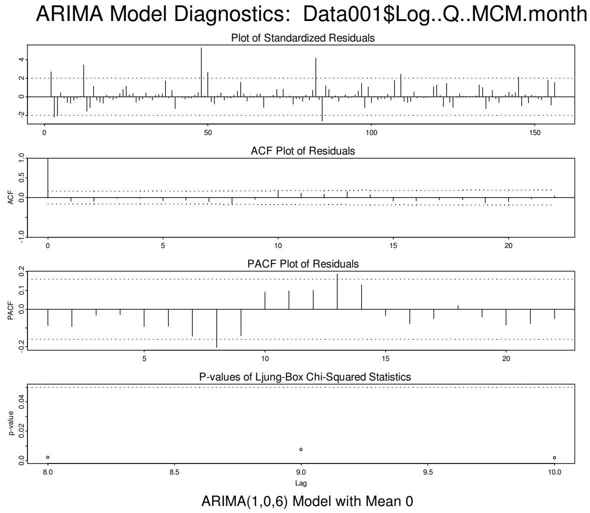


Figure (79): ARIMA (1,0,6) Model Diagnostic for log(Q)

Comment: Figure (8)

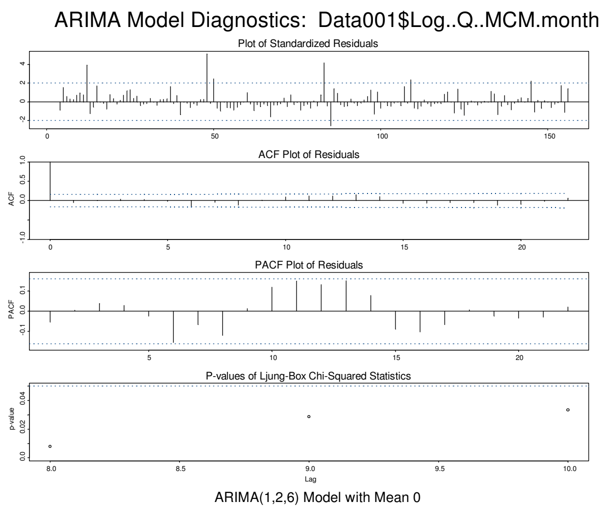


Figure (80): ARIMA (1,2,6) Model Diagnostic for log(Q)

3.5.6.6 forecasting future values

The following procedure will be used in the forecasting: The values of the data collected will be divided into two parts, the first part consists of 90% of the real data, and this data will be analyzed and predicted. And the second part consists of the last

10% of the real data, and this part will be compared with the predicted values in the mean. The best model is the one that gives the least error in mean.

A- deterministic forecasting

A1- linear regression model

The regression of the additive linear trend is shown in Figure (81).

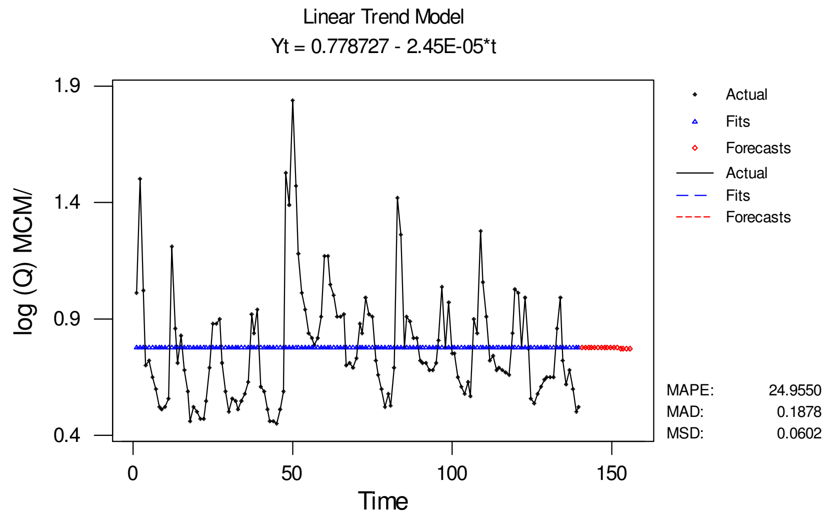


Figure (81):Trend Analysis for log (Q) MCM/month

It can be observed from the above figure and equation of the linear trend that the data is decreasing very slowly. Table (79) shows the linear prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (79): The values of the predicted and actual data by linear regression for log (Q) variable

Row	Period (months)	Forecasted (MCM)	Actual (MCM)
1	141	0.7753	0.55
2	142	0.7752	0.64
3	143	0.7752	0.67
4	144	0.7752	0.74
5	145	0.7752	1.09
6	146	0.7752	0.82
7	147	0.7751	0.85

8	148	0.7751	0.63
9	149	0.7751	0.63
10	150	0.7751	0.60
11	151	0.7750	0.56
12	152	0.7750	0.52
13	153	0.7750	0.53
14	154	0.7750	0.85
15	155	0.7749	0.64
16	156	0.7749	0.94

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 9.2%, that the linear trend model has satisfied the forecasting for log (Q) variable.

A2- quadratic regression model

The regression of the additive quadratic trend is shown in Figure (82).

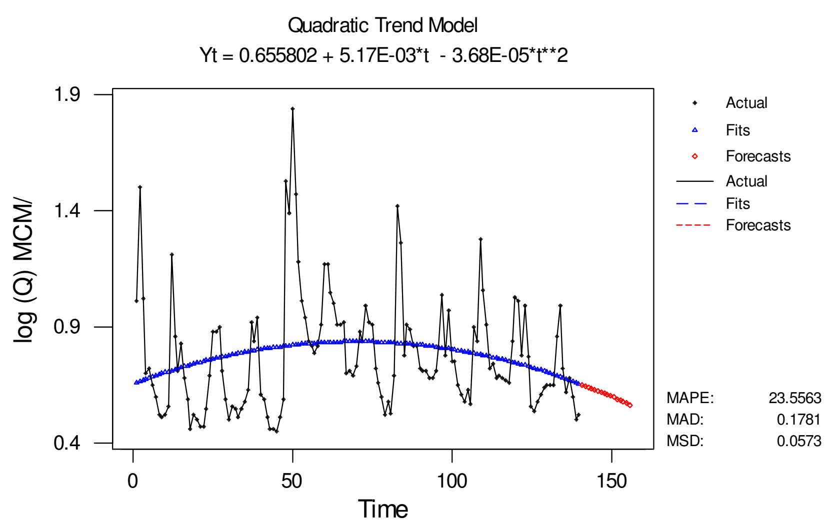


Figure (82): Trend Analysis for log (Q) MCM/month

It can be observed from Figure (82) and the equation of the quadratic trend that the data is increasing upward and then it is decreasing. Table (80) shows the quadratic prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (80): The values of the predicted and actual data by quadratic regression for log (Q) variable

Row	Period (months)	Forecasted (MCM)	Actual (MCM)
1	141	0.6523	0.55
2	142	0.6471	0.64
3	143	0.6418	0.67
4	144	0.6364	0.74
5	145	0.6309	1.09
6	146	0.6253	0.82
7	147	0.6197	0.85
8	148	0.6140	0.63
9	149	0.6082	0.63
10	150	0.6024	0.60
11	151	0.5965	0.56
12	152	0.5905	0.52
13	153	0.5844	0.53
14	154	0.5783	0.85
15	155	0.5721	0.64
16	156	0.5658	0.94

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 15.3%, that the quadratic trend model did not satisfy the forecasting for log (Q) variable.

A3- exponential growth regression model

The regression of the additive exponential growth trend model is shown in Figure (83).

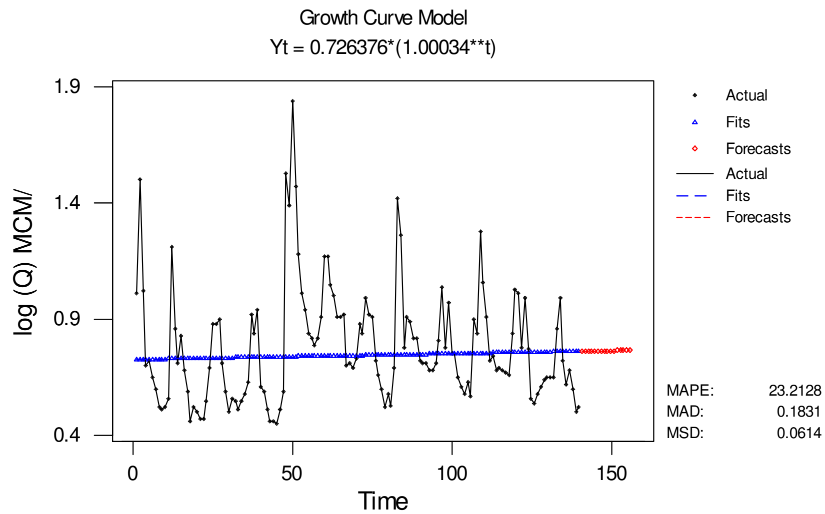


Figure (83): Trend Analysis for log (Q) MCM/month

It can be observed from the above figure and equation of the exponential growth trend that the data has an increasing (very slowly) trend. Table (81) shows the exponential growth prediction of the next 10% of the predicted and real data, which equals to 16 observations.

Table (81): The values of the predicted and actual data by exponential growth regression for log (Q) variable

Row	Period (months)	Forecasted (MCM)	Actual (MCM)
1	141	0.7620	0.55
2	142	0.7623	0.64
3	143	0.7625	0.67
4	144	0.7628	0.74
5	145	0.7630	1.09
6	146	0.7633	0.82
7	147	0.7636	0.85
8	148	0.7638	0.63
9	149	0.7641	0.63
10	150	0.7643	0.60
11	151	0.7646	0.56
12	152	0.7649	0.52
13	153	0.7651	0.53
14	154	0.7654	0.85
15	155	0.7656	0.64
16	156	0.7659	0.94

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 7.9%, that the exponential growth trend model has satisfied the forecasting for the log (Q) variable.

A4- single exponential smoothing model

The regression of the additive single exponential smoothing trend model is shown in Figure (84).

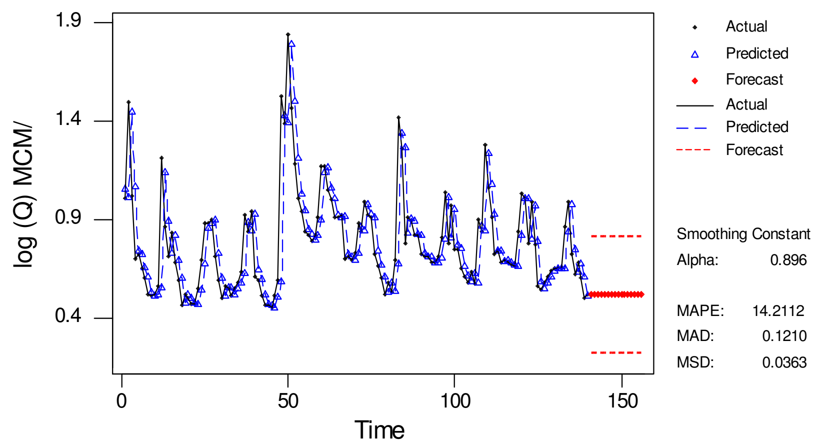


Figure (84): Single Exponential Smoothing for log (Q) MCM/month

Table (82) shows the single exponential smoothing prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (82) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (82): Forecasted, lower, upper and actual values by single exponential smoothing for log (Q) variable

Row	Period (month)	Forecast (MCM)	Lower (MCM)	Upper (MCM)	Actual (MCM)
1	141	0.5191	0.2226	0.8156	0.55
2	142	0.5191	0.2226	0.8156	0.64
3	143	0.5191	0.2226	0.8156	0.67

4	144	0.5191	0.2226	0.8156	0.74
5	145	0.5191	0.2226	0.8156	1.09
6	146	0.5191	0.2226	0.8156	0.82
7	147	0.5191	0.2226	0.8156	0.85
8	148	0.5191	0.2226	0.8156	0.63
9	149	0.5191	0.2226	0.8156	0.63
10	150	0.5191	0.2226	0.8156	0.60
11	151	0.5191	0.2226	0.8156	0.56
12	152	0.5191	0.2226	0.8156	0.52
13	153	0.5191	0.2226	0.8156	0.53
14	154	0.5191	0.2226	0.8156	0.85
15	155	0.5191	0.2226	0.8156	0.64
16	156	0.5191	0.2226	0.8156	0.94

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 35.6%, that the simple exponential smoothing trend model did not satisfy the forecasting for the Q variable.

B- stochastic forecasting

B1- auto regression model

Table (83) shows the AR(1) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (83) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (83): Forecasted, lower, upper and actual values by AR(1) for Q variable

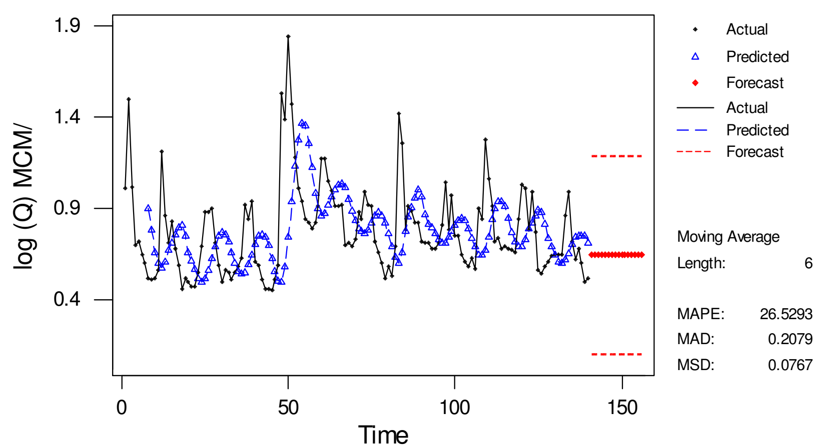
Row	Period (month)	Forecast (MCM)	Lower (MCM)	Upper (MCM)	Actual (MCM)
1	141	0.5975	0.2487	0.9463	0.55
2	142	0.6516	0.2262	1.0770	0.64
3	143	0.6894	0.2313	1.1474	0.67
4	144	0.7157	0.2425	1.1889	0.74
5	145	0.7341	0.2537	1.2145	1.09
6	146	0.7470	0.2631	1.2308	0.82
7	147	0.7559	0.2704	1.2414	0.85

8	148	0.7622	0.2758	1.2485	0.63
9	149	0.7665	0.2798	1.2533	0.63
10	150	0.7696	0.2827	1.2565	0.60
11	151	0.7717	0.2847	1.2587	0.56
12	152	0.7732	0.2861	1.2602	0.52
13	153	0.7742	0.2872	1.2613	0.53
14	154	0.7750	0.2879	1.2620	0.85
15	155	0.7755	0.2884	1.2626	0.64
16	156	0.7758	0.2887	1.2629	0.94

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 4.9%, that the AR(1) trend model has satisfied the forecasting for the Q variable.

B2- moving average regression model

The regression of the additive MA (6) trend model is shown in Figure (85).



Figure(85): Moving Average for log (Q) MCM/month

Table (84) shows the MA(6) prediction for the next 10% of the predicted and real data, which equals to 16 observations. In Table (84) it can be seen the upper and lower values, the forecasted value was the average between the upper and lower values.

Table (84): Forecasted, lower, upper and actual values by MA(6) for log (Q) variable

Row	Period (month)	Forecast (MCM)	Lower (MCM)	Upper (MCM)	Actual (MCM)
1	141	0.6458	0.1031	1.1885	0.55
2	142	0.6458	0.1031	1.1885	0.64
3	143	0.6458	0.1031	1.1885	0.67
4	144	0.6458	0.1031	1.1885	0.74
5	145	0.6458	0.1031	1.1885	1.09
6	146	0.6458	0.1031	1.1885	0.82
7	147	0.6458	0.1031	1.1885	0.85
8	148	0.6458	0.1031	1.1885	0.63
9	149	0.6458	0.1031	1.1885	0.63
10	150	0.6458	0.1031	1.1885	0.60
11	151	0.6458	0.1031	1.1885	0.56
12	152	0.6458	0.1031	1.1885	0.52
13	153	0.6458	0.1031	1.1885	0.53
14	154	0.6458	0.1031	1.1885	0.85
15	155	0.6458	0.1031	1.1885	0.64
16	156	0.6458	0.1031	1.1885	0.94

Comparing the actual values with the predicted ones, one can conclude, after calculating the predication error, which equals to 9.0%, that the MA(6) trend model has satisfied the forecasting for the Q variable.

B3- ARIMA modeling

The software (Minitab 13) cannot estimate the value of ARIMA (1,2,6), because each of ARIMA parameter should not exceed 5.

3.5.6.7 results of prediction

The results of error are summarized in the following Table (85), which provides a summary of the models name used in the prediction and also it provides the percentage error.

Table (85) : Percentage of error of each model for log (Q) variable

Model	Percentage of Mean Error
Linear Method	9.2 %
Quadratic Method	15.3 %
Exponential Growth Method	7.9 %
Simple Exponential Smoothing	35.6 %
Auto Regression, AR(1)	4.9 %
Moving Average, MA(6)	9.0 %
ARIMA (1,2,6)	--

The previous Table (85) shows that the methods, which have satisfied the 10% acceptable prediction limits, are linear, exponential growth, Auto regression AR(1), and moving average MA(5). The best model is AR(1) model with a least error of 4.9%.

3.6 Cross and Distance Correlation

In this section, the relations between Zarka's River variables, which is expressed in term of cross variable and then the relation between each variable in Zarka River and Samra's effluent, the figures of each relation is shown in Appendix (4).

3.6.1 Cross correlation in Zarka River variables:

The cross correlation procedure will be implemented between all the six variables; Flow, TSS, BOD_5, COD, T-P, and T-N, in Zarka River.

3.6.1.1 cross correlation between Zarka River flow in MCM/month and TSS in mg/l

The relation between flow and TSS has a cyclic shape. From the figure in Appendix4 (A4.1) it can be seen that the relation between the flow and TSS in Zarka River varies every about six months.

3.6.1.2 cross correlation between Zarka River flow in MCM/month and BOD_5 in mg/l

The relation between flow and BOD_5 has a cyclic shape. From the figure in Appendix4 (A4.2) it can be seen that the relation between the flow and BOD_5 in Zarka River is negative in most of the times. But there are some months where the relation between them is positive.

3.6.1.3 cross correlation between Zarka River flow in MCM/month and COD in mg/l

The relation between flow and COD has a cyclic shape. From the figure in Appendix4 (A4.3) it can be seen that the relation between the flow and COD in Zarka River is negative in most of the times. But there are some months where the relation between them is positive.

3.6.1.4 cross correlation between Zarka River flow in MCM/month and T-P in mg/l

The relation between flow and T-P has a cyclic shape. From the figure in Appendix4 (A4.4) it can be seen that the relation between the flow and T-P in Zarka River varies every about six months. Also it can be seen that the relation is negative in most of the times.

3.6.1.5 cross correlation between Zarka River flow in MCM/month and T-N in mg/l

The relation between flow and T-N has a cyclic shape. From the figure in Appendix4 (A4.5) it can be seen that the relation between the flow and T-N in Zarka River varies every about six months. Also it can be seen that the relation is negative in most of the times.

3.6.1.6 cross correlation between Zarka River TSS in mg/l and BOD_5 in mg/l

The relation between TSS and BOD_5 has a cyclic shape. From the figure in Appendix4 (A4.6) it can be seen that the relation between the TSS and BOD_5 in Zarka

River varies every about six months. Also it can be seen that the relation is positive in most of the times.

3.6.1.7 cross correlation between Zarka River TSS in mg/l and COD in mg/l

The relation between TSS and COD has a cyclic shape. From the figure in Appendix4 (A4.7) it can be seen that the relation between the TSS and COD in Zarka River varies every about six months. Also it can be seen that the relation is positive in most of the times.

3.6.1.8 cross correlation between Zarka River TSS in mg/l and T-P in mg/l

The relation between TSS and T-P has a cyclic shape. From the figure in Appendix4 (A4.9) it can be seen that the relation between the TSS and T-P in Zarka River varies every about six months. Also it can be seen that the correlation in the positive direction is more than in the negative direction.

3.6.1.9 cross correlation between Zarka River TSS in mg/l and T-N in mg/l

It can be seen from the figure in Appendix4 (A4.9) that when TSS increases the amount of T-N increases. So the relation between the TSS and T-N is increasing in most of the times.

3.6.1.10 cross correlation between Zarka River BOD_5 in mg/l and COD in mg/l

It can be seen from the figure in Appendix4 (A4.10) that when BOD_5 increases the amount of COD increases, and that the correlation between them is high. So the relation between the BOD_5 and COD is an increasing regression.

3.6.1.11 cross correlation between Zarka's River BOD_5 in mg/l and T-P in mg/l

It can be seen from the figure in Appendix4 (A4.11) that when BOD_5 increases the amount of T-P increases, and that the correlation between them is high. So the relation between the BOD_5 and T-P is an increasing regression.

3.6.1.12 cross correlation between Zarka River BOD_5 in mg/l and T-N in mg/l

It can be seen from the figure in Appendix4 (A4.12) that when BOD_5 increases the amount of T-N increases, and that the correlation between them is high. So the relation between the BOD_5 and T-N is an increasing regression.

3.6.1.13 cross correlation between Zarka River COD in mg/l and T-P in mg/l

It can be seen from the figure in Appendix4 (A4.13) that when COD increases the amount of T-P increases, and that the correlation between them is high. So the relation between the COD and T-P is an increasing regression.

3.6.1.14 cross correlation between Zarka River COD in mg/l and T-N in mg/l

It can be seen from the figure in Appendix4 (A4.14) that when COD increases the amount of T-N increases, and that the correlation between them is high. So the relation between the COD and T-N is an increasing regression.

Figure (86): Example on the Cross Correlation Function: BOD5 mg/l; COD mg/l

```
CCF - correlates BOD5    mg/l(t) and COD    mg/l(t+k)

          -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
            +----+----+----+----+----+----+----+----+----+----+
-22  0.086                        XXX
-21  0.046                        XX
-20  0.044                        XX
-19  0.060                        XXX
-18  0.099                        XXX
-17  0.053                        XX
-16  0.047                        XX
-15  0.041                        XX
-14  0.028                        XX
-13  0.117                        XXXX
-12  0.177                        XXXXX
-11  0.233                        XXXXXXX
-10  0.226                        XXXXXXX
 -9  0.194                        XXXXXX
 -8  0.141                        XXXXX
 -7  0.153                        XXXXX
 -6  0.209                        XXXXXX
 -5  0.172                        XXXXX
 -4  0.085                        XXX
 -3  0.171                        XXXXX
 -2  0.166                        XXXXX
 -1  0.251                        XXXXXXX
  0  0.453                        XXXXXXXXXXXX
  1  0.403                        XXXXXXXXXXX
  2  0.351                        XXXXXXXXXX
  3  0.271                        XXXXXXXX
  4  0.285                        XXXXXXXX
  5  0.279                        XXXXXXXX
  6  0.264                        XXXXXXXX
  7  0.296                        XXXXXXXX
  8  0.262                        XXXXXXXX
  9  0.146                        XXXXX
 10  0.226                        XXXXXXX
 11  0.203                        XXXXXX
 12  0.319                        XXXXXXXXX
 13  0.320                        XXXXXXXXX
 14  0.210                        XXXXXX
 15  0.262                        XXXXXXXX
 16  0.236                        XXXXXXX
 17  0.162                        XXXXX
 18  0.265                        XXXXXXXX
 19  0.203                        XXXXXX
 20  0.180                        XXXXXX
 21  0.194                        XXXXXX
 22  0.178                        XXXXX
```

3.6.1.15 cross correlation between Zarka River T-P in mg/l and T-N in mg/l

It can be seen from the figure in Appendix4 (A4.15) that when T-P increases the amount of T-N increases, and that the correlation between them is so high. So the relation between the T-P and T-N is an increasing regression.

3.6.2 Distance cross correlation between Zarka River and Samra's effluent

The distance cross correlation procedure will be implemented between each variable; Flow, TSS, BOD_5, COD, T-P, and T-N, in Zarka River and its arbitrary in Samra's effluent.

3.6.2.1 distance cross correlation function: Zarka Flow MCM/month; Samra flow MCM/month

It can be seen from the figure in Appendix4 (A4.16) that the relation between the Zarka River flow and Samra's effluent is decreasing regression.

3.6.2.2 distance cross correlation function: Zarka TSS mg/l; TSS Samra mg/l

The relation between TSS in Zarka River and TSS in Samra's effluent has a cyclic shape. From the figure in Appendix4 (A4.17) it can be seen that the relation between the TSS in Zarka and Samra varies every about six months. Also it can be seen that the relation is negative in most of the times.

3.6.2.3 distance cross correlation function: Zarka BOD_5 mg/l; BOD_5 Samra mg/l

The relation between BOD_5 in Zarka River and Samra's effluent is positively in most of the times, see Appendix4 (A4.18).

3.6.2.4 distance cross correlation function: Zarka COD mg/l; COD Samra mg/l

The relation between COD in Zarka River and Samra's effluent is positively in most of the times see Appendix4 (A4.19).

3.6.2.5 distance cross correlation function: Zarka T-P mg/l; T-P Samra mg/l

The relation between T-P in Zarka River and Samra's effluent is positively in most of the times, but the correlation is not very high between these two variables see Appendix4 (A4.20).

3.6.2.6 distance cross correlation function: Zarka T-N mg/l; T-N Samra mg/l

The relation between T-N in Zarka River and Samra's effluent is positively in most of the times, but the correlation is not very high between these two variables see Appendix4 (A4.21).

Figure (87): Example on the Distance Correlation Function: BOD5 in Zarka River mg/l; BOD5 Samra mg/l

```
CCF - correlates BOD5    mg/l(t) and BOD5  Samra   mg/l_1(t+k)

          -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
            +----+----+----+----+----+----+----+----+----+----+
-22  -0.000                            X
-21   0.039                            XX
-20   0.101                            XXXX
-19   0.096                            XXX
-18   0.179                            XXXXX
-17   0.152                            XXXXX
-16   0.193                            XXXXXX
-15   0.173                            XXXXX
-14   0.093                            XXX
-13   0.028                            XX
-12  -0.016                            X
-11  -0.043                           XX
-10   0.008                            X
 -9   0.037                            XX
 -8   0.047                            XX
 -7   0.215                            XXXXXX
 -6   0.214                            XXXXXX
 -5   0.273                            XXXXXXXX
 -4   0.282                            XXXXXXXX
 -3   0.211                            XXXXXX
 -2   0.192                            XXXXXX
 -1   0.110                            XXXX
  0   0.188                            XXXXXX
  1   0.019                            X
  2   0.053                            XX
  3   0.100                            XXX
  4   0.138                            XXXX
  5   0.145                            XXXXX
  6   0.216                            XXXXXX
  7   0.267                            XXXXXXXX
  8   0.245                            XXXXXXX
  9   0.211                            XXXXXX
 10   0.084                            XXX
 11   0.095                            XXX
 12   0.030                            XX
 13   0.070                            XXX
 14   0.062                            XXX
 15   0.067                            XXX
 16   0.205                            XXXXXX
 17   0.185                            XXXXXX
 18   0.249                            XXXXXXX
 19   0.297                            XXXXXXXX
 20   0.279                            XXXXXXXX
 21   0.245                            XXXXXXX
 22   0.186                            XXXXXX
```

4. DISCUSSION, CONCLUSION, AND RECOMENDATION

4.1 Discussion & conclusion:

At the beginning of the analysis, the data had some missing values; these missing values in Zarka River were less than the values in Samra's effluent. The only missing data in Zarka River were in December 1999 and it was in the TSS variable, but in Samra's effluent there were nine missing data, and all of them were only in the year 2000 and they were in different types of Samra's effluent variables.

Missing data were estimated by taking the average of the same months. After substituting the missing data, the data of Zarka River were plotted against time in a scatter diagram. The data showed some abnormal observations, which had to be discussed either as real or not. Some observations were real data and kept as is. Other data could not be explained and were considered as outliers.

The abnormality in the real abnormal observed data were explained, this explanation was due to an external condition that influenced the value of the data, such condition is the huge amount of rainfall that the kingdom received in some months during 1992 and other years (see appendix (1)). In these months, the rainfall was mixed with water in Zarka River & decreased its variables, such as; BOD5, COD, TSS, T-P, and T-N, in the other way it increased the amount of Zarka River flow. The other observed data, which could not have any explanation why this value is abnormal, was treated as an outlier and assumed that the abnormal value was due to human error, such as: human error in reading or calculating the value, equipment error, or any other errors related to human, and it was treated as missing data (average of same months).

Next, normality test was used to plot the normality test of the data; Weibull distribution was used in plotting the histogram of the normality test for each variable. There are other normality histogram plots that could have been used, such as California, Hazen, Beard and other distributions. But Weibull distinguished in its simplicity and it does not have a 100% probability.

The other three normality tests were; the coefficient of variation, which is a preliminary test to determine the horizontal skewness, the Kurtosis test, which is a good test in determining the data normality in the vertical direction, and finally the Shapiro-Wilk test; which showed good results in the normality analysis.

From the Kurtosis and Shapiro –Wilk tests, it has been evident that the two tests had almost identical results in most data quarters, such as; the first quarter of TSS variable, the fourth quarter of BOD5, and other quarters. This proves that the Shapiro-Wilk test is a good method for testing the normality in the vertical direction

The five variables (TSS, BOD5, COD, T-P, and T-N) had a normal distribution, but the flow variable showed abnormality due to the variation in data and the large amount of abnormal observations (9 in total). These abnormal observations were kept and used as it is since they were real data. A lognormal test was made to the flow variable; the lognormal test showed a big difference in data normality between the flow and the log flow. The Weibull distribution model was skewed to right in the flow variable distribution, the coefficient of variation was more than one, and the Kurtosis coefficient reached a value of 20. The lognormal test decreased the skewness and made the data of the flow variable better in normality.

After analyzing the data and finding 16-forecasted values for each variable, the percentage of mean error was calculated. For the TSS variable it was shown that the least percentage of mean error was in the exponential growth method, with an error of 1.7%, but this model's method is a deterministic one, the least percentage of mean error in the stochastic model for TSS variable is an AR(1), which equals to 5.4%, and this was the best method to be used. ARIMA (1,0,4) was also a good model in forecasting the TSS with a mean error of 8.2%.

In forecasting the BOD5 variable, none of the models have satisfied the 10% error. The least percentage of mean error for BOD_5 variable was in ARMA(1,3), which equals to 16.1%.

The stochastic model for COD variable did not satisfy the 10% error, thus, none of the stochastic methods was used in forecasting the COD variable, but in the deterministic modeling, the quadratic method gave the least percentage of mean error which equaled to 3.8%, and it is the best model to be used in forecasting the COD.

For the T-P variable, the only method that satisfied the 10% error was the moving average, MA(4), which gave a percentage of mean error equals to 8.5%, so the moving average MA(4) is the best method to be used in forecasting the T-P variable. In forecasting the T-N variable, many methods have satisfied the 10% of the mean error, the method that gave the least error was the linear method with an error of 3.3%, but the best model to be used in forecasting the T-N variable is ARIMA(1,2,5), which gave 4.8% of mean error.

Finally, for the flow variable, the forecasting was made for log (Q) variable. Many models have satisfied the 10% mean error. However, the computer software (Minitab13) could not estimate the values of ARIMA (1,2,6), because the coefficients of ARIMA model (p,d,q) should be less than 5 in order to be estimated. The best method that gives the least error and to be used in forecasting is AR(1), which gives mean error of 4.9%.

From the above results it can be concluded that the ARIMA model has satisfied the forecasting model for most of the variables. The least amount of mean error resulting from ARIMA model was when calculating the percentage of mean error for ARIMA(1,2,5), which was 4.8%.

The cross and distance correlation gave a good indication about the relations between variables in Zarka River it self and between the Zarka River with Samra's effluent.

- From the relation between the flow and the TSS in Zarka River, it can be concluded that the relation depends on the season, whether it's summer or winter. At winter, the amount of TSS exerted from rainfall in the run off of Zarka River catchment area was high. But in summer, the amount of TSS was low since the suspended solids had been precipitated.

- From the relation between the flow and BOD5 in Zarka River, it can be concluded that the flow coming from the northern south direction has the ability to dilute the Samra's affluent and decrease the amount of BOD5.

- From the relation between the flow and COD in Zarka River, it can be seen that it has the same relation between the flow and BOD5, but with less correlation, that is because the BOD5 measures only the organic matters. Where as, the COD measures organic and inorganic matters, so the variation in COD with flow would be less.

- The relation between the flow and T-P has the same correlation with the flow and TSS. This indicates that most of the phosphorus in Zarka River is in the form of suspended solids.

- From the relation between the flow and T-N, it can be seen that it has the same correlation between the flow rate with BOD_5 and COD. This is because the source of T-N, BOD5, and COD are from the domestic waste. In some cases the relation between the flow and T-N is proportional because some T-N is exerted in the flow of Samra's effluent.

- From the relation between TSS with BOD5 and COD, it can be concluded that most of the organic and inorganic matters are in the form of solids. The correlation between TSS and COD is less than correlation between TSS and BOD5, because the COD variable has more components than the BOD5 variable.

- From the relation between TSS and T-P, it can be concluded that in winter most of the phosphorus is in a solid form. Where as in summer it will be dissolved in Samra's affluent.

- From the relations between the BOD5, COD, T-P, and T-N in Zarka River, it can be seen that all of them have a proportional correlation, so this ensures that the source of these variables is mostly the wastewater.

- From the relation between the flow in Zarka River and the flow from Samra's effluent, it can be concluded that the time of the huge flow in Zarka River, which is mostly between December and March, differs from the huge releases in Samra's effluent, which is mostly between June and July.
- From the relations between the BOD5, COD, T-P, and T-N in Zarka River and Samra's effluent, it can be seen that in each variable the quality of water in Samra's effluent affects positively the quality of water in Zarka River.

4.2 Recommendations:

1. Improving the quality of wastewater in Samra's effluent, which affects mainly Zarka River water quality. This improvement could be achieved by increasing the capacity of Al Samra WWTP with decreasing the concentration of the variables (TSS, BOD5, COD, T-P, and T-N) of the Samra's effluent.
2. Finding alternatives and solutions for the usage of Al Samra's effluent and trying to decrease the amount of the effluent that goes to the Zarka River
3. The necessity of studying the environmental impact assessment (EIA) for future projects in Zarka River catchment area, which could influence the water quality in this River; such as: Industrial factories, agricultural projects, and wastewater treatment plants.
4. Decreasing the amount of TSS in Zarka River exerted from soil erosion by rainfall. Also cleaning the precipitated solids in King Talal Reservoir, taking into consideration that the amount of precipitated solids reach up to 13.0 million cubic meters in the year 2000. (RSS reports, 2001) .
5. A study should be made to find the correlation between the cross and distance correlation in the vertical (depth) and horizontal directions, so that the variables

achieved from one site could be used to know the similar variables in any other site along the whole track of the flow.

6. Decreasing the amount of phosphorus and Nitrogen, because there presence in water in a concentration of 300 mg/m3 for phosphorus and 5 mg/m3 for nitrogen will create the eutrophication process and encourage the algae to build up. The algae is already exist in Zarka River, decreasing the concentration of T-P and T-N using Macrophytes, which lives on algae will help in minimizing the amount of algae.

7. For the time series analysis, as increasing the amount of months, as the time series forecasting will be more accurate. A thirteen years data collected is not enough to show the trend and seasonal affect in a proper way.

8. The values of the variable should be very accurate, since the samples are taken to Amman and then tested there, this can decrease the accuracy of the data. A site laboratory will increase the accuracy of the data.

9. The time of sample should be in the peak hours; usually the samples were taken in the working day (between 8 and 3), which is not necessary to be the peak hours. Taking three or more daily readings for each point will increase the accuracy of the data.

10. More information about any strange reading should be recorded, so that any reading to be considered as an outlier should be justified by comprehensive reasons.

11. Decreasing the amount of missing data will help in decreasing the randomness in data.

12. There should be Jordanian specifications for Irrigation and Agriculture instead of using the FAO specifications or any other specifications in analyzing the water quality.

APPENDICES

Appendix (1): The amount of rainfall in Al Zarka River catchment area

Year	Amount of Rainfall in MCM/Month					
	Jan.	Feb.	March	April	May	June
1988	0.051	0.207	0.135	0.020	0.000	0.000
1989	0.073	0.050	0.093	0.000	0.000	0.000
1990	0.102	0.045	0.084	0.035	0.000	0.000
1991	0.070	0.048	0.058	0.015	0.000	0.000
1992	0.267	0.466	0.054	0.000	0.003	0.008
1993	0.115	0.493	0.045	0.002	0.014	0.000
1994	0.164	0.054	0.076	0.008	0.000	0.007
1995	0.021	0.129	0.033	0.018	0.000	0.000
1996	0.157	0.015	0.218	0.016	0.000	0.000
1997	0.152	0.190	0.178	0.012	0.014	0.000
1998	0.184	0.065	0.258	0.000	0.000	0.000
1999	0.128	0.095	0.049	0.017	0.000	0.000
2000	0.126	0.046	0.054	0.000	0.000	0.000

Year	Amount of Rainfall in MCM/Month					
	July	Aug.	Sep.	Oct.	Nov.	Dec.
1988	0.020	0.000	0.000	0.003	0.026	0.149
1989	0.000	0.000	0.000	0.003	0.031	0.059
1990	0.000	0.000	0.000	0.005	0.016	0.001
1991	0.000	0.000	0.000	0.000	0.032	0.162
1992	0.000	0.000	0.000	0.000	0.088	0.225
1993	0.000	0.000	0.000	0.004	0.012	0.011
1994	0.000	0.000	0.000	0.005	0.200	0.142
1995	0.000	0.000	0.000	0.000	0.032	0.034
1996	0.000	0.000	0.000	0.013	0.023	0.071
1997	0.000	0.000	0.000	0.030	0.068	0.154
1998	0.000	0.000	0.000	0.000	0.000	0.009
1999	0.000	0.000	0.000	0.000	0.002	0.007
2000	0.000	0.000	0.000	0.027	0.006	0.063

Appendix (2):Coefficients (a_{N-I+1}) for Shapiro-Wilk *W*-test of normality

i/n	2	3	4	5	6	7	8	9	10
1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739
2	-	0.0000	0.1677	0.2413	0.2806	0.0310	0.3164	0.3244	0.3291
3	-	-	-	0.0000	0.0875	0.1401	0.1743	0.1976	0.2141
4	-	-	-	-	-	0.0000	0.0561	0.0947	0.1224
5	-	-	-	-	-	-	-	0.0000	0.0399

i/n	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	0.3315	0.3325	0.3325	0.3318	0.3306	0.3290	0.3273	0.3253	0.3232	0.3211
3	0.2260	0.2347	0.2412	0.2460	0.2495	0.2521	0.2540	0.2553	0.2561	0.2565
4	0.1429	0.1586	0.1707	0.1802	0.1878	0.1939	0.1988	0.2027	0.2059	0.2085
5	0.0695	0.0922	0.1099	0.1240	0.1353	0.1447	0.1524	0.1587	0.1641	0.1686
6	0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7	-	-	0.0000	0.0240	0.0433	0.0593	0.0725	0.0837	0.0932	0.1013
8	-	-	-	-	0.0000	0.0196	0.0359	0.0496	0.0621	0.0711
9	-	-	-	-	-	-	0.0000	0.0163	0.0303	0.0422
10	-	-	-	-	-	-	-	-	0.0000	0.0140

Cont. Appendix (2):Coefficients (a_{N-I+1}) for Shapiro-Wilk *W*-test of normality

i/n	21	22	23	24	25	26	27	28	29	30
1	0.4643	0.4590	0.4542	0.4493	0.4450	0.4407	0.4366	0.4328	0.4291	0.4254
2	0.3185	0.3156	0.3126	0.3098	0.3069	0.3043	0.3018	0.2992	0.2968	0.2944
3	0.2578	0.2571	0.2563	0.2554	0.2543	0.2533	0.2522	0.2510	0.2499	0.2487
4	0.2119	0.2131	0.2139	0.2145	0.2148	0.2151	0.2152	0.2151	0.2150	0.2148
5	0.1736	0.1764	0.1787	0.1807	0.1822	0.1836	0.1848	0.1857	0.1864	0.1870
6	0.1399	0.1443	0.1480	0.1512	0.1539	0.1563	0.1584	0.1601	0.1616	0.1630
7	0.1092	0.1150	0.1201	0.1245	0.1283	0.1316	0.1346	0.1372	0.1395	0.1415
8	0.0804	0.0878	0.0941	0.0997	0.1046	0.1089	0.1128	0.1162	0.1192	0.1219
9	0.0530	0.0618	0.0696	0.0764	0.0823	0.0876	0.0923	0.0965	0.1002	0.1036
10	0.0263	0.0368	0.0459	0.0539	0.0610	0.0672	0.0728	0.0778	0.0822	0.0862
11	0.0000	0.0122	0.0228	0.0321	0.0403	0.0476	0.0540	0.0598	0.0650	0.0697
12	-	-	0.0000	0.0107	0.0200	0.0284	0.0358	0.0424	0.0483	0.0537
13	-	-	-	-	0.0000	0.0094	0.0178	0.0253	0.0320	0.0381
14	-	-	-	-	-	-	0.0000	0.0084	0.0159	0.0227
15	-	-	-	-	-	-	-	-	0.0000	0.0076

Cont. Appendix (2):Coefficients (a_{N-I+1}) for Shapiro-Wilk *W*-test of normality

i/n	31	32	33	34	35	36	37	38	39	40
1	0.4220	0.4188	0.4156	0.4127	0.4096	0.4068	0.4040	0.4015	0.3989	0.3964
2	0.2921	0.2898	0.2876	0.2854	0.2834	0.2813	0.2794	0.2774	0.2755	0.2737
3	0.2475	0.2463	0.2451	0.2439	0.2427	0.2415	0.2403	0.2391	0.2380	0.2368
4	0.2145	0.2141	0.2137	0.2132	0.2127	0.2121	0.2116	0.2110	0.2104	0.2098
5	0.1874	0.1878	0.1880	0.1882	0.1883	0.1883	0.1883	0.1881	0.1880	0.1878
6	0.1641	0.1651	0.1660	0.1667	0.1673	0.1678	0.1683	0.1686	0.1689	0.1691
7	0.1433	0.1449	0.1463	0.1475	0.1487	0.1496	0.1503	0.1513	0.1520	0.1526
8	0.1243	0.1265	0.1284	0.1301	0.1317	0.1331	0.1344	0.1356	0.1366	0.1376
9	0.1066	0.1093	0.1118	0.1140	0.1160	0.1179	0.1196	0.1211	0.1250	0.1237
10	0.0899	0.0931	0.0961	0.0988	0.1013	0.1036	0.1056	0.1075	0.1092	0.1108
11	0.0739	0.0777	0.0812	0.0844	0.0873	0.0900	0.0924	0.0947	0.0967	0.0896
12	0.0585	0.0629	0.0669	0.0706	0.0739	0.0770	0.0798	0.0824	0.0848	0.0870
13	0.0435	0.0485	0.0530	0.0572	0.0610	0.0645	0.0677	0.0706	0.0733	0.0759
14	0.0289	0.0344	0.0395	0.0441	0.0484	0.0523	0.0559	0.0592	0.0622	0.0651
15	0.0144	0.0206	0.0262	0.0314	0.0361	0.0404	0.0444	0.0481	0.0515	0.0546
16	0.0000	0.0068	0.0131	0.0187	0.0239	0.0287	0.0331	0.0372	0.0409	0.0444
17	-	-	0.0000	0.0062	0.0119	0.0172	0.0220	0.0264	0.0305	0.0343
18	-	-	-	-	0.0000	0.0057	0.0110	0.0158	0.0203	0.0244
19	-	-	-	-	-	-	0.0000	0.0053	0.0101	0.0146
20	-	-	-	-	-	-	-	-	0.0000	0.0049

Cont. Appendix (2):Coefficients (a_{N-I+1}) for Shapiro-Wilk *W*-test of normality

i/n	41	42	43	44	45	46	47	48	49	50
1	0.3940	0.3917	0.3894	0.3872	0.3850	0.3830	0.3808	0.3789	0.3000	0.3751
2	0.2719	0.2701	0.2684	0.2667	0.2651	0.2635	0.2620	0.2604	0.2589	0.2574
3	0.2357	0.2345	0.2334	0.2323	0.2313	0.2302	0.2291	0.2281	0.2271	0.2260
4	0.2091	0.2085	0.2078	0.2072	0.2065	0.2058	0.2052	0.2045	0.2038	0.2032
5	0.1876	0.1874	0.1871	0.1868	0.1865	0.1862	0.1859	0.1855	0.1851	0.1847
6	0.1693	0.1694	0.1695	0.1695	0.1695	0.1695	0.1695	0.1693	0.1692	0.1691
7	0.1531	0.1535	0.1539	0.1542	0.1545	0.1548	0.1550	0.1551	0.1553	0.1554
8	0.1384	0.1392	0.1398	0.1405	0.1410	0.1415	0.1420	0.1423	0.1427	0.1430
9	0.1249	0.1259	0.1269	0.1278	0.1286	0.1293	0.1300	0.1306	0.1312	0.1317
10	0.1123	0.1136	0.1149	0.1160	0.1170	0.1180	0.1189	0.1197	0.1205	0.1212
11	0.1004	0.1020	0.1035	0.1049	0.1062	0.1073	0.1085	0.1095	0.1105	0.1113
12	0.0891	0.0909	0.0927	0.0943	0.0959	0.0972	0.0986	0.0998	0.1010	0.1020
13	0.0782	0.0804	0.0824	0.0842	0.0860	0.0876	0.0892	0.0906	0.0919	0.0932
14	0.0677	0.0701	0.0724	0.0745	0.0775	0.0785	0.0801	0.0817	0.0832	0.0846
15	0.0575	0.0602	0.0628	0.0651	0.0673	0.0694	0.0713	0.0731	0.0748	0.0764
16	0.0476	0.0506	0.0534	0.0560	0.0584	0.0607	0.0628	0.0648	0.0662	0.0685
17	0.0379	0.0411	0.0442	0.0471	0.0497	0.0522	0.0546	0.0568	0.0588	0.0608
18	0.0283	0.0318	0.0352	0.0383	0.0412	0.0439	0.0465	0.0489	0.0511	0.0532
19	0.0188	0.0227	0.0263	0.0296	0.0328	0.0357	0.0385	0.0411	0.0436	0.0459
20	0.0094	0.0316	0.0175	0.0211	0.0245	0.0277	0.0307	0.0335	0.0361	0.0386
21	0.0000	0.0045	0.0087	0.0126	0.0163	0.0197	0.0229	0.0259	0.0288	0.0314
22	-	-	0.0000	0.0042	0.0081	0.0118	0.0153	0.0185	0.0215	0.0244
23	-	-	-	-	0.0000	0.0039	0.0076	0.0111	0.0143	0.0174
24	-	-	-	-	-	-	0.0000	0.0037	0.0071	0.0104
25	-	-	-	-	-	-	-	-	0.0000	0.0035

Appendix (3): Percentage points of the Shapiro-Wilk W-test

	1 – Confidence Interval			1 – Confidence Interval	
n	**0.01**	**0.05**	**n**	**0.01**	**0.05**
3	0.753	0.767	27	0.894	0.923
4	0.687	0.748	28	0.896	0.924
5	0.686	0.762	29	0.898	0.926
6	0.713	0.788	30	0.900	0.927
7	0.730	0.803	31	0.902	0.929
8	0.749	0.818	32	0.904	0.930
9	0.764	0.829	33	0.906	0.931
10	0.781	0.842	34	0.908	0.933
11	0.792	0.850	35	0.910	0.934
12	0.805	0.859	36	0.912	0.935
13	0.814	0.866	37	0.914	0.936
14	0.825	0.874	38	0.914	0.938
15	0.835	0.881	39	0.917	0.939
16	0.844	0.887	40	0.919	0.940
17	0.851	0.892	41	0.920	0.941
18	0.858	0.897	42	0.922	0.942
19	0.863	0.901	43	0.923	0.943
20	0.868	0.905	44	0.924	0.944
21	0.873	0.908	45	0.926	0.945
22	0.878	0.911	46	0.927	0.945
23	0.881	0.914	47	0.928	0.946
24	0.884	0.916	48	0.929	0.947
25	0.888	0.918	49	0.929	0.947
26	0.891	0.920	50	0.930	0.947

Appendix (4):

A4.1: Cross Correlation Function: Zarka Flow MCM/month; TSS Zarka mg/l

CCF - correlates Zarka Flow MCM/month(t) and TSS Zarka mg/l(t+k)

```
          -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
            +----+----+----+----+----+----+----+----+----+----+
-22  -0.220                     XXXXXX
-21  -0.161                      XXXXX
-20  -0.031                        XX
-19   0.069                        XXX
-18   0.198                        XXXXXX
-17   0.197                        XXXXXX
-16   0.299                        XXXXXXXX
-15   0.208                        XXXXXX
-14   0.154                        XXXXX
-13  -0.005                        X
-12  -0.165                      XXXXX
-11  -0.199                     XXXXXX
-10  -0.250                    XXXXXXX
 -9  -0.185                    XXXXXX
 -8  -0.056                        XX
 -7   0.012                        X
 -6   0.097                        XXX
 -5   0.202                        XXXXXX
 -4   0.209                        XXXXXX
 -3   0.171                        XXXXX
 -2   0.034                        XX
 -1  -0.044                       XX
  0  -0.164                     XXXXX
  1  -0.166                     XXXXX
  2  -0.109                      XXXX
  3   0.031                        XX
  4   0.162                        XXXXX
  5   0.220                        XXXXXXX
  6   0.209                        XXXXXX
  7   0.232                        XXXXXXX
  8   0.206                        XXXXXX
  9   0.088                        XXX
 10   0.021                        XX
 11  -0.113                      XXXX
 12  -0.210                     XXXXXX
 13  -0.213                     XXXXXX
 14  -0.189                     XXXXXX
 15  -0.146                      XXXXX
 16  -0.074                       XXX
 17   0.091                        XXX
 18   0.194                        XXXXXX
 19   0.133                        XXXX
 20   0.157                        XXXXX
 21   0.149                        XXXXX
 22   0.018                        X
```

Cont. Appendix (4):

A4.2: Cross Correlation Function: Zarka Flow MCM/month; BOD5 Zarka mg/l

CCF - correlates Zarka Flow MCM/month(t) and BOD5 Zarka mg/l(t+k)

```
             -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
               +----+----+----+----+----+----+----+----+----+----+
 -22  -0.151            XXXXX
 -21  -0.088              XXX
 -20   0.005                X
 -19   0.031                XX
 -18  -0.015                X
 -17  -0.025               XX
 -16   0.036                XX
 -15   0.005                X
 -14  -0.000                X
 -13  -0.041               XX
 -12  -0.086              XXX
 -11  -0.117             XXXX
 -10  -0.110             XXXX
  -9  -0.083              XXX
  -8  -0.024               XX
  -7  -0.018                X
  -6  -0.047               XX
  -5  -0.015                X
  -4   0.013                X
  -3  -0.025               XX
  -2  -0.095              XXX
  -1  -0.126             XXXX
   0  -0.250          XXXXXXX
   1  -0.224          XXXXXXX
   2  -0.098              XXX
   3  -0.053               XX
   4  -0.020               XX
   5   0.062                XXX
   6   0.003                X
   7  -0.029               XX
   8  -0.096              XXX
   9  -0.156            XXXXX
  10  -0.145            XXXXX
  11  -0.193           XXXXXX
  12  -0.171            XXXXX
  13  -0.093              XXX
  14  -0.009                X
  15   0.077                XXX
  16   0.096                XXX
  17   0.167                XXXXX
  18   0.160                XXXXX
  19   0.119                XXXX
  20   0.073                XXX
  21   0.096                XXX
  22   0.060                XXX
```

Cont. Appendix (4):

A4.3: Cross Correlation Function: Zarka Flow MCM/month; COD Zarka mg/l

CCF - correlates Zarka Flow MCM/month(t) and COD Zarka mg/l(t+k)

```
          -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
            +----+----+----+----+----+----+----+----+----+----+
-22  -0.142                    XXXXX
-21  -0.103                     XXXX
-20  -0.024                       XX
-19   0.008                        X
-18   0.038                        XX
-17   0.053                        XX
-16   0.053                        XX
-15   0.061                        XXX
-14   0.085                        XXX
-13   0.005                        X
-12  -0.050                       XX
-11  -0.160                    XXXXX
-10  -0.169                    XXXXX
 -9  -0.141                    XXXXX
 -8  -0.060                      XXX
 -7  -0.027                       XX
 -6   0.014                        X
 -5  -0.003                        X
 -4  -0.010                        X
 -3  -0.042                       XX
 -2  -0.149                    XXXXX
 -1  -0.110                     XXXX
  0  -0.261                 XXXXXXXX
  1  -0.342               XXXXXXXXXX
  2  -0.238                  XXXXXXX
  3  -0.170                    XXXXX
  4  -0.130                     XXXX
  5  -0.061                      XXX
  6  -0.071                      XXX
  7  -0.049                       XX
  8  -0.045                       XX
  9  -0.044                       XX
 10  -0.084                      XXX
 11  -0.140                    XXXXX
 12  -0.190                   XXXXXX
 13  -0.183                   XXXXXX
 14  -0.144                    XXXXX
 15  -0.050                       XX
 16  -0.004                        X
 17   0.030                        XX
 18   0.028                        XX
 19   0.037                        XX
 20   0.041                        XX
 21   0.022                        XX
 22  -0.028                       XX
```

Cont. Appendix (4):

A4.4: Cross Correlation Function: Zarka Flow MCM/month; T-P Zarka mg/l

CCF - correlates Zarka Flow MCM/month(t) and T-P Zarka mg/l(t+k)

```
       -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
        +----+----+----+----+----+----+----+----+----+----+
-22  -0.127                  XXXX
-21  -0.079                   XXX
-20   0.015                     X
-19   0.052                     XX
-18   0.101                     XXXX
-17   0.110                     XXXX
-16   0.099                     XXX
-15   0.121                     XXXX
-14   0.090                     XXX
-13   0.024                     XX
-12  -0.103                  XXXX
-11  -0.208                XXXXXX
-10  -0.185                XXXXXX
 -9  -0.159                XXXXX
 -8  -0.116                 XXXX
 -7  -0.028                   XX
 -6   0.009                   X
 -5   0.075                   XXX
 -4   0.063                   XXX
 -3   0.054                   XX
 -2  -0.079                  XXX
 -1  -0.250               XXXXXXX
  0  -0.409            XXXXXXXXXXX
  1  -0.462           XXXXXXXXXXXXX
  2  -0.368             XXXXXXXXXX
  3  -0.309             XXXXXXXXX
  4  -0.195               XXXXXX
  5  -0.126                XXXX
  6  -0.048                 XX
  7  -0.008                  X
  8  -0.000                  X
  9  -0.047                 XX
 10  -0.141                XXXXX
 11  -0.159                XXXXX
 12  -0.222               XXXXXXX
 13  -0.233               XXXXXXX
 14  -0.221               XXXXXXX
 15  -0.116                 XXXX
 16  -0.032                  XX
 17   0.068                  XXX
 18   0.128                  XXXX
 19   0.149                  XXXXX
 20   0.115                  XXXX
 21   0.110                  XXXX
 22   0.032                  XX
```

Cont. Appendix (4):

A4.5: Cross Correlation Function: Zarka Flow MCM/month; T-N Zarka mg/l

CCF - correlates Zarka Flow MCM/month(t) and T-N Zarka mg/l(t+k)

```
         -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
           +----+----+----+----+----+----+----+----+----+----+
-22  -0.130                    XXXX
-21  -0.142                   XXXXX
-20  -0.124                    XXXX
-19  -0.148                   XXXXX
-18  -0.148                   XXXXX
-17  -0.156                   XXXXX
-16  -0.069                     XXX
-15  -0.022                      XX
-14   0.013                       X
-13  -0.008                       X
-12  -0.058                      XX
-11  -0.101                    XXXX
-10  -0.147                   XXXXX
 -9  -0.155                   XXXXX
 -8  -0.179                   XXXXX
 -7  -0.184                  XXXXXX
 -6  -0.154                   XXXXX
 -5  -0.130                    XXXX
 -4  -0.094                     XXX
 -3  -0.036                      XX
 -2  -0.035                      XX
 -1  -0.114                    XXXX
  0  -0.214                  XXXXXX
  1  -0.259                 XXXXXXX
  2  -0.211                  XXXXXX
  3  -0.230                 XXXXXXX
  4  -0.206                  XXXXXX
  5  -0.232                 XXXXXXX
  6  -0.246                 XXXXXXX
  7  -0.164                   XXXXX
  8  -0.116                    XXXX
  9  -0.108                    XXXX
 10  -0.126                    XXXX
 11  -0.151                   XXXXX
 12  -0.151                   XXXXX
 13  -0.153                   XXXXX
 14  -0.137                    XXXX
 15  -0.122                    XXXX
 16  -0.103                    XXXX
 17  -0.060                      XX
 18  -0.034                      XX
 19  -0.017                       X
 20   0.024                       XX
 21   0.069                       XXX
 22   0.069                       XXX
```

Cont. Appendix (4):

A4.6: Cross Correlation Function: TSS mg/l ; BOD5 mg/l for Zarka River

```
CCF - correlates TSS      mg/l(t) and BOD5    mg/l(t+k)

       -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
         +----+----+----+----+----+----+----+----+----+----+
-22   0.153                   XXXXX
-21   0.008                   X
-20  -0.028                  XX
-19  -0.021                  XX
-18   0.016                   X
-17   0.037                   XX
-16   0.210                   XXXXXX
-15   0.105                   XXXX
-14   0.175                   XXXXX
-13   0.138                   XXXX
-12   0.181                   XXXXXX
-11   0.088                   XXX
-10  -0.047                  XX
 -9  -0.026                  XX
 -8   0.036                   XX
 -7   0.049                   XX
 -6   0.035                   XX
 -5   0.088                   XXX
 -4   0.121                   XXXX
 -3   0.159                   XXXXX
 -2   0.251                   XXXXXXX
 -1   0.291                   XXXXXXXX
  0   0.394                   XXXXXXXXXXX
  1   0.130                   XXXX
  2  -0.001                   X
  3   0.082                   XXX
  4   0.008                   X
  5  -0.031                  XX
  6   0.039                   XX
  7   0.119                   XXXX
  8   0.125                   XXXX
  9   0.096                   XXX
 10   0.246                   XXXXXXX
 11   0.184                   XXXXXX
 12   0.155                   XXXXX
 13   0.118                   XXXX
 14   0.049                   XX
 15  -0.082                  XXX
 16  -0.043                  XX
 17   0.024                   XX
 18   0.055                   XX
 19   0.073                   XXX
 20   0.042                   XX
 21   0.196                   XXXXXX
 22   0.214                   XXXXXX
```

Cont. Appendix (4):

A4.7: Cross Correlation Function: TSS mg/l; COD mg/l in Zarka River

CCF - correlates TSS mg/l(t) and COD mg/l(t+k)

```
       -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
         +----+----+----+----+----+----+----+----+----+----+
-22   0.122                      XXXX
-21   0.011                      X
-20  -0.018                      X
-19  -0.026                     XX
-18   0.008                      X
-17  -0.049                     XX
-16   0.044                      XX
-15   0.051                      XX
-14   0.134                      XXXX
-13   0.121                      XXXX
-12   0.147                      XXXXX
-11   0.135                      XXXX
-10   0.071                      XXX
 -9   0.006                      X
 -8  -0.018                      X
 -7  -0.030                     XX
 -6  -0.039                     XX
 -5  -0.045                     XX
 -4  -0.009                      X
 -3   0.064                      XXX
 -2   0.114                      XXXX
 -1   0.254                      XXXXXXX
  0   0.338                      XXXXXXXXX
  1   0.272                      XXXXXXXX
  2   0.093                      XXX
  3   0.021                      XX
  4   0.070                      XXX
  5  -0.057                     XX
  6  -0.140                   XXXXX
  7  -0.024                     XX
  8   0.009                      X
  9  -0.058                     XX
 10   0.109                      XXXX
 11   0.228                      XXXXXXX
 12   0.205                      XXXXXX
 13   0.204                      XXXXXX
 14   0.064                      XXX
 15   0.029                      XX
 16  -0.082                    XXX
 17  -0.053                     XX
 18  -0.040                     XX
 19  -0.023                     XX
 20  -0.174                   XXXXX
 21  -0.037                     XX
 22   0.024                      XX
```

Cont. Appendix (4):

A4.8: Cross Correlation Function: TSS mg/l; T-P mg/l in Zarka River

```
CCF - correlates TSS     mg/l(t) and T-P    mg/l(t+k)

          -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
            +----+----+----+----+----+----+----+----+----+----+
-22   0.153                         XXXXX
-21   0.037                         XX
-20   0.085                         XXX
-19  -0.038                        XX
-18  -0.058                        XX
-17  -0.073                       XXX
-16  -0.002                         X
-15   0.099                         XXX
-14   0.171                         XXXXX
-13   0.279                         XXXXXXXX
-12   0.203                         XXXXXX
-11   0.235                         XXXXXXX
-10   0.226                         XXXXXXX
 -9   0.141                         XXXXX
 -8  -0.068                       XXX
 -7  -0.071                       XXX
 -6  -0.081                       XXX
 -5  -0.085                       XXX
 -4  -0.135                      XXXX
 -3   0.054                         XX
 -2   0.160                         XXXXX
 -1   0.291                         XXXXXXXX
  0   0.352                         XXXXXXXXXX
  1   0.272                         XXXXXXXX
  2   0.154                         XXXXX
  3   0.078                         XXX
  4   0.007                         X
  5  -0.102                      XXXX
  6  -0.155                     XXXXX
  7  -0.110                      XXXX
  8  -0.036                        XX
  9   0.056                         XX
 10   0.200                         XXXXXX
 11   0.216                         XXXXXX
 12   0.247                         XXXXXXX
 13   0.233                         XXXXXXX
 14   0.169                         XXXXX
 15   0.080                         XXX
 16  -0.054                        XX
 17  -0.088                       XXX
 18  -0.151                     XXXXX
 19  -0.098                       XXX
 20  -0.074                       XXX
 21  -0.048                        XX
 22   0.071                         XXX
```

Cont. Appendix (4):

A4.9: Cross Correlation Function: TSS mg/l; T-N mg/l in Zarka River

```
CCF - correlates TSS      mg/l(t) and T-N    mg/l(t+k)

        -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
          +----+----+----+----+----+----+----+----+----+----+
-22   0.141                     XXXXX
-21   0.072                     XXX
-20   0.095                     XXX
-19   0.090                     XXX
-18   0.076                     XXX
-17   0.021                     XX
-16   0.032                     XX
-15   0.012                     X
-14   0.060                     XXX
-13   0.098                     XXX
-12   0.150                     XXXXX
-11   0.158                     XXXXX
-10   0.219                     XXXXXX
 -9   0.152                     XXXXX
 -8   0.210                     XXXXXX
 -7   0.147                     XXXXX
 -6   0.166                     XXXXX
 -5   0.117                     XXXX
 -4   0.056                     XX
 -3   0.076                     XXX
 -2   0.096                     XXX
 -1   0.159                     XXXXX
  0   0.252                     XXXXXXX
  1   0.151                     XXXXX
  2   0.183                     XXXXXX
  3   0.164                     XXXXX
  4   0.213                     XXXXXX
  5   0.140                     XXXX
  6   0.065                     XXX
  7   0.047                     XX
  8   0.023                     XX
  9   0.054                     XX
 10   0.101                     XXXX
 11   0.103                     XXXX
 12   0.169                     XXXXX
 13   0.155                     XXXXX
 14   0.119                     XXXX
 15   0.099                     XXX
 16   0.125                     XXXX
 17   0.109                     XXXX
 18   0.082                     XXX
 19  -0.001                     X
 20  -0.007                     X
 21  -0.000                     X
 22   0.083                     XXX
```

Cont. Appendix (4):

A4.10: Cross Correlation Function: BOD5 mg/l; COD mg/l in Zarka River

```
CCF - correlates BOD5    mg/l(t) and COD    mg/l(t+k)

         -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
           +----+----+----+----+----+----+----+----+----+----+
-22   0.086                        XXX
-21   0.046                        XX
-20   0.044                        XX
-19   0.060                        XXX
-18   0.099                        XXX
-17   0.053                        XX
-16   0.047                        XX
-15   0.041                        XX
-14   0.028                        XX
-13   0.117                        XXXX
-12   0.177                        XXXXX
-11   0.233                        XXXXXXX
-10   0.226                        XXXXXXX
 -9   0.194                        XXXXXX
 -8   0.141                        XXXXX
 -7   0.153                        XXXXX
 -6   0.209                        XXXXXX
 -5   0.172                        XXXXX
 -4   0.085                        XXX
 -3   0.171                        XXXXX
 -2   0.166                        XXXXX
 -1   0.251                        XXXXXXX
  0   0.453                        XXXXXXXXXXXX
  1   0.403                        XXXXXXXXXXX
  2   0.351                        XXXXXXXXXX
  3   0.271                        XXXXXXXX
  4   0.285                        XXXXXXXX
  5   0.279                        XXXXXXXX
  6   0.264                        XXXXXXXX
  7   0.296                        XXXXXXXX
  8   0.262                        XXXXXXXX
  9   0.146                        XXXXX
 10   0.226                        XXXXXXX
 11   0.203                        XXXXXX
 12   0.319                        XXXXXXXXX
 13   0.320                        XXXXXXXXX
 14   0.210                        XXXXXX
 15   0.262                        XXXXXXXX
 16   0.236                        XXXXXXX
 17   0.162                        XXXXX
 18   0.265                        XXXXXXXX
 19   0.203                        XXXXXX
 20   0.180                        XXXXXX
 21   0.194                        XXXXXX
 22   0.178                        XXXXX
```

Cont. Appendix (4):
A4.11: Cross Correlation Function: BOD5 mg/l; T-P mg/l in Zarka River

CCF - correlates BOD5 mg/l(t) and T-P mg/l(t+k)

```
       -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
         +----+----+----+----+----+----+----+----+----+----+
-22  0.159                    XXXXX
-21  0.157                    XXXXX
-20  0.191                    XXXXXX
-19  0.141                    XXXXX
-18  0.123                    XXXX
-17  0.063                    XXX
-16  0.119                    XXXX
-15  0.075                    XXX
-14  0.083                    XXX
-13  0.276                    XXXXXXXX
-12  0.312                    XXXXXXXXX
-11  0.334                    XXXXXXXXX
-10  0.352                    XXXXXXXXXX
 -9  0.349                    XXXXXXXXXX
 -8  0.335                    XXXXXXXXX
 -7  0.285                    XXXXXXXX
 -6  0.282                    XXXXXXXX
 -5  0.216                    XXXXXX
 -4  0.150                    XXXXX
 -3  0.273                    XXXXXXXX
 -2  0.289                    XXXXXXXX
 -1  0.368                    XXXXXXXXXX
  0  0.451                    XXXXXXXXXXXX
  1  0.437                    XXXXXXXXXXXX
  2  0.453                    XXXXXXXXXXXX
  3  0.379                    XXXXXXXXXX
  4  0.290                    XXXXXXXX
  5  0.288                    XXXXXXXX
  6  0.304                    XXXXXXXXX
  7  0.263                    XXXXXXXX
  8  0.152                    XXXXX
  9  0.215                    XXXXXX
 10  0.275                    XXXXXXXX
 11  0.317                    XXXXXXXXX
 12  0.276                    XXXXXXXX
 13  0.339                    XXXXXXXXX
 14  0.263                    XXXXXXXX
 15  0.278                    XXXXXXXX
 16  0.293                    XXXXXXXX
 17  0.203                    XXXXXX
 18  0.202                    XXXXXX
 19  0.191                    XXXXXX
 20  0.180                    XXXXXX
 21  0.118                    XXXX
 22  0.183                    XXXXXX
```

Cont. Appendix (4):

A4.12: Cross Correlation Function: BOD5 mg/l; T-N mg/l in Zarka River

CCF - correlates BOD5 mg/l(t) and T-N mg/l(t+k)

```
        -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
          +----+----+----+----+----+----+----+----+----+----+
-22   0.178                      XXXXX
-21   0.171                      XXXXX
-20   0.170                      XXXXX
-19   0.224                      XXXXXXX
-18   0.209                      XXXXXX
-17   0.213                      XXXXXX
-16   0.188                      XXXXXX
-15   0.145                      XXXXX
-14   0.249                      XXXXXXX
-13   0.280                      XXXXXXXX
-12   0.309                      XXXXXXXXX
-11   0.338                      XXXXXXXXX
-10   0.352                      XXXXXXXXXX
 -9   0.345                     XXXXXXXXXX
 -8   0.401                     XXXXXXXXXXX
 -7   0.371                     XXXXXXXXXX
 -6   0.448                     XXXXXXXXXXXX
 -5   0.392                     XXXXXXXXXXX
 -4   0.365                     XXXXXXXXXX
 -3   0.426                     XXXXXXXXXXXX
 -2   0.373                     XXXXXXXXXX
 -1   0.494                     XXXXXXXXXXXXX
  0   0.561                     XXXXXXXXXXXXXXX
  1   0.482                     XXXXXXXXXXXXX
  2   0.521                     XXXXXXXXXXXXXX
  3   0.440                     XXXXXXXXXXXX
  4   0.498                     XXXXXXXXXXXXX
  5   0.471                     XXXXXXXXXXXXX
  6   0.536                     XXXXXXXXXXXXXX
  7   0.425                     XXXXXXXXXXXX
  8   0.409                     XXXXXXXXXXX
  9   0.391                     XXXXXXXXXXX
 10   0.406                     XXXXXXXXXXX
 11   0.463                     XXXXXXXXXXXXX
 12   0.420                     XXXXXXXXXXX
 13   0.436                     XXXXXXXXXXXX
 14   0.397                     XXXXXXXXXXX
 15   0.429                     XXXXXXXXXXXX
 16   0.479                     XXXXXXXXXXXXX
 17   0.488                     XXXXXXXXXXXXX
 18   0.481                     XXXXXXXXXXXXX
 19   0.394                     XXXXXXXXXXX
 20   0.360                     XXXXXXXXXX
 21   0.335                     XXXXXXXXX
 22   0.355                     XXXXXXXXXX
```

Cont. Appendix (4):

A4.13: Cross Correlation Function: COD mg/l; T-P mg/l in Zarka River

```
CCF - correlates COD    mg/l(t) and T-P    mg/l(t+k)

        -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
          +----+----+----+----+----+----+----+----+----+----+
-22   0.146                  XXXXX
-21   0.185                  XXXXXX
-20   0.193                  XXXXXX
-19   0.159                  XXXXX
-18   0.150                  XXXXX
-17   0.178                  XXXXX
-16   0.156                  XXXXX
-15   0.278                  XXXXXXXX
-14   0.302                  XXXXXXXXX
-13   0.389                  XXXXXXXXXXX
-12   0.339                  XXXXXXXXX
-11   0.379                  XXXXXXXXXX
-10   0.333                  XXXXXXXXX
 -9   0.283                  XXXXXXXX
 -8   0.203                  XXXXXX
 -7   0.218                  XXXXXX
 -6   0.206                  XXXXXX
 -5   0.264                  XXXXXXXX
 -4   0.280                  XXXXXXXX
 -3   0.351                  XXXXXXXXXX
 -2   0.505                  XXXXXXXXXXXXXX
 -1   0.557                  XXXXXXXXXXXXXXX
  0   0.644                  XXXXXXXXXXXXXXXXX
  1   0.474                  XXXXXXXXXXXXX
  2   0.419                  XXXXXXXXXXX
  3   0.392                  XXXXXXXXXXX
  4   0.375                  XXXXXXXXXX
  5   0.243                  XXXXXXX
  6   0.198                  XXXXXX
  7   0.195                  XXXXXX
  8   0.222                  XXXXXXX
  9   0.228                  XXXXXXX
 10   0.287                  XXXXXXXX
 11   0.327                  XXXXXXXXX
 12   0.370                  XXXXXXXXXX
 13   0.313                  XXXXXXXXX
 14   0.211                  XXXXXX
 15   0.225                  XXXXXXX
 16   0.155                  XXXXX
 17   0.094                  XXX
 18   0.125                  XXXX
 19   0.119                  XXXX
 20  -0.017                  X
 21   0.009                  X
 22   0.075                  XXX
```

Cont. Appendix (4):

A4.14: Cross Correlation Function: COD mg/l; T-N mg/l in Zarka River

CCF - correlates COD mg/l(t) and T-N mg/l(t+k)

```
          -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
            +----+----+----+----+----+----+----+----+----+----+
-22  0.299                          XXXXXXXX
-21  0.319                          XXXXXXXXX
-20  0.324                          XXXXXXXXX
-19  0.371                          XXXXXXXXXX
-18  0.392                          XXXXXXXXXXX
-17  0.329                          XXXXXXXXX
-16  0.351                          XXXXXXXXXX
-15  0.315                          XXXXXXXXX
-14  0.336                          XXXXXXXXX
-13  0.408                          XXXXXXXXXXX
-12  0.403                          XXXXXXXXXXX
-11  0.380                          XXXXXXXXXXX
-10  0.374                          XXXXXXXXXX
 -9  0.410                          XXXXXXXXXXX
 -8  0.455                          XXXXXXXXXXXX
 -7  0.480                          XXXXXXXXXXXXX
 -6  0.460                          XXXXXXXXXXXXX
 -5  0.469                          XXXXXXXXXXXXX
 -4  0.472                          XXXXXXXXXXXXX
 -3  0.463                          XXXXXXXXXXXXX
 -2  0.482                          XXXXXXXXXXXXX
 -1  0.567                          XXXXXXXXXXXXXXX
  0  0.561                          XXXXXXXXXXXXXXX
  1  0.489                          XXXXXXXXXXXXX
  2  0.461                          XXXXXXXXXXXXX
  3  0.473                          XXXXXXXXXXXXX
  4  0.447                          XXXXXXXXXXXX
  5  0.458                          XXXXXXXXXXXX
  6  0.422                          XXXXXXXXXXXX
  7  0.352                          XXXXXXXXXX
  8  0.351                          XXXXXXXXXX
  9  0.332                          XXXXXXXXX
 10  0.369                          XXXXXXXXXX
 11  0.420                          XXXXXXXXXXXX
 12  0.408                          XXXXXXXXXXX
 13  0.294                          XXXXXXXX
 14  0.282                          XXXXXXXX
 15  0.288                          XXXXXXXX
 16  0.307                          XXXXXXXXX
 17  0.269                          XXXXXXXX
 18  0.245                          XXXXXXX
 19  0.129                          XXXX
 20  0.106                          XXXX
 21  0.082                          XXX
 22  0.128                          XXXX
```

Cont. Appendix (4):

A4.15: Cross Correlation Function: T-P mg/l; T-N mg/l in Zarka River

```
CCF - correlates T-P    mg/l(t) and T-N    mg/l(t+k)

          -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
            +----+----+----+----+----+----+----+----+----+----+
-22   0.276                      XXXXXXXX
-21   0.251                      XXXXXXX
-20   0.319                      XXXXXXXXX
-19   0.300                      XXXXXXXXX
-18   0.274                      XXXXXXXX
-17   0.256                      XXXXXXX
-16   0.273                      XXXXXXXX
-15   0.252                      XXXXXXX
-14   0.302                      XXXXXXXXX
-13   0.382                      XXXXXXXXXXX
-12   0.391                      XXXXXXXXXXX
-11   0.458                      XXXXXXXXXXXX
-10   0.443                      XXXXXXXXXXXX
 -9   0.467                     XXXXXXXXXXXXX
 -8   0.468                     XXXXXXXXXXXXX
 -7   0.514                     XXXXXXXXXXXXXX
 -6   0.459                     XXXXXXXXXXXX
 -5   0.468                     XXXXXXXXXXXXX
 -4   0.438                     XXXXXXXXXXXX
 -3   0.486                     XXXXXXXXXXXXX
 -2   0.536                     XXXXXXXXXXXXXX
 -1   0.620                     XXXXXXXXXXXXXXXXX
  0   0.688                     XXXXXXXXXXXXXXXXXX
  1   0.628                     XXXXXXXXXXXXXXXXX
  2   0.604                     XXXXXXXXXXXXXXXX
  3   0.559                     XXXXXXXXXXXXXXX
  4   0.530                     XXXXXXXXXXXXXX
  5   0.518                     XXXXXXXXXXXXXX
  6   0.456                     XXXXXXXXXXXX
  7   0.432                     XXXXXXXXXXXX
  8   0.382                     XXXXXXXXXXX
  9   0.382                     XXXXXXXXXXX
 10   0.427                      XXXXXXXXXXXX
 11   0.435                      XXXXXXXXXXXX
 12   0.457                      XXXXXXXXXXXX
 13   0.429                      XXXXXXXXXXXX
 14   0.383                      XXXXXXXXXXX
 15   0.365                      XXXXXXXXXX
 16   0.340                      XXXXXXXXXX
 17   0.327                      XXXXXXXXX
 18   0.284                      XXXXXXXX
 19   0.241                      XXXXXXX
 20   0.212                      XXXXXX
 21   0.178                      XXXXX
 22   0.231                      XXXXXXX
```

Cont. Appendix (4):

A4.16: Distance Correlation Function: Zarka Flow MCM/month; Samra Flow MCM/month

CCF - correlates Zarka Flow MCM/month(t) and Samra Flow MCM/month(t+k)

```
       -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
         +----+----+----+----+----+----+----+----+----+----+
-22  -0.101                 XXXX
-21  -0.128                 XXXX
-20  -0.118                 XXXX
-19  -0.113                 XXXX
-18  -0.124                 XXXX
-17  -0.120                 XXXX
-16  -0.125                 XXXX
-15  -0.113                 XXXX
-14  -0.109                 XXXX
-13  -0.086                  XXX
-12  -0.086                  XXX
-11  -0.099                  XXX
-10  -0.112                 XXXX
 -9  -0.118                 XXXX
 -8  -0.119                 XXXX
 -7  -0.130                 XXXX
 -6  -0.116                 XXXX
 -5  -0.108                 XXXX
 -4  -0.098                  XXX
 -3  -0.075                  XXX
 -2  -0.042                   XX
 -1  -0.051                   XX
  0  -0.013                    X
  1  -0.024                   XX
  2  -0.041                   XX
  3  -0.055                   XX
  4  -0.040                   XX
  5  -0.031                   XX
  6  -0.035                   XX
  7  -0.040                   XX
  8  -0.035                   XX
  9  -0.042                   XX
 10  -0.015                    X
 11  -0.017                    X
 12  -0.042                   XX
 13  -0.036                   XX
 14  -0.031                   XX
 15  -0.034                   XX
 16  -0.033                   XX
 17  -0.041                   XX
 18  -0.038                   XX
 19  -0.045                   XX
 20  -0.039                   XX
 21  -0.032                   XX
 22  -0.003                    X
```

Cont. Appendix (4):

A4.17: Distance Correlation Function: TSS Zarka River mg/l; TSS Samra mg/l

```
CCF - correlates TSS      mg/l(t) and TSS  Samra     mg/l_1(t+k)

        -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
          +----+----+----+----+----+----+----+----+----+----+
-22   0.059                      XX
-21  -0.058                     XX
-20  -0.207                  XXXXXX
-19  -0.259                 XXXXXXX
-18  -0.202                  XXXXXX
-17  -0.150                   XXXXX
-16  -0.190                  XXXXXX
-15  -0.111                    XXXX
-14  -0.023                     XX
-13   0.035                      XX
-12   0.039                      XX
-11  -0.032                     XX
-10  -0.030                     XX
 -9  -0.205                 XXXXXX
 -8  -0.315               XXXXXXXXX
 -7  -0.368              XXXXXXXXXX
 -6  -0.272               XXXXXXXX
 -5  -0.281               XXXXXXXX
 -4  -0.162                 XXXXX
 -3  -0.020                    XX
 -2   0.114                     XXXX
 -1   0.225                     XXXXXXX
  0   0.302                     XXXXXXXXX
  1   0.268                     XXXXXXXX
  2   0.131                     XXXX
  3  -0.036                    XX
  4  -0.120                  XXXX
  5  -0.183                 XXXXXX
  6  -0.347              XXXXXXXXXX
  7  -0.297               XXXXXXXX
  8  -0.130                  XXXX
  9  -0.029                    XX
 10  -0.018                     X
 11   0.090                      XXX
 12   0.137                      XXXX
 13   0.164                      XXXXX
 14   0.041                      XX
 15  -0.103                   XXXX
 16  -0.143                  XXXXX
 17  -0.209                 XXXXXX
 18  -0.155                  XXXXX
 19  -0.165                  XXXXX
 20  -0.139                   XXXX
 21  -0.018                      X
 22   0.136                      XXXX
```

Cont. Appendix (4):
A4.18: Distance Correlation Function: BOD5 in Zarka mg/l; BOD5 Samra mg/l

CCF - correlates BOD5 mg/l(t) and BOD5 Samra mg/l_1(t+k)

```
         -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
           +----+----+----+----+----+----+----+----+----+----+
-22  -0.000                        X
-21   0.039                        XX
-20   0.101                        XXXX
-19   0.096                        XXX
-18   0.179                        XXXXX
-17   0.152                        XXXXX
-16   0.193                        XXXXXX
-15   0.173                        XXXXX
-14   0.093                        XXX
-13   0.028                        XX
-12  -0.016                        X
-11  -0.043                       XX
-10   0.008                        X
 -9   0.037                        XX
 -8   0.047                        XX
 -7   0.215                        XXXXXX
 -6   0.214                        XXXXXX
 -5   0.273                        XXXXXXXX
 -4   0.282                        XXXXXXXX
 -3   0.211                        XXXXXX
 -2   0.192                        XXXXXX
 -1   0.110                        XXXX
  0   0.188                        XXXXXX
  1   0.019                        X
  2   0.053                        XX
  3   0.100                        XXX
  4   0.138                        XXXX
  5   0.145                        XXXXX
  6   0.216                        XXXXXX
  7   0.267                        XXXXXXXX
  8   0.245                        XXXXXXX
  9   0.211                        XXXXXX
 10   0.084                        XXX
 11   0.095                        XXX
 12   0.030                        XX
 13   0.070                        XXX
 14   0.062                        XXX
 15   0.067                        XXX
 16   0.205                        XXXXXX
 17   0.185                        XXXXXX
 18   0.249                        XXXXXXX
 19   0.297                        XXXXXXXX
 20   0.279                        XXXXXXXX
 21   0.245                        XXXXXXX
 22   0.186                        XXXXXX
```

Cont. Appendix (4):

A4.19: Distance Correlation Function: COD in Zarka River mg/l; COD Samra mg/l

```
CCF - correlates COD    mg/l(t) and COD  Samra    mg/l_1(t+k)

          -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
            +----+----+----+----+----+----+----+----+----+----+
-22  -0.024                     XX
-21   0.006                      X
-20   0.029                      XX
-19   0.124                      XXXX
-18   0.139                      XXXX
-17   0.088                      XXX
-16   0.040                      XX
-15   0.064                      XXX
-14  -0.011                      X
-13   0.024                      XX
-12   0.105                      XXXX
-11   0.037                      XX
-10   0.061                      XXX
 -9   0.172                     XXXXX
 -8   0.220                     XXXXXX
 -7   0.205                     XXXXXX
 -6   0.173                     XXXXX
 -5   0.153                     XXXXX
 -4   0.148                     XXXXX
 -3   0.088                     XXX
 -2   0.082                     XXX
 -1   0.090                     XXX
  0   0.075                     XXX
  1   0.023                     XX
  2   0.085                     XXX
  3   0.193                     XXXXXX
  4   0.135                     XXXX
  5   0.118                     XXXX
  6   0.127                     XXXX
  7   0.084                     XXX
  8   0.060                     XXX
  9   0.015                     X
 10  -0.007                      X
 11  -0.067                    XXX
 12  -0.037                     XX
 13   0.011                      X
 14   0.013                      X
 15   0.108                      XXXX
 16   0.187                      XXXXXX
 17   0.205                      XXXXXX
 18   0.179                      XXXXX
 19   0.178                      XXXXX
 20   0.063                      XXX
 21   0.039                      XX
 22  -0.083                    XXX
```

Cont. Appendix (4):

A4.20: Distance Correlation Function: T-P in Zarka River mg/l; T-P Samra mg/l

CCF - correlates T-P mg/l(t) and T-P Samra mg/l_1(t+k)

```
         -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
           +----+----+----+----+----+----+----+----+----+----+
-22   0.115                      XXXX
-21  -0.006                      X
-20   0.023                      XX
-19   0.076                      XXX
-18   0.078                      XXX
-17   0.071                      XXX
-16   0.074                      XXX
-15   0.062                      XXX
-14   0.112                      XXXX
-13   0.171                      XXXXX
-12   0.201                      XXXXXX
-11   0.121                      XXXX
-10   0.145                      XXXXX
 -9   0.200                      XXXXXX
 -8   0.156                      XXXXX
 -7   0.134                      XXXX
 -6   0.088                      XXX
 -5   0.110                      XXXX
 -4   0.127                      XXXX
 -3   0.097                      XXX
 -2   0.094                      XXX
 -1   0.180                      XXXXX
  0   0.279                      XXXXXXXX
  1   0.309                      XXXXXXXXX
  2   0.199                      XXXXXX
  3   0.074                      XXX
  4   0.040                      XX
  5   0.040                      XX
  6  -0.009                      X
  7   0.059                      XX
  8   0.015                      X
  9  -0.003                      X
 10   0.072                      XXX
 11   0.162                      XXXXX
 12   0.162                      XXXXX
 13   0.094                      XXX
 14   0.051                      XX
 15   0.018                      X
 16  -0.001                      X
 17  -0.010                      X
 18   0.061                      XXX
 19   0.022                      XX
 20  -0.002                      X
 21  -0.035                     XX
 22   0.049                      XX
```

Cont. Appendix (4):

A4.21: Distance Correlation Function: T-N in Zarka River mg/l; T-N Samra mg/l

CCF - correlates T-N mg/l(t) and T-N Samra mg/l_1(t+k)

```
           -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
             +----+----+----+----+----+----+----+----+----+----+
-22   0.049                        XX
-21   0.034                        XX
-20   0.050                        XX
-19   0.037                        XX
-18   0.032                        XX
-17   0.012                        X
-16   0.006                        X
-15   0.034                        XX
-14   0.033                        XX
-13   0.011                        X
-12   0.016                        X
-11   0.065                        XXX
-10   0.117                        XXXX
 -9   0.101                        XXXX
 -8   0.067                        XXX
 -7   0.074                        XXX
 -6   0.097                        XXX
 -5   0.110                        XXXX
 -4   0.116                        XXXX
 -3   0.094                        XXX
 -2   0.102                        XXXX
 -1   0.101                        XXXX
  0   0.222                        XXXXXXX
  1   0.221                        XXXXXXX
  2   0.162                        XXXXX
  3   0.129                        XXXX
  4   0.125                        XXXX
  5   0.078                        XXX
  6   0.086                        XXX
  7   0.103                        XXXX
  8   0.096                        XXX
  9   0.128                        XXXX
 10   0.126                        XXXX
 11   0.143                        XXXXX
 12   0.171                        XXXXX
 13   0.153                        XXXXX
 14   0.118                        XXXX
 15   0.046                        XX
 16   0.057                        XX
 17   0.063                        XXX
 18   0.085                        XXX
 19   0.094                        XXX
 20   0.109                        XXXX
 21   0.095                        XXX
 22   0.116                        XXXX
```

REFERENCES

Anderson T. W., 1971. *The Statistical Analysis of Time Series*. Wiley. USA

Berthouex, P.M. and Box, G.E. 1996. Time Series Models For Forecasting Wastewater Treatment Plant Performance. *Jour. Water Research* 30, No. 8: 1865-1875.

Boyd C. 2000. *Water Quality: An Introduction*. Kluwer Academic Publishers, Boston-USA.

Chatfield C. 1984. *The Analysis of Time Series: An Introduction. 3rd edition.* Chapman and hall, New York – USA.

Chatfield. C. 1978. *Statistics for Technology A Coarse in Applied Statistics, 2nd edition*. John Wiley & Sons, New York – USA.

Clinton I. Chase. 1976. *Elementary Statistical Procedure, 2nd edition.* McGraw-Hill Book Company, New Delhi – India.

Conrads, P.A., 1998, Effects of model output time-averaging on the determination of the assimilative capacity of the Waccamaw River and Atlantic Intracoastal Waterway near Myrtle Beach, South Carolina. Subcommittee on Hydrology of the Interagency Advisory Committee on Water *Data*, p. 2-93 to 2-100.

Elias Salameh. 1996. *Water Quality Degradation in Jordan: Impact on Environment, Economy and Future Generations Resources Base*. Jordan.

Ellis, G.W., Ge, X. and Grasso, D. 1990. Time Series Analysis of Wastewater Quality. Proceedings of the fifth IAWPRC Workshop, Japan, 26 July-3 August. NewYork: Pergamon Press: 441-448.

General Statistical Department. *The Annual Statistical Book for the year 1999*. G.S.D 2000.

Green J.R, and Magerison D 1978. *Statistical Treatment of Experimental data*. New York-USA.

Hammett, K.M. 1988. Land Use, Water Use, Streamflow, and Water Quality Characteristics of the Carlotte Harbor Inflow Area, Florida. USGS Open Files. Report: 87-472.

Hare, S.R. and R.C. Francis. 1994. Climate change and salmon production in the northeast Pacific Ocean. Publ. Fish. Aquat. Sci. 121, p. 357-372.

Koning N. , and Roos JC 1999. The continued influence of organic pollution on the water quality of the turbid Modder River. Water SA Vol. 25 No. 3. P:285-292

LaValle, P.D., Lakhan, V.C., and Trenhaile A.S. 2000. Short Term Fluctuations of Lake Erie Wate Levels and the El Nio/Southern Oscillation. The Great Lakes Geographer, Vol.7, No.1

Levine D. M., Patricia P. Ramsey, and Smidt R. K.. 2001. *Applied Statistics for Engineers and Scientists Using Microsoft Excel and Minitab.* Prentice-Hall, New Jersey-USA.

Lincoin L. Chao, 1974. *Statistics Methods and Analyses, 2nd edition*. McGraw-Hill Kogukusha, LTD, California-USA.

Lu Guanghua, Tang Jie, Yuan Xing, and Zhao Yuanhui, 2001. Correlation for the structure and biodegradability of substituted benzenes in the Songhua river water. The National Natural Science Foundation of China 29877004. Vol.3 No.7 P.34.

McBean E. A., and Rovers F. A. 1998. *Statistical Procedures for Analysis of Environmental Monitoring Data and Risk Assessment*. Prentice-Hall, New Jersey-USA.

McLeod A.I. & Hipel K.W., 1994. Exploratory Spectral Analysis of Hydrological Time Series. Ontario, Canada. *www.stats.vwo.ca*

Metcalf and Eddy, 1991. *Wastewater Engineering: Treatment, Disposal, and Reuse*, *3rd edition*. McGraw-Hill, Inc., Singapore.

Miao-Hsiang P. and Jin-King L., 2000. Groundwater Level Forecasting with Time Series Analysis. HSIN-Chu, Taiwan, *www.gisdevelopment.net*

Ministry of Water and Irrigation, Water Authority of Jordan, Operational Manual, Amman- Jordan, 1998.

Montgomery, D., and Johnson L. 1976. *Forecasting and Time Series Analysis*. McGraw-Hill Book Company, Japan.

Phil Spector. 1994. *An Introduction to S and S-Plus*. Duxbury Press an Imprint of Wadsworth Publishing Company. Belmont, California.

Pillai R.S.N. and Bagavathi V. 1997. *Statistics: Theory and Practice. 9th edition*. Ram Nagar, New Delhi – India.

Royal Scientific Society, *Monitoring the Water in King Talal Dam*. Amman-Jordan. RSS Reports 1988-2001.

Royal Scientific Society. *National Project for Assessment of Water Quality in Jordan*. RSS Report 1998.

Robert B. Miller. 1988. *Minitab Handbook for Business and Economics*. PWS-Kent Publishing Company. Boston-USA.

Roedel, R. 1997. Time Series Analysis of Rising Underground Salt-Water from the Abandoned Werra Potash Mining District. *Jour. Wasser und Boden* 49, No.12: 27-30.

Sakakini A. E., 2001. *Forecasting the Characteristics of Jordanian Domestic Wastewater Using Time Series*, Master Thesis. University of Jordan. Amman-Jordan.

Solomatine D. P., Rojas C. J., Velickov S. and Wüst J. C., 2000. Chaos theory in predicting surge water levels in the North Sea. Proc. 4-th International Conference on Hydroinformatics, Iowa, USA, July 2000. p. 1-8. *www.ihe.nl/hi/sol /papers/HI2000_ chaos.pdf.*

Stegmann, R., Ehrig, H. and Liem, P. 1978. The Application of a Time-Series-Analysis in Water Quality Management. *Jour. Wasser und Boden* 30, No. 3: 50-54.

Stien M., and Lloret J. 1999. Forecasting of air and water temperatures for fishery with selected examples from northwest Atlantic. J. Northw. Atl. Fish. Sci., Vol. 29: 23-30

Warren Viessman, Jr. and Gray L. Lewis, 1996. *Introduction to Hydrology, 4th edition*. Addison-Wesley Educational Publishers, Inc., USA.

Warren Viessman, Jr. and Mark J. Hammer, 1985. *Water Supply and Pollution Control, 4th edition*. Harper and Roe Publishers, Inc., Florida-USA.

التحليل والتنبؤ بكمية ونوعية المياه الداخلة إلى سد الملك طلال

إعداد
أحمد عبدالرزاق أحمد بني هاني

المشرف
الدكتور أحمد الجمرة

مـــلـــخــــــــــص

يعتبر سد الملك طلال من أهم المشاريع المائية التي تم إنشائها لزراعة مناطق شاسعة في وادي الأردن. و بالتالي فإن دراسة نوعية و كمية المياه في سد الملك طلال و المياه التي تصب فيه يجب أن تكون من الأولويات في الأردن.

تتركز هذه الدراسة في دراسة و تنبؤ نوعية و كمية المياه في نهر الزرقاء و الذي يعتبر الرافد الرئيسي لسد الملك طلال.عبّر عن كمية المياه المستعملة في هذا البحث بالتدفق, ونوعية المياه بالمواد العالقة الكلية, BOD_5, COD, الفسفور الكلي, و النيتروجين الكلي في نهر الزرقاء. البيانات التي جمعت لكل متغير سجلت خلال ١٥٦ شهراً من العام ١٩٨٨ و حتى نهاية العام ٢٠٠٠. الطريقة المستخدمة في تحليل الستة متغيرات في نهر الزرقاء هي من خلال علاقة المتغير بنفسه (autocorrelation) و المتغير مع متغير أخر عند نقطة معينة (cross correlation) و علاقة المتغير مع نفسه و لكن بعد مسافة افقية معينة (distance correlation). و قد تم استخدام التنبؤ الحتمي (deterministic) و الإحتمالي (stochastic) لستة متغيرات لإيجاد أفضل نموذج للتنبؤ.

نتائج الدراسة مؤشر إلى أن نموذج ARIMA نموذجاً جيداً في التنبؤ بمعظم الستة متغيرات. في تنبؤ قيم ال BOD_5 لم يحقق أي من النماذج أقل من ١٠% من خطأ المتوسط الحسابي, ومع ذلك فأن نموذج ARIMA أعطى أفضل نموذج وأقل نسبة من خطأ المتوسط الحسابي, أما في قيم ال COD فإن نموذج ARIMA لم يعطِ أفضل النتائج. أقل نسبة من خطأ المتوسط الحسابي, والتي أعطيت من خلال نماذج ARIMA, كانت تعادل نسبة ٤,٨% في قيم الفسفور الكلي. أما في علاقة المتغيرات ببعض ففد تمت معرفة مدى ترابط المتغيرات ببعض, هيئة المتغيرات في نهر الزرقاء, مصدر المتغيرات, و معلومات أخرى عنها.