American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-7, Issue-6, pp-190-199
www.ajer.org
Research Paper Open Access

# Frequency Analysis of Some of Water Quality Parameters of Shatt Al-Hilla River, Iraq

Prof. Dr. Rafea H Al-Suhili
*Prof. at the City College of New York, Cuny, New York, USA.*
*Corresponding Author: Prof. Dr. Rafea*

***ABSTRACT:*** *Different frequency distributions models were fitted to the monthly data of some of the water quality parameters of Shatt Al-Hilla River located, south Iraq. Five water quality parameters were investigated for the period (2000-2013). The data collected are the monthly series for Magnesium, Alkalinity, Total Hardness, Sulphate and Calcium. The distributions models fitted are the Normal, Log-normal, Weibull, Exponential and Two parameters Gamma type. The Kolmogorov-Smirnov test was used to evaluate the goodness of fit. The fittingswere done for the data period, of 9-years (2000-2008). The best fitted distributions for Magnesium, Alkalinity, Total Hardness, are the Gamma, Log-Normal, and the Normal distributions, respectively. No distribution of the tried ones can fit for Sulphate and Calcium. The best fitted distributions for the first three parameters were used to forecast three sets of monthly data for each one and compared to the observed ones for the last 5- years of data(2009-2013). The Kolmogorov-Smirnov test results indicate the capability of these models to produce data that has the same frequency distribution of the observed one.Lag-Correlation analysis was done for all of the variables to investigate the reasons of the non-fit of Sulphate and Calcium. This analysis indicates that these two variables exhibit relatively strong lag-correlation than the other three variables and hence low randomness behavior.*

***KEYWORDS:*** *Frequency Distribution, Kolmogorov-Smirnov test, Shatt Al-Hilla River, Iraq,Water quality parameters,Forecasting*

---

Date of Sumisión: 1-06-2018 Date of aceptante: 16-06-2018

---

## I. INTRODUCTION

The identification of the range and the likelihood of water quality parameters of a certain water bodies like rivers and lakes, can provide the designers, planners and decision maker's useful guides to prepare for and deal with the consequences of water quality problems. Models of water quality parameters probability distributions over various timescales can provide useful information for proper planning, operation and design of water quality enhancement projects.The development of a frequency model starts with the historical data acquisition. These collected data sample can provide the necessary information about the time evolution variability of these water quality variables if the size of the collected measured sample is of representative size. The available samples of these historical data should be divided into two sub-samples, the first one should be used for the estimation of the parameters,while the second subsample should be used for verification.When the parameters of the best fitted frequency distribution are estimated and the model is verified, then the model results could be considered dependable and can be used for variety of applications in water resources systems planning and design for quality assurance. Among these applications for example,locating areas of potential low water quality, feasibility of constructing water treatment plants, …..etc. Different probability distributions could be found in literature, such as normal, log-normal, Weibull, exponential, Gamma, log-Person and others. When the parameters of any distribution are found , then the probability of any event r>ro, where ro is a threshold value selected by the analyzer ,could be found, and hence provide useful information about the most probable events that could occur and plan and design to overcome it's consequences. Many researches had been conducted to find the probability distribution function parameters of water quality parameters.

Hassan et al.(2008), had studied the chemical and physical properties in the ecological system of Shatt Al-Hilla in Babylon Governorate in Iraq and it's effect on phytoplankton population. In this context, several limn logical parameters were evaluated during the period from December 2003 through November 2004 from four sampling stations sited along Shatt Al-Hilla River. The physical parameters included: are temperature, turbidity and electrical conductivity. The chemical parameters included: are pH, alkalinity dissolved oxygen, total hardness and the concentrations of nitrite, nitrate, phosphate and Sulphate. A total of 154 species were recorded. Ninety-seven species of the total belong to Bacillariophyceae, 37 species belong to Chlorophyceae, 13 species to Cyanophyceae, 5 species to Chrysophyceae, and 2 species to Euglenophyceae. Bimodal variation of phytoplankton was observed. Five genus of phytoplankton were the highest number of species ( Nitzschia, Navicula, Gomphonema, Cymbella and Scendesmus). Some species was occurred continuously during study period such as, Cyclotellaocellata, Cyclotellameneghiniana, Aulacoseriadistans, and Gomphonemaabbreviatum. The phytoplankton communities at all sampling sites showed a clear seasonal variation in phytoplankton cell number. However, no significant correlation between total cell number of phytoplankton and nutrient concentration was observed. The study was revealed the city sewage discharge, agriculture and urban run-off were affecting the water quality of Shatt Al-Hilla River.

Obais and Al-Fatlawi(2012) had developed a study of the water pollution for Shatt Al-HillaRiver in Middle Euphrates region of Iraq (Babylon and Al-Diwaniya governorates) occurs in both rural and urban areas. An environmental database was constructed and applied. To evaluate the pollutant concentrations, regression models were obtained by Data Fit Software program (version 8.0). The results were compared with the Iraqi and WHO standards for domestic and irrigation purposes to determine pollution extend and suggest suitable solutions. The results of the program are verified with data of year 2008 which is not included in regression model. This verification shows a good agreement with coefficient of determination ranged between "0.927 to 0.996".

Jehad(2014) had conducted a study to assess the quality of water of the Euphrates River (Shatt Al-Hillah) for civilian use (drinking water) and agricultural (irrigation water). Models collected on a monthly basis for a period of two years 2010 - 2011 of four stations on the Euphrates River in the province of Babylon, stations are Hindia barrage, Musayyib, Hilla and Al- Hashemeyeh. Samples taken by repeat each month and adopted average. The results showed that the quality of water of the Euphrates River (Shatt Al- Hillah) for civilian use meets the requirements of standard specification by the Iraqi high and the requirements of WHO, with the exception of sulfates, while failing to achieve the requirements of the Environmental Protection Agency and the U.S. EPA. As for the quality of water for agricultural purposes, the study has shown that there is problem salinity by different classifications, while there is no problem regarding the risk of danger sodium chlorine, the water is considered mild -to-moderate damage.

Shat Al-HillaRiver is located about 100 Km south of Baghdad city, the capital of Iraq. Recently continuous developments of water resources projects are under design and planning phases. Due to the reduced surface water that coming from near countries of Turkey and Iran. The water quality of almost all Iraqi rivers exhibits lower water quality. This led to the need of studies of the required actions that is to be taken to reduce the water quality deterioration of these waters. Providing models to estimate the probable expected water quality parameters of this river, will be a useful tool for better studies and planning of the required actions to overcome this situation. This paper is one of the contributions that could provide future estimation of the probability of some of the water quality parameters of Shat Al-Hilla River,and its results could be useful for those studies for planners and decision makers of these projects. Hence this river was selected for the present analysis.

## II. THEORETICAL ASPECTS OF FREQUENCY DISTRIBUTIONS:

The estimation of the probabilities of any measured variableneeds the mathematical definition of the probability density function(PDF) of the selected distribution. The mathematical description of these density functions could be found in the literature. Followings are only those used in this paper.

1- Normal Distribution:

For a variable x normally distributed the probability density function PDF is given as below:

$$f(x|\sigma,\mu)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (1)$$

Where: μ is the mean value and σ is the standard deviation, called as the Location and Scale parameters of the distribution, respectively.

2- Log-Normal Distribution:

For a variable x that Log(x) is normally distributed the PDF is:

$$f(x|\sigma,\mu) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(logx-\mu)^2}{2\sigma^2}} \quad (2)$$

Where: μ is the mean value of log(x) and σ is the standard deviation of this variable, called as the Scale and Shape parameters of the distribution,respectively.

3- Exponential Distribution:

For a variable x that distributed exponentially the PDF is:

$$f(x|\mu) = \frac{1}{\mu} e^{\frac{x}{\mu}} (3)$$

Where, μ is the Scale parameter of the distribution.

4- Weibull Distribution:

The PDF of this distribution is given by:

$$f(x|a,b) = \frac{a}{b^{-a}} x^{a-1} e^{-(\frac{x}{b})^a} \quad (4)$$

Where: a, and b are the Shape and Scale parameters respectively.

5- The Gamma Distribution PDF is:

$$f(x|a,b) = \frac{1}{b^a} (\frac{1}{r(a)})\, x^{a-1} e^{(\frac{x}{b})} \quad (5)$$

Where: a and b are the Shape and Scale parameters respectively, and $r(a)$ is the Gamma Function.

The parameters estimation methods of these distributions are well known and could be found in relevant text books. The most used two methods are the moments and the maximum likelihood methods. The parameters were estimated in this paper using the SPSS Software , Version 20.

## III. DATA AND METHODS

Al-Hilla River lies in the city of Babylon, it is branched out from Euphrates River, at Al-Hindiya dam. Shatt Al-Hilla situates in Hillacity, 100 Km south of Baghdad city in Iraq. It is the main channel that is branched from the left side of Euphrates River, just upstream the Hindiya Barrage. It is the largest channel withdrawing water from the pool upstream the Hindiya Barrage. The reach considered of Shatt Al-Hilla has length of 33 Km and starts from the intake point [head regulator of Shatt AlHilla at Km (00+000)] to the gauge station in Hilla city at Km (33+000), Alghazali et al. (2013).

The models were applied to the data of the case study described above. The length of record for the five water quality parameters of Shat Al-Hilla river, Magnesium, Alkalinity, Total Hardness, Sulphate and Calciumare of (14) years, (2000-2013). The data for the first (9), years(2000-2008) were used for models building, while the last 5years data were used for verification,(2009-2013). It is worth to mention that the data are on monthly basis.Unfortunately there is no updated reliable records of the water quality parameters beyond year 2013.

Table (1) shows the north and East coordinates of the metrological stations selected for analysis and their locations are shown in figure(1).

As mentioned above the data used for estimating the models parameters are those for the 9 years of (2000-2008). The goodness of fit test used in the present analysis is the non-parametric Kolmogorov-Smirnov test, which is as follows:

$$D_n = Max|F_n(y) - F(y)| \quad (6)$$

Where:

$$F_n(y) = \frac{\Pi(\{i \in \{1,2,.,.,n\}: y_i \le y\})}{n} \quad (7)$$

$$F(y) = \int_0^y f(x)dx \quad (8)$$

Where

$F_n(y)$is the empirical cumulative probability of observing a value less than or equal y as shown in equation (7).

F(y) is the theoretical cumulative probability at y estimated by the frequency distribution function f(x), with the estimated parameters. A smaller value of D means a better fit between the observed and theoretical distributions, for a fixed number n of observations.

## IV. RESULTS AND DISCUSSION

The SPSS software gives the SIG. value of the test which is a test parameter if greater than 0.05 then the null hypothesis that the data fits the used distribution is not rejected, otherwise the distribution tried do not fit the data. As the Kolmogrov-Smirnov test value D decreased SIG increased and hence higher SIG values gives better fitting.

A P-P plot (probability-probability plot or percent-percent plot) is a probability plot for assessing how closely two data sets agree, which plots the two cumulative distribution functions against each other.

Figures (2,3,4,5 and 6) shows the fitting P-P plots of the monthly water quality data for the five water quality parameters of Shat Al-Hilla river for the Normal, Log-Normal, Exponential,Weibull and the Gamma frequency distributions,respectively. Those figures are indicative for the best fit distribution to the data of the first 9 years of these parameters (2000-2008). Table (1) shows the parameters estimated for each variable and each distribution, and the SIG values of the Kolmogorov-Smirnov test. The shaded values indicate the maximum SIG value for each parameter which indicates that the best fitted distributions are of Gamma(12.252,0.349) for Magnesium, LN(122.87,0.171) for Alkalinity and Normal(381.62,92.43) for Total Hardness,while non of the distributions can fit for Sulphate and calcium. Moreover it is shown that for Magnesium the test results indicates that three distributions fits the Normal, Log-Normal, and the Gamma since the SIG value for all of these three distribution is >0.05, but the last one was selected since it gives the highest SIG value. Similar observations were found for Alkalinity (Normal, and Log-Normal), and Hardness(Normal, Weibull, and, Gamma). For Sulphate and Calcium all the SIG values are <0.05.

As mentioned above the data available was subdivided into two sub-samples , the second subsample data for 5- years (2009-2013), were left for models verification. For each parameterof those that fits a distribution (Magnesium, Alkalinity, and Hardness), three randomly generated series were obtained using the best fitted distributions found above. Table (2) shows the Kolmogorov-Smirnov independent two samples test results, that checks the best of fit of the generated series with the observed ones.All of the SIG values shown in this table are > 0.05 which indicates that the generated series are of the same frequency distribution of the observed ones. This indicates that these models are capable of generating data series that are probable to occur for these water quality parameters,and hence can be used by planners,decision makers and designer. Table (3) shows a comparison between the means and standard deviations of the observed and generated series for the period (2009-2013),for Magnesium, Alkalinity and Total Hardness. The t-test for means and F-test for variances results shows insignificant differences for all of the generated series. Tables (4) ,(5), and (6) shows the minimum and maximum values and a probability analysis for Magnesium, Alkalinity, and Total Hardness respectively. These results indicate that the models always generate data of a little wider ranges than that for the observed one. The probability analysis indicates the ability of the models to generate data series that give very close probabilities to the obtained ones.

Table (7) shows the serial Lag correlation coefficients up to lag 5, for each of the water quality parameters investigated herein. The results indicate that both Sulphate and Calcium exhibits strong lag correlation behavior, while Magnesium, Alkalinity, and Total Hardness, exhibit weak lag correlation behavior. This is the reason of obtaining good fit probability distribution models for these last three parameters, while none of the tried probability distributions can fit the Sulphate and Calcium data. This is due to the fact that the modeling process of fitting probability distributions is suitable for random variables only, i.e variables that exhibits more random behavior. Table (7) shows also the multiple correlation coefficient between each variable and it's five first lags series. It is obvious that both Sulphate and Calcium data gives high values for multiple correlation coefficient (0.833,and 0.747),respectively, while for the other three variables this coefficient values are relatively low.

## V. CONCLUSIONS

The analysis of the models developed and verified in this paper indicates that the monthly Magnesium ,Alkalinity and Hardness can be fitted by the Gamma,Log-normal, and Normalfrequency distributions, with distribution parameters of (12.252,0.349),(122.87,0.171), and (381.62,92.43),respectively. Non of the tried frequency distributions can fit the data of the Sulphate, and Calcium. The generated sequences of the first three parameters are representative of the real observed values that were occurred in the period left for verification.The Kolmogorov-Smirnov test values for the comparison of the frequency distribution of the three generated series with the observed one indicates that those series has the same frequency distributions as the observed ones. The t-test for differences between means and the F-test for differences between variances of the observed and generated series using the fitted models, indicates that there is no significant differences for both means and variances, since all the test values are less than the corresponding critical test values at the 95% level of significance. The probability analysis of the observed and generated series gives very close results. The serial lag correlation analysis and the estimated multiple correlation coefficients between each variable and it 5 lags series had indicated that both Sulphate and Calcium had exhibited low random behavior and hence could not fit any of the tried frequency distribution. This analysis indicates that the three best fitted variables, Magnesium, Alkalinity and Total Hardness had high random behavior , and hence more suitable for the used method of modeling.

## REFERENCES:


[1]. Hassan F. M.,Kathim N.F., and Hussein F.H.,"Effect of Chemical and Physical Properties of River Water in Shatt Al-Hilla on Photoplankton Communities", Eniveronmental Journal of Chemistry, Vol. 5(2008), Issue 2,pp 323-330.

[2]. Obais A. A., and Al-Fatlawi A. H., "Assessment and Monitoring of Shatt Al-Hilla River in the Middle Euphrates Region", Journal of Babylon University Engineering Science, No. 4,Vol. 20(2012),pp 994-1004.

[3]. Algazali N. O. S., Alkhaddar R.M., and Hadi H.A., "The use of SCADA System in Water Resources Management of Shatt Al-Hilla in Iraq as a Case Study", International Journal of Environmental Monitoring and Analysis, (1),(5),(2013),pp237-247.

[4]. Jehad U.A., "Study of the Quantitave Assessment of the Euphrates River (Shatt Al-Hilla ) for Civil Purposes and Agriculture", International Journal of Civil Engineering and Technology, Vol 5,Issue 1, Jan (2014),pp 35-46.



**Fig. 1 The location of the Shat AL-Hilla river in Iraq.**

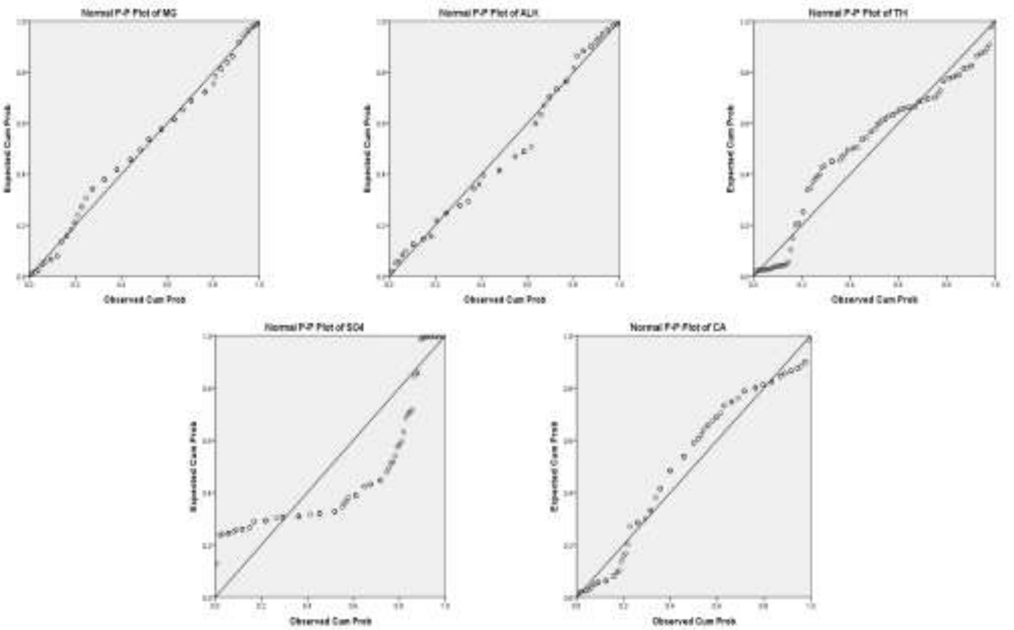



**Fig. 2 P-P plots for Normal Distribution Fitting for Magnesium, Alkalinity, Total Hardness, Sulphate and Calcium, Shatt Al-Hilla water,(2000-2008).**

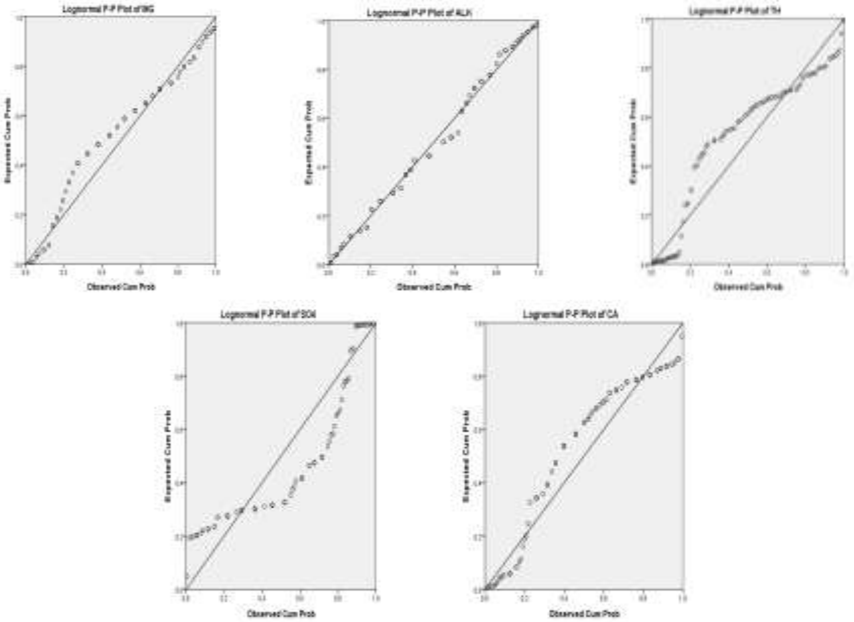
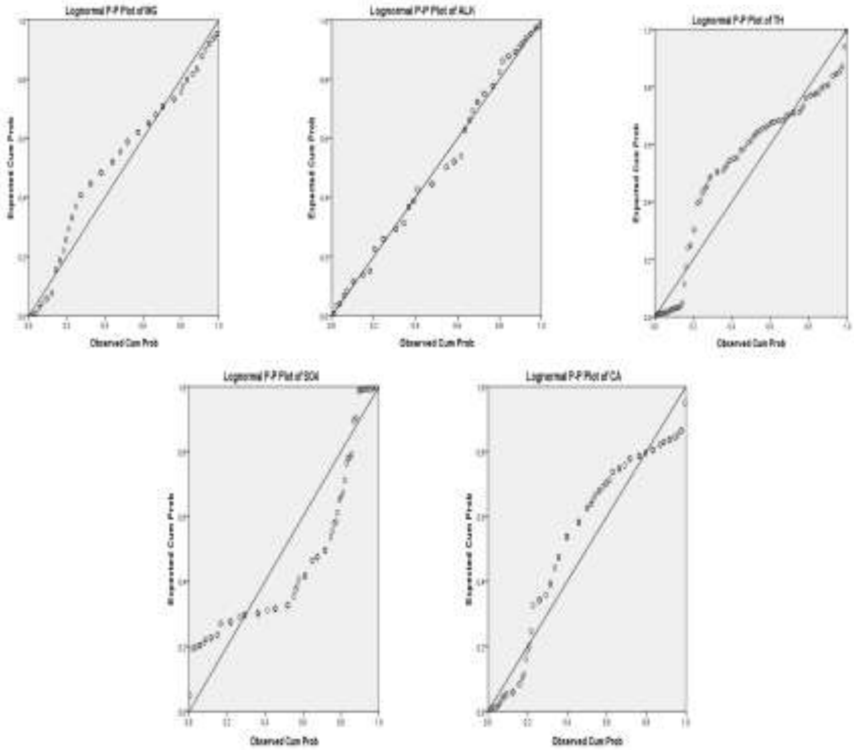



**Fig. 3 P-P plots for Log-Normal Distribution Fitting for Magnesium, Alkalinity, Total Hardness, Sulphate and Calcium, Shatt Al-Hilla water,(2000-2008).**




.

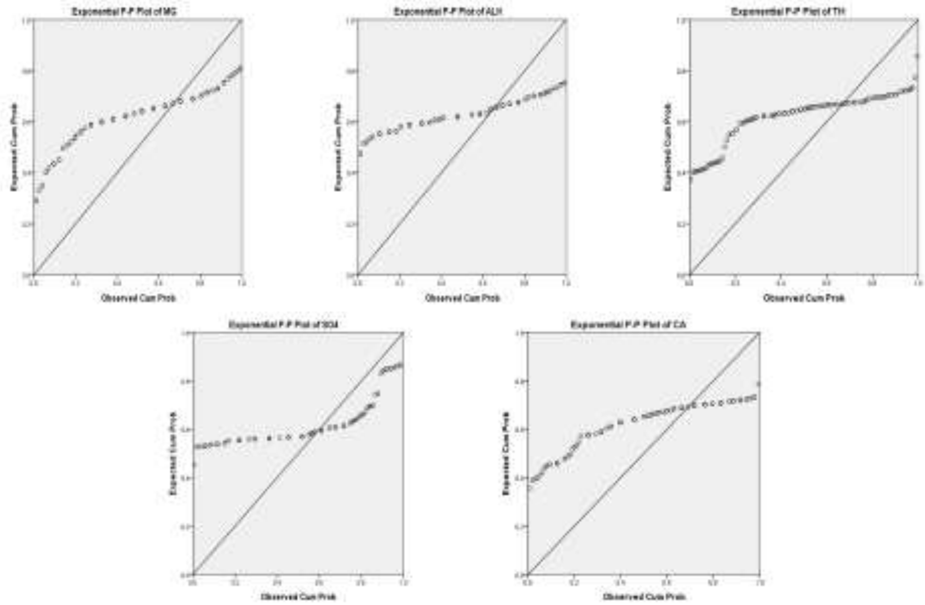



**Fig. 4 P-P plots for Exponential Distribution Fitting for Magnesium, Alkalinity, Total Hardness, Sulphate and Calcium, Shatt Al-Hilla water,(2000-2008).**

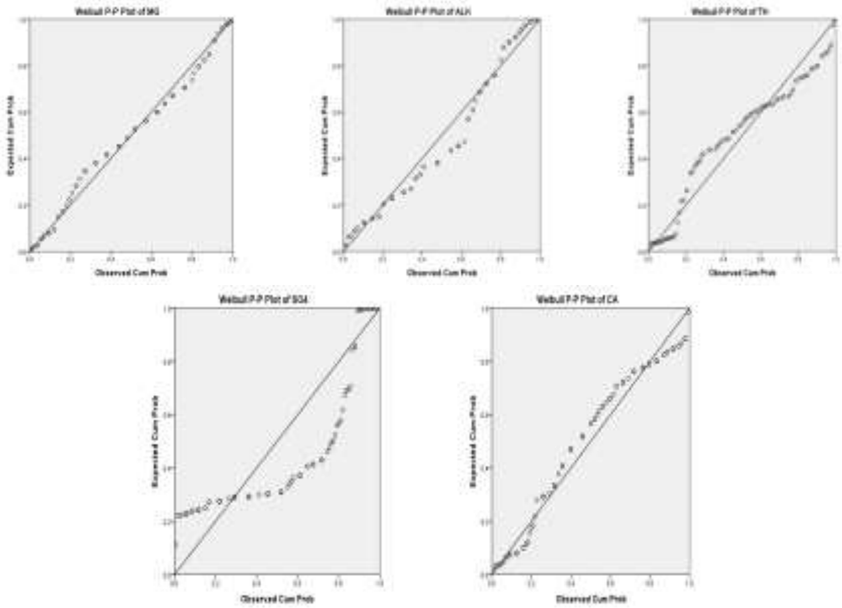


**Fig. 5 P-P plots for Weibull Distribution Fitting for Magnesium, Alkalinity, Total Hardness, Sulphate and Calcium, Shatt Al-Hilla water,(2000-2008).**

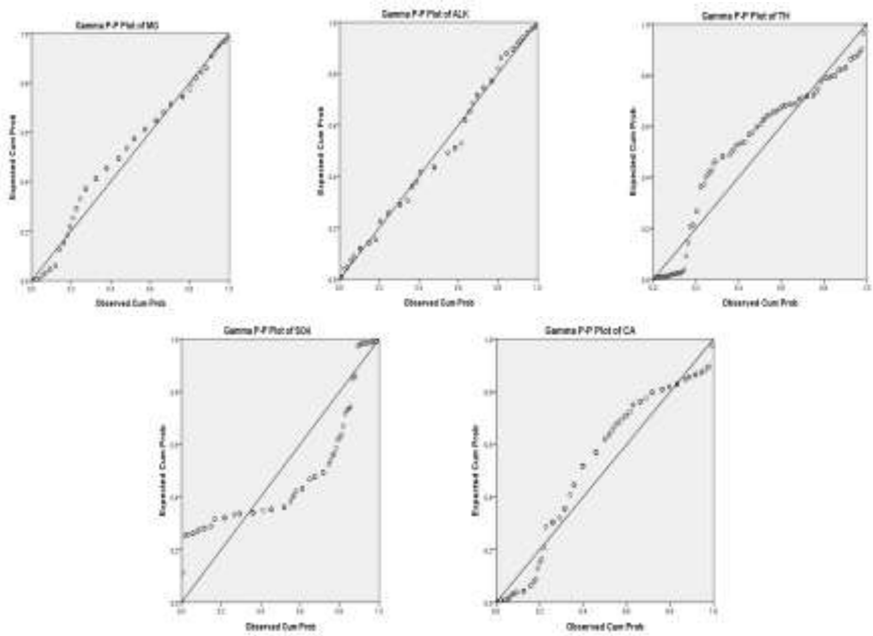


**Fig. 6 P-P plots for Gamma Distribution Fitting for Magnesium, Alkalinity, Total Hardness, Sulphate and Calcium, Shatt Al-Hilla water,(2000-2008).**

**Table 1. Estimated Parameters and SIG values for Kolmogorov-Smirnove Test for the Selectedfive Water Quality Parametersof Shatt Al-Hilla River, Iraq.**

| Distribution | Parameters | Magnesium | Alkalinity | Total Hardness | Sulphate | Calcium |
|---|---|---|---|---|---|---|
| Normal | µ | 35.092 | 124.67 | 381.62 | 290.75 | 90.9 |
| | σ | 10.025 | 21.42 | 92.43 | 102.06 | 22.69 |
| | SIG | 0.1 | 0.187 | 0.25 | 0.000 | 0.01 |
| Log Normal | µ | 33.458 | 122.87 | 368.89 | 277.88 | 87.61 |
| | σ | 0.328 | 0.171 | 0.275 | 0.281 | 0.286 |
| | SIG | 0.138 | 0.25 | 0.01 | 0.000 | 0.01 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Weibull | a | 38.9 | 133.19 | 420.36 | 328.94 | 100.126 |
| | b | 3.77 | 7.04 | 4.731 | 3.373 | 4.267 |
| | SIG | 0.187 | 0.01 | 0.071 | 0.000 | 0.01 |
| Exponen-tial | μ | 0.028 | 0.008 | 0.003 | 0.003 | 0.011 |
| | SIG | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Gamma | a | 12.252 | 33.865 | 17.046 | 8.116 | 16.05 |
| | b | 0.349 | 0.272 | 0.045 | 0.028 | 0.177 |
| | SIG | 0.744 | 0.004 | 0.049 | 0.000 | 0.01 |

**Table 2 Kolmogorov-Smirnov Test SIG values for three generated series of the three Water quality parameters of Shatt Al-Hilla River, using the frequency distribution models and that observed for years(2009-2013).**

| Generated Series No. | MgGamma(12.25,0.349) | Alkalinity LN(122.87,0.171) | Total Hardness N(381.62,102.06) |
|---|---|---|---|
| 1 | 0.51 | 0.81 | 0.266 |
| 2 | 0.05 | 0.08 | 0.660 |
| 3 | 0.12 | 0.12 | 0.375 |

**Table 3 Comparison of the Descriptive Statistics between the three generated series of each of the three water quality parameters of Shatt Al-Hilla River, using the fitted models, and those observed for years(2009-2013),(tc=1.98,Fc=1.54,at 95%, significance level )**

| Data | N | Mean | Std. Deviation | T-Test for Mean | F-Test for Variance |
|---|---|---|---|---|---|
| Mg Generated 1 | 60 | 36.1653 | 6.89202 | 0.800 | 1.086 |
| Mg Generated 2 | 60 | 33.1162 | 6.81548 | 1.306 | 1.115 |
| Mg Generated 3 | 60 | 34.3756 | 6.28489 | 0.563 | 1.102 |
| Alkalinity Generated 1 | 60 | 130.3001 | 18.61614 | 0.481 | 1.010 |
| Alkalinity Generated 2 | 60 | 126.6469 | 21.03712 | 1.473 | 1.210 |
| Alkalinity Generated 3 | 60 | 126.9750 | 21.16161 | 1.292 | 1.225 |
| Hardness Generated 1 | 60 | 371.9359 | 97.12119 | 0.846 | 1.221 |
| Hardness Generated 2 | 60 | 375.8565 | 92.39124 | 0.624 | 1.105 |
| Hardness Generated 3 | 60 | 374.1661 | 92.66139 | 0.786 | 1.111 |
| Mg Observed | 60 | 35.0833 | 6.34139 | -------------- | -------------- |
| Alkalinity Observed | 60 | 131.9000 | 19.11934 | -------------- | -------------- |
| Hardness Observed | 60 | 386.9667 | 87.88879 | -------------- | -------------- |

**Table 4 Frequency Analysis Comparison between Observed and Generated Monthly Magnesium of Shatt Al-Hilla River,(2009-2013).**

| Data | Min. | Max. | P>20 | P>30 | P>40 | P>50 | P>60 |
|---|---|---|---|---|---|---|---|
| Observed | 20 | 53 | 0.983 | 0.783 | 0.133 | 0.016 | 0 |
| Generated 1 | 18.2 | 67.72 | 0.983 | 0.766 | 0.133 | 0.033 | 0.01 |
| Generated 1 | 12.86 | 61.72 | 0.967 | 0.783 | 0.150 | 0.016 | 0.01 |
| Generated 1 | 17.74 | 61.08 | 0.967 | 0.766 | 0.150 | 0.033 | 0.01 |

**Table 5 Frequency Analysis Comparison between Observed and Generated Monthly Alkalinity of Shatt Al-Hilla River,(2009-2013).**

| Data | Min. | Max. | P>100 | P>120 | P>140 | P>160 | P>180 |
|---|---|---|---|---|---|---|---|
| Observed | 95 | 174 | 0.933 | 0.700 | 0.350 | 0.050 | 0 |
| Generated 1 | 92.17 | 192.8 | 0.950 | 0.683 | 0.333 | 0.050 | 0.016 |
| Generated 1 | 92.12 | 192.04 | 0.917 | 0.716 | 0.350 | 0.066 | 0.016 |
| Generated 1 | 89 | 189.4 | 0.933 | 0.716 | 0.333 | 0.050 | 0.016 |

**Table 6 Frequency Analysis Comparison between Observed and Generated Monthly Total Hardness of Shatt Al-Hilla River,(2009-2013).**

| Data | Min. | Max. | P>100 | P>200 | P>300 | P>400 | P>500 |
|---|---|---|---|---|---|---|---|
| Observed | 178 | 535.0 | 1 | 0.950 | 0.900 | 0.517 | 0.017 |
| Generated 1 | 69 | 564.8 | 0.983 | 0.966 | 0.883 | 0.533 | 0.033 |
| Generated 1 | 74 | 572.4 | 0.983 | 0.966 | 0.900 | 0.517 | 0.017 |
| Generated 1 | 160 | 580.0 | 1 | 0.950 | 0.016 | 0.500 | 0.017 |

**Table 7 The Lag-Correlation and Multiple Correlation Coefficients of the Water Quality Parameters of Shatt Al-HillaRiver,Iraq,(2000-2013).**

| Parameter | Lag 1 | Lag2 | Lag3 | Lag4 | Lag5 | Multiple Correlation Coefficient |
|---|---|---|---|---|---|---|
| $So_4$ | 0.816 | 0.744 | 0.686 | 0.583 | 0.492 | 0.833 |
| Mg | 0.403 | 0.169 | 0.059 | 0.08 | 0.067 | 0.409 |
| Ca | 0.712 | 0.645 | 0.513 | 0.487 | 0.449 | 0.747 |
| Alkalinity | 0.491 | 0.273 | 0.225 | 0.007 | 0.024 | 0.532 |
| Total Ha. | 0.35 | 0.209 | 0.212 | 0.0431 | 0.0345 | 0.521 |

Prof. Dr. Rafea." Frequency Analysis of Some of Water Quality Parameters of Shatt Al-Hilla River, Iraq."American Journal Of Engineering Research (AJER), Vol. 7, No. 6, 2018, PP.190-199.