Interpretation Of Ground Water Quality Data Variation In Erbil City, Northern Iraq

Mus'ab A. Al-Tamir

Civil Engg. Dept. University of Mosul, Mosul, Iraq

Abstract

In Erbil city; more than 30% of the water supply is derived from wells. Since the wells are located through the city, the quality of their waters may have widely variation.

Principal components analysis technique (PCA) was used to processing the physical, chemical and biological data of several wells at different parts of Erbil city to define the components or factors that responsible to the main variance in Erbil ground water quality. The correlation matrix also adopted in data analysis to determine the relationships of each parameter with the others.

The results of PCA showed the domination of three factors that responsible of about 68% of the ground water variation these factors are: change of the rock nature with 31%, human activities impact with 20.9%, and 16.8% of variation is according to the agricultural and storm water effects.

The correlation matrix had shown that there are two strong direct correlation between TDS and TH with 0.73, and opposite correlation between pH and Coliform bacteria with 0.53 the correlation between other water quality parameters is weak and less than 0.40.

Keywords: Ground water quality, Principal component analysis, Factor analysis, Erbil.

تفسير تغاير بيانات نوعية المياه الجوفية في آبار مدينة أربيل، شمال العراق

مصعب عبد الجبار التمر

قسم الهندسة المدنية/ جامعة الموصل

الخلاصة

أكثر من 30% من مياه الإسالة المستخدمة في مدينة أربيل مصدرها مياه الآبار المنتشرة هناك بكثرة. وبما أن الآبار موجودة داخل المدينة فإن نوعيتها تتغاير بشكل واسع.

استخدمت تقنية تحليل المركبات الأساسية في معاملة البيانات الفيزيائية والكيميائية والبيولوجية لمجموعة من الآبار في مناطق متفرقة من مدينة أربيل لتعريف العوامل المسؤولة عن التغيرات الرئيسة في نوعية الماء الجوفي للمنطقة، هذا فضلاً عن استخدام مصفوفة الارتباط لتحديد علاقة كل متغير مع بقية المتغيرات.

بينت نتائج تحليل المركبات الأساسية وجود ثلاث عوامل رئيسة مسؤولة عن حوالي 68% من التغاير في نوعية المياه الجوفية هذه العوامل: التباين في طبيعة صخور المنطقة بنسبة31%، الأنشطة البشرية في المدينة بنسبة 20.9%، 16.8% من التغاير في المياه الجوفية للمنطقة يعزى إلى عامل النشاطات الزراعية ومياه السيح.

كما أظهرت مصفوفة الارتباط وجود ترابط طردي قوي بين المواد الصلبة الذائبة الكلية والعسرة الكلية بمقدار 0.73، وترابط أخر عكسي قوي بين الأس الهيدروجيني وبكتريا الكوليفورم بمقدار 0.53 أما الترابط بين بقية المتغيرات فكان ضعيفاً ولا يتجاوز 0.4%.

الكلمات الدالة: نوعية المياه الجوفية، تحليل المركبات الأساسية، التحليل العاملي، أربيل.

Received 6 May 2007 **Accepted 12Aug. 2007**

Introduction

The history of ground water utilization in northern Iraq begins in the antiquity (about 7000 year B.C.). In the mountain regions springs were exploited for water supply, Irrigation or orchards and for animals. Wells were mainly utilized for water supply and at times used as strategic war tool. Underground tunnels locally called Kahreez were also developed to fulfill the needs of water, the remains of which are still found in the region[1].

The ground water quantity can play an important role in satisfying the different water uses and importance social and economic circumstances of the people. Yet the quality of such water resources may be of equal importance to its quantity if not exceeding it[2].

As water quality involves many parameters, the monitoring produces a large number of variables. To identify a set of dimensions and the major sources of quality variation, statistical analysis methods may be applied to reach to these points[3].

Habib et al., (1990)[4] use a correlation matrix, trilinear diagram and several comparing methods for interpretation the ground water quality in Erbil city north of Iraq. Mahmoud And Zangana, (1990)[5] made a simple comparison of their data with the water quality criteria to state the suitability of the ground water quality for human drinking use. Many researches in the world have used principal components analysis (PCA) in ground water quality management, Jayakumar and Siraz (1997)[6] in India, Eraifej and Abu-jaber (1999)[7] in Jordan, Suk and Lee (1999)[8] in Koria, Khatttab (2004)[9] in Hit-Cubaisa area west of Iraq, Olobaniyi and Owoyemi (2006)[10] in western Niger delta in Nigeria, Sarkar et al., (2006)[11] in India.

Many factors contribute to the variation in ground water quality, in Erbil city these include the geological formation, agricultural activities percolation of runoff to the geological stratification, and human activities and so forth in interaction between them by different percentages.

This paper try to identify the sources of water quality variation and pollution and the relationship for each water quality parameter with the others for the ground water quality in Erbil city, the correlation matrix and principal components analysis loading matrix were adopted to achievement this goal.

Geology of the area

Fig.1 show the geological map of Erbil basin [4], Geologically Erbil area is part of the low folded bet of northern Iraq. It lies within a structural trough with a NW-SE axial trend. The area extension of this basin corers a wide syncline bounded by the Perman Dagh anticline in the NE and by Kirkuk anticlinal structure in the SW.

The oldest outcrop cock in the area is of an upper Miocene age. More than 80% of the area is covered by the Quaternary deposits, composed mainly of soils, which is underlain by the Bakhtiari formation of Pliocene age[12]. This formation is composed mainly of gravels and conglomerates with an intercalation of sands, clay and silt. It is exposed at the eastern and western parts of the basin near the recharge and discharge zones [4].

Hydrogeologically the basin is bordered by the greater Zab and lesser Zab from NW and SE respectively. The basin is divided into three sub-basins, these are; Kapran at the north, central (includes Erbil city), and Bash Tepa in the south. The greater Zab river drains both the Kapran and central sub-basin, while the lesser Zab river drains the Bash Tepa sub-basin[4].

The Bakhtiari sediments make up the aquifer system in Erbil basin. The productive thickness of these sediments may reach up to 200m at the central and bash Tepa sub-basins, while it is in the range of 100m at the Kaprn sub-basin, also the ground water flow follows the topography of the region.

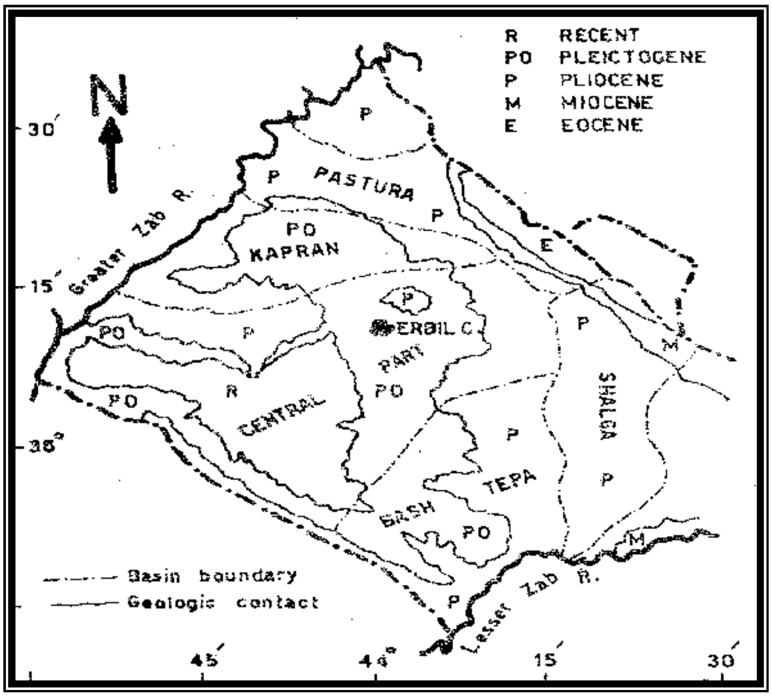

Figure 1: Geological map of Erbil basin.[4]

Methodology

To define the interaction between the water quality parameters and to identify the sources of pollution or the main sources of the ground water quality variation in Erbil city principle components analysis (PCA) applied on the water quality data obtained from references [4] and [5] for Erbil city northern Iraq that its location showed in Fig.2 [4],[5]. The data includes the following physical, chemical and biological properties: pH, Total Hardness (TH), Turbidity (Turb.), Total Dissolved Solids (TDS), Nitrate (NO_3^{-1}), Sulfate (SO_4^{-2}), Chlorides (Cl^{-1}), and most probable number for Coliform bacteria MPN/100ml. Table 1 presented the descriptive statistics of these data set.

Table 1: The descriptive statistics of the data.

Parameter / Descriptive	pH Unit	Turb. Unit	TDS mg/ℓ	TH mg/ℓ	NO_3^{-1} mg/ℓ	SO_4^{-2} mg/ℓ	Cl^{-1} mg/ℓ	MPN / 100 ml
Mean	7.38	2.79	674	251.1	4.7	171.3	157.4	42.6
Median	7.35	2.7	625	225	4.7	122.5	124.5	41
Mode	7.3	2.7	580	220	5.4	30	140	40
Standard Deviation	0.45	1.38	270.7	98.7	2.8	129.2	111.9	17.8
Variance	0.20	1.91	73277.9	9732.6	7.8	16695	12524	315.7
Kurtosis	-0.86	-1.09	-0.72	0.65	0.6	-0.5	3.2	0.4

Skewness	0.17	0.29	0.57	1.0	0.7	0.8	1.8	0.6
Range	1.5	4.3	870	367	10.5	420	440	70
Minimum	6.7	0.8	330	113	0.7	30	50	15
Maximum	8.2	5.1	1200	480	11.2	450	490	85

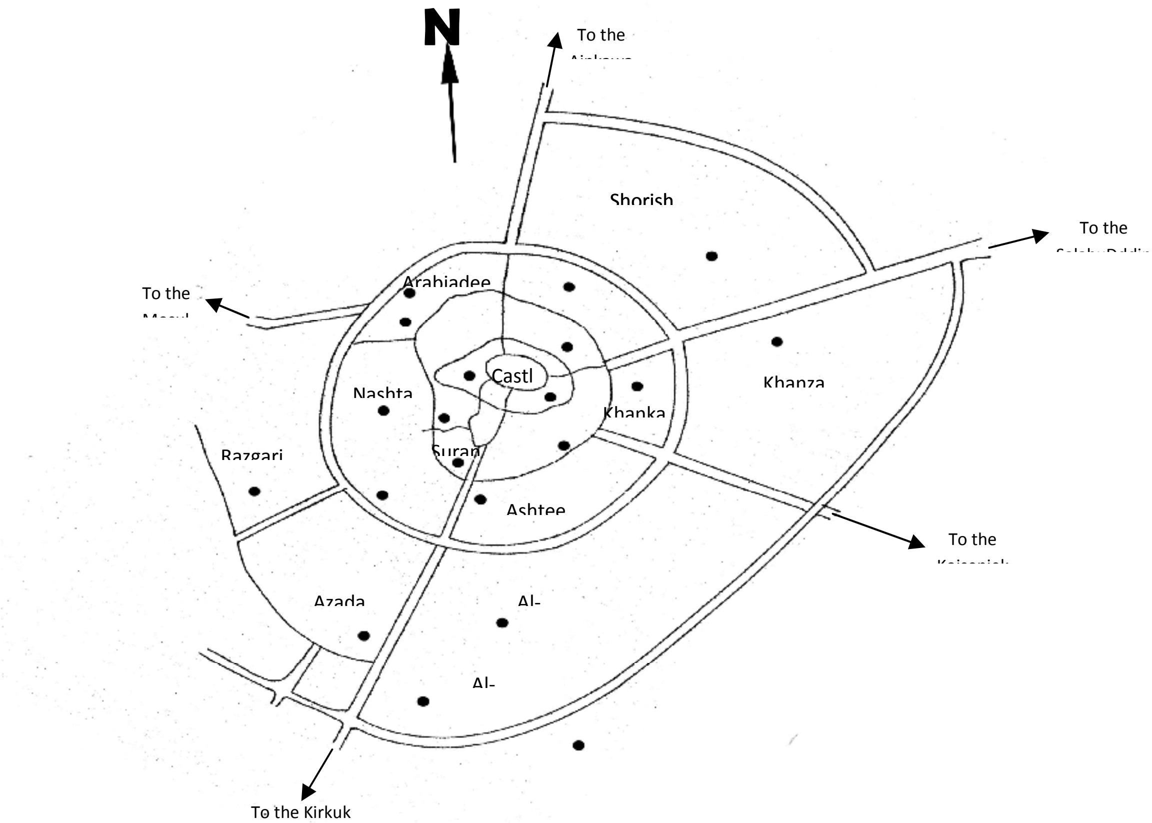

Figure 2: Erbil city map with Wells positions. [4], [5]

The data were statistically analyzed using the software SPSS program, the raw data were standardized as water quality parameters have different magnitudes and scales of measurement as in the formula [13]:

$$\text{Standirdized value} = \frac{\text{original value - mean of the set}}{\text{Standar divation of the set}} \ldots\ldots\ldots\ldots\ldots 1$$

The PCA procedure transforms an original set of variables into a new set of principal components, which are at right angles (i.e. uncorrelated) to each other. Specifically, it decomposes a 2-D matrix X(I,J), consisting of I measurements of J variables into scores and loadings matrices, as given by:

$$X=TP'+E \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 2$$

Where each column of the matrix X is men centered (mean =0) and variance scaled (Standard deviation=1); T(I,J) represents the matrix of J principal components scores (each column of matrix T refers to a principal components); P' denotes the transpose of original data. In general, first R principal components are sufficient to describe a large amount of variance in the original matrix, X, and thus equ.2 Can be recast as:

$$\mathbf{X} = \sum_{R=1}^{R} \mathbf{tr(Pr)}^{T} + \mathbf{E'} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3$$

Where tr represents the rth I dimensional score vector; $(Pr)^T$ denotes the transpose of rth J dimensional loading vector (Pr) and E', the residual matrix. As can be seen from equ.3, the original matrix X can be represented in terms of only R score vectors (R<J) and thus data

reduction is achieved. The sum of squares of elements of score vector (tr) relates to the Eigenvalue (also termed as "Trace") associated with that score vector and is a measure of the variance captured by that principal component [14].

The principal components with Eigenvalues equal or greater than one are rotated using Kaiser criterion varimax rotation to yield a simpler components structure [15].

Discussion

Principal components analysis (PCA) extracted three factors with Eigenvalues greater than 1 that account 68.86% of the variance of water quality, (i.e. the 8 variables were concentrated into 3 components). Eigenvalues their loadings after rotation, percentage of variances and communalities showed in table 2.

Communalities provide an index to the efficiency of degree of contribution of each variable in the selected three components, the lower communality was for sulfate (0.579).

For components loadings the values which are greater than the radius of the equilibrium circle of contribution radius =(no. of extracted components/ no. of variables)$^{0.5}$=(3/8)$^{0.5}$=0.612, So the loaded greater than 0.612 are considered significant see table 2 and Fig. 2.

Table 2: Rotated principal components loadings.

Parameters	**Components**			**Communalities**
	I	**II**	**III**	
TH	**0.893**	-0.028	0.159	0.633
TDS	**0.868**	0.00	-0.04	0.736

			7	
Cl^{-1}	**0.613**	-0.388	-0.243	0.755
pH	0.116	**-0.792**	0.356	0.824
Turbidity	-0.319	**0.729**	0.009	0.628
Coliform Bacteria	0.382	**0.727**	0.248	**0.579**
NO_3^{-1}	0.279	-0.019	**-0.742**	0.586
SO_4^{-2}	0.215	-0.060	**0.727**	0.768
Eigenvalues	2.483	1.675	1.352	-
% of Variance	31.03	20.94	16.89	-
% Cumulative	31.03	51.97	68.86	-

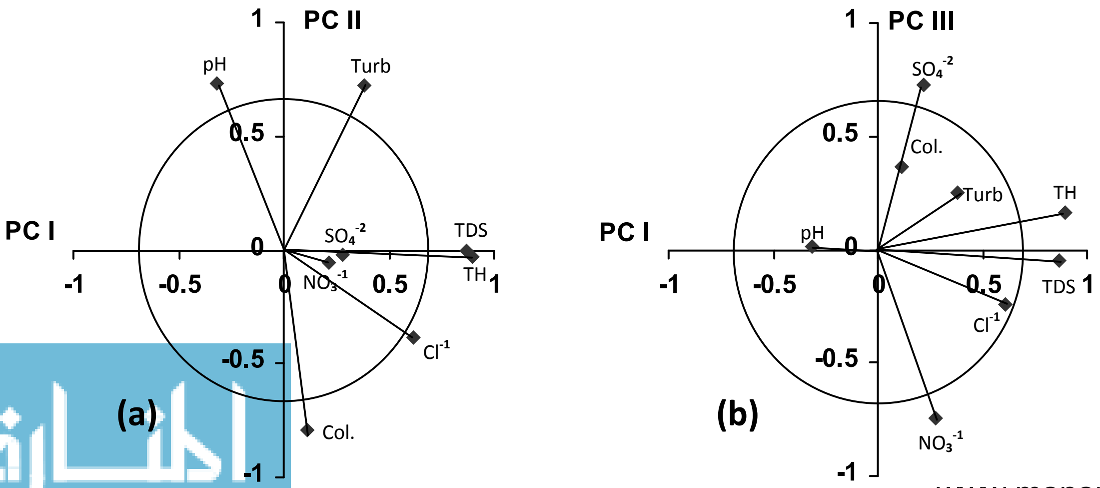

Fig. 2: Correlations of water quality parameters with 95% confidence level origocentered

Principle component I account 31.032% of the variance in ground water quality of Erbil city within the study area; has high loading of TH, TDS and Cl^{-1}, as seen from Fig. 2a all of the three dominated variables in PC1 are in the positive pole of the component that mean the positive correlated with each other. As its well known the hardness of the water on the region is Ca and Mg hardness (Habib et al., 1990); So this component represents the dissolution of rocks composed from marine minerals as $CaCl_2$ and $MgCl_2$ during the deposition of Bakhtari formation [16].

Principal component II account 20.94% of the total variance in water quality. It is dominated by Coliform bacteria and turbidity on its positive pole and pH on its negative pole Fig. 2a., as it is known that coliform bacteria represent an indication to the fresh domestic pollution, and the turbidity it expect to be colloidal particles (as dissolved) that come from domestic wastewater that drains into deep vertical pits or septic tanks from home and industries within study area these pits have been dugged to more than 20m until reaching a sandstone layer, once the wastewater reaches the sandstone layer, they may inter the flow system and migrate to the water table [4]. therefore this component indicate to the domestic and industrial pollution effects on the Erbil city. Principal component III account for 16.89% of the variance in water quality. It is dominated by NO_3^{-1} on its negative side and SO_4^{-2} on its positive side Fig. 2b, the fertilizer and storm water that percolating with the irrigation and rain water to the ground water can effect to the variation of NO_3^{-1} and SO_4^{-2} concentrations, therefore this component can reflect the agricultural and storm water effects on the ground water on the region. The relationships between the water quality variables are shown from the correlation matrix given in Table 3. It is seen in general that there is no relationships between most of the variables. However there is a strong direct relationship between TH and TDS (r=0.731) this may suggest that the TDS is from mineral origin that derived from the geology of the regions. Also there is a strong opposite relationship between coliform bacteria and pH (r=0.508), this indicate to the fresh pollution of ground water with the domestic waste from the region. The lack of the strong relationship between the other analyzed properties is most probably due to contamination [4].

Table 3: Below side of Correlation matrix.

Parameters	**pH Unit**	**Turb. mg/ℓ**	**TDS mg/ℓ**	**TH mg/ℓ**	**NO_3^{-1} mg/ℓ**	**SO_4^{-2} mg/ℓ**	**Cl^{-1} mg/ℓ**	**MPN/100 ml**
pH (unit)	**1**							
Turb.(unit)	0.219	**1**						
TDS (mg/ℓ)	-0.255	0.221	**1**					
TH (mg/ℓ)	-0.279	0.296	**0.731**	**1**				
NO_3^{-1}	0.05	-	0.17	0.12	**1**			

(mg/ℓ)	1	0.097	5	7				
SO_4^{-2} (mg/ℓ)	0.09	0.096	0.054	0.231	-0.192	**1**		
Cl^{-1} (mg/ℓ)	-0.401	-0.107	0.404	0.395	0.302	0.142	**1**	
Coliform Bacteria	**-0.508**	-0.366	0.093	0.209	-0.077	0.282	0.156	**1**

Conclusion

Principle components analysis gives effective identification of the main sources of water quality variations in the study area. Such variation cannot be easily detected with its percent effect. This shows the importance of such technique for water quality management agencies.Principal components analysis explain 63% of variation in water quality in three components, from the results of analysis, it is seen that rock dissolution, human effectiveness, and agricultural wastes with storm percolation to ground water are the main factors contribute at different degrees to such variation.

From the results of the correlation matrix it appears that there are direct significant correlation between TH and TSD with 0.731 and indirect significant correlation between coliform bacteria and pH with 0.508.

References

1. Mahmoud, T.A., 1981. Ground water quality in northern Iraq, Quality of ground water, proceeding of an International symposium, noordwijkerhout, The Netherlands, 23-27 march 1981, W. Van Duijvenbooden, P.Glasbergenand H.Van Lelyveld (Eds.), Studies in Environmental Science, Vol. 17, Elsevier Scientific publishing company-printed in the Netherlands.

2. Al-Rawi, S. M., Al-Azzo, S. I. and Abbawi, S. A., 1990. Hydrogeochemical evaluation of groundwater in some parts of Mosul city and suitability for irrigation, Proceeding of 2nd Sci. Conf. of SDRC,Mosul University, Mosul, Iraq, 18-20 March 1990.
3. Al-Rawi, S.M. and Shihab, A.S., 2005. Application of factor analysis al a tool for water quality management of Tigris river within Mosul city, Raf. Jour. Sci., Vol. 16, No. 1 Geology Special Issue, pp.56-64.
4. Habib, H. R., Al-Saigh, N. H. and Hassan, Z. M., 1990. Geochemistry of underground water in Erbil city, Iraq, Proceeding of 2nd Sci. Conf. of SDRC, Mosul University, Mosul, Iraq, 18-20 March 1990.
5. Mahmoud, F. Y. and Zangana, J. M., 1990. Ground water quality in Erbil city, Proceeding of 2nd Sci. Conf. of SDRC, Mosul University, Mosul, Iraq, 18-20 March 1990. (in Arabic).
6. Jayakumar, R. and Siraz, L., 1997. Factor analysis in hydrogiochemistry of coastal aquifers- a preliminary study, Environmental geology, 31(3/4)June, Springer-Verlag, pp.174-177.
7. Eraifej, N. and Abu-Jaber, N., 1999. Geochemistry and pollution of shallow aquifers in the Mafraq area, north Jordan, Environmental Geology 37(1-2), January Springer-Verlag, pp. 162-170.
8. Suk, H. and Lee, K., 1999. Characterization of aground water hydrochemical system through multivariate analysis: Clustering into ground water zones, J Ground water, Vol.37, No.3, pp.358-363.
9. Khattab, M.F., 2004. Factor analysis of water quality in Hitt-Cubaisa Area, west of Iraq, Raf. Jour. Sci., Vol.4, No.1, pp.98-110.(in Arabic)
10.Olobaniyi, S.B. and Owoyemi, F.B., 2006. characterization by factor analysis of the chemical facies of ground water in the deltaic plain sands aquifer of Warri, western Niger delta, Nigeria, AJST, Vol. 7, No.1, pp.73-81.
11.Sarkar, B. et al., 2006. Geo-environmental quality assessment in Jharia coalfield, India, using multivariate statistics and geographic information system, Environmental Geol., DOI 10.1007/s00254-006-0409-8, Springer-Verlag.
12.Al-Talabani, N.J. and Salih, A.Y., 1983. Land slope and hypsometric analysis of Erbil plane, Spec. Rep. Agr. And water, reaserch center, 28p.
13.Davis, J.C., 1973. Statistics and data analysis in geology, John Wiley and Sons, inc., Newyork, 550p.
14.Iyer, C.S., Sindhu, M., Kulkarni, S.G., Tambe, S.S. and Kulkarni, B.D., 2003. Statistical analysis of the physico-chemical data on the coastal waters of Cochin, J. Environ. Monit., 2003, 5, pp. 324-327.
15.Kleinbaum, D.G.; Kupper, L.L. and Muller, K.E., 1988. Applied Regresion Analysis and Other Multivariable Methods, PWS-KENT Publishing Comp., 2nd ed., USA., 718p.
16.El-Kiki, F.E., Al-Din, T.S. and Hassan, H.A., 1974. Hydrological condition of Bakhtiari aquifer in Shari lake basin, J. of Geol. Soc. Of Iraq, Vol. 7, pp. 111-127.