*JKAU: Earth Sci.*, **Vol. 20**, No. 2, pp: 1-10 (2009 A.D. / 1430 A.H.)

# Determination of Dominant Chemical Processes in an Alluvial Aquifer System of Wadi Ghiran Using Multivariate Statistical Analysis

**Mahmoud S. Alyamani** and **Abdullah A. Sabtan**

*Hydrogeology Department and Engineering Geology Department,*
*Faculty of Earth Sciences,*
*P.O. Box 80206, Jeddah 21589, Saudi Arabia*

Received: 2/8/2008 Accepted: 2/11/2008

*Abstract.* Factor analysis (FA) was applied to groundwater chemistry data of Al-Barzah town located in the upper part of Wadi Ghiran western Saudi Arabia to assess the usefulness of such method for interpreting commonly collected groundwater quality data and relating the data to specific hydrochemical processes. Six factors were retained; The first three factors were adequate for describing the groundwater chemistry processes that have taken place in the area: Factor 1 represents the degree of the overall mineralization of the groundwater due to intensive evaporation rates. Factor 2 reflects the dissolution and precipitation of carbonate minerals, Factor 3 may reflect groundwater pollution by on-site sewage disposal systems used in these regions and Factor 4 displays the buffering of the $HCO_3$ content by $CO_2$ exchange between groundwater and the atmosphere. Factors 5 and 6 represent single variables, dissolved oxygen (DO) and groundwater temperature and both show high association in these two factors. This confirm that the two variables are largely independently controlled and describe the unique factors assumed to be uncorrelated with each other and with common factors.

## Introduction

Due to the complexity of the chemical evolution of groundwaters and the substantially large amount of basic information available, investigators are often unable to obtain a clear picture of the system under study. In general, the measurements, chemicals and non chemicals carried out both in the field and laboratories, might reflect a number of variables which together make up the groundwater chemistry. Such obtained variables could cause rather complexity and confusion for investigation to have a clear picture of the existing system (Usunoff and Guzman-Guzman, 1989).

Several methods of data analysis have been devised to simplify interpretation and presentation such as Trilinear, Semilogarithmic, Duorv diagrams….*etc.* (Lloyd and Heathcote, 1985). The existing methods may provide some information. Nevertheless, these conventional techniques are generally limited to major constituents ions. They ignore many parameters which are otherwise important for hydrochemical studies. The limitation that is coupled for using the traditional graphical methods were discussed by several authors (Sen and Al-Dakheel, 1986; and Liu *et al*., 2003). In view of the limitation of the existing methods and increasing number of chemical and physical variables measured in groundwater investigations, multivariate analysis comes into play as a rather essential tool for explaining groundwater chemistry conditions. Factor analysis (FA) is considered the most widely used in groundwater chemistry. The aim of using factor analysis is to simplify the quantitative description of a system by determining the minimum number of new variables necessary to reproduce various attributes of the data. These procedures reduce the original data matrix from one having n variables necessary to describe the N samples to a matrix with m factors ($m<n$) for each of the N samples (Davis, 1973; and Johnston, 1980).

This study attempts to employ the factor analysis (FA) to describe the hydrochemical processes that controls the groundwater chemistry in Wadi Ghiran in western part of Saudi Arabia from commonly collected groundwater chemistry data.

## Location and Geologic Setting

The study was conducted in Al-Barzah town that lay on a shallow aquifer. It is nearly located in western part of Saudi Arabia (Fig. 1) at the upper reaches of the Wadi Ghiran drainage, that is 365 meters high above the sea level. The climate is typically arid and the average annual rainfall is about 70 mm. The rain is irregular and of torrential nature when it occurs. In the area, the agricultural activities are wide spread. Groundwater occurs with the alluvial deposits of the wadi system. These deposits are mainly derived from the weathering of the parent rocks. The wadi deposits are of pebbles alternating with coarse sand and gravel that serve as a permeable conduit for the percolation of surface water into the aquifers. The depth to the water level varied between 25 to 30 m**.** The existing aquifer is mainly recharged through runoff events.

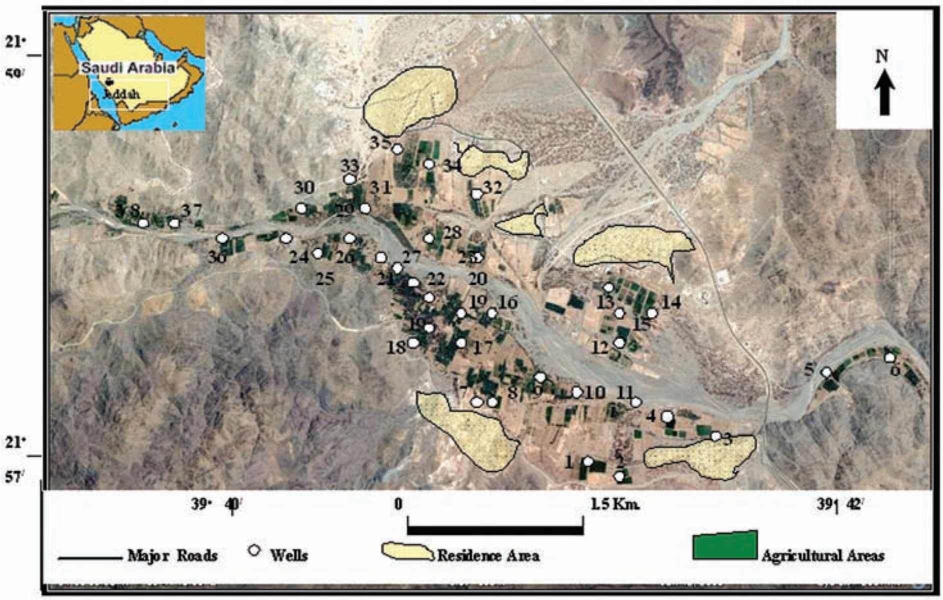


**Fig. 1. Location map.**

Geologically, from youngest to oldest four major geologic units in the area: (1) Quaternary superficial deposits that consist mainly of sand and gravel; (2) Tertiary volcanics occur as a series of basalt flows; (3) Tertiary sedimentary cover that consists of continental sediments, and (4) basement complex which is represented by Precambrian group of meta volcanic rocks.

## Procedures

The 38 groundwater samples described herein were taken from private wells that are situated in Al-Barzah town (Fig. 1). These water samples were collected in May 2008. Samples were analyzed for the major ions. Electrical conductivity (EC) in µS/cm, pH, Dissolved oxygen (DO) and temperature ($^{o}$C) were measured at all the wells after it had been pumped for at least 5-10 minutes. Variables used are $Ca^{2+}$, $Mg^{2+}$,

$Na^+$, $K^+$, $HCO_3^-$, $Cl^-$, $SO_4^{2-}$, $NO_3^-$, pH, EC, DO, and groundwater temperature. All chemical constituents mentioned above were in ppm except where noted. In addition, saturation indices of calcite, dolomite, gypsum and partial pressure of carbon dioxide ($P_{co2}$) were also determined by using PHREEQC program (2001). Factor analysis was performed with help of SPSS program (1990).

## Factor Analysis (Q-mode)

Prior to factor analysis, the data were standardized to remove the effects of using different units in describing the various variables, and so the variables are put in standard form. The standardized value then has a mean value of zero and standard deviation of unity. The first stage in the FA is to compute a correlation matrix, which is a measure of interrelation for all variables. These correlation coefficients are then arranged and allowed FA to compare and group the variables according to linear correlation coefficient. This type of analysis is called R-mode factor analysis, which is employed in this study, with iterations of Nie *et al.* (1975). Kaiser normalization and varimax rotation were also used. This is considered one of the most commonly used methods for orthogonal rotation, and results in factors that are uncorrelated. Finally, the factors scores were calculated for each sample and reflect the importance of given factor at that sample site.

## Results and Discussion

The output results for the chemical data are given in Table 1. Four factors had eigenvalue more than 1, while the other two factors (5 and 6) are of eigenvalue of less than 1. However, the "sky-slope" diagram suggest that six factors should be retained (Fig. 2), where the experimenter selects factors upto the point which the diagram shows a break of slope. The first four factors account for 76.1% of the total variance while the last two factors (5 and 6) represent 7.1% of the total variance. Therefore, these four factors are assumed to represent, adequately, the overall variance of the data set. On the other hand, the factor loadings in Table 1 indicates that the first three factors (1, 2 and 3) are hydrochemically meaningful, which seem to describe the existing conditions of groundwater chemistry. For obtaining an interpretation of the nature of the retained factors, these factors will be discussed below:

**Table 1. Variables and factor loading after varimax rotation.**

| Variables | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Factor 5 | Factor 6 |
|---|---|---|---|---|---|---|
| EC | <u>0.938</u> | 0.173 | 0.109 | 0.077 | 0.066 | –0.051 |
| $SO_4$ | <u>0.917</u> | 0.125 | 0.051 | 0.090 | 0.061 | 0.003 |
| Cl | <u>0.885</u> | 0.141 | <u>0.593</u> | 0.119 | 0.120 | –0.006 |
| Ca | <u>0.872</u> | 0.127 | 0.226 | –0.286 | 0.240 | –0.012 |
| Sat. Ind (Gypsum) | <u>0.851</u> | 0.152 | 0.072 | 0.028 | 0.021 | 0.049 |
| Mg | <u>0.845</u> | 0.121 | 0.119 | 0.188 | 0.007 | 0.008 |
| Na | <u>0.838</u> | 0.105 | 0.018 | 0.103 | 0.038 | –0.023 |
| K | <u>0.746</u> | 0.076 | 0.082 | 0.107 | 0.021 | 0.004 |
| pH | –0.173 | <u>0.947</u> | –0.008 | –0.029 | 0.046 | 0.007 |
| $P_{co2}$ | –0.056 | <u>–0.914</u> | 0.170 | <u>0.431</u> | –0.193 | –0.045 |
| Sat. Ind.( calcite) | –0.181 | <u>0.889</u> | 0.040 | 0.035 | –0.023 | 0.051 |
| at. Ind. (Dolomite) | –0.395 | <u>0.852</u> | 0.032 | 0.165 | –0.018 | 0.043 |
| $NO_3$ | 0.141 | –0.193 | <u>0.914</u> | 0.005 | 0.002 | 0.006 |
| $HCO_3$ | –0.091 | 0.003 | –0.018 | <u>0.813</u> | 0.043 | 0.058 |
| Dis. Oxyg. | –0.067 | –0.012 | –0.336 | 0.061 | <u>0.731</u> | –0.067 |
| Temp. | 0.061 | –0.044 | 0.226 | 0.056 | 0.096 | <u>0.662</u> |
| **Eigenvalue** | **6.220** | **3.420** | **1.470** | **1.035** | **0.668** | **0.458** |
| **Percent of Variance** | **39.0** | **21.4** | **9.19** | **6.47** | **4.18** | **2.9** |

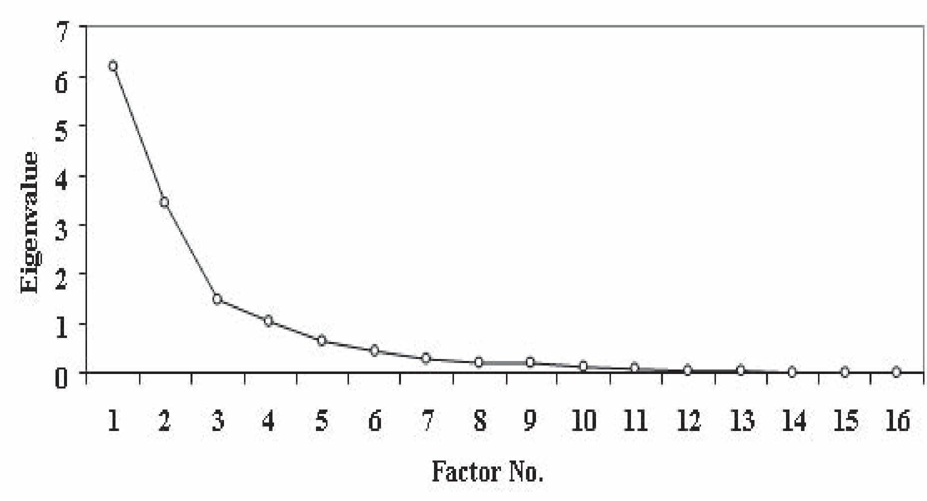


**Fig. 2. Number of factors plotted against the proportion of variance it extracted (Eigenvalue).**

## *Factor 1*

It accounts for 39.0% of the total variance. It has high loading for EC, $SO_4$, Cl, Ca, saturation index of Gypsum, Mg, K, Na and K. High values of these constituents most likely result from evaporation. All are positively correlated with factor 1 weighting. Thus, factor 1 represents the degree of overall mineralization of the groundwater. This in turn is affected by mineral dissolution and evaporation, of which evaporation is believed to be the most important.

### *Factor 2*

It accounts for 21.4% of the total variance and has a high loading of pH, partial pressure of carbon dioxide ($P_{co2}$) and the saturation indices of calcite and dolomite. These variables control and reflect the dissolution and precipitation of carbonate minerals. The sign values of the factor loadings indicate that pH and $P_{co2}$ values are inversely proportional, which is expected since when $P_{co2}$ decreases the pH increases. Surprisingly, $HCO_3$ is not significantly loaded on this factor.

### *Factor 3*

It accounts for 9.2% of the total variance and has a high loading of $NO_3$ and moderate one of Cl. It may indicate that the groundwater has been affected by on-site sewage disposal systems used in the study area. In unsewered residential area and shallow groundwater (Fig. 1), both $NO_3^-$ and Cl are the most significant contaminants associated with domestic wastewater (EPA, 2004). Disposal of sewage within the wadi basin has been accomplished almost exclusively through the use of conventional on-site sewage systems. A traditional on-site sewage system consists of a cesspool and a subsurface absorption system. Chloride ($Cl^-$) is a good indicator parameter of sewage impacts because it is not subject to adsorption, ion exchange, or oxidation-reduction "redox" reactions. Recently, Saudi Geological Survey (2008) shown that many of the domestic water supply wells, particularly in the upper reaches of the wadi extract waters with high $NO_3^-$ concentration (161 mg/l) which is greater than the acceptable maximum contamination level (MCL) of 45 mg/l (10 mg/l of nitrate-nitrogen) by the World Health Organization (1998, 2000). To differentiate the potential sources of $Cl^-$ and $NO_3^-$ in the groundwater of the study area; the plots of $NO_3^-$ *vs* $Cl^-$ is used. The relationship between $NO_3^-$ and $Cl^-$ in an area of known cesspool system contamination is shown in (Fig. 3). It appears that Cl almost increases linearly with increasing $NO_3^-$ with correlation coefficient ( $r^2$ ) of about 0.61.

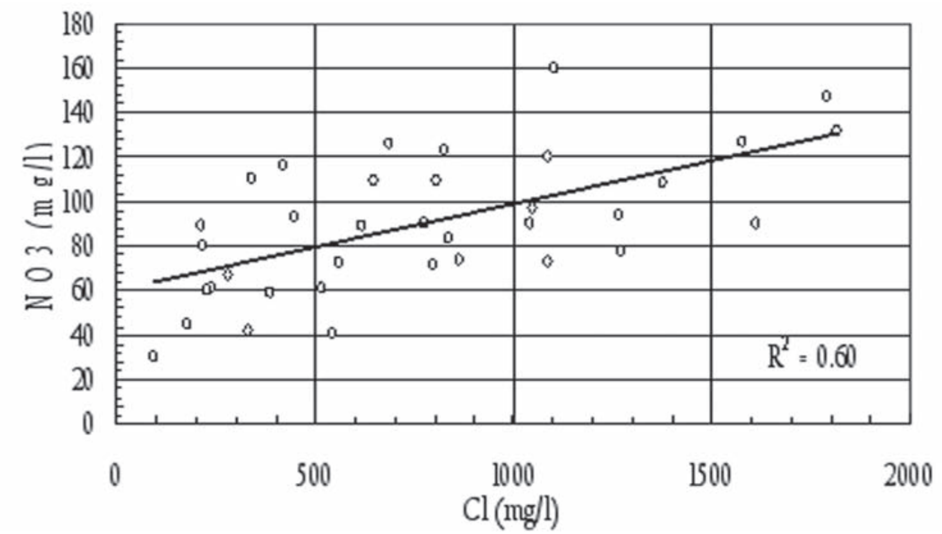


**Fig. 3. The relationship between $NO_3$ and Cl.**

### *Factor 4*

It account for 6.5% of the total variance and has a high loading of $HCO_3$ and secondary one of ($P_{co2}$). It may reflect the initial dissolution of minerals by groundwater. The low variance associated with this factor may indicate buffering of the $HCO_3$ content by $CO_2$ exchange between groundwater and the atmosphere, and perhaps the precipitation of calcite and dolomite. Geochemically, both factors 2 and 4 are associated with $CO_2$ – $H_2CO_3$ – $HCO_3$ – $H_2O$ chemistry of the groundwater. However, such condition usually results from the involvement of more than one process.

### *Factors 5 and 6*

Both represent single variables, dissolved oxygen and groundwater temperature respectively. The fact that dissolved oxygen shows only secondary association in factor 3 and high association with an orthogonal factor of its own confirms that its concentration is largely independently controlled. This may also apply on Cl behavior, which is loaded by both factors 1 and 3.

By examining the factor scores, the spatial importance of the new variables and its distribution can be mapped. The spatial distribution of the first major factor obtained for the chemical data was used to construct

Fig. 4. In term of the distribution of high scores (+ve), the distribution of the factor somewhat matches that of the groundwater salinity. Figure 5 outlined the area affected by on-site sewage disposal systems used in the study area.

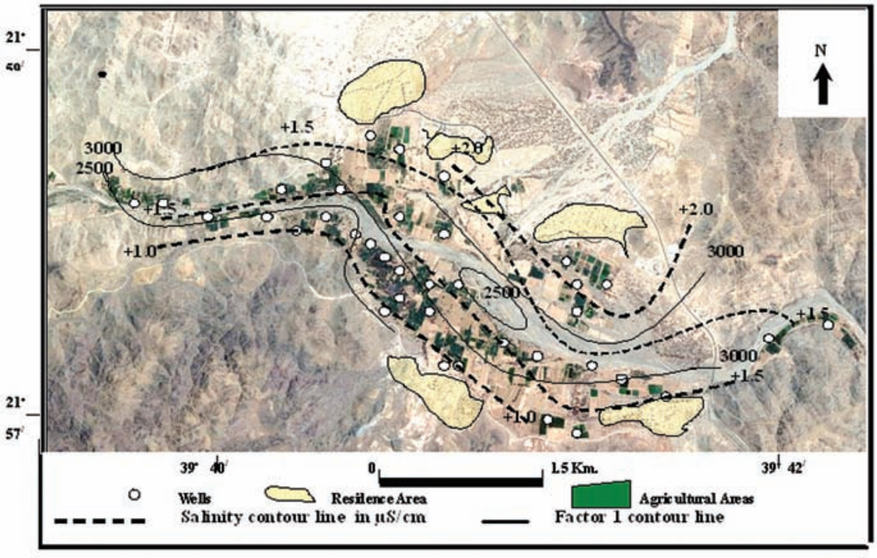


**Fig. 4. Factor 1 distribution in the area.**

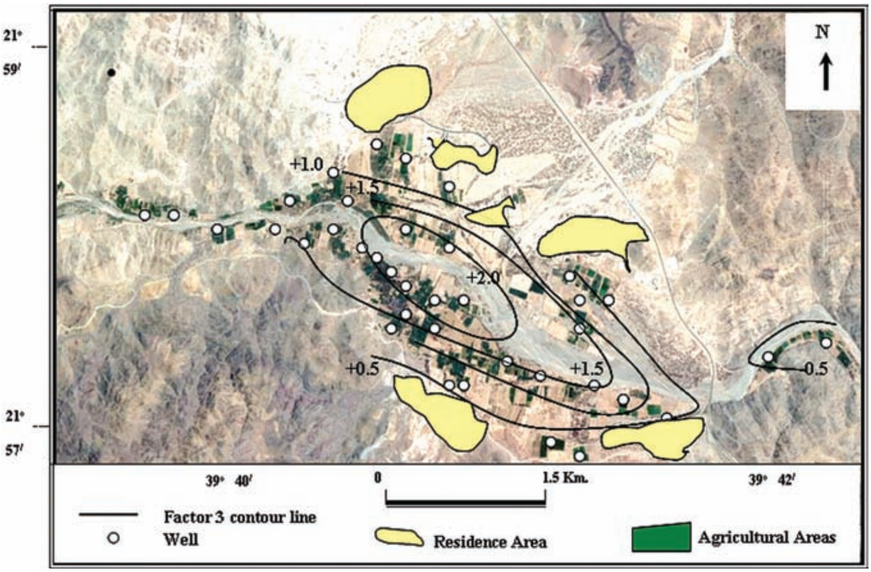


**Fig. 5. Factor 3 distribution in the area.**

## Conclusion

The results of factor analysis have rather clarified several aspects of groundwater condition from the hydrochemical point of view, and have led to the following conclusions that just three factors (1, 2, and 3) can describe the processes that are taking place in the hydrochemistry system. They represent three dominant groups of processes: Degree of mineralization, which is in turn controlled mainly by evaporation and precipitation and dissolution of carbonate minerals and finally, the contamination taken place by on-site sewage water. The distribution of the three factors also help in the identification of highly affected areas by these chemical processes.

### *Acknowledgement*

The authors would like to thank the Deanship of Scientific Research, King Abdulaziz University, for their financial support of the project No. 201/428.

### References

**Davis, J.C.** (1973) *Statistical and Data Analysis in Geology*, John Wiley and Sons, Inc., New York, 550 p.

**Environmental Protection Agency (EPA)** (2004) *Edition of Drinking Water Standards and Health Advisories*, EPA **822**-R-04-005.

**Johnston, R.J.** (1980) *Multivariate Statistical Analysis in Geography*, Longman Inc., New York, 280 p.

**Liu, CW., Lin, KH.** and **Kuo, YM.** (2003) Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan, *The Science of the Total Environment*, **313:** 77-89.

**Lloyd, J.W.** and **Heathcote, J.A.** (1985) *Natural Inorganic Hydrochemistry in Relation to Groundwater*, An Introduction, Oxford University Press, New York, 294 p.

**Nie, N.J., Hull, C.H., Jenkins, J.G., Steinbrenner, K.** and **Brent, D.H.** (1975) *Statistical Package for the Social Sciences*, 2nd ed., McGraw Hill Book Co., Inc., New York, 675 p.

**Phreeqc Interactive Alpha (PHREEQC)** (2001) *US Geological Survey*, V. **2**, Denver, USA. SPSS Inc. (1990). SPSS/PC+ Advanced Statistics, Marketing Department, Chicago, II., 60611.

**Saudi Geological Survey** (2008) *Strategic Groundwater Storage in Wadi Khulays-Makkah Region*, Saudi Arabia. Saudi Geological Survey, Internal Technical Report.

**Sen, Z.** and **Al-Dakheel, A.** (1986) Hydrochemical facies evaluation in mm Er Radhuma limestone, eastern Saudi Arabia, *Groundwater,* **24**(5): 626- 635.

**SPSS Inc.** (1990) *SPSS/PC Advanced Statistics*, Marketing Department, Chicago, II, 60611.

**Usunnoff, E.J.** and **Guzman-Gusman, A**. (1989) Multivariate analysis in hydrochemistry. An example of the use of factor analysis and correspondence analysis, *Groundwater*, **27**(1): 27-33.

**World Health Organization (WHO)** (1998) *Guidelines for Drinking Water – Water Quality*, 2nd ed. Addendum to volume 1, Recommendations, Geneva.

**World Health Organization (WHO)** (2000) *Reported of Drinking Water Quality Committee Meeting*, Berlin.

# تحديد العمليات الكيميائية السائدة في المتكون المائي لوادي غران باستخدام التحليل العاملي عديد المتغيرات

**محمود سعيد اليماني، و عبدالله عبدالعزيز سبتان**[*]

*قسم جيولوجيا المياه، و قسم الجيولوجيا الهندسية،*

*كلية علوم الأرض، جامعة الملك عبدالعزيز*

*جـــدة – المملكة العربية السعودية*

*المستخلص.* طبق التحليل العاملي على بيانات لكيمياء المياه من قرية البرزة، الواقعة في الجزء العلوي من وادي غران بغرب المملكة العربية السعودية، بغرض تقييم مدى فائدة هذه الطريقة في تفسير العمليات الكيميائية المائية. أوضحت الدراسة ظهور ستة عوامل: الثلاثة العوامل الأولى كانت كافية لوصف العمليات الكيميائية للمياه الجوفية التي تحدث في منطقة الدراسة. فالعامل (١) يمثل تمعدن الماء الجوفي نتيجة معدلات البخر الشديدة. والعامل (٢) يعكس عمليات إذابة وترسيب المعادن الجيرية. ومن المعتقد أن العامل (٣) يعكس تلوث المياه الجوفية بواسطة مواقع مخلفات الصرف الصحي المستخدمة في هذه المناطق، والعامل (٤) يبين تأثير محتوى $HCO_3$ بواسطة غاز ثاني أكسيد الكربون $CO_2$، نتيجة التبادل بين المياه الجوفية والغلاف الجوي. العاملان (٥) و(٦) يمثلان عاملين مستقلين هما الأكسجين الذائب، ودرجة حرارة المياه الجوفية. وهذا يؤكد أن هذين المتغيرين هما عاملان وحيدان ليسا مرتبطين بباقي العوامل الأربعة السابقة.