Available online at **www.joac.info**

ISSN: 2278-1862

Journal of Applicable Chemistry

2021, 10 (4): 466-471

(International Peer Reviewed Journal)

Factor Analysis of Ground Water Quality and Its Interpretation

V. Jena[1] and K. S. Patel[2]*

1. Department of Chemistry, Govt. NPG College of Science, Raipur, **INDIA**
2. Department of Chemistry, Pt Ravishankar Shukla University, Raipur, CG, **INDIA**

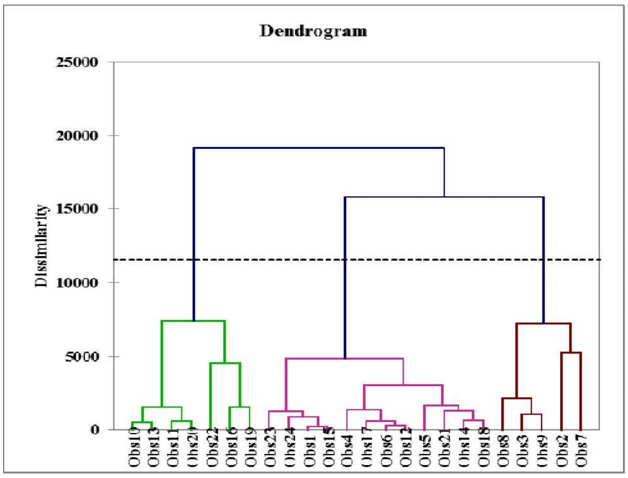
Email: jenavinod@gmail.com

Accepted on 10th July, 2021

ABSTRACT

In this study, the factor analysis techniques are applied to ground water quality data sets obtained from the different villages of Kanker North Bastar, Chhattisgarh India. The data obtained were standardized and subjected to Factor extraction to simplifying its interpretation and to define the parameters responsible for the main variability in water quality. The objective is to evaluate the mutual correlations among the various water quality parameters to reveal the primary factors that affect reservoir water quality, and the differences among the various water quality parameters. The factor analysis resulted in three factors explaining more than 60% of the total variation in ground water quality data set. The first factor indicates the variation in water quality is due to anthropogenic sources and second factor shows variation in water quality due to organic sources that are taking place in the system. Finally, the results of factor analysis reflect a good look on the water quality monitoring and interpretation of the ground water. Multivariate statistical techniques are potential tools and provide greater precision for identifying contaminant parameters linkages with groundwater chemistry in the area.

Graphical Abstract:



Dendrogram for 21 samples from cluster analysis in Q-mode

Keywords: Factor Analysis, Ground water, Water quality, Sources of pollution

INTRODUCTION

A thorough understanding of the nature and extent of contamination in an area requires detailed hydro chemical data [1]. Water quality is influenced by natural and anthropogenic effects including local climate, geology and irrigation practices [2]. Conventional techniques including Stiff and Piper plots only consider major and minor ions to assess the chemical quality of water, whether surface or groundwater [3]. Considering the limitations of these traditional methods to express the water quality and also the recent advances in analytical capabilities and the availability of larger numbers of chemical parameters, wide ranging statistical techniques are now needed to assess the water quality, nature and extent of contamination. In this regard, factor analysis is useful for interpreting groundwater quality data and relating those data to specific hydro-geologic and anthropogenic processes [4]. In particular many examples can be found of the application of multivariate analysis to sets of variables collected for surface and ground waters. Multivariate statistical techniques, cluster analysis (CA) and factor analysis (FA), are effective means of manipulating, interpreting and representing data concerning groundwater pollutants and geochemistry. They are frequently employed to characterize the quality of groundwater. Surface water, groundwater quality assessment and environ mental research employing multi-component techniques are well described in the literature [5-7]. Multivariate statistical approaches allow deriving hidden information from the data set about the possible influences of the environment on water quality [8-10]. In the present study, a water quality data matrix, obtained has been subjected to Factor analysis. The main objectives of this paper are: (1) to assess the chemistry of groundwater and (2) to identify the geologic factors that presently affect the water chemistry in the region by using multivariate statistical techniques.

MATERIALS AND METHODS

Collection of samples: The water samples (n=25) were collected from different stations. Exact locations for all sampled sites were determined using global positioning system and entered into a geographical information system for data processing.

Chemical and reagents: The ICP multi-element standard (E. Merck) and the European standard 13346:2000 EN 13346:46 were used for the quality control of the data.

Analysis of sample: The Varian Liberty AX Sequential ICP-AES, Flame GBC AAS 932 and HG-AAS were used for analysis of the trace metals in the samples.

Multivariate Factor Analysis: The hydro chemical data were statistically analyzed. The first step was to standardize the raw data to form table 2. Standardization tends to increase the influence of variables whose variance is small, and reduce the influence of variables whose variance is large. Factor analysis takes data contained in a correlation matrix and rearranges them in a manner that better explains the structure of the underlying system that produced the data. Eigen values and eigenvectors were calculated for the covariance matrix. Then, the data were transformed into factors. table 3 presents the Eigen value and the percentages of variance associated with each factor. These values are summed to express as a cumulative Eigen value and percentage of variance, respectively. Factor analysis attempts to explain the correlations between the observations in terms of the underlying factors, which are not directly observable [11-12]. There are three stages in factor analysis [13]. For all the variables a correlation matrix is generated Factors are extracted from the correlation matrix based on the correlation coefficients of the variables To maximise the relationship between some of the factors and variables, the factors are rotated. A first step is the determination of the parameter correlation matrix. It is used to account for the degree of mutually shared variability between individual pairs of water quality variables. Then, Eigen values and factor loadings for the correlation matrix are determined. Eigen values correspond to an Eigen factor which identifies the groups of variables that are highly correlated among them. Lower Eigen values may contribute little to the explanatory ability of the data. Only the first few factors are needed to account for much of the parameter variability. Once the correlation matrix and Eigen values are obtained, factor loadings are used to measure the correlation between the variables and factors. Factor rotation is used to facilitate interpretation by providing a simpler factor structure. After rotation of the factor loading matrix (e.g. by varimax rotation, which involves scaling the loadings by dividing them by the corresponding communality), the factors can often be interpreted as origins or common sources [**14-18**].

RESULTS AND DISCUSSION

The analytical results of multivariate analysis were performed for the set of 15 samples and 7 variables. The variables cluster into three major groups. From this dendrogram one can find the relationship between different variables, the dendrogram shows a high correlation between major ions (Ca^{2+}, Mg^{2+}, Na^{+} , HCO_3^{-} , SO_4^{2-}), which indicated surface water recharge and water-rock interaction. Other group of total hardness, bicarbonate and salinity are highly associated with each other indicated recharge zone area. The third cluster shows the similarity between trace elements, nitrate, potassium and manganese as one group, which probably represents the effects of weathering of the rich feldspars and mica, in addition of agriculture fertilizers. Figure 1 shows the clustering of the basins in Q-mode according the similarity. Two main groups can be verified in this figure. The first cluster shows a high similarity between the wells in the same geological formation (Figure 1). The second cluster shows a similarity between other wells which have been drilled in different geological formations.

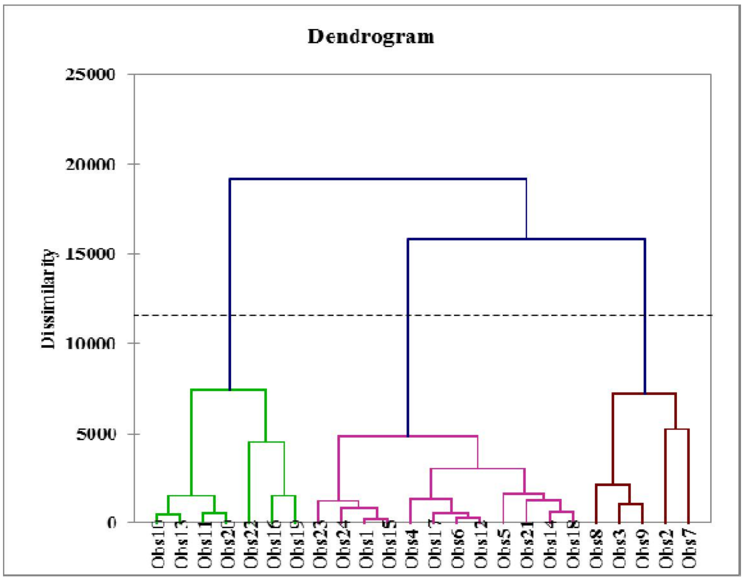


Figure 1. Dendrogram for 21 samples from cluster analysis in Q-mode.

Figure 2 shows the clustering of the basins in Q-mode according the similarity. Two main groups can be verified in this figure. The first cluster shows a high similarity between the wells in the same geological formation. The second cluster shows a similarity between other wells which have been drilled in different geological formations. Table 1 shows the Eigen values and cumulative variance for each factor. Distribution of samples with respect of F1 and F2 is shown in Figure 3 and Figure 4. Factor 1 distinguishes samples in relation to the enrichment in cations (Ca, Mg, Na) and anions (HCO_3^{-} , SO_4^{2-}). Four samples, with the higher negative score are located in recharge zone with the lowest salinity among the samples and it comes from the most superficial well. Figure 3 shows plots of samples on F1 versus F3. The first factor appears in separating the calcium bicarbonate water (Group B), where Group A is less concentrated. Total five samples show the highest calcium bicarbonate with positive scores in F1 and F3 whereas ten samples shows negative scores. Varimax rotation of the three factors was performed, Table 3 and Figure 4 shows the calculated scores for each sample.

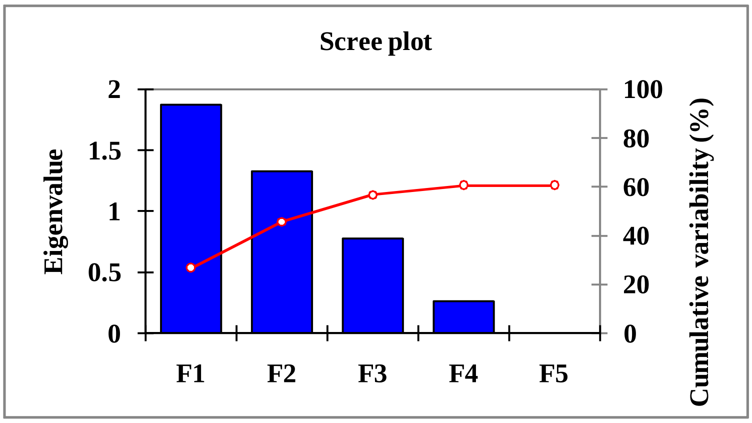


Figure 2. Plot of Eigen value.

Table 1. Correlation matrix (Pearson (n))

Variables	Ca	Mg	Na	K	Cl^-	HCO_3^-	SO_4^{2-}
Ca	1						
Mg	-0.102	1					
Na	0.247	0.002	1				
K	-0.278	0.048	0.452	1			
Cl^-	-0.293	0.184	-0.165	-0.160	1		
HCO_3^-	0.067	**0.386**	-0.357	-0.547	0.317	1	
SO_4^{2-}	**0.426**	0.121	0.250	-0.329	0.218	**0.309**	1

Values in bold are different from 0 with a significance level alpha=0.05

Factor I: It accounts for 26.79% of the variance in water quality of ground water within the study area figure 3. It represents the strong loadings of Ca, Cl-, SO_4^{2-} and HCO_3^-, which originates from anthropogenic causes like house hold materials and agricultural pollution in surface water. So we named this factor as „anthropogenic origin‟. Factor II: It account for > 19% of the variance in water quality figure 3. It is mainly dominated by Na and K. Table 3 gives the rotated factor loadings, communalities, Eigen values and the percentage of variance explained by these factors. In order to reduce the number of factors and enhance the interpretability, the factors are rotated. The rotation usually increases the quality of interpretation of the factor. All the parameters and Factor loadings are given also in table 1 and table 2.

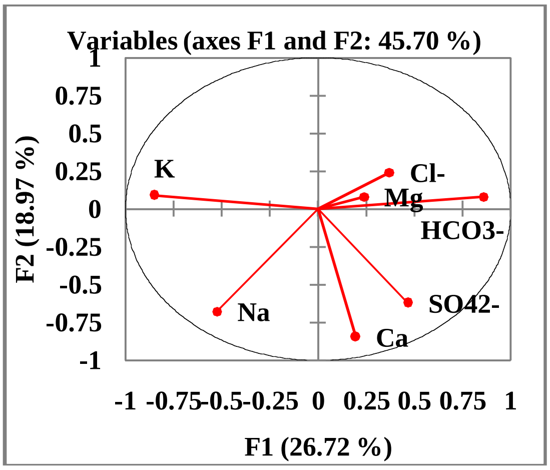


Figure 3. Plot of Factor 1 versus Factor 2 scores

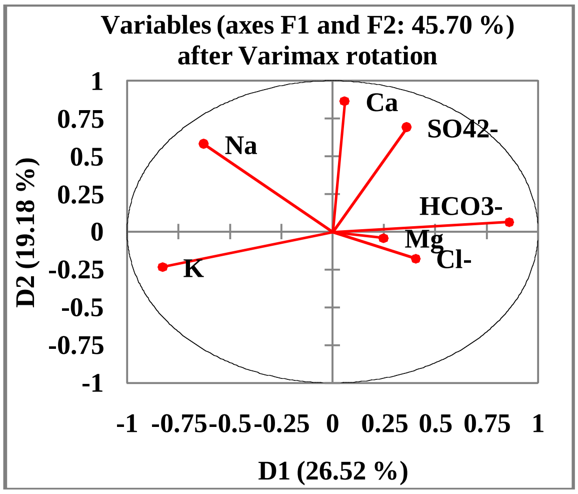


Figure 4. Plot of Factor 1 versus Factor 2 scores after Varimax rotation.

Table 2. Factor Loadings for water quality parameter

Parameters	F1	F2	F3	Initial communality	Final communality	Specific variance
Ca	0.18	0.76	0.30	0.41	0.70	0.30
Mg	0.22	0.07	0.42	0.26	0.23	0.77
Na	0.49	0.61	0.38	0.47	0.75	0.25
K	0.79	0.08	0.34	0.56	0.74	0.26
Cl^-	0.34	0.22	0.36	0.29	0.30	0.70
HCO_3^-	0.79	0.07	0.22	0.52	0.68	0.32
SO_4^{2-}	0.43	0.56	0.26	0.45	0.57	0.43

Table 3. Rotation matrix

	F1	F2
F1	0.987	0.162
F2	0.162	-0.987

APPLICATION

Factor analysis is suitable statistical technique for identification of contamination water. In this study, we applied this method for the source apportionment of contamination of ground water.

CONCLUSION

Factor analysis was applied to ground water samples collected and the results showed that a three factor which explains >60% of the variation in quality of ground water. The data indicate the ground water of studied locations is highly deteriorated as it is polluted with high amount of anthropogenic causes agricultural. This study shows that Primary analysis of major ions shows that the groundwater is Calcium-bicarbonate water. Q-mode distinguishes between wells according to the similarity of geological formations. This could also provide crude guidelines for selecting the priorities of possible preventative measures in the proper management of surface water.

REFERENCES

[1]. B. A. Helena, M. Vega, E. Barrado, R. Pardo, L Farnandez, A Case of Hydro chemical Characterization of An Alluvial Aquifer Influenced by Human Activities, *Water Air Soil Pollut.*, **1999**, 112, 365.

[2]. K. P. Singh, A. Malik, D. Mohan, S. Sinha, Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)-a case study. Water Res. **2004**, 38, 3980-3992.
[3]. K. Zeinalzadeh, E. Rezaei, Determining spatial and temporal changes of surface water quality using principal component analysis, Jo*urnal of Hydrology: Regional Studies*, **2017**, 13, 1-10
[4]. P. Praus, Water quality assessment using SVD based principal component analysis of hydrological data, *Water SA*, **2005**, 417-422.
[5]. P. Karthikeyan, R. Venkatachalapathy, G. Vennila, Multivariate analysis for river water quality assessment of the Cauvery River, Tamil Nadu, India, *Ind. J. Geo Mar. Sci*., **2017**, 46, 785-790.
[6]. Hussein Ewaida, Salwan Ali Abed, Water quality index for Al-Gharraf River, southern Iraq, *Egyptian Journal of Aquatic Research*, **2017**, 43, 117-122.
[7]. K. Bengraine, T. F. Marhaba, Using principal component analysis to monitor spatial and temporal changes in water quality, *J. Hazard. Mat*., **2003**, 100, 179-195.
[8]. E. de Miguel, J. F. Llamas, E. Chacon, L. F. Mazadiego, Sources and pathways of trace elements in urban environments: a multi elemental quantitative approach, *Sci. Total Environ*., **1999**, 235, 355–357.
[9]. Mansi Tripathi, Sunil Kumar Singal, Use of *Principal Component Analysis* for parameter selection for development of a novel *Water Quality* Index: A case study of river Ganga India, *Ecological Indicators*, 96, Part 1, **2019**, 430-436..
[10]. Jun Xiao, Lingqing Wang, Li Deng, Zhangdong Jin, Characteristics, sources, *water quality* and health risk assessment of trace elements in river *water* and well *water* in the Chinese Loess Plateau, S*cience of The Total Environment*, 650, Part 2, **2019**, 2004-2012.
[11]. V. K. Jena, Assessment of Ground Water Quality and Source of Contamination by Multivariate Factor analysis, *Int. J. Water Res*., **2013**, 1, 43-46.
[12]. F. T. Patel, S. M. Mitchla, A. S. Patel and G. M. Malik, Seasonal Variation of Groundwater Quality of Coastal Villages of Olpad and Choryasi Taluka, Surat District, Gujarat, India, *J.Applicable Chem.*, **2019**, 8(4), 1946-1952.
[13]. Debin Tian, Cuixia Li, Risk assessment of raw milk *quality* and safety index system based on primary component analysis, Sustainable Computing: Informatics and Systems, **2019**, 21, 47-55.
[14]. R. T. Raghunath, R. Sreedhara Murthy, B. R. Raghavan, The utility of Multivariate statistical techniques in hydrogeochemical studies: an example from Karnataka, India, *Water Res*., **2002**, 36, 2437–2442.
[15]. Salam Hussein Ewaid, Salwan Ali Abed, Safaa A. Kadhum Predicting the Tigris River water quality within Baghdad, Iraq by using water quality index and regression analysis, *Environ. Technol. Innovat.* **2018**, 11, 390-398.
[16]. M. K. Ghosh, Removal of Toxic Metal Lead from the Surface Water of Naya Raipur and Its Impact on Human Health, *J. Applicable. Chem*., **2021**, 10(2), 212-219.
[17]. WHO Guidelines for drinking water quality, World Health Organisation, Washington DC, 333-335, **1984**.
[18]. J. Suresh P. Venkateswarlu, A. V. L. N. S. H. Hariharan, Studies on the Phisico-chemical characteristics of water in Pidaparthi village [Anaparthi mandal] E.G.Dt., A.P., *J. Applicable. Chem*., **2012**, 1(3), 427-432.