The area of study receives great quantities of agricultural, industrial, municipal and domestic wastes and the rainfall. The present study aims to estimate the distribution of some heavy metals in the groundwater that is useful in assessing possible chemical pollution and perceiving pollution sources. Groundwater samples were collected from 32 water wells in the area of study. Concentrations of some heavy metals (Pb, Mn, Cd, Cu, Zn, and Fe) were determined in these samples. Salient mean concentration levels were recorded for: Fe (174.1 µg/L), Mn (79.5 µg/L), Zn (28.48 µg/L), Cu (12.02 µg/L), Pb (6.277 µg/L) and Cd (1.006 µg/L). Overall, the decreasing metal concentration order was: Fe > Mn > Zn > Cu > Pb > Cd. Positive correlations (r) were found between Zn-Cu (r = 0.809), Cd-Pb (r = 0.471) at p 0.01 and Fe-Pb (r = 0.392) at p 0.05 level. The contour maps showed the spatial distribution of the studied heavy metals in the study area. The results of correlation analyses suggest that geographical as well as chemical factors may influence metal distribution in the groundwater samples.

**Keywords:** Groundwater, Correlation analysis, Cluster analysis, Heavy metals, Chemical pollution

**INTRODUCTION**

Groundwater is an important water supply source worldwide. It is the major source of drinking water in urban and rural areas in Yemen. Different studies on groundwater quality with respect to drinking water have been carried out in the area of study by Van der Welle (1997), Metwali (2003), Saeed (2004) and El Sharabi (2011). Groundwater in the area of study has been contaminated by enormous amounts of agricultural wastewater that carry various chemical pollutants related to the widespread use of fertilizers and pesticides. Besides, great quantities of industrial, municipal and domestic wastes and street dusts are drained, indirectly, into the groundwater. Hence, groundwater quality is closely linked to the point in time of recharge and the residence time of groundwater in the subsurface. In this paper an attempt has been made to evaluate some heavy metals in the groundwater aiming to recognize their environmental impact in the area of study.
However, in the wake of recent industrialization and fast urbanization, the quality of groundwater has become an increasing concern due to contamination by various toxic substances (Mance, 1987; Aiyesanmi et al., 2004; Amajor, 1986; Ezeigbo, 1989; Calderon, 2000; Jha et al., 1990; Ramesh et al., 2000). Many multivariate analysis techniques have been applied in hydrological studies: R-mode analysis in groundwater quality studies (Grande et al., 1996; Chen-Wuung et al., 2003; Panagopoulos et al., 2004; Garcia-Rodriguez et al., 2007); R-mode, Q-mode and cluster analysis to assess surface-groundwater interaction and groundwater mixing (Reg hunath et al., 2002).

LOCATION

The disposition of study area is given as Figure 1. The Taiz city, lying close to the Red Sea, is home to around 700,000 permanent residents. Taiz, a culturally and economically developed city and experiences a tropical climate with an average annual rainfall of 600 mm. The intensity of rainfall differs with topographic position in the area of study. Geologically, the Tertiary volcanic rocks (Tb and Tr) and Quaternary alluvium deposits (Qw) covered most of the area added to the Tertiary granite (Tgr) which intruded south of the area (Figure 1). According to Kruck et al. (1996), the Tertiary volcanic rocks ranges between basic and acidic facies such as basalt, rhyolite, dacite, ignimbrite, and ash flow deposits. The Tertiary granite plutons of J Sabir was intruded in the late Miocene. Quaternary deposits are represented by alluvium deposits derived as a result of weathering and erosion of the surrounding volcanic and granitic rocks and occur in the bottoms of wadis which dissected the area of study.
MATERIALS AND METHODS
A total of 32 samples (Figure 1) were collected from the water wells. Concentrations of some heavy metals were measured using Buck Scientific Atomic Absorption Spectrophotometer 210 VGP with graphite furnace. All reagents used were of AAS grade (certified purity > 99.9%). Distilled water was used both for the preparation of standards and the dilution of samples. Triplicate sub-samples of each sample were aspirated separately to compute average metal concentrations in a given water sample. The concentrations of the elements were expressed in micrograms per litre (µg/L). Statistical analysis of data, including a bivariate method (Pearson correlation analysis) and two multivariate techniques (Principal Component Analysis and hierarchical cluster analysis) were carried out using the software, SPSS (ver. 16.0) software for windows and graphs plotted using Sigma Plot 11.0 software. Pearson correlation was used to study the correlation structure between variables to understand the distribution of parameters. The Principal Component Analysis (PCA) has emerged as a useful tool for better understanding the relationships among the variables (metal concentrations in the present study) and for revealing groups (or clusters) that are mutually correlated within a data body. On the other hand, Hierarchical agglomerative Cluster Analysis (CA) was performed on the normalized data set by means of Ward’s method, using squared Euclidean distances as a measure of similarity.

RESULTS AND DISCUSSION
Table 1 provides mean metal concentrations in the selected groundwater samples, along with other relevant statistical distribution parameters. The decreasing trend of average metal levels was as follows: Fe > Mn > Zn > Cu > Pb > Cd. A comparison of the listed means, medians and standard deviations are shown in Table 1, we see a large dispersion around the mean of the metal levels which is indicated that Zn, Cu, Cd, Mn and Pb have almost randomized distribution. Also, the differences in the parameters explain the data are not distributed as a normal distribution. The substantial differences in the symmetry parameters in the case of Fe indicated a non-normal distribution, thus supporting a possibility of random infiltration of the metals from some anthropogenic source. Zn, Fe, Cu, Mn and Pb shows large standard deviations revealed their randomly fluctuating concentration levels in the groundwater. Correlations (r) between heavy metals were carried out using Pearson correlation (Table 2). A correlation analysis is a bivariate method applied to describe the degree of relation between the variables.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Median</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zn</td>
<td>8.69</td>
<td>169.75</td>
<td>25.63</td>
<td>28.48</td>
<td>26.60</td>
</tr>
<tr>
<td>Fe</td>
<td>24.0</td>
<td>590.0</td>
<td>142.0</td>
<td>174.1</td>
<td>120.0</td>
</tr>
<tr>
<td>Cu</td>
<td>4.17</td>
<td>49.67</td>
<td>9.25</td>
<td>12.02</td>
<td>8.64</td>
</tr>
<tr>
<td>Cd</td>
<td>0.470</td>
<td>1.978</td>
<td>0.973</td>
<td>1.006</td>
<td>0.299</td>
</tr>
<tr>
<td>Mn</td>
<td>22.0</td>
<td>360.0</td>
<td>71.5</td>
<td>79.5</td>
<td>60.4</td>
</tr>
<tr>
<td>Pb</td>
<td>0.052</td>
<td>10.820</td>
<td>6.625</td>
<td>6.277</td>
<td>2.640</td>
</tr>
</tbody>
</table>

This article can be downloaded from http://www.ijerst.com/currentissue.php
between two parameters. A high correlation coefficient (near 1 or −1) indicates a good relationship between two variables whereas its value around zero means no relationship between them at a significant level (<0.05). More precisely, Parameters showing \( r > 0.7 \) are considered to be strongly correlated while \( r \) between 0.5 and 0.7 shows moderate correlation.

Table 2 shows strong positive correlation between Cu and Zn at 0.01 level, indicating the existence of a common source/origin of these metals in groundwater samples under investigation. This provided the primary evidence for a mutual concentration dependence of the metals in the aquatic system (Tariq et al., 2008).

Principal Component Analysis (PCA) procedure reduces overall dimensionality of the linearly correlated data by using a smaller number of new independent variables, called Principal Components (PC), each of which is a linear combination of originally correlated variables. On the other hand, CA exclusively classifies a set of observations into two or more unknown groups based on combination of internal variables. The purpose of CA is to discover a system of organized observations where a number of groups/variables share properties in common, and it is cognitively easier to predict mutual properties based on an overall group membership (Everitt, 1993). This helps define source profiles of variables, such as metal concentrations, and their interpretation in terms of possible sources (Jobson, 1991).

Accordingly, using Varimax normalized rotation, the PCA was conducted for source identification. The rotated Principal Component Loadings are given in Table 3. Three Principal Components (Eigenvalues >1) emerged with more than 80% of cumulative variance. The first PC with 31.9% variance showed higher loadings for Cu and Zn with significant contributions from Pb and Cd. The second PC, with 26.2% of total variance, had significant contribution from Cu and Zn. The third PC had higher loadings for Cd at 22.3% of total variance. All of these components could be conceived to mainly originate from industrial effluents and influent waste discharged from other industries in the area.

The following cluster (Hierarchial analysis) analysis dendogram (Figure 2) used for examining the similarity among the samples of Q- mode cluster and for evaluating interaction among the variable, R- mode cluster was used. Figure 2a shows the Q-mode cluster grouping all 32 water samples.
**Table 3: Varimax Normalized Rotated Principal Component Loadings of Selected Metals in Groundwater Samples (n=32)**

<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pb</td>
<td>0.383</td>
<td>-0.551</td>
<td>0.142</td>
</tr>
<tr>
<td>Mn</td>
<td>0.083</td>
<td>-0.138</td>
<td>-0.719</td>
</tr>
<tr>
<td>Cd</td>
<td>0.265</td>
<td>-0.394</td>
<td>0.530</td>
</tr>
<tr>
<td>Cu</td>
<td>0.609</td>
<td>0.340</td>
<td>-0.081</td>
</tr>
<tr>
<td>Fe</td>
<td>0.145</td>
<td>-0.550</td>
<td>-0.416</td>
</tr>
<tr>
<td>Zn</td>
<td>0.620</td>
<td>0.323</td>
<td>-0.042</td>
</tr>
<tr>
<td>Eigen value</td>
<td>1.9139</td>
<td>1.5742</td>
<td>1.3398</td>
</tr>
<tr>
<td>% Total variance</td>
<td>31.9</td>
<td>26.2</td>
<td>22.3</td>
</tr>
<tr>
<td>Cumul. Eigen value</td>
<td>1.9139</td>
<td>3.4881</td>
<td>4.8279</td>
</tr>
<tr>
<td>Cumulative %</td>
<td>31.9</td>
<td>58.1</td>
<td>80.5</td>
</tr>
</tbody>
</table>

**Figure 2: Dendogram Showing Q Mode (a) and R Mode (b) Cluster of the Selected Metals in Groundwater Samples**

wells where sampling sites are located into three significant clusters. It is observed that each group represents similar hydro-chemical environment. Samples belonging to cluster 1 (samples no.3, 10, 12, 13, 23, 27, 28, 29 and 30) corresponds to highly polluted sites. Cluster 2 (Samples no. 1, 5, 6, 9, 11, 14, 20, 21, 22, 24, 25, 26 and 31) corresponds to the moderately polluted sites. Cluster 3 (Samples no. 2, 4, 7, 8, 15, 16, 17, 18, 19 and 32) is very much close to cluster 2 with corresponds to relatively less polluted sites excepting that concentrations of certain trace elements is higher. Some water samples in these sites have shown elevated concentrations of Fe,
Figure 3: Spatial Distribution of the Heavy Metals

- **Fe**
- **Cu**
- **Cd**
- **Mn**
- **Pb**
- **Zn**
Mn, Pb, and Zn mainly from anthropogenic sources and some water wells receive pollution from industrial activities. The most common sources of iron and manganese in the selected groundwater samples in the area of study are due to industrial effluent, acid-mine drainage, sewage and landfill leachate. Concentrations that exceed the drinking water guideline can occur locally anywhere in the province. On the other hand, the mainly source of lead in highly values in some of the analyzed water samples is due to waste incinerating and waste of lead-acid batteries which release lead into the air and soil.

R-mode cluster shows 2 major clusters, cluster 1 and cluster 2. Cluster 1 has 2 subdivisions, shows Pb, Cd, Mn and Fe form one group and the rest of the parameters form separate cluster (2), (Figure 2b). The R- mode cluster showed good agreement with PCA confirmed that the metals with common PC and strong mutual correlations formed primary clusters.

The following contour maps (Figure 3) show the spatial distribution of the analyzed heavy metals, in which an anomalous area of higher values can be distinguished. Based on the obtained information from the spatial distribution maps, it is possible to design a future, optimal sampling strategy for monitoring programs, which could reduce the number of sampling points and associated costs.

CONCLUSION

A comparison of the listed means, medians and standard deviations have a large dispersion around the mean of the metal levels which is indicated that Zn, Cu, Cd, Mn and Pb have almost randomized distribution. Strong positive correlation between Cu and Zn at 0.01 level, indicating the existence of a common source/origin of these metals in groundwater samples under investigation. Multivariate statistical methods of factor analysis are shown to be an important tool for characterizing hydrogeochemical processes and clustering groundwater according to their shared hydrochemical characteristics.

The PCA and CA identified industrial effluents, to be the probable source of the observed metal pollution of the groundwater. Some water samples have shown elevated concentrations of Fe, Mn, Pb, and Zn mainly from anthropogenic sources, arising from agricultural activities, Electroplating materials and lubricants used near the area of study. It is anticipated that the results of the present study would provide the requisite basis towards formulating a future industrial pollution abatement program by the concerned authorities.

REFERENCES


