ISSN: 2639-7455

Research Article

Earth & Environmental Science Research & Reviews

Using Statistical Analysis to Assess Shallow Groundwater Quality in Squatter Settlement of Dar es Salaam and Pwani Regions in Tanzania

NAJAT ZAHOR SAID[1], HASSAN RASHID ALI[1], LYDIA NDIBALEMA[2] AND SAID SULEIMAN BAKARI[1]

[1]*State University of Zanzibar (SUZA), Department of Natural Science P.O BOX 146, Zanzibar- Tanzania.*

[2]*Water Supply Project Development Manager, Dar es Salaam Water Supply and Sewerage Authority (DAWASA) P.O.BOX 1573, Dar es Salaam, Tanzania.*

[*]Corresponding Author
SAID SULEIMAN, BAKARI, State University of Zanzibar (SUZA) Department of Natural Science P.O BOX 146, Zanzibar- Tanzania.

Submitted: 2023, June 17; **Accepted**: 2023, July 21; **Published**: 2023, Aug 09

Citation: SAID, N. Z., ALI, H. R., NDIBALEMA, L., BAKARI, S. S. (2023). Using Statistical Analysis to Assess Shallow Groundwater Quality in Squatter Settlement of Dar es Salaam and Pwani Regions in Tanzania. *Envi Scie Res & Rev, 6*(4), 554-565.

Abstract
In Dar es Salaam and the Coastal regions, groundwater is the primary source of freshwater. Urbanization, industrialization, and agricultural activities have all contributed to the overuse of this resource, resulting in a decline in quality. In this study, we employ statistical analysis as a means to improve our understanding on groundwater quality. Seventy-one groundwater samples in squatter areas of Dar es Salaam and Coastal regions were used with 10 chemical parameters. Based on their dominant chemical compositions, hierarchical cluster analysis (HCA) classified samples into three main clusters. The clusters were then characterized with the help of the results from principal component analysis (PCA), Pearson's correlation coefficient of the chemical variables and piper diagrams. Cluster one was characterized by high salinity as a result of seawater intrusion. Cluster two was characterized by an intermediate salinity and high contamination. This was a result of sewage contamination and a moderate effect of seawater intrusion. With a low level of salinity and contamination, cluster three was found to be suitable for human consumption. The information obtained from this study demonstrates that the quality of the water has declined in some areas within the chosen regions. Therefore, prompt action is required to prevent further deterioration to the aquifers in the region.

Keywords: Groundwater; Multivariate statistical analysis; Dar es Salaam and Pwani Regions; Seawater intrusion.

1. Introduction

Many large cities in developing countries face significant challenges when it comes to access to clean water. Dar es Salaam and Coastal regions are the fastest growing city in Tanzania with population of more than 5 million. Over 65 percent of the population are living in squatter areas where water supply and other infrastructures are inadequately. The pressure on the existing water supply in the city is ever increasing due to the rapid urbanization and high population growth. Due to the shortage of the clean water, people have been looking for different sources including surface and groundwater. Being a water resource that a large number of people depend on it, groundwater needs to be preserved so that it continues to be useful. This demands regular analysis of its quality so as to detect any pollutants and pollution sources. In addition, recharging of groundwater is a long process, hence, the best option is to minimize and control the sources of contaminations. The information from the analysis of groundwater quality are useful for the water managers as they can monitor the level of pollution in a groundwater. This will help them to know what kind of treatment to use, or how safe it is to be used by humans. Despite of this importance of groundwater resource, water quality assessment and its suitability are not sufficiently monitored. Recharging zone of these resources is exposed with complex waste materials and contaminants, such as agricultural activities, and accumulation of wastes disposed on land. Furthermore, poor drainage system is the main sources of water pollution in squatter areas of Dar es Salaam and Coastal Regions [1]. These contaminants jeopardize the quality of the groundwater resources and its future use as the main source of water for human consumption. The combination of these problems increasing the pressure on groundwater quality and the needs for regular monitoring of water resources. On study conducted by [1]. revealed that some shallow water were polluted by both biological and chemical contaminants and hence are not suitable for human consumption.

Therefore, the aim of this study is to assess the status of shallow groundwater quality in squatter settlement of Dar es Salaam and Coastal regions by using statistical methods. Multivariate statistical methods are useful tools in studying the quality of groundwater. They have been successfully used to assess hydro-geochemical data. Examples of the effective utilization of multivariate statistical methods in hydro-geochemical studies are contained in [2] and [3]. To achieve that goal, we have the following specific objectives: classify groundwater samples into groups that are governed by the same hydro geochemical processes, identify the main factors affecting groundwater quality, determine the inter-relationship of the chemical parameters of the groundwater samples.

2. Materials and methods

2.1 Study site

Dar es Salaam is the largest commercial city of Tanzania located beteen latitude 6.36° and 7.0° to the south of equator and longitudes 39.0° and 44.4° the east of Greenwich. It is bounded by the Indian Ocean on the eastern side and the rest is bounded by the coastal region.

The city experiences a modified type of equatorial climate which is normally hot and humid throughout the year, with an average temperature of 29℃. The total surface area of this city is 1800 km^2 with a population of 4.36 million [4]. It is the most populated among the 31 regions of Tanzania, which accounts for 10% of the total population for the Tanzania mainland [5]. It is expected to expand by more than 85% through 2025 and could reach up to 21.4 million people by 2052 [6].

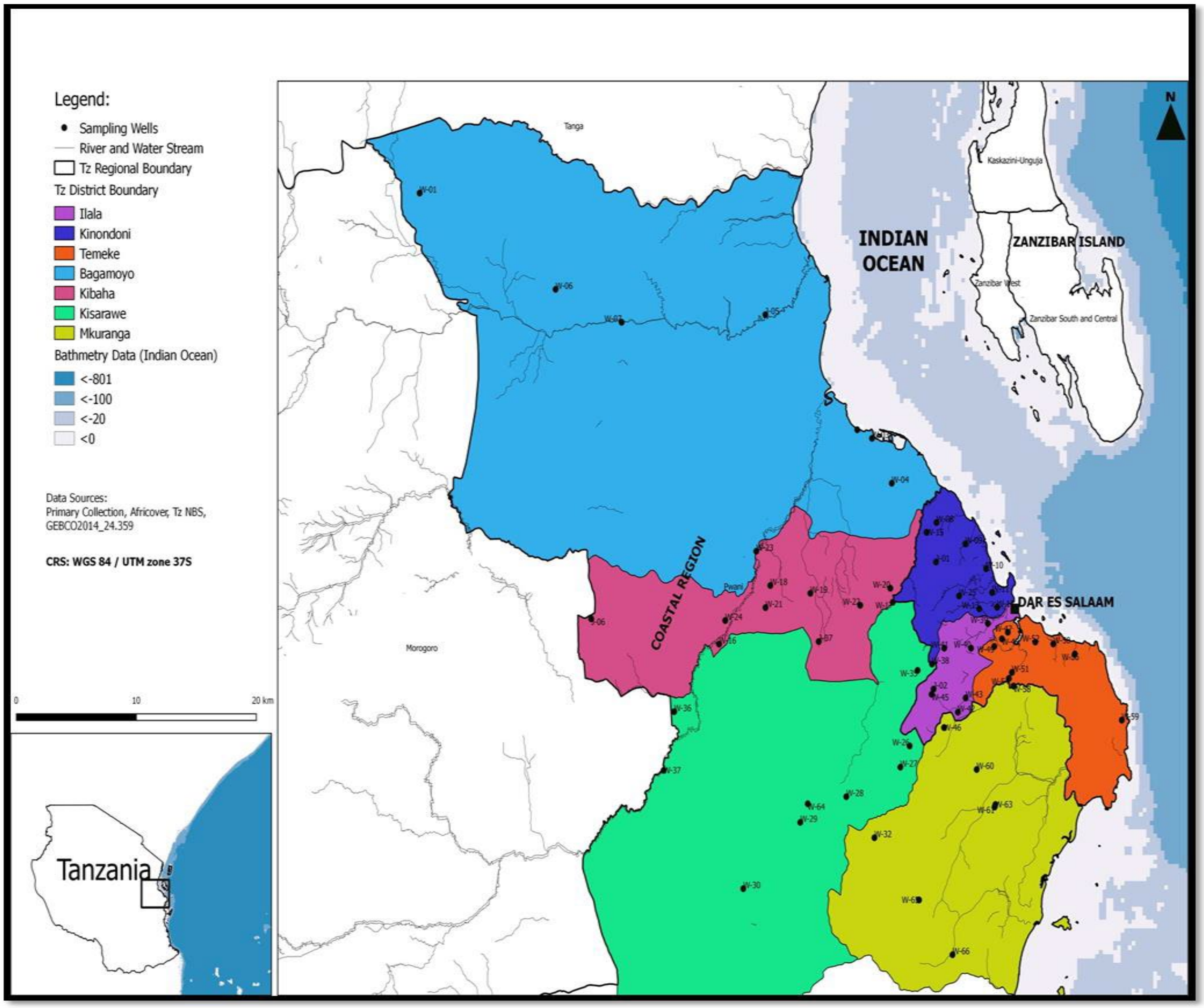

Figure 1: Map showing sampling locations for wells. The selected wells are from Squatter areas and along coastal part of Dar-es-salaam. Colors used to separate the districts and they are defined in the legend part of the figure above.

2.2 Source of data

The data used for this study were the measurements of chemical variables of groundwater for the seventy one samples collected by Dar es Salaam Water and Sewerage Authority (DAWASA). These samples were taken from shallow wells with depth ranging from 30 meters to 91 meters. The sampling points were far away from each other and so they were considered independent (Figure 1). All the measurements were taken in the same season so that there was no temporal variation. The physicochemical variables used were temperature of water, depth of the wells, pH, electrical conductivity (measure of groundwater salinity, EC), sodium (Na), magnesium (Mg), calcium (Ca), potassium (K), chloride (Cl), sulphate (SO_4), nitrate (NO_3) and bicarbonate (HCO_3). Electrical conductivity, pH and temperature of water samples were measured immediately after sampling in the field. Each of the 71 samples was measured at Ardhi University of Dar es Salaam for their major anions and cations. All the measurements were checked for their accuracy by calculating percent charge balance error (%CBE) using the following equation,

$$\%CBE = \frac{\sum cations - \sum anions}{\sum cations + \sum anions} \times 100.$$

All samples had errors within ±10%, hence all 71 samples were used in the analysis.

3. Theory

As a branch of mathematics, statistics has been used widely to solve real life problems. Being a daily life issue that needs to be solved, groundwater quality classification is a newly emerging issue where statistics can be applied. Many variables involved in analyzing per groundwater sample makes it an intrinsically a multivariate problem.

The multivariate statistical analysis involves the study on how the variables are related to one another and how they work in combination. There are many statistical techniques for conducting multivariate analysis and the best appropriate techniques for a given study varies with the type of study and the key objectives of the research. In this research essay, the multivariate statistical analysis techniques were implemented to assess the shallow groundwater in the squatter areas of Dar es Salaam and Coastal Regions. Hierarchical cluster analysis and principal component analysis were used to classify groundwater quality by using R statistical software package. The dataset, in this case, can be summarized by data matrix X or simply called designed matrix. Hence,

$X = (x_{ij})_{i=1},\ldots,n; j=1,\ldots,p.$

Where *n* is the number of groundwater samples and p is the number of chemical variables.

3.1 Hierarchical Cluster Analysis (HCA)

Cluster analysis deals with discovering groupings within data matrix, X. The main goal is to find an optimal grouping for the observations or objects within each cluster. Grouping data is important because it can reveal useful information about the data. There are several types of cluster analysis but for this case we use hierarchical agglomerative (bottom-up) clustering. It treats each observation as a singleton cluster and then successively joins the most similar observations to form new clusters until all clusters are merged into a single cluster (that contains all observations).

The distribution of the dataset used in this study is positively skewed (Table 1). Therefore, the first thing was to do the log transformation followed by standardization. Standardization simply means rescaling of data to have a mean of zero and a standard deviation of one. The main reason for standardization is to make sure that all variables contribute evenly to a scale when items are added.

3.2 Principal Component Analysis (PCA)

The principal component is a technique of identifying patterns in data and expressing the data in a way that similarities and differences among the dataset are highlighted. These patterns help to compress the data in the sense that the number of dimensions is reduced without losing much information.

Mathematically, principal components are new variables formed from the linear combination of the original variables. The aim of this analysis is to generate new variables (and fewer) that are capable of explaining most of its variability [7]. The new variables (components) are called the eigenvectors of the covariance matrix (or correlation matrix) and are orthogonal (perpendicular to one another). To get these principal components, correlation matrix was formed from the designed matrix X by using R statistical software package.

4. Results

4.2 Physicochemical Characteristics of Groundwater

The physicochemical compositions of the water samples are summarized in Table 1. The chemical variables (except pH) have values that are sparsely spread out within them. Additionally, the chemical variables are all positively skewed which indicates that the data are not normally distributed. This reflects the variations in the groundwater quality of the study area. The water samples had average temperatures during sampling of around 29℃ which reflects the mean annual temperature of the study area. pH values ranging from 5 to 8.4 were found, with a mean pH of 6.7 indicating that the groundwater samples were slightly acidic to alkaline.

The order of abundance of the cations is $Na^+ > C^{a2+} > Mg^{2+} > K+$ and the order of abundance of the anions is $Cl^- > SO_4^{2-} > HCO_3^- > NO_3^-$. Additionally, 29% and 18.3% of all samples exceeding the desirable limit of Na^+ (200mg/L) and Mg^{2+} (50mg/L) respectively; and 35.2%, 7% and 1.4% of all samples exceeding the permissible limit of Cl^- (250mg/L), NO_3^- (50mg/L) and SO_4^{2-} (250mg/L) respectively as given by WHO (WHO, 1996) for drinking water. This means that the groundwater quality in some areas is not suitable for human consumption.

4.2 Multivariate Statistical Analysis of Groundwater Samples

4. 2.1 Hierarchical Cluster Analysis

HCA was performed in order to group the samples into a fnite number of clusters with similar hy-drochemical compositions and hence inferred similar processes governing their physiochemistry. In this study, HCA was performed by using a combination of Euclidean distances as a measure of dissimilarity and the Ward's linkage method.

Table 1: Descriptive statistics for the 71 groundwater samples. Each sample has been measured its Temperature and other chemical variables presented in this Table.

Parameter	Unit	Mean	Maximum	Minimum	Standard deviation	Skewness
Temp	(℃)	28.6	40	25	2.18	
pH		6.7	8.3	5.2	0.8	0.05
EC	(μ/cm)	2088.1	9490	330	1849.1	2.64
Na^+	(mg/L)	306.8	1700	50	318.7	3.40
K^+	(mg/L)	11.1	50	3.5	9.9	2.94
Mg^{2+}	(mg/L)	45.3	163	3.2	42.8	1.43
Ca^{2+}	(mg/L)	73.1	184	4.8	60	0.35
NO_3^-	(mg/L)	23.7	316	1.3	59.9	4.83
SO_4^{2-}	(mg/L)	159	1600	30	293.4	4.50
HCO_3^-	(mg/L)	114.1	588	12	130.3	2.26
Cl^-	(mg/L)	559.2	1970	78	420.5	1.46
Depth	(m)	58.4	91	30	17.1	0.02

Based on the measurements of the chemical variables, the groundwater samples were classified by HCA into two main clusters (cluster one and cluster two) conforming to their dominant chemical compositions (Figure 3). The dendrogram was plotted using R statistical software. The number of clusters was chosen based on the minimum number of clusters that describe most of the variation in hydro-geochemical properties of the water samples.

Based on concentrations of major ions and EC values, cluster one is separated from clusters two and three, indicating that samples in this cluster are distinct from the rest. It contained 27% of all groundwater samples with Na^+ , Ca^{2+}, Mg^{2+}, Cl^- and SO_4^{2-} as highly dominated ions with also high values of EC (Table 1 and Figure 2). Therefore, this cluster is characterized by high salinity.

Cluster two is formed with 51% of water samples having intermediate EC values, other ions and high concentration of NO_3^- . 21% of all the samples in this cluster have exceeded the permissible limit of NO_3^- (50mg/L) for drinking water, as given by WHO [8]. Hence, this cluster can be categorized as contaminated group and has intermediate salinity.

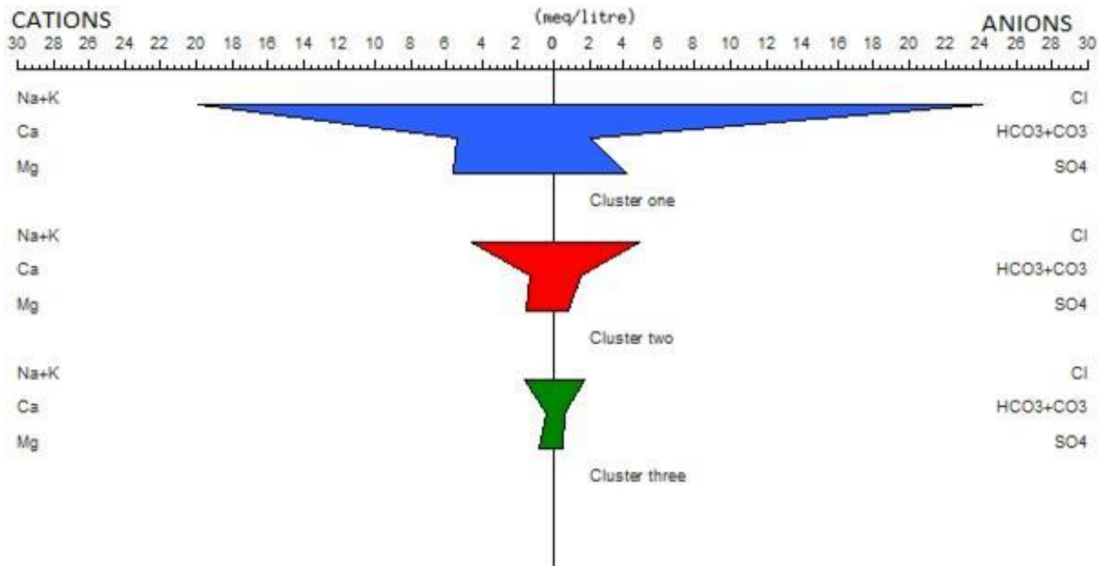

Figure 2: Stiff diagram showing chemical variations of samples within three clusters formed by HCA

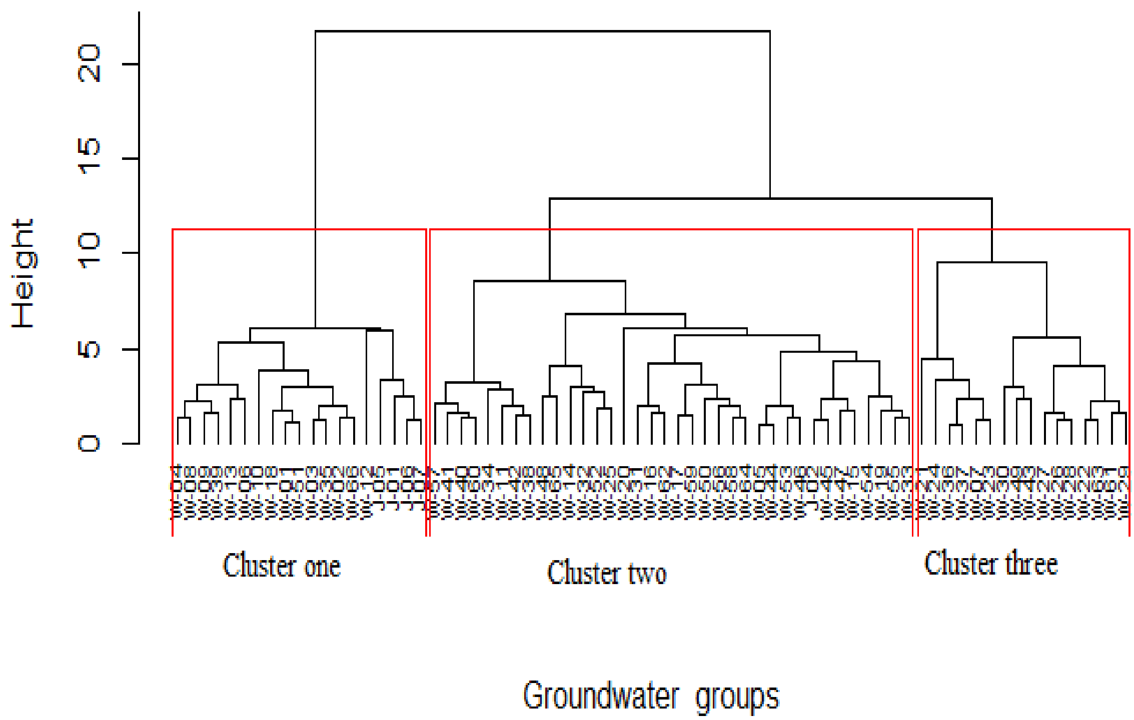

Figure 3: Dendrogram Showing Three Clusters of Groundwater Samples. Clearly, Three Clusters Were Formed After The Analysis

The third cluster contains lowest concentrations of major ions, EC values and all concentration of NO_3^- are less than 50mg/L compared to the other groups. Hence, samples in this cluster have the lowest salinity and are less contaminated.

4.2.2 Principal Component Analysis

The PCA was performed on a standardized data set of the water samples with 10 chemical parameters to identify the main variables affecting groundwater quality. In this study, the FactoMineR package which is the latest package for multivariate data analysis in R statistical software was used. The main feature of this package is the ability to consider different types of variables like quantitative or categorical, a different type of structure on the data (a partition on the variables, a hierarchy on the variables and a partition on the individuals) and lastly supplementary information (supplementary individuals and variables). For purposes of this study, clusters produced by HCA were entered as qualitative supplementary variables, EC as a quantitative supplementary variable, pH and measurement of ions as active variables, and all groundwater samples as active individuals.

In FactoMineR, the correlation coefficient (and its significance test) between the variable and the coordinates of the individuals on the axis is first calculated for the quantitative variable. Then, the significant variables are sorted out from the most to least correlated and each dimension is well described by the variables. For categorical variables, each category is described by coordinates. The category with a coordinate significantly greater than zero for a particular PC, shows that it is well described by that PC.

By default, two graphs are produced by PCA which are Individual factor map and Variables factor map. Individual factor map represents rows (observations) as points and each point is coloured to indicate certain cluster membership. Clearly, separation by shape of our samples shows that there are three main clusters produced by HCA in the first two dimensions (Error! Reference source not found.). Variables factor map is the correlation circle showing the correlation between variables and the dimensions (principal components).

In this project, the number of PCs to be extracted is based on Kaiser criterion that explains only components with eigenvalues greater than or equal to 1 and the variables with p-values less than 0.05 were retained (Figure 5). Thus, to meet the mentioned criteria, three PCs were extracted which accounted for 78.31% of the total variance as presented in Table 2. The first two PCs: PC1 and PC2 explain 50.13% and 16.56% of the variance respectively, to account for the majority of the variability in the original data set (Figure 4 and Figure 6). The third principal component, PC3 explains 11.61% of the variance.

Table 2: Results for PCA of the groundwater samples. The p-value selected for the analysis was < 0.05

	PC1		PC2		PC3	
Quantitative variables	Correlation	p-value	Correlation	p-value	Correlation	p-value
EC	0.89	< 0.05				
Na^+	0.92	< 0.05				
K^+	0.63	< 0.05			0.63	< 0.05
Mg^{2+}	0.85	< 0.05				
Ca^{2+}	0.85	< 0.05				
NO_3^-			0.47	< 0.05	0.80	< 0.05
SO_4^{2-}	0.81	< 0.05	-0.33	< 0.05		
HCO_3^-	0.48	< 0.05	0.66	< 0.05	-0.34	< 0.05
Cl^-	0.94	< 0.05	-0.28	< 0.05		
pH	0.24	< 0.05	0.76	< 0.05		
Categorical variables	Coordinate		Coordinate		Coordinate	
Cluster one	2.75		-0.08		-0.33	
Cluster two	-0.20		0.34		0.34	
Cluster three	-2.64		-0.70		-0.42	
Eigenvalue	**4.512**		**1.490**		**1.045**	
% of variance	**50.133**		**16.599**		**11.614**	
Cumulative % of variance	**50.133**		**66.691**		**78.306**	

PC1 shows strong positive correlation with Cl^-, Na^+, EC, Ca^{2+}, Mg^{2+}, K^+ and SO_4^{2-}, moderate positively correlated with HCO^{3-}, and weak positively correlated with pH (Figure 7). These chemical parameters that govern 50% of the main variation in the original data set of groundwater samples (PC1) are dominant solutes in brackish and saline water. Hence, this principal component may be used to account for groundwater salinisation processes. Therefore, this component (PC1) describes the salinisation effect.

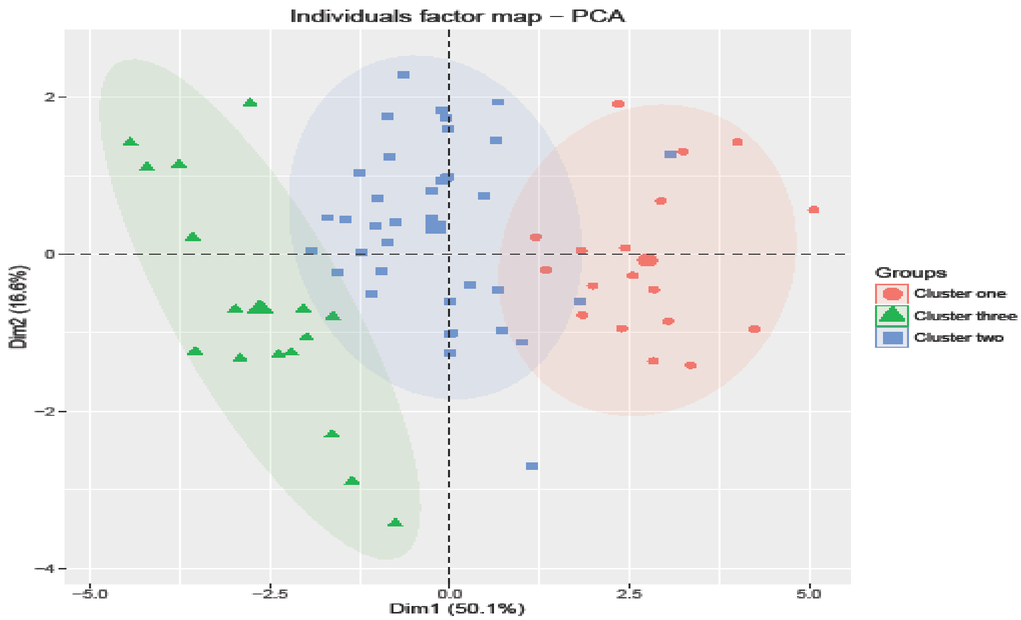

Figure 4: Individual factor map of PC1 and PC2.

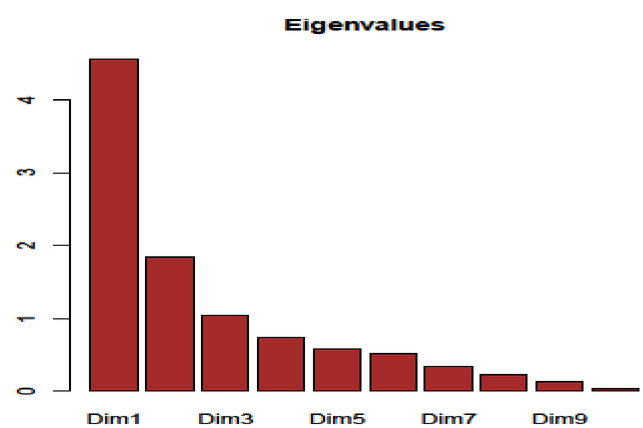

Figure 5: Barplot of the eigenvalues showing the first three dimensions having eigenvalue greater than one

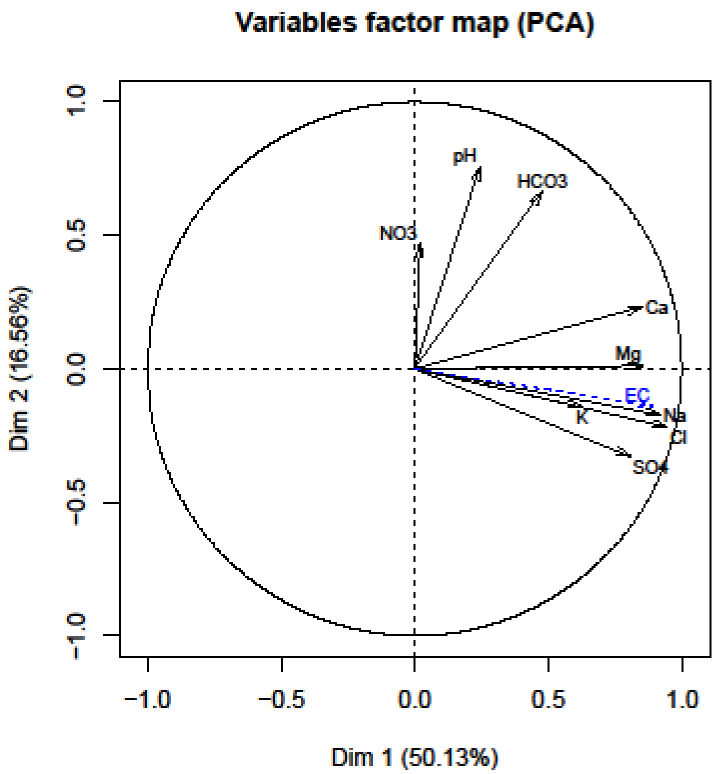

Figure 6: Variables factor map of PC1 and PC2.

Cluster one category has a coordinate significantly greater than zero in this dimension (), thus, PC1 is well described by this cluster. This means that most of the samples in cluster one are characterized by the ions that are correlated with PC1.

PC2 is characterized by strong positive correlation in HCO_3^- and pH, moderately correlated with NO_3^- and negatively weak correlated with SO_4^{2-} . HCO_3^- is an alkaline anion and so its pH value is higher than 7. The most predominant mineral compound causing alkalinity is Calcium carbonate ($CaCO_3$) which might come from rocks such as limestone or can be leached from dolomite and calcite in the soil. Therefore, PC2 can be termed a dimension that describes mineralization of groundwater by calcite weathering from rocks such as limestone or through soil-water interaction. The majority of the samples in cluster two describe this dimension as its coordinate is significantly greater than zero in this principal component.

PC3 exhibits strongly positive correlation in NO_3^- and K^+ but moderate negatively correlated in HCO_3^- . NO_3^- and K^+ are the main pollutants in sewage; they can be a result of poor septic tank systems that result in leakage of sewage which then enters groundwater directly since in the study area there are no agricultural activities. Negative correlation with HCO_3^- shows that there is a dilution of groundwater by recharged water. Therefore, the PC3 can be termed the dimension that describes sewage contamination and dilution of groundwater through recharge. The categorical variable that describes this dimension well is cluster two as its coordinate is greater than zero in this dimension.

4.2.3 Relationship among Chemical Parameters

Pearson's correlation coefficient was computed in order to find the inter-relationship among chemical parameters of the groundwater samples. The correlation coefficient of each pair of chemical variables was calculated and represented by an upper triangular correlation matrix. The corrplot package from R statistical software was used to display this correlation matrix in a graph called correlogram. Correlations with p-value greater than 0.01 were considered as insignificant and their values were left blank. Therefore, the results described in this section are significant only at 1% significance level.

From Error! Reference source not found., it can be observed that all correlation coefficients that are significant for this case are positive. Major ions including Na^+, Mg^{2+}, Ca^{2+}, Cl^-, SO_4^{2-} are strong correlated with each other (except Ca^{2+} and SO_4^{2-}). This indicates that they are derived from the same source of saline waters. Since both of these ions are strong correlated with EC then this is the characteristic of seawater intrusion. The non-significant relationship between Ca^{2+} and SO_4^{2-} indicates the absence of permanent hardness in water caused by Calcium sulphate ($CaSO_4$). Furthermore, there is a strong relationship between Mg^{2+} and Ca^{2+} and between Mg^{2+} and HCO_3^- which indicates dissolution of carbonate rocks. From Figure 7, it can be seen that there is no significant relationship between NO_3^- and other ions except K^+. This shows that K^+ and NO_3^- come from a different source of pollution and hence, there is more than one source of pollution in the groundwater. Since there are no agricultural activities in the study area, the main sources of nitrate in the groundwater samples could be waste water disposals like pit latrines and leaky sewers, solid waste disposal and industrial activities.

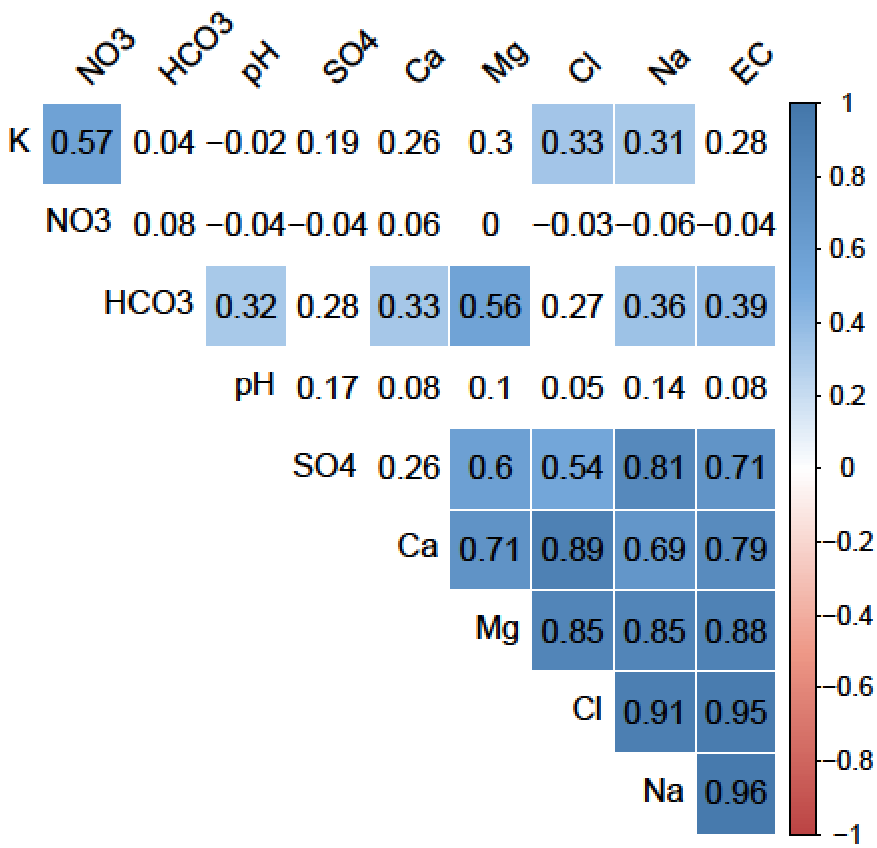

Figure 7: Correlogram for the chemical parameters of the groundwater samples.

4.2.4 Seawater Intrusion Indicator

The process by which sea water penetrates coastal groundwater and hence mixes with limited fresh water supply is referred to as seawater intrusion [9]. Normally HCO_3^- is abundant in groundwater while Cl- is found in small amounts in groundwater, but abundant in sea water. To identify whether the fresh groundwater has been intruded by sea water, one can compute the Cl-/HCO_3^- ratio recommended by Revelle in 1941 as cited in [10]. This ratio is considered an indicative measure of saltwater intrusion and it ranges between 0.13 and 198.70. Groundwater salinization can then be categorized according to the value of the ratio. For the groundwater having Cl-/HCO_3^- ratio less than 0.5, is considered unaffected, 0.5 - 6.6 for moderately affected and greater than 6.6 for strongly affected groundwater with saline.

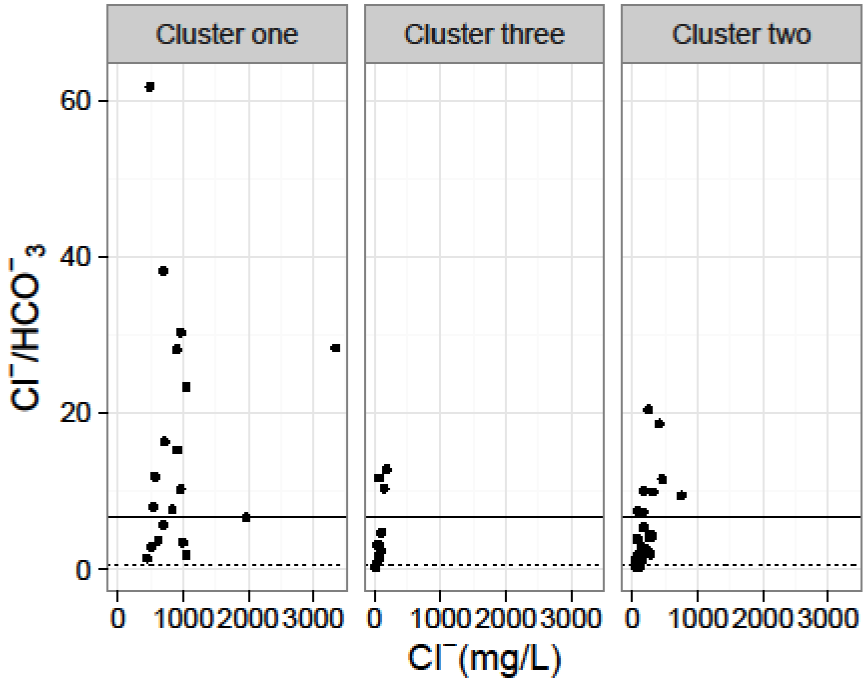

Figure 8: Ionic ratio of Cl-/HCO_3^- of the groundwater samples

In this study, 11.27% of groundwater samples has a Cl^-/HCO_3^- ratio of less than 0.5 (below the dotted line in Figure 8), 56.37% of groundwater samples has a Cl^-/HCO_3^- ratio between 0.5 and 6.6 (between the solid and dotted line in Figure 8) and 32.39% has Cl^-/HCO_3^- ratio greater than 6.6 (above the solid line in Figure 8). This indicates that 89% of the groundwater samples has been contaminated by sea water. The order of affected clusters by seawater intrusion is : Cluster one > Cluster two > Cluster three.

4.2.5 Hydro geochemical Facies of Groundwater Samples

Figure 9 reveals that almost all the groundwater samples in cluster one is of Na-K-Cl water type as they are found in lower right corners of anions and cations triangles. Additionally, it shows that samples in cluster one also contained mixed cations--Cl water type which includes Na-Mg-Ca-Cl, Na-Mg-Cl, Na-Ca-Cl and very few are of mixed cations--SO4 water type. On the other hand, cluster two contained groundwater samples which are of mixed--type; contained mixed cations and anions. This is clearly visible in the piper diagram in Figure 9. Additionally, we observe that all the samples in cluster two are located in the central zone of the diamond. This shows that there is cation exchange reaction process that occurs in the groundwater. Lastly, cluster three contains mixed--type of water with less concentration of Cl^- compared to cluster one and cluster two as its samples are scattered in the diamond.

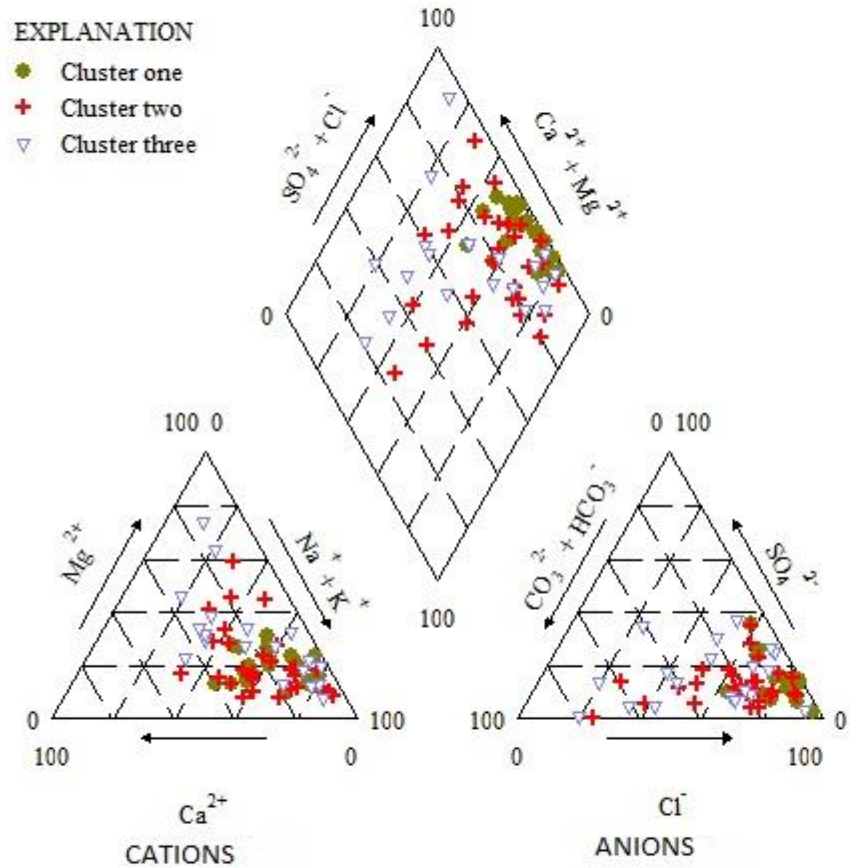

Figure 9: Piper diagram showing hydrochemical facies of study area. Groundwater samples are grouped in clusters generated by HCA.

5. Discussion

Based on the results from the groundwater samples, the study area is composed of rocks such as limestone, dolomite and gypsum that contribute to the concentration of Ca^{2+}, Mg^{2+} and HCO_3^- . The classification of groundwater samples generated by HCA indicated that there were three main groups of water according to their ions concentrations. Cluster one which was different from the other two clusters, contained 27% of all samples having a high proportion of chlorine, characterized by high salinity. The analysis showed that the salinisation process was due to the seawater intrusion. This reflects that shallow wells which formed cluster one, could be older wells that have been over exploited, since excessive pumping of groundwater increases the risk of seawater encroachment.

Cluster two had 51% of water samples with mixed--type of anions and cations. This cluster was characterized by intermediate salinity and high concentration of NO_3^- . From Figure 8, it can be seen that few samples from this cluster were highly affected by seawater intrusion.

Furthermore, there were some samples that exceeded the permissible limit of NO_3^- (50mg/L) as given by WHO [8], ranging from 61.1--316 mg/L (Table 1). This shows that these samples are highly contaminated and if not treated, then they are harmful and not recommended for human consumption. Drinking water with an excess level of nitrate can cause methemoglabinemia (blue baby disease) [11].

The rest of the samples formed cluster three which was characterized with low salinity and low NO_3^- concentration. This shows that the groundwater sources from this cluster are relatively suitable for human consumption.

6. Conclusions

The study has examined groundwater quality in squatter areas of Dar es Salaam and Pwani Regions in Tanzania. The groundwater was classified by the use of multivariate statistical techniques to determine the variability of groundwater quality. It also helped to identify major variables affecting groundwater quality. Three main clusters with different characteristics were discovered after performing HCA. Each cluster was then described according to the samples it contained. Cluster one with 27% of the samples was characterized by high salinity. Cluster two formed with 51% of water samples was characterized by moderate salinity and high contamination. Lastly, the samples forming cluster three were less contaminated and so are of better quality compared to other clusters.

PCA was then used to investigate the source of each water quality parameters due to nature and anthropogenic activities based on three clusters. Three PCs were extracted accounting for 78.31% of the total variance. Principal component one accounted for 50.13% of the variation of the data, appears to be from water quality parameters associated with saltwater intrusion process. Cluster one was found to describe this dimension well. The second largest source of variation (16.56%), PC2 is from water quality parameters associated with the dissolution of rocks. The third component described sewage contamination process. Cluster two was found to have a better description of the principal components two and three. On that account, it is noteworthy that PCA is very useful in determining sources of pollution.

Therefore, the results of this study show that the multivariate statistical analysis is a useful tool in classifying groundwater quality and identifying the major factors (pollution sources) that affect the groundwater quality.

7. Acknowledgements

We sincerely express our gratitude to the African Institute for Mathematical Sciences (AIMS) for the scholarship that allowed to do this project. We acknowledge DAWASA for their generous supply of data needed to accomplish this study. We also thank Ardhi University of Dar es Salaam for their help on chemical analysis of the data.

References

1. Saria, J., & Thomas, I. (2012, January 13). Water Quality in Selected Shallow Wells In Dar es Salaam. Huria Journal, XI.
2. Hosseinimarandi, H., Mahdavi, M., Ahmadi, H., Motamedvaziri, B., & Adelpur, A. (2014). Assessment of groundwater quality monitoring network using cluster analysis, Shib-kuh plain, Shur watershed, Iran. Journal of Water Resource and Protection, 6(6), 618.
3. Usman, U. N., Toriman, M. E., Juahir, H., Abdullahi, M. G., Rabiu, A. A., & Isiyaka, H. (2014). Assessment of Groundwater Quality Using Multivariate Statistical Techniques in Terengganu. Science and Technology, 4(3), 42-49.
4. NBS and OCGS. (2013). 2012 Population and Housing census. Dar es salaam: Population Distribution by Administrative Areas, United Republic of Tanzania.
5. NBS, & OCGS. (2013). 2012 Population Housing and Census. Dar es salaam: United Republic of Tanzania.
6. Abebe, F. (2011). Modelling Informal Settlement Growth in Dar es Salaam, Tanzania. PhD Thesis, University of Twente, Faculty of Geo-information Science and Earth Observation.
7. Rencher, A. C. (2003). Methods of Multivariate Analysis (Vol. 492). John Wiley & Sons.
8. WHO. (1996). Guidelines for Drinking-water Quality. World Health Organization, 2.
9. ACASA. (2011). Saltwater intrusion and climate change : A primer for local and provincial decision-makers. New Brunswick: Prince Edward Island, Department of Environnmet, Labour and Justice.
10. Sarada, P., & Bhushanavathi, P. (2015). Cl/(CO3 2-+ HCO3-) ratio to evaluate salt water intrusion: A case study of gnanapuram area of visakhapatnam, ap, India.
11. WHO. (2011). Nitrate and nitrite in drinking water. 1211 Geneva 27: WHO Press