Assessment of Nitrate Occurrence in the Shallow Groundwater of Merimandroso Area, Analamanga Region, Madagascar Using Multivariate Analysis

Mamiseheno Rasolofonirina[1,*], Voahirana Ramaroson[1,2], Solofonirina Dieudonné Ravelomanantsoa[1]

[1]Department of Physics, Faculty of Sciences, University of Antananarivo, Antananarivo, Madagascar
[2]Department of Isotope Hydrology, INSTN-Madagascar, Antananarivo, Madagascar
*Corresponding author: mrasolofonirina@yahoo.com

Abstract Knowing the groundwater quality is important for the drinking water supply in the highland area of Madagascar, including Merimandroso Commune insofar that groundwater is the main source of drinking water for a large number of Malagasy people. In this way, this study assessed the shallow groundwater quality with special focus on nitrate occurrence using multivariate statistical techniques such as cluster analysis (CA) and principal component analysis (PCA). That was to determine the similarities among the water samples in terms of hydrochemical features and to identify the different mechanisms involved in the shallow groundwater hydrochemistry. The study was conducted on twenty-one water samples collected from dug wells. Cluster analysis grouped the water-sampling points into two main clusters: a highly nitrate polluted group (concentration greater than 50 mg/l) and a non-nitrate polluted group. The results showed a spatial variation of the groundwater chemistry processes, while no such variability was found temporally for water samples collected at different periods. Principal component analysis extracted three principal components accounted for over 82% of the total variance. It attributed the hydrochemical features of the water samples of high nitrate content to the nitrate pollution mechanisms along with the weathering of feldspar and ferromagnesian minerals. For some of the latter water samples, the water chemistry is likely affected by igneous rock weathering. This study confirmed both the usefulness and powerfulness of multivariate statistical techniques in water quality assessment, since they helped get a proper understanding of processes controlling the shallow groundwater chemistry.

Keywords: assessment, groundwater quality, nitrate, cluster analysis, principal component analysis

Cite This Article: Mamiseheno Rasolofonirina, Voahirana Ramaroson, and Solofonirina Dieudonné Ravelomanantsoa, "Assessment of Nitrate Occurrence in the Shallow Groundwater of Merimandroso Area, Analamanga Region, Madagascar Using Multivariate Analysis." *American Journal of Water Resources*, vol. 6, no. 1 (2018): 39-47. doi: 10.12691/ajwr-6-1-5.

1. Introduction

Access and use of safe drinking water constitute the cornerstone for a healthy life, as waterborne diseases still hit many parts of the world, especially the developing countries, resulting in deaths, mostly among children. Diarrheal disease causes about 1.5 million deaths per year in the world [1].

Though nitrate is considered relatively non-toxic, exposure to elevated nitrate levels in drinking water can have adverse health effects on people, particularly on infants of age less than three months [2]. The consumption of water containing high level of nitrate can cause methemoglobinemia in infants, the so-called "blue-baby" syndrome. Methemoglobinemia is a potentially fatal condition, occurring when normal hemoglobin in red blood cells is oxidized to methemoglobin, which is unable to transport oxygen to the tissues, resulting in oxygen deprivation [3]. This may lead to unusual blue-gray skin color (cyanosis) and in some cases, to asphyxiation and deaths, when methemoglobin level goes beyond fifty percent [4].

Natural nitrate concentrations are usually less than 2 mg/l in groundwater [5]. However, nitrate is of the primary interest nutrient in many studies due to the prevalence of its contamination in groundwater. The occurrence of high nitrate level in groundwater derives from fertilizer use, land application of manure and organic wastes, leachate from sanitation systems, atmospheric deposition from fossil-fuel combustion, and disposal of urban and domestic sewage [6]. Nitrate is part of nitrogen cycle. It is the most common form of nitrogen found in water, as it is highly soluble in there. Nitrate is a chemically unreactive element in natural groundwater conditions. As a result, when nitrate is introduced into the groundwater, it will be contaminated for years or decades. It will take time before nitrate is naturally attenuated through bacterial processes (denitrification) or dilution.

Nitrate contamination of groundwater is of great concern throughout the whole world, particularly in areas where groundwater is the major source of drinking water. In the United States, about twenty-two percent of the wells in the agricultural areas have nitrate concentrations above the maximum contaminant level of 45 mg/l. High nitrate

levels are found in groundwater in the Southern African sub-continent such as South Africa, Namibia and Botswana [7]. In South Africa, elevated nitrate levels are recorded in groundwater of two areas in the Springbok Flats. In Botswana, a study carried out in Ramotswa areas showed that about thirty-six percent of sampled boreholes are of nitrate concentrations exceeding the guideline value of 50 mg/l for drinking water [2,8].

Even though nitrate is a common groundwater contaminant, Smedley [9] reported that few data on nitrate occurrence in groundwater are available in Madagascar. Nevertheless, study conducted in Mahitsy city located in the region of Analamanga indicates that thirteen out of fifteen sampled wells have excessive nitrate concentrations beyond the WHO (World Health Organization) recommended value [10].

Recently, multivariate statistical techniques such as principal component analysis (PCA) and cluster analysis (CA) have been utilized in water quality assessment. The use of PCA and CA enables to evaluate the water quality, including the assessment of water pollution sources [11,12,13]. When assessing the groundwater quality in Terengganu area located in Malaysia, Usman et al. [14] employed multivariate statistical techniques to determine the spatial variability of groundwater quality and to identify the sources of groundwater contaminant. Adebola et al. [15] performed multivariate analysis to form clusters based on similarities of the water characteristics of Ogun River and to extract factors controlling the water quality variability over the years.

A quality assessment of groundwater is of critical interest to the rural population in Merimandroso, which rely on groundwater wells as main sources of drinking water. In this sense, the aims of this study were to assess the physico-chemical quality of the shallow groundwater wells with special focus on nitrate and to describe the variability of the nitrate concentration in the groundwater wells within the study area. CA and PCA were used to explore the water quality datasets in order to assess the similarities or dissimilarities between the sampling points based on physico-chemical parameters, as well as to identify factors controlling the variability of water sample characteristics.

2. Study Area

The study was conducted in the Fokontany of Merimandroso, Alatsinainy Merimandroso, Belanitra and Ambodivona within the Commune of Merimandroso, located in the district of Ambohindratrimo within the Analamanga region, at a distance of around 30 km northwards from the Capital city, Antananarivo. The study area is situated in the northern portion of the central highlands plateau of Madagascar. It extends between longitudes 47° 29'E and 47° 31'E, and latitudes 18° 44'S and 18° 46'S. The area elevations vary from 1120 m to 1300 m above mean sea level.

The high plateau zone experiences a subtropical highland climate with mild and dry winters, and warm and rainy summers. Total annual precipitation averages between 1200 mm and 1400 mm. The highlands plateau area receives practically its annual rainfall between November and April. The average annual temperature ranges from 16°C to 24°C. The lowest temperature is recorded between June and August [16].

The study area lies on the Precambrian crystalline basement rocks [17]. It is composed of late Archaean, and Neoproterozoic gneisses, and granitoids [18]. The basement rocks are covered by a layer of weathered material, constituted typically by 10–40 m thick of red lateritic soil [19,20].

In weathered zones, groundwater is produced in both weathered laterite and a semi-confined aquifer located in fissure granitic basement, which are separated by clay layer [21].

3. Materials and Methods

3.1. Sampling

Field sampling for this study was conducted in March 2005 (summer season) and in July 2005 (winter season). Figure 1 displays the sampling site locations.
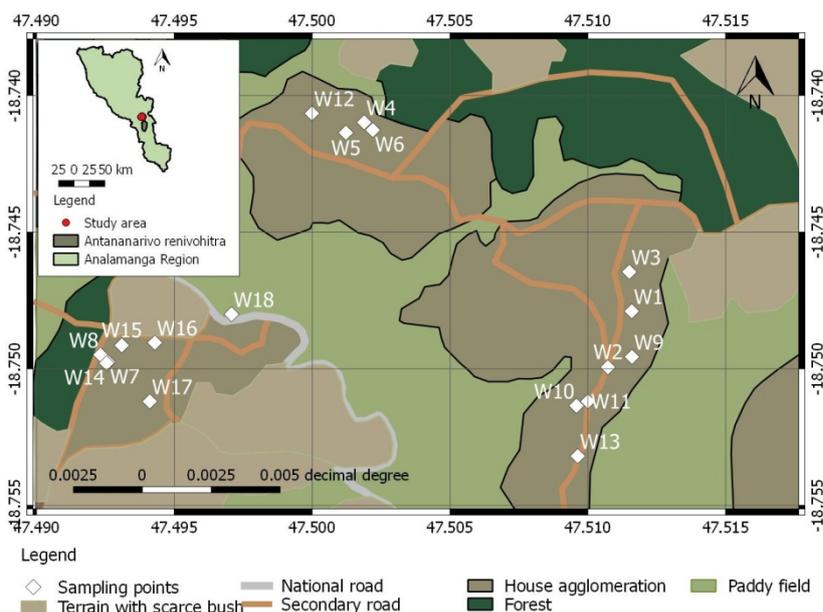
Figure 1. Location of the water-sampling sites

Eight (8) wells were sampled during the first campaign (W1 to W8), whereas thirteen (13) samples (W9 to W18) were collected during the second campaign. Among those collected within the first campaign, three (3) wells (W4, W6 and W7) were resampled in scope of the second sampling campaign.

Water samples were collected from private dug wells using clean bucket attached to a rope. Sample containers were rinsed at least three times with water from the wells before sampling. Most of the sampled wells are indoors installations. The wells have a large diameter of about 1 m, as they are excavated by hand shovel. Although, they do not have casing pipe as for driven or drilled wells, they are lined with brick or stones to prevent collapse. The groundwater table ranged from 4.7 m to 17.7 m and from 3.9 m to 20.4 m under the surface land for the first and second sampling campaigns, respectively.

All samples collected were filtered with a 0.2 μm cellulose nitrate membrane in the field. Filtered water samples were collected in two separate pre-rinsed (100 ml) polyethylene (PE) bottles for anion and cation measurements for each of sampled wells. Cation samples were preserved with HNO_3 after filtration.

3.2. Field Measurement

Field water quality parameters such as pH, Electrical Conductivity (EC), Dissolved Oxygen (DO), Temperature (T) and Total Dissolved Solids (TDS) were determined using portable multimeter (Multi 340i model), whereas the measurement of Total Dissolved Solids was performed using Hach conductimeter (SensION 5). Alkalinity in term of bicarbonate concentration was measured by acidic titration (H_2SO_4) using digital titrator. The field parameters were determined within few minutes after collecting samples for *in situ* measurements. They were measured using aliquots of sampled waters. The groundwater level measurement was performed using water level meter before purging wells.

3.3. Anion and Cation Measurement

Concentrations of ions such as ammonium (NH^{4+}), sodium (Na^+), potassium (K^+), magnesium (Mg^{2+}), calcium (Ca^{2+}), chloride (Cl^-), bromide (Br^-), nitrate (NO_3^-), and sulfate (SO_4^{2-}) were determined by Dionex Ion Chromatograph system DX-120. The Chromatograph systems are equipped with AS 14A and CS 12A as analytical columns for anion and cation measurements, respectively. 20 mmole/L of CH_3SO_3H and a combination of 8 mmole/L of Na_2CO_3 with 1 mmole/L of $NaHCO_3$ were used as eluent for cationic and anionic measurements, respectively. Eluents were pumped through the columns at 1 mL/minute. All water samples were measured in duplicate. The precision of analytical results was typically less than 5%.

3.4. Multivariate Statistical Method

3.4.1. Cluster Analysis

Cluster analysis is a numerical technique of classification, which consists in grouping into the same class called cluster a set of objects (observation or samples) of similar features [22,23]. Data clustering groups objects possessing similar characteristics used in the analysis, whilst it separates into different clusters those of dissimilar characteristics. In this study, Ward linkage method was used along with the squared Euclidean distance as a measure of similarity among the sampling points. The calculation of distance between the sampling points (water samples) refers to the values of the physico-chemical parameters. This study used hierarchical agglomerative method to form cluster, since it is the most widespread clustering technique in the area of water quality assessment [24,25,26]. The analytical data were log-transformed and normalized using z-score method prior to performing clustering analysis to balance the influence of variable variance [27].

3.4.1. Principal Component Analysis

Principal component analysis is a dimension-reduction method that is used to convert a large number of related original variables into a smaller number of uncorrelated variables called principal components, which are linear combinations of the original variables [28]. The aim is to reduce the number of variables of interest whilst retaining as much of the original information as possible. This study performed KMO (Kaiser-Meyer-Olkin) & Bartlett's Test of Sphericity and the analysis of correlation matrix to examine the adequacy of factorization prior to extracting the constructs. Bartlett's Test is esteemed as statistically significant at level of $p < 0.05$. As for determining the number of principal components of interest, this study considered all components with eigenvalue greater than 1 to ensure that the extracted component explains large amounts of variance than an individual variable [29]. The principal components retained account for at least 80% of the variation to get an adequate representation of the data. In addition, the Cattell's scree plot method was displayed to check the validity of the principal components to be considered in the analysis. The PCA was performed using the correlation matrix, as it treats all variables in an equal weight. Retained principal components were subjected to varimax rotation with Kaiser Normalization to make the rotated principal components as simple as possible to interpret.

This study used the Statistical Package for the Social Sciences (SPSS) version 19 for data processing and analysis.

4. Results and Discussion

4.1. Shallow Groundwater Quality

Measured shallow groundwater temperatures varied between 20.8°C and 22.1°C, averaging 21.4°C. They show the seasonal fluctuation of the ambient air temperature. Sampled water temperatures averaged 21.5°C and 21.3°C for the first and second sampling campaigns, respectively. Groundwater is warm in March, whereas it is a little bit cooler in July.

pH values ranged from a minimum of 4.6 up to 6.4 for the two sampling campaigns (Table 1). These values indicated that sampled waters were acidic, as pH values were less than 7. Low pH values result likely from the acidifying influences of natural soil and atmospheric processes, since the sampled wells are at a shallow depth.

They were out of the range given by the WHO standards (within 6.5 to 8.5). Although pH is considered as an aesthetic parameter in drinking water, it plays important roles in the chemical reactions occurring within the natural water body.

Measured Total Dissolved Solids (TDS) values varied in the range of 3.1 - 438 mg/l for both sampling campaigns (Table 1). They averaged 150 mg/L and 106 mg/L for the first and second sampling campaigns, respectively. All sampled waters had TDS concentrations less than 600 mg/l, which is considered as a recommended value for the palatability of drinking water [2]. The maximum TDS value was recorded in W1 well water collected during the first campaign, while the minimum value was measured in W15 well water.

Concentrations of Dissolved Oxygen (DO) were of range between 2.3 and 6.0 mg/l for all collected samples (Table 1). The mean values during the first and second sampling campaign were 3.8 and 4.5 mg/l, respectively. All sampled wells were of aerobic redox conditions, since the DO concentrations recorded were greater than 1 mg/l [30].

Waters sampled within this study had small quantities of bicarbonates (less than 41 mg/l). Sample waters collected from the first and second campaigns presented mean values of 15.1 and 4 mg/l as bicarbonates, respectively. The maximum value of bicarbonate concentration (40.9 mg/l) was recorded in well water W6.

Sodium concentrations ranged between 0.9 and 52.7 mg/l with the average values of 17.6 mg/l and 10.3 mg/l for the first and second sampling campaigns, respectively. Waters sampled contained potassium of concentration varying from 0.3 to 133 mg/l. The concentration mean was of 23.5 mg/l for the first sampling campaign, whereas it was 17.6 mg/l for the second campaign.

Chloride contents of sampled waters fluctuated between 0.3 and 78 mg/l, averaging 18.3 mg/l for both the first and the second sampling campaigns. Nitrate concentrations of sampled waters showed significant variations, ranging from 0.7 - 352 mg/l. Nitrate contents averaged 102 mg/l and 60 mg/l for the first and second sampling campaigns, respectively. Nitrate was present in higher concentrations exceeding the maximum allowed values recommended by WHO (50 mg/l as NO_3^-) for drinking water in the 38% and 46% of cases for well waters sampled during the first and second campaigns, respectively. Large amounts of nitrate in well waters could result from latrine or livestock waste infiltration into the soil. The sampled wells were found within the range of 2 – 20 m from the pit latrines, poultry/livestock breeding location or organic manure storage in most cases.

Table 2 provides the Pearson correlation coefficients between the water chemistry parameters. Table 2 indicates that there was a very strong and significant correlation at $p < 0.01$ level between TDS and nitrate (r=0.954), chloride (r=0.906), sodium (r=0.936), magnesium (r=0.814) and calcium (r=0.758). pH was positively correlated to bicarbonate at $p < 0.05$ level (r=0.459), revealing to some extent the pH control by reactions involving in the carbonate system. Sodium had positive and strong correlations with magnesium (r=0.859), calcium (r=0.738), chloride (r=0.760) and potassium (r=0.643) at $p < 0.01$ level. As for the nitrate, strong positive correlations were observed between nitrate and sodium (r=0.964), magnesium (r=0.922), calcium (r=0.833) and chloride (r=0.789) at $p < 0.01$ level.

Table 1. Physico-chemical characteristics of sampled waters

Water Point	Na⁺ [mg/l]	K⁺ [mg/l]	Mg²⁺ [mg/l]	Ca²⁺ [mg/l]	Cl⁻ [mg/l]	NO₃⁻ [mg/l]	SO₄²⁻ [mg/l]	HCO₃⁻ [mg/l]	pH	TDS [mg/l]	Temp [°C]	DO [mg/l]
W1	48.2	133	9.7	11.0	63.7	240	28.5	7.7	4.73	438	21	3.07
W2	4.3	4.3	2.8	1.6	2.6	25.4	1.5	8.5	5.04	25.3	21.5	2.88
W3	1.6	2.1	0.3	1.7	0.9	3.6	0.4	9.8	5.94	13	21.7	4.15
W4	4.3	3.9	4.0	6.7	10.6	21.5	0.9	19.6	5.81	57.8	21.6	3.84
W5	17.2	3.3	11.6	24.7	19.4	134	2.1	15.9	6.13	215	21.1	4.44
W6	3.8	10.9	2.9	5.1	1.3	8.1	0.8	40.9	5.54	9.5	21.8	3.95
W7	8.6	4.0	1.1	0.6	0.5	30.2	0.4	7.3	4.63	44.3	22.1	4.13
W8	52.7	26.4	23.5	32.3	44.7	352	0.3	11.0	4.82	399	21.5	3.94
W9	0.9	0.7	1.3	1.0	0.3	0.7	0.9	9.8	5.71	17.3	21	4.33
W10	17.1	41.6	3.9	7.6	37.2	91.3	1.1	1.2	4.66	170	21	4.32
W11	8.0	16.4	3.0	5.2	16.5	52.3	1.3	1.2	5	87.6	21	4.31
W12	24.3	13.1	7.7	10.1	26.0	119	0.8	1.2	5.43	157	21.6	3.8
W13	14.7	26.3	3.4	8.9	39.4	57.6	8.8	0.0	5.19	138	21	2.31
W14	8.9	9.6	1.9	1.7	1.4	39.0	0.7	1.2	5.32	19.9	21.4	5.06
W15	2.0	1.3	0.3	0.3	0.3	7.5	0.3	2.4	5.24	3.1	21.6	5.12
W16	1.1	0.7	0.3	0.5	0.3	0.9	0.3	5.5	4.97	3.2	21.2	2.42
W17	20.3	106	6.6	12.7	78.0	185	2.3	0.0	5.23	331	20.8	4.18
W18	3.6	2.4	1.0	1.9	4.9	15.2	0.3	1.2	5.4	4.6	21.3	5.16
W4'	4.6	3.8	4.0	6.5	11.1	34.6	0.9	4.9	6.07	58.5	21.3	6.03
W5'	19.9	3.8	11.8	25.1	20.9	137	2.6	23.2	6.41	185	21.4	5.89
W7'	8.4	3.6	1.2	0.8	1.2	36.0	0.7	0.0	4.55	40	21.7	5.11

Table 2. Correlation matrix of water chemistry

	pH	TDS	DO	Na$^+$	K$^+$	Mg^{2+}	Ca^{2+}	Cl$^-$	NO$_3^-$	SO$_4^{2-}$	HCO$_3^-$
pH	1	-.131	.417	-.272	-.346	.073	.261	-.206	-.174	-.220	.459*
TDS		1	-.167	.936**	.779**	.814**	.758**	.906**	.954**	.589**	-.045
DO			1	-.176	-.279	.005	.087	-.242	-.086	-.359	.023
Na$^+$				1	.643**	.859**	.738**	.760**	.964**	.568**	-.034
K$^+$					1	.302	.229	.874**	.611**	.759**	-.175
Mg^{2+}						1	.950**	.573**	.922**	.209	.216
Ca^{2+}							1	.556**	.833**	.135	.277
Cl$^-$								1	.789**	.552**	-.198
NO$_3^-$									1	.416	-.013
SO$_4^{2-}$										1	-.042
HCO$_3^-$											1

*Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

4.2. Data Clustering

This study used the hydrochemical parameters such as Na$^+$, K$^+$, Mg^{2+}, Ca^{2+}, Cl$^-$, SO$_4^{2-}$, HCO$_3^-$ and NO$_3^-$ for the hierarchical agglomerative cluster analysis [31]. Data clustering categorized the water sampling points into two (2) main cluster groups (A and B) according to the dendrogram displayed in Figure 2. The first cluster A was made up of nine (9) water-sampling sites with water of mineralization ranging from 87 mg/l to 438 mg/l in terms of TDS. The water samples within the cluster A were of moderate (50 – 150 mg/l) to high (over 150 mg/l) levels of nitrate. The group cluster B was composed of twelve (12) water-sampling points whose water had a very low mineralization (TDS < 60 mg/l). Figure 2 depicts that the cluster A was split into two (2) sub-clusters (1 and 2). The sub-cluster 1 consisted of three (3) water-sampling points (W5, W5' and W8), whereas the sub-cluster 2 was formed by six (6) water-sampling stations such as W1, W10, W11, W12, W13 and W17. The cluster B contained three (3) sub-clusters, in this case sub-clusters 3, 4 and 5. The sub-cluster 3 was composed of four (4) water-sampling points (W3, W9, W15 and W16), whilst the sub-clusters 4 also included four (4) water-sampling sites (W2, W4, W4' and W6). In addition, the sub-cluster 5 comprised four (4) water-sampling points namely W7, W7', W14 and W18.

In sub-cluster 1, calcium/magnesium or sodium/magnesium were considered as the dominant cations in relative amounts more than 30%, while chloride was found to be the dominant anion. The dominance of chloride was accompanied by an increased content of nitrate (134 – 352 mg/l) with a positive strong correlation (r=0.999), suggesting that chloride and nitrate could be from a common source. As for the sub-cluster 2, alkali metals (Na$^+$ and K$^+$) were the dominant cations, having the relative amounts in the range of 27 - 57%. The dominant anion was the chloride with the relative amount greater than 70%. All water samples in the sub-cluster 2 contained high nitrate concentration exceeding the WHO maximum recommended value (50 mg/l as NO$_3^-$). The water samples within the sub-cluster 3 were characterized by the anion dominance of bicarbonate (more than 75% as relative amount) as well as the cation dominance of the mixtures of calcium/sodium or magnesium/calcium, except for the water sample from W15. This last was a sodium-bicarbonate water type (Table 3). The nitrate concentration of the water samples in the sub-cluster 3 was the lowest among the sampled water points (less than 8 mg/l).
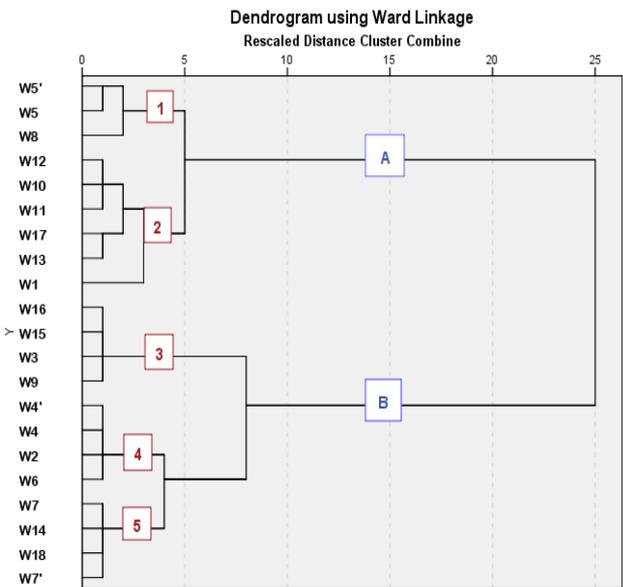
Figure 2. Water samples dendrogram

Table 3. Hydrochemical facies of the water samples

Hydrochemical facies	Water sampling point
Ca – Mg – HCO$_3$ (mixed)	W4
Ca – Na – HCO$_3$ (mixed)	W3
K – Ca – HCO$_3$ (mixed)	W6
Mg – Ca – HCO$_3$ (mixed)	W9
Mg – Na – HCO$_3$ (mixed)	W2
Na – Ca – HCO$_3$ (mixed)	W16
Na – HCO$_3$	W7, W15
Ca – Mg – Cl (mixed)	W4', W5, W5'
K - Cl	W17
K – Na – Cl (mixed)	W1, W10, W11, W13
Na – Ca – Cl (mixed)	W18
Na – Cl	W7', W14
Na – Mg – Cl (mixed)	W8, W12

For the sub-cluster 4, the water samples had bicarbonate as the dominant anion, excepting that of W4', which was of chloride dominance. In addition, the dominant cations were the mixtures of the alkali metals (Na^+ and K^+) and earth (Ca^{2+} and Mg^{2+}), depending on their relative amounts in the water samples. The nitrate content varied in the range of 8.1 – 34.6 mg/l for the water samples in the sub-cluster 4. As far as the sub-cluster 5 was concerned, chloride was found as the dominant anion in the relative amount higher than 53% for all samples in this sub-cluster, excluding the water sample from W7, which was a bicarbonate water type. The water samples in the sub-cluster 5 had sodium as the dominant cation with the presence of calcium in the relative amount of 25% for the water sampled fromW18.

The cluster analysis indicated that the water-sampling points located at the same Fokontany belonged to more than one sub-cluster, suggesting that the nature of their waters appears to originate from more than one hydrochemical features. This revealed that there were various mechanisms involved in the mineralization of the shallow groundwater in the study area.

As for the temporal variability of the hydrochemical data of the sampled waters, the cluster analysis showed that the water samples collected from the same water points (W4/W4', W5/W5' and W7/W7') during the first and second campaigns belonged to the same sub-clusters. This fact implied that the waters from re-sampled water points displayed similar hydrochemical features throughout both sampling campaigns, even if the hydrochemical data registered to some extent a slight increase. That was the case for the nitrate concentration, showing a slight upward variation from the range of 21.5 – 134 mg/l to 34.6 – 137 mg/l. It was also accompanied by the increased content of chloride, suggesting that the increase of nitrate and chloride concentrations seemed to share somehow the same origin.

4.3. Principal Component Extraction

PCA was performed to determine the major water parameters influencing the water features in the water quality assessment. In this study, twelve (12) variables were used for the principal component analysis. It was dissolved oxygen (DO), pH, total dissolved solid (TDS), sodium, potassium, calcium, magnesium, chloride, sulfate, nitrate, bicarbonate and silica. Eigenvalues of 1 or greater were considered as significant for measuring the amount of variance explained by the principal component. The analysis of the correlation matrix (Table 2) showed that many variables were correlated since their correlation coefficients were greater than 0.5. In addition, Bartlett's test of sphericity had an associated p value less than < 0.001 (Table 4). These results indicated that there were appropriate correlations to conduct the factor analysis, even if the KMO value (0.488) was miserable [32].

Table 4. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.488
Bartlett's Test of Sphericity	Approx. Chi-Square	379.580
	Degree of freedom	66
	Sig.	.001

Three major Principal Components (PCs) were extracted, explaining over 82% of the total variance in the original data structure. To confirm the number of components to keep, this study used the Cattell's scree plot method. Figure 3 displayed that the eigenvalues begun to level off after the third principal component. These three first components were retained, as they explained for themselves the adequate amount of variance in the dataset. The remaining components accounted for a small proportion of the variability (less than 18%) of water quality data.
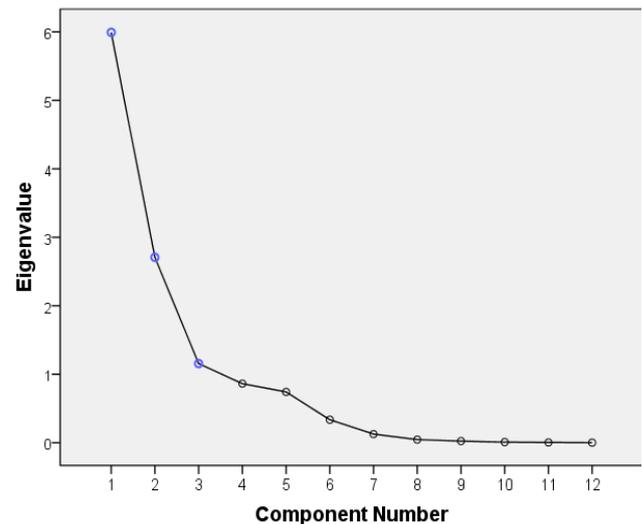
Figure 3. Scree plot of water quality data

Table 5. Component loadings and eigenvalues for the extracted principal components after Varimax rotation

Variables	PC1	PC2	PC3
Nitrate	0.980	-0.143	-0.034
Magnesium	0.928	-0.067	0.247
TDS	0.928	0.357	-0.017
Sodium	0.923	0.270	-0.100
Calcium	0.878	-0.122	0.413
Chloride	0.769	0.475	-0.096
Silica	0.270	-0.014	0.819
pH	-0.076	-0.307	0.824
Bicarbonate	-0.037	0.038	0.747
Sulfate	0.331	0.784	-0.079
DO	0.076	-0.770	0.017
Potassium	0.541	0.685	-0.252
Eigenvalue	5.386	2.238	2.230
% of Variance	44.879	18.652	18.584
Cumulative %	44.879	63.531	82.115

As provided in Table 5, PC1 accounted for about 44.88% of the total variance after Varimax rotation. It showed a strong positive loading on nitrate, magnesium, TDS, sodium, calcium and chloride. This first component seems to identify the nitrate pollution and the feldspar and ferromagnesian weathering processes of the shallow groundwater. The nitrate was the most important variable to the PC1, as it had the highest component loading

(0.980). In addition, the increase of nitrate content was accompanied by that of magnesium, TDS, sodium, calcium and chloride. For the second component (PC2), it accounted for 18.65% of the total variance. PC2 had strong positive loadings on sulfate and potassium, whereas it displayed a strong negative loading on DO. This indicated that the higher the sulfate and potassium concentrations, the lower the DO concentration. PC2 related to the tendency to reducing conditions in the shallow groundwater and to the igneous rock weathering. As for the third component (PC3), it explained about 18.58% of the total variability in the dataset. PC3 was made up of silica, pH and bicarbonate with component loadings greater than 0.7. It was viewed as related to silicate weathering process.

As shown in Figure 4, the first component (PC1) separated water samples from W1, W5, W5', W8, W10, W12 and W17 from W2, W3, W4, W4', W6, W7, W7', W9, W11, W13, W14, W15, W16 and W18. The second component (PC2) differentiated the samples from W1, W2, W3, W4, W10, W13, W16 and W17 to those from W4', W5, W5', W7, W7', W8, W9, W11, W12, W14, W15 and W18. Figure 5 displayed that the third component (PC3) separated water samples from W2, W3, W4, W4', W5, W5', W6 and W13 from W1, W7, W7', W8, W9, W10, W11, W12, W14, W15, W16, W17 and W18.

PC1 contrasted all samples having nitrate concentration greater than 50 mg/l, excepting those from W11 and W13 with the water samples of low or moderately low level of nitrate. These two last samples contained nitrate of concentration less than the average value for all samples (75.8 mg/l as nitrate).
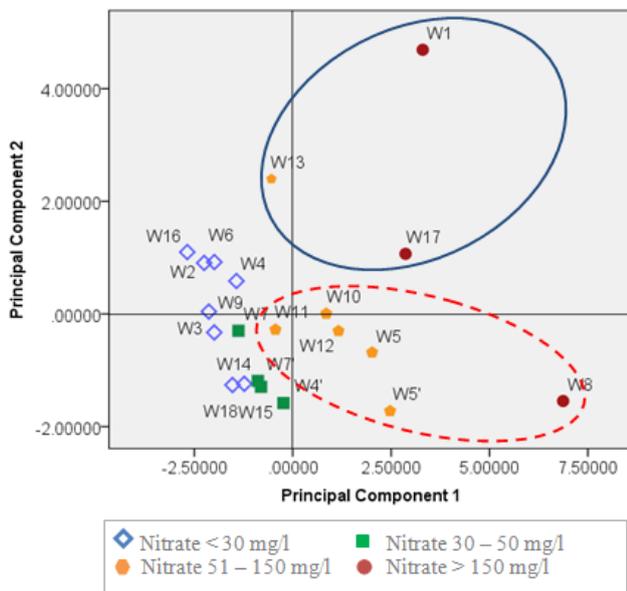
Figure 4. Component scores of the water samples plotted on the PC1 and PC2

The examination of the samples location shown in Figure 4 and Figure 5 revealed that the water samples of high level of nitrate had high scores for PC1. It suggested that their hydrochemical features were firstly controlled by the nitrate pollution mechanisms, including probably the infiltration from the pit latrines, the poultry/cattle barn and/or the manure storage coupled with the chemical weathering of feldspar and ferromagnesian minerals.

Likewise, the first group of the water samples composed of W1, W13 and W17 (encircled in red in Figure 4) had high scores for PC2, whereas the second group (W5, W5', W8, W10, W11 and W12) displayed low scores within PC2. It indicated that the hydrochemical characteristics appeared also to be affected by the igneous rock weathering under the tendency to low DO, as far as the first group of the water samples with high nitrate concentration was concerned. As for the water samples in the second group, they were characterized by high DO level, which makes the shallow groundwater an oxidizing environment that is not conducive to denitrification process. The water sampled from W8 was located far away within the second group in the biplot between PC1 and PC2 (Figure 4), suggesting that its hydrochemical feature was more completely explained by the first component (PC1).
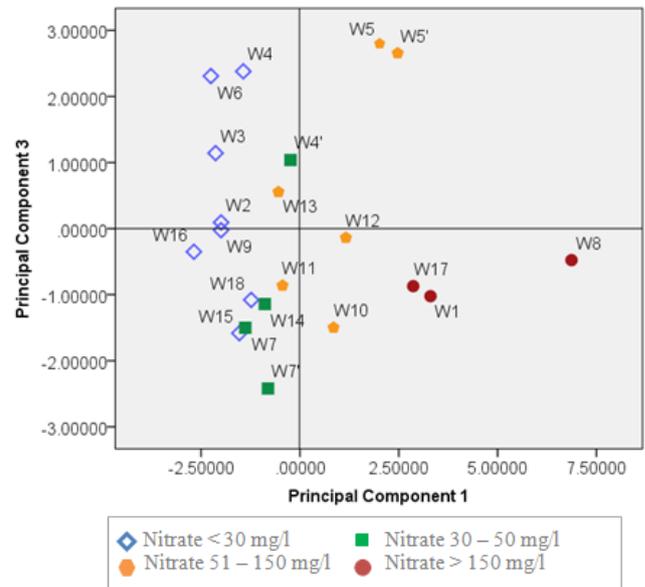
Figure 5. Component scores of the water samples plotted on the PC1 and PC3

The waters sampled from the sampling points W2, W3, W4, W4', W6 and W13 had their hydrochemical features affected by the silicate weathering, as they displayed high scores for PC3 according to Figure 5. In addition, the weathering of igneous rock appeared to occur in the shallow groundwater where were located the sampling points W2, W4, W6 and W16, since their water samples were found in the left-hand part of Figure 4.

5. Conclusion

This study assessed the quality of the shallow groundwater in the area of Merimandroso using the Cluster Analysis (CA) and the Principal Component Analysis (PCA) techniques. CA enabled to classify the water-sampling points into two main clusters (A and B) according to the nitrate content and TDS value of their water samples. The water samples belonged to the cluster A were considered as highly polluted in terms of nitrate, whereas those grouped in cluster B were found to be less polluted. Additionally, the use of CA revealed that there were numerous processes involved in the hydrochemistry of the shallow groundwater in the study area. The water samples

from the same water-sampling points did not register a substantial temporal variability in terms of hydrochemical characteristics, even though the water quality parameters displayed to some extent a slight increase.

PCA extracted three principal components accounted for over 82% of the total variance of the original water quality dataset. The first component (PC1) explained about 45% of the total variability of the dataset. It appeared to originate from the nitrate pollution and the weathering of feldspar and ferromagnesian mineral processes. The second component (PC2) and the third component (PC3) accounted for about 18.7% and 18.6% of the total variability of the water quality dataset, respectively. The extraction of these three principal components provided an evidence of multiple mechanisms involved in the shallow groundwater chemistry. The water samples of high nitrate concentration (greater than 50 mg/l) seemed to have their hydrochemical features controlled by the nitrate pollution mechanism along with the weathering process of feldspar and ferromagnesian minerals. In addition to that, their water chemistry was likely to be affected by the igneous rock weathering for some of them. The silicate weathering process came across as controlling the hydrochemistry of some water samples considered as less polluted by nitrate coupled with the feldspar and ferromagnesian weathering.

The multivariate analysis technique such as CA and PCA is both a useful and powerful method for exploring and analyzing the dataset within the water quality assessment. It enables to group a large number of observations or variables into meaningful categories that are simpler to interpret. However, this study demonstrates the need of combining the use of multivariate statistical analysis with a good knowledge of the geological formation settings of the study area while being used to evaluate the groundwater quality in order to get better understanding of the extracted component meanings. That will enhance the result accuracy when identifying the various processes involved in the hydrochemistry of the groundwater.

Acknowledgements

References

[1]   World Health Organization (WHO), Global Health Observatory Data Repository. Estimates for 2000 – 2012. Geneva. 2012.

[2]   World Health Organization (WHO), Guidelines for Drinking Water Quality (4th ed.). Geneva. 2011.

[3]   Agency for Toxic Substances and Disease Registry (ATSDR), Nitrates and Nitrites. CAS # 84145-82-4, 14797-65-0. Atlanta, GA: U.S. Department of Public Health and Human Services, Public Health Service. [Online] Available at http://www.atsdr.cdc.gov/toxfaqs/tfacts204.pdf. [Accessed Aug. 10, 2017]

[4]   Knobeloch, L., Salna, B., Hogan, A., Postle, J. and Anderson, H., "Blue babies and nitrate-contaminated well water", Environ Health Perspect, 2000, 108, 675-678.

[5]   Mueller, D.K., Hamilton, P.A., Helsel, D.R., Hitt, K.J., and Ruddy, B.C., Nutrients in ground water and surface water of the United States--an analysis of data through 1992: U.S. Geological Survey Water-Resources Investigations Report 95-4031, 74 p. 1995.

[6]   Dylan, B., Aaron, K., Giorgos, K., Katherine, L., Megan, M., Graham, E. F. and Thomas, H., Groundwater Nitrate Occurrence: Technical Report 4, 2012.

[7]   Tredoux, G. and Talama, A. S., Groundwater Pollution in Africa. Nitrate Pollution of groundwater in southern Africa, Ed. Taylor&Francis/Balkema, 15-36. 2006.

[8]   Staudt, M., Environmental Hydrogeology of Ramotswa. Report by the Environmental Geology Division, 47 p., Dept of Geological Survey (DGS), Lobatse, Botswana. 2003.

[9]   Smedley, P.L., Groundwater Quality: Madagascar. British Geological Survey.NERC 2002. [Online]. [Online]. Available at http://www.bgs.ac.uk/sadcreports/madagascar2002smedleybgsgroundwaterquality.pdf. [Accessed Oct. 5, 2016]

[10]   Rasolofonirina, M., Ramaroson, V., Raoelina Andriambololona, "On-site Sanitation Influence on Nitrate Occurrence in the Shallow Groundwater of Mahitsy City, Analamanga Region, Madagascar", Environment and Pollution, 2015, 4(2), 55-65.

[11]   Varol, M., Gökot, B., Bekleyen, A. and Şen, B., "Spatial and Temporal Variations in Surface Water Quality of the Dam Reservoirs in the Tigris River Basin, Turkey", Ca-tena, 2012, 92, 11-21.

[12]   Kazi, T.G., Arain, M.B., Jamali, M.K., Jalbani, N., Afridi, H.I., and Sarfraz, R.A., "Assessment of water quality of polluted lake using multivariate statistical techniques: A case study", Ecotox Environ Safe, 2009, 72, 301-309.

[13]   Andrade, E. M., Palácio, H. A. Q., Souza, I. H., Leão, R. A. and Guerreiro, M. J., "Land Use Effects in Groundwater Composition of an Alluvial Aquifer (Trussu River, Brazil) by Multivariate Techniques", Environmental Research, 2008, 106(2), 170-177.

[14]   Usman Nasiru Usman, Mohd Ekhwan Toriman, Hafizan Juahir, "Assessment of Groundwater Quality Using Multivariate Statistical Techniques in Terengganu", Science and Technology, 2014, 4(3), 42-49.

[15]   Adebola, A. Oketola, S. M., Adekolurejo, Oladele Osibanjo, "Water Quality Assessment of River Ogun Using Multivariate Statistical Techniques", Journal of Environmental Protection, 2013, 4, 466-479.

[16]   UN. Madagascar. In: Groundwater in Eastern,Central and Southern Africa, Natural Resources/Water Series No. 19, United Nations, New York, 1989, 133-149.

[17]   Besairie, H., "Précis de géologie malgache", Annales géologiques de Madagascar, 1973, 36, 109-134.

[18]   Goncalves, P., Nicollet, C., and Lardeaux, J. M, "Finite strain pattern in Andriamena unit (North Central Madagascar): evidence for late Neoproterozoic-Cambrian thrusting during continental convergence", Precambrian Research, 2003, 123, 135-157.

[19]   Jeffrey, D., Hydrogeological mapping of north central of Madagascar using limited data, 2009. [Online]. [Online]. Available at http://nora.nerc.ac.uk/id/eprint/9062. [Accessed Sept. 25, 2015].

[20]   Grillot, J-C and Ferry, L., "Approche des échanges surface-souterrain en milieu cristallin altéré aquifère", Hydrol. Continent, 1990, 5(1), 3-12.

[21]   Grillot, J. C., Blavoux, B., Rakotondrainibe, J. H., Raunet, M., and Randrianarisoa, N., "A propos des aquifères d'altérites sur les hauts plateaux cristallophylliens de Madagascar", Comptes Rendus de L'Académie des Sciences Paris Hydrogéologie, 1987, 305(Série II), 1471-1476.

[22]   Brian, S. Everitt, Landau, S., Leese, M. and Stahl, D., Cluster Analysis, 5th Edition. John Wiley & Sons, Ltd. 2011

[23]   Shrestha, S. and Kazama, F., "Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji River basin, Japan", Environ. Model. Softw, 2007, 22, 464-475.

[24]   Wang, Y., Wang, P., Bai, Y., Tian, Z., Li, J., Shao, X., Mustavich, L.F. and Li, B.-L., "Assessment of surface water quality via multivariate statistical techniques: A case study of the Songhua River Harbin region, China", J. Hydro-Environ. Res., 2013, 7, 30-40.

[25] Zhaoa, Y., Xia, X.H., Yang, Z.F. and Wang, F., "Assessment of water quality in Baiyangdian Lake using multivariate statistical techniques", Procedia Environmental Sciences, 2012, 13, 1213-1226.

[26] Huang, F., Wang, X., Lou, L., Zhou, Z. and Wu, J., "Spatial variation and source apportionment of water pollution in Qiantang River (China) using statistical techniques", Water Res., 2010, 44, 1562-1572.

[27] Ryberg, K.R., Cluster Analysis of Water-Quality Data for Lake Sakakawea, Audubon Lake, and McClusky Canal, Central North Dakota, 1990-2003. U.S. Geological Survey Scientific Investigations Report 2006-5202, 38 p. 2006.

[28] Abdi, H. and Williams, Lynne J., Principal component analysis. John Wiley & Sons, Inc. Vol.2. 2010.

[29] Kaiser, H.F., "The application of electronic computers to factor analysis", Educational and Psychological Measurement, 1960, 20, 141-151.

[30] Dinicola, R.S., Continued biodegradation of chloromethane compounds in ground water at Operable Unit 1, Naval Undersea Warfare Center, Division Keyport, Washington. U.S. Geological Survey Scientific Investigations Report 2006-5056, 42 p. 2006.

[31] Rasolofonirina, M., Ramaroson, V. and Ravelomanantsoa, S.D., "Use of Cluster Analysis Technique for Water Quality Assessment in Mahitsy Commune, Central Highland of Madagascar", Int J Recent Sci Res. 2018, 9(1). 23191-23196.

[32] Kaiser, H F and Rice, J., "Little Jiffy, Mark IV," Educational and Psychological Measurement, 1974, 34(1), 111-117.