Relationship Between Aquifer Processes and Groundwater Quality: A Case of Olbanita Aquifer System, Lower Baringo Basin, Kenya Rift

Authors: Sosi, Benjamin, Getabu, Albert, Maobe, Samson, and Barongo, Justus

Source: Air, Soil and Water Research, 12(1)
Published By: SAGE Publishing
URL: https://doi.org/10.1177/1178622119889910
Relationship Between Aquifer Processes and Groundwater Quality: A Case of Olbanita Aquifer System, Lower Baringo Basin, Kenya Rift

Benjamin Sosi1,2, Albert Getabu2, Samson Maobe2 and Justus Barongo3

1Department of Natural Resources, Egerton University, Nakuru, Kenya. 2Department of Natural Resources, Kisii University, Kisii, Kenya. 3Department of Geology, University of Nairobi, Nairobi, Kenya.

ABSTRACT: A hydrogeochemical relation has been hypothesized through the analyses of physiochemical data of a fractured volcanic rock aquifer located in the Lower Baringo Basin, Kenyan Rift. Data sets included 15 individual metrics determined in 42 dry and wet season water samples obtained from 6 boreholes in the area. Aquifer evolutionary theory was postulated using sequential principal component analysis (PCA) and hierarchical cluster analysis. To eliminate the effects of scale dimensionality, PCA decomposed the variable data into 4 factors, namely, electrical conductivity, salinity, alkalinity, and carbonate equilibrium with external pH control for the dry season and salinity, carbonate equilibrium with external pH control, alkalinity, and electrical conductivity for the wet season. The main result depicted a major shift in the variability factor from electrolytic conductivity (34.8%) in the dry season to salinity (23.5%) in the wet season. Ward's linkage cluster analysis partitioned the aquifer into 2 spatially discrete associations; the western and the eastern entities, respectively, in spite of their shared recharge area. These agglomerative scheduling validated in an integrative approach (with groundwater flow predictions using a calibrated petrophysical groundwater model for the area) linked the 4 factors to aquifer processes and 3 pathways: fault permeability, weathering processes, and water-rock interaction. Statistical approaches are, therefore, useful in the conceptualization of pollutant sources and their attenuation for effective groundwater quality management.

KEYWORDS: Groundwater evolutionary processes, water quality characterization, Olbanita, Kenya Rift

Introduction

Groundwater quality deterioration is a fundamental concern world over. Its quality characteristics in sub-Saharan Africa exhibit a wide spatial and temporal variability.1,2 The high volcanic fluoride (>1.5 mg/L) according to World Health Organization3 depicts widespread variability in groundwater aquifers of the African Rift.4,5 Continued urbanization is also rapidly causing increasingly poorer quality groundwater.1,2 However, the resource is strategically most reliable to exterminate the emergent demand.5 Improved hydrochemical characterization of aquifers in such highly heterogeneous areas, decoupling the relation between groundwater quality and aquifer lithology, could enhance their effective management.6–8 Pazand9 observed that sustainable development and management of the quality of groundwater resources in arid and semi-arid lands is achievable through improved understanding of geochemical evolution and groundwater processes. The determination of groundwater evolutionary processes concerning groundwater quality in fractured confined aquifers is complex. Large field experimental and geochemical laboratory data sets require proper management, elaboration, and interpretation tools and techniques.6 Güler et al10 noted that the interpretation of associations exclusively based on graphical techniques is fraught with difficulties owing to the simultaneous nature of aquifer processes and their effects. Recent studies by Nwankwoala,11 Pazand,9 Moghimi,12 and Rotiroti et al13 established the value of multivariate techniques such as principal component analysis (PCA) and hierarchical clustering analysis (HCA) in the provision of quantitative measures of correlation between water quality parameters and fundamental aquifer processes. PCA is widely utilized to characterize groundwater pollution sources,9 while HCA is used to deduce spatial variability among sampling sites.12 An integrative application of PCA and HCA was employed by Rotiroti et al13 and Moghimi11 to interpret processes affecting groundwater hydrochemistry and by Yidana et al,7 to evaluate evolutionary trends of groundwater dynamics and by Majeed et al14 to assess spatial patterns of pollutants in water. Several investigators15,16 utilized z-score or log conversions to normalize physical and chemical data acquired from groundwater sources. However, Everitt et al17 noted that such transformations could eliminate legitimate data values from the analysis. It is therefore suggested that Ward’s Linkage agglomeration procedure could automatically rescale metric data into a range and yield good recovery of clusters.

Despite the significance of groundwater,2 the reported high fluoride variability,4,5 and emergent urbanization1,2 in the region, knowledge on groundwater evolution and quality remain elusive. In this study, we postulate an evolutionary trend of a fractured volcanic aquifer based on multivariate statistics of groundwater quality data of the area. These joint hydrochemical evaluations of
relationships among deep groundwater wells and time-series groundwater quality variations are a fundamental step needed vis-à-vis variability in groundwater processes.

Materials and Methods

Study area

The Olbanita aquifer system is located in the lower Baringo basin of the Kenyan Rift (Figure 1). It is characterized by proximity to the equator and by its elevation which ranges from 1750 meters to 1880 meters above sea level. According to National Council for Population and Development,\textsuperscript{18} the demographic profile for Nakuru County (wherein the Olbanita aquifer system is located) in 2009 was 1602637 and is projected to increase to 2400367 in 2030 and 3013869 in 2050. The city of Nakuru, situated in the upper Lake Nakuru Basin in the south, relies heavily on the 8 wells drilled in the adjacent Lower Baringo basin for potable water supply. As noted by GOK,\textsuperscript{19} water availability problems in the greater Nakuru and Baringo basins had negatively impacted on resident communities and the regional gross domestic product. Within the city, where the abstracted water is used, high incidences of

Figure 1. Location of Olbanita aquifer in Kenya showing sampled boreholes (blue circled dots).
diarrhoeal infections become significant within the overall disease burden.20

The hydrogeology of the area comprises fractured and weathered volcanic rocks and lacustrine sediments. The weathered tuffs sandwiched between the Samburu basalts and the Wasagess flows (phonolites and trachytes) of the Rumuruti group form the best aquifers in Olbanita area. The near N-S trending fault systems interrupting the aquifer system at 5 sites provide groundwater high porosity pathways. At a regional semi-arid climate, the open-faulted drainage, together with the semi-arid climate, renders groundwater in the area indispensable for human subsistence. The confined aquifer in the area recharges from the high altitude areas of Bahati (east of the area) and Menengai, where ecological conditions permit households to grow coffee, tomato, and maize as high-value enterprises.

**Sampling and analysis methods**

The sampling design in this study involved 6 out of 8 sampling sites constituted of boresholes in the Olbanita aquifer system located in the Kenya Rift (Figure 1). The sampling sites were located using a Garmin GPSmap60CSx model. At each sampling site, 4 replicate samples were taken monthly in 250-mL high-density plastic bottles which had already been precleaned by using concentrated nitric acid and drenched in deionized water. The samples were screened through Millipore membrane filters of pore size 0.2 micrometers to remove suspended solids. Thereafter, the water samples for sulfate analysis were precipitated using 0.2 M Zn(CH₃COO)₂ while samples for metal analysis were acidified in concentrated HCl or HCO₃, 1mL per 100mL sample. In situ parameters (pH, electrical conductivity, and total dissolved solids) were measured using Hanna Instruments multiparameter meter model HI98194. Out of the 4 samples taken in plastic bottles, 2 replicate samples were stored in ice and taken to the laboratory for determination of geochemical parameters, namely Na⁺, K⁺, Ca²⁺, Cl⁻, F⁻, SO₄²⁻, H₂S, Total CO₂, CO₃²⁻, HCO₃⁻, H₂CO₃, and NH₄⁺. The concentrations of sulfate and chloride components in this study were quantified using ultraviolet-visible spectrophotometer model UV-1800 while the ions of Na, K, Ca, and NH₄ were quantified using an Ion-Selective Electrode (ISE) model ELIT 9801 as detailed by Victoria and the US Environmental Protection Agency (US EPA).21 Concentrations of nonionic components namely total carbonate carbon (TCC) and H₂S were quantified using titrimetric methods as also described in Victoria and the US EPA.21 Before ISE analysis, the acidified samples were re-digested using a strong alkali solution. TCC was further speciated into CO₃²⁻, CO₂, H₂CO₃, and HCO₃⁻.

To evaluate the appropriateness of geochemical data for factor analysis, Kaiser-Mayer-Olkin (KMO) measure of sampling adequacy and Barlett’s test of sphericity were performed.22,23 The KMO index was found to be 0.6 indicating a moderate degree of correlation among the variables could be appropriately carried out. In addition, Barlett’s test of sphericity was significant (Barlett χ² = .570, 66 df; P < .001), confirming the suitability of factor analysis on the data collected in this study. The results of the assessment depict a lower number of variables (factors) that may be used to elucidate the variability in the hydrochemical data. The correlation matrix was further inspected for correlation coefficients greater than 0.3.24 Statistically significant correlations in physical-chemical data sets for each season were identified through the application of the Spearman’s rho for nonnormally distributed data. Most of the correlation coefficients are over 0.3 indicating that factor analysis may be utilized to provide significant reductions in data dimensionality. Helsel25 and Rangeti et al.26 warn about the dangers of such simplified substitutions for nondetectable concentrations (by using numerical surrogates such as one-half the detection limit). Data were, therefore, automatically rescaled for HCA (because each predictor variable adopted a different scale of measurement), though exempted data sets for PCA. The laboratory geochemical data sets were utilized for PCA considering that the bottom-line correlation matrix (based on the Spearman rank) has the effect of standardizing the variable data.27 Therefore, it is a more robust estimation technique that is less responsive to outliers compared with the widely used Pearson correlation matrix.27

Groundwater quality data were further subjected to PCA and HCA. All statistical computations were executed using of MS Excel spreadsheet and SPSS software version 20.0. For the PCA matrix, orthogonalization of factors was based on the rotated varimax method (with significant eigenvalue loadings > ±0.5) and a derived scree plot (with the criterion of eigenvalues >1) was inspected for purposes of extracting varifactors.28 For HCA, the software’s algorithm utilized Euclidean distances and “sum of squared errors” to minimize the criterion function.

**Results and Discussion**

**General consideration of data sets**

Based on the test of skewness which uses the arithmetic mean and standard deviation, the physical-chemical parameters such as total dissolved solids (TDS), pH, electrical conductivity (EC), and the ions such as Cl⁻, Na⁺, F⁻, Ca²⁺, SO₄²⁻, H₂CO₃, CO₃²⁻, and NH₄⁺ are considered to follow a normal frequency distribution across the study area and between the sampling seasons. On the contrary, CO₂ and NH₄ depicted Gaussian distribution (i.e., a significant departure from normality with a skew statistic >2) (Table 1). The binomial nature of the former is an indication of natural substandard waters characterized by high mineralization. Supersaturation of Olbanita groundwater may be linked to processes like the dissolution of halide, ion exchange, and weathering of sodium-rich plagioclases (usually giving rise to clay mineralogy). The high concentration of H₂S is an indication of very deep circulation attaining anoxic conditions accompanied by abundant bacterial activity.

To offset the effects on mineral dissolution caused by dilution by meteoric waters from the analyses, seasonal data sets were analyzed independently. The initial step in the analysis...
was to account for the extent of mutual variability between individual pairs of water quality variables during the separate seasons. The inter-item correlation matrix of the measured parameters during the dry season and the wet season is provided in Tables 2 and 5, respectively.

Generally, the rotated component matrix was found to contain both positive and negative loadings (Tables 4 and 7). The work of Liu et al.30 as enumerated in Mohapatra et al.31 observed that eigenvalue loadings near ±1 designate a strong association between a variable and a principal component (PC); eigenvalues exceeding ±0.75 represent strong correlation, eigenvalues between ±0.5 and ±0.74 represent moderate correlation and those approaching 0 depict weak correlations. Each PC was attributed to a process owing to which the corresponding variables are probably linearly linked. The underlying processes occurring within the aquifer as construed from the consequent eigenvalue loadings are presented in Tables 4 and 7 where significant eigenvalue loadings are indicated by using the * mark.

**Dry season water quality parameters**

For the dry season data, the pairs pH-F, TDS-EC, TDS-CO2, TDS-HCO3, EC-EC, EC-HCO3, Cl-SO4, Ca-K, Ca-NH4, F-NH4, CO2-HCO3, as well as HCO3-CO3 showed strongly significant relationships. The pairs pH-Na, pH-Ca, pH-H2CO3, pH-NH4, TDS-Na, EC-Na, K-Cl, Cl-CO2, Cl-HCO3, Na-F, Na-HCO3, SO4-CO2, SO4-HCO3, Ca-F, NH4-H2CO3, and HCO3-H2S depicted moderate correlations (Table 2).

### Principal components extracted.

Based on the eigenvalues > 1 criterion, 4 principal components explained variability in groundwater quality at the site (Figure 2). Based on the cumulative variance of the rotation sums of squared loadings of the dry season, the retained latent constructs account for 90.1% of the variance in the data set (Table 3). PC1 with the largest eigenvalue accounted for a maximum of the total variability (34.8%). PC 2 accounted for the total variation of 21%. The third PC explained 20.8% of the total variance, whereas final PC explained 13.6% of the remaining variation in the data. Observed eigenvalue decomposition corresponds to earlier observations by Hossain et al.32 that after the first PC, the second PC explains the greatest of the residual variance and so forth. The first principal component, PC1 (the conductivity component), is associated with significantly high concentrations of electrolytic ions indicated by TDS, EC, Na+, Cl−, SO42−, CO2, and HCO3−. TDS depicted a strong correlation with EC (+0.99, α = 0.01) due to the extensive range in the solubility of/and mineral diversity within the aquifer system. Under sluggish flow during the dry season, groundwater can attain chemical saturation regarding TDS.

### Table 1. Descriptive statistical data for all the parameters for the sampling period (n = 6, N = 42).

| N STATISTIC | RANGE | MINIMUM | MAXIMUM | MEAN | SD | SKEWNESS |
|-------------|-------|---------|---------|------|----|----------|
| pH          | 42    | 1.670   | 7.400   | 9.070 | 8.16452 | 0.441495 | 0.105 | 0.365 |
| TDS         | 42    | 156.000 | 242.000 | 398.000 | 320.1667 | 45.152272 | −0.231 | 0.365 |
| EC          | 42    | 312.000 | 482.000 | 794.000 | 641.80952 | 92.302324 | −0.162 | 0.365 |
| Cl          | 42    | 10.640  | 10.180  | 20.820 | 14.69536 | 2.686202 | 0.948 | 0.365 |
| Na          | 42    | 127.420 | 134.320 | 261.740 | 173.22667 | 35.178236 | 1.232 | 0.365 |
| SO4         | 42    | 19.890  | 0.900   | 20.790 | 7.85314   | 6.578745 | 0.705 | 0.365 |
| F           | 42    | 9.580   | 3.230   | 12.810 | 6.36310   | 3.310061 | 0.804 | 0.365 |
| Ca          | 42    | 15.190  | 2.150   | 17.340 | 6.40024   | 4.675027 | 1.659 | 0.365 |
| K           | 42    | 11.500  | 1.970   | 13.470 | 6.56071   | 3.822435 | 0.284 | 0.365 |
| CO2         | 42    | 168.206 | 134.404 | 302.610 | 230.40724 | 52.722064 | 0.151 | 0.365 |
| H2CO3       | 42    | 28.816  | 0.758   | 29.574 | 8.04064   | 7.570068 | 1.358 | 0.365 |
| HCO3        | 42    | 220.384 | 181.518 | 401.903 | 306.08954 | 69.488132 | 0.118 | 0.365 |
| CO3         | 42    | 33.031  | 0.490   | 33.522 | 5.42143   | 6.540314 | 2.469 | 0.365 |
| H2S         | 42    | 0.720   | 0.000   | 0.720 | 0.03374   | 0.109123 | 6.360 | 0.365 |
| NH4         | 42    | 0.530   | 0.360   | 0.890 | 0.56310   | 0.159381 | 0.441 | 0.365 |
| Valid N (listwise) | 42 |        |         |       |       |         | 0.365 |

Units: ppm (except for pH and EC. EC in μS/cm).
Table 2. Spearman rank correlation coefficients of the physicochemical parameters of groundwater for the dry season (November 2017–February 2018).

|       | PH   | TDS  | EC    | Cl−   | Na+   | SO4²− | F−    | Ca²+  | K+   | CO₂   | HCO₃⁻ | CO₃²⁻ | H₂S   | NH₄⁻  |
|-------|------|------|-------|-------|-------|--------|-------|-------|------|-------|-------|-------|-------|-------|
| pH    | 1.000| −0.009| 0.067 | −0.156| 0.533**| −0.003 | 0.763**| −0.525**| −0.216| −0.109| −0.590**| 0.018 | 0.611**| −0.161| −0.683**|
| TDS   | 1.000|      | 0.985**| −0.463*| 0.702**| −0.446*| 0.374 | 0.278 | −0.017| 0.813**| 0.226 | 0.756**| −0.015| −0.373 | 0.147 |
| EC    | 1.000|      |       | −0.459*| 0.735**| −0.404 | 0.430*| 0.247 | −0.031| 0.779**| 0.182 | 0.756**| 0.034 | −0.341| 0.098 |
| Cl−   | 1.000|      |       |       | −0.373 | 0.842**| −0.200| 0.288 | 0.526**| −0.676**| 0.230 | −0.692**| −0.371| 0.383 | 0.037 |
| Na+   | 1.000|      |       |       | −0.400 | 0.679**| −0.065| −0.018| 0.499* | −0.151| 0.549**| 0.298 | −0.221| −0.233|
| SO4²− | 1.000|      |       |       |       | −0.464*| 0.679**| −0.714**| 0.093 | −0.687**| −0.222| 0.437* | −0.223|
| F−    | 1.000|      |       |       |       |       | −0.546**| −0.425*| 0.106 | −0.337| 0.202 | 0.416*| −0.004| −0.752**|
| Ca²+  | 1.000|      |       |       |       |       |       | −0.203| 0.472*| 0.126 | −0.424*| −0.072| 0.774**|
| K+    | 1.000|      |       |       |       |       |       |       | −0.229| 0.294 | −0.235| −0.314| 0.082 | 0.513*|
| CO₂   | 1.000|      |       |       |       |       |       |       |       | 0.089 | 0.929**| 0.120 | −0.486*| 0.331|
| HCO₃⁻ | 1.000|      |       |       |       |       |       |       |       |       | −0.966**| −0.358| 0.545**|
| CO₃²⁻ | 1.000|      |       |       |       |       |       |       |       |       |       | 0.149 | −0.593**| 0.273|
| H₂S   | 1.000|      |       |       |       |       |       |       |       |       |       |       | −0.492*|       |
| NH₄⁻  | 1.000|      |       |       |       |       |       |       |       |       |       |       |       | 0.235 |

The strong correlations between variables are specified by coefficients in bold fonts.

*Correlation is significant at α = 0.05 level (2-tailed).

**Correlation is significant at α = 0.01 level (2-tailed).
associate Na+ with the dissolution or chemical weathering of sodium-rich plagioclases (to produce clay minerals) or the dissolution of halide. The moderate correlation observed between Na and F (+0.7) and Ca-F (-0.5) coupled with a relatively weak association between the pairs; Na-Cl (-0.4) supports partial derivation of electrolytic ions from weathering of Na-rich feldspatic rocks, dissolution of accessory mineral apatite as well as carbonate materials as opposed to dissolution of either halide rocks. The insoluble products of rock weathering such as Cl− and SO42− show a strong positive correlation (+0.84, α = 0.01) but cumulatively tend to inhibit the electrical conductivity of groundwater. Within the hydro-geological framework, the pattern and, therefore, the provenance of weathering can be accounted for by the roughly N-S fracture-fissure zones (Figure 1).

The second principal component, PC 2 (the salinity component), is depicted mainly by Cl−, SO42−, Ca2+, and K+ ions in water. The anomalous distribution of Cl−, SO42−, and Ca2+ is attributable to ion exchange mechanisms in saturated aquifer zones. The relatively strong association between the pairs Cl-SO4 (+0.8), and the moderate association between the pairs Ca-F, K-Cl, SO4-CO2, Cl-HCO3, and SO4-HCO3, indicate that aquifer water salinity is chiefly attributed to geologic derivation. In addition, agriculture is equally a major pollution source owing to the skewed distribution of Ca2+ which depicts strong correlations with K+ ions (+0.8) and NH4+ (+0.8) at α = 0.01.

The third principal component, PC 3, is the alkalinity component indicated by pH, and Na+, F−, H2CO3, and NH4+. The pH, Na, and F have positive loadings, whereas H2CO3 and NH3 have moderate negative loading on this PC. H2CO3 and NH3 are slightly broken down (ie, at about 25°C) to release H+, HCO3−, CO32−, and NH4+ thereby reducing the pH. Reduced pH significantly increases the rates of weathering introducing more Na and F from geologic sources.

The fourth PC represents carbonate equilibrium with exogenic pH control. The external factor controlling pH is H2S. Sulfide in borehole waters is probably due to inorganic and bacterial changes in the deep aquifer under low dissolved oxygen, optimum growth range in pH (between 5.5 and 8.5) and optimum temperature (between 24°C and 42°C). Besides the authigenic derivation, meteoric waters may also leach agricultural SO42− which may also undergo reduction at depth to sulfides. Unfortunately, the latter may effectively be ruled out due to a complete lack of strong correlation neither with sulfide nor with ammonia (Table 2). Deeper depths within the bedrock are readily flushed by low mobility, oxygen-deficient groundwater; effects that are conditioned by low permeability at depth. The sulfate-reducing bacteria, Desulfovibrio desulfuricans obtain energy via the interconversion between sulfates and sulfides within the larger sulfur cycle in the aquifer system. Under these conditions, sulfates may be reduced to sulfides producing metallic sulfide which is again changed to H2S under the action of H2CO3. The dissociation of H2CO3 yields the CO32− and donates H+ which consequently reduces sulfates to sulfides.

### Wet season water quality parameters

During the wet sampling season, the pairs pH-F, pH-H2CO3, pH-CO3, TDS-EC, Na-F, F-H2CO3, F-CO3, CO32-H2CO3, and H2CO3-CO3 showed strong positive correlations at α = 0.01 (Table 5). The pairs pH-Na, pH-NH3, TDS-Cl, TDS-Na, TDS-F, EC-Cl, EC-Na, Cl-Na, K-Cl, Na-CO3, SO4-CO2, SO4-HCO3, SO4-H2S, Ca-F, Ca-CO3, and Ca-NH3 depicted moderate correlations at α = 0.01.

### Table 3. Total variance explained for the dry season data.

| COMPONENT | INITIAL EIGENVALUES | EXTRACTION SUMS OF SQURED LOADINGS | ROTATION SUMS OF SQURED LOADINGS |
|-----------|---------------------|------------------------------------|----------------------------------|
|           | TOTAL | % OF VARIANCE | CUMULATIVE % | TOTAL | % OF VARIANCE | CUMULATIVE % | TOTAL | % OF VARIANCE | CUMULATIVE % |
| 1         | 6.265 | 41.763       | 41.763       | 6.265 | 41.763       | 41.763       | 5.226 | 34.842       | 34.842       |
| 2         | 3.611 | 24.077       | 65.840       | 3.611 | 24.077       | 65.840       | 3.143 | 20.955       | 55.797       |
| 3         | 2.082 | 13.882       | 79.722       | 2.082 | 13.882       | 79.722       | 3.117 | 20.783       | 76.580       |
| 4         | 1.565 | 10.434       | 90.156       | 1.565 | 10.434       | 90.156       | 2.036 | 13.576       | 90.156       |

Figure 2. Cattel scree plot with eigenvalues > 1 criteria.
Principal components extracted. Based on the cumulative variance of the rotation sums of squared loadings of the wet season, the retained latent constructs account for 91.1% of the variability of the data set (Table 6).

PC1 with the largest eigenvalue accounted for the maximum of the total variability (23.5%). PC2 accounted for the total variation of 22.8% and corresponds in concept to the first PC. The third and fourth PCs explained 22.8% and 22% of the total variance, respectively. Cattel’s scree test plot is presented in Figure 2.

The first principal component, PC1 (the salinity component), is depicted mainly by $K^+$, $Ca^{2+}$, $Cl^-$, $SO_4^{2-}$, $H_2S$, and $NH_4^-$ ions in water. As the component explains the largest variance in the data, it can be inferred that the groundwater in the study area is mainly saline. As nearly all carbonate ions depict a negative correlation in explaining PC1 and the non-carbonate alkali exceeds 50% in all (except BH7) boreholes (Figure 3), it can be deduced that NaCl constitutes the primary salinity at Olbanita. Abnormal distribution of $Ca^{2+}$ and $SO_4^{2-}$ may be characteristically associated with the ion exchange mechanisms in saturated aquifer zones. At the wet season pH, for instance, the monovalent $Na^+$ is depleted from groundwater as it substitutes for divalent $Ca^{2+}$ on exchangeable micro-pore surface water interfaces (Table 7). This inference is demonstrated by BH 7 where none of the cation or anion pairs exceed 50% setting forth a secondary salinity characterized by Ca-Mg-$SO_4$ and/or chloride mixed type waters (Figure 3). The ions $SO_4^{2-}$ and $H_2S$ depict a strong positive correlation ($+0.8$ at $\alpha=0.01$) indicating either geo-logic provenance, probably due to deep circulation of oxygen-saturated waters causing aerobic conditions (Table 5), or anthropogenic derivation, probably due to their simultaneous positive contribution to PC1 along with ammonia (Table 7). The moderate correlations within the pairs $NH_3-Ca$ ($+0.7$) and $NH_3-H_2CO_3$ ($+0.6$) imply anthropogenic contributions to the pH controls on aquifer water salinity. During the wet season, decreases in $H^+$ cause a reduction in the water concentrations of $K^+$, $Ca^{2+}$, and $Cl^-$. Aquifer salinity, therefore, is chiefly a construct of water-rock interactions and to a lesser extent anthropogenic inputs.

The second principal component, PC 2, depicts carbonate equilibrium with external pH control. The external factors controlling pH are $SO_4^{2-}$ and $H_2S$, which are strongly correlated ($+0.8$, $\alpha=0.01$). Sulfate and sulfide transformations in borehole waters are invariably mediated by bacterial changes under anaerobic conditions. $D$ desulfuricans may also produce $H_2S$ under conditions of the measured pH range (optimum growth range in pH is between 5.5 and 8.5) and

### Table 4. Extracted factor loadings of the measured parameters during the dry season which suited the provisions of orthogonal varimax rotation.

| COMPONENT | 1    | 2    | 3    | 4    | COMMUNALITIES |
|-----------|------|------|------|------|---------------|
| pH        | 0.037| -0.230| 0.885*| 0.087| 0.844         |
| TDS       | 0.968*| -0.112| 0.091| -0.003| 0.957         |
| EC        | 0.959*| -0.083| 0.156| 0.012| 0.952         |
| Cl        | -0.583*| 0.763*| -0.126| -0.132| 0.955         |
| Na        | 0.725*| 0.094| 0.613*| -0.104| 0.921         |
| SO_4^{2-} | -0.548*| 0.775*| 0.086| -0.111| 0.920         |
| F^-       | 0.373| -0.053| 0.897*| -0.056| 0.950         |
| Ca^{2+}   | -0.205| 0.889*| -0.299| -0.022| 0.923         |
| $K^+$     | 0.134| 0.870*| -0.234| -0.183| 0.862         |
| $CO_2$    | 0.952*| -0.174| -0.111| 0.111| 0.961         |
| $H_2CO_3$ | 0.219| 0.204| -0.555*| -0.543| 0.693         |
| $HCO_3^-$ | 0.943*| -0.184| -0.077| 0.056| 0.932         |
| $CO_3^{2-}$ | 0.139| -0.168| 0.254| 0.904*| 0.930         |
| $H_2S$    | 0.063| -0.045| -0.185| 0.902*| 0.854         |
| $NH_4^+$  | 0.370| 0.440| -0.728*| -0.094| 0.869         |

Note: Significant eigenvalue loadings are indicated by using the * mark.
Table 5. Spearman rank correlation coefficients for physiochemical parameters of groundwater for the wet season (March-May 2018).

|     | PH  | TDS | EC   | Cl⁻  | Na⁺  | SO₄²⁻ | F⁻   | Ca²⁺ | K⁺  | CO₂ | H₂CO₃ | HCO₃⁻ | CO₃²⁻ | H₂S  | NH₃  |
|-----|-----|-----|------|------|------|-------|------|------|-----|-----|-------|-------|-------|------|------|
| pH  | 1.00| .010| −.007| −.239| .621**| .225  | .873**| −.557*| −.161| −.317| −.980**| −.240 | .978**| −.041| −.666**|
| TDS | 1.00| .991**| −.636**| .704**| −.235| .258  | .003 | −.019| .231 | .045 | .287  | .072  | −.350 | .134 |
| EC  | 1.00| −.645**| .686**| −.226| .224 | .015  | −.011| .249 | .067 | .298 | .053  | −.319 | .124 |
| Cl⁻ | 1.00| −.600**| .538* | −.265| .112 | .629**| −.406| .166 | −.459| −.259| .575* | .272 |
| Na⁺ | 1.00| −.261| .788**| −.370| −.150| .119  | −.542*| .193 | .697**| −.433| −.280 |
| SO₄²⁻| 1.00| .082| −.022| .534*| −.686**| −.335 | −.699**| .108 | .760**| −.081|
| F⁻  | 1.00| −.705**| −.208| −.381| −.853**| −.309 | .867**| −.169| −.581*|
| Ca²⁺| 1.00| .461| .598**| .558*| .556*| −.490*| .123 | .656**|
| K⁺  | 1.00| .012| .159| .007| −.106| .511* | .337 |
| CO₂ | 1.00| .418| .979**| −.189| −.385| .225 |
| H₂CO₃| 1.00| .362| −.930**| −.068| .637**|
| HCO₃⁻| 1.00| −.098| −.454| .181 |
| CO₃²⁻| 1.00| −.143| −.579*|
| H₂S | 1.00| .010|
| NH₃ | 1.00|

The strong correlations between variables are specified by coefficients in bold fonts.

*Correlation is significant at the .05 level (2-tailed).
**Correlation is significant at the .01 level (2-tailed).
optimum temperature between 24°C and 42°C. Under these conditions, sulfides (in the form of metal sulfide) undergo oxidation to sulfates which are again converted to H₂S under the action of H₂CO₃. The dissociation of H₂CO₃ yields the ions CO₃²⁻ and H⁺ which causes the reduction of sulfates to sulfides. Due to protracted mobility (diminutive residence period) of oxygen-saturated water in the wet season, reduced dissociation of H₂CO₃ yields less CO₃²⁻ and H⁺ which consequently causes oxidation of sulfides to sulfates. BH 7 depicts the mixing effects probably conditioned by such pH shifts resulting from SO₄²⁻ and H₂S variability (Figure 3).

The third principal component, PC 3, is the alkalinity component as indicated by pH, CO₃²⁻, H₂CO₃, F⁻ and NH₄⁺. The pH, CO₃²⁻, and F have positive loading, whereas H₂CO₃ and NH₄⁺ have negative loading on this PC. The PC explains the dissolution of fluoride through microbial activity releasing ammonia and thereby decreasing the pH. In addition, weak acids such as H₂CO₃ ionize sequentially releasing CO₃²⁻, which again raises the pH of water.

The fourth PC is the electrolytic conductivity component which is associated with significantly high concentrations of electrolytic ions indicated by TDS, EC, Na⁺, and F⁻. Statistical analyses (eg, Pazand⁹) associate F⁻ with weathering of the fluoro-apatite and silicate mineralogy, whereas Moghimi¹² linked Na⁺ with the dissolution or weathering of sodium-rich plagioclases (clay mineralogy) or the dissolution of halide. The strong correlation observed between Na⁺ and F⁻ (0.9) and the weak association between Na⁺ and Cl⁻ (-0.6) supports weathering plagioclase feldspars as the chief source as opposed to the dissolution of halide. Na and F ions

Table 6. Cumulative variance explained for the wet season data.

| COMPONENT | INITIAL EIGENVALUES | ROTATION SUMS OF SQUARED LOADINGS |
|-----------|---------------------|----------------------------------|
|           | TOTAL % OF VARIANCE | CUMULATIVE %                      | TOTAL % OF VARIANCE | CUMULATIVE %                      |
| 1         | 5.691               | 37.938                           | 37.938              | 3.519                              | 23.460                           | 23.460                           |
| 2         | 4.423               | 29.489                           | 67.427              | 3.427                              | 22.844                           | 46.304                           |
| 3         | 2.273               | 15.154                           | 82.581              | 3.420                              | 22.799                           | 69.103                           |
| 4         | 1.278               | 8.521                            | 91.102              | 3.300                              | 21.999                           | 91.102                           |

Figure 3. Major composition in groundwater during the wet season.
are the intrinsic constructs responsible for electrolytic conduction, as supported by their strong correlations with TDS. Within the hydrogeological framework, the pattern and, therefore, the provenance of weathering can be accounted for by the roughly N-S fracture-fissure zones. The component accounts for the lowest variability because of the reduced residence time of groundwater during the wet season exerting substantial reduction in TDS.

The dry to wet seasonal shift in variability from electrolytic conduction to low alkali salinity, respectively, is probably due to differential seasonal rates of weathering, flow and dilution processes in the aquifer. Significant correlations (at \( \alpha = 0.01 \)) in the pairs Cl-SO\(_4\)\(^{2-}\), Ca-K, Ca-NH\(_3\), F-NH\(_3\), CO\(_2\)-HCO\(_3\), HCO\(_3\)-CO\(_3\), Na-F, and F-H\(_2\)CO\(_3\) indicate that the water is alkaline to mildly acidic which are a manifestation of authigenic and to a lesser scale anthropogenic imprints. Key groundwater evolutionary trends suggest that silicate, carbonate, and/or accessory mineral apatite dissolution, as well as ion exchange at sorption sites with the clay-water interface, are the central sources of variability in the groundwater chemistry of the aquifer. Protracted mobility of oxygen-saturated water in the wet season, probably conditioned seasonal by SO\(_4\)\(^{2-}\) and H\(_2\)S variability, though being an exogenic factor, restrained the pH constraint and hence carbonate equilibrium mixing reactions.

### Spatial variability between sampling sites

The results of hierarchical clustering procedures were discrete clusters presented graphically in the form of a dendrogram by an averaging algorithm (Figure 4).

Based on rescaled Euclidean distances and the “sums of squared errors,” 2 main borehole clusters are conspicuous in the area. [Where to “prune” the tree (eg, using the continuous bold lines) is a vital factor in interpreting the results of the analysis. The within-cluster medium depict translational invariance in sample composition. Alternate shading was thus introduced to facilitate review]. The first cluster (forming the left-hand group) consists of the western zone cases (boreholes 2, 3, 4, 5 and 6), whereas the second cluster (forming the right-hand group) consists of the eastern zone borehole cases (boreholes 1, 7, and 7A). The former boreholes were deciphered to be hydraulically connected by a major inferred north west – south east fault which corresponds to a calibrated transmissivity-formation resistivity

| COMPONENT | 1 | 2 | 3 | 4 | COMMONALITIES |
|-----------|---|---|---|---|---------------|
| pH        | -0.166 | -0.169 | 0.948* | 0.157 | 0.980 |
| TDS       | -0.148 | 0.292 | 0.060 | 0.934* | 0.983 |
| EC        | -0.194 | 0.280 | 0.010 | 0.921* | 0.964 |
| Cl\(^-\)  | 0.830* | -0.292 | -0.155 | -0.379 | 0.942 |
| Na\(^+\)  | -0.071 | -0.203 | 0.285 | 0.925* | 0.983 |
| SO\(_4\)\(^{2-}\) | 0.694* | -0.702* | 0.043 | -0.059 | 0.980 |
| F\(^-\)   | -0.071 | -0.246 | 0.735* | 0.596* | 0.961 |
| Ca\(^{2+}\) | 0.865* | 0.002 | -0.191 | -0.307 | 0.879 |
| K\(^+\)  | 0.929* | -0.040 | -0.198 | 0.148 | 0.926 |
| CO\(_2\)  | -0.088 | 0.967* | -0.147 | 0.092 | 0.974 |
| H\(_2\)CO\(_3\) | 0.102 | 0.485 | -0.762* | -0.023 | 0.827 |
| HCO\(_3\) | -0.102 | 0.988* | -0.122 | 0.100 | 0.973 |
| CO\(_3\)\(^{2-}\) | -0.118 | 0.237 | 0.944* | 0.061 | 0.966 |
| H\(_2\)S | 0.572* | -0.503* | 0.081 | -0.237 | 0.643 |
| NH\(_4\) | 0.519* | 0.367 | -0.529* | -0.038 | 0.685 |

| Aquifer salinity | Carbonate equilibrium with external pH controls | Alkalinity of water | Electrolytic conduction |
|-------------------|-----------------------------------------------|---------------------|-------------------------|

Note: Significant eigenvalue loadings are indicated by using the * mark.
model for the aquifer. Intrinsic permeability was empirically elevated along major fracture traces, consequently increasing yields of the affected boreholes. In addition, some wet season samples obtained from boreholes 4 (sample Nos. 27 and 33), 5 (sample Nos. 28, 34 and 40), and 6 (sample Nos. 29, 35 and 41) formed a mini-cluster within the former main group. The samples represent the effects of a high permeability fault/fracture structure which accentuates deep circulation of oxygen-saturated waters from recent precipitation events coupled with dilution within the corresponding season. We suggested low residence times of groundwater in the zone. Worth noting was the strong indication of low-carbonate alkali to mildly acidic mixed waters in the western zone at BH 6, which transits to become Ca-Mg sulfate and/or chloride water types of at BH 7 in the eastern periphery in the wet season.

The latter cluster contains boreholes located in the eastern compartment of the aquifer. Elevated values of Cl$^-$, SO$_4^{2-}$, and exceptionally low values of EC, TDS, pH, Na$^+$, T-CO$_2$, and HCO$_3^-$ were recorded in these groundwater boreholes. However, values of EC, TDS, Ca$^{2+}$, SO$_4^{2-}$, and K$^+$ depicted an upward trend, whereas those of Na$^+$, T-CO$_2$, CO$_3^{2-}$, and HCO$_3^-$ showed a downward trend from the dry season toward the wet season for these boreholes. By contrast, boreholes in the zone are not hydraulically connected via major fault structures. We, therefore, suggested that pore-level adsorption/desorption processes in which vast quantities of monovalent ions such as Na$^+$ and K$^+$ are removed from the groundwater in exchange for divalent ions control the observed variability in groundwater hydrochemistry in this zone. Lack of aquifer-scale hydraulic networks and the presence of clay micropores signify extended groundwater residence time favoring ion exchange reactions in the zone. Borehole 8 drilled in this zone dried-up after its completion because a clay layer was inadvertently targeted for production.

**Conclusions**

PCA and HCA are robust methods for establishing aquifer evolutionary structures. In this study, the PCA technique condensed multidimensional data into factors that explained seasonal variability in groundwater aquifer trends and quality. The dry to wet seasonal shift in variability from electrolytic conduction to salinity, respectively, is probably due to differential seasonal rates of weathering, flow, and dilution processes in the aquifer. The authors believe that key groundwater evolutionary trends, water-rock interactions, as well as ion exchange at sorption sites with the clay-water interface are the central restrictions of groundwater chemistry variability. HCA partitioned the aquifer into 2 discrete spatial associations, in spite of their indicated shared recharge area. These agglomerative scheduling validated in an integrative approach (with groundwater flow predictions using a calibrated petrophysical groundwater model for the area), linked each aquifer compartment to aquifer spatial heterogeneities and processes. The authors incontrovertibly deciphered groundwater residence periods for each compartment, diminutive for the western zone and protracted for the eastern zone. It is convincingly essential, therefore, based on pH shifts per season to design a groundwater quality monitoring plan and policy that reduces the number of measured parameters purposely to provide an opportunity cost in terms of resources for measurements elsewhere. A sustainable alternative would be to measure (as surrogate for the presence of the remaining parameters) EC during the dry season and K$^+$, Ca$^{2+}$, and Cl$^-$ during the wet seasons. The additional analyses may be required during extended dry periods accompanied by an upward trend in EC measurements.

**Author Contributions**

Conceptualization, B.S. and A.G.; methodology, B.S. and A.G.; software, B.S.; validation, B.S., A.G., S.M. and J.B.; formal analysis, B. S. and A.G.; investigation, B.S.; resources,
A.G., S.M. and J.B.; data curation, B.S. and A.G.; writing, B.S.; writing—review and editing, B.S., A.G., S.M., and J.B.; visualization, B.S. and A.G.; supervision, A.G., S.M., and J.B.; project administration, B.S. and S.M.; funding acquisition, A.G., S.M., and J.B.

Data availability
The data sets used and generated in this study are available upon request to sosibenjamin@yahoo.com.

ORCID iD
Benjamin Sosi https://orcid.org/0000-0001-6104-7011

REFERENCES
1. Nyika J, Onyari E. Hydrogeochemical analysis and spatial distribution of groundwater quality in roundhill landfill vicinity of South Africa. Air Soil Water Res 2019;12:1-8.
2. Xu Y, Seward P, Gaye C, Olago D. Preface: groundwater in sub-Saharan Africa. Hydrog Earth J 2019;27:815-822.
3. World Health Organization (WHO). Guidelines for Drinking-Water Quality. Geneva, Switzerland: WHO; 2011.
4. Raj D, Shaji E. Fluoride contamination in groundwater resources of Alleppey, southern India. Geosci Front 2017;8:117-124.
5. Olaka LA, Wilke FD, Olago DO, Odada EO, Mulch A, Musolf A. Groundwater fluoride enrichment in an active rift setting: Central Kenya Rift case study. Sci Total Environ 2016;545-546:641-653.
6. Ratnottor M, Zanotti C, Fumagalli L, et al. Multivariate statistical analysis supporting the hydrochemical characterization of groundwater and surface water: a case study in northern Italy. Rend Online Soc Geol Ital 2019;47:90-96.
7. Yidana SM, Bawoyobie P, Sakyi P, Fynn OF. Evolutionary analysis of groundwater flow: application of multivariate statistical analysis to hydrochemical data in the Denso Basin, Ghana. J Afr Earth Sci 2018;138:167-176.
8. Ravikumar P, Somashekar RK. Principal component analysis and hydrochemical facies characterization to evaluate groundwater quality in Varahi river basin, Karnataka state, India. Appl Water Sci 2017;7:745-755.
9. Pazand K. Geochemistry and multivariate statistical analysis for fluoride occurrence in groundwater in the Kuhbanan basin, Central Iran. Med Earth Syst Environments 2016;2:72.
10. Güler C, Thyne GD, Tağa H, Yıldırım Ü. Processes governing alkaline groundwater quality in the eastern parts of the Kula basin, Turkey. Geofluids 2017;2:1-21.
11. Nwanwkoala HO. Interpretation of hydro -geochemical characteristics of deep aquifers in parts of Port Harcourt, Eastern Niger Delta. Stand Sci Res Essays 2013;1:134-163. http://www.standresjournals.org/journals/SSRE. Accessed February 16, 2018.
12. Moghimi H. The study of processes affecting groundwater hydrochemistry by multivariate statistical analysis (case study: coastal aquifer of Ghaemshahr, North Iran). Open J Geol 2017;7:830-846.
13. Esmaeili S, Moghaddam AA, Barzegar R, Tziritis E. Multivariate statistics and hydrogeochmical modeling for source identification of major elements and heavy metals in the groundwater of Qareh-Ziaeddin plain, NW Iran. Arab J Geosci 2018;11:5.
14. Majed S, Rashid S, Qadir A, Mackay C, Hayat F. Spatial patterns of pollutants in water of metropolitan drain in Lahore, Pakistan, using multivariate statistical techniques. Environ Monit Assess 2018;190:128.
15. Smeti EM, Golfinopoulos SK. Characterization of the quality of a surface water resource by multivariate statistical analysis. Anal Lett 2016;49:1032-1039.
16. Cloutier V, Lefebvre R, Therrien R, Savarini MM. Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. J Hydrol 2008;353:294-313. doi:10.1016/j.jhydrol.2008.02.015.
17. Everitt BS, Landau S, Leese M, Stahl D. Hierarchical clustering. Cluster Anal 2013;1:71-110.
18. National Council for Population and Development (NCPD). 2015 Kenya National Adolescents’ and Youth Survey (NAYS). Nairobi, Kenya: NCPD, 2017.
19. GOK. Rift valley water supply and sanitation project appraisal report. Working paper. Nakuru, Kenya: Rift Valley Water Services Board; 2004.
20. Muchakura E, Grenier PR. Social determinants of health and health inequities in Nakuru (Kenya). Int J Equity Health 2009;8:16.
21. Victoria E. Industrial waste resource guidelines sampling and analysis of waters, wastewaters, soils and wastes. Environmental Protection Authority. Victoria; 2009.
22. Voza D, Vuković M. The assessment and prediction of temporal variations in surface water quality—a case study. Environ Monit Assess 2018;190:434.
23. Joshi RR, Mulay P. Deep incremental statistical closeness factor-based algorithm DIS-CFBA; to assess diabetes mellitus. Biomed 2018;115:210.
24. Tabachnick BG, Fidell LS. Using Multivariate Statistics. Boston, MA: Pearson Education, 2007.
25. Heele D. Much ado about nothing: incorporating nondetects in science. Ann Occup Hyg 2010;54:257-262. doi:10.1093/anoxy/mep092.
26. Rangeti I, Dawairo B, Barratt GJ, Otieno FA. Validity and errors in water quality data: a review. Res Dentis Endodontics 2016;1:1-6. doi:10.1166/rde.2015.036.
27. Kim HY. Statistical use of principal component analysis: assessing normal distribution (2) using skewness and kurtosis. Res Dentis Endodontics 2013;38:52-54.
28. Liu CW, Lin KH, Kuo YM. Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan. Sci Total Environ 2003;315:77-89.
29. Mobapara PK, Vijay R, Pujari PR, Sundaray SK, Mohanty BP. Determination of processes affecting groundwater quality in the coastal aquifer beneath Puripuri city, India: a multivariate statistical approach. Water Sci Technol 2011;64:809-817. doi:10.2166/wst.2011.605.
30. Hossain M, Patras A, Barry-Ryan C, Martin-Dirana A, Brunton N. Analysis of principal component analysis, hierarchical cluster analysis to classify different species based on in vitro antioxidant activity and individual polyphenolic antioxidant compound. J Funct Food 2011;3:179-189.
31. Tostevin R, Craw D, Van-Hale R, Vaughan M. Sources of environmental sulfur in the groundwater system, southern New Zealand. Appl Geochem 2016;1:1-6.
32. Stanley W, Southam G. The effect of gram-positive (Desulfosporosinus orientis) and gram-negative (Desulfovibrio desulfuricans) sulfate-reducing bacteria on iron sulfide mineral precipitation. Can J Microbiol 2018;64:629-637.
33. Nanatharaman K, Brown CT, Hug LA, et al. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nat Commun 2016;7:13219.