Groundwater for Sustainable Development 16 (2022) 100696

Contents lists available at ScienceDirect

Groundwater for Sustainable Development

journal homepage: www.elsevier.com/locate/gsd

Research paper

Estimation of natural background and source identification of nitrate-nitrogen in groundwater in parts of the Bono, Ahafo and Bono East regions of Ghana

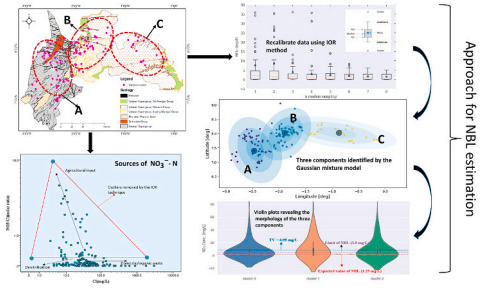
Evans Manu [a,b,c,*], George Yamoah Afrifa [d], Theophilus Ansah-Narh [f], Frederick Sam [e], Yvonne Sena Akosua Loh [d]

[a] GFZ German Research Centre for Geosciences, Telegrafenberg, 14473, Potsdam, Germany
[b] University of Potsdam, Institute of Geosciences, Potsdam, Germany
[c] Council for Scientific and Industrial Research - Water Research Institute, Ghana
[d] Department of Earth Science, University of Ghana, Box LG 58, Legon-Accra, Ghana
[e] Department of Physics, University of Cape Coast, Ghana
[f] Ghana Space Science and Technology Institute (GSSTI), Ghana Atomic Energy Commission (GAEC), Box LG 80, Legon-Accra, Ghana

HIGHLIGHTS

- Natural background levels (NBL) are estimated using iterative outlier removal (IOR) method.
- IOR provides better estimates of NBLs in less heterogeneous datasets.
- NBL is validated by the robust Gaussian mixture model and the violin plot.
- Agriculture, domestic and denitrification are major NO_3^-–N contaminant sources.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Nitrate-nitrogen
Gaussian mixture model
Iterative outlier removal
Co-kriging
Threshold value
Semivariogram

ABSTRACT

The identification of the chemical status of groundwater is a prerequisite for the sustainable management and development of groundwater resources. A better assessment of the chemical status of the groundwater requires the knowledge and understanding of the natural background concentrations to establish threshold values of chemical pollutants in groundwater. The aim of this study is to estimate the natural background levels, the threshold value (TV) of nitrate-nitrogen (NO_3^--N) and to identify its source in the groundwater in the Bono, Ahafo and Bono East regions of Ghana. A total of 165 groundwater samples were taken from the crystalline and sedimentary aquifers in the study area. Iterative outlier removal technique and a Gaussian mixture model were then used to assess the natural background and threshold nitrate-nitrogen concentrations in the groundwater. Chloride data was then used to trace the NO_3^-–N source in the groundwater. The estimated NBL of NO_3^-–N in the area ranges from [0.001–3.9] mg/L with an expected value of 1.25 mg/L and calculated TV of 6.95 mg/L. The data suggest that NO_3^-–N concentrations are homogeneous across all the lithologies underlying the study area. The results showed that agricultural, domestic, and denitrification contribute significantly to the loading of NO_3^-–N concentration in the groundwater. The estimated threshold range of NO_3^-–N provides the baseline

* Corresponding author. GFZ German Research Centre for Geosciences, Telegrafenberg, 14473, Potsdam, Germany.
E-mail address: emanu@gfz-potsdam.de (E. Manu).

https://doi.org/10.1016/j.gsd.2021.100696
Received 5 April 2021; Received in revised form 23 October 2021; Accepted 28 October 2021
Available online 30 October 2021

nitrate-nitrogen concentration for future studies in the region. However, the results are inconclusive, and we recommend using isotopic detection ($^{15}N\text{-}NO_3^-$ and $^{18}O\text{-}NO_3^-$) in future studies through comprehensive and sustainable regional monitoring of the aquifer system in order to further limit the source of nitrate-nitrogen in the groundwater system.

1. Introduction

Groundwater is undoubtedly the most abstracted resource on Earth and very important for the well-being of many ecosystems (Jarvis, 2012). It has not only been an alternative to providing water for domestic use, food production and industrial use (Conti et al., 2016; Llamas and Martínez-Santos, 2005) but instrumental in human development, poverty eradication, human dignity, and general well-being. Meanwhile, the availability of fresh groundwater in many parts of the world is decreasing rapidly and deteriorating in quality due to population growth, industrialization, and agriculture (Bulut et al., 2020; Guo and Bartram, 2019; Niu et al., 2019).

At the heart of the SDGs i.e., 1,2,3,4 and 6 is the urgent demand to ensure clean water for everyone by 2030. In order to achieve these goals and to better assess the geochemical status of groundwater, it is essential to make estimates of the natural background level (NBLs) and set threshold values (TVs) of chemical pollutants in the groundwater. Taking cognizance of the NBL in groundwater enables us to understand the concentration of the geochemical compound free from anthropogenic activities (De Caro et al., 2017; Edmunds, 2008). The NBLs are not only derived from pollutants with purely anthropogenic origin but also from geochemical species with both geogenic and human impacts. A list of such pollutants to consider for NBL determination are discussed in Pauwels et al. (2007).

In this study, nitrate-nitrogen (NO_3^--N) is considered for the estimation of the NBL, because of extensive agricultural activities coupled with the indiscriminate disposal of solid and liquid wastes by the growing population. Nitrate-nitrogen ions occur naturally as part of the nitrogen cycle. However, its increasing tendency in natural waters has made it a pollutant of concern (Anornu et al., 2017). Nitrate-nitrogen comes from a variety of sources, including synthetic and natural fertilization, bacterial production, atmospheric deposition, and leaky septic systems (Anornu et al., 2017; Bordeleau et al., 2008). Biological processes such as nitrite (NO_2^-), ammonium (NH_4^+), and ammonia (NH_3) also convert other forms of nitrogen into NO_3^-–N (Chandran and Smets, 2000). Its occurrence in groundwater poses a serious health risk for consumers (Zhai et al., 2017). For example, high concentration of NO_3^-–N concentration in drinking water can lead to methemoglobinemia/blue baby syndrome, potentially carcinogenic effects, NO_3^-–N poisoning in animals, excessive plants, and algae growth (Anornu et al., 2017; Sadler et al., 2016; Murgulet and Tick, 2009; Burkart et al., 2008). The wide range of high variation of NO_3^-–N sources and the dynamics of their transformation make separating geogenic and anthropogenic sources a daunting task (Panno et al., 2006). However, earlier studies by scientists (Paronne et al., 2019; Sellerino et al., 2019; Geranian et al., 2019; Zhou et al., 2018; Dalla Libera et al., 2017; Kim et al., 2015; Preziosi et al., 2014; Molinari et al., 2012; Panno et al., 2006; Muller et al., 2006) have used various graphical methods, which are divided into graphic (such as probability plot, box plots etc.), descriptive statistics, and pre-selection techniques to determine NBLs and TVs. One of the key limitations of adopting only the existing methods is the subjective determination of NBL and TV which makes it very tedious to measure the uncertainties in the method. With a large amount of data in our time, it is practically impossible to unravel the complex patterns in an observation with the current approach alone. Therefore, a more robust technique is needed to augment the existing methods to address the problem, and this is exactly what is presented. In this paper, we therefore focus on: (i) estimation of the NBL in the groundwater nitrate-nitrogen of the proposed study area. Here, we apply a bootstrap technique to successively remove anomalies in the measured data in order to obtain the NBL values, (ii) calculate the TVs using with the given NBLs in step (i), (iii) test the accuracy and validity of the values gleaned for the NBL and TVs using the robust Gaussian mixture model and finally (iv) develop a spatial NO_3^-–N map to identify zones with high concentrations.

Meanwhile, the study area, Bono, Ahafo and Bono East Regions of Ghana are farming communities noted for several cash crops, food crops and animals. Recently, it has been reported that large river systems in the country, and the study area in particular, are heavily polluted due to factors such as illegal mining, excessive fertilization on farms, indiscriminate disposal of sewage and household waste (Affum et al., 2016; Bempah and Ewusi, 2016; Tay et al., 2014; Armah et al., 2014). Reports of high bacteria load counts in wells in some parts of the study area are also attributed to the septic and contaminant transport through the unsaturated zone (Tekpor et al., 2017).

2. Geology and hydrogeology

Geologically, the study area is dominated by six lithological units: Birimian Supergroup, Tarkwain formation, Oti-Pendjari, Obosom, Kwahu-Morago Group, and Eburnean Plutonic Suite (Fig. 1). The Birimian Supergroup, Tarkwain formation with its intruder granitoid, is part of the Precambrian Basement Rock Complex in Ghana, while the Oti-Pendjari, Obosom, and the Kwahu-Morago Groups are also part of the Voltaian Supergroup. The Birimian Supergroup comprises meta-sediments and meta-volcanic rocks, with greywackes, turbidite features, phyllites, slates, schists, weakly metamorphosed tuffs and sandstones make up the meta-sediments. The Tarkwain formation unconformably overlies the Birimian Supergroup. The Tarkwaian comprises clastic sediments–conglomerates, sandstones, and slates (Leube et al., 1990). Oti-Pendjari Group is made up of two sub-formations, which include the Afram sandstone formation and the Ejura sandstone formation. The Afram sandstone formation comprises mudstone, siltstone, micaceous, and sporadic limestone beds, while the Ejura sandstone formation is characterized by sandstone, thinly bedded, fine to medium-grained, quartzose. The Obosom Group constitutes the Densubon sandstone formation which mainly contains fine to medium grained sandstone, mudstone, and siltstone. The Anyaboni sandstone, Yabraso sandstone and the Damongo sandstone constitute the Kwahu-Morago Group. The hydrogeological terrain of Ghana has been divided into five provinces: Coastal sedimentary Province, Voltaian Province, Pan African Province, Crystalline Basement Granitoid Complex Province, and the Birimian Province (Banoeng-Yakubo et al., 2010). The study area falls within three major Provinces namely, the Birimian Province, Crystalline Basement Granitoid Complex Province and Volatian Province. Together, the rocks underlying the Birimian and Crystalline Basement Granitoid Complex and the Voltaian Provinces are inherently impermeable; however, secondary porosities have developed and serve as the primary control factor for groundwater flow in the aquifers. The saprolite, saprock and fractured bedrock represent the most productive zones within the Crystalline Basement Province (Carrier et al., 2008). In the Voltaian Province, groundwater is mainly controlled by fracture intensity and bedding planes. This results from the reduced primary porosity of some of these rocks due to consolidation and cementation (Banoeng-Yakubo et al., 2010). According to Dapaah-Siakwan and Gyau-Boakye (2000), the success rate of drilling a productive well in the Volatain Province is 55% (Banoeng-Yakubo et al., 2010; Dapaah-Siakwan and Gyau-Boakye, 2000); However, with an improved method,

such as the application of 2D electrical resistivity imaging, the success rate could be increased by 65% (Mainoo et al., 2019).

3. Materials and methods

3.1. Location, vegetation, soil, and climate

The study area (Fig. 2) comprises three newly created regions: Bono East Region (A), Bono Region (B) Ahafo Region (C) formerly called Brong Ahafo Region. It shares an international border with Cote d'Ivoire to the west, Ashanti Region of Ghana to the south, the Savannah region to the north and Oti region in the east. The three regions together have a total population of 2,610,332 (Ghana, 2010). Nine rivers drain the area: The Black Volta, Tano, Bia, Pru, Afram, West Bia, Obosum and the Volta Lake (Fig. 2). The vegetation cover of the study area consists of a moist semi-deciduous forest and a Guinea savannah forest (Fig. 3). The Savannah land cover consists of close and open cultivated savannah woodland, widely open cultivated savannah, and open savannah woodlands (Fig. 3). The semi-deciduous forest comprises closed forest and open forest, moderately closed trees, moderately dense herb/bush with scattered trees (Fig. 3). These two types of vegetation are very rich and support a wide variety of cash and food crops. With its rich vegetation, most of the four regions farm for a living. Recently there has been an increase in mining activities, including large scale (mainly gold, bauxite, manganese) and small-scale mining (mainly gold), which are also carried out by part of the population. The region is characterized by its rich soils, including Acrisols, Lixisols, Luvisols, and Planosols (Fig. 4). The Eburnean Plutonic Suite comprises k-feldspar-rich granitoids – mainly granite and monzonite from the 'Banso'/Bongo type, biotite granitoid, muscovite granite and minor granodiorite and local leucogranite (Kesse, 1985). The study area falls under the tropical climate regime and is characterized by bimodal precipitation with an average annual precipitation amount of 1088 mm–1197 mm (Dickson et al., 1988). The temperature in the area is high, averaging over 23.9 °C all year round. Relative humidity is high in the wet months and low in the dry months, with an average annual humidity of over 75%.

3.2. Data collection and analysis

A total of 165 water samples were taken both from both shallow wells (mainly hand-dug wells with depths <20 m) and from deep (mainly domestic boreholes with depths >30 m) aquifers from the crystalline bedrock and sedimentary rocks underlying the Brong Ahafo region in Ghana. Out of 165 samples, 84 samples were taken from the Birimian Supergroup and its intruded granitoid (crystalline basement), while81 were sourced from the Kwahu-Morago Group (42 samples) and the Obosom Group (39 samples), and these make up the sedimentary aquifers. The coordinate of each sampling location was noted using the Garmin Global Positioning System device with an error of ±2 m. The sample bottles were pre-cleaned in the laboratory by immersing them in a 5% nitric acid solution and rinsing them three times with distilled water to remove any possible contamination from the bottles. Flushing was carried out for between 10 and 15 min prior to sampling in order to remove the standing water from the borehole column.

The groundwater was then sampled after filtration with a 0.45 μm cellulose acetate membrane, and the bottles were rinsed with portions of the filtrate. The samples were immediately put in an ice chest under a temperature of less than 4 °C and later transported to the Environmental Chemistry Laboratory of the Water Research Institute in Accra, Ghana for analysis. The concentrations of chloride (Cl^-) and nitrate-nitrogen (NO_3^--N) were analyzed colorimetrically with a UV–visible spectrophotometer using the Spectroscan 60 DV model. Physical parameters,

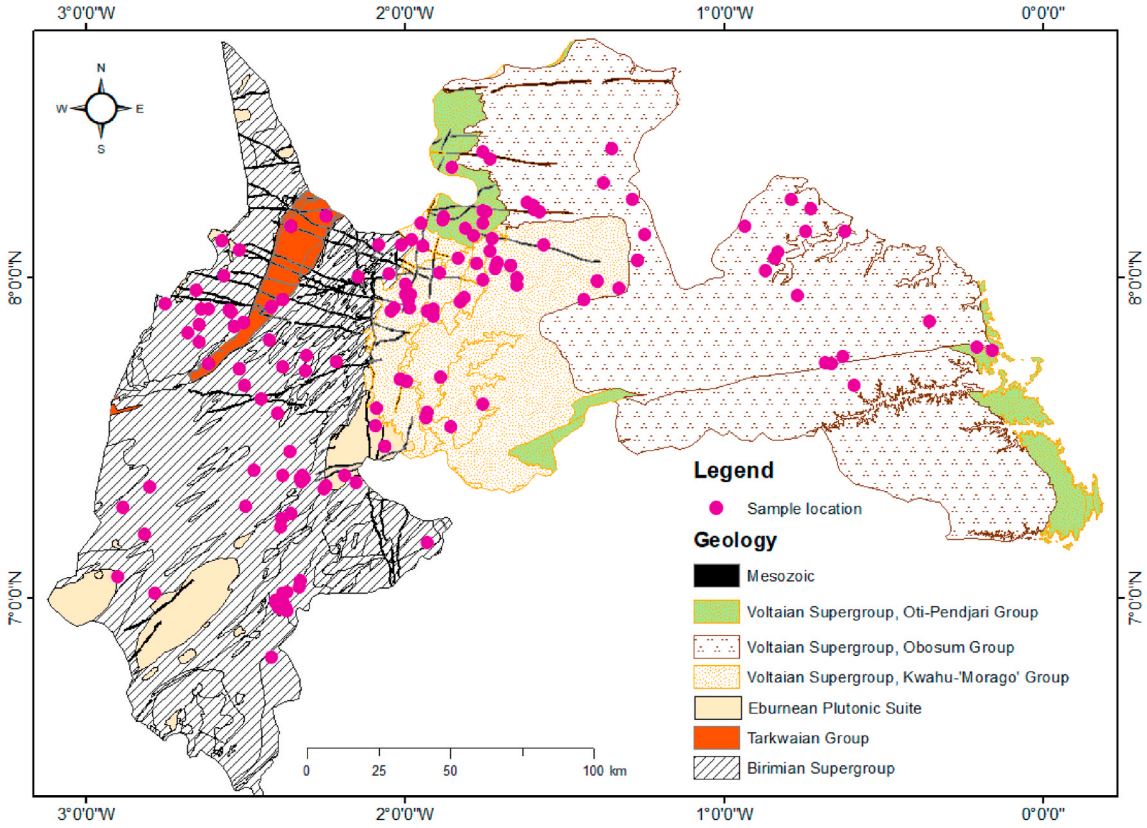


Fig. 1. A map of the study area showing the six geological terrains; Birimian Supergroup (greywackes with turbidite features, phyllites, slates, schists), Tarkwaian formation (clastic sediments-conglomerate, shales), Oti-Pendjari Group (mudstone, siltstone, micaceous and sporadic limestone), Obosom Group (Sandstone, fine to medium-grained, arkosic, lithic, siltstone and mudstone). Kwahu-Morago Group (sandstone, dune-bedded to cross-bedded, medium-grained, arkosic, with mudstone to-wards base, flaggy, quartzose), Eburnean (mainly granite and monzonite, biotite granitoid, muscovite granite and minor granodiorite and locally leucogranite).

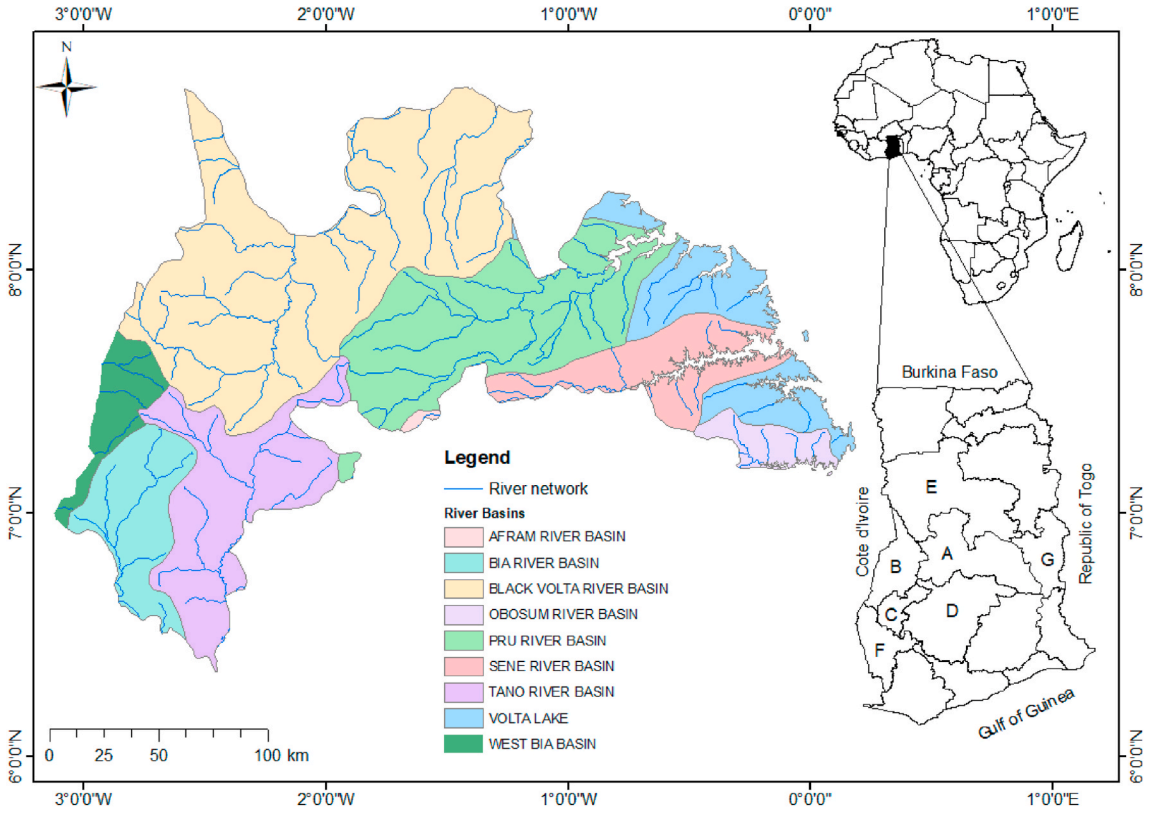


Fig. 2. The study area shows how the nine river basins are connected. The Africa map (top right corner) shows the geographical location of Ghana. Also, the Ghana map inserted (bottom right) shows the international and local boundaries. The Brong Ahafo region includes the regions; Bono east (A), Brong Ahafo (B), Ahafo (C) and Western North (F). Regional boundaries comprise the Ashanti region (D), Savannah (E), Western North (F) and Oti (G). The international boundary includes Cote D'Ivoire (to the west).

including pH, temperature, and electrical conductivity, were measured at each sampling location in the field.

3.3. Probability plot

The probability plot (PP) method is based on the assertion that the chemical composition of groundwater can contain one or more populations that may be from geogenic or anthropogenic sources (Panno et al., 2006; Sellerino et al., 2019). A straight-line PP indicates the sampled data is normally distributed, and an abrupt rupture or curvature of the lines reveals inflection points or threshold values between populations. Inflection points indicate the upper limit of the background and can be considered the threshold between natural and influenced concentration (Preziosi et al., 2014; Sellerino et al., 2019). Searching for breaks and inflection points is subjective and relies on the researcher's own experience (Reimann et al., 2005). When a single population is recognized in a data set, it is prudent to assume a few highest values as anomalous. Hence, the sum of the two standard deviations and the mean is considered the natural background (Sinclair, 1974). The PP plots were built with all the groundwater NO_3^-–N data set, and the plot was drawn using lognormal distribution.

3.4. Iterative outlier removal (IOR) method

An IOR is an iterative process to locate and eliminate "extreme" values in the sample data (Nakić et al., 2007). The technique is very useful when the observational data follows heavy tail distributions like the nitrate data used in this work. The following workflow (Fig. 5) outline how an observation point that is distant from other observations can be statistically reduced from the data set:

The technique assumes that the background concentration can be obtained by attaining a normal distribution. However, this method's effectiveness also largely depends on the sample size of the data set and can be realized only when the number of removed outliers is small relative to the total sample size (Lindley, 1957; Qian and Lyons, 2006; Kim et al., 2015). The IOR technique is suitable for estimating the threshold values as the upper limit of the background levels (Nakić et al., 2007) and has the added advantage of eliminating anomalies below the lower limit (Sakyi et al., 2018; Akoto and Adiyiah, 2007; Yidana et al., 2012; Banoeng-Yakubo et al., 2009).

3.5. Gaussian mixture model (GMM)

The GMM is an unsupervised statistical learning technique used for clustering and mining data (Sarkar et al., 2020; Adhikari et al., 2016). The statistical model considers that the observation data points are produced from finite Gaussian distributions with given parameters. The key significance of taking this mixture model is that it does not require any prior knowledge of the subpopulation to which a data point belongs. Therefore, the detection of the subpopulations can be learned automatically, making the learning process unsupervised. In addition, the model can estimate the relevant expected values and the uncertainties in each subpopulation, which is what we are interested in this work.

As the name suggests, the model assumes that each point in the data set is drawn from a multivariate Gaussian probability distribution, denoted as $N(\mu_\lambda, \Sigma_\lambda)$ and it is parameterized by the co-variance matrix Σ_λ and the mean vector μ_λ. The subscript 'λ' indicates that there are Λ finite components of Gaussian distributions that are $(\mu_\lambda, \Sigma_\lambda)|_{1\ll\Lambda}$ in the data set. This probabilistic approach ensures that soft assignments are introduced in each cluster, in contrast to the K-Means method, which uses the hard assignment scheme (Capó et al., 2018).

So each point Z_i is independently drawn from the Λ finite compo-

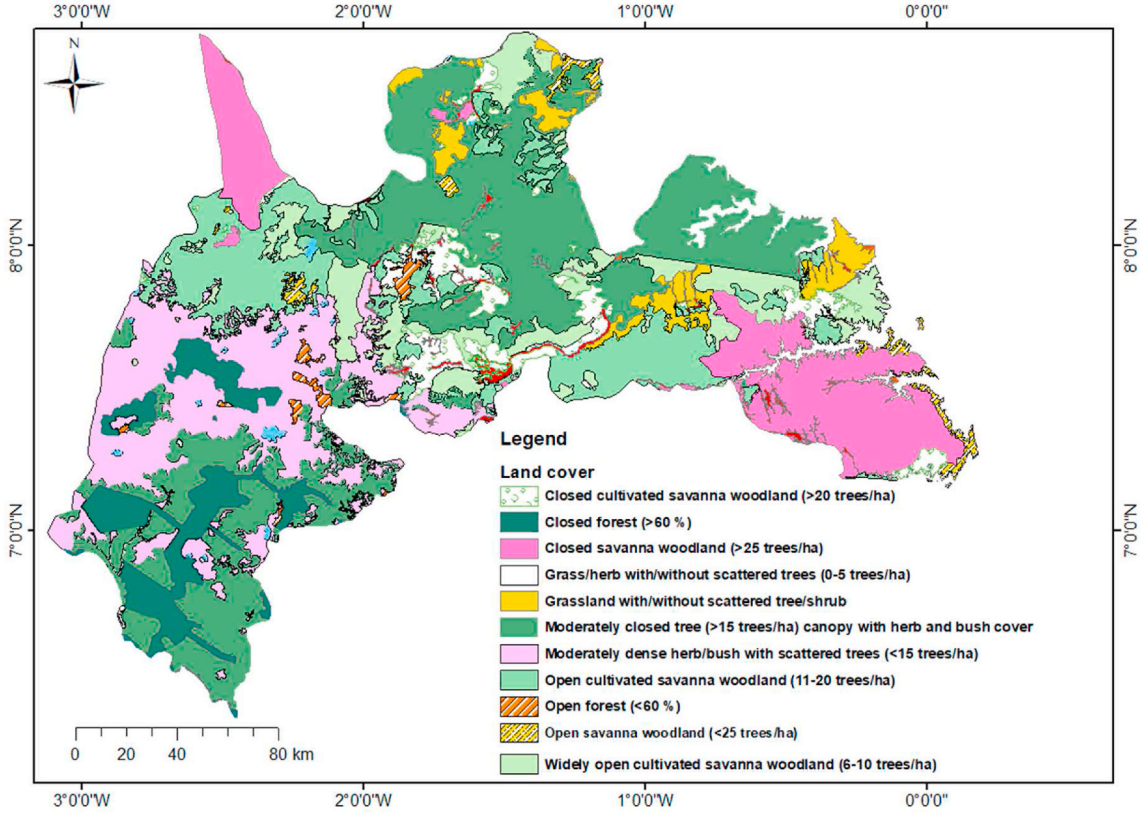


Fig. 3. Land cover of the study area (modified after FAO (2012)).

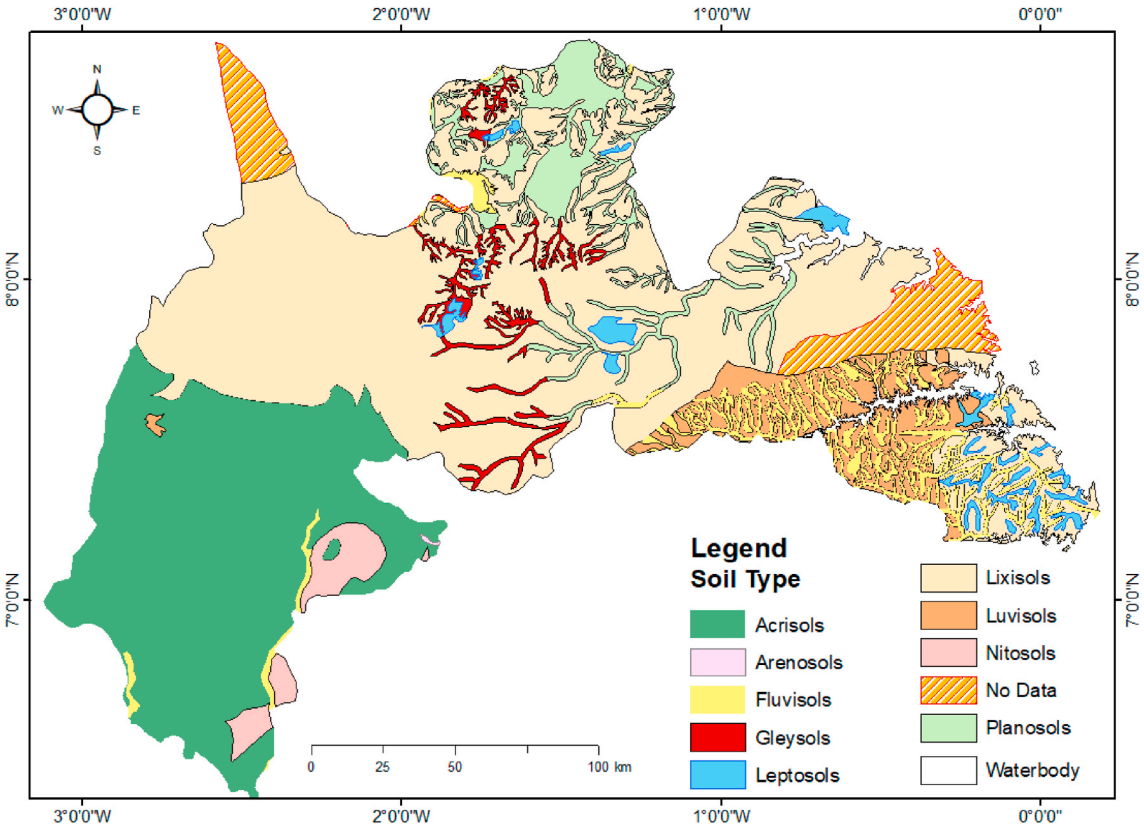


Fig. 4. The soil cover of the study area. The area is predominantly covered with the Acrisols to the southwest, Lixisols in the central portion and a mixture of Luvisols and Leptosols to the east.

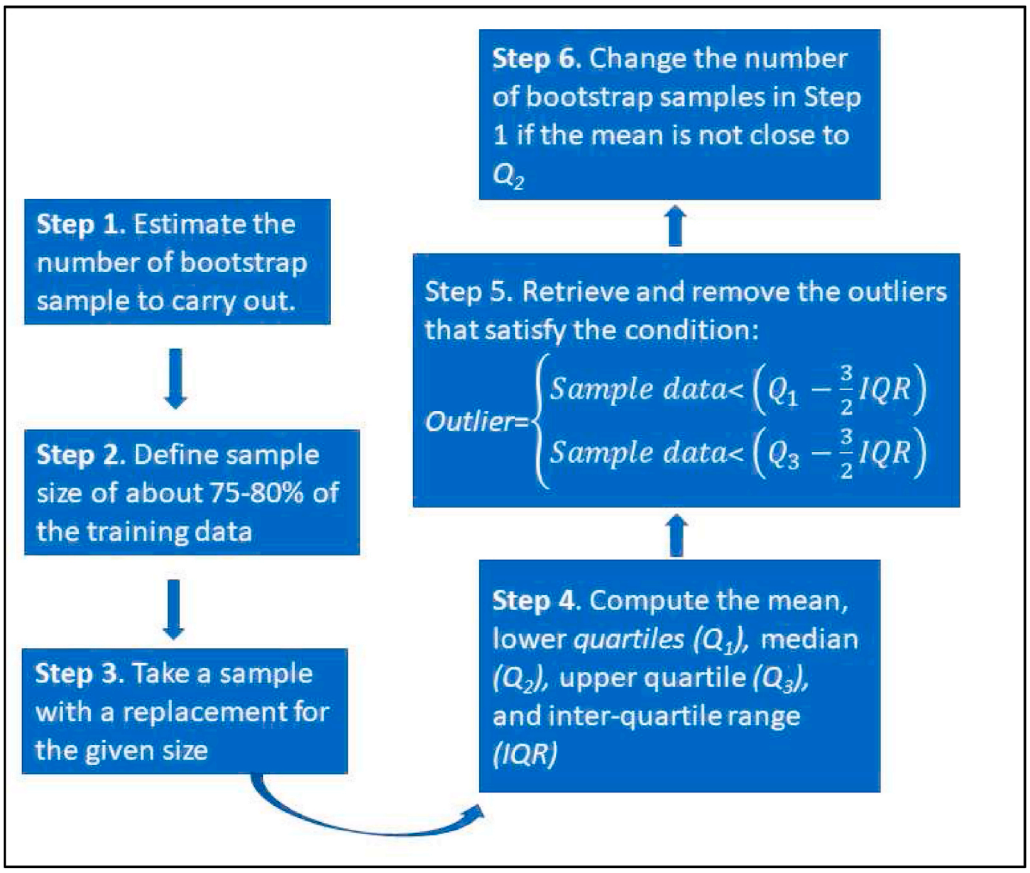


Fig. 5. Workflow showing the steps followed in executing the outlier removal techniques.

nents of Gaussians with the probability of,

$$\psi(\mu_\lambda, \Sigma_\lambda) = \frac{1}{\sqrt{|2\pi\Sigma_\lambda|}} \exp\left[-\frac{1}{2}(z_i - \mu_\lambda)^T \Sigma_\lambda^{-1} z_i - \mu_\lambda\right)\Big] \quad (1)$$

The maximization of the above probability distribution for all points assuming finite Gaussian distributions in the entire data set are obtained by the following steps with some initial Gaussian distributions:

i. Find the Gaussian probability that is most likely to generate each point from $(\mu_\lambda, \Sigma_\lambda)|_{1\ll\Lambda}$. Here, the similarity for each component in the data is measured by Equation (1).

ii. ii. Estimate the parameters $(\mu_\lambda, \Sigma_\lambda)$, for each $1 < \lambda < \Lambda$, using the data points for each cluster.

Note that the Maximum Likelihood Estimation (MLE) approach cannot be directly computed, since the GMM is nonlinear. Therefore, for this work, the Expectation-Maximization (EM) algorithm presented in Vu et al. (2019) was used to estimate the model parameters in Step ii above. Fig. 6 shows the clustering patterns of the NO_3^-–N positions. The three spatial segregations of the nitrate-nitrogen data set are clustered, representing peculiar geological terrain. The first component (violet sampling points) is captured within the western part of the study area, mainly representing the Birimian Supergroup. The second component

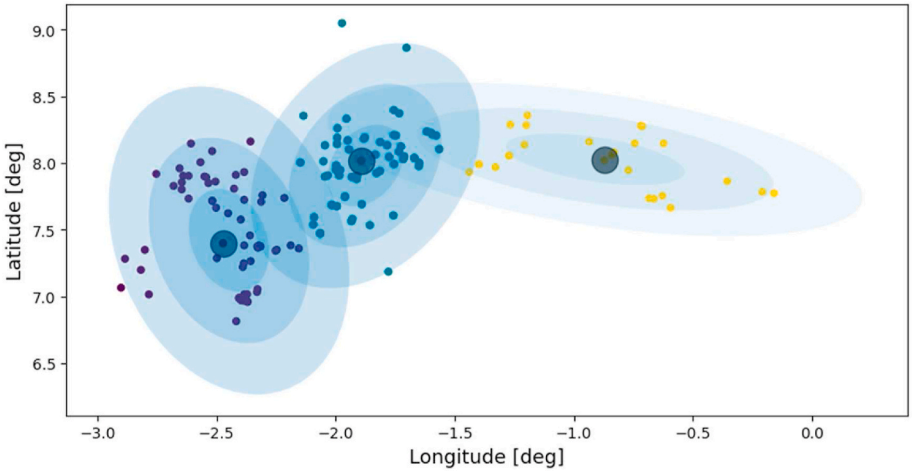


Fig. 6. A two-dimensional Gaussian mixture model displaying the spatial distributions of three components. The first component (violet points) was captured within the western part (Birimian Supergroup), the Second component (teal points) depicted at the central portion is Kwahu-Morago Group, and the third component (yellow points) represents the Eastern portion (Obosum Group). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(teal sampling points) is also identified in the central parts of the study area and represents the Kwahu-Morago Group of the Voltaian Supergroup. Finally, the third component is recorded in yellow sampling points in the eastern part, which mainly represents the Obosum Group. These clusters were used to estimate the NBL of the sample data and then compared to the NBL assessed by the NBL evaluated with the method in Section 3.3.

3.6. Estimation of threshold values (TV)

In TV estimation, the method requires the determination of reference values (REF), which is essential. These REFs are determined based on groundwater use for either drinking, irrigation, or industrial use. However, when the groundwater quality is unknown, drinking water quality standards are used as surrogates (Bulut et al., 2020). Since the groundwater in the study area is primarily used for human consumption, the WHO standard for drinking water (10 mg/L NO_3^-–N) was used as REF. The BRIDGE project proposed and adopted by Molinari et al. (2012) and Bulut et al. (2020) such that if NBL < REF, then TV = $(REF + NBL)/2$; else $TV = NBL$.

4. Results and discussion

4.1. NBL and sources of nitrate-nitrogen in groundwater

Table 1 shows the descriptive statistics of the NO_3^-–N and Cl^- found in the groundwater. The NO_3^-–N concentration is within a range of [0.01, 225] mg/L with an average value of 7.54 mg/L and a median of 1.76 mg/L. The chloride concentration also ranges from [0.037, 2978] mg/L with an average of 80 mg/L and a median of 28.8 mg/L. The standard deviations for both elements are relatively higher than the estimated mean values suggesting large variation among the data points. The dispersion of NO_3^-–N and Cl^- concentrations are attributed to natural dynamics leading to the percolation of chemical species into the subsurface. Some of these natural phenomena include periodic precipitation patterns, groundwater recharge rate, evapotranspiration process, etc. Other natural factors responsible for the spatial variations of NO_3^-–N concentration in groundwater include hydraulic conductivity, soil water holding capacity, rainfall intensity, depth of water table, aquifer media (Ahada and Suthar, 2018; Minet et al., 2017; Buvaneshwari et al., 2017; Nakagawa et al., 2017; Shalev et al., 2015). There is no petrophysical evidence of halites in the area, which could not account for the spread of chloride. We can infer that the mode of spread suggests multiple sources of chloride in the area. These sources may be due to sea aerosol spray, chlorine-containing fertilizers, and domestic and industrial waste effluent.

Fig. 7A exhibits a cumulative distribution of the nitrate-nitrogen concentration. The histogram shows that the nitrate-nitrogen concentration is normally distributed after the data has been log-transformed. The kernel density function (KDE) also exhibits a single major inflection point with minor, not legible inflection points. This minor modulation is attributed to anomalies in the data set, which can be discarded.

Table 1
Statistical summary of chemical parameter concentrations (NO_3^--N and Cl^-(mg/l)) of groundwater.

Statistical Parameters	NO_3^--N (mg/L)	Cl^-(mg/L)
Count	165	165
Mean	7.54	80
standard deviation	20.9	281
Minimum	0.01	0.037
Maximum	225	2978
First quartile	0.15	10.9
Median	1.76	28.8
Third quartile	5	52
P90	22.4	213.6

The probability plot (Fig. 7B) also confirms this assertion. The PP plot shows that most of the sample points are plotted on the theoretical normal distribution line (red) with a negligible inflection point. The NO_3^-–N concentration homogeneity means there is only one population suggesting a common nitrate source in the study area. It also suggests that the factors that control the formation and accumulation of NO_3^-–N in the area are weakly variable and homogeneous in the study area. In our case, extreme values influence the mean; therefore, using this means for simple descriptive statistics will result in an incorrect NBL.

In addition, the median statistic does not take into account the actual values of each observation and all possible information that is available in the data set. We therefore, evaluate the expected value of the NBL based on the results presented in Figures; 8 and 9, respectively.

The data was further subjected to the iterative outlier removal (IOR) technique to eliminate the anomalous concentration from anthropogenic sources (such as fertilizers). Fig. 8 shows successive iteration to eliminate anomalous values. In succeeding iterations, the mean NO_3^-–N concentration values exhibit a successive decrease of the mean gradient concentration. In the succeeding iteration from the box plot, the skewness also gradually decreases, leading to the symmetric distribution of the data set. The nitrate-nitrogen sub-population obtained in the 8th iteration step has the lowest positive skewness, so the concentration distribution becomes close to symmetrical. This symmetric distribution is further confirmed by the lack of outliers in this subpopulation and the lowest difference between the estimated mean and the median.

Finally, the iterative technique eliminates the disturbances originating from the local environment conditions at sampling sites. Hence the mean value obtained for the 8th iteration is closest to that of the true background population and, consequently, attributed to the NBL. As shown in Fig. 8, the background values of the NO_3^-–N is a range of values evaluated as the lowest value to the 90th percentile of the 8th iteration [0.001–3.9] mg/L with an expected value of 1.25 mg/L. A study by Akoto and Adiyiah (2007) recorded NO_3^-–N concentrations from [0.09, 0.99] mg/L with a mean value of 0.32 mg/L. In this work, there is a significant variation in the nitrate-nitrogen concentrations due to anthropogenic activities, such as applying nitrogen-rich organic and inorganic fertilizers used for crop production in the area.

Similarly, Yidana et al. (2012) found the NO_3^-–N concentration in some Voltain and Birimian aquifers in the range of [0.01, 27.1] mg/L with a mean and median value of 1.38 mg/L and 0.25 mg/L, respectively. This buttresses the earlier suggestion that there is a significant increase in nitrate-nitrogen concentration derived from anthropogenic activities. To obtain the threshold above which the concentrations are anomalous or the upper bound of the NBL, the 90th of the normalized data set was considered (3.9 mg/L). When translated into the TV of the area for regulation of the aquifer system, this value gives 6.95 mg/L of NO_3^-–N (30.58 mg/L for NO_3^-).

The violin plots (Fig. 9) describe the respective distributions of the NO_3^-–N concentration for the geo-location clustering presented in Fig. 6, using the Gaussian mixture algorithm. The white dots in the violin plots represent the median. The thick grey bar in the centre represents the interquartile range, and the thin grey line represents the rest of the distribution. On each side of the grey line is a KDE to show the distribution shape of the data. Here, the shape of the violin plots has wider sections indicating that the NO_3^-–N concentration for each cluster is homogeneous about the median. The median of the NO_3^-–N concentration for each lithology underlying the study area is the same.

From the PP distribution (Fig. 7B), we observed a normal distribution of the nitrate-nitrogen data in the study area. To further identify the influence of the geology on the nitrate distribution through infiltration to the water table, geostatistical modelling using the Co-Kriging interpolation model (Fig. 10) and a 2D spatial Gaussian mixture model (Fig. 6) were employed to learn the spatial nitrate-nitrogen distribution.

The Ordinary Co-Kriging method was selected for spatial modelling because of the poor distribution of the data points across the area. The robustness of the Co-Kriging method made it possible to estimate the

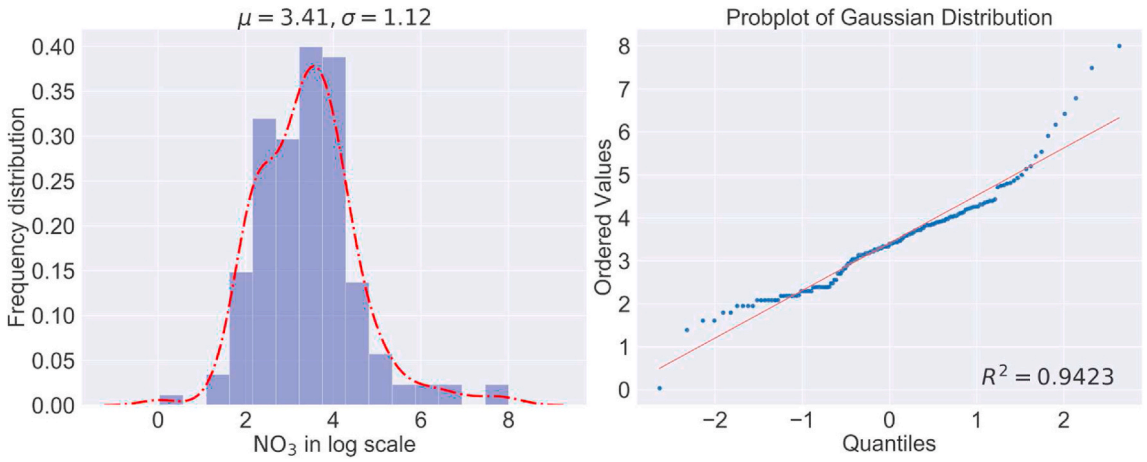


Fig. 7. A cumulative distribution of the nitrate-nitrogen concentration shown on a logarithmic scale. The left panel (A) shows a histogram representation of the nitrate-nitrogen data superimposed with a KDE curve. The right panel (B) is a representation of the data plotted against a theoretical normal distribution to assess whether the data points (in blue) approximate to a straight line (in red). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

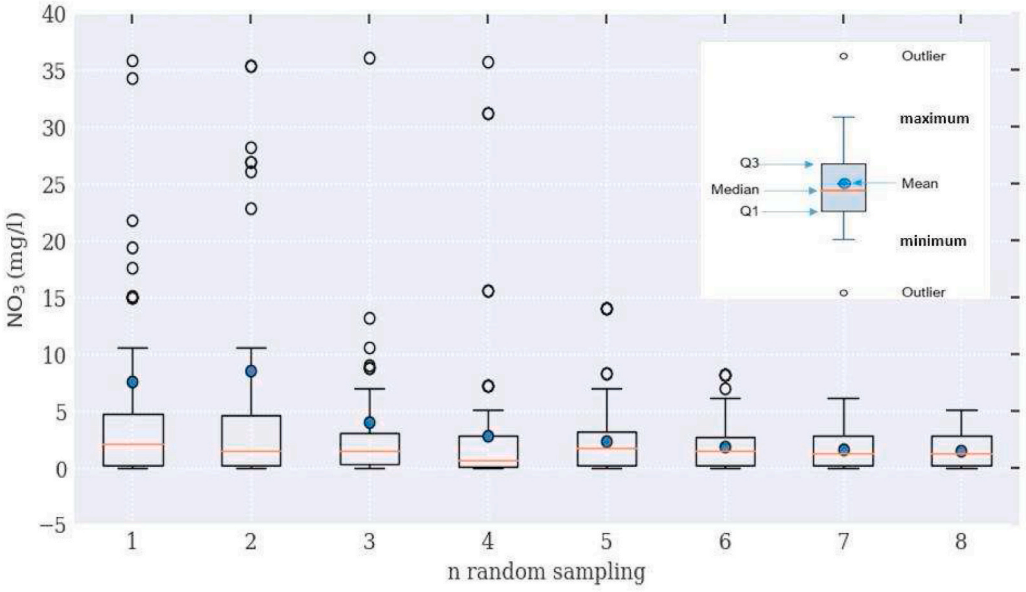


Fig. 8. Comparative boxplot of NO_3^-–N concentrations in successive iterations showing the variations in mean and median concentration in successive iterations. The NBL distribution is considered the distribution with the lowest difference between the calculated mean and the median and confirmed by the one with a lack of outliers (8th).

uncertainty error between the predicted and the observed concentrations. The semivariogram with the spherical model was used to fit the model for the predicted map. The model was subjected to a cross-validation in which the measured and the predicted values estimate the possible permissible uncertainty in the model. Because of the lack of evenly distributed spatial data, the semivariogram approach was chosen for spatial analysis. A simple inverse distance weighting (IDW) will not deliver the best quality results if the data is not fairly distributed spatially (Sutton et al., 2009), which is the case in this study. An optimal semivariogram model for estimating the spatial nitrate-nitrogen model was achieved through a rigorous cross-validation exercise using the kriging interpolation. A good fit of the experimental and the fitted semivariogram models was achieved by adjusting the variogram parameters until an optimal model was realized (Fig. 10). The results of the semivariogram gave a range of 3.01, a sill of 2.86 and a nugget of 0.443 (Fig. 10). The observed nugget value (0.443) suggests that the nitrate-nitrogen varies spatially at smaller and larger distances. Various factors may contribute to the observed large nugget and include geological and sampling (Simmonds, 2009). The geological component can arise because of the heterogeneous fractures that control groundwater movement in the study area. Groundwaters underlying areas with high fracturing intensities will be characterized by a relatively high nitrate-nitrogen load. The contribution of sampling to the nugget effect can be introduced at the sampling point, and during the laboratory analysis. A cursory examination of the spherical semivariogram model reveals an oscillation around the sill and reflects the periodicity in the data. The relatively low Residual Sum of Squares (RSS) recorded in this study (RSS = 0.549) shows a good model fit that can be used to perform nitrate-nitrogen spatial map analysis.

The spatially predicted map of this study is acceptable with a mean error of −0.0059 and a Root Mean Square Standardized Error (RMSSE) of 0.922. The mean error indicates a good agreement between the predicted and the measured values and can explain the spatial variability in the nitrate-nitrogen distribution over the area. However, the negative

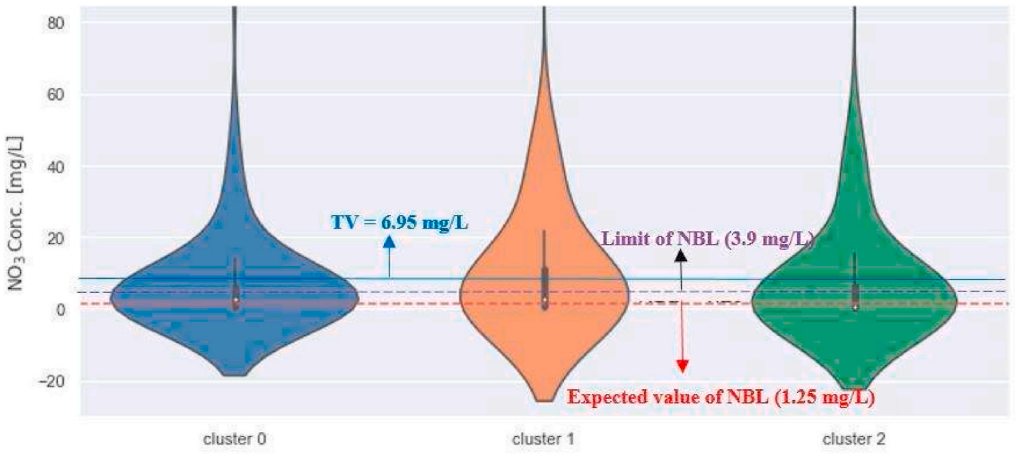


Fig. 9. Violin plot revealing the morphology in the nitrate-nitrogen data set, components based on spatial distribution representing different geological terrain. Cluster 0 - represent the Birimian supergroup largely, cluster 1 - represent entirely Kwahu-Morago Group and cluster 2 - represent mainly the Obosum Group.

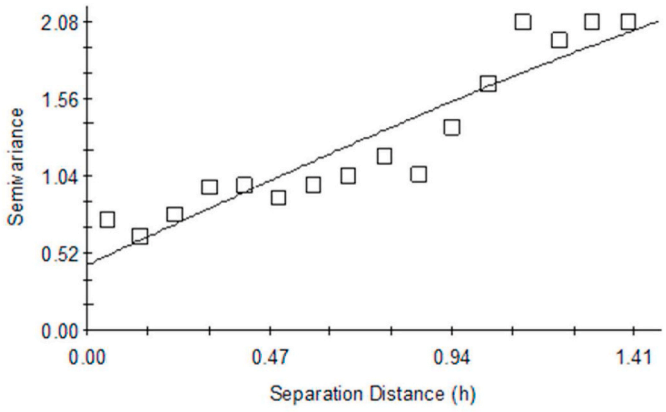


Fig. 10. The final semivariogram fitted with the spherical model with model statistics; Co = 0.443, Co + C = 2.869, Ao = 3.01, r2 = 0.86 and RSS = 0.549. Co is the Nugget Effect, Co + C is the Sill (also known to describe the total variance), Ao is the Range, RSS is the Residual Sum of Square.

indicates that the predicted value is 0.0059 less than the actual value, and also indicates that the under- and over-predictions cancel each other out, making the model unbiased. According to Marc et al. (2017), an RMSSE less than 1 indicates an overestimation of the variability and an RMSSE greater than one indicates an underestimation of the variability. In this study, an RMSSE value of 0.922 indicates a strong correlation between the predicted and the measured. Fig. 11 shows how NO_3^-–N concentration is spatially distributed. The associated probability map with the spatial threshold concentration acceptable for drinking water (WHO, 2017) is also shown in Fig. 12. According to Fig. 11, we observed that a relatively low nitrate-nitrogen variations characterize the Voltaian Supergroup samples, for instance, areas around Atebubu, Kwame Danso, Yeji and Nkoranza (Fig. 11). The communities mentioned above, and their environs are agricultural communities where various types of nitrate-rich fertilizers are applied. Because of this, nitrate-nitrogen concentrations in groundwater in these areas is expected to be significantly higher than what was measured. However, the thick clay overburden material which serves as an impervious layer restricted the movement of the nitrate-nitrogen into the underlying aquifer. It is again observed that samples sourced from the Birimian (areas around Nsawkaw, Berekum, Goaso, Tachiman and Dormaa Ahenkro) showed relatively high NO_3^-–N concentration compared to the Voltaian Supergroup (Fig. 11). Unlike the Voltaian Supergroup, the situation is different in the Birimian Supergroup characterized by fractured crystalline aquifers. Although the Birimian rocks are inherently impermeable, groundwater movement is controlled by secondary porosity development due to shearing, jointing, faulting, and fracturing. The high hydraulic conductivity can best explain the relatively high concentration amongst the fractured aquifer system in the Birimian, contributing significantly to enhancing the movement of NO_3^-–N contaminated water from the surface to the aquifer. Fig. 12 shows the probability spatial distribution of threshold nitrate-nitrogen concentration <10 mg/L acceptable for drinking water by the WHO. It is observed that a significant portion of the region falls within the acceptable threshold of nitrate-nitrogen <10 mg/L for drinking water.

A few areas around Nsawkaw, Tachiman, Berekum, Kintampo, Dormaa Ahenkro and Goaso show relatively high nitrate-nitrogen variation above the WHO acceptable limit. The areas mentioned above are highly populated areas where several anthropogenic activities such as indiscriminate waste discharge into rivers, inappropriate agricultural practices (e.g., excessive fertilizer application) and unregulated human waste disposal may have caused leaching nitrate into the aquifer system. Fig. 13 shows a relation between NO_3^-–N/ Cl^- mass ratio and Cl^-. The NO_3^-–N/ Cl^- technique can be used to differentiate the effect of the denitrification process and normalize evapotranspiration. It also helps to discriminate NO_3^-–N from atmospheric precipitation and other sources (Jiang et al., 2016; Gates et al., 2008; Anornu et al., 2017). In general, low NO_3^-–N/Cl^-ratio against high Cl^- are attributed to domestics (municipal) and organic waste, while agricultural inputs are typically associated with low Cl^- and high NO_3^-–N/Cl^- ratio (Liu et al., 2006; Anornu et al., 2017). From Fig. 13, samples that show considerably higher Cl^- concentrations with lower NO_3^-–N/Cl^- ratio suggest domestic effluent and organic matter as the probable sources of nitrate-nitrogen.

Likewise, the denitrification process plays a significant role in nitrate removal in the area. In the water samples, the NO_3^-–N/Cl^- ratio varies from 0.01 to 58.4, with a mean value of 0.78. Due to the absence of reference value of NO_3^-–N/Cl^- determined in the sub-region, the reference value in the range of 0.05 and 0.22 without human activities reported in the North-western region of China (Ma et al., 2009; Li, 2014) was adopted and used by (Anornu et al., 2017; Egbi et al., 2020) who did similar work in Ghana. The fiducial triangle in red shows the variation in NO_3^-–N/Cl^- ratio. It can be observed that a significant number of the samples fell below this triangle. The results showed that about 76% of the samples fell below this ratio which is considered the reference value of NO_3^-–N/Cl^- for this study. Samples found to have been affected by agricultural activities in the red triangle make up approximately 24% of the data set. The IOR techniques removed these samples from the data set, which aided in the NBL determination. This implies that a larger

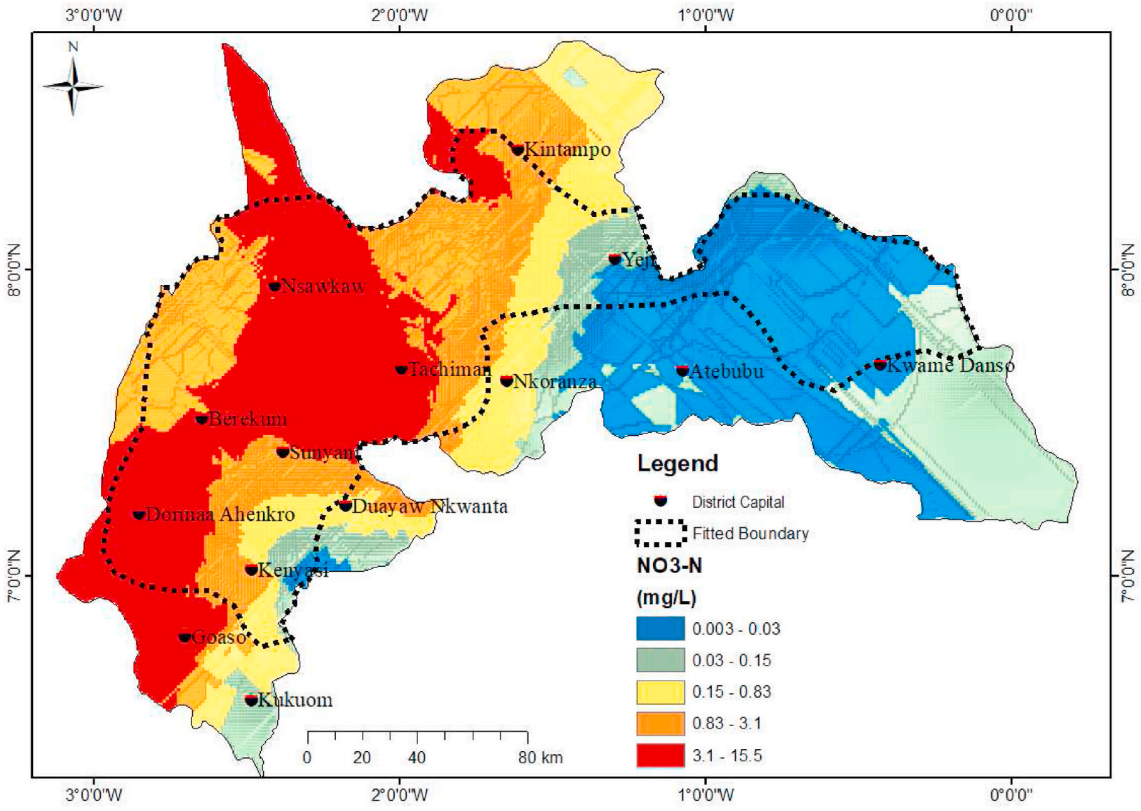


Fig. 11. Predicted spatial map showing NO_3^-–N concentration across the study area using the geostatistical modelling (co-Kriging). The red colour indicates relatively high NO_3^-–N concentration in the area; conversely, large areas are characterized by low concentration (blue and pale green) and very few (yellow and orange) characterized as intermediate concentration. The predicted model is robust within the fitted boundary. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

population of the groundwater resource in the study area is mainly controlled by a natural reduction process (denitrification) due to microbial activities within the subsurface before recharging. The above-mentioned NO_3^-–N/Cl^- accordingly supports that the water sources in the area are less polluted by human activities.

4.2. Implication for groundwater resource management

According to the results, the natural background value of the NO_3^-–N concentration in the groundwater of the study area is estimated to be in a range of [0.001, 3.9] mg/L, which is well below the acceptable limit of 10 mg/L of NO_3^-–N in drinking water from the WHO (2017). It is observed that 85% of the samples are within the acceptable range of NO_3^-–N in drinking water, suggesting that the groundwater in this area is suitable for drinking. However, there are isolated cases in which the nitrate-nitrogen concentrations deviated significantly from the estimated natural background and the acceptable limit values of the WHO (2017). The range of deviation varies between 13 and 48 mg/L and makes up 15% of the samples. Anthropogenic activities such as agricultural practices involving the application of nitrate-rich fertilizers, the discharge of household and untreated industrial wastewater, and indiscriminate waste disposal are the potential sources of the high levels of NO_3^-–N in groundwater in the region. In Fig. 11, we can observe that locations around Nsawkaw, Tachiman, Berekum, Kintampo, Dormaa Ahenkro and Goaso show significantly high nitrate concentrations compared to the background value. This level of nitrate-nitrogen concentrations is unacceptable for drinking water suggested by WHO (2017). The above mentioned towns in the study area are noticeably agricultural areas in which fertilizers and animal manure are constantly being applied in order to increase soil fertility.

The chemical composition of the commonly used fertilizers is enriched with NO_3^-–N and can undoubtedly get into the groundwater system. It is proposed that fertilizer applications be regulated to the normal and routine sampling and analysis of water from the wells within these agricultural communities in order to mitigate future NO_3^-–N pollution of groundwater. Domestic influence, indiscriminate waste disposal and untreated sewage contribute significantly to the NO_3^-–N and chloride levels in the groundwater. Recent reports of high bacteria load counts in boreholes located in some parts of the region have been attributed to septic and contaminant transport through the unsaturated zone (Tekpor et al., 2017). For example, Escherichia coli and Enterococci faecalis were detected in drinking water collected from some hand-dug wells in the Sunyani metropolis in the study area and were attributed to pit latrine Gyasi et al. (2017). A significant correlation regarding E. coli contamination in the water sample was found between the hand-dug well (shallow aquifer) and the nearest pit latrine (Gyasi et al., 2017). Disposal of liquid and solid waste is a challenge in the study area, and the easiest way to dispose of waste is through river channels and dugouts. Therefore, in some of these locations, the groundwater becomes vulnerable to pollution as these chemical components of the waste can enter the aquifer system. This study has helped to identify the possible sources of nitrate concentrations in groundwater and to provide the natural background value in order to ensure sustainable groundwater resource management in the examined region. This study will provide the background groundwater nitrate-nitrogen concentration in groundwater for current and future studies of nitrate in this area and similar environments.

5. Conclusion

The nitrate-nitrogen in the groundwater of some crystalline and sedimentary aquifers in the Brong Ahafo region of Ghana was assessed to

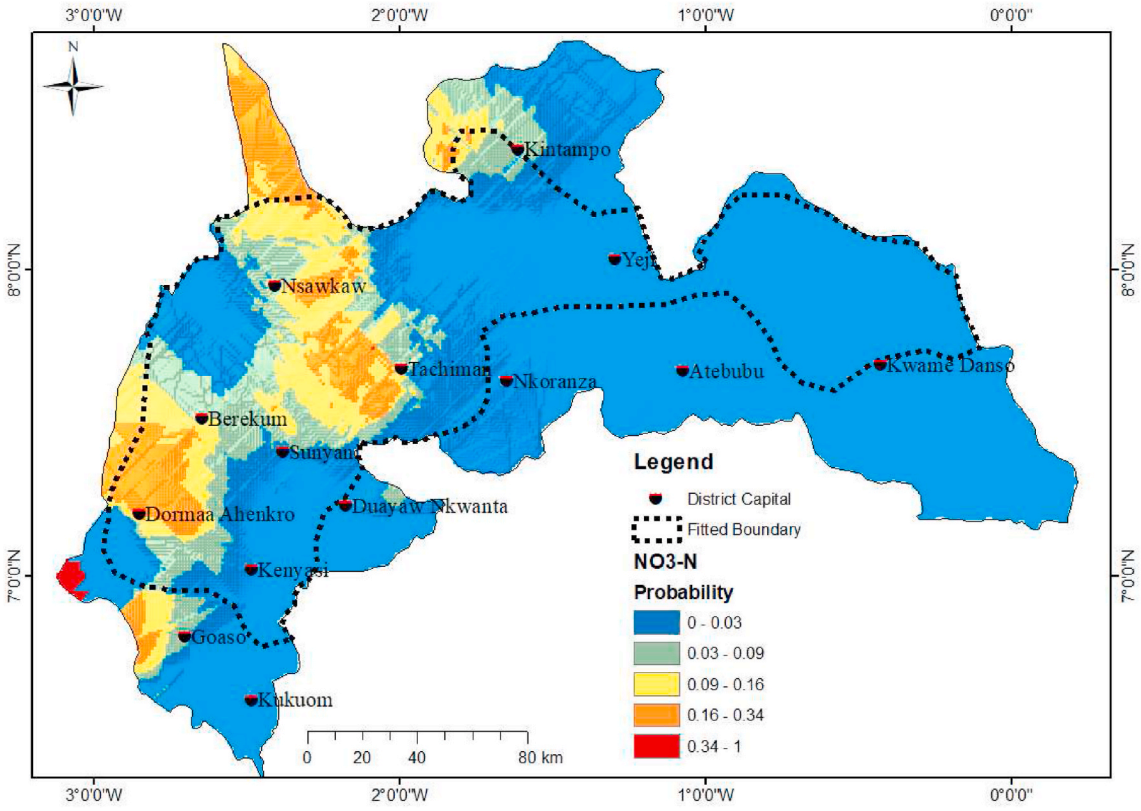


Fig. 12. Spatial predicted probability map showing the acceptable threshold (<10 mg/L) of NO_3^-–N concentration for drinking water in the area. The predicted map was generated using the co-Kriging geostatistical interpolation method. Light blue indicates NO_3^-–N < 10 mg/L and yellow, brown, and red indicate NO_3^-–N > 10 mg/L. The model is robust within the fitted boundary, where most of the points are concentrated. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

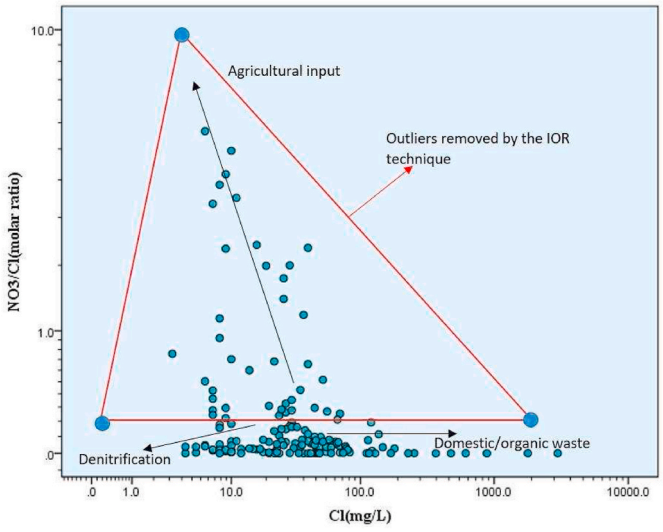


Fig. 13. Relation between NO_3^-–N/Cl^- mass ratios and the Cl^- in the Brong Ahafo data sets; the samples showing significant low ratios are mainly observed in low nitrate levels below the proposed NBL, which indicate that the natural attenuation process may be controlling NO_3^-–N in the study area. Samples found in the red triangle were samples removed by the IOR technique indicating agricultural input. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

determine the NBL and its possible sources. The estimated NBL of NO_3^-–N in the area's groundwater of the area is between 0.001 mg/L and 3.9 mg/L with an expected mean of 1.25 mg/L and a calculated TV of 6.95 mg/L (Figs. 8 and 9). This indicates that the background concentration of NO_3^-–N in the groundwater is well below the drinking limit acceptable by the WHO. In a few individual cases (e.g., areas around Nsawkaw, Tachiman, Berekum, Kintampo, Dormaa Ahenkro and Goaso), however, NO_3^-–N N values were measured above the level acceptable for drinking by the WHO (Fig. 11). The relationship diagram of NO_3^-–N/Cl^- (Fig. 13) highlights agricultural, domestic/organic waste and denitrification as the potential sources of nitrate-nitrogen in the groundwater in the study area. A major limitation of this study is the lack of time series nitrate data to investigate the potential temporal variations in the nitrate concentrations in groundwater in the study area. However, the results provide the background NO_3^-–N concentration in groundwater and potential sources for current and future studies of nitrate-nitrogen pollution studies in this area and similar environments. This study represents the preliminary assessment of NO_3^-–N in the region and serves as a benchmark for further studies. We recommend using the isotopic detection (^{15}N-$NO_3^-$and ^{18}O-NO_3^-) in future studies through comprehensive and sustainable regional monitoring of the aquifer system. The seasonal variation of NO_3^-–N concentration should also be considered in future work in order to understand the dynamics of the seasonal variability of the NO_3^-–N in the groundwater for the study area.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors wish to thank all staff at the Water Research Institute of Ghana, who made data collection and laboratory analysis possible.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gsd.2021.100696.

Author contributions

EM and GYA conceived the experiments which were carried out by EM, YSAL and FS contributed to the site selection and provided logistical support. T.A-N worked on the statistical models and was also involved in writing the Python codes for the analysis. GYA, T.A-N, and EM also handled the data interpretation and co-wrote the manuscript. Finally, YSAL and FS proofread the entire manuscript.

References

Adhikari, P.R., Vavpetič, A., Kralj, J., 2016. Explaining mixture models through semantic pattern mining and banded matrix visualization. Mach. Learn. 105, 3–39. https://doi.org/10.1007/s10994-016-5550-3.

Ahada, C.P., Suthar, S., 2018. Groundwater nitrate contamination and associated human health risk assessment in southern districts of Punjab, India. Environ. Sci. Pollut. Control Ser. 25, 25336–25347. https://doi.org/10.1007/s11356-018-2581-2.

Akoto, O., Adiyiah, J., 2007. Chemical analysis of drinking water from some communities in the Brong Ahafo region. Int. J. Environ. Sci. Tech 4, 211–214.

Anornu, G., Gibrilla, A., Adomako, D., 2017. Tracking nitrate sources in groundwater and associated health risk for rural communities in the white Volta River Basin of Ghana using isotopic approach (X15n, X18ono3 and 3h). Sci. Total Environ. 603, 687–698. https://doi.org/10.1016/j.scitotenv.2017.01.219.

Armah, F.A., Quansah, R., Luginaah, I., 2014. A systematic review of heavy metals of anthropogenic origin in environmental media and biota in the context of gold mining in Ghana. International Scholarly Research Notices. https://doi.org/10.1155/2014/252148.

Banoeng-Yakubo, B., Yidana, S., Ajayi, J., Loh, Y., Asiedu, D., 2010. Hydrogeology and Groundwater Resources of Ghana: a Review of the Hydrogeology and Hydrochemistry of Ghana, Potable Water and Sanitation.

Banoeng-Yakubo, B., Yidana, S.M., Anku, Y., Akabzaa, T., Asiedu, D., 2009. Water quality characterization in some Birimian aquifers of the Birim Basin, Ghana. KSCE Journal of Civil Engineering 13, 179–187. https://doi.org/10.1007/s12205-009-0179-4.

Bempah, C.K., Ewusi, A., 2016. Heavy metals contamination and human health risk assessment around Obuasi gold mine in Ghana. Environ. Monit. Assess. 188 (5), 261. https://doi.org/10.1007/s10661-016-5241-3.

Bordeleau, G., Savard, M.M., Martel, R., Ampleman, G., Thiboutot, S., 2008. Determination of the origin of groundwater nitrate at an air weapons range using the dual isotope approach. J. Contam. Hydrol. 98, 97–105. https://doi.org/10.1016/j.jconhyd.2008.03.004.

Bulut, O.F., Duru, B., Çakmak, Ö., Günhan, Ö., Dilek, F.B., Yetis, U., 2020. Determination of groundwater threshold values: a methodological approach. J. Clean. Prod. 253, 120001. https://doi.org/10.1016/j.jclepro.2020.120001.

Burkart, M.R., Stoner, J.D., 2008. Nitrogen in groundwater associated with agricultural systems. In: Nitrogen in the Environment. Elsevier, pp. 177–202. https://doi.org/10.1016/B978-0-12-374347-3.00007-X.

Buvaneshwari, S., Riotte, J., Sekhar, M., Kumar, M.M., Sharma, A.K., Duprey, J.L., Audry, S., Giriraja, P., Praveenkumarreddy, Y., Moger, H., 2017. Groundwater resource vulnerability and spatial variability of nitrate contamination: insights from high density tube well monitoring in a hard rock aquifer. Sci. Total Environ. 579, 838–847. https://doi.org/10.1016/j.scitotenv.2016.11.017.

Capó, M., Pérez, A., Lozano, J.A., 2018. An Efficient K -means Clustering Algorithm for Massive Data arXiv e-prints, arXiv:1801.02949arXiv:1801.02949.

Carrier, M.A., Lefebvre, R., Racicot, J., Asare, E.B., 2008. Northern Ghana Hydrogeological Assessment Project.

Chandran, K., Smets, B.F., 2000. Single step nitrification models erroneously describe batch ammonia oxidation profiles when nitrite oxidation becomes rate limiting. Biotechnol. Bioeng. 68, 396–406.

Conti, K., Velis, M., Antoniou, A., Nijsten, G.J., 2016. Groundwater in the Context of the Sustainable Development Goals: Fundamental Policy Considerations. International Groundwater Resources Assessment Centre (IGRAC), Delft.

Dalla Libera, N., Fabbri, P., Mason, L., Piccinini, L., Pola, M., 2017. Geo-statistics as a tool to improve the natural background level definition: an application in groundwater. Sci. Total Environ. 598, 330–340. https://doi.org/10.1016/j.scitotenv.2017.04.018.

Dapaah-Siakwan, S., Gyau-Boakye, P., 2000. Hydrogeologic framework and borehole yields in Ghana. Hydrogeol. J. 8, 405–416. https://doi.org/10.1007/PL00010976.

De Caro, M., Crosta, G.B., Frattini, P., 2017. Hydrogeochemical characterization and natural background levels in urbanized areas: milan metropolitanarea (northern Italy). J. Hydrol. 547, 455–473. https://doi.org/10.1016/j.jhydrol.2017.02.025.

Dickson, K.B., Benneh, G., Essah, R., 1988. A New Geography of Ghana, ume 34. Longman London.

Edmunds, W., 2008. Groundwater baseline quality. In: en "natural groundwater quality" Edmunds y shand. blackwell publishing.

Egbi, C.D., Anornu, G.K., Ganyaglo, S.Y., Appiah-Adjei, E.K., Li, S.L., Dampare, S.B., 2020. Nitrate contamination of groundwater in the lower volta river basin of Ghana: sources and related human health risks. Ecotoxicol. Environ. Saf. 191, 110227. https://doi.org/10.1016/j.ecoenv.2020.110227.

Gates, J.B., Böhlke, J.K., Edmunds, W.M., 2008. Ecohydrological factors affecting nitrate concentrations in a phreatic desert aquifer in northwestern China. Environ. Sci. Technol. 42, 3531–3537. https://doi.org/10.1021/es702478d.

Geranian, H., Mokhtari, A.R., Cohen, D.R., 2019. Bivariate probability plots: a method for delineating different populations in soil geochemical data. Sci. Total Environ. 671, 1047–1055. https://doi.org/10.1016/j.scitotenv.2019.03.295.

Ghana, G., 2010. Population and Housing Census: National Analytical Report. Accra-Ghana: Ghana Statistical Service.

Guo, A.Z., Bartram, J.K., 2019. Predictors of water quality in rural health- care facilities in 14 low-and middle-income countries. J. Clean. Prod. 237, 117836. https://doi.org/10.1016/j.jclepro.2019.117836.

Gyasi, S., Abdul-Rashid, H., Boakye, T., Abugre, R., Korda, H., 2017. Microbial contamination of hand-dug wells and pit latrines in Fiapre in Sunyani, Ghana. Journal of Energy and Natural Resource Management (JENRM) 4. https://doi.org/10.26796/jenrm.v4i3.107.

Jarvis, W.T., 2012. Integrating groundwater boundary matters into catchment management. In: The Dilemma of Boundaries. Springer, pp. 161–176.

Jiang, W., Wang, G., Sheng, Y., Zhao, D., Liu, C., Guo, Y., 2016. Enrichment and sources of nitrogen in groundwater in the turpan-hami area, northwestern China. Exposure and Health 8, 389–400. https://doi.org/10.1007/s12403-016-0209-7.

Kesse, G., 1985. The Mineral and Rock Resources of Ghana Aa Balkema (Rotterdam).

Kim, K.H., Yun, S.T., Kim, H.K., Kim, J.W., 2015. Determination of natural backgrounds and thresholds of nitrate in south Korean groundwater using model-based statistical approaches. J. Geochem. Explor. 148, 196–205. https://doi.org/10.1016/j.gexplo.2014.10.001.

Leube, A., Hirdes, W., Mauer, R., Kesse, G.O., 1990. The early Proterozoic Birimian supergroup of Ghana and some aspects of its associated gold mineralization. Precambrian Res. 46, 139–165. https://doi.org/10.1016/0301-9268(90)90070-7.

Li, P., 2014. Research on Groundwater Environment under Human Interferences: a Case Study from Weining Plain, Northwest China. Changan University (Doctoral Thesis), Changan.

Lindley, D.V., 1957. A statistical paradox. Biometrika 44, 187–192.

Liu, C.Q., Li, S.L., Lang, Y.C., Xiao, H.Y., 2006. Using X15n-and X18o-values to identify nitrate sources in karst ground water, guiyang, southwest China. Environ. Sci. Technol. 40, 6928–6933. https://doi.org/10.1021/es0610129.

Llamas, M.R., Martínez-Santos, P., 2005. Intensive Groundwater Use: Silent Revolution and Potential Source of Social Conflicts.

Ma, J., Ding, Z., Wei, G., Zhao, H., Huang, T., 2009. Sources of water pollution and evolution of water quality in the wuwei basin of shiyang river, northwest China. J. Environ. Manag. 90, 1168–1177. https://doi.org/10.1016/j.jenvman.2008.05.007.

Mainoo, P.A., Manu, E., Yidana, S.M., Agyekum, W.A., Stigter, T., Duah, A.A., Preko, K., 2019. Application of 2d-electrical resistivity tomography in delineating groundwater potential zones: case study from the Voltaian super group of Ghana. J. Afr. Earth Sci. 160, 103618. https://doi.org/10.1016/j.jafrearsci.2019.103618.

Minet, E.P., Goodhue, R., Meier-Augenstein, W., Kalin, R., Fenton, O., Richards, K., Coxon, C.E., 2017. Combining stable isotopes with contamination indicators: a method for improved investigation of nitrate sources and dynamics in aquifers with mixed nitrogen inputs. Water Res. 124, 85–96. https://doi.org/10.1016/j.watres.2017.07.041.

Molinari, A., Guadagnini, L., Marcaccio, M., Guadagnini, A., 2012. Natural background levels and threshold values of chemical species in three large-scale groundwater bodies in northern Italy. Sci. Total Environ. 425, 9–19. https://doi.org/10.1016/j.scitotenv.2012.03.015.

Müller, D., Blum, A., Hart, A., Hookey, J., Kunkel, R., Scheidleder, A., Tomlin, C., Wendland, F., 2006. D18: Final Proposal for a Methodology to Set up Groundwater Threshold Values in Europe. BRIDGE project, Background Criteria for the Identification of Groundwater Thresholds, 6th Framework Programme Contract 6538.

Murgulet, D., Tick, G.R., 2009. Assessing the extent and sources of nitrate contamination in the aquifer system of southern Baldwin County, Alabama. Environ. Geol. 58, 1051–1065. https://doi.org/10.1007/s00254-008-1585-5.

Nakagawa, K., Amano, H., Takao, Y., Hosono, T., Berndtsson, R., 2017. On the use of coprostanol to identify source of nitrate pollution in groundwater. J. Hydrol. 550, 663–668. https://doi.org/10.1016/j.jhydrol.2017.05.038.

Nakić, Z., Posavec, K., Bačani, A., 2007. A visual basic spreadsheet macro for geochemical background analysis. Groundwater 45, 642–647.

Niu, J., Zhu, X.G., Parry, M.A., Kang, S., Du, T., Tong, L., Ding, R., 2019. Environmental burdens of groundwater extraction for irrigation over an inland river basin in northwest China. J. Clean. Prod. 222, 182–192. https://doi.org/10.1016/j.jclepro.2019.03.075.

Panno, S., Kelly, W., Martinsek, A., Hackley, K.C., 2006. Estimating background and threshold nitrate concentrations using probability graphs. Groundwater 44, 697–709. https://doi.org/10.1111/j.1745-6584.2006.00240.x.

Parrone, D., Ghergo, S., Preziosi, E., 2019. A multi-method approach for the assessment of natural background levels in groundwater. Sci. Total Environ. 659, 884–894. https://doi.org/10.1016/j.scitotenv.2018.12.350.

Pauwels, H., Muller, D., Griffioen, J., Hinsby, K., T en Brower, R., 2007. Bridge: Background Criteria for the Identification of Groundwater Thresholds: Publishable FInal Acitivity Report (Contract).

Preziosi, E., Parrone, D., Del Bon, A., Ghergo, S., 2014. Natural background level assessment in groundwaters: probability plot versus preselection method. J. Geochem. Explor. 143, 43–53. https://doi.org/10.1016/j.gexplo.2014.03.015.

Qian, S.S., Lyons, R.E., 2006. Characterization of background concentrations of contaminants using a mixture of normal distributions. Environ. Sci. Technol. 40, 6021–6025. https://doi.org/10.1021/es0606071.

Reimann, C., Filzmoser, P., Garrett, R.G., 2005. Background and threshold: critical comparison of methods of determination. Sci. Total Environ. 346, 1–16. https://doi.org/10.1016/j.scitotenv.2004.11.023.

Sadler, J., Webb, G.E., Leonard, N.D., Nothdurft, L.D., Clark, T.R., 2016. Reef core insights into mid-holocene water temperatures of the southern great barrier reef. Paleoceanography 31, 1395–1408. https://doi.org/10.1002/2016PA002943.

Sarkar, S., Melnykov, V., Zheng, R., 2020. Gaussian mixture modeling and model-based clustering under measurement inconsistency. Adv Data Anal Classif 14, 379–413. https://doi.org/10.1007/s11634-020-00393-9.

Sakyi, P.O., Alfred Bienibuor, K., Robert, A.K., 2018. Assessing the physico-chemical parameters of water sources in the fiapre vicinity, a suburb of Sunyani in the Brong Ahafo region of Ghana. International Journal of Advances in Scientific Research and Engineering 4, 32890. https://doi.org/10.31695/IJASRE.2018.32890.

Sellerino, M., Forte, G., Ducci, D., 2019. Identification of the natural background levels in the phlaegrean fields groundwater body (southern Italy). J. Geochem. Explor. 200, 181–192. https://doi.org/10.1016/j.gexplo.2019.02.007.

Shalev, N., Burg, A., Gavrieli, I., Lazar, B., 2015. Nitrate contamination sources in aquifers underlying cultivated fields in an arid region–the arava valley, Israel. Appl. Geochem. 63, 322–332. https://doi.org/10.1016/j.apgeochem.2015.09.017.

Simmonds, S., 2009. The Nugget Effect: in Describing the Variability of an (Doctoral Dissertation. Rhodes University.

Sinclair, A., 1974. Selection of threshold values in geochemical data using probability graphs. J. Geochem. Explor. 3, 129–149.

Sutton, M.A., Orteu, J.J., Schreier, H., 2009. Image Correlation for Shape, Motion and Deformation Measurements: Basic Concepts, Theory and Applications. Springer Science & Business Media.

Tay, C.K., Kortatsi, B.K., Hayford, Hodgson, I.O., 2014. Origin of major dissolved ions in groundwater within the Lower Pra Basin using groundwater geochemistry, source-rock deduction and stable isotopes of ^{2}H and ^{18}O. Environ Earth Sci 71, 5079–5097. https://doi.org/10.1007/s12665-013-2912-z.

Tekpor, M., Akrong, M., Asmah, M., Banu, R., Ansa, E., 2017. Bacteriological quality of drinking water in the Atebubu-Amantin district of the Brong-Ahafo region of Ghana. Applied Water Science 7, 2571–2576. https://doi.org/10.1007/s13201-016-0457-5.

Vu, T.K., Hoang, M.K., Le, H.L., 2019. An em algorithm for gmm parameter estimation in the presence of censored and dropped data with potential application for indoor positioning. ICT Express 5, 120–123. https://doi.org/10.1016/j.icte.2018.08.001.

WHO, 2017. Guidelines for Drinking-Water Quality: FIrst Addendum to the, fourth ed.

Yidana, S., Banoeng-Yakubo, B., Aliou, A.S., Akabzaa, T., 2012. Groundwater quality in some Voltaian and Birimian aquifers in northern Ghana—application of multivariate statistical methods and geographic in-formation systems. Hydrol. Sci. J. 57, 1168–1183.

Zhai, Y., Zhao, X., Teng, Y., Li, X., Zhang, J., Wu, J., Zuo, R., 2017. Groundwater nitrate pollution and human health risk assessment by using HHRA model in an agricultural area, NE China. Ecotoxicol. Environ. Saf. 137, 130–142. https://doi.org/10.1016/j.ecoenv.2016.11.010.

Zhou, S., Zhou, K., Wang, J., Yang, G., Wang, S., 2018. Application of cluster analysis to geochemical compositional data for identifying ore-related geochemical anomalies. Front. Earth Sci. 12, 491–505.