



A multi-model study for understanding the contamination mechanisms, toxicity and health risks of hardness, sulfate, and nitrate in natural water resources

Johnbosco C. Egbueri¹

Received: 29 September 2022 / Accepted: 7 March 2023 / Published online: 17 March 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Several water quality contaminants have attracted the attention of numerous researchers globally, in recent times. Although the toxicity and health risk assessments of sulfate and water hardness have not received obvious attention, nitrate contamination has gained peculiar research interest globally. In the present paper, multiple data-driven indexical, graphical, and soft computational models were integrated for a detailed assessment and predictive modeling of the contamination mechanisms, toxicity, and human health risks of natural waters in Southeast Nigeria. Majority of the tested physicochemical parameters were within their satisfactory limits for drinking and other purposes. However, total hardness (TH), SO_4 , and NO_3 were above stipulated limits in some locations. A nitrate health risk assessment revealed that certain areas present a chronic health risk to children, females, and males due to water intake. However, the dermal absorption route was found to have negligible health risks. SO_4 in some locations was above the 100 mg/L Nigerian limit; thus, heightening the potential health effects due to intake of the contaminated water resources. Most samples had low TH values, which exposes users to health defects. There are mixed contamination mechanisms in the area, according to graphical plots, R-mode hierarchical dendrogram, factor analysis, and stoichiometry. However, geogenic mechanisms predominate over human-related mechanisms. Based on the results, a composite diagrammatic model was developed. Furthermore, predictive radial basis function (RBF) and multiple linear regression (MLR) models accurately predicted the TH, SO_4 , and NO_3 , with the RBF outperforming the MLR models. Insights from the RBF and MLR models were useful in validating the results of the hierarchical dendrogram, factor, stoichiometric, and graphical analyses.

Keywords Health risk assessment · Multiple regression · Nitrate · Radial basis function neural network · Sulfate · Total hardness

Introduction

Life on earth depends on water. The two main components of water resources are groundwater and surface water. In recent times, the pressure on groundwater resources (comprising of boreholes and hand-dug wells) has increased, due to easy contamination and pollution of surface water bodies (Egbueri and Agbasi 2022a). Nevertheless, many surface water resources are free of pollution and useful to attend to

the needs of humans. Groundwater is used to meet municipal, agricultural, and recreational needs in both urban and rural regions. Groundwater is also used for industrial and commercial purposes in many nations across the world. In the present time, groundwater is an alternative supply of drinking water in regions where surface water is sparse or extremely polluted (Wang et al. 2020; Omeka et al. 2022). However, there have been high evidences of increasing groundwater pollution across the globe (Li and Wu 2019; Adimalla 2020; Agbasi and Egbueri 2022). The continuous pollution of groundwater, which had been regarded as a safer source of water, has become a global threat to homes, communities, industries, and agricultural sector. In dry and semi-arid areas, groundwater is an essential supply of water for domestic, industrial, and agricultural purposes (Ukah et al. 2019; Wu et al. 2017; Li et al. 2017a). Around

Responsible Editor: Xianliang Yi

✉ Johnbosco C. Egbueri
johnboscoegbueri@gmail.com; jc.egbueri@coou.edu.ng

¹ Department of Geology, Chukwuemeka Odumegwu Ojukwu University, Uli, Anambra State, Nigeria

the world, groundwater is used to produce around 40% of food and about 30% of drinking water (Amiri et al. 2021a; Nickson et al. 2005). However, the quick economic growth and population growth raise the demand for freshwater and degrade the quality of the water.

Numerous pollutants are introduced into the soil and water environments (which are in steady interaction), endangering both ecological stability and human health (Amiri et al. 2021b; Egbueri and Agbasi 2022b). Nonpoint sources of pollution, including septic tank leaks, home wastes, animal manure, and urban runoffs, contaminate both groundwater and surface water resources (Karunanidhi et al. 2020a). When sources of groundwater recharge are contaminated, it seeps into the groundwater through the earth, rendering the water unsafe for human use (Karunanidhi et al. 2020a). Increase in human activities, such as agricultural and industrial practices, as a result of growing populations, cause a series of changes to the environment. These human-induced changes cause severe water emergencies and gradually deteriorate water quality over time, which causes a number of environmental problems as well as human health issues (Ruidas et al. 2022). Increases in water use and contamination are directly correlated with dense population and quick unplanned urbanization in urban areas. Due to overuse and rapid contamination of groundwater, the availability of groundwater that is safe for consumption has given rise to major concerns in many regions of the world (Subba Rao and Chaudhary 2019; Al-Abadi 2017). As a result, the safety of the groundwater supply depends on the quality of the groundwater (Wei et al. 2021). Access to sufficient and safe water is currently a challenge for many nations. Global climate change, the growth of the human population, and decline in precipitation continuously contribute to an ongoing rise in groundwater demand. Therefore, one of the most crucial and required steps to be performed to maintain safe water supply and public health of consumers is the quantitative and qualitative management of natural water resources (Amiri et al. 2021b; Egbueri 2020). Sustainable management of water resources is primarily based on comprehensive water quality research.

Undoubtedly, the management and conservation of water resources can be accelerated and improved with the help of water quality studies. For various reasons, water evaluation and management in Nigeria and other regions of the world have been investigated by numerous authors. Information regarding water types, classification, and suitability for various uses, including irrigation, drinking, bathing, and industry, can be found in hydrogeochemical data (Snousy et al. 2021; Egbueri et al. 2021). The geochemical processes that occur in aquifer materials affect the chemistry of groundwater. When groundwater flows from recharge to discharge locations, its physical and chemical properties are altered, making the water unfit

for human use (Farid et al. 2017; Ali et al. 2018). Surface processes, on the other hand, impact surface water bodies, which are the ready recharge systems of groundwater aquifers. Additionally, hydrogeochemical analysis reveals the water's chemistry in relation to source, climate, geologic, and anthropogenic inputs (Karmakar et al. 2021; Omeka et al. 2022). The weathering of rock-forming minerals and secondary factors like unrestricted water use, a lack of water recharge, wastewater discharge, industrial and agricultural activities, excessive fertilizer use, significant evaporation, and little rainfall have recently put both surface and underground water resources in danger (Egbueri et al. 2019; Karmakar et al. 2021; Ruidas et al. 2022). For determining the hydrogeochemical mechanisms and assessing water quality, the measurement of water physicochemical parameters, such as temperature, pH, main ions, and electrical conductivity, is crucial. In recent years, aside from the hydrogeochemical assessment and modeling of water resources, there has also been a lot of interest in researching the effects of water quality on human health (Li and Wu 2019; Snousy et al. 2021; Egbueri and Mgbenu 2020; Ayejoto et al. 2022; Agbasi and Egbueri 2022).

Access to safe drinking water is a known human right (regardless of wealth, age, tribe, and race), and could promote public health. The World Health Organization (WHO) reports that water demand and extraction from these various sources have dramatically grown in recent decades, impacting water quality and raising health risks. Due to the toxicity and carcinogenic character of the pollutants, a large number of people are exposed to a range of health concerns. It has been reported that pollutants in water are responsible for 70% of diseases and 20% of cancer cases worldwide (Ruidas et al. 2022). While many trace elements necessary for human health are present in groundwater, excessive consumption of these elements can have negative consequences on the body, including reproductive issues, neurological disorders, breathing issues, lung illness, cancer, and hypertension (Ukah et al. 2019; Egbueri 2020). For instance, nitrate (NO_3) is a crucial compound for plant growth; however, its excessive amounts in water can result in methemoglobinemia (Wongsanit et al. 2015; Egbueri 2020). The health of several human populations is often risked due to lack of effective testing and mitigation strategy for drinking water supplies. However, accurate water testing and human health risk evaluation can be very helpful in identifying problematic locations. The human health risk is a key assessment tool that may link a number of hydrogeochemical characteristics to human health and can pinpoint an area's potential health risk zones. To evaluate the health hazards posed by pollution, hazard quotient, hazard index, and total hazard quotient values are frequently utilized (Baig et al. 2011;

Egbueri 2020; Ukah et al. 2019; Wei et al. 2021; Ayejoto et al. 2022).

Several other methods have been used by researchers to study the contamination and management of natural water resources. With respect to the degree of contamination, many scientists classify water using a variety of indexing techniques. Numerous researchers have used diverse strategies up to this point (such as water quality indices, multivariate statistical techniques, and data-driven artificial intelligence techniques) (Islam et al. 2017; Wu et al. 2014, 2020; Liu and Han 2020; Botheju et al. 2021; Egbueri and Agbasi 2022a, b, c). Even though the use of data-driven artificial intelligence (which is emerging as a capable and promising multi-functional approach to all research fields) is relatively new in water research, the interpolation method is still utilized in water quality evaluation. Due to their improved performance over conventional statistical analysis, various data-driven artificial intelligence and machine learning techniques (including regression models, artificial neural networks, support vector machine, Monte Carlo simulation, etc.) have been used widely during the past three decades to evaluate the quality of water resources (Onjia et al. 2022; Ozel et al. 2020; Egbueri and Agbasi 2022a, b, c). Due to their high precision in predictive modeling, many artificial intelligence and machine learning algorithms are widely applicable. These algorithms have been rated as reliable alternative and economical approach for water quality monitoring and assessment (Egbueri 2021; Ozel et al. 2020).

In Nigeria, several investigations on water quality have revealed that surface water and groundwater sources have been polluted and remarked as unsuitable for human use (Ighalo et al. 2020; Ighalo and Adeniyi 2020). However, the application of data-driven artificial intelligence and machine learning algorithms in water quality studies across the country is very limited. The available research studies have primarily concentrated on investigating the hydrogeochemistry and health impacts of potentially toxic elements/metals on humans and ecosystems. Overall, the studies that unraveled hydrogeochemical processes and evolution are far more in number than those that calculated the risks posed to human health by contaminated water. Few studies have been conducted on human health risk based on physical and ionic properties of water, despite the fact that several studies have assessed water quality in Nigeria. Ingestion of NO_3 -contaminated water resources can often pose a serious threat to people, particularly through two exposure routes: the first is through drinking water or other oral channels, and the second is through cutaneous (dermal) interactions (Adimalla 2020; Adimalla and Qian 2019; Rahman et al. 2021). Several studies across the globe have calculated the health risks of NO_3 -contaminated water (Adimalla and Qian 2021; Karunanidhi et al. 2019; Chen et al. 2017a, b). However, only few studies did so in Nigeria (Unigwe et al. 2022). On a

global scale, studies that have investigated the health risks of sulfate (SO_4) and TH of water resources are scarce, despite the fact that these physicochemical parameters are often tested by water quality researchers during hydrogeochemical studies. Generally, in Nigeria and many other parts of the world, there is a dearth of study on water quality in relation to studies of human health risks due to NO_3 , SO_4 , and TH.

Therefore, by developing several investigative graphical, conceptual, and computational models for the Ojoto Region (in southeastern Nigeria), the aim of this comprehensive research is to identify the key contamination mechanisms of NO_3 , SO_4 , and TH in water and establish their connections with human health risks. In this study, health hazard appraisal models were used to assess the toxicity of these contaminants based on the pathways that the water is utilized and the rate at which each contaminant is ingested in relation to various population groups. The specific objectives are to (1) combine descriptive and computational approaches in assessing the toxicity and human health risks of NO_3 , SO_4 , and TH; (2) explore the efficacy of graphical methods, Varimax-rotated factor analysis, and R-mode hierarchical clustering algorithm for unraveling their contamination mechanisms; (3) develop a conceptual diagrammatic model that explains the relationships between geochemical processes, anthropogenic perturbations, water quality, and human health risks; and (4) develop computational predictive multiple linear regression (MLR) and radial basis function neural network (RBF-NN) models for estimating NO_3 , SO_4 , and TH in the water resources. Most previous studies did not compute hazard index values for the different exposure pathways as done in the present paper. This study also introduced cumulative hazard index for wholistic nitrate risk assessment. Moreover, this research is the first of its kind to combine MLR and RBF-NN in the prediction of NO_3 , SO_4 , and TH. Many researchers and policymakers would be able to learn more from baseline information provided in this paper. This study would also serve as a theoretical foundation for the prevention and management of the unsafe, contaminated water resources by local decision-makers.

Materials and methods

Study area description

The present research study was conducted in Ojoto and its neighborhood within Anambra State in the southeastern part of Nigeria (Fig. 1). Onitsha, a major industrial and commercial hub in Nigeria, lies a few kilometers distant from the study area. Agricultural lands, commercial and industrial zones, and human settlements are common land use patterns in the region. Parts of the study area are now home to businessmen and industrialists from Onitsha city, accelerating

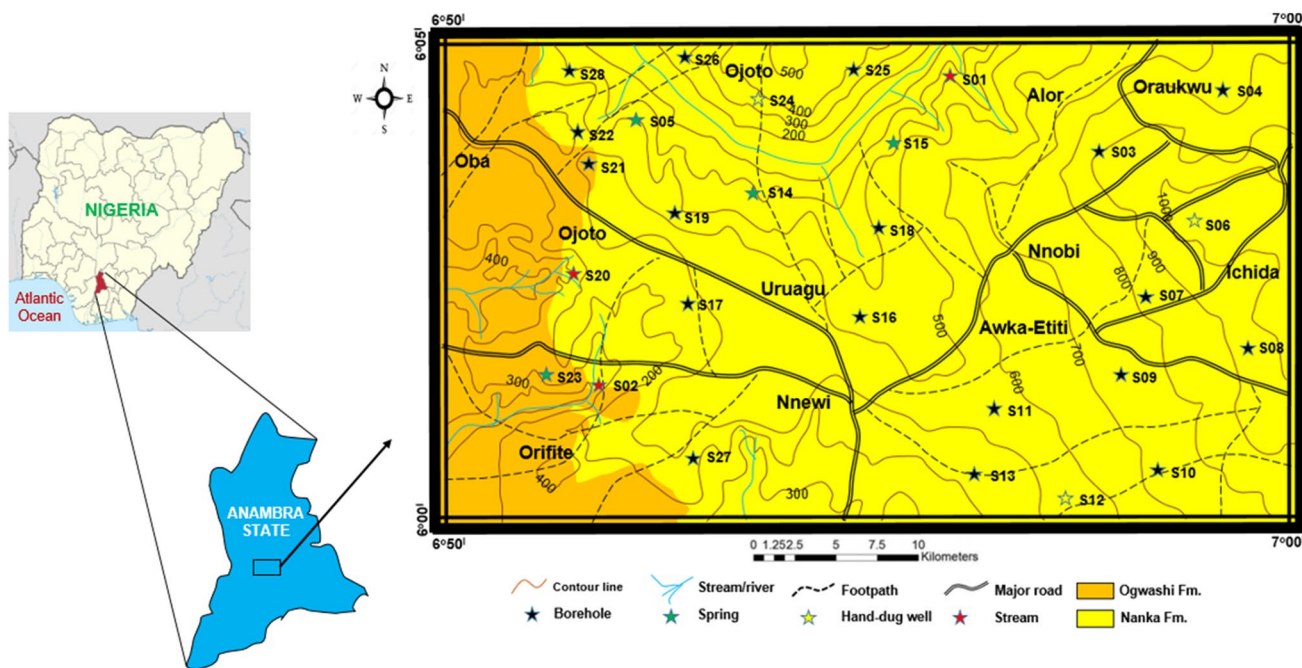


Fig. 1 Maps showing the study area location and the dominant underlying geologic formations

the area’s population expansion. Among the rising sectors in this region are those related to food, water, automobile parts, and petrochemicals. Water is typically distributed in the research region using a variety of plumbing materials consisting of steel, iron, polymer of vinyl chloride, and plastic, in both homes, industries, and commercial venues (including markets). In the research area, there are two main atmospheric seasons: the wet season (which is known to usually spans from April to October) and the dry season (which is known to usually spans from November to March, every year). When the weather is dry, temperatures can get to 32 °C or more (Egbueri et al. 2019). In rainy seasons, annual rainfall is between 1500 and 2000 mm, although the driest months in this season might see little precipitation.

The research region is underlain by two lithostratigraphic, geologic units (Fig. 1): the Nanka Formation (dated Eocene in age) and the Ogwashi Formation (dated Oligocene–Miocene in age) (Nwajide 2013). Friable sands, light-colored mudrocks, sandy shale, thin limestone beds, and ironstones are the dominating lithologies within the Nanka Formation (Nwajide 2013). Over 80% of the research region is covered by the Nanka Formation, which is a regressive deposit made up of the Abakaliki Anticlinorium’s folding-related parts of the 2nd stage of the Campanian-Eocene compressive tectonic motion (Nwajide 2013; Nwachukwu 1972). The Ogwashi Formation was deposited as a result of the 3rd stage of the tectonic motion, which is usually thought to have started at the end of the Eocene. The Ogwashi Formation is mostly made up of bands of limestone, ironstones, and

lignite seams, as well as coarse-textured weakly to moderately cemented sandstone units (Nwajide 2013). Additionally, it has been determined that these two geological formations are abundantly producing aquiferous units with different depths to the aquifer systems (Akpoborie et al. 2011; Okoro et al. 2010a, b; Nfor et al. 2007). Nanka Sandstone’s aquifer’s typical depth has been estimated to be over 20 m in some areas of Anambra State (Nfor et al. 2007), whereas the Ogwashi Formation is estimated to be around 7 m (Okoro et al. 2010a, b). However, there is a dearth of literature describing a more current estimate of the aquifer depths across the studied area. Future hydrogeologic studies may have to consider this aspect of research in the area.

Water sampling, physicochemical testing, and data processing

The study’s objectives were met by randomly selecting twenty-eight water samples from springs, hand-dug wells, streams, and boreholes situated within the study area. Using a portable Geographical Positioning System, the coordinates of each sample location were captured and the spatial view of the sampling stations are shown in Fig. 1. In 1-L plastic bottles, the water samples were gathered and preserved prior to laboratory testing. The collection of water sample from every station was done after some few minutes the water was pumped from the well, to ensure the freshness of the samples. Sufficient quantity of water was collected from each station. The standard operating procedures of the American

Public Health Association (APHA 2005) were adhered to during collection, sample preservation, and analysis in order to guarantee data quality and consistency. Pre-washing the water sample bottles with distilled water and acidifying them with weak hydrochloric acid were both done before the laboratory testing. The acidification, using 1 ml of concentrated HNO₃, was to avoid cationic precipitation. Prior to physicochemical investigation, all of the collected samples were kept in refrigeration condition. A 0.45-µm milli pore cellulose acetate filter was used to filter the water samples before the analysis. The methods of sample preservation employed were designed to stop any unintended chemical reactions or biological growths in the water, before the physicochemical investigation.

In the water samples for this investigation, several variables, like pH, temperature (T), electrical conductivity (EC), K⁺, Na⁺, Ca²⁺, Mg²⁺, Cl⁻, HCO₃⁻, TH, SO₄²⁻, and NO₃⁻, were tested. A handheld analysis kit (HM Digital COM-100) was used to assess the pH and EC at the sampling locations. The TH was tested using ethylenediaminetetraacetic acid (EDTA) titrimetric method. NO₃ concentration was tested with spectrophotometry technique. The HCO₃⁻ was measured using H₂SO₄-based water titration, whereas the Cl⁻ was measured using an AgNO₃-based titration. The concentrations of SO₄²⁻ were determined using the turbidity technique with BaCl₂. The Ca²⁺ and Mg²⁺ were estimated using volumetric techniques (0.05N and 0.01N EDTA), whereas the K⁺ and Na⁺ contents were measured using a Systronics Flame Photometer. Additionally, the laboratory analysis also followed the standards and guidelines recommended by APHA (2005) for the analysis of each parameter. Ion balance error

(IBE) estimated the validity of physicochemical data (Eq. 1). IBE values ≤ 5% were obtained.

$$BE = \frac{\sum cations - \sum anions}{\sum cations + \sum anions} \times 100 \quad (1)$$

Health risk assessment

Health risk due to nitrate

The comprehensive technique used in this study for the assessment of human health risk was initially proposed by the US Environmental Protection Agency (USEPA 1989, 1997). In this study, three categories of the exposed population, including children, females, and males, were assessed for risk.

The following formulas were used to calculate the non-carcinogenic health risk from oral ingestion of NO₃ in the water resources (Adimalla and Qian 2021; Karunanidhi et al. 2019; Chen et al. 2017a, b; Unigwe et al. 2022).

$$CDI = \frac{C \times EF \times ED \times IR}{ABW \times AET} \quad (2)$$

$$HQ_{oral} = \frac{CDI}{RfD} \quad (3)$$

The explanations of the parameters in Eqs. 2 and 3 are presented in Table 1.

Table 1 Values and units of parameters for nitrate health risks assessment (oral and dermal routes)

Parameter	Full meaning	Unit	Child	Female	Male
CDI	Chronic daily intake	mg/kg/day	–	–	–
C	Nitrate/sulfate concentration in water	mg/L	–	–	–
IR	Daily ingestion rate of water	L/day	1	2.5	2.5
EF ^a	Exposure frequency	Days/year	365	365	365
ED ^b	Exposure duration	Year	12	67	64
ABW	Average body weight	kg	15	55	65
AET	Average exposure time	Days	4,380	24,455	23,360
RfD ^c	Reference dose of nitrate	mg/kg/day	1.6	1.6	1.6
DAD	Dermal absorbed dose	mg/kg × day	–	–	–
TC	Contact time	h/day	0.4	0.4	0.4
Ki	Dermal adsorption parameters	cm/h	0.001	0.001	0.001
CF	Conversion factor	–	0.00134	0.00134	0.00134
EV	Bathing frequency	Times/day	1	1	1
SSA ^d	Skin surface area	cm ²	12,000	16,600	16,600

^aUSEPA (1989)

^bAdimalla (2020); Karunanidhi et al. (2020c); Rahman et al. (2021); Unigwe et al. (2022)

^cUSEPA (1989, 1997)

^dUSEPA (1997, 2002)

The following formulas were used to determine the non-carcinogenic health risk associated with cutaneous contact of NO_3 in the water resources (Adimalla and Qian 2021; Karunanidhi et al. 2019; Chen et al. 2017a, b; Unigwe et al. 2022).

$$\text{DAD} = \frac{C \times \text{TC} \times K_1 \times \text{CF} \times \text{EV} \times \text{ED} \times \text{EF} \times \text{SSA}}{\text{ABW} \times \text{AET}} \quad (4)$$

$$\text{HQ}_{\text{dermal}} = \frac{\text{DAD}}{\text{RfD}} \quad (5)$$

The HI_{oral} of nitrate was calculated using Eq. 6 whereas the $\text{HI}_{\text{dermal}}$ of nitrate was obtained using Eq. 7. It is pertinent to mention that most previous studies did not compute HI values for the different exposure pathways.

$$\text{HI}_{\text{oral}} = \sum_{i=1}^n (\text{HQ}_{\text{oral (child)}} + \text{HQ}_{\text{oral (female)}} + \text{HQ}_{\text{oral (male)}}) \quad (6)$$

$$\text{HI}_{\text{dermal}} = \sum_{i=1}^n (\text{HQ}_{\text{dermal (child)}} + \text{HQ}_{\text{dermal (female)}} + \text{HQ}_{\text{dermal (male)}}) \quad (7)$$

The total hazard index of nitrate in the water resources for infants, females, and males was computed using Eq. 8.

$$\text{HI}_{\text{Total (nitrate)}} = \sum_{i=1}^n (\text{HQ}_{\text{oral}} + \text{HQ}_{\text{dermal}}) \quad (8)$$

The hazard index, abbreviated HI, indicates the non-carcinogenic health risk and is represented by the value of HI. A HI number less than 1 implies an acceptable level of human health risk, while an HI value more than 1 indicates a potential risk to human health from nitrate contamination (USEPA 1997, 2002).

Health risk due to sulfate and total hardness

In this study, the human health risks due to sulfate (SO_4) and TH were evaluated using qualitative and descriptive approach, and not computational approach. This is due to lack of standard health-based guidelines and mathematical explanations of key parameters for possible computational approach for SO_4 and TH health risks assessment. However, there are some documented research works in the literature that have investigated and reported the health concerns of these parameters (Sharma and Kumar 2020; Peterson 1951; Morris and Levy 1983; Cocchetto and Levy 1981).

Computational algorithms for unraveling contamination mechanisms

R-mode hierarchical clustering

The R-mode agglomerative hierarchical clustering has been discovered to offer insights into the degree of correlation between examined water quality metrics, much

like the factor analysis and Pearson's correlation analysis. In several previous studies, water quality contamination mechanisms and source apportionment have effectively been achieved by numerous research scholars using this clustering technique (Egbueri 2018). In this modeling, the Ward (1963) linkage approach and squared Euclidean distance were applied. Additionally, the datasets were normalized using the *Z-score* transformation technique (Eq. 9). The same cluster of water quality variables is therefore assumed to have similar or identical anthropogenic or geological influences or closer sources of pollution. In this investigation, only one R-mode dendrogram was used to identify the contamination mechanisms and sources. All of the tested physicochemical variables for the water quality assessment were used in the R-mode hierarchical clustering. IBM SPSS (v. 22) was used for this investigation.

$$Z_z = \frac{X - \mu_x}{\sigma_x} \quad (9)$$

where X represents the individual variable value, μ_x represents the average value from the row, and σ_x represents the standard deviation of variable values in the row.

Varimax-rotated factor analysis

Datasets are frequently intricate and multifaceted in character, like the water quality dataset. As a result, it is frequently challenging to immediately understand such datasets. But dimensionality reduction machine learning tools have frequently shown to be helpful for more understandable evaluation and interpretation of water quality. For the purpose of dimensionality reduction and interpretation of water quality datasets, principal component extractions were generated by Varimax-rotated factor analysis. This analysis was done with SPSS (v. 22) to complement the result of the R-mode hierarchical cluster analysis. First, the dataset's appropriateness for principal component analysis (PCA) was evaluated using Bartlett's test of sphericity (BTS), which identifies correlation matrices, and Kaiser–Meyer–Olkin (KMO) score, which assesses sampling adequacy. For PCA to be effective, studies have stated that the KMO sampling adequacy for the set of variables must be more than 0.50 (Liu et al. 2003). If the KMO is greater than 0.5, the factor analysis is appropriate to provide a significant reduction in data size.

The KMO value to confirm the sampling adequacy in the present study was 0.673. For this water quality investigation, the BTS (χ_2) agrees with $p(\text{Sig.})$ 0.0001 (i.e., $\text{BTS} \approx 0$). According to these results, the PCA-based Varimax-rotated factors can be implemented and the water quality dataset is

valid (Ustaoglu and Tepe 2019; Aguirre et al. 2019; Long and Luo 2020). Based on eigenvalues > 1, significant factor classes crucial for the interpretation of the water quality dataset were identified. Twelve factor classes were identified in this study. However, only five classes with eigenvalues > 1 were selected as illustrated in the scree plot of Fig. 2. It is important to note that the scores from the factor analysis were divided into three groups. Factor loadings below 0.50 were regarded as weak and inconsequential, but loadings between 0.50 and 0.75 were regarded as moderate and substantial. Strong loadings were encountered in ranges greater than 0.75, though. Strong loadings are typically considered to be very significant and to explain more specifics about a particular dataset (Egbueri and Agbasi 2022a).

Computational algorithms for predictive modeling of TH, SO₄, and NO₃

Radial basis function neural networks

In an RBF-NN, learning happens in two stages. The output layer comes first, followed by the hidden layer. Instead of being mapped from the input to the output, the parameters of the hidden layer are decided by the distribution of the inputs (Emami et al. 2020). On the other hand, the output layer uses supervised learning to determine its parameters. A single variable makes up each input unit in the hidden unit (Niroobakhsh et al. 2012). The hidden layer is where RBF-NN and MLP-NN differ most significantly from one another. Both structural and functional differences exist in this scenario. The radial functions of the neurons in the hidden layer account for the functional difference, while the training procedure and the number of neurons account for the structural difference (Ozel et al. 2020). Contrary to the

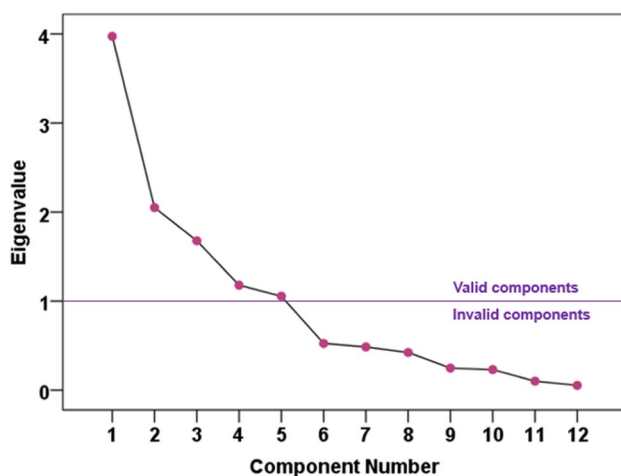


Fig. 2 Scree plot showing that only five components (factor classes) with eigenvalue > 1 are valid for this study

MLP-NN, the RBF-NN does not need parameter learning. A linear weight change is carried out for the radial functions. A trial-and-error approach is used to determine the proper neurons' number in the hidden layer and the spread numbers in return (Ozel et al. 2020).

An RBF-NN has a different architecture than other neural networks. In recent decades, RBF-NN has seen increasing use in load forecasting, pattern recognition, signal processing, and water quality predictions, etc., due to its straightforward structural design and training effectiveness (Panda et al. 2004; Egbueri and Agbasi 2022c). In comparison to back-propagation neural networks, RBF-NN provides a number of benefits, including a guaranteed learning procedure and a high processing speed because it has just one hidden layer (Panda et al. 2004). However, different hidden layers in the RBF-NN architecture can be employed to carry out the error back-propagation function (Panda et al. 2004). These layers, however, are optional and hardly used. The fact that curve-fitting approximation is used to build RBF-NN in a highly dimensional space to provide the best fit to training data, as opposed to back-propagated networks, is another advantage of RBF-NN over back-propagated networks (Panda et al. 2004). Using exponentially degraded nonlinear functions, RBF-NN creates approximations to nonlinear input–output mapping, whereas MLP-NN can only produce a wider approximation. Thus, the mapping function of an RBF-NN is primarily composed of Gaussians (Eq. 10) rather than sigmoid as it is in the case of an MLP-NN (Ozel et al. 2020; Niroobakhsh et al. 2012; Egbueri and Agbasi 2022c). Localized approximation in input–output mapping is always preferred (Haykin 1999) in RBF-NN. Net convergence is not guaranteed in an MLP-NN, although it is in an RBF-NN. As a result, the RBF-NN is growing in acceptance despite the complexity involved with its computation and challenging ideal parameterization (Panda et al. 2004).

$$f(x) = \exp\left(-\frac{\|x - c_i\|^2}{\sigma_i^2}\right), i = 1, 2, \dots, k \quad (10)$$

where x represents input variable; c represents the distance from the gaussian function center; and σ represents the spread variable, which substantially influences the performance of the neural network.

Three layers make up a standard RBF-NN: input, hidden, and output layers. These layers are all connected by feedforward arcs in their entirety (Panda et al. 2004). For a specific number of tuned units that are completely coupled to a linear output layer, an RBF-NN's structure consists of a single hidden layer. Literature searches in the course of the present study indicated that RBF-NN has not been used to model the three water quality parameters (TH, SO₄, and NO₃) before. Using the methodological instructions listed in Table 2, the RBF-NN modeling was used in this study

to simulate and predict the concentrations of TH, SO₄, and NO₃. For this purpose, three feedforward RBF-NN models were created.

Multiple linear regression models

A common algorithmic technique for parameter prediction is MLR model. In several academic disciplines, the concept of soft computing has been heavily utilized. It has also been used to forecast parameters and indices linked to water quality (Egbueri and Agbasi 2022a, b, c). MLR modeling was basically included into this work to support and complement the RBF-NN modeling and forecast the TH, SO₄, and NO₃ by focusing on the linear relationships between the input and output variables. The combination of the RBF-NN and MLR models also reduces the biases that could be obtained when a single (standalone) predictive model is used (Egbueri 2022). Moreover, MLR is suitable for this study because of its capability to shed insight on the contributions made by each input parameter in predictive research (Egbueri et al. 2022). When compared to other soft computing methods, the MLR is also quite easy to implement. This suggests that even someone without specialized artificial intelligence knowledge may use it to assess and forecast water quality. MLR requires little computational resources, suggesting that it is simple to compute even in fundamental, low-memory systems. MLR is adaptable in addition to being straightforward. This suggests that it can allow several input parameters to participate in regression modeling at once. The mathematical formulation of the MLR model is expressed as Eq. 11 (Weisberg 1985; Gaya et al. 2020; Egbueri and Agbasi 2022a, b, c).

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_ix_i + \epsilon \tag{11}$$

where x_i represents the value of the i th predictor, b_0 represents the regression constant, and b_i represents the i th predictor’s coefficient. Finally, ϵ represents the error for each unique i (Agbasi and Egbueri 2022; Chen and Liu 2015).

In this modeling, the three focal parameters (TH, SO₄, and NO₃) were predicted and three MLR models were consequently produced for each of the water quality parameters. The models were all developed using the IBM SPSS (v. 22).

Verification of RBF-NN and MLR models

Residual error, sum of the square errors (SSE), relative error (RE), and R^2 (determination coefficient) were used in this work to evaluate the accuracy and reliability of the RBF-NN models. Equations 12 through 15 provide mathematical expressions for each of these statistical model performance indicators. The three RBF-NN frameworks created in this study would give water managers and hydrogeologists the ability to choose the best analytical parameters for sustainable management of the drinking water resources using the principles of artificial intelligence.

$$\text{Residual error} = X_{\text{measured value}} - X_{\text{predicted value}} \tag{12}$$

$$\text{SSE} = \sum_{i=1}^n (X_{\text{measured value}} - X_{\text{predicted value}})^2 \tag{13}$$

$$\text{RE} = \frac{(X_{\text{measured value}} - X_{\text{predicted value}})}{X_{\text{measured value}}} \tag{14}$$

Table 2 Essential methodological commands for the development of the RBF-NN models

Model parameter	Command/activity report
Input variables	The independent parameters are pH, T, EC, TH, Na, K, Ca, Mg, Cl, SO ₄ , HCO ₃ , and NO ₃ . For the predictions of TH, SO ₄ , and NO ₃ , the predicted is made the dependent variable and all others used as predictors
Output variables	The dependent parameters are TH, SO ₄ , and NO ₃
ANN type	Radial basis function (RBF)
Rescaling of covariates	Normalized
Partitioning of dataset	Randomly assigned cases based on relative number of cases: In all scenarios the portioning for training datasets was ≥ 70% while that for the testing datasets was ≤ 30%. The validity of all cases was 100%
Number of hidden layers	One (1)
Hidden layer activation function	Softmax
Output layer activation function	Identity
Number of units	Automatically computed
Rescaling of scale dependent variables	Standardized
Activation function	Identity

$$R^2 = 1 - \frac{\sum_{i=1}^n (X_{\text{predicted value}} - X_{\text{measured value}})^2}{\sum_{i=1}^n (X_{\text{predicted value}} - X_{\text{average value}})^2} \quad (15)$$

where n represents the number of data points.

Furthermore, several statistical methods were also used to evaluate the accuracy, validity, and reliability of the three MLR models developed in this study. They include R^2 (Eq. 15), adjusted R -square (R^2_{adj} , Eq. 16), standard error of estimate (SEE, Eq. 17), and multiple correlation coefficient (R , Eq. 18). It is important to mention that the SEE is the same thing as the root mean square error.

$$R^2_{\text{adj}} = 1 - \left[\frac{(1 - R^2)(n - 1)}{n - k - 1} \right] \quad (16)$$

where n represents the number of data points; k represents the number of input parameters.

$$\text{SEE} = \sqrt{\sum_{i=1}^n \frac{(X_{\text{measured value}} - X_{\text{predicted value}})^2}{n}} \quad (17)$$

$$R = \frac{n \sum (y \cdot y') - (\sum y)(\sum y')}{\sqrt{[n(\sum y^2) - (\sum y)^2][n(\sum y'^2) - (\sum y')^2]}} \quad (18)$$

where y represents the measured value; y' represents the predicted value; n represents the number of data points.

The R , R^2 , and/or R^2_{adj} statistics can be used to represent the variance and correlation between the predicted variables and the measured variables. The model's performance, accuracy, and dependability are higher the closer R , R^2 , and R^2_{adj} values are to 1 (Egbueri and Agbasi 2022a; Menard 2000). The difference between the predicted variable values and the measured variable values of the RBF-NN models is described by the residual error scores, which are frequently visualized as residual plots. The RBF-NN prediction's accuracy with respect to 0 increases as it approaches 0. Thus, residuals of 0 have perfect model precision. However, the model performs worse the larger the values on the y -axis of the residual error plot. The SSE conveys the accumulation of the magnitude of mistake (error) related to model prediction. The SSE, which is typically a positive number, is hence the sum squared difference between the predicted variable values and the measured variable values. Typically, the model performs better, is more accurate, and is more reliable the lower the SSE value. The RE can also be used to convey how closely the predicted values match the values of the measured variables. Similar to the SSE metrics, the lower the RE, the higher the model performance and reliability. Further, an idea of how well an MLR model fits a given set of data

could be gotten from the SEE. Statistically, higher SEE values indicate bad MLR models while lower SEE values (that are much closer to 0) often indicate better models.

Results and discussion

Statistical distributions and general water quality characteristics

The traditional method for determining a general understanding of the quality of water resources involves the analysis of a number of physicochemical parameters. Understanding the statistical distribution and the peculiar characteristics of water resources is paramount to effective and sustainable management of water. The physicochemical compositions of the water in the research region and the WHO (2017) permissible limits of the physicochemical parameters are summarized in Table 3. Box-whisker plots were created to explain the statistical distributions and variations of the physicochemical variables. As can be seen from Fig. 3, there are different levels of variances in the concentrations of these physical and chemical parameters across the water samples. The SO_4^{2-} content, in particular, showed the most striking variance between the sampling stations (Fig. 3). The research area's pH ranged from 4.00 to 6.30, as shown in Table 3. Thus, it is adjudged that all the samples are categorized as weakly acidic water resources. This suggests that they may have a tendency to corrode metals, water distribution routes, and aesthetics (Egbueri 2022). Additionally, the acidic waters may have a negative impact on a person's health, including gastrointestinal problems and degeneration of the mucous membranes (Egbueri et al. 2020). With respect to the information obtained from Fig. 3 a, it was noticed that the pH of the water resources had the lowest statistical variation in the study area. Humic acid from vegetative (organic) decays, as well as other geochemical processes, are the most likely influencers of the waters' acidity (Egbueri et al. 2020).

Water temperature (T), which ranged from 23 to 29 °C (Table 3), indicated that the analyzed water resources were in warm conditions. This could be a favorable state for some geochemical reactions to take place in the water environment. In general, drinking chilly water is considerably preferable to warm water (WHO 2017). Warm water is said to speed up the dissolving of minerals and promote the growth of potentially dangerous bacteria, which could affect the water's color, taste, odor, and susceptibility to corrode (Egbueri 2019). Additionally, investigations have found that warmer seas typically have less oxygen, indicating anaerobic conditions (Egbueri 2019). In such a scenario, nitrates could be converted to nitrites by nitrate-reducing microorganisms and increasing the water's iron concentration,

Table 3 Water quality parameters for the present investigation

Sample ID	Source	pH	T	EC	TH	Na ⁺	K ⁺	Ca ²⁺	Mg ²⁺	Cl ⁻	SO ₄ ²⁻	HCO ₃ ⁻	NO ₃ ⁻
S01	Stream	5.40	26.00	43.00	17.00	14.00	6.00	8.00	0.50	2.00	10.00	1.20	0.60
S02	Stream	6.30	24.00	29.00	33.00	13.00	0.00	13.00	0.90	15.00	14.00	4.00	0.00
S03	BH	5.80	25.00	42.00	10.00	15.00	5.00	2.00	0.20	2.00	70.00	2.00	0.90
S04	BH	4.60	24.00	22.00	13.00	13.00	4.00	4.00	0.30	4.00	10.00	0.00	0.00
S05	Spring	5.20	24.00	13.00	10.00	25.00	6.00	7.00	0.20	25.00	24.00	0.00	0.04
S06	HW	4.80	28.00	69.00	16.00	17.00	8.00	8.00	0.60	8.00	115.00	3.20	0.02
S07	BH	4.70	25.00	73.00	45.00	10.00	3.00	20.00	0.70	4.00	48.00	0.00	0.03
S08	BH	4.80	23.00	33.00	16.00	8.00	7.00	8.00	0.50	2.00	13.00	0.00	1.90
S09	BH	4.80	26.00	16.00	18.00	10.00	4.00	3.00	0.60	21.00	24.00	0.40	0.03
S10	BH	5.90	27.00	102.00	52.00	27.00	7.00	25.00	16.00	20.00	67.00	3.50	0.00
S11	BH	4.30	25.00	19.00	8.00	15.00	3.00	6.00	0.40	12.00	8.00	3.20	0.20
S12	HW	5.10	25.00	25.00	8.00	15.00	0.00	4.00	0.20	3.00	40.00	3.20	0.00
S13	BH	6.00	27.00	42.00	24.00	16.00	4.00	11.00	0.60	5.00	25.00	2.00	0.30
S14	Spring	4.50	26.00	35.00	12.00	14.00	0.00	6.00	0.00	62.00	55.00	0.50	0.00
S15	Spring	4.60	27.00	79.00	50.00	28.00	8.00	23.00	1.20	41.00	60.00	0.50	0.00
S16	BH	4.00	25.00	53.00	33.00	15.00	12.00	10.00	0.00	3.00	33.00	0.40	18.48
S17	BH	5.00	24.00	10.00	6.00	17.00	0.00	2.00	0.50	3.00	18.00	0.00	8.40
S18	BH	6.40	23.00	43.00	32.00	12.00	5.00	8.00	0.60	4.00	26.00	3.00	0.06
S19	BH	4.90	27.00	41.00	8.00	38.00	9.00	4.00	0.40	33.00	72.00	5.00	4.20
S20	Stream	5.00	26.00	38.00	22.00	30.00	8.00	8.00	0.50	18.00	64.00	3.20	0.40
S21	BH	5.30	25.00	32.00	12.00	12.00	8.00	3.00	0.30	25.00	8.00	0.40	1.68
S22	BH	4.60	23.00	19.00	14.00	14.00	3.00	3.00	0.30	23.00	25.00	0.60	0.03
S23	Spring	4.40	24.00	24.00	12.00	14.00	5.00	6.00	0.30	2.00	7.00	0.30	0.10
S24	HW	4.10	29.00	40.00	16.00	20.00	8.00	7.00	0.30	9.00	130.00	0.60	0.03
S25	BH	4.40	28.00	37.00	8.00	11.00	3.00	5.00	0.00	10.00	64.00	0.00	1.80
S26	BH	4.80	25.00	12.00	10.00	12.00	11.00	7.00	1.00	21.00	15.00	0.00	14.40
S27	Spring	4.20	25.00	23.00	12.00	21.00	4.00	4.00	0.80	2.00	23.00	0.00	0.10
S28	BH	5.40	26.00	21.00	10.00	14.00	8.00	2.00	0.30	3.00	70.00	1.50	1.24
WHO (2017) limit	–	6.5–8.5	Ambient	1000	100	200	12	75	50	250	250	250	50

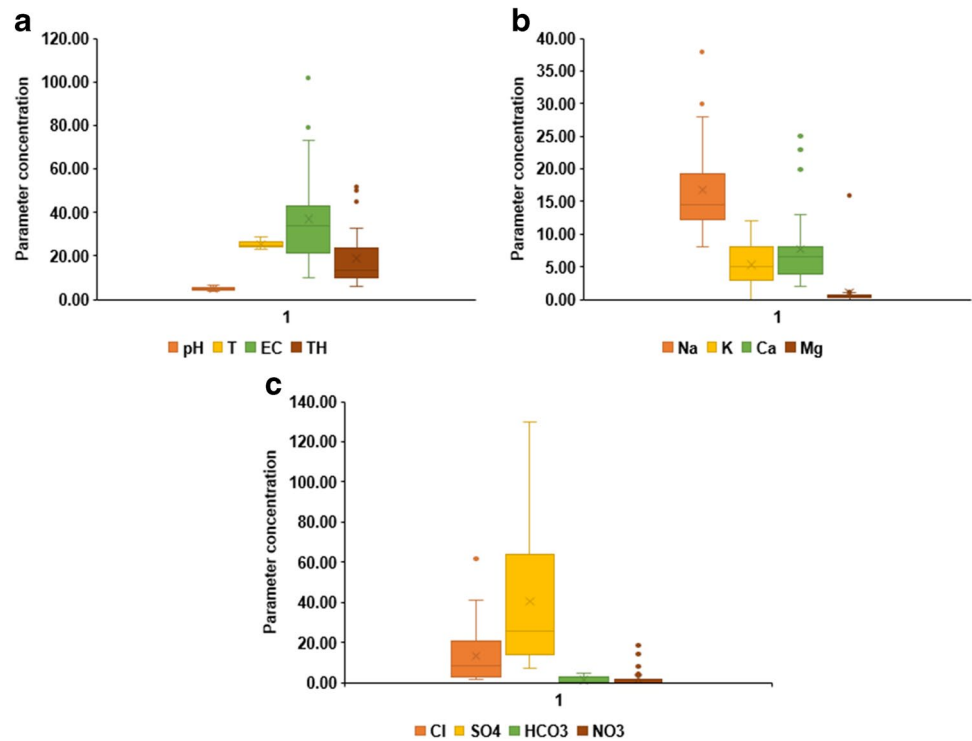
BH borehole, *HW* hand-dug well. Except pH, T (measured in °C), and EC (measured in µS/cm), all other parameters were measured in mg/L

which could then result in the discoloration of aerated pipe-borne water (Egbueri 2019). Water's EC typically serves as a measure of the ions and dissolved solids that are present. Attributable to low level or rate of cationic dissolution in the water resources, the EC was seen to be well below the SON (2015) and WHO (2017) benchmark, ranging between 10.00 and 102.00 µS/cm (Table 3). The statistical information presented in Fig. 3 a indicate that the EC had the second highest level of statistical variation across the samples. The lack of mineralization in the water caused by geogenic and anthropogenic processes is indicated by the low EC values. The conductivity of each water sample is generally low. This is consistent with the water resources' relatively low concentrations of salt-forming ions (particularly sodium and chloride ions), which are below their SON and WHO benchmarks (Table 3). All of the analyzed samples are excellent freshwaters for drinking, according to the low EC values (Langenegger 1990). Moreover, it is indicated that the

inhabitants are not at risk due to high water conductivity, since the EC values are well within acceptable limits and standards. The values of TH of the local water resources, which frequently reflects the levels of Ca²⁺ and Mg²⁺ and the ability of a water sample to neutralize soap (Wu et al. 2017) during laundry, ranged from 6.00 to 52.00 mg/L (Table 3). All of the obtained TH values in this investigation were found to be within the national and international standards (SON 2015; WHO 2017).

Na⁺, K⁺, Ca²⁺, and Mg²⁺ values in the area ranged from 8.00 to 38.00 mg/L, 0.00 to 12.00 mg/L, 2.00 to 25.00 mg/L, and 0.00 to 16.00 mg/L, respectively. According to the maximum concentrations reported in mg/L, the relative abundance of the principal cations in the water resources is in the following order, in terms of their highest concentration values: Na⁺ > Ca²⁺ > Mg²⁺ > K⁺. The box plots of the cations, which show their graphical-statistical distributions and variations, are presented in Fig. 3 b. Na⁺ has the highest

Fig. 3 Box and whisker plots showing the statistical distribution and variation of **a** pH, T, EC, and TH; **b** Na⁺, K⁺, Ca²⁺, and Mg²⁺; **c** Cl⁻, HCO₃⁻, and NO₃⁻



average value (of 16.79 mg/L) among the major cations, which is well below the 200 mg/L permissible level recommended by the SON (2015) and WHO (2017) for drinking water. According to several studies, sewage irrigation may be the cause of a water sample's high Na⁺ concentration (Wu et al. 2013, 2017). However, the lower concentration of Na⁺ in the water in the current investigation is thought to be partly due to minimal water evaporation (Wei et al. 2021). Maximum K⁺ concentration is just 12 mg/L, which is likely due to limited mobility brought on by biological activities (Wu et al. 2017). While sodium is a crucial mineral supplement for the human diet, calcium is crucial for healthy bone and teeth growth (Egbueri 2019; Karunanidhi et al. 2020a). However, high sodium levels are linked to an increased risk of hypertension, heart problems, and nervous system disorders. Similar to this, drinking water with K⁺ concentrations over the recommended 12 mg/L (WHO 2017) (as in sample S16) can cause nervous system diseases (Egbueri et al. 2020; Karunanidhi et al. 2020a).

Cl⁻, SO₄²⁻, HCO₃⁻, and NO₃⁻ have concentration ranges of 2.00–62.00 mg/L, 7.00–130.00 mg/L, 0.00–5.00 mg/L, and 0.00–18.48 mg/L, respectively, for the major anions (Table 3). The study's findings show that all of the samples' anion concentrations were within permissible ranges (SON 2015; WHO 2017). Therefore, the consumption of too many major anions does not put water resource consumers at risk for health problems. The order of the main anions, measured in mg/L, is SO₄²⁻ > Cl⁻ > NO₃⁻ > HCO₃⁻. The average SO₄²⁻ concentration in the region is 40.64 mg/L, which is

far lower than the national and international standards for the quality of drinking water (250 mg/L for WHO (2017) and 100 mg/L for SON (2015)). Drinking water could have a bad taste if there is SO₄²⁻ in it (SON 2015; Egbueri and Mgbenu 2020). Additionally, too much SO₄²⁻ in water can clog pipes and make people sick with diarrhea. Also, too much HCO₃⁻ can affect how acidic water is. However, the extremely low quantities of HCO₃⁻ may be a sign that the location has little soil organic matter decomposing, releasing CO₂, which is a significant source of water HCO₃⁻. Concentrations of Cl⁻ and SO₄²⁻ are related to geogenic processes and evaporation rates. Overall, it was discovered that the major ions in the study region include SO₄²⁻, Cl⁻, Na⁺, and Ca²⁺, which may be the result of mixed geochemical processes occurring in the aqueous environment of the study area. This is consistent with a previous study conducted in the region by Egbueri et al. (2019), which reported that the research area is predominantly characterized by Na–Ca–SO₄ water type and Na–SO₄ water facies. A variety of human activities, however, have also been well-documented to have a significant impact on the quality of water suitable for human use. For instance, despite the comparatively low NO₃⁻ and Cl⁻ concentrations in the study area, it is likely that domestic, agricultural, and industrial activities play contributory roles in the NO₃⁻ contamination (Egbueri and Agbasi 2022a, b; Wei et al. 2021). The WHO's (2017) and SON's (2015) safe drinking water standard for nitrate is 50 mg/L. When it occurs in excess of this value, nitrate can threaten and impact the health conditions of both children and adults (males and

females). Nevertheless, studies have shown that children are the human population mostly affected by nitrate contamination and pollution (Unigwe et al. 2022; Panneerselvam et al. 2021; Egbueri and Mgbenu 2020).

Toxicity and health risk assessments of NO₃, SO₄, and TH

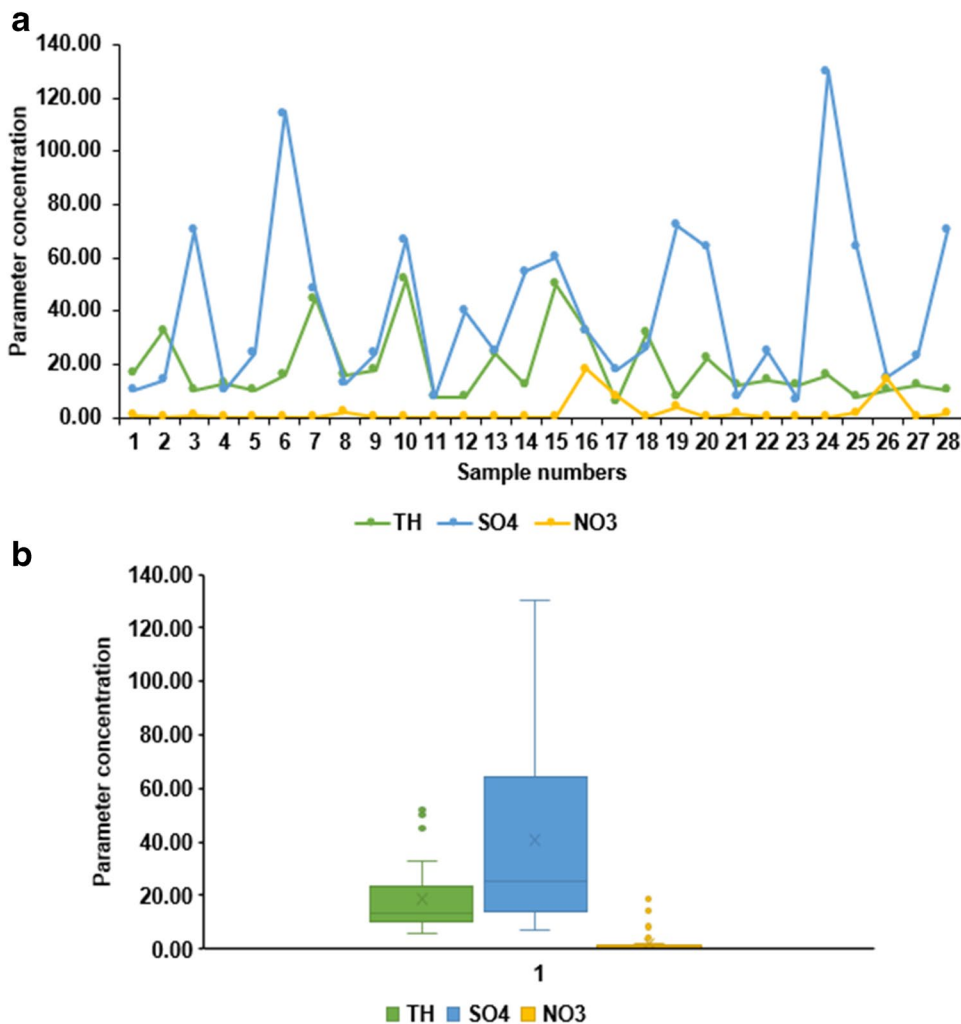
It is evident that variations in geochemical processes, anthropogenic activities, and contamination of water resources are related to changes in human (and even animal) health. All over the world (especially in under-developed and developing semi-arid and arid regions), these stand out as some of the biggest challenges people, researchers, and policy-makers continuously have to deal with. The toxicity and human health risks of NO₃, SO₄, and TH were evaluated and described in this study. These three are the focus parameters in the current investigation. Although by comparison to their respective standard limits the water resources were generally said to be pollution-free, the contamination levels in some samples are a threat to human health. Line graphs and box

plots were used to demonstrate the statistical trends and variations of the NO₃, SO₄, and TH (Fig. 4). It was noticed that the concentrations of these parameters across the samples varied in trends and distributions. However, as seen in Fig. 4, SO₄ and TH had similar trends in some samples whereas NO₃ had a striking variation in trend.

Toxicity and health risk due to nitrate

Since nitrogen molecules are crucial for plant growth, agricultural productivity, and food security, they are required components of water. But nitrate (NO₃⁻), a well-known environmental pollutant, is produced both naturally and as a result of a number of human activities. In the majority of natural water environments, NO₃⁻ contributes significantly to the ionic charges. Thus, international guidelines and regulations utilize NO₃⁻ ions due to its negative effects on people of different sexes and age groups (Torres-Martinez et al. 2020). However, prolonged exposure to high nitrate concentrations can result in major health risks for adults as well as children, including methemoglobinemia

Fig. 4 Depiction of the statistical trends and variations of the focus parameters, TH, SO₄, and NO₃, using **a** line graphs and **b** box and whisker plots



and stomach cancer (SON 2015; WHO 2017; Datta et al. 1997; Unigwe et al. 2022; Karunanidhi et al. 2019; Panneerselvam et al. 2021). Results of the computed human health risks due to nitrate-contaminated water ingestion and dermal contact are presented in Tables 4 and 5, respectively. The computed total health hazard index (THI) of NO₃ in the study area is summarized in Table 6. The CDI, HQ, and HI values for oral and dermal health risks and the THI were calculated for males, females, and children in the study area. Bar charts showing the demarcation of the water quality status, with respect to the NO₃ HI (risk due to oral and dermal pathways) and THI, are represented by Fig. 5 a and b, respectively.

Traditionally, HQ and HI values < 1 indicate that the analyzed water is safe and pose low and negligible non-carcinogenic health risk (Ayejoto et al. 2022; Agbasi and Egbueri 2022). On the other hand, HQ and HI values > 1 indicate substantial human health risk due to use of NO₃-contaminated water (Unigwe et al. 2022; Panneerselvam et al. 2021; Karunanidhi et al. 2019; USEPA 1997, 2002). Children's HQ

values for the health risk due to ingestion (oral route) varied from 0.00 to 0.7700, with a mean of 0.0818. (Table 4). The average HQ value for females was 0.0557, with values ranging from 0.00 to 0.5250. (Table 4). Males' HQ values varied from 0.00 to 0.1770, with an average of 0.0189 (Table 4). Additionally, a careful analysis indicated that the HQ results for children, females, and males were < 1 and within USEPA's (1997, 2002) acceptable limits. However, higher HQ values were observed in S16 and S26 (Table 4). This indicates that they were possibly the most predisposed water sources that received the highest impact from anthropogenic activities and related geochemical reactions that affect nitrate enrichment. According to the HQ values, children > females > men is the order for non-carcinogenic risk of exposure in the research area. Similar trends have been reported by previous researchers from different parts of the world (Unigwe et al. 2022; Panneerselvam et al. 2021; Chen et al. 2017a, b; Adimalla and Qian 2021; Karunanidhi et al. 2019), indicating that children are the most vulnerable. The HI for oral pathway was computed by summing the oral HQs

Table 4 Human health risk results due to nitrate ingestion by children, females, and males

Sample ID	Source	CDI (child)	HQ _{oral} (child)	CDI (female)	HQ _{oral} (female)	CDI (male)	HQ _{oral} (male)	HI _{oral}
S01	Stream	0.04000	0.02500	0.02727	0.01705	0.00923	0.00577	0.0478
S02	Stream	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S03	BH	0.06000	0.03750	0.04091	0.02557	0.01385	0.00865	0.0717
S04	BH	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S05	Spring	0.00267	0.00167	0.00182	0.00114	0.00062	0.00038	0.0032
S06	HW	0.00133	0.00083	0.00091	0.00057	0.00031	0.00019	0.0016
S07	BH	0.00200	0.00125	0.00136	0.00085	0.00046	0.00029	0.0024
S08	BH	0.12667	0.07917	0.08636	0.05398	0.02923	0.01827	0.1514
S09	BH	0.00200	0.00125	0.00136	0.00085	0.00046	0.00029	0.0024
S10	BH	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S11	BH	0.01333	0.00833	0.00909	0.00568	0.00308	0.00192	0.0159
S12	HW	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S13	BH	0.02000	0.01250	0.01364	0.00852	0.00462	0.00288	0.0239
S14	Spring	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S15	Spring	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S16	BH	1.23200	0.77000	0.84000	0.52500	0.28431	0.17769	1.4727
S17	BH	0.56000	0.35000	0.38182	0.23864	0.12923	0.08077	0.6694
S18	BH	0.00400	0.00250	0.00273	0.00170	0.00092	0.00058	0.0048
S19	BH	0.28000	0.17500	0.19091	0.11932	0.06462	0.04038	0.3347
S20	Stream	0.02667	0.01667	0.01818	0.01136	0.00615	0.00385	0.0319
S21	BH	0.11200	0.07000	0.07636	0.04773	0.02585	0.01615	0.1339
S22	BH	0.00200	0.00125	0.00136	0.00085	0.00046	0.00029	0.0024
S23	Spring	0.00667	0.00417	0.00455	0.00284	0.00154	0.00096	0.0080
S24	HW	0.00200	0.00125	0.00136	0.00085	0.00046	0.00029	0.0024
S25	BH	0.12000	0.07500	0.08182	0.05114	0.02769	0.01731	0.1434
S26	BH	0.96000	0.60000	0.65455	0.40909	0.22154	0.13846	1.1476
S27	Spring	0.00667	0.00417	0.00455	0.00284	0.00154	0.00096	0.0080
S28	BH	0.08267	0.05167	0.05636	0.03523	0.01908	0.01192	0.0988

Table 5 Human health risk results due to nitrate dermal contact by children, females, and males

Sample ID	Source	DAD (child)	HQ _{dermal} (child)	DAD (female)	HQ _{dermal} (female)	DAD (male)	HQ _{dermal} (male)	HI _{dermal}
S01	Stream	0.00026	0.00016	0.00010	0.00006	0.00008	0.00005	0.0003
S02	Stream	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S03	BH	0.00039	0.00024	0.00015	0.00009	0.00012	0.00008	0.0004
S04	BH	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S05	Spring	0.00002	0.00001	0.00001	0.00000	0.00001	0.00000	0.0000
S06	HW	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000	0.0000
S07	BH	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000	0.0000
S08	BH	0.00081	0.00051	0.00031	0.00019	0.00026	0.00016	0.0009
S09	BH	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000	0.0000
S10	BH	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S11	BH	0.00009	0.00005	0.00003	0.00002	0.00003	0.00002	0.0001
S12	HW	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S13	BH	0.00013	0.00008	0.00005	0.00003	0.00004	0.00003	0.0001
S14	Spring	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S15	Spring	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
S16	BH	0.00792	0.00495	0.00299	0.00187	0.00253	0.00158	0.0084
S17	BH	0.00360	0.00225	0.00136	0.00085	0.00115	0.00072	0.0038
S18	BH	0.00003	0.00002	0.00001	0.00001	0.00001	0.00001	0.0000
S19	BH	0.00180	0.00113	0.00068	0.00042	0.00057	0.00036	0.0019
S20	Stream	0.00017	0.00011	0.00006	0.00004	0.00005	0.00003	0.0002
S21	BH	0.00072	0.00045	0.00027	0.00017	0.00023	0.00014	0.0008
S22	BH	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000	0.0000
S23	Spring	0.00004	0.00003	0.00002	0.00001	0.00001	0.00001	0.0000
S24	HW	0.00001	0.00001	0.00000	0.00000	0.00000	0.00000	0.0000
S25	BH	0.00077	0.00048	0.00029	0.00018	0.00025	0.00015	0.0008
S26	BH	0.00617	0.00386	0.00233	0.00146	0.00197	0.00123	0.0065
S27	Spring	0.00004	0.00003	0.00002	0.00001	0.00001	0.00001	0.0000
S28	BH	0.00053	0.00033	0.00020	0.00013	0.00017	0.00011	0.0006

for the three populations. The HI result is also presented in Table 4 and graphically illustrated in Fig. 5 a. The oral pathway's HI value ranged from 0.00 to 1.4727, with an average value of 0.1564 (Table 4). It was realized that HI levels for water samples S16 and S26 were greater than 1, indicating that they are unsuitable due to the greater content of NO_3^- (Fig. 5a). These samples represent the most long-term risk to the area's water resource consumers.

The HQ values for the children population varied from 0.00 to 0.00495 with a mean of 0.00053 for human health risk owing to absorption (dermal pathway) (Table 5). For females, the HQ values ranged from 0.00 to 0.00187, with an average of 0.0002 (Table 5). The mean HQ value for males was 0.00017, and the range for HQ values was 0.00 to 0.00158 (Table 5). Similar to the oral health risk exposures, a careful analysis indicated that the dermal HQ results for children, females, and males were < 1 and within the acceptable limits. Moreover, higher HQ values were also observed in S16 and S26 (Table 5). Meanwhile, the risk due to dermal route is far lesser than that due to oral route.

Nevertheless, the chronic risk level for the exposed people in the study region is still in the following descending order: children $>$ females $>$ males, with regard to the dermal HQ values. Similar observations were also reported by previous studies (Unigwe et al. 2022; Panneerselvam et al. 2021; Chen et al. 2017a, b; Karunanidhi et al. 2019). The HI for dermal pathway was computed by summing the dermal HQs of the three populations and the result is also shown in Table 5 and graphically illustrated in Fig. 5 a. The average HI value for the dermal route is 0.0009, ranging from 0.00 to 0.0084 (Table 5). Water samples S16 and S26 had the highest HI values (Table 5; Fig. 5a), indicating that they are the most predisposed that could get unfit for domestic, laundry, and sanitary purposes, which should be effective remedial measures against nitrate pollution not implemented in the study area.

The THI was computed by summing the oral and dermal health risks, to provide a more detailed view of the nitrate health risk for each group in the study region. The THI results for children, females, and males are represented in

Table 6 Total hazard indexes for infants, females, and males due to nitrate contamination

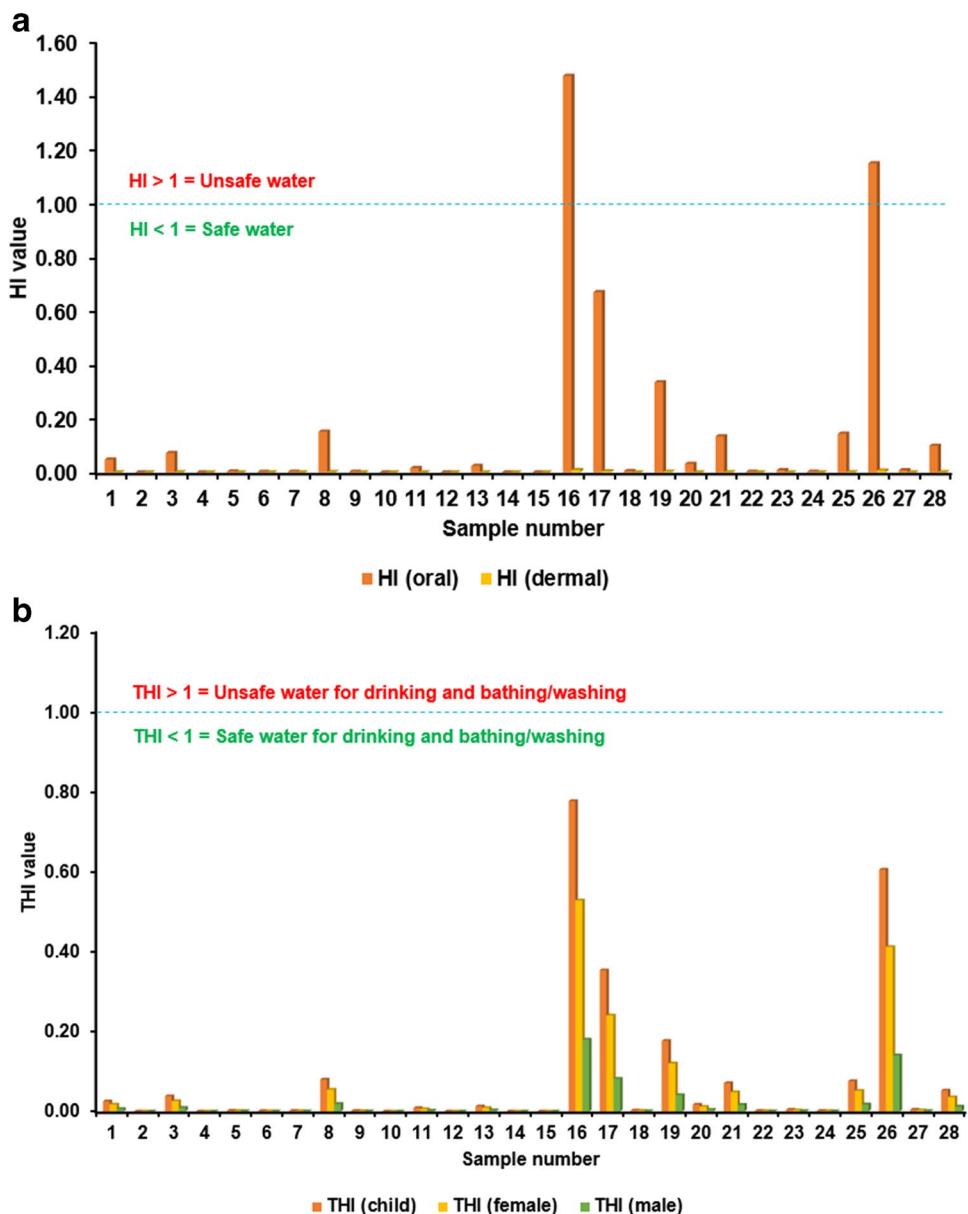
Sample ID	Source	THI (child)	THI (female)	THI (male)	THI (cumulative)
S01	Stream	0.0252	0.0171	0.0058	0.0481
S02	Stream	0.0000	0.0000	0.0000	0.0000
S03	BH	0.0377	0.0257	0.0087	0.0721
S04	BH	0.0000	0.0000	0.0000	0.0000
S05	Spring	0.0017	0.0011	0.0004	0.0032
S06	HW	0.0008	0.0006	0.0002	0.0016
S07	BH	0.0013	0.0009	0.0003	0.0024
S08	BH	0.0797	0.0542	0.0184	0.1523
S09	BH	0.0013	0.0009	0.0003	0.0024
S10	BH	0.0000	0.0000	0.0000	0.0000
S11	BH	0.0084	0.0057	0.0019	0.0160
S12	HW	0.0000	0.0000	0.0000	0.0000
S13	BH	0.0126	0.0086	0.0029	0.0240
S14	Spring	0.0000	0.0000	0.0000	0.0000
S15	Spring	0.0000	0.0000	0.0000	0.0000
S16	BH	0.7750	0.5269	0.1793	1.4811
S17	BH	0.3523	0.2395	0.0815	0.6732
S18	BH	0.0025	0.0017	0.0006	0.0048
S19	BH	0.1761	0.1197	0.0407	0.3366
S20	Stream	0.0168	0.0114	0.0039	0.0321
S21	BH	0.0705	0.0479	0.0163	0.1346
S22	BH	0.0013	0.0009	0.0003	0.0024
S23	Spring	0.0042	0.0029	0.0010	0.0080
S24	HW	0.0013	0.0009	0.0003	0.0024
S25	BH	0.0755	0.0513	0.0175	0.1443
S26	BH	0.6039	0.4105	0.1397	1.1541
S27	Spring	0.0042	0.0029	0.0010	0.0080
S28	BH	0.0520	0.0354	0.0120	0.0994

Table 6 and graphically depicted in Fig. 5 b. Between 0.00 and 0.7750, with a mean of 0.0823, were the THI values for the population of children (Table 6). The average THI for females was 0.0559, while the range for THI values was 0.00 to 0.5269 (Table 6). With a mean of 0.0190, the THI scores for men varied from 0.00 to 0.1793 (Table 6). It is seen from Table 6 and Fig. 5 b that all the three population groups had THI values < 1. However, cumulative THI values were computed by simple summation of the THI values for the three population groups. This was done to generate an overall, wholistic view of the health risk of nitrate in all the water resources in the study area. From the $THI_{cumulative}$ result (Table 6), it was realized that, overall, S16 and S26 had values > 1. This is consistent with the HI_{oral} result. Meanwhile, it was also noticed from Pearson's correlation analysis that the results of the HQ, HI, and THI maintained perfect agreements in their trends (Table 7), even though varied input parameters were used for the health risk computations for the three population groups. Additionally, simple linear regression analysis performed on the health parameters (graphs not presented in this paper) was consistent and

confirmed the findings of the Pearson's correlation matrix. The R^2 values of the linear regression models were all equal to 1.

So far, varied degrees of nitrate contamination have been recorded in the study area. Thus, the users of the water resources are also predisposed to varied levels of toxicity and health risk. Methemoglobinemia, also known as “blue baby syndrome,” a potentially lethal condition, and elevated NO_3^- contents in water are linked (Egbueri and Mgbenu 2020). The “blue baby syndrome” condition is caused by infants consuming excessive amounts of nitrate-contaminated water compared to adults (Tian and Wu 2019; Egbueri and Agbasi 2022b). The “blue baby syndrome” condition typically affects infants younger than 6 months, and it results in bluish skin because nitrate reduces hemoglobin's ability to transfer oxygen through the blood (Skold and Klein 2011). The body weights of babies and children may also decrease significantly due to the elevated NO_3^- concentration in samples S16 and S26 (Tables 3 and 4; Fig. 5). Moreover, the presence of NO_3^- in these samples can encourage the growth of microorganisms, which could have toxic effects on human health (Egbueri 2020). High infant

Fig. 5 Charts showing water resources in safe or unsafe category based on nitrate, **a** HI assessed for oral/dermal, **b** THI assessed for various age groups



mortality, rashes, severe cyanosis, thyroid issues, goiter, congenital deformities, chromosomal anomalies, hypertension, and headaches are some other health consequences linked to excessive long-term nitrate intake through water (Rahman et al. 2021; Egbueri 2020). Unfortunately, the non-carcinogenic risk of NO₃⁻ in drinking water is almost indomitable in many regions of the world where a sizable population relies only on a variety of water sources without first checking for suitability and health-related risks (Rahman et al. 2021).

Toxicity and health risk due to sulfate

The minerals barite (BaSO₄), gypsum (CaSO₄·2H₂O), and epsomite (MgSO₄·7H₂O) all naturally contain sulfates. Many drinking waters’ mineral concentration is influenced by these

dissolved minerals. Sulfate is virtually always present in natural water, where it can be found in varied quantities, ranging between a few tenth mg/L and a several thousand mg/L (WHO 2004). Freshwater bodies can become salinized by sulfate, which is why it is regarded as a pressing environmental issue (Cañedo-Argüelles et al. 2013). However, the problems associated with sulfates in water are completely ignored in many countries of the world; therefore, the impact has remained understated. According to Sharma and Kumar (2020), at a global scale, research on sulfate contamination in water system is limited to few numbers only. Although no health-based recommendation values have been proposed for SO₄²⁻ in drinking water, as concentrations in the water rise above 500 mg/L, there is a greater chance that complaints may arise from discernible bitter taste (WHO 2004). In this

Table 7 Pearson's correlation matrix showing the perfect agreements in trends of the key health risk parameters for nitrate

Parameter	HQ _{oral} (infant)	HQ _{oral} (female)	HQ _{oral} (male)	HI (oral)	HQ _{dermal} (infant)	HQ _{dermal} (female)	HQ _{dermal} (male)	HI (dermal)	THI (child)	THI (female)	THI (male)	THI (cumulative)
HQ _{oral} (infant)	1	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**
HQ _{oral} (female)		1	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**
HQ _{oral} (male)			1	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**
HI (oral)				1	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**
HQ _{dermal} (infant)					1	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**
HQ _{dermal} (female)						1	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**
HQ _{dermal} (male)							1	1.000**	1.000**	1.000**	1.000**	1.000**
HI (dermal)								1	1.000**	1.000**	1.000**	1.000**
THI (child)									1	1.000**	1.000**	1.000**
THI (female)										1	1.000**	1.000**
THI (male)											1	1.000**
THI (cumulative)												1

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

study, the health-related concerns of SO_4^{2-} in drinking water resources are evaluated. As have been mentioned earlier, the concentrations of SO_4^{2-} in the water resources ranged between 7.00 and 130.00 mg/L (Table 3). In this area, it was noticed that SO_4 ions were above the national SON (2015) standard limit of 100 mg/L in some water samples (S06 and S24). It was also observed that samples S3, S19, and S28 had elevated SO_4^{2-} concentrations close to the 100 mg/L benchmark. Meanwhile, all of the samples were found within the acceptable limit of 250 mg/L by WHO (2017).

The analyzed water resources, which have been contaminated with SO_4^{2-} , may leave a bad taste when ingested. However, the USEPA and other recognized bodies do not employ organoleptic qualities, such as taste, odor, and color, when creating primary water standards. Organoleptic characteristics can, however, be employed in the creation of secondary water standards for drinking purpose. Based on taste qualities, the USEPA (1984) set a secondary sulfate standard of 250 mg/L for drinking water. The taste threshold concentrations for various common sulfate salts have been reported, notwithstanding the dearth of genuine experimental evidence on the taste threshold for sulfate (WHO 2004; Zoeteman 1980; NAS 1977; Whipple 1907). Sulfate may also aggravate concrete deterioration and distribution system corrosion (WHO 2004; Sharma and Kumar 2020; Egbueri 2022; Omeka et al. 2022). Scaling in water pipes from sulfate can lead to pipes with reduced diameter (i.e., clogged pipes). Cleaning fabrics/cloths becomes more challenging when sulfate is combined with chlorine bleach. As a result, excessive sulfate in water supplies causes issues for both materials and human health.

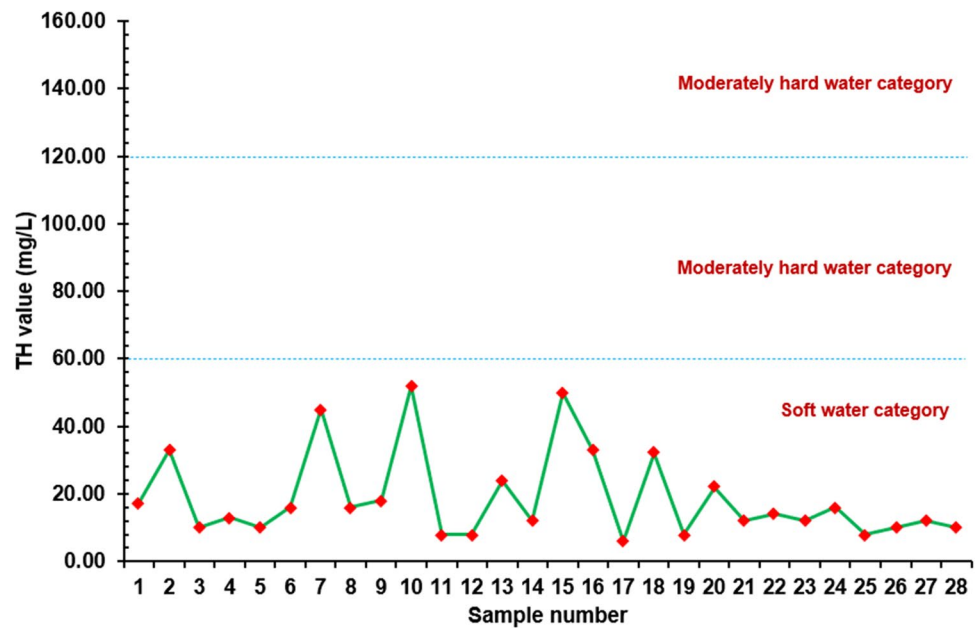
With respect to the Nigerian standard limit, it is adjudged that the water samples with elevated SO_4^{2-} could pose some health issues to people who use the water. The most typical side effect of being exposed to high SO_4^{2-} concentrations is a laxative impact (Moore 1952; WHO 2004; Sharma and Kumar 2020). Sulfate's laxative effects could be more dangerous for infants. As a result, water containing sulfates should not be used for baby purposes, such as drinking and feeding (Sharma and Kumar 2020). Unfortunately, acute SO_4^{2-} exposures can cause laxative effects, although how severe they are may depend on the sulfate salt and how the SO_4^{2-} amount is taken into the human system. For instance, Na_2SO_4 has a weaker laxative effect than MgSO_4 (Morris and Levy 1983). However, research by Peterson (1951) revealed that the SO_4^{2-} concentration that was connected to a laxative reaction reduced as the Mg^{2+} content in water increased. Although SO_4^{2-} is not generally known to contribute to toxicity, it can lead to a number of health problems such as dehydration, catharsis, and diarrhea when consumed in larger amounts through water ingestion or food absorption. When the intestinal contents' osmolality is higher than that of the interstitial fluids, osmotic diarrhea (soft stool) is

the most outcome. The osmotic-induced diarrheas would be frequently accompanied by poorly absorbed dissolved substances such MgSO_4 , sorbitol, or lactulose (Stipanuk 2000). The presence of Mg^{2+} and SO_4^{2-} in the water supply under study would exacerbate the diarrheal response because both ions have osmotically active properties. Also, sulphaemoglobin and methaemoglobin levels in the human body system may vary as a result of increased SO_4 absorption from the water resources (Sharma and Kumar 2020). The amount of SO_4^{2-} consumed affects how much of it is absorbed from the human intestine. According to Morris and Levy (1983) and Cocchetto and Levy (1981), the kind of cation associated with the SO_4^{2-} may also affect absorption, with MgSO_4 having lesser chances of being absorbed than Na_2SO_4 . High dosages of SO_4^{2-} that cause catharsis, however, have also been documented to frequently surpass intestinal absorption capacity and can be eliminated through human feces (WHO 2004). The primary determinants of the amount of SO_4^{2-} reaching the human colon include the nature of the diet and intestinal absorption (WHO 2004), mostly carried out via the upper digestive system (Florin et al. 1991). With the continuous ingestion of the sulfate-contaminated water in the study area, human blood, but not tissues, may readily distribute inorganic SO_4 (WHO 2004). This is because the majority of SO_4 present in human tissues is organic and biosynthetically integrated into macromolecules. However, studies have also shown that with increments in time, humans may adjust to increasing SO_4^{2-} amounts (USEPA 1985).

Toxicity and health risk due to total hardness

The phrase “water hardness” refers to the quantity of ions, particularly sulfate and/or carbonate salts of Mg and Ca, that are present in water. The human body uses hardness minerals (e.g., Ca, Mg, Cl, SO_4) for a number of vital metabolic processes or functions. Thus, some experts who study water quality disagree that water hardness poses a health risk. But because of mineral buildup on plumbing fixtures and sub-par soap efficacy, hard water is often considered a nuisance. Before using water as industrial water, drinking water, or public water, this property is crucial to investigate. The water TH in this study ranged between 6.00 and 52.00 mg/L (Table 3). A number of TH classification systems have been reported in the literature. The most popular classification system seems to be the one reported by Sawyer and McCarthy (1967), McGowan (2000), and Botheju et al. (2021), illustrated in Fig. 6. Based on this classification scheme, all the water samples were classified as soft water resources. The U.S. Department of Interior and the Water Quality Association reported a scheme for classifying TH, having five categories as thus: soft water (TH = 0–17.1), slightly hard water (TH = 17.1–60 mg/L), moderately hard water

Fig. 6 A TH classification system showing that the water samples are soft waters



(TH = 60–120 mg/L), hard water (TH = 120–180 mg/L), and very hard water (TH > 180 mg/L). With respect to this system, 67.86% of the water samples were grouped as soft waters whereas 32.14% were grouped as slightly hard waters. Ca and Mg bicarbonate salts, as opposed to chloride, sulfate, nitrate, phosphate, and silicate salts, are what give water its temporary hardness level (Rapant et al. 2020; Botheju et al. 2021). However, samples of hard water have higher concentrations of Ca and Mg chloride, sulfate, nitrate, phosphate, and silicate salts. Also, Rapant et al. (2020) reported a hardness classification as thus: soft water (Mg^{2+} 6.05 mg/L and Ca^{2+} 20.70 mg/L) and hard water (Mg^{2+} 26.40 mg/L and Ca^{2+} 70.12 mg/L).

The majority of earlier research have interpreted water hardness based on secondary standards including the color, odor, taste, corrosivity, staining, and foaming qualities of water because there are no clear basic standards for health-related concerns of water hardness. Nevertheless, there are also several studies that have investigated the positive and negative human health implications of water hardness on heart disease, stroke, respiratory tract, gastrointestinal and cerebrovascular diseases, and cancers, etc. Over 60 years ago, the connection between drinking water hardness and the occurrence and death due to cardiovascular disease was first reported (Kobayashi 1957). Numerous pieces of international literature have corroborated this association throughout the last few decades (Rubenowitz–Lundin and Hiscock 2005; Catling et al. 2008; Gianfredi et al. 2017; Jiang et al. 2016; Rapant et al. 2017, 2020). Recent epidemiological research by Rapant et al. (2020) examined the ecological type and connection between drinking water hardness, public health status, and average life expectancy of Slovak

Republic residents. Additionally, numerous research works have discussed the beneficial impacts of Ca and Mg in water on the occurrences and mortality of the metabolic syndrome or diabetes mellitus (Yang et al. 1999; Joslyn et al. 1990; Naumann et al. 2017). The current study has indicated that the analyzed water resources are in two main categories—soft water and slightly hard water—in terms of hardness values. Insights drawn from previous studies indicate that both soft water and hard water have their human health implications. For instance, the inhabitants who use the waters and are diagnosed with cancer could have their ailment impacted by the soft drinking waters (i.e., the low hardness waters) analyzed in this study (Ahn et al. 2007; Yang et al. 1998, 2000; Butler et al. 2010; Rapant et al. 2020). Moreover, the soft waters could influence amyotrophic lateral sclerosis, high blood pressure, neurological disturbances, preeclampsia in pregnant women (Rosborg and Kozisek 2020; Rapant et al. 2020), bone/teeth development issues in children, and bone fracturing (Huang et al. 2018, 2019; Dahl et al. 2013).

On the other hand, the analyzed water resources classified in the slightly hard water category could provide much more health benefits than the soft waters. Previous studies have indicated that natural waters with elevated hardness could improve the health status of humans than low hardness natural waters, in terms of cardiovascular, digestive tract, oncological, and respiratory system well-being (Rapant et al. 2020). It has also been reported that longer life expectancy is associated to drinking water with elevated hardness than with low hardness drinking water (Rapant et al. 2020). This discrepancy could be linked to deficiency of Ca^{2+} and Mg^{2+} mineral salts in soft waters (Rapant et al. 2020). Previously, Mukate et al. (2019) reported that not only is excess of physicochemical

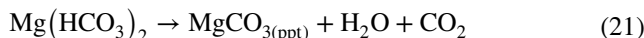
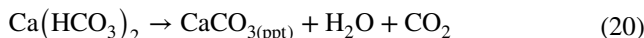
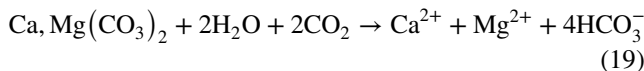
parameters in water hazardous to human health, but also is the deficiency of some mineral elements. With this perspective, they developed an integrated water quality index that considers the excessive and deficient concentration levels of elemental and mineral compositions of water.

Computational and graphical methods for unraveling contamination mechanisms

R-mode hierarchical clustering

Various contamination sources influence the suitability of the drinking water resources in the study area. It is revealed in this study that the plotted R-mode dendrogram has four main contamination cluster groups as illustrated in Fig. 7. The first cluster (C I) comprised of TH, Ca, EC, and Mg. This physicochemical cluster represents a group predominantly influenced by geogenic processes. The strong association between these variables indicates that the bivalent cations (Ca²⁺ and Mg²⁺) were possibly originated from the weathering and dissolution of silicate rocks and minerals in the study area. Similar observations have been reported in the literature (Omeka and Egbueri 2022 2022; Subba Rao and Chaudhary 2019; Mukate et al. 2019). While Ca and Mg are known to influence the TH of the water resources, the Ca, Mg ions, and TH seem to influence the water conductivity

(Egbueri and Agbasi 2022a). Nevertheless, as indicated by the two sub-clusters formed in the first cluster (Fig. 7), Ca²⁺ concentration is thought to influence the TH and EC more than the Mg²⁺ concentration. pH and HCO₃⁻ constituted the second cluster (C II), indicating that the HCO₃⁻ might be influencing the water acidity. The interaction between the pH and HCO₃⁻ in this cluster could also favor carbonic and silicate weathering and the elevation of the water hardness in terms of Ca(HCO₃)₂, CaCO₃, and MgCO₃ (Eqs. 19–21).



The third cluster (C III) is made up of K and NO₃ (Fig. 7). This cluster typically represents human-induced and nutrient enrichment processes. The possible origins of this cluster include organic wastes (e.g., sewage) and agricultural inputs (e.g., inorganic NPK fertilizers) from the inhabitants of the research area. Previous studies reported that human-dominated areas are particularly prone to having high levels of water nitrate, derived from sources such as animal waste, industrial waste, residential wastewater, septic tanks, and nitrogen-rich fertilizers (He et al. 2018; Egbueri 2020; Egbueri and Agbasi 2022b; Li et al. 2017b; Karunanidhi et al. 2020b). Other anthropogenic activities' sources of nitrate in water include fossil fuel combustion (which occurs through atmospheric deposition) and nitrogen-fixing crops in naturally vegetated areas (Ward et al. 2018; Gutiérrez et al. 2018). Owing to the association observed in this cluster, the formation of KNO_{3(aq)} is anticipated in the water environment. Also, nitrification processes could also play a contributory role in the NO₃ formation. Ammonia and ammonium from sewage and certain inorganic wastes could also react to facilitate the production of NO₂⁻, HNO_{3(aq)}, and NO₃⁻ in the water in oxic environment (Eqs. 22–26).

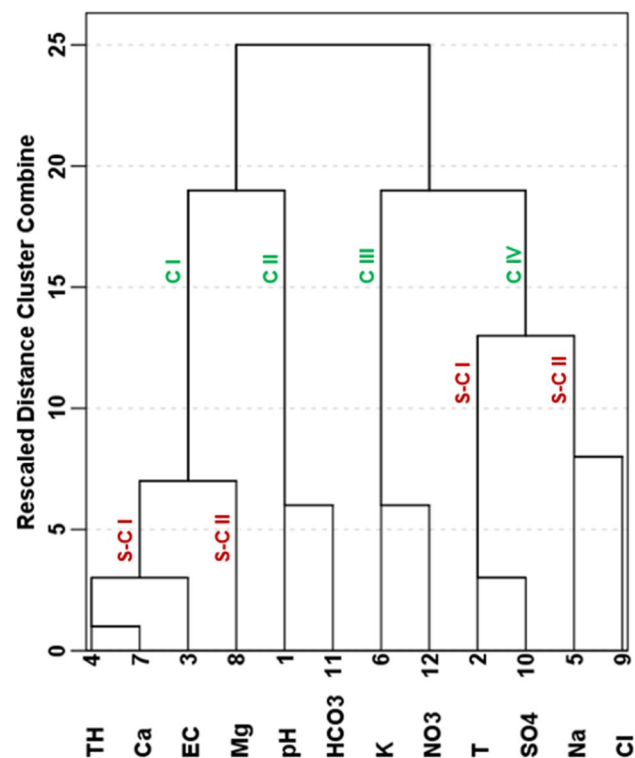
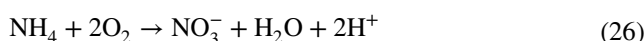
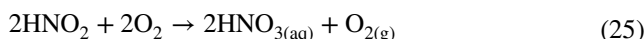
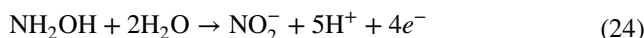
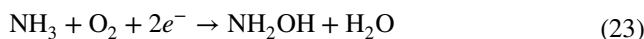
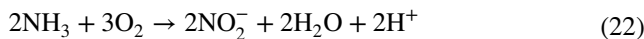
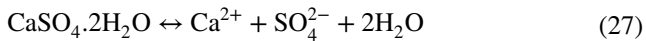


Fig. 7 R-mode hierarchical dendrogram showing different clusters of the physicochemical parameters

The fourth cluster (C IV) is composed of water temperature (T), SO₄²⁻, Na⁺, and Cl⁻ (Fig. 7). This association

is typically geogenic and is divided into two sub-clusters– (T–SO₄²⁻) and (Na⁺–Cl⁻). The sub-cluster 1 is apparently indicative of the role T plays in sulfate enrichment processes within the water environment. For instance, it is well known that temperature enhances most chemical reactions. It has also been indicated in this study that SO₄²⁻ is the most abundant anion whereas Na⁺ and Ca²⁺ are the predominant cations. The addition of SO₄²⁻ to the water from varied sources (such as sulfide oxidation, biogeochemical degradation of organic matter, and acid precipitation from humic soils in the area) could be facilitated by increasing the T values. Also, the formation of Na₂SO₄ and CaSO₄ salts and their dissociation to form Na–SO₄ and Ca–SO₄ water types in the area (Egbueri et al. 2019) could have been much impacted by the T. For instance, the reversible reaction (Eq. 27) illustrating the linkage between CaSO₄ and water could performance better with warmer temperature.



The sub-cluster 2 of the C IV (represented as Na⁺–Cl⁻) suggests that salinization reaction takes place in the water environment. The associations of T, SO₄²⁻, Na⁺, and Cl⁻ in the fourth cluster also indicates that T and SO₄²⁻ influence the water salinization to substantial extents (Cañedo-Argüelles et al. 2013). Moreover, ion exchanges between the SO₄²⁻ and Cl⁻, related to the cations (Na⁺, Ca²⁺, and Mg²⁺), could also be impacted upon by the T.

It is important to recall that the pH of all the samples indicated that they are all acidic. NO_x and SO₂ emissions in the area could also be caused by rising fossil fuel burning. Because of this, acidic particulate matter deposition and acid rain in the form of HNO₃ and H₂SO₄ could be frequent in the studied area (Han and Liu 2006; Liu and Han 2020; Ding et al. 2013; Wu et al. 2012). With excessive HNO₃ and H₂SO₄ in the water environment, the breakdown of rocks and minerals can be facilitated (Wu et al. 2012; Xu and Liu 2007; Han et al. 2019), which could result in the transfer and accumulation of more loads of ions and other substances into the water systems.

Varimax-rotated factor analysis

In this research, five factor classes were generated following the hydrogeochemical dataset transformation using the orthogonal Varimax rotation plus the Kaiser normalization technique. The five factor classes had their cumulative variance at 82.792%. Factor classes F1, F2, F3, F4, and F5 varied by 26.440%, 18.542%, 13.960%, 12.833%, and 11.017%, respectively (Table 8(a)). Table 8(b) represents the principal component transformation matrix showing the associations of the five factor classes while Fig. 8 is a 3D component plot depicting the spatial associations between the physico-chemical parameters in the various factor classes. Overall, it was realized that the results of the Varimax-rotated factor analysis are consistent with those of the R-mode hierarchical

Table 8 Factor classes, their extractions and transformation matrix

sParameter	a) Factor classes and their extractions						b) Component transformation matrix showing the associations of factor classes					
	F1	F2	F3	F4	F5	Extraction commonality	Factor class	F1	F2	F3	F4	F5
pH	0.242	-0.263	0.758	-0.215	-0.186	0.783	F1	0.795	0.472	0.299	0.071	0.224
T	0.144	0.888	-0.026	0.033	0.132	0.829	F2	-0.404	0.675	-0.315	0.444	0.291
EC	0.848	0.425	0.103	0.048	0.004	0.912	F3	0.411	-0.325	-0.569	0.600	-0.206
TH	0.951	-0.011	0.043	0.026	0.014	0.907	F4	-0.185	-0.091	0.690	0.638	-0.273
Na	0.122	0.384	0.409	0.249	0.624	0.782	F5	-0.031	-0.456	0.111	0.176	0.865
K	0.147	0.289	0.006	0.833	0.025	0.800						
Ca	0.956	0.069	0.007	0.001	0.116	0.932						
Mg	0.700	0.059	0.311	0.031	0.100	0.601						
Cl	0.097	0.038	-0.113	-0.110	0.915	0.874						
SO ₄	0.105	0.922	0.079	0.014	0.054	0.870						
HCO ₃	0.076	0.229	0.876	-0.053	0.114	0.841						
NO ₃	-0.067	-0.206	-0.192	0.847	-0.049	0.803						
Eigenvalue	3.173	2.225	1.675	1.540	1.322	-						
% of variance	26.440	18.542	13.960	12.833	11.017	-						
Cumulative %	26.440	44.982	58.942	71.776	82.792	-						

Extraction method: Principal component analysis. Rotation method: Varimax with Kaiser Normalization

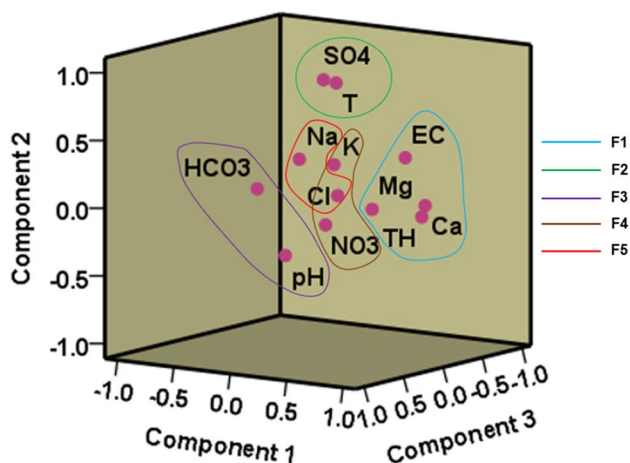


Fig. 8 3D component plot showing the spatial association between the physicochemical parameters

dendrogram. The observed consistencies are summarized as follows:

- The first factor class (F1) has high loadings on EC, TH, Ca, and Mg (Table 8(a); Fig. 8). The F1 is a typical representation of the C1 in the hierarchical dendrogram (Fig. 7). In line with the C1, the water contamination mechanism represented by the F1 is mainly influenced by geogenic processes.
- The second factor class (F2) has high loadings on T and SO_4 (Table 8(a); Fig. 8). The F2 typically represents the sub-cluster 1 of the C4 in the hierarchical dendrogram (Fig. 7). In line with the sub-cluster 1 of the C4, the water contamination mechanism represented by the F2 is mainly influenced by geogenic processes. With reference to the common lithologies in the study area, it suffices to say that carbonate and evaporitic weathering do not have much controls on SO_4 addition to the waters. Also, since Na is the most common cation and highly soluble and reactive than Ca, the predominance of Na_2SO_4 over CaSO_4 and their water types is expected in the study area (Egbueri et al. 2019).
- The third factor class (F3) has high loadings on pH and HCO_3 (Table 8(a); Fig. 8). The F3 represents, typically, the C2 in the hierarchical dendrogram (Fig. 7). In line with the C2, the water contamination mechanism represented by the F3 is mainly influenced by geogenic processes.
- The fourth factor class (F4) has high loadings on K and NO_3 (Table 8(a); Fig. 8). The F4 is a typical representation of the C3 in the hierarchical dendrogram (Fig. 7). In line with the C3, the water contamination mechanism represented by the F4 is mainly influenced by anthropogenic processes and inputs. Moreover, the F4 was noticed to have a closer proximity with Cl^- (Fig. 8). This suggests that Cl^- in the water could also be influenced by

anthropogenic sources such as industrial waste and agricultural inputs (Egbueri 2018; Ge et al. 2021).

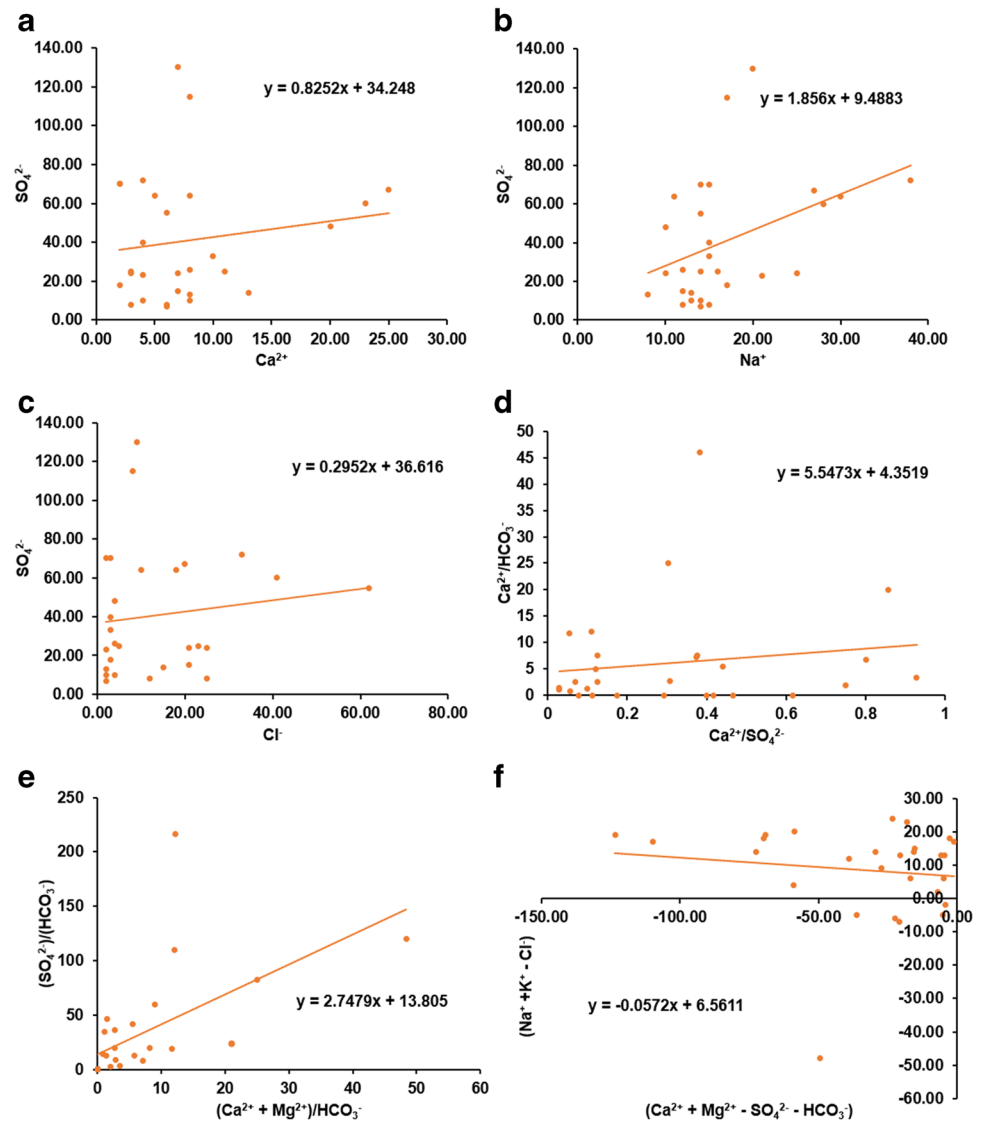
- The fifth factor class (F5) has high loadings on Na and Cl (Table 8(a); Fig. 8). The F5 typically represents the sub-cluster 2 of the C4 in the hierarchical dendrogram (Fig. 7). In line with the sub-cluster 2 of the C4, the water contamination mechanism represented by the F5 is mainly influenced by geogenic processes like salinization and weathering of sodic minerals in the underlying geologic units. However, with F5 having the least percentage of variance and for the fact that there are no known evaporite deposits in the study area, it is reasoned that evaporitic processes do not control much of the Na^+ and Cl^- addition to the water resources.

Nevertheless, it is important to mention that, with respect to the percentages of variance of the five factor classes, the importance and significances of the contamination mechanisms in the study area increased in the order $\text{F5} < \text{F4} < \text{F3} < \text{F2} < \text{F1}$. This order also represents the trend by which the important information in the hydrogeochemical data is explained. So far, it has been indicated that geogenic (natural) processes are much more predominant in the study area than anthropogenic influencers. This fact is reflected in the concentration levels of the physicochemical parameters shown in Table 3. In a scenario where human-induced activities predominate over geogenic processes, the concentrations of these variables could be much higher. Also, it is thought that the anthropogenic influences that were reflected by the C3 and F4 are still not intense in the study area. However, with continuous increase in urbanization and population growth rates, plus lack of effective water resources management strategies, in the study area, the imprints of human activities could greatly be pronounced in the area's hydrogeochemistry. With increments in the variables' concentration values, their toxicity and human health risks would adjust accordingly.

Bivariate/multivariate graphical plots

As shown in Fig. 9, ten graphical models were utilized to express and unravel the relationships between the analyzed hydrogeochemical variables and further provide insights on the water contamination mechanisms of the study area. Linear equations of the considered variable sets were also presented in each graphical model. Careful observations across the graphs confirms that multiple and mixed processes are responsible for the contamination and evolution of the natural water resources. The common hydrogeochemical processes indicated by the graphical plots are silicate rock weathering and mineral dissolution

Fig. 9 Graphical bivariate/multivariate plots for unravelling the trends between the parameters (measured in mg/L) and water contamination mechanisms



(Fig. 9a–h); carbonic acid reactions and silicate hydrolysis (Fig. 9d, e, f); neutralization, redox reactions, and ion exchange processes (Fig. 9a, b, d, e–h), and human-induced activities in the study area (Fig. 9i, j).

Conceptual diagrammatic model of processes affecting contamination mechanisms

A diagrammatic model of the processes affecting the contamination mechanisms of the water resources is shown in Fig. 10, based on the information drawn from the R-mode hierarchical cluster analysis, Varimax-rotated factor analysis, and graphical plots. A careful observation made from Fig. 10 confirms that natural geogenic processes are the most common in the area. The interactions between the various variables and processes are indicated in the model.

Computational algorithms for predictive modeling of TH, SO_4 , and NO_3

Radial basis function neural networks

Due to their simplicity of use and effectiveness in predicting significant traits and variables in situations with limited datasets, ANNs (such as RBF-NN) and MLR are usually seen as more cost-effective methods in prediction research (Ozel et al. 2020; Egbueri 2022). As a result, such modeling can cover a large number of data points while saving a lot of money. The performance summary result of the RBF-NN modeling of the TH, SO_4 , and NO_3 contents in the water resources is contained in Table 9. Scatterplots showing the relationships between the measured and predicted variables and their R^2 values are illustrated in Fig. 11. The SSE and RE values are shown in Table 9

Fig. 9 (continued)

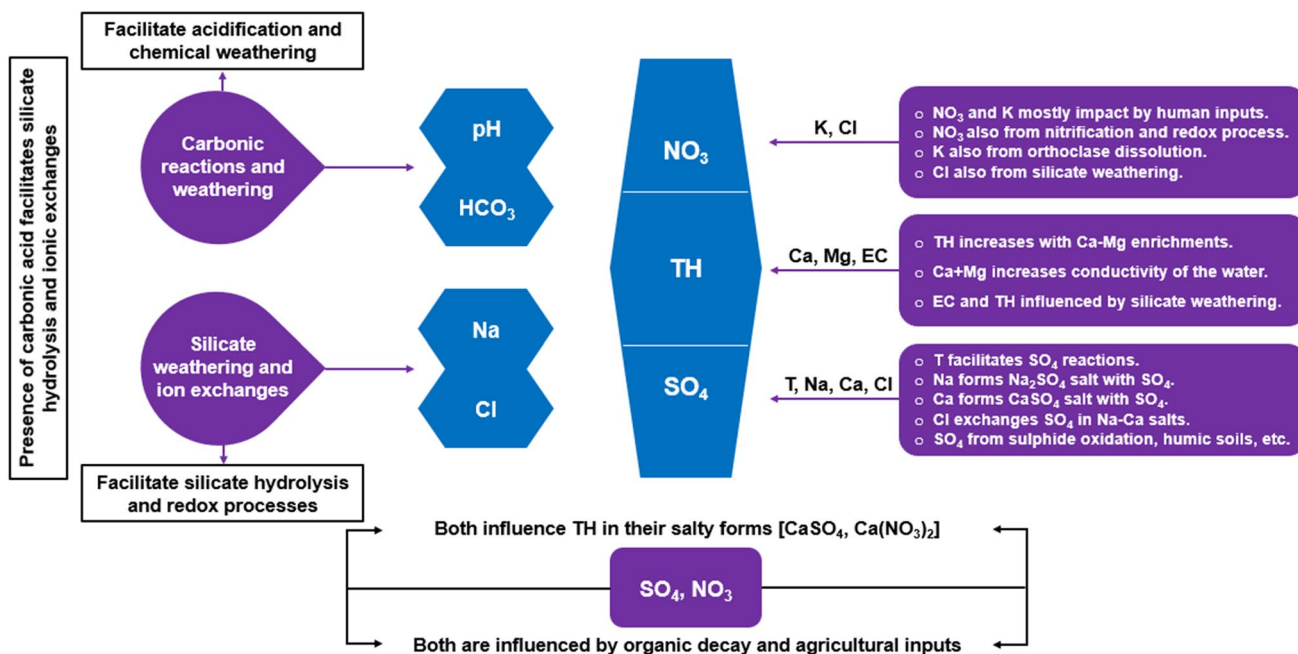
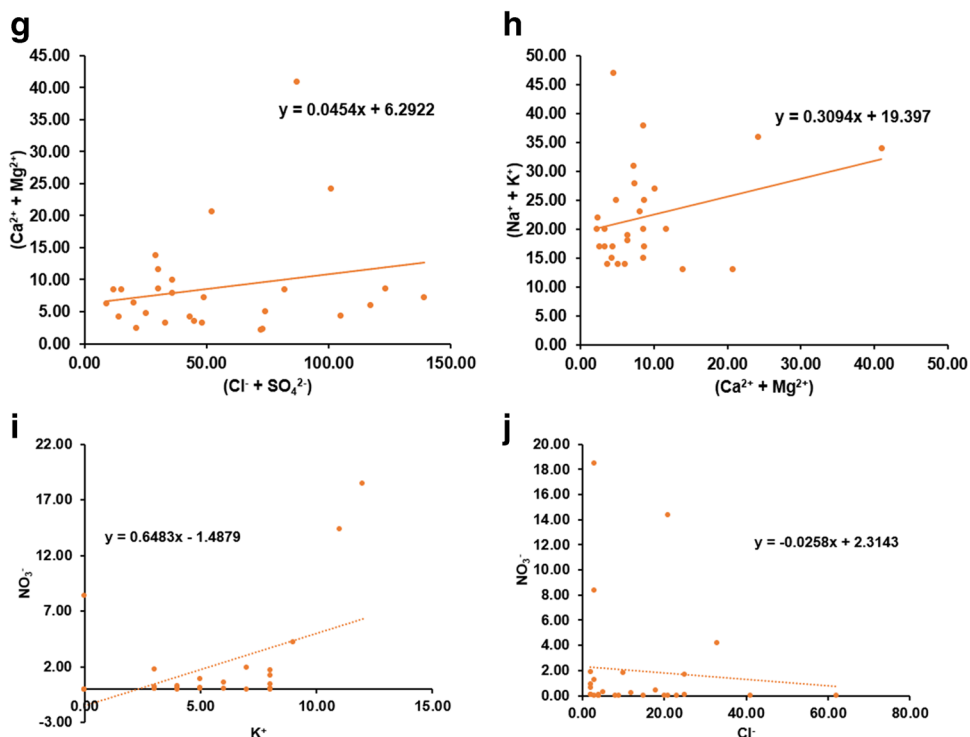


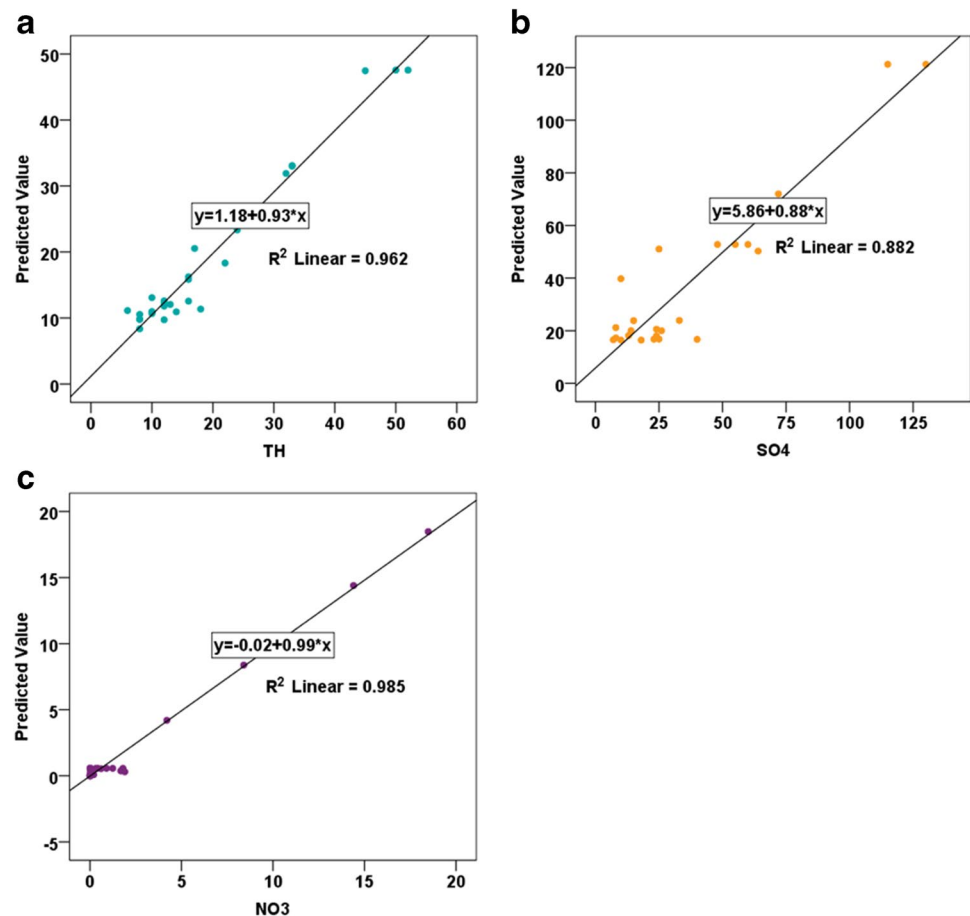
Fig. 10 Conceptual diagrammatical model of the main processes affecting the water contamination

whereas the residual error plots are depicted in Fig. 12. Interestingly, all the statistical error values proved that the RBF-NN predictive modeling generally had low error values. However, it was noticed that the predictions of NO₃ and TH outperformed that of SO₄, marginally (Table 9). Meanwhile, it is apparent that the predictions performed very well, with all R² values seen

to be above 0.800. Additionally, sensitivity analysis was carried out to comprehend how the input parameters affected how well the RBF-NN models performed. For TH prediction, Ca, EC, and Mg (Ca > EC > Mg) were the most sensitive elements that impacted on the TH RBF-NN structure and performance (Fig. 13a). This observation is seen to be consistent with the

Table 9 Statistical metrics for validating the RBF-NN and MLR models

Parameter	RBF-NN training		RBF-NN testing			Residual error plot	MLR			
	SSE	RE	SSE	RE	R^2		R^2	R^2_{adj}	SEE	R
TH	0.239	0.024	0.372	0.071	0.962	Figure 12a	0.897	0.826	5.435	0.947
SO ₄	0.983	0.089	0.455	0.409	0.882	Figure 12b	0.697	0.489	23.043	0.835
NO ₃	0.079	0.008	0.076	1.007	0.985	Figure 12c	0.378	0.049	4.589	0.615

Fig. 11 Scatterplots and R^2 values for the RBF-NN prediction of **a** TH, **b** SO₄, and **c** NO₃

previous findings of the hierarchical cluster analysis and factor analysis, which had suggested that the TH was noticed to be mainly influenced by the Ca, EC, and Mg.

The SO₄ prediction, architecture, and performance were majorly influenced by the T and Ca (Fig. 13b). This finding is also consistent with the reports of the cluster and factor analyses. T has been shown to influence the presence of SO₄²⁻ in the water resources. Based on the obvious association noticed between the SO₄ RBF-NN and Ca, formation of CaSO₄ could also be impacted by T. CaSO₄ has also been reported to influence water TH (Fig. 10). Finally, as graphically illustrated in Fig. 13 c, the NO₃ prediction and its model performance were mostly determined by K, Mg, HCO₃, SO₄, Ca, and TH (K > Mg > HCO₃ > SO₄ > Ca > T H). Again, the reports of the cluster and factor analyses

are seen to be consistent with this sensitivity analysis. K⁺ was identified to have the highest influence on the NO₃ RBF-NN prediction, confirming that they both have very close association. Nevertheless, the NO₃ RBF-NN predictive model indicated that other active variables could have linkages with the presence of NO₃⁻ in the water resources. For instance, nitrate, in the form of Ca(NO₃)₂, has also been reported to influence the water TH (Fig. 10). The association seen between NO₃ and SO₄ in this RBF-NN model could be indicative of similar origins such as organic wastes and agricultural inputs (Fig. 10).

Fig. 12 Residual error plots for the RBF-NN prediction of **a** TH, **b** SO₄, and **c** NO₃

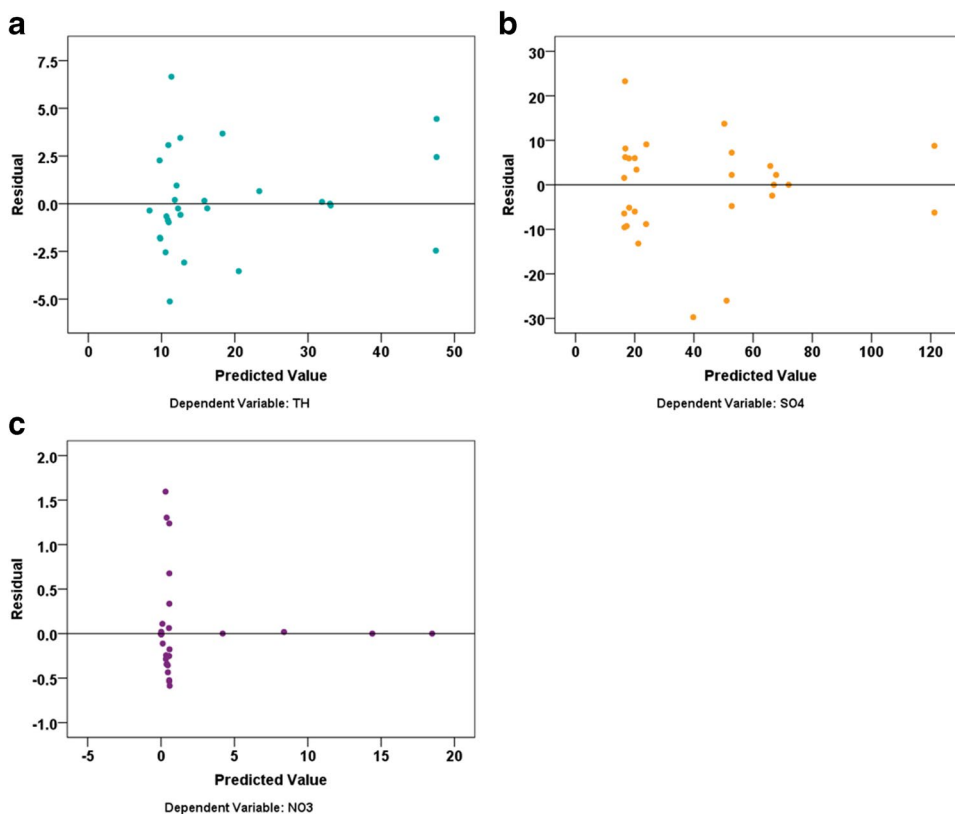
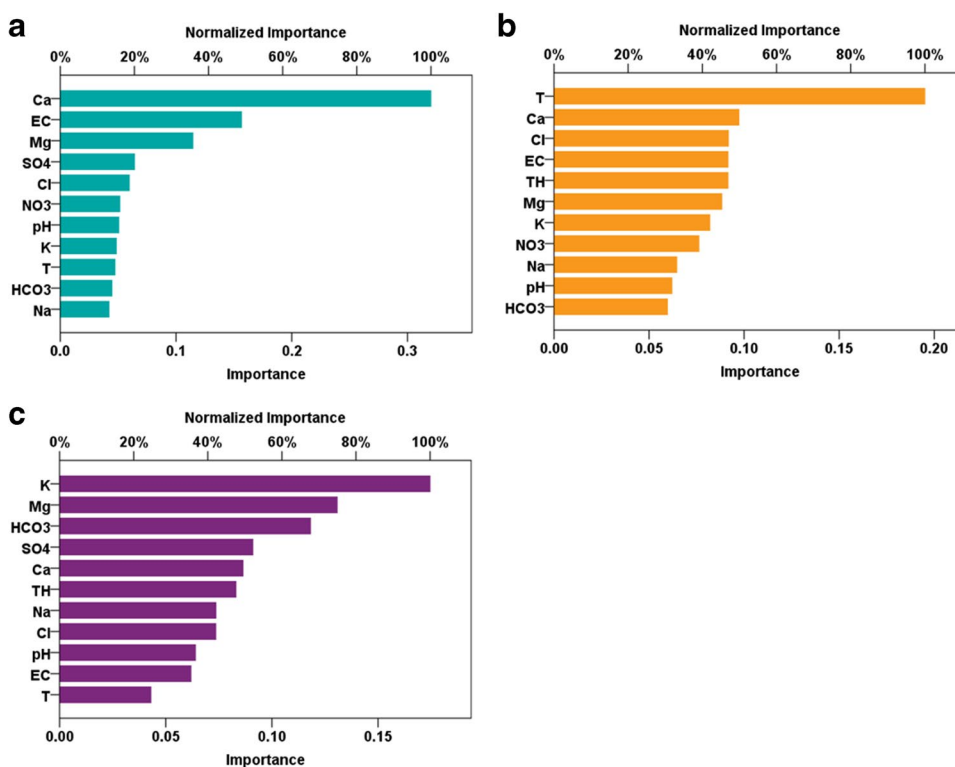


Fig. 13 Bar charts illustrating the sensitivities of input variables on the RBF-NN prediction of **a** TH, **b** SO₄, and **c** NO₃



Multiple linear regression models

Table 9 also includes the results of the statistical indicators (metrics) used to summarize the MLR models' performances.

Figure 14 shows the R^2 parity graphs for the models. Moderate–high R , R^2 , and R^2_{adj} values were achieved, notably for TH and SO_4 , demonstrating that the MLR models were effective in the parameter prediction. The obtained R , R^2 , and R^2_{adj} values demonstrated that the MLR models’ performances followed the following order: $TH > SO_4 > NO_3$ (Table 9). Mathematically, the MLR models created in this study are represented by Eqs. 28–30.

$$y_{TH} = 26 + 2.171pH + (-1.441T) + 0.152EC + (-0.076Na) + 0.117K + 1.719Ca + (-0.349Mg) + 0.002Cl + 0.019SO_4 + (-0.206HCO_3) + 0.099NO_3 \tag{28}$$

$$y_{SO_4} = -244.658 + (-3.713pH) + 10.972T + 0.843EC + 0.599Na + 0.182K + (-0.295Ca) + (-0.875Mg) + (-0.033Cl) + 1.094HCO_3 + (-0.626NO_3) + 0.346TH \tag{29}$$

$$y_{NO_3} = 11.728 + (-1.779pH) + (-0.071T) + (-0.053EC) + (-0.098Na) + 0.801K + (-0.059Ca) + 0.107Mg + (-0.004Cl) + 0.232HCO_3 + 0.071TH + (-0.025SO_4) \tag{30}$$

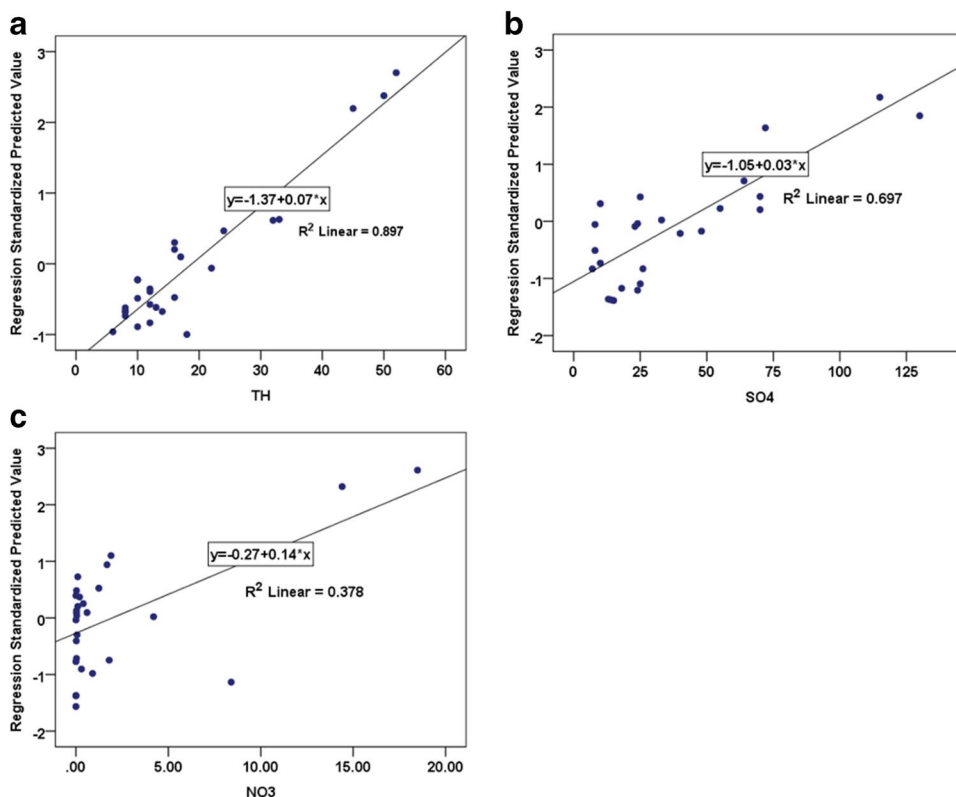
These equations show how the input parameters affect or are sensitive to the MLR predictions. Based on Eq. 28, it was realized that pH, Ca, and EC influenced the model more than the other variables. The linkages between TH, Ca, and

EC are again reflected in this MLR model. For the SO_4 prediction, T, HCO_3 , and EC influenced the MLR model, with T having the most impact (Eq. 29). This result also reflects the consistency between the hierarchical clustering, factor analysis, and RBF-NN in linking SO_4 and T closely than other parameters. Lastly, the NO_3 MLR prediction (Eq. 30) shows that the model was impacted by K, HCO_3 , and Mg, with K having the highest impact factor on the water NO_3 . Interestingly, this result also demonstrates the consistency of the RBF-NN, factor analysis, and hierarchical clustering in correlating SO_4 and T more closely than other parameters.

Comparison of the performances of the RBF-NN and MLR models

In order to lower the cost of monitoring and assessing water quality, it is essential to develop trustworthy models that can predict the characteristics of interest. The creation of precise predictive models is also essential since they serve as crucial decision-making tools for sustainable water quality management for a variety of uses. The RBF-NN and MLR models created for this study functioned well, as shown thus far by the validation metrics. Nevertheless, there were some differences among the modeling approaches. The MLR models produced superior forecasts for the water quality parameters in the following order: $TH > SO_4 > NO_3$. The RBF-NN models, on the other hand, offered higher forecasts for the water

Fig. 14 Scatterplots and R^2 values for the MLR prediction of **a** TH, **b** SO_4 , and **c** NO_3



quality parameters in the following order: $\text{NO}_3 > \text{TH} > \text{SO}_4$. RBF-NN > MLR is the model performance order based on R^2 values and other model error measures (Table 9). This can be because the two modeling methodologies have different sensitivity levels to the physicochemical input parameters used in the predictive modeling. Compared to the MLR, the RBF-NN appears to be more sensitive. Similar observation had been reported by Egbueri and Agbasi (2022c), who predicted irrigation water quality indices using MLR and RBF-NN. Regardless of the performance variations seen in this study, it is important to recall that there are noted significant agreements in their predictions and sensitivity analysis. This suggests that the both models will be useful for water quality monitoring and assessment in the study area. The models also provide useful insights for national and international researchers.

Conclusions

For the comprehensive assessment and predictive modeling of the contamination mechanisms, toxicity, and human-related health risks of water quality in Ojoto region in southeast Nigeria, multiple data-driven indexical, graphical, and soft computational models have been successfully integrated and implemented. Descriptive study approach was also adopted. The study revealed that majority of the physicochemical parameters tested for were within their SON (2015) and WHO (2017) satisfactory limits for natural drinking water other purposes. However, TH, SO_4 , and NO_3 concentrations in some water locations were not within the stipulated satisfactory limits. The nitrate health risk assessment indicated that some locations pose substantial chronic health risk due to water ingestion to the children, females, and males in the study area. However, insignificant health risk was recorded for dermal absorption route for all the populations. Overall, the health risk due to ingestion was higher than the risk due to absorption. Also, the human populations in the area are predisposed to nitrate health risk in the following order: children > females > males. The newly introduced cumulative nitrate health risk maintained the same trend. SO_4 concentrations in some locations were above the 100 mg/L limit by SON (2015); thus, the affected water resources pose some negative health effects on the water users. Low TH values were recorded in most of the samples. This also predisposes the water users to some health defects. R-mode hierarchical clustering, Varimax-rotated factor analysis, and bivariate/multivariate graphical plots consistently revealed that there are several, mixed water contamination mechanisms in the study area. Nevertheless, their findings posit that geogenic mechanisms predominate over human-related mechanisms. Moreover, these findings were used to develop a composite diagrammatic model that showcases the prevalent contamination sources and mechanisms. Stoichiometric

equations also aided the assessment of the contamination sources and mechanisms. Furthermore, the predictive RBF-NN and MLR models produced in this study accurately predicted the TH, SO_4 , and NO_3 , with the RBF-NN model performing better than the MLR models. The RBF-NN and MLR models also provided insights that aided the validations of the findings from the cluster, factor, and graphical analyses. Conclusively, the outcomes of this research satisfied the research objectives and can be very helpful in identifying areas with health risks and the corresponding actions that need to be performed to control and reduce these risks in the research region. This research also provides theoretical foundations for the effective management and prevention of unsafe, contaminated water supplies by local decision-makers. It will encourage sustainable development of high-quality natural water resources, while promoting ecological protection in the area. For the local and regional authorities to take the necessary precautions in the reduction of health risks through sustainable water management, the provided models and results from this research are extremely important.

Limitations of study, perspectives for future research, and recommendations

There are some noted shortcomings that could be addressed in subsequent research, despite the fact that the current study was founded on strong scientific principles and took into account a number of water quality parameters. Also provided are some vital recommendations:

- The USEPA provided the majority of the values for the health risk parameters that were employed, which could have a negative impact on the nitrate health risk model's accuracy. Therefore, more research is required to produce data that is particular to the Nigerian context.
- In the current modeling, the seasonal variations in water quality were not taken into account. Thus, future studies in the region are needed to investigate that.
- In the research area, anthropogenic activities generally correlate with nitrate content in the studied water resources. The potential of non-carcinogenic health impacts on humans could dramatically increase when the water is utilized for drinking, cooking, bathing, and laundry in the affected areas. The same is true for the sulfate and water hardness. Thus, the poor-quality water resources spotted in parts of the area is to be appropriately adjusted in order to make them suitable for human use.
- Future research could make use of sophisticated deep learning and hybrid predictive modeling tools to enhance or validate the results of the current study.

- The water management strategy should be improved. This can be achieved by encouraging public education about the effects of nitrates, sulfates, and hardness on water quality. Such awareness may help to prevent water contamination in the first place.
- Effective legislation and enforcement strategies, related to anthropogenic activities that threaten the environmental hygiene and water quality, should be introduced as this would help in ensuring that further quality deterioration is minimized and water resources protected.

Author contribution Not applicable.

Manuscript written, reviewed and revised by J. C. Egbueri.

Data availability Not applicable.

All the data used in this paper are provided herein.

Declarations

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent to publish Not applicable.

Competing interests The authors declare no competing interests.

References

- Adimalla N (2020) Controlling factors and mechanism of groundwater quality variation in semiarid region of South India: an approach of water quality index (WQI) and health risk assessment (HRA). *Environ Geochem Health* 42(6):1725–1752
- Adimalla N, Qian H (2019) Groundwater quality evaluation using water quality index (WQI) for drinking purposes and human health risk (HHR) assessment in an agricultural region of Nanganur, south India. *Ecotoxicol Environ Saf* 176:153–161
- Adimalla N, Qian H (2021) Groundwater chemistry, distribution and potential health risk appraisal of nitrate enriched groundwater: a case study from the semi-urban region of South India. *Ecotoxicol Environ Saf* 207:111277
- Agbasi JC, Egbueri JC (2022) Assessment of PTEs in water resources by integrating HHRISK code, water quality indices, multivariate statistics, and ANNs. *Geocarto Int* <https://doi.org/10.1080/10106049.2022.2034990>
- Aguirre BP, Masachessi G, Ferreyra LJ et al (2019) Searching variables to assess recreational water quality: the presence of infectious human enterovirus and its correlation with the main variables of water pollution by multivariate statistical approach in Córdoba, Argentina. *Environ Sci Pollut Res* 26:6586–6601. <https://doi.org/10.1007/s11356-019-04124-2>
- Ahn J, Albanes D, Peters U, Schatzkin A, Lim U, Freedman M et al (2007) Dairy products, calcium intake, and risk of prostate cancer in the prostate, lung, colorectal, and ovarian cancer screening trial. *Cancer Epidemiol Biomark Prev* 16(12):2623–2630. <https://doi.org/10.1158/1055-9965.EPI-07-0601>
- Al-Abadi AM (2017) Modeling of groundwater productivity in north-eastern Wasit Governorate, Iraq using frequency ratio and Shannon's entropy models. *Appl Water Sci* 7(2):699–716
- Ali S, Shekhar S, Bhattacharya P, Verma G, Chandrasekhar T, Chandrashekhar AK (2018) Elevated fluoride in groundwater of Siwani Block, Western Haryana, India: a potential concern for sustainable water supplies for drinking and irrigation. *Groundw Sustain Dev* 7:410–420. <https://doi.org/10.1016/j.gsd.2018.05.008>
- Amiri V, Li P, Bhattacharya P, Nakhaei M (2021a) Mercury pollution in the coastal Urmia aquifer in northwestern Iran: potential sources, mobility, and toxicity. *Environ Sci Pollut Res* 28:17546–17562. <https://doi.org/10.1007/s11356-020-11865-y>
- Amiri V, Nakhaei M, Lak R, Li P (2021b) An integrated statistical-graphical approach for the appraisal of the natural background levels of some major ions and potentially toxic elements in the groundwater of Urmia aquifer. *Iran Environ Earth Sci* 80:432. <https://doi.org/10.1007/s12665-021-09733-0>
- Ayejoto DA, Agbasi JC, Egbueri JC, Echefu KI (2022) Assessment of oral and dermal health risk exposures associated with contaminated water resources: an update in Ojoto area, southeast Nigeria. *Int J Environ Anal Chem*. <https://doi.org/10.1080/03067319.2021.2023515>
- Baig JA, Kazi TG, Shah AQ, Afridi HI, Khan S, Kolachi NF, Kandhro GA, Wadhwa SK, Shah F (2011) Evaluation of toxic risk assessment of arsenic in male subjects through drinking water in southern Sindh Pakistan. *Biol Trace Elem Res* 143(2):772–786. <https://doi.org/10.1007/s12011-010-8933-5>
- Botheju WSM, Liyanage JA, Kannangara SDP (2021) The groundwater geochemistry and the human health risk assessment of drinking water in an area with a high prevalence of chronic kidney disease of unknown etiology (CKDu), Sri Lanka. *J Chem* 2021:1755140. <https://doi.org/10.1155/2021/1755140>
- Butler LM, Wong AS, Koh WP, Wang R, Yuan JM, Yu MC (2010) Calcium intake increases risk of prostate cancer among Singapore Chinese. *Cancer Res* 70:4941–4948. <https://doi.org/10.1158/0008-5472.CAN-09-4544>
- Cañedo-Argüelles M, Kefford BJ, Piscart C, Prat N, Schäfer RB, Schulz CJ (2013) Salinisation of rivers: an urgent ecological issue. *Environ Pollut* 173:157–167
- Catling LA, Abubakar I, Lake IR, Swift L, Hunter PR (2008) A systematic review of analytical observational studies investigating the association between cardiovascular disease and drinking water hardness. *J Water Health* 6(4):433–442. <https://doi.org/10.2166/wh.2008.054>
- Chen J, Qian H, Wu H (2017a) Nitrogen contamination in groundwater in an agricultural region along the New Silk Road, northwest China: distribution and factors controlling its fate. *Environ Sci Pollut Res* 24(15):13154–13167
- Chen J, Wu H, Qian H, Gao Y (2017b) Assessing nitrate and fluoride contaminants in drinking water and their health risk of rural residents living in a semiarid region of northwest China. *Expo Health* 9(3):183–195
- Chen W, Liu W (2015) Water quality modeling in reservoirs using multivariate linear regression and two neural network models. *Adv Artif Neural Syst*. <https://doi.org/10.1155/2015/521721>
- Cocchetto DM, Levy G (1981) Absorption of orally administered sodium sulfate in humans. *J Pharm Sci* 70(3):331–333
- Dahl C, Sogaard AJ, Tell GS, Flaten TP, Hongve D, Omsland TK et al (2013) Nationwide data on municipal drinking water and hip fracture: could calcium and magnesium be protective? A norepros Study. *Bone* 57(2):84–91. <https://doi.org/10.1016/j.bone.2013.06.017>
- Datta PS, Deb DL, Tyagi SK (1997) Assessment of groundwater contamination from fertilizers in the Delhi area based on NO₃⁻ and K⁺ composition. *J Contam Hydrol* 27(3):249–262

- Ding H, Lang Y-C, Liu C-Q, Liu T-Z (2013) Chemical characteristics and $\delta^{34}\text{S}_{\text{SO}_4^{2-}}$, of acid rain: anthropogenic sulfate deposition and its impacts on CO_2 consumption in the rural karst area of southwest China. *Geochem J* 47:625–638. <https://doi.org/10.2343/geochemj.2.0293>
- Egbueri JC (2019) Water quality appraisal of selected farm provinces using integrated hydrogeochemical, multivariate statistical, and microbiological technique. *Model Earth Syst Environ* 5(3):997–1013. <https://doi.org/10.1007/s40808-019-00585-z>
- Egbueri JC (2020) Heavy metals pollution source identification and probabilistic health risk assessment of shallow groundwater in Onitsha, Nigeria. *Anal Lett* 53(10):1620–1638. <https://doi.org/10.1080/00032719.2020.1712606>
- Egbueri JC (2021) Prediction modeling of potentially toxic elements' hydrogeopollution using an integrated Q-mode HCs and ANNs machine learning approach in SE Nigeria. *Environ Sci Pollut Res* 28(30):40938–40956. <https://doi.org/10.1007/s11356-021-13678-z>
- Egbueri JC (2022) Incorporation of information entropy theory, artificial neural network and soft computing models in the development of integrated industrial water quality index. *Environ Monit Assess* 194(10):693. <https://doi.org/10.1007/s10661-022-10389-x>
- Egbueri JC, Mgbenu CN (2020) Chemometric analysis for pollution source identification and human health risk assessment of water resources in Ojoto Province, southeast Nigeria. *Appl Water Sci* 10(4):98. <https://doi.org/10.1007/s13201-020-01180-9>
- Egbueri JC, Mgbenu CN, Chukwu CN (2019) Investigating the hydrogeochemical processes and quality of water resources in Ojoto and environs using integrated classical methods. *Model Earth Syst Environ* 5(4):1443–1461. <https://doi.org/10.1007/s40808-019-00613-y>
- Egbueri JC, Agbasi JC (2022a) Data-driven soft computing modeling of groundwater quality parameters in south-east Nigeria: comparing the performances of different algorithms. *Environ Sci Pollut Res* 1–28. <https://doi.org/10.1007/s11356-022-18520-8>
- Egbueri JC, Agbasi JC (2022b) Combining data-intelligent algorithms for the assessment and predictive modeling of groundwater resources quality in parts of southeastern Nigeria. *Environ Sci Pollut Res* 1–25. <https://doi.org/10.1007/s11356-022-19818-3>
- Egbueri JC, Agbasi JC (2022c) Performances of MLR, RBF-NN, and MLP-NN in the evaluation and prediction of water resources quality for irrigation purposes under two modeling scenarios. *Geocarto Int.* <https://doi.org/10.1080/10106049.2022.2087758>
- Egbueri JC, Ezugwu CK, Unigwe CO, Onwuka OS, Onyemesili OC, Mgbenu CN (2020) Multidimensional analysis of the contamination status, corrosivity and hydrogeochemistry of groundwater from parts of the Anambra Basin, Nigeria. *Anal Lett.* <https://doi.org/10.1080/00032719.2020.1843049>
- Egbueri JC, Mgbenu CN, Digwo DC, Nnyigide CS (2021) A multi-criteria water quality evaluation for human consumption, irrigation and industrial purposes in Umunya area, southeastern Nigeria. *Int J Environ Anal Chem.* <https://doi.org/10.1080/03067319.2021.1907360>
- Egbueri JC, Igwe O, Omeka ME, Agbasi JC (2022) Development of MLR and variedly optimized ANN models for forecasting the detachability and liquefaction potential index of erodible soils. *Geosyst Geoenviron.* <https://doi.org/10.1016/j.geogeo.2022.100104>
- Egbueri JC (2018) Assessment of the quality of groundwaters proximal to dumpsites in Awka and Nnewi metropolises: a comparative approach. *Int J Energy Water Res.* <https://doi.org/10.1007/s42108-018-0004-1>
- Emami S, Emami H, Choopan Y, Parsa J, Jahandideh O (2020) Modeling groundwater quality using three novel hybrid support vector regression models. *Adv Environ Technol* 2:99–110
- Farid H, Mahmood-Khan Z, Ali A, Mubeen M, Anjum M (2017) Site-specific aquifer characterization and identification of potential groundwater areas in Pakistan. *Poli J Environ Stud* 26(1):17–27. <https://doi.org/10.15244/pjoes/64645>
- Florin T, Neale G, Gibson GR, Christl SU, Cummings JH (1991) Metabolism of dietary sulphate: absorption and excretion in humans. *Gut* 32:766–773
- Gaya MS, Abba SI, Abdu AM, Tukur AI, Saleh AM, Esmaili P, Wahab NA (2020) Estimation of water quality index using artificial intelligence approaches and multi-linear regression. *IAES Int J Artif Intell* 9(1):126–134. <https://doi.org/10.11591/ijai.v9.i1.pp126-134>
- Ge X, Wu Q, Wang Z, Gao S, Wang T (2021) Sulfur isotope and stoichiometry-based source identification of major ions and risk assessment in Chishui River Basin, Southwest China. *Water* 13:1231. <https://doi.org/10.3390/w13091231>
- Gianfredi V, Bragazzi NL, Nucci D, Villarini M, Moretti M (2017) Cardiovascular diseases and hard drinking waters: implications from a systematic review with meta-analysis of case-control studies. *J Water Health* 15(1):31–40. <https://doi.org/10.2166/wh.2016.131>
- Gutiérrez M, Biagioni RN, Alarcón-Herrera MT, Rivas-Lucero BA (2018) An overview of nitrate sources and operating processes in arid and semiarid aquifer systems. *Sci Total Environ* 624:1513–1522
- Han GL, Liu CQ (2006) Strontium isotope and major ion chemistry of the rainwaters from Guiyang, Guizhou Province, China. *Sci Total Environ* 364:165–174. <https://doi.org/10.1016/j.scitotenv.2005.06.025>
- Han G, Tang Y, Wu Q, Liu M, Wang Z (2019) Assessing contamination sources by using sulfur and oxygen isotopes of sulfate ions in Xijiang River Basin, Southwest China. *J Environ Qual* 48(5):1507–1516. <https://doi.org/10.2134/jeq2019.03.0150>
- Haykin S (1999) *Neural Networks: a comprehensive foundation*, 2nd edn. Prentice Hall Inc, Upper Saddle River, pp 26–32
- Huang Y, Wang J, Tan Y, Wang L, Lin H, Lan L et al (2018) Low-mineral direct drinking water in school may retard height growth and increase dental caries in schoolchildren in China. *Environ Int* 115:104–109. <https://doi.org/10.1016/j.envint.2018.02.021>
- Huang Y, Ma X, Tan Y, Wang L, Wang J, Lan L et al (2019) Consumption of very low mineral water is associated with lower bone mineral content in children. *J Nutr* 149:1994–2000. <https://doi.org/10.1093/jn/nxz161>
- Ighalo JO, Adeniyi AG (2020) A comprehensive review of water quality monitoring and assessment in Nigeria. *Chemosphere.* <https://doi.org/10.1016/j.chemosphere.2020.127569>
- Ighalo JO, Adeniyi AG, Adeniran JA, Oggunniyi S (2020) A systematic literature analysis of the nature and regional distribution of water pollution sources in Nigeria. *J Clean Prod.* <https://doi.org/10.1016/j.jclepro.2020.124566>
- Isiuku BO, Enyoh CE (2020) Pollution and health risks assessment of nitrate and phosphate concentrations in water bodies in South Eastern. *Nigeria Environ Adv* 2:100018
- Jiang L, He P, Chen J, Liu Y, Liu D, Qin G et al (2016) Magnesium levels in drinking water and coronary heart disease mortality risk: a meta-analysis. *Nutrients* 8(1):5. <https://doi.org/10.3390/nu8010005>
- Joslyn S, Lynch C, Wallace R, Olson D, Van Hoesen C (1990) Relationship between diabetes mellitus mortality rates and drinking water magnesium levels in Iowa. *Magnes Trace Elem* 9(2):94–100
- Karmakar B, Singh MK, Choudhary BK, Singh SK, Egbueri JC, Gautam S, Rawat SK (2021) Investigation of the hydrogeochemistry, groundwater quality and associated health risks in industrialized regions of tripura, northeast India. *Environ Forensics.* <https://doi.org/10.1080/15275922.2021.2006363>
- Karunanidhi D, Aravinthasamy P, Subramani T, Wu J, Srinivasamoorthy K (2019) Potential health risk assessment for fluoride and nitrate contamination in hard rock aquifers of Shanmuganadhi River basin, South India. *Hum Ecol Risk Assess* 25:250–270

- Karunanidhi D et al (2020a) Evaluation of non-carcinogenic risks due to fluoride and nitrate contaminations in a groundwater of an urban part (Coimbatore region) of south India. *Environ Monit Assess* 192(2):102
- Karunanidhi D, Aravinthasamy P, Kumar D, Subramani T, Roy PD (2020) Sobol sensitivity approach for the appraisal of geomedi-cal health risks associated with oral intake and dermal pathways of groundwater fluoride in a semi-arid region of south India. *Ecotoxicol Environ Saf* 194:110438. <https://doi.org/10.1016/j.ecoenv.2020.110438>
- Karunanidhi D, Aravinthasamy P, Deepali M, Subramani T, Bel-lows BC, Li P (2020a) Groundwater quality evolution based on geochemical modeling and aptness testing for ingestion using entropy water quality and total hazard indexes in an urban-indus-trial area (Tiruppur) of Southern India. *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-020-10724-0>
- Kobayashi J (1957) On geographical relationship between the chemical nature of river water and death-rate from apoplexy. *Berichte d Ohara Inst f Landwirtsch Biologie* 11:12–21
- Langenegger O (1990) Groundwater quality in rural areas of western Africa. UNDP project INT/81/026:10
- Li P, Wu J (2019) Drinking water quality and public health. *Expo Health* 11:73–79. <https://doi.org/10.1007/s12403-019-00299-8>
- Li P, He S, He X, Tian R (2017a) Seasonal hydrochemical charac-terization and groundwater quality delineation based on matter element extension analysis in a paper wastewater irrigation area, northwest China. *Expo Health* 10:241–258. <https://doi.org/10.1007/s12403-017-0258-6>
- Li P, Tian R, Xue C, Wu J (2017b) Progress, opportunities and key fields for groundwater quality research under the impacts of human activities in China with a special focus on western China. *Environ Sci Pollut Res* 24(26):13224–13234. <https://doi.org/10.1007/s11356-017-8753-7>
- Liu CW, Lin KH, Kuo YM (2003) Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. *Sci Total Environ* 313:77–89. [https://doi.org/10.1016/S0048-9697\(02\)00683-6](https://doi.org/10.1016/S0048-9697(02)00683-6)
- Liu J, Han G (2020) Distributions and source identification of the major ions in Zhujiang River, Southwest China: examining the relationships between human perturbations, chemical weathering, water quality and health risk. *Expo Health*<https://doi.org/10.1007/s12403-020-00343-y>
- Long J, Luo K (2020) Elements in surface and well water from the central North China Plain: enrichment patterns, origins, and health risk assessment. *Environ Pollut* 258:113725. <https://doi.org/10.1016/j.envpol.2019.113725>
- McGowan W (2000) Water processing: residential, commercial, light industrial, 3rd edn. Water Quality Association, Lisle
- McGuire MJ et al (1984) Controlling attached blue-green algae with copper sulphate. *J Am Water Works Assoc* 76:60
- Menard S (2000) Coefficients of determination for multiple logistic regression analysis. *Am Stat* 54(1):17–24
- Mondal D, Dutta G, Gupta S (2016) Inferring the fluoride hydrogeo-chemistry and effect of consuming fluoride-contaminated drink-ing water on human health in some endemic areas of Birbhum district, West Bengal. *Environ Geochem Health* 38:557–576. <https://doi.org/10.1007/s10653-015-9743-7>
- Moore E (1952) Physiological effects of the consumption of saline drink-ing water. A progress report to the Subcommittee on the Water Sup-ply of the Committee on Sanitary Engineering and Environment. National Academy of Sciences, Washington DC Appendix B 1–2
- Morris ME, Levy G (1983) Absorption of sulfate from orally admin-istered magnesium sulfate in man. *J Toxicol-Clin Toxicol* 20(2):107–114
- Mukate S, Wagh V, Panaskar D, Jacobs JA, Sawant A (2019) Devel-opment of new integrated water quality index (IWQI) model to evaluate the drinking suitability of water. *Ecol Indic* 101:348–354. <https://doi.org/10.1016/j.ecolind.2019.01.034>
- NAS (1977) Drinking water and health. DC, National Research Coun-cil, National Academy of Sciences, Washington
- Naumann J, Biehler D, Lu'ty T, Sadaghiani C. (2017) Prevention and therapy of type 2 diabetes – what is the potential of daily water intake and its mineral nutrients? *Nutrients* 9(8):914. <https://doi.org/10.3390/nu9080914>
- Nfor B, Olobaniyi S, Ogala J (2007) Extent and distribution of ground-water resources in parts of Anambra State, Southeastern, Nigeria. *J Appl Sci Environ Manag* 11(2). <https://doi.org/10.4314/jasem.v11i2.55050>
- Nickson RT, McArthur JM, Shresthn B, Kyaw- Nyint TO, Lowry D (2005) Arsenic and other drinking water quality issues, muzaf-fargarh district, pakistan. *Appl Geochem* 20(1):55–68. <https://doi.org/10.1016/j.apgeochem.2004.06.004>
- Niroobakhsh M, Musavi-Jahromi SH, Manshouri M, Sedghi H (2012) Prediction of water quality parameter in Jajrood River basin: application of multilayer perceptron (MLP) perceptron and radial basis function networks of artificial neural networks (ANNs). *Afr J Agric Res* 7(29):4131–4139
- Nwachukwu SO (1972) The tectonic evolution of the southern portion of the Benue Trough, Nigeria. *Geol Mag* 109:411–419
- Nwajide CS (2013) Geology of Nigeria's sedimentary basins. CSS Press, Lagos
- Okoro EI, Egboka BCE, Anike OL, Enekeuchi EK (2010a) Evalua-tion of groundwater potentials in parts of the Escarpment area of southeastern Nigeria. *Int J Geomat Geosci* 1(3):544–551
- Okoro EI, Egboka BCE, Onwuemesi AG (2010b) Evaluation of the aquifer characteristics of the Nanka Sand using hydrogeological method in combination with vertical electric sounding (VES). *J Appl Sci Environ Manag* 14(2):5–9
- Omeka ME, Egbueri JC, Unigwe CO (2022) Investigating the hydro-geochemistry, corrosivity and scaling tenden-cies of groundwa-ter in an agrarian area (Nigeria) using graphical, indexical and statistical modelling. *Arab J Geosci* 15(13):1233. <https://doi.org/10.1007/s12517-022-10514-7>
- Omeka ME, Egbueri JC (2022) Hydrogeochemical assessment and health-related risks due to toxic element ingest-ion and dermal contact within the Nnewi-Awka urban areas, Nigeria. *Environ Geochem Health*. <https://doi.org/10.1007/s10653-022-01332-7>
- Onjia A, Huang X, Trujillo González JM, Egbueri JC (2022) Editorial: Chemometric approach to distribution, so-urce apportionment, ecological and health risk of trace pollutants. *Front Environ Sci* 10:1107465. <https://doi.org/10.3389/fenvs.2022.1107465>
- Ozel HU, Gemici BT, Gemici E, Ozel HB, Cetin M, Sevik H (2020) Application of artificial neural networks to predict the heavy metal contamination in the Bartin River. *Environ Sci Pollut*. <https://doi.org/10.1007/s11356-020-10156-w>
- Panda SS, Garg V, Chaubey I (2004) Artificial neural networks appli-cation in lake water quality estimation using satellite imagery. *J Environ Inform* 4(2):65–74
- Panneerselvam B, Muniraj K, Pande C, Ravichandran N, Thomas M, Karuppannan S (2021) Geochemical evaluation and human health risk assessment of nitrate-contaminated groundwater in an industrial area of South India. *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-021-17281-0>
- Peterson NL (1951) Sulfates in drinking water. *Official Bulletin N.D. Water Sewage Works* 18:11–12
- Rahman A, Mondal NC, Tiwari KK (2021) Anthropogenic nitrate in groundwater and its health risks in the view of background concentration in a semi-arid area of Rajasthan. *India Sci Rep* 11:9279. <https://doi.org/10.1038/s41598-021-88600-1>
- Rapant S, Cvečková V, Fajčíková K, Sedláková D, Stehlíková B (2017) Impact of calcium and magnesium in groundwater and drinking

- water on the health of inhabitants of the Slovak Republic. *Int J Environ Res Public Health* 14:278. <https://doi.org/10.3390/ijerph14030278>
- Rapant S, Letkovičová A, Jurkovičová D, Kosmovský V, Kožíšek F, Jurkovič L (2020) Differences in health status of Slovak municipalities supplied with drinking water of different hardness values. *Environ Geochem Health* <https://doi.org/10.1007/s10653-020-00664-6>
- Rosborg I, Kozisek F (eds) (2020) Drinking water minerals and mineral balance. Importance, health significance, safety precautions, 2nd edn. Springer International Publishing Switzerland, Springer Verlag
- Rubenowitz-Lundin E, Hiscock K (2005) Water hardness and health effects. *Essential of Medical Geology*. Elsevier Academic Press, Amsterdam, pp 331–345
- Ruidas D, Pal SC, Islam ARMT, Saha A (2022) Hydrogeochemical evaluation of groundwater aquifers and associated health hazard risk mapping using ensemble data driven model in a water scares Plateau Region of Eastern India. *Expo Health* <https://doi.org/10.1007/s12403-022-00480-6>
- Sawyer GN, McCarthy DL (1967) Chemistry of sanitary engineers, 2nd edn. McGraw Hill, New York
- Schofield R, Hsieh D (1983) Criteria and recommendations for standards for sulfate in military field supplies. Lawrence Livermore National Laboratory, Livermore (**Contract no. UCRL-53481-4**)
- Sharma MK, Kumar M (2020) Sulphate contamination in groundwater and its remediation: an overview. *Environ Monit Assess* 192:74. <https://doi.org/10.1007/s10661-019-8051-6>
- Skold ACD, Klein R (2011) Methemoglobinemia: pathogenesis, diagnosis, and management. *South Med J* 104(11):757–761
- Snousy MG, Wu J, Su F, Abdelhalim A, Ismail E (2021) Groundwater quality and its regulating geochemical processes in Assiut Province, Egypt. *Expo Health*. <https://doi.org/10.1007/s12403-021-00445-1>
- SON (Standard Organization of Nigeria) (2015) Nigerian-standard for drinking-water-quality-NIS-554–2015 pp. 1–28, Abuja
- Stipanuk MH (2000) Biochemical and physiological aspects of human nutrition. W. B. Saunders Co., Philadelphia, p 706
- Subba Rao N, Chaudhary M (2019) Hydrogeochemical processes regulating the spatial distribution of groundwater contamination, using pollution index of groundwater (PIG) and hierarchical cluster analysis (HCA): A case study. *Groundw Sustain Dev* 9:100238. <https://doi.org/10.1016/j.gsd.2019.100238>
- Tian R, Wu J (2019) Groundwater quality appraisal by improved set pair analysis with game theory weightage and health risk estimation of contaminants for Xuecha drinking water source in a loess area in Northwest China. *Hum Ecol Risk Assess* 25:132–157. <https://doi.org/10.1080/10807039.2019.1573035>
- Tokatli C (2021) Application of nutrient pollution index and water pollution index to evaluate the drinking water quality of the villages located in Edirne, Turkey. In: International Agricultural, Biological & Life Science Conference, Edirne 1–3, pp 554–559
- Torres-Martínez JA, Mora A, Knappett P, Ornelas-Soto N, Mahlknicht J (2020) Tracking nitrate and sulfate sources in groundwater of an urbanized valley using a multi-tracer approach combined with a Bayesian isotope mixing model. *Water Res* 182:115962
- Ukah BU, Egbueri JC, Unigwe CO, Ubido OE (2019) Extent of heavy metals pollution and health risk assessment of groundwater in a densely populated industrial area, Lagos, Nigeria. *Int J Environ Res Public Health* 16(4):291–303. <https://doi.org/10.1007/s42108-019-00039-3>
- Unigwe CO, Egbueri JC, Omeka ME (2022) Geospatial and statistical approaches to nitrate health risk and groundwater quality assessment of an alluvial aquifer in SE Nigeria for drinking and irrigation purposes. *J Indian Chem Soc* 99:100479. <https://doi.org/10.1016/j.jics.2022.100479>
- USEPA (1985) National primary drinking water regulations; synthetic organic chemicals, inorganic chemicals and microorganisms; proposed rule. US Environ Protect Agency Fed Regist 50(219):46936
- USEPA (2002) Supplemental guidance for developing soil screening levels for superfund sites. Environmental Protection Agency, office of emergency and remedial response, Washington DC
- USEPA (1984) National Secondary Drinking Water Regulations. EPA570/9–76–000. Office of Water, Washington DC
- USEPA (1989) Risk assessment guidance for superfund. Human Health Evaluation Manual (Part A): Office of Emergency and Remedial Response 1: Washington DC
- USEPA (1997) Exposure factors handbook, volume 1: General Factors. U. S. Environmental Protection Agency, office of research and development, Washington DC
- USEPA (2003) Drinking Water Advisory: Consumer Acceptability Advice and Health Effects Analysis on Sulfate. U.S. Environmental Protection Agency, EPA 822-R-03–007, Office of Water (4304T), Health and Ecological Criteria Division, Washington, DC
- Ustaoglu F, Tepe Y (2019) Water quality and sediment contamination assessment of Pazarsuyu stream, Turkey using multivariate statistical methods and pollution indicators. *Int Soil Water Cons Res* 7:47–56. <https://doi.org/10.1016/j.iswcr.2018.09.001>
- Wang D, Wu J, Wang Y, Ji Y (2020) Finding high-quality groundwater resources to reduce the hydatidosis incidence in the Shiqu County of Sichuan Province, China: analysis, assessment, and management. *Expo Health* 12(2):307–322. <https://doi.org/10.1007/s12403-019-00314-y>
- Ward JH (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58(301):236–244
- Ward M et al (2018) Drinking water nitrate and human health: an updated review. *Int J Environ Res Public Health* 15:1557
- Wei M, Wu J, Li W, Zhang Q, Su F, Wang Y (2021) Groundwater geochemistry and its impacts on groundwater arsenic enrichment, variation, and health risks in Yongning County, Yinchuan Plain of Northwest China. *Expo Health*, <https://doi.org/10.1007/s12403-021-00391-y>
- Weisberg S (1985) Applied Regression, 2nd edn. John Wiley & Sons, New York
- Whipple GC (1907) The value of pure water. John Wiley, New York (**cited in Committee on Water Quality Criteria, 1972**)
- WHO (2017) Guidelines for drinking water quality: fourth edition incorporating the first addendum. World Health Organization, Geneva
- WHO (2004) Sulfate in drinking-water: background document for development of WHO guidelines for drinking-water quality. World Health Organization, Geneva. WHO/SDE/WSH/03.04/114, 16p
- Wongsanit J, Teartisup P, Kerdsueb P, Tharnpoophasiam P, Worakunpiset S (2015) Contamination of nitrate in groundwater and its potential human health: a case study of lower Mae Klong River basin, Thailand. *Environ Sci Pollut Res* 22(15):11504–11512. <https://doi.org/10.1007/s11356-015-4347-4>
- Wu QX, Han GL, Tao FX, Tang Y (2012) Chemical composition of rainwater in a karstic agricultural area, Southwest China: the impact of urbanization. *Atmos Res* 111:71–78. <https://doi.org/10.1016/j.atmosres.2012.03.002>
- Wu J, Li P, Qian H, Duan Z, Zhang X (2014) Using correlation and multivariate statistical analysis to identify hydrogeochemical processes affecting the major ion chemistry of waters: Case study in Laoheba phosphorite mine in Sichuan, China. *Arab J Geosci* 7(10):3973–3982. <https://doi.org/10.1007/s12517-013-1057-4>
- Wu J, Wang L, Wang S, Tian R, Xue C, Feng W, Li Y (2017) Spatiotemporal variation of groundwater quality in an arid area experiencing long-term paper wastewater irrigation, northwest China. *Environ Earth Sci* 76(13):460. <https://doi.org/10.1007/s12665-017-6787-2>

- Wu J, Li P, Wang D, Ren X, Wei M (2020) Statistical and multivariate statistical techniques to trace the sources and affecting factors of groundwater pollution in a rapidly growing city on the chinese loess plateau. *Hum Ecol Risk Assess* 26(6):1603–1621. <https://doi.org/10.1080/10807039.2019.1594156>
- Xu Z, Liu C-Q (2007) Chemical weathering in the upper reaches of Xijiang River draining the Yunnan-Guizhou Plateau, Southwest China. *Chem Geol* 239:83–95. <https://doi.org/10.1016/j.chemgeo.2006.12.008>
- Yang CY, Cheng MF, Tsai SS, Hsieh YL (1998) Calcium, magnesium, and nitrate in drinking water and gastric cancer mortality. *Jpn J Cancer Res* 89:124–130. <https://doi.org/10.1111/j.1349-7006.1998.tb00539.x>
- Yang CY, Chiu HF, Cheng MF, Tsai SS, Hung CF, Tseng YT (1999) Mg in drinking water and the risk of death from diabetes mellitus. *Magnes Res* 12:131–137
- Yang CY, Chiu HF, Cheng BH, Hsu TY, Cheng MF, Wu TN (2000) Calcium and magnesium in drinking water and risk of death from breast cancer. *J Toxicol Environ Health Part A Curr Issues* 60(4):231–241. <https://doi.org/10.1080/00984100050027798>
- Zoeteman BCJ (1980) *Sensory assessment of water quality*. Pergamon Press, New York

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.