Scholarly Publisher
RS Global Sp. z O.O.
ISNI: 0000 0004 8495 2390

Dolna 17, Warsaw, Poland 00-773
Tel: +48 226 0 227 03
Email: editorial_office@rsglobal.pl

JOURNAL	International Journal of Innovative Technologies in Social Science
p-ISSN	2544-9338
e-ISSN	2544-9435
PUBLISHER	RS Global Sp. z O.O., Poland

ARTICLE TITLE	GROUNDWATER POLLUTION SOURCE USING PRINCIPAL COMPONENT ANALYSIS IN GUELMA PLAIN, NORTHEAST ALGERIA
AUTHOR(S)	Baazi Houria, Nemili Zohra, Necer AbdeLdjabar, Saoudi Messaoud
ARTICLE INFO	Baazi Houria, Nemili Zohra, Necer AbdeLdjabar, Saoudi Messaoud. (2024) Groundwater Pollution Source Using Principal Component Analysis in Guelma Plain, Northeast Algeria. *International Journal of Innovative Technologies in Social Science*. 3(43). doi: 10.31435/rsglobal_ijitss/30092024/8248
DOI	https://doi.org/10.31435/rsglobal_ijitss/30092024/8248
RECEIVED	15 August 2024
ACCEPTED	27 September 2024
PUBLISHED	29 September 2024
LICENSE	This work is licensed under a **Creative Commons Attribution 4.0 International License.**

GROUNDWATER POLLUTION SOURCE USING PRINCIPAL COMPONENT ANALYSIS IN GUELMA PLAIN, NORTHEAST ALGERIA

Baazi Houria
Natural Hazards and Territory Planning Laboratory (LRNAT), Earth Sciences and Universe Institute, University of Batna 2, Fesdis 05078, Batna, Algeria

Nemili Zohra
Natural Hazards and Territory Planning Laboratory (LRNAT), Faculty of Technology, University of Batna 2, Fesdis 05078, Batna, Algeria

Necer AbdeLdjabar
Laboratory of Thermodynamics and Energetic Systems, USTHB university, BP 32 Bab Ezzouar, 16111 Algeirs, Faculty of Natural and Life Sciences, University of Batna 2, Fesdis 05078, Batna, Algeria

Saoudi Messaoud
Laboratory of Biodiversity, Biotechnology, and Sustainable Development, Faculty of Natural and Life Sciences, University of Batna 2, Fesdis 05078, Batna, Algeria

DOI: https://doi.org/10.31435/rsglobal_ijitss/30092024/8248

ARTICLE INFO

Received 15 August 2024
Accepted 27 September 2024
Published 29 September 2024

KEYWORDS

Guelma, Pollution, Groundwater Quality, PCA, Correlation Matrix, Piper Diagram.

ABSTRACT

This study was conducted in the agricultural region of Guelma, located in northeastern Algeria, where groundwater is the main source of water for human consumption, agriculture, and industry. The objective was to characterize groundwater quality and identify potential pollution sources using multivariate statistical methods, including Principal Component Analysis (PCA), correlation matrix, and the Piper diagram. The analyses revealed strong correlations between certain ions, such as sodium and chloride, suggesting carbonate dissolution processes, such as calcite and dolomite, leading to increased water hardness. The Piper diagram allowed for the classification of water types based on the relative concentrations of major cations (Ca^{2+}, Mg^{2+}, Na^{+}+K^{+}) and anions (Cl^{-}, SO_4^{2-}, CO_3^{2-}+HCO_3^{-}), showing a predominance of mixed hydrochemical types influenced by both natural and anthropogenic processes.

PCA then simplified the interpretation by identifying the most influential variables that could serve as key indicators for the continuous monitoring of water quality. Overall, the results indicate that groundwater chemistry in this region is strongly influenced by human activities and local geological conditions, highlighting the need for sustainable management and continuous monitoring to protect this vital resource.

Citation: Baazi Houria, Nemili Zohra, Necer AbdeLdjabar, Saoudi Messaoud (2024) Groundwater Pollution Source Using Principal Component Analysis in Guelma Plain, Northeast Algeria. *International Journal of Innovative Technologies in Social Science*. 3(43). doi: 10.31435/rsglobal_ijitss/30092024/8248

1. Introduction.

Groundwater is a crucial water resource for potable water supply. Similarly, groundwater resources provide significant natural wealth for the socio-economic development of a country (Flörke, Schneider et al. 2018). Irrigation and industrial uses in many countries, especially in arid and semi-arid areas(Moukhliss, Koubi et al. (2023); Ji, Wu et al. (2020)) where surface water resources and precipitation are limited. Moreover, natural factors such as climate, the nature of aquifer rocks, and interactions between groundwater and these rocks, as well as anthropogenic factors like agriculture and fertilizer use, significantly influence the quality of groundwater in the short or long term (Alabjah, Amraoui et al. (2018);Moukhliss and Taleb (2021)), thus endangering these vital resources, especially those considered a safe source for consumption, domestic, industrial, and agricultural uses(Li, Tian et al. 2017). Additionally, industry and urban development influence the quality of groundwater in a sustainable or unsustainable manner (Wu, Li et al. 2020) Also, waste has contributed to the deterioration of groundwater quality by releasing heavy metals, organic, and inorganic chemical substances into the aquifer (Islam, Shen et al. 2018). Although monitoring data (physical and chemical parameters) cannot explain, interpret, and identify the source of pollution due to their complexity, it is proven that multivariate statistical approaches are an effective means for interpreting complex data matrices. Indeed, several multivariate statistical methods have been employed to estimate water quality by detecting pollution sources and their moderating parameters (Wu, Li et al. (2014); Zhang, Wang et al. (2017); Gulgundi and Shetty (2018)). Although several multivariate statistical methods include cluster analysis (CA), factor analysis (FA), principal component analysis (PCA), and multiple linear regression (MLR) (Moukhliss, Taleb et al. (2023); Li, Zhang et al. (2020)), PCA is the most effective and widely used multivariate statistical method to identify explanatory factors for recognizing potential pollution sources by specifying the main components of groundwater data sets (Qilin, Xiaodan et al. (2020) ; Bhutiani, Kulkarni et al. (2016); BENMARCE, CHOUABBI et al. (2014)), as a means of reducing statistical data. PCA is generally coupled with the hydrochemical approach to define the anthropogenic and natural processes that contribute to the degradation of groundwater quality (Ayed, Jmal et al. 2017).

The study was conducted in the agricultural region of Guelma in northeastern Algeria (Baazi 2023), where groundwater is the main source of water for human consumption, agriculture, and industry. The use of fertilizers and long-term wastewater threatens groundwater quality in this region. Indeed, previous studies conducted in the study area have highlighted agricultural-origin pollution of the aquifer, using groundwater vulnerability assessment methods DRASTIC and SI (Baazi 2023) and the GOD method (Houria and Naima (2023); Latifi and Chaab (2017)). Consequently, the vulnerability maps obtained by the three methods revealed three classes (high, medium, and low), proving the contamination of the study area. PCA was used in this study to characterize groundwater quality and identify pollution sources.

2. Stady Area.

The Guelma-Boumahra plain, covering an area of approximately 122 km², stretches over about twenty kilometers from east to west and 3 to 10 kilometers from north to south. Integrated into the vast hydrographic basin of the Oued Seybouse, it is bordered by significant mountain ranges: the Houara and Djebel Bousbaa to the north, the Mahouna and Beni Marmis to the south, Djebel Arar to the west, and the Nador massif to the east (Gouaidia Samira 2019). This region, with an average altitude of 227 meters, is primarily dedicated to agriculture and relies on the Guelma alluvial aquifer (Houria and Naima 2023). The geology of the Guelma region is composed of three main formations : pre-aquifer layers, Miocene and Pliocene sediments (originating from the Guelma basin), and more recent Pliocene and Quaternary deposits. The neritic domains of Djebel Debagh, Héliopolis, and the southern part of Guelma all belong to the Guelma region. This Jurassic-Cretaceous carbonate basement includes several thrust sheets and has undergone significant tectonic events (Houria and Naima 2023). The Oued Seybouse is formed by alluvium (Vila 1980) between Nador and Medjez Amar, consisting of gypsiferous marls and Quaternary deposits (heterogeneous alluvial terraces).

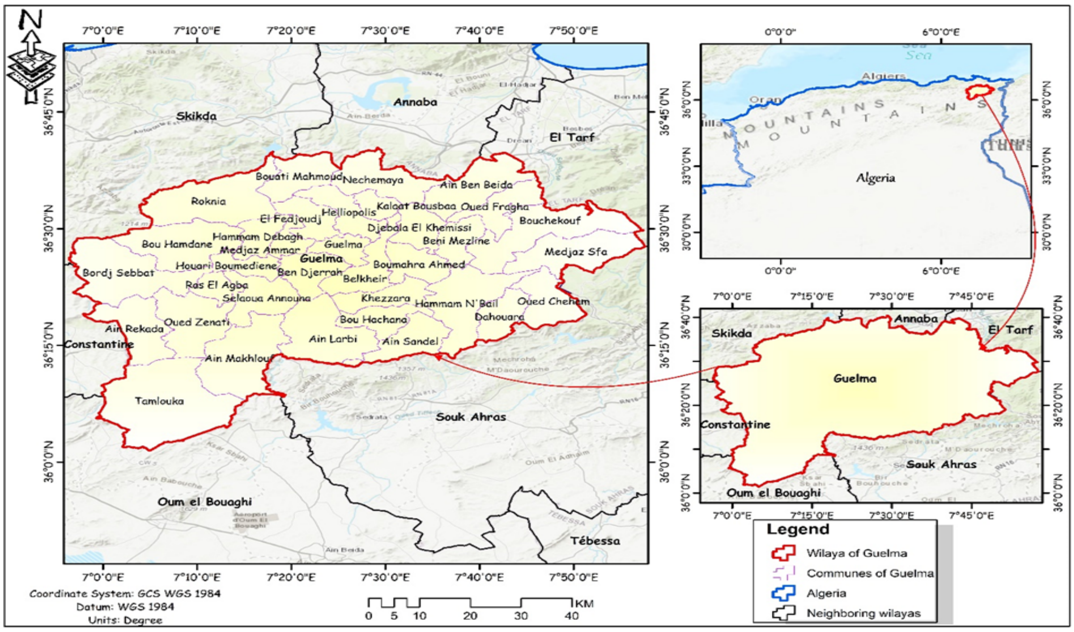

Figure 1. Map of geographical location of study area.

3. Materials and methods.

3.1. Sampling and Analysis.

In this study, a total of 60 sample analysis results were obtained from the National Agency of Water Resources (A.N.R.H.) in Guelma. The analysis results span a 9-year period (from 2013 to 2019) and are distributed across 10 wells. The results include concentrations of physical parameters such as electrical conductivity (EC) and total hardness (TH), major elements like calcium (Ca^{2+}), magnesium (Mg^{2+}), potassium (K^{+}), sodium (Na^{+}), bicarbonates (HCO_3^{-}), sulfates (SO_4^{2-}), and chloride (Cl^{-}), as well as pollution indicators such as nitrates (NO_3^{-}), ammonium (NH_4^{+}), iron (Fe_2^{+}), and orthophosphates (PO_4^{3-}).

Table 1. Descriptive Statistics of chemical elements.

Variables	Cond	HCO_3.	TH	Ca2.	Mg2.	Cl.	$SO4^{2-}$	NO_3	K.	Na.
Min	563	134.1	0.43	110.5	8.46	127.8	48.48	1.942	0.6004	51.15
1st qu	1027	278.2	42.85	129.0	20.65	156.2	85.91	4.836	1.4522	63.59
Median	1131	301.3	45.10	138.4	25.87	184.6	103.01	9.190	2.2999	70.82
Mean	1218	302.5	45.78	140.4	26.54	216.9	113.97	11.399	2.4685	82.82
3rd qu	1284	317.2	49.40	147.5	32.13	258.9	116.50	13.940	2.8240	94.59
Max	2980	484.3	88.20	40.082	69.14	461.5	323.00	40.082	10.6996	194.39

3.2. Statistical analysis.

Correlation matrix analysis (CMA) and principal component analysis (PCA) for groundwater physicochemical parameters are multivariate statistical methods performed using RStudio. Indeed, CMA was carried out to determine the degree of correlation between each pair of water quality parameters. The degree of dependence of one parameter on another is analyzed using Pearson's correlation coefficient (r), which varies from -1 (negative correlation) to 1 (positive correlation). If the value of r is close to zero, it means that there is no correlation between the variables (Strickert, Schleif et al. 2009). Pearson's correlation coefficient can be expressed as follows:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{(X_i - \bar{X})^2}\sqrt{(Y_i - \bar{Y})^2}}$$

where $\bar{X}$,$\bar{Y}$ are the means of X and Y, respectively; i is the total number of variables. Principal Component Analysis (PCA) is a statistical method used to reduce the dimensionality of a dataset while preserving as much of the variability present in this data as possible(Wu, Li et al. (2014); Li, Tian et al. (2019)). It is particularly useful in situations where there are a large number of variables and it allows for a simpler analysis and interpretation of the data, which explains its use in the field of water quality characterization. The analysis transforms the original variables into uncorrelated (orthogonal) principal components, which are expressed as follows:

$$z_{IJ} = a_{i1}x_{1j} + a_{i2}x_{2j} + a_{i3}x_{3j} + \cdots + a_{im}x_{mj}$$

where z is the component ordre; a is the component loading; x is the measured value of the variable; i is the component number; j is the sample number; and m is the total number of variables.

In this Research, 10 groundwater quality parameters were exploited. Data verification was performed to meet the Kaiser-Meyer-Olkin (KMO) criteria (Noori, Abdoli et al. 2009). The minimum KMO value is 0.5 for the reliability of applying PCA (Williams, Onsman et al. 2010). Similarly, Bartlett's sphericity test was used to verify the suitability of applying PCA. If the significance level of Bartlett's test is sufficiently low (< 0.05), it means that the correlation matrix is not an identity matrix, thus indicating that PCA is applicable to the correlation matrix (Ayed, Jmal et al. 2017). Principal components (PCs) were selected according to Kaiser's criterion. Principal components (PCs) with eigenvalues greater than 1.0 were retained. Eigenvalues reflect the importance of each PC ; thus, a PC with higher eigenvalues is considered more significant (Gulgundi and Shetty 2018). Eigenvalues were derived from the covariance matrix of the initial variables (Chabukdhara and Nema 2012). The first component captured the most variance in the dataset, followed by the second component, and so on. Similarly, to make the factors more interpretable without modifying the initial mathematical dataset, the extracted PCs were rotated. In order to better understand the factors influencing groundwater quality, we applied a varimax rotation to the data. This statistical method makes the results more readable by rotating the factor axes in a way that maximizes the variance explained by each factor. Thus, the least importantes variables are less loaded on the main factors, which facilitates the identification of pollution sources. The factor loadings obtained after rotation indicate the degree of correlation between each parameter and each factor. The higher a loading, the more the corresponding factor contributes to explaining the variation of the parameter in question (Ielpo et al., 2012; Felipe-Sotelo et al., 2007)., 10 groundwater quality parameters were exploited. Data verification was performed to meet the Kaiser-Meyer-Olkin (KMO) criteria (Noori, Abdoli et al. 2009). The minimum KMO value is 0.5 for the reliability of applying PCA (Williams, Onsman et al. 2010). Similarly, Bartlett's sphericity test was used to verify the suitability of applying PCA. If the significance level of Bartlett's test is sufficiently low (< 0.05), it means that the correlation matrix is not an identity matrix, thus indicating that PCA is applicable to the correlation matrix (Ayed, Jmal et al. 2017). Principal components (PCs) were selected according to Kaiser's criterion. Principal components (PCs) with eigenvalues greater than 1.0 were retained. Eigenvalues reflect the importance of each PC ; thus, a PC with higher eigenvalues is considered more significant (Gulgundi and Shetty, 2018). Eigenvalues were derived from the covariance matrix of the initial variables (Chabukdhara and Nema, 2012). The first component captured the most variance in the dataset, followed by the second component, and so on. Similarly, to make the factors more interpretable without modifying the initial mathematical dataset, the extracted PCs were rotated. In order to better understand the factors influencing groundwater quality, we applied a varimax rotation to the data. This statistical method makes the results more readable by rotating the factor axes in a way that maximizes the variance explained by each factor. Thus, the least importantes variables are less loaded on the main factors, which facilitates the identification of pollution sources. The factor loadings obtained after rotation indicate the degree of correlation between each parameter and each factor. The higher a loading, the more the corresponding factor contributes to explaining the variation of the parameter in question (Ielpo et al., 2012 ; Felipe-Sotelo et al., 2007).

4. Results and discussion.

4.1 Physicochemical parameters.

The descriptive statistics of the physicochemical parameters of groundwater samples collected from 60 wells during the period (2013-2019) are presented in Table 1. During this period, some physicochemical parameters showed a wide variation. The minimum and maximum concentrations of Ca^{2+}, Mg^{2+}, and TH ranged from 110.5 mg/L, 8.46 mg/L, and 0.43 mg/L to 237.9 mg/L, 69.14 mg/L, and 88.20 mg/L, respectively. Meanwhile, the concentrations of Na^{+}, Cl^{-}, and SO_4^{2-} ranged from 51.15 mg/L, 127.8 mg/L, and 48.48 mg/L to 194.39 mg/L, 461.5 mg/L, and 323 mg/L, respectively. The values of electrical conductivity and HCO_3^- varied from 563 µS/cm to 2890 µS/cm and from 134.1 mg/L to 484.3 mg/L.

The Piper diagram was used to represent the groundwater samples. (Fig. 2) to determine the dominant hydrochemical types of groundwater in the study area. **Cations (Bottom left triangle: Ca^{2+}, Mg^{2+}, Na^{+}+K^{+}):** The majority of points are located near the Ca^{2+} axis, indicating that calcium is the dominant cation in these water samples. There are also points in the region between Ca^{2+} and Mg^{2+}, suggesting a notable presence of magnesium but still with a dominance of calcium. The water samples are predominantly "Calcareous" or "Calcium-rich," with some samples showing an influence of magnesium, which could indicate waters originating from rock formations containing limestone or dolomite. **Anions (Bottom right triangle : Cl^{-}+NO_3^{-}, CO_3^{2-}+HCO_3^{-}, SO_4^{2-}) :** The points are mostly concentrated towards the central area but tend to align towards Cl^{-}+NO_3^{-}, indicating a certain predominance of chlorides and nitrates in the water samples. There are also points close to CO_3^{2-}+HCO_3^{-}, indicating a significant presence of bicarbonates.

Central diamond: The points are predominantly located in the "Sodium and potassium chloride or sodium sulfate" zone, classifying the water as sodium and potassium chloride, which means that sodium (Na^{+}) and potassium (K^{+}) ions are dominant, in association with chlorides (Cl^{-}). This composition is typical of water from aquifers influenced by geological formations containing salt or by saline contamination, such as in coastal areas.

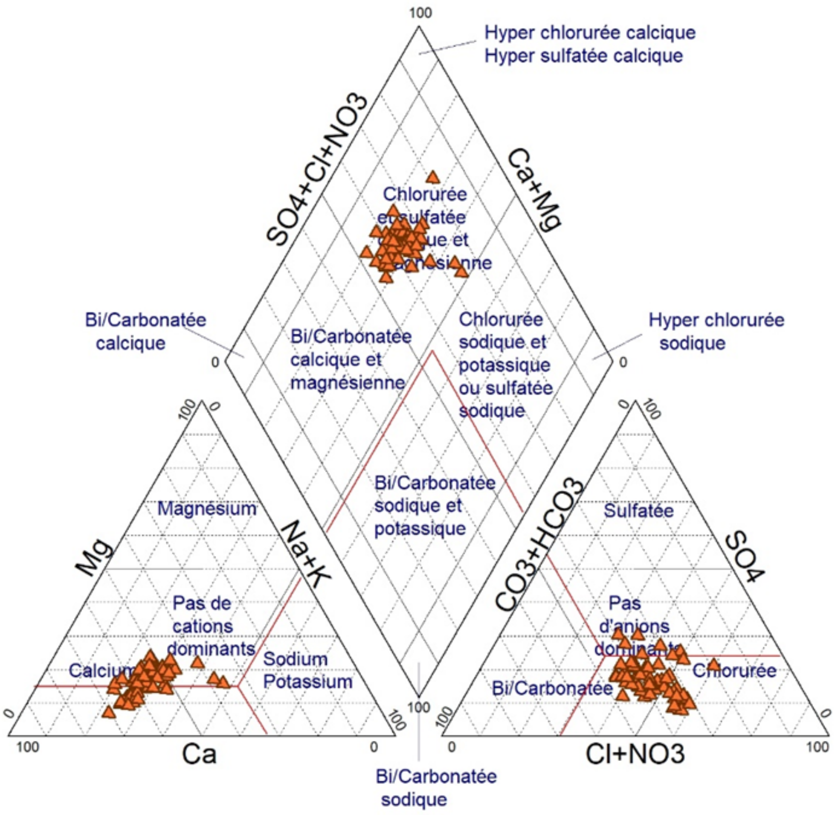

Figure 2. Piper diagram showing the chemical compositions of groundwater.

4.2 Correlation matrix analysis.

The correlation matrix (fig.3) provides a synthetic view of the relationships between different analysis parameters. The strong correlations between conductivity, magnesium, calcium, potassium, sodium, and bicarbonates suggest intense carbonate dissolution, such as calcite and dolomite, leading to an increase in water hardness. Additionally, significant links observed between bicarbonates, magnesium, chloride, and sulfates indicate complex geochemical processes related to the dissolution of carbonate and evaporite minerals, as well as ionic exchanges. Similarly, the strong correlation between sodium and chloride is mainly due to the fact that they are often present together in the form of dissolved salt in water (NaCl), which confirms the linear relationship observed in the correlation matrix. However, the weak correlation observed between sulfates and chlorides suggests that these elements could originate from different sources or have undergone distinct geochemical processes.

Lastly, the very narrow relationship between nitrates and other elements, with this weak correlation generally indicating that nitrates have their own dynamics, different from those of other elements. Additionally, nitrogen fertilizers are a major source of nitrates in groundwater, and their application can vary considerably in time and space, which may weaken the correlations with other parameters.

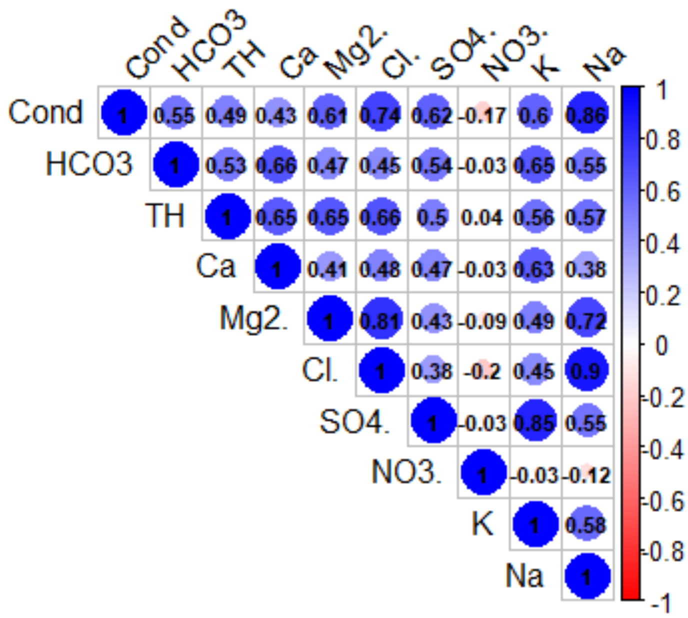

Figure3. correlation matrix of the groundwater quality parameters.

4.3 Principal component analysis.

The overall Kaiser-Meyer-Olkin (KMO) value is 0.706, which is less than 0.5, confirming the validity of applying PCA. According to the selection criteria, factors with eigenvalues greater than 1.0 were considered in this study. Two principal components (PC1, PC2) were extracted from the groundwater quality parameters, accounting for 70.35% of the total variance.
The eigenvalues, proportions of total variance, and cumulative proportions of variance associated with each principal component are tabulated in Table 2. The rotated factor loadings for the initial two principal components are graphically represented in Figures 4.

Table 2. Calculated eigenvalues and variance of each variances.

	PC1	*PC2*
Eigenvalues	5.66207164	1.27293290
% of variance	56.6207164 %	13.7293290 %
Cumulative %	56.62072 %	70.35005 %

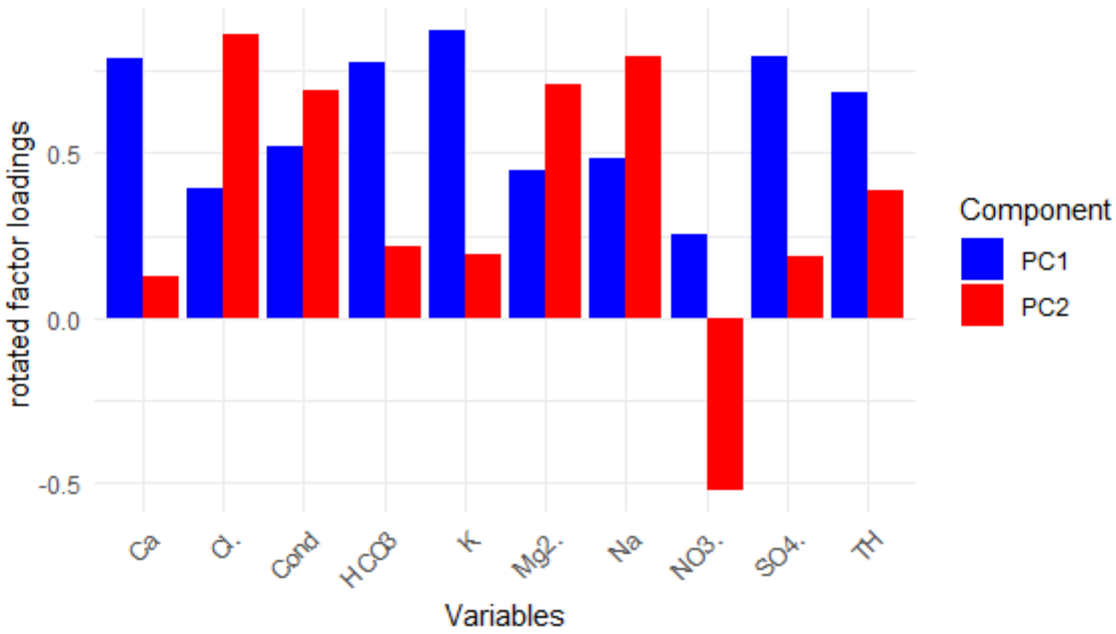

Figure.4 Variance of rotated factor loadings in Varimax rotated principal component analysis.

La PC1 représentait 56.6207% de la variance totale. Elle était fortement influencée par le K, les HCO_3^-, le SO_4, TH et le Ca^{2+}, la combinaison de ces facteurs expliquait la plus grande part de la variance dans l'ensemble de données. Ces variables sont peut-être influencées par des processus géochimiques similaires ou par des sources communes de pollution ou de minéralisation, ce qui entraîne leur forte influence collective sur la première composante principale (PC1) (Soltani, Asghari Moghaddam et al. 2017).

La PC2 révèle un autre schéma de variation dans les données, distinct de celui de la PC1, et met en évidence l'influence de ces ions et de la conductivité électrique sur la qualité de l'eau analysée, par exemple, la présence de Cl-, Na2+, et une conductivité élevée pourraient être associées à une origine commune, comme l'intrusion d'eau salée ou une source de pollution industrielle. De même, les concentrations de Mg2+ et NO3- pourraient être liées à des processus agricoles ou à la dissolution de minéraux spécifiques.

5. Conclusion.

This work analyzed the groundwater quality in the Guelma region of Algeria using multivariate statistical methods, including the correlation matrix, Piper diagram, and principal component analysis (PCA). The study revealed complex interrelations between various physicochemical parameters, suggesting geochemical processes influenced by natural and anthropogenic factors. Strong correlations between certain ions, such as sodium and chloride, as well as the identification of mixed hydrochemical types, indicate potential pollution sources, such as saltwater intrusion or agricultural contamination. The PCA simplified these complexities by highlighting the most influential variables, which could serve as key indicators for monitoring water quality. Overall, the results suggest that the groundwater chemistry in this region is significantly impacted by human activities and local geological conditions, calling for sustainable management and continuous monitoring to protect this vital resource.

6. Conflicts of Interest .

The authors declare no conflict of interest.

REFERENCES

1. Alabjah, B., F. Amraoui, M. Chibout and M. Slimani (2018). "Assessment of saltwater contamination extent in the coastal aquifers of Chaouia (Morocco) using the electric recognition." Journal of Hydrology 566: 363-376.

2. Ayed, B., I. Jmal, S. Sahal, N. Mokadem, S. Saidi, E. Boughariou and S. Bouri (2017). "Hydrochemical characterization of groundwater using multivariate statistical analysis: the Maritime Djeffara shallow aquifer (Southeastern Tunisia)." Environmental Earth Sciences 76: 1-22.
3. Baazi, H. (2023). "Aquifer Vulnerability Assessment by DRASTIC and SI methods: The case study of Guelma Plain, Northeast Algeria." Engineering, Technology & Applied Science Research 13(1): 10045-10050.
4. BENMARCE, K., A. CHOUABBI and L. BOUABSA (2014). "Variations spatiales des caractéristiques physico-chimiques des eaux souterraines de la plaine alluvionnaire de Guelma,(Nord-est algérien)." Bulletin des Sciences Géographiques-N 29(1er).
5. Bhutiani, R., D. B. Kulkarni, D. R. Khanna and A. Gautam (2016). "Water quality, pollution source apportionment and health risk assessment of heavy metals in groundwater of an industrial area in North India." Exposure and Health 8: 3-18.
6. Chabukdhara, M. and A. K. Nema (2012). "Assessment of heavy metal contamination in Hindon River sediments: a chemometric and geochemical approach." Chemosphere 87(8): 945-953.
7. Flörke, M., C. Schneider and R. I. McDonald (2018). "Water competition between cities and agriculture driven by climate change and urban growth." Nature Sustainability 1(1): 51-58.
8. Gulgundi, M. S. and A. Shetty (2018). "Groundwater quality assessment of urban Bengaluru using multivariate statistical techniques." Applied water science 8: 1-15.
9. Houria, B. and D. Naima (2023). "Assessment of Groundwater Vulnerability using GOD Method." International Journal of Innovative Studies in Sociology and Humanities 8: 185-189.
10. Islam, A. T., S. Shen, M. A. Haque, M. Bodrud-Doza, K. Maw and M. A. Habib (2018). "Assessing groundwater quality and its sustainability in Joypurhat district of Bangladesh using GIS and multivariate statistical approaches." Environment, development and sustainability 20: 1935-1959.
11. Ji, Y., J. Wu, Y. Wang, V. Elumalai and T. Subramani (2020). "Seasonal variation of drinking water quality and human health risk assessment in Hancheng City of Guanzhong Plain, China." Exposure and health 12: 469-485.
12. Latifi, S. and S. Chaab (2017). "Evaluation et cartographie de la vulnérabilité à la pollution des eaux souterraines, selon les méthodes DRASTIC et GOD, dans la plaine alluviale de Guelma, Nord-Est Algérien." Synthèse: Revue des Sciences et de la Technologie 34: 48-62.
13. Li, P., R. Tian and R. Liu (2019). "Solute geochemistry and multivariate analysis of water quality in the Guohua phosphorite mine, Guizhou Province, China." Exposure and Health 11: 81-94.
14. Li, P., R. Tian, C. Xue and J. Wu (2017). "Progress, opportunities, and key fields for groundwater quality research under the impacts of human activities in China with a special focus on western China." Environmental Science and Pollution Research 24: 13224-13234.
15. Li, Q., H. Zhang, S. Guo, K. Fu, L. Liao, Y. Xu and S. Cheng (2020). "Groundwater pollution source apportionment using principal component analysis in a multiple land-use area in southwestern China." Environmental Science and Pollution Research 27: 9000-9011.
16. Moukhliss, M. and A. Taleb (2021). "Assessment of groundwater quality in Berrechid Aquifer, Morocco." Moroccan Journal of Chemistry 9(4): J. Chem. 9 N° 4 (2021) 2800-2812.
17. Moukhliss, M., A. Taleb, A. Ouallali, V. Spalevic, N. Mazigh, B. El Fathi and S. Souabi (2023). "Assessment of Groundwater Quality in the Berrechid Aquifer, Central Morocco, Using Multivariate Statistics and Water Quality Indices." Ecological Engineering & Environmental Technology 24..
18. Moukhliss, Y., Y. Koubi, M. Alaqarbeh, H. Muzzammel Rehman, H. Maghat, A. Sbai, M. Bouachrine and T. Lakhlifi (2023). "Computational and retrosynthetic investigation of isoxazole-bearing chalcones as antioxidant activate compounds." ChemistrySelect 8(7): e202203908.
19. Qilin, L., C. Xiaodan, H. Shunsheng, H. Biao, R. Jinghua, G. Xueyuan, F. Jian and X. Hongting (2020). "Element geochemistry of selenium-enriched soil and its main sources in Jiangsu Province." Geology in China 47(6): 1813-1825.
20. Soltani, S., A. Asghari Moghaddam, R. Barzegar, N. Kazemian and E. Tziritis (2017). "Hydrogeochemistry and water quality of the Kordkandi-Duzduzan plain, NW Iran: application of multivariate statistical analysis and PoS index." Environmental monitoring and assessment 189: 1-20.
21. Strickert, M., F.-M. Schleif, T. Villmann and U. Seiffert (2009). "Unleashing Pearson correlation for faithful analysis of biomedical data." Similarity-Based Clustering: Recent Developments and Biomedical Applications: 70-91.
22. Vila, J.-M. (1980). "La chaîne alpine de l'Algérie orientale et des confins algéro-tunisiens." These de Doctorat-es-sciences, Universite Pierre et Marie curie.
23. Williams, B., A. Onsman and T. Brown (2010). "Exploratory factor analysis: A five-step guide for novices." Australasian journal of paramedicine 8: 1-13.
24. Wu, J., P. Li, H. Qian, Z. Duan and X. Zhang (2014). "Using correlation and multivariate statistical analysis to identify hydrogeochemical processes affecting the major ion chemistry of waters: a case study in Laoheba phosphorite mine in Sichuan, China." Arabian Journal of Geosciences 7: 3973-3982.

25. Wu, J., P. Li, D. Wang, X. Ren and M. Wei (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." Human and Ecological Risk Assessment: An International Journal.
26. Zhang, Q., H. Wang, Y. Wang, M. Yang and L. Zhu (2017). "Groundwater quality assessment and pollution source apportionment in an intensely exploited region of northern China." Environmental Science and Pollution Research 24: 16639-16650.