Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Disentangling Natural Mineralization from Nitrate Pollution in a Semi-Arid Alluvial Aquifer: A Log-Transformation-Based Multivariate and GIS Framework (Tébessa-Morsott Basin, Algeria)

Mehdi Bendekkoum
University of Tebessa
Karima Seghir
karima.seghir@univ-tebessa.dz

University of Tebessa
Ali Hadjela
University of Tebessa
Messaoud Abidi Saad
University of Tebessa
Vincent Valles
Avignon University

Research Article

Keywords: Groundwater vulnerability, Nitrate contamination, Log-transformation, PCA, HCA, GIS, Recharge zones, Tébessa basin, Algeria

Posted Date: July 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10264686/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Disentangling Natural Mineralization from Nitrate Pollution in a Semi-Arid Alluvial Aquifer: A Log-Transformation-Based Multivariate and GIS Framework (Tébessa-Morsott Basin, Algeria)

Mehdi Bendekkoum1, Karima Seghir[1] Ali Hadjela[1] Messaoud Abidi Saad[1] and Vincent Valles[2,3]

[1].Laboratoire Environnement Sédimentaires, Ressources Minérales et Hydriques de l'Algerie Orientale Université Echahid Cheikh Larbi Tébessi, Route de Constantine, 12002, Tébessa, Algeria ;

mehdi.bendekkoum@univ-tebessa.dz , Karima.seghir@univ-tebessa.dz , ORCID 0009-0003-6753-5720

ali.hadjela@univ-tebessa.dz , messaoud.abidisaad@univ-tebessa.dz

[2].Mixed Research Unit EMMAH (Environnement Méditerranéen et Modélisation des Agro-Hydrosystèmes), HydrogeologyLaboratory, Avignon University, Avignon, 84916, France ; vincent.valles@outlook.fr

[3].Faculté des Sciences et Techniques de Beni Mellal (FSTBM), BP 523, 23000 Beni Mellal Morocco

Corresponding author: K. Seghir, **email**; karima.seghir@univ-tebessa.dz

ORCID 0009-0003-6753-5720

Abstract

Groundwater quality in semi-arid Mediterranean aquifers reflects the joint action of geogenic mineralization and anthropogenic contamination, yet few studies statistically separate the two. This study addresses that gap in the shallow alluvial aquifer of Tébessa (NE Algeria) using a multi-temporal dataset (2005–2018) from wells, boreholes, springs, and surface water. Measured electrical conductivity ranged from 442 to 9120 µS cm^{-1}, whereas nitrate concentrations varied between 1.52 and 117.97 mg L^{-1}, with several groundwater samples exceeding the WHO drinking-water guideline of 50 mg L^{-1}.

Because raw hydrochemical variables were non-normally distributed, logarithmic transformation was applied before Principal Component and Hierarchical Cluster Analyses. This step proved decisive: nitrate consistently separated from mineralization-related variables (TDS, Cl^-, Na^+, SO_4^{2-}) only after transformation, confirming its anthropogenic origin. Carbonate and evaporite dissolution govern natural mineralization, whereas nitrate enrichment is principally associated with agricultural return flow and wastewater infiltration affecting shallow recharge zones. GIS mapping further identifies the recharge zones most exposed to contamination.

This log-transformation-based framework offers a replicable diagnostic pathway for distinguishing natural evolution from pollution in comparable semi-arid aquifers.

Keywords: Groundwater vulnerability; Nitrate contamination; Log-transformation; PCA; HCA; GIS; Recharge zones; Tébessa basin; Algeria

Authors's email; mehdi.bendekkoum@univ-tebessa.dz, Karima.seghir@univ-tebessa.dz , ali.hadjela@univ-tebessa.dz , messaoud.abidisaad@univ-tebessa.dz ,vincent.valles@outlook.fr

1. Introduction

Water security in semi-arid Mediterranean environments rests heavily on groundwater, precisely because surface-water availability is erratic. Rainfall arrives irregularly, evaporation is intense, and domestic and agricultural demand keeps climbing — a combination that leaves shallow aquifers exposed. Their limited depth and close hydraulic connection to the surface mean that whatever happens above ground is quickly transmitted below it (Scanlon et al., 2012; FAO, 2020, Bendekkoum, M., 2025)

Across the southern Mediterranean circumference, including North Africa, this exposure is intensifying. Climatic variability, recurrent drought, demographic growth, and the steady expansion of irrigated land are together reducing natural recharge while increasing the flux of dissolved

contaminants toward aquifer systems. Nitrate occupies a special place among these contaminants: unlike many geogenic species, it travels easily through the unsaturated zone, resists degradation, and is almost always traceable to human activity — nitrogen fertilization, irrigation drainage, or poorly managed sewage. Because rural communities frequently depend on untreated groundwater for drinking, nitrate loading is as much a public-health issue as an environmental one (Ward et al., 2018; World Health Organization, 2017, 2023).

Tébessa, on Algeria's northeastern border, embodies these pressures concretely. Its population and agriculture rely on a shallow alluvial aquifer fed by fractured carbonate recharge zones. Limited natural recharge, combined with sustained groundwater abstraction, leaves this system structurally vulnerable to surface-derived contamination (Coleman et al., 2013; Abidi Saad et al., 2024a; Abidi Saad et al., 2024b, Bendekkoum, M., 2025).

A substantial body of prior work on northeastern Algerian aquifers has already established that groundwater chemistry there results from interplay of natural water-rock interaction and human land use. What most of that literature has not done, though, is connect hydrochemical characterization directly to the spatial geometry of recharge — that is, to show where the aquifer is being recharged and whether those same zones are the ones absorbing nitrate loading (Marjoua et al., 1997; Fakir et al., 2001; El Moujabber et al., 2006; Ben Hamouda et al., 2009; Kouzana et al., 2009; Mondal et al., 2010; Malana & Khosa, 2011; Belghiti et al., 2013; Yu et al., 2013; Kim et al., 2014; Kumar, 2014, Bendekkoum, M., 2025). Bridging hydrochemistry, multivariate statistics, and spatial recharge mapping into a single diagnostic exercise remains comparatively rare, even though each piece individually is well established in the groundwater-vulnerability literature (Aghzar, 2002; Wakida & Lerner, 2005; Rivett et al., 2008).

This is precisely the gap the present study targets. We integrate hydro-chemical characterization with multivariate statistics (PCA, HCA) and GIS-based spatial analysis into one coherent framework, applied to the Tébessa shallow aquifer, with four connected aims: (i) to trace how groundwater quality shifts across space and across four sampling campaigns; (ii) to isolate the natural versus human-driven controls on water chemistry; (iii) to test whether nitrate distribution aligns with — or diverges from — the aquifer's hydrogeological vulnerability; and (iv) to translate these findings into management recommendations suited to semi-arid conditions.

What distinguishes this work methodologically is the deliberate use of *logarithmic transformation* as a diagnostic device rather than a mere normalization step. Because groundwater chemical variables are almost never normally distributed, working with raw data risks letting a handful of extreme samples dominate the statistical structure and obscure the true relationships among variables. By systematically testing normality, transforming the dataset logarithmically, and only then running PCA and HCA, we obtain a cleaner separation between the geogenic mineralization signal and the anthropogenic nitrate signal — a separation that the raw-data analysis does not deliver with the same clarity. Coupling that statistical gain with GIS mapping of recharge pathways lets us pinpoint, rather than merely describe, the sectors where geology and agriculture jointly maximize contamination risk. We see this pairing — log-transformed multivariate discrimination plus recharge-zone mapping — as a transferable diagnostic pathway for other vulnerable semi-arid Mediterranean aquifers, and as the main original contribution of this study.

2. Study Framework and Analytical Procedures

2.1. Study area

The study area sits astride Algeria's northeastern border with Tunisia, occupying the central portion of the Tébessa-Morsott graben. It spans roughly 140 km² between 35°10′ and 35°30′ N and 7°50′ and 8°20′ E, within the wider Medjerda-Méllègue catchment. Tébessa city, the region's administrative and demographic center, sits near the plain's eastern edge at an elevation close to 820 m (Figure 1).

A semi-arid Mediterranean regime dominates locally, marked by irregular rainfall, high summer evaporation, and recurring seasonal water deficits (Seghir, 2008; Seghir, 2014; Abidi Saad et al., 2014a; Abidi Saad et al., 2024b;Bendekkoum, M.,2025)

Long-term rainfall averages around 397 mm/year, concentrated in short wet spells; the remaining, longer dry periods drive evapotranspiration losses that consume roughly 91.5% of annual precipitation, leaving only a small fraction to feed aquifer recharge (Bendekkoum, 2017; Seghir, 2008; Bendekkoum, M.,2025). Combined with sustained groundwater withdrawal for irrigation and domestic use, this weak recharge regime leaves the shallow aquifer structurally exposed: pollutants originating at the surface, particularly from agriculture, can reach groundwater with little natural attenuation.

The low recharge potential, combined with increasing groundwater abstraction for irrigation and domestic supply, makes the shallow aquifer particularly sensitive to contamination. Under these conditions, pollutants originating from agricultural activities and surface sources can rapidly affect groundwater quality [Bendekkoum, M., 2025]

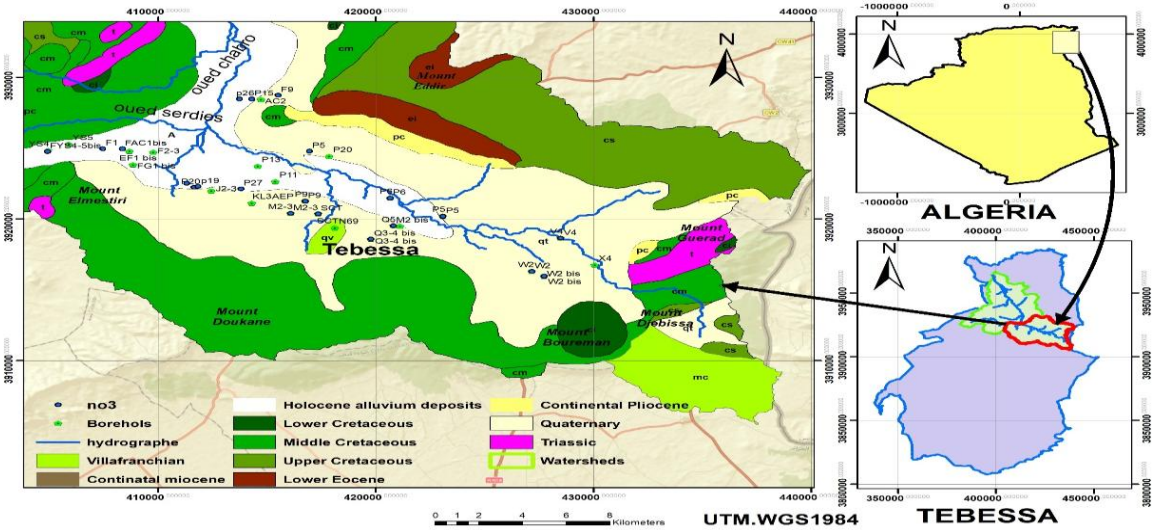


Figure 1. Geographical setting of the study area [Bendekkoum, M.,2025]

2.2. Geological and Hydrogeological Framework

The Tébessa-Morsott basin is a tectonic graben shaped by successive deformation phases from the Mesozoic through the Quaternary, bounded by carbonate ranges and dissected by fault systems that steer groundwater movement. To the north, the Tébessa Boulhaf-Dyr high is framed by the Djebel Dyr and Djebel Guenifida escarpments; to the south, the Dj.Doukane, Dj.Mistiri, Dj.Tazent, and Dj.Bouziane massifs complete the enclosure the northern peaks exceeding 1,400 m near Kef Ahmed, the southern ones surpassing 1,600 m at Djebel Doukane, both sharply carved by erosion.

These surrounding reliefs consist mainly of fractured Cretaceous-Eocene limestones and dolomites. Their fracture-and-karst networks make them the aquifer's principal recharge conduits, channeling infiltrated rainfall toward the central depression (Durozoy, 1949) (Figure 2). That depression itself is filled with Mio-Pliocene to Quaternary alluvium — an uneven mix of gravel, sand, silt, and clay — whose heterogeneity governs how water is stored, how easily it moves, and how contaminants migrate once they enter the system. The Kebir-Chabro-Ksob wadi, draining the watershed along a gentle southeast-northwest slope, is the plain's principal surface conduit.

The aquifer itself is an unconfined shallow system hosted in these alluvial deposits, shallowest near the drainage network and deepening toward the basin margins, with flow following the same southeast-northwest trend imposed by regional structure. Recharge arrives mainly as lateral inflow from the surrounding fractured carbonates, supplemented occasionally by direct infiltration from the wadi during high-flow events. It is this combination active recharge pathways, permeable alluvium, and a shallow water table that renders the aquifer particularly susceptible to nitrate infiltration from the surface [Bendekkoum, M., 2025]

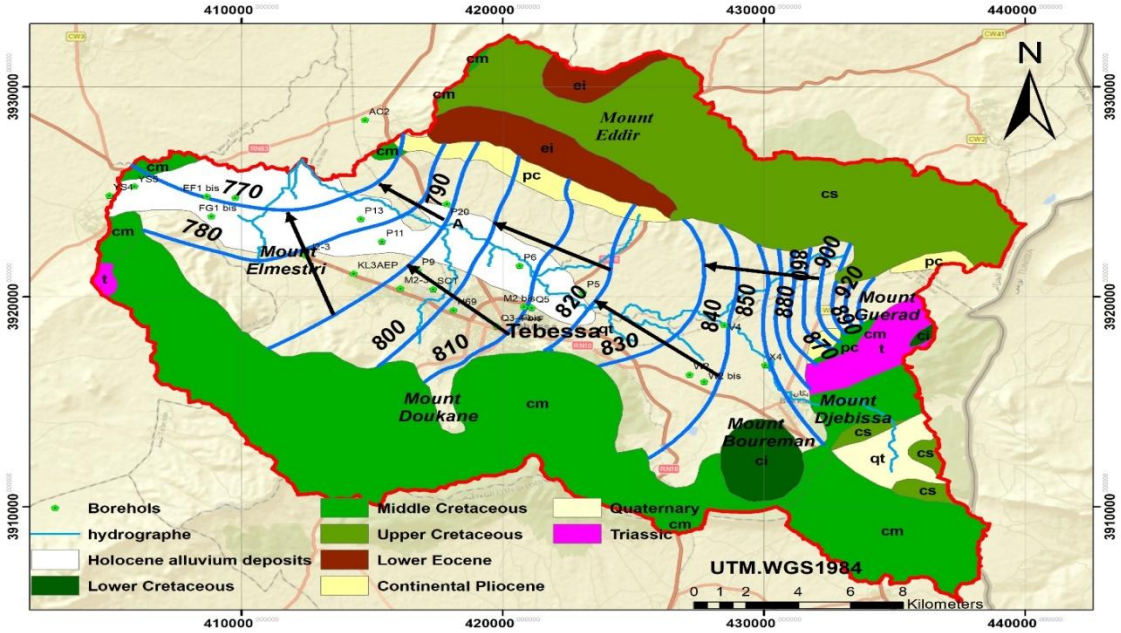


Figure 2. Geological and piezometric map (May, 2021) of the study area [Bendekkoum, M.,2025]

2.3. Water Sample Collection

Sampling targeted the full range of water types present in the basin — shallow wells, boreholes, springs, and wadi stations — so as to capture hydrogeological contrasts across the plain. Four campaigns (May 2005, November 2005, May 2006, and 2018) allowed both seasonal and multi-year hydrochemical shifts to be tracked, with sites chosen according to their hydrogeological setting, groundwater depth, and use.

Boreholes were purged of stagnant water before sampling to ensure the aliquot represented actual aquifer conditions; well and spring samples followed standard collection procedures. All samples were filtered through 0.45 µm membranes prior to laboratory transfer. Field campaigns used calibrated instruments to record pH, temperature, conductivity, and alkalinity on site, while major cations (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}) were quantified by ICP-OES and major anions (Cl^{-}, SO_4^{2-}, HCO_3^{-}, NO_3^{-}) by ion chromatography in the laboratory. Calibration standards, duplicate runs, and ionic-balance checks (samples exceeding ±5% error were discarded) safeguarded analytical reliability.

2.4. Statistical and spatial analysis

Before multivariate analysis, the distribution of hydro-chemical variables was evaluated using normality tests. Because groundwater chemical parameters commonly exhibit asymmetric distributions caused by multiple interacting processes, logarithmic transformation was applied when necessary to improve statistical performance.

To separate natural geochemical drivers from anthropogenic ones and pinpoint the dominant controls on water chemistry, Principal Component Analysis was employed; sampling points were subsequently grouped by chemical resemblance using Hierarchical Cluster Analysis.

GIS-based spatial analysis was used to map the distribution of major chemical parameters and nitrate concentrations and to identify sectors where recharge conditions, aquifer characteristics, and human activities combine to increase groundwater vulnerability.

3. Results and discussion

3.1. 3.1 Evolution of Hydro-chemical Facies

Several interacting factors — recharge dynamics, lithology, residence time, and human pressures — jointly shape the hydro-chemical signature observed across the Tébessa-Morsott basin. The comparison between surface water, shallow wells, and boreholes highlights a clear spatial variability in groundwater evolution within the aquifer system (Table 1).

Across the sampled network, pH values cluster tightly between 7.01 and 8.33 — a near-neutral to mildly alkaline signature consistent with a system dominated by carbonate buffering; calcite and dolomite dissolution keep pH stable regardless of sampling point.

Electrical conductivity tells a different, more heterogeneous story, ranging from 442 to 9120 µS/cm. Surface water remains comparatively fresh (average 456.66 µS/cm), boreholes occupy an intermediate position (average 1067.62 µS/cm), while shallow wells show markedly elevated mineralization (average 2414.11 µS/cm) — a gradient suggesting that shallow groundwater undergoes far more intense mineral enrichment than water tapped at depth. Prolonged water-rock contact, dissolution of soluble evaporitic minerals, and evaporative concentration under the region's semi-arid climate together explain the elevated conductivity observed locally; the Triassic evaporitic formations flanking the basin supply much of the dissolved sulfate and chloride responsible for this pattern.

Total dissolved solids follow the same gradient — low in surface water, climbing to 7968 mg/L in shallow groundwater — reinforcing the picture of a shallow aquifer where limited recharge and slow turnover allow dissolved constituents to accumulate.

Table 1. Mean physico-chemical parameters measured for groundwater across the study zone

Parameter	pH	Sal. g/l	Cond. µS/cm	TDS mg/l	Ca^{2+}	Mg^{2+}	Na^{+}	K^{+}	HCO_3^{-}	NO_3^{-}
Springs (avg)	7.80	0	456.66	203	75.48	7.33	9.12	0.82	316.76	23.03
Wells (avg)	7.42	1.12	2414.11	2087.91	172.78	107.62	482.42	2.17	361.47	38.67
Boreholes (avg)	7.47	0.33	1067.62	675.65	93.38	26.77	182.55	0.5	312.01	28.87
WHO Standard	>6.5<9.5	-	<400	1500	100	50	150	12	-	50

3.2. Nitrate contamination and anthropogenic impact

Nitrate concentrations swing widely — from 1.52 to 117.97 mg/L — with a meaningful fraction of samples breaching the WHO's 50 mg/L guideline in specific sectors of the aquifer (World Health Organization, 2017, 2023). Surface water carries comparatively modest nitrate loads (16.37–27.45 mg/L); shallow wells, by contrast, register the highest values in the dataset, while borehole water generally — though not universally — shows lower concentrations (Figure 3).

This shallow-versus-deep contrast is itself diagnostic of how nitrate travels through the system. Shallow wells sit close to surface recharge pathways and benefit from little natural filtration, so contamination reaches them

quickly; deeper groundwater, having percolated through more aquifer material over longer residence times, shows attenuated — though not absent — signals (Rivett et al., 2008).

Spatially, nitrate loading tracks agricultural land use and shallow-water-table zones rather than appearing at random. Excess nitrogen fertilization, irrigation return flow, and untreated wastewater infiltration stand out as the principal sources, amplified in the Tébessa plain by the coincidence of intensive cultivation and a shallow, poorly protected water table (Rouabhia et al., 2010; Baazi Houria et al., 2020). That said, human activity alone does not fully explain the pattern: permeable alluvium, shallow water tables, and active recharge pathways are hydrogeological preconditions that determine how readily surface-derived nitrate actually reaches the aquifer, independent of land-use intensity (Strebel et al., 1989; Wakida & Lerner, 2005).

However, nitrate distribution cannot be explained only by human activities. Hydrogeological factors strongly control contaminant migration. Areas characterized by permeable alluvial deposits, shallow water tables, and active recharge pathways present higher vulnerability because pollutants can rapidly reach groundwater before significant attenuation occurs.

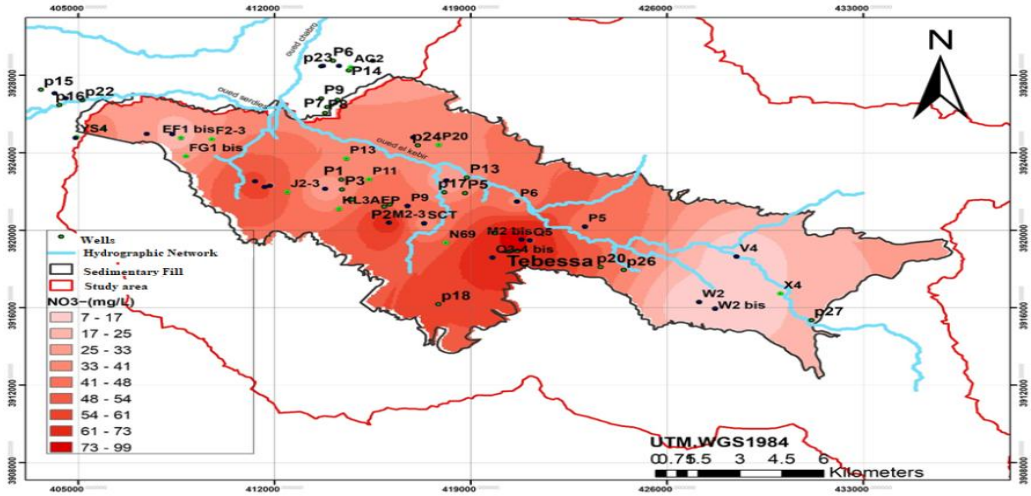


Figure 3.Spatial distribution of nitrate concentrations, highlighting sectors vulnerable to surface-derived contamination [Bendekkoum; M.2025]

3.3. Major ions distribution and geochemical processes

Chloride (Cl^-) and Sodium (Na^+)

Chloride shows marked enrichment in both the western and eastern sectors of the basin, near Aïn Chabro and Bekkaria respectively. Because chloride behaves conservatively — it neither precipitates nor is readily taken up biologically — its spatial pattern is a reliable tracer of solute origin. Here, that pattern points to a mixed origin: evaporitic mineral dissolution contributes a natural baseline, while agricultural return flow and wastewater infiltration add an anthropogenic increment in more intensively farmed sectors. The parallel between chloride enrichment and elevated conductivity suggests that both mineral dissolution and evaporative concentration jointly drive the basin's salinization under semi-arid conditions (Figure 4).

Nitrate and chloride enrichment are thus both linked, at least partly, to rigorous agriculture: extreme nitrogen fertilization and unrestrained irrigation applies are known to damage groundwater quality through diffuse nitrate and chloride contamination elsewhere in comparable settings, posing a genuine public-health risk where rural communities depend on untreated groundwater (World Bank, 1995; Collin, 2004, Bendekkoum, M., 2025).

Sodium (Figure 5) mirrors chloride closely enough to imply a shared evaporitic origin, though cation exchange within clay-rich alluvial sediments — which releases Na^+ while retaining Ca^{2+} and Mg^{2+} — likely modifies concentrations further along the flow path. Sodium enrichment therefore reflects a layered process: initial mineral dissolution, progressive chemical evolution along groundwater flow, and secondary ion-exchange effects.

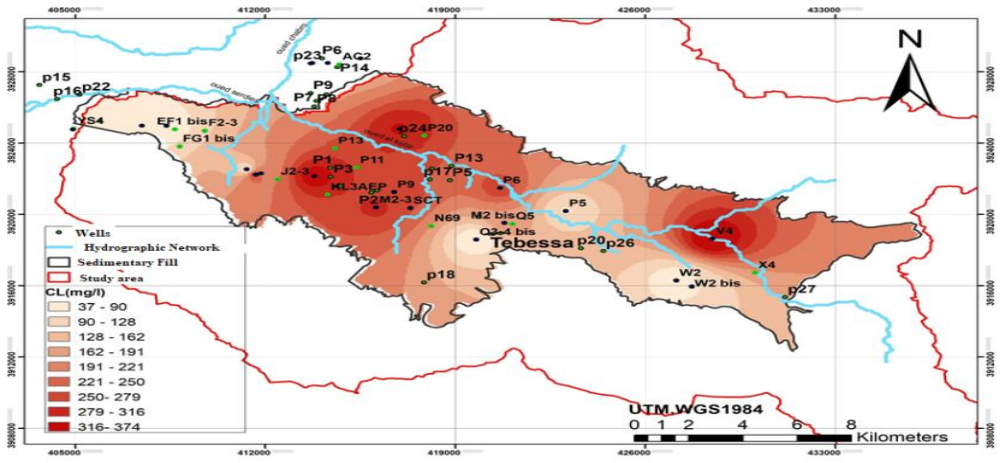


Figure 4: Spatial distribution of chloride concentrations[Bendekkoum; M.2025]

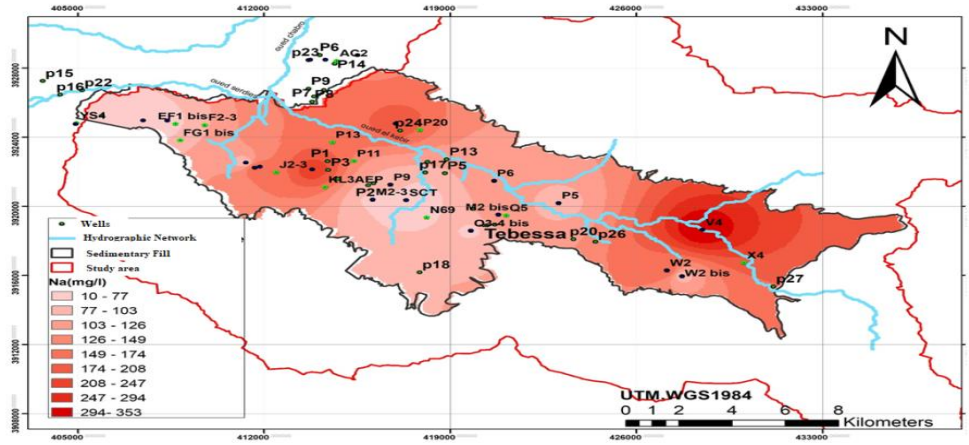


Figure 5: Spatial distribution of sodium concentrations[Bendekkoum; M.2025]

Sulfate (SO_4^{2-})

Sulfate distribution is strongly controlled by geological conditions. The highest concentrations occur near areas influenced by Triassic evaporitic formations, especially gypsum-bearing units (Figure 6).

Sulfate distribution correlates spatially with the Triassic evaporitic outcrops — gypsum-bearing units near the mountains Dj. Djebissa and Dj. Belkefif in the east to the west respectively— pointing to gypsum and anhydrite dissolution as the dominant geogenic control. Agricultural amendments and evaporative concentration can locally intensify sulfate levels, particularly in shallow groundwater, but the primary driver remains geological rather than anthropogenic[Bendekkoum; M.2025] .

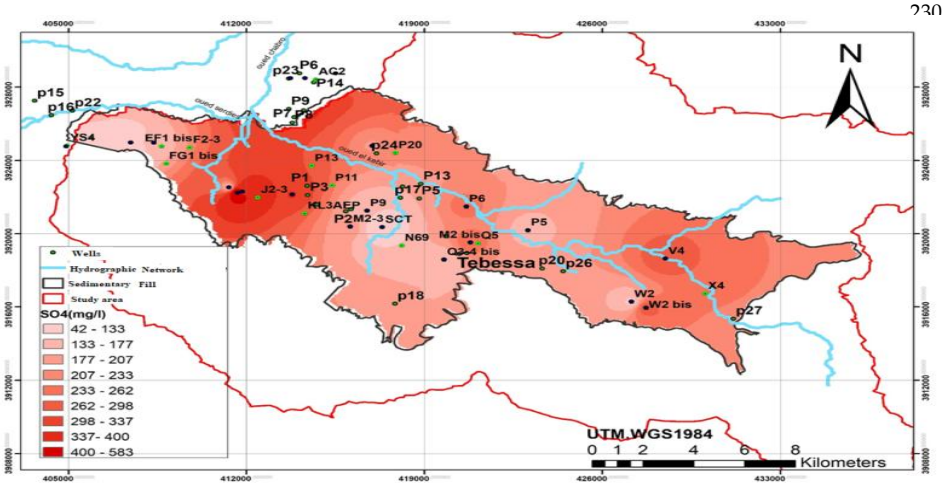


Figure 6: Spatial distribution of sulfate concentrations[Bendekkoum; M.2025]

Taken together, the spatial maps of NO_3^-, Cl^-, SO_4^{2-}, and Na^+ reveal a basin whose chemistry is shaped simultaneously by geology, hydrogeological setting, climate, and human land use. Where multiple vulnerability factors — permeable lithology, shallow water table, proximity to evaporitic outcrops — match, the risk of nitrate reaching the groundwater rises accordingly (Strebel et al., 1989; Demarets et al., 1995; Robert et al., 1994, Bendekkoum, M., 2025)

Deep boreholes and springs generally retain a calcium-bicarbonate signature indicative of better-preserved quality, whereas shallow wells and wadi water show the clearest pollution signals. In the northwestern, downstream sectors, unprocessed urban wastewater and agro-industrial overflow — both rich in nitrogen complexes — compound the agricultural contribution to nitrate loading, while prolonged drought further limits the dilution and renewal capacity of the aquifer [Bendekkoum, M., 2025]

Ultimately, the basin's hydrochemical profile confirms that natural geological inputs and human-driven contamination act jointly, with the shallow alluvial aquifer bearing the brunt of both because it is simultaneously the most actively recharged and the most exposed compartment of the system.

3.4. Multivariate statistical analysis and identification of groundwater controlling processes

3.4.1. Normality Test and Data Transformation

Before any multivariate treatment, we screened each hydrochemical variable for normality. Groundwater parameters routinely fail this test because several overlapping processes — geogenic and anthropogenic alike — shape their distributions simultaneously; ignoring that fact and running PCA directly on raw data risks letting a few extreme samples dictate the entire factorial structure.

Normality-test outcomes appear in Tables 2 through 4, while the corresponding P-P and Q-Q diagrams — covering TDS, sulfate, and nitrate — are grouped in Figures 7 to 9.

Table 2. *Normality test results for TDS.*

Test	Raw data	Log-transformed	Interpretation
Shapiro-Wilk	< 0.0001	0.039	Improved
Anderson-Darling	< 0.0001	0.190	Improved

Lilliefors	< 0.0001	0.590	Improved
Jarque-Bera	< 0.0001	0.025	Improved

Table 3. *Normality test results for NO_3^-.*

Test	Raw data	Log-transformed	Interpretation
Shapiro-Wilk	< 0.0001	< 0.0001	Persistent non-normality
Anderson-Darling	< 0.0001	< 0.0001	Persistent non-normality
Lilliefors	< 0.0001	< 0.0001	Persistent non-normality
Jarque-Bera	< 0.0001	< 0.0001	Persistent non-normality

Table 4. *Normality test results for SO_4^{2-}.*

Test	Raw data	Log-transformed	Interpretation
Shapiro-Wilk	< 0.0001	0.125	Improved
Anderson-Darling	< 0.0001	0.049	Improved
Lilliefors	< 0.0001	0.118	Improved
Jarque-Bera	< 0.0001	0.268	Improved

We therefore applied a logarithmic transformation systematically, and then re-tested normality using four independent tests (Shapiro-Wilk, Anderson-Darling, Lilliefors, Jarque-Bera) applied in parallel to guard against any single test's blind spots. TDS and SO_4^{2-} responded well: p-values rose substantially after transformation, and Q-Q/P-P plots aligned much more closely with a normal distribution. Nitrate behaved differently — normality remained rejected even after transformation — which we interpret not as a statistical failure but as a signal in itself: NO_3^- concentrations are governed by localized, irregular anthropogenic inputs (specific fields, specific wastewater sources) rather than by the smoother, basin-wide geogenic processes that govern TDS or sulfate (Figure 7a,8a,9a) and (Figure 7b, 8b;9b)

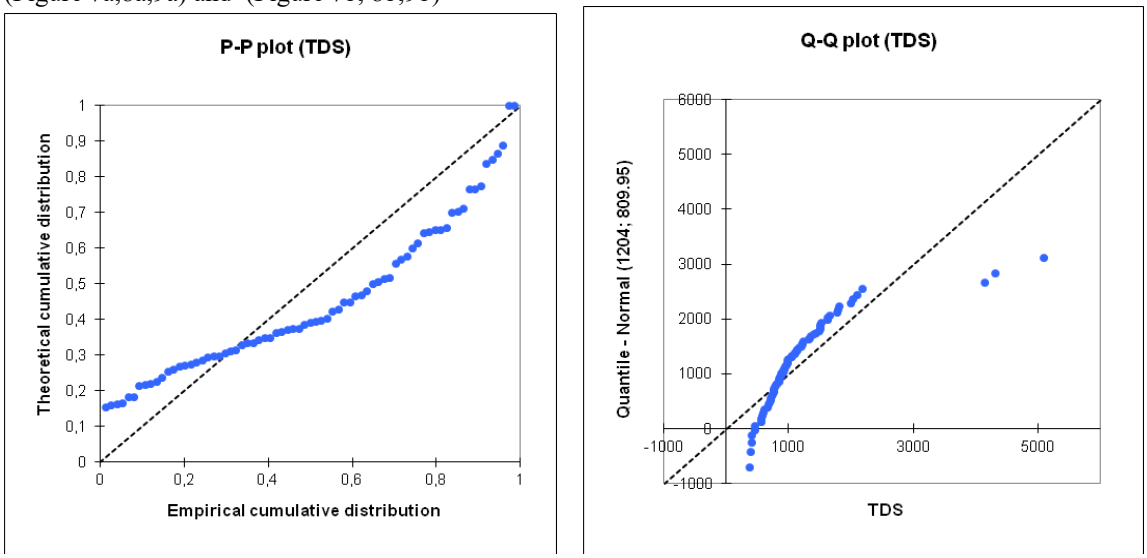


Figure 7a.; Q-Q plot and P-P plot for TDS (normal data) **[Bendekkoum, M., 2025]**

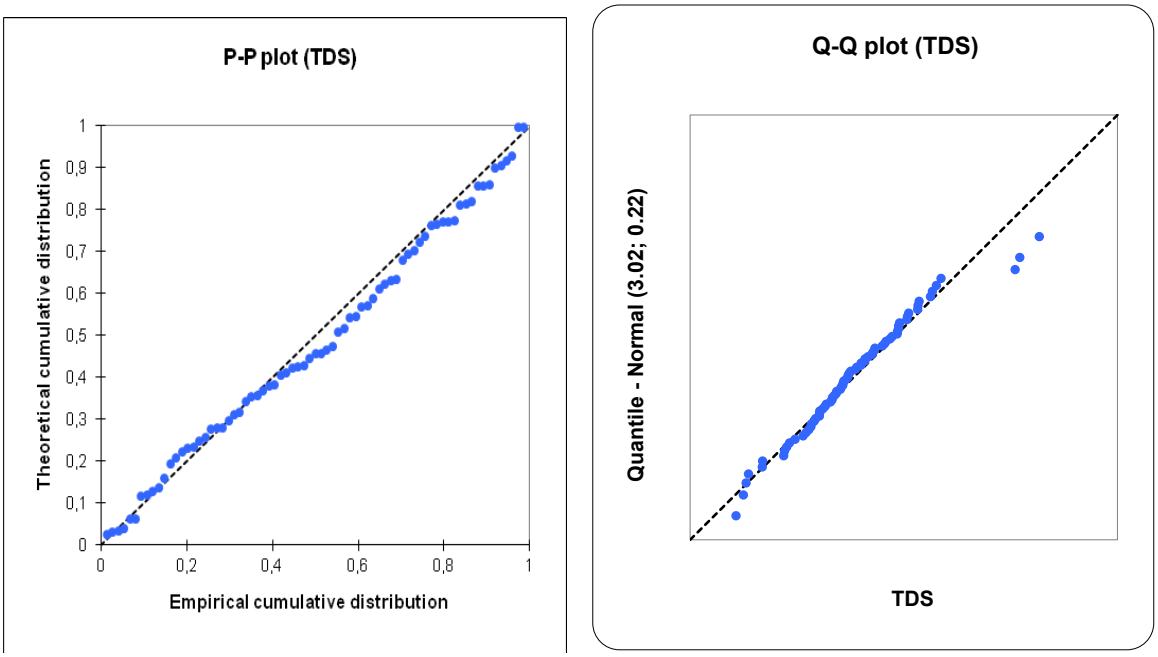


Figure 7b. : Q-Q plot and P-P plot for TDS (log data)[Bendekkoum, M., 2025]

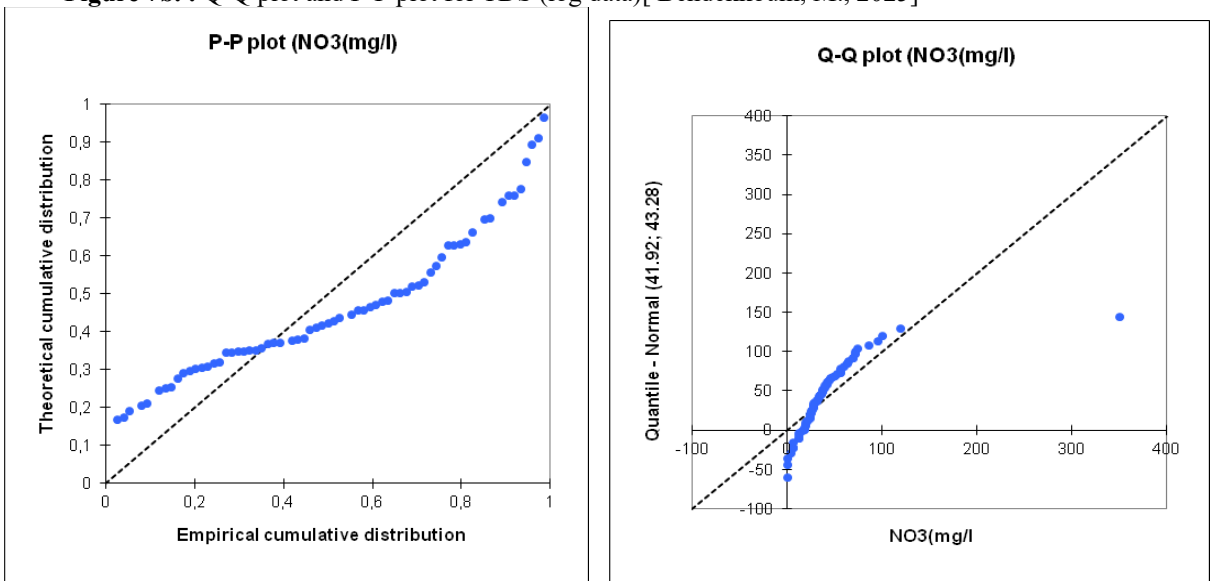


Figure 8a. Q-Q plot and P-P plot for normal NO3 data[Bendekkoum, M., 2025]

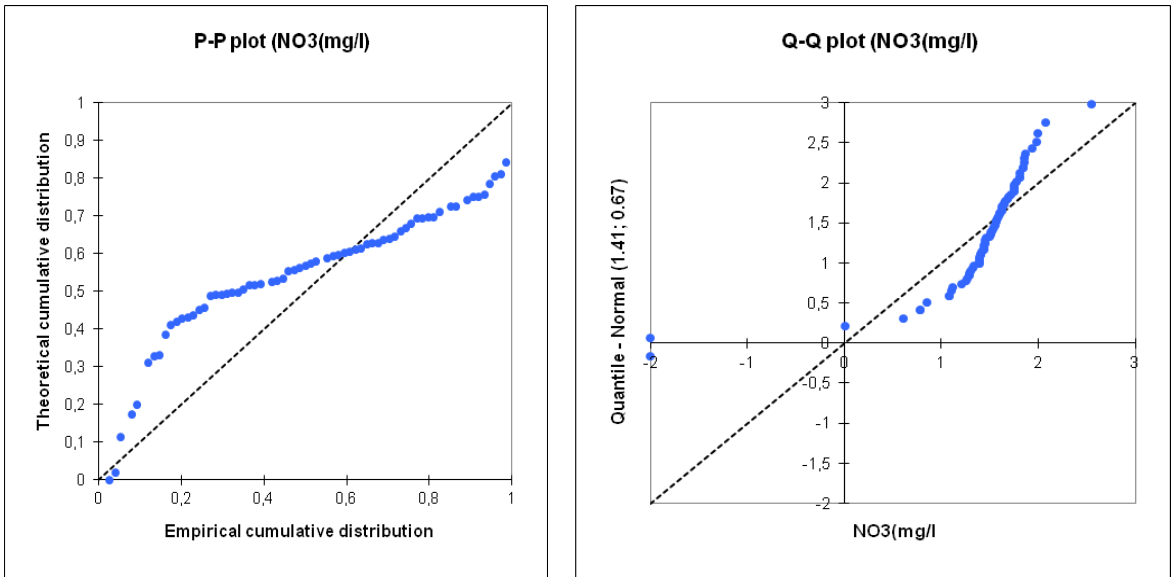


Figure 8b Q-Q plot. and P-P plot for log NO3 data**[Bendekkoum, M., 2025]**

This asymmetry — most variables normalizing under log-transformation, nitrate resisting it — is precisely what justifies adopting log-transformed data for the PCA and HCA that follow.

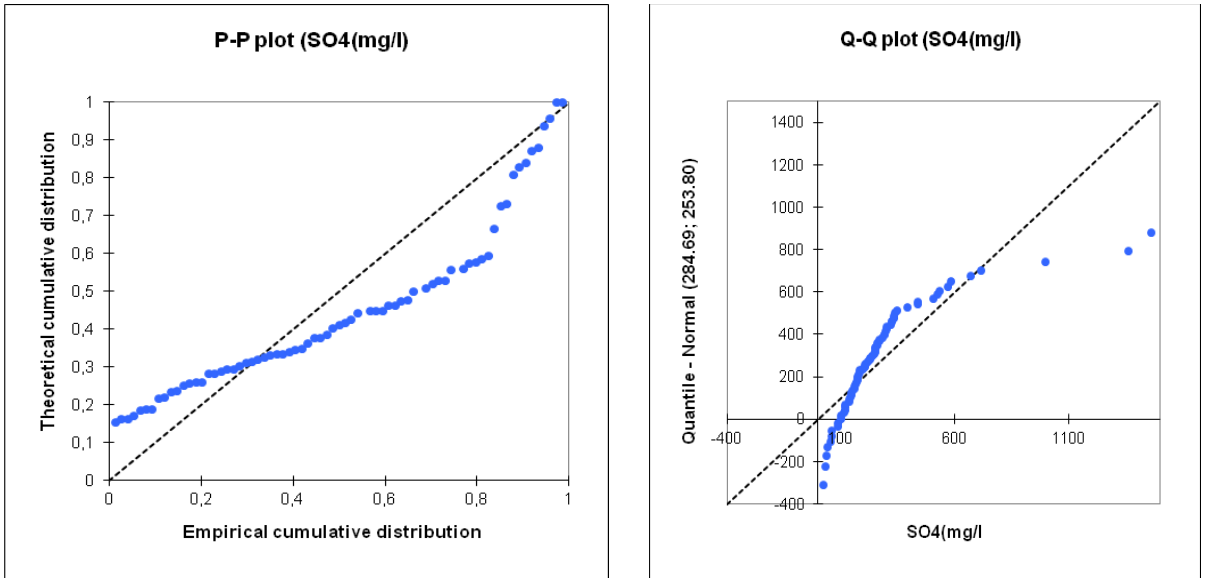


Figure 9a. : Q-Q plot and P-P plot for normal SO4 data [Bendekkoum, M., 2025]

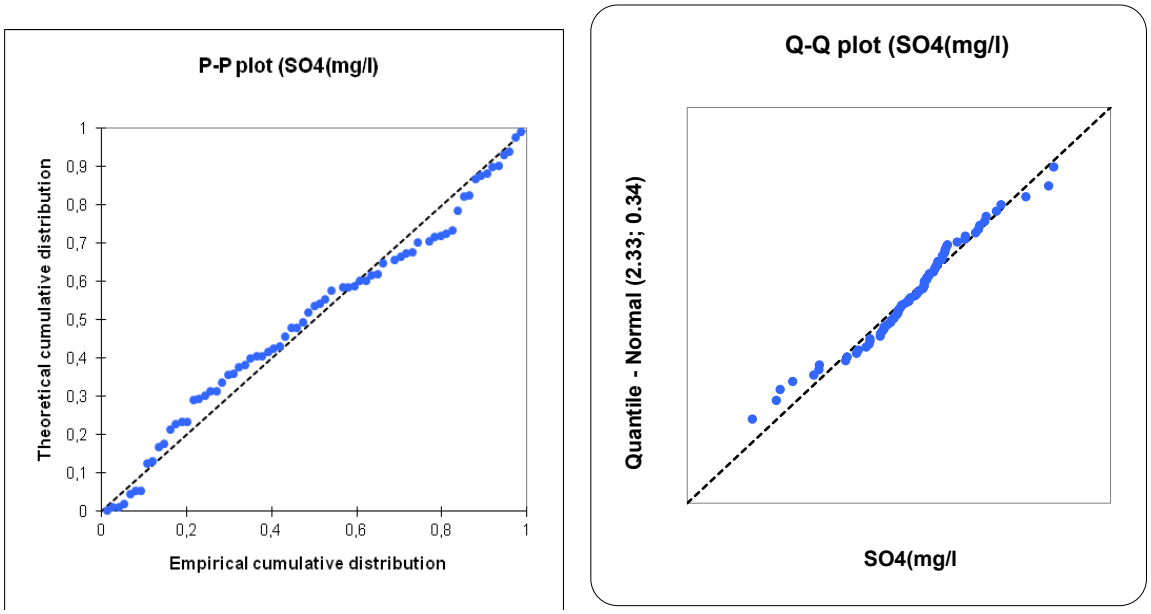


Figure 9b.: Q-Q plot and P-P plot for log SO4 data[Bendekkoum, M., 2025]

Beyond the p-values, the practical benefit is visible in the factorial plane itself: with raw data, three samples (AC2, P1, P1) sit as clear outliers, pulling the rest of the point cloud into a single crowded sector and obscuring genuine structure. Once log-transformed, the same samples distribute evenly across the plane, letting the underlying hydrochemical relationships emerge clearly. We regard this transformation step — applied deliberately and validated through multiple normality tests rather than assumed — as a central methodological contribution of this work, since it is what ultimately allows nitrate's independent, anthropogenic behavior to be statistically isolated from the basin's natural mineralization trend (Figures 10a, 10b).

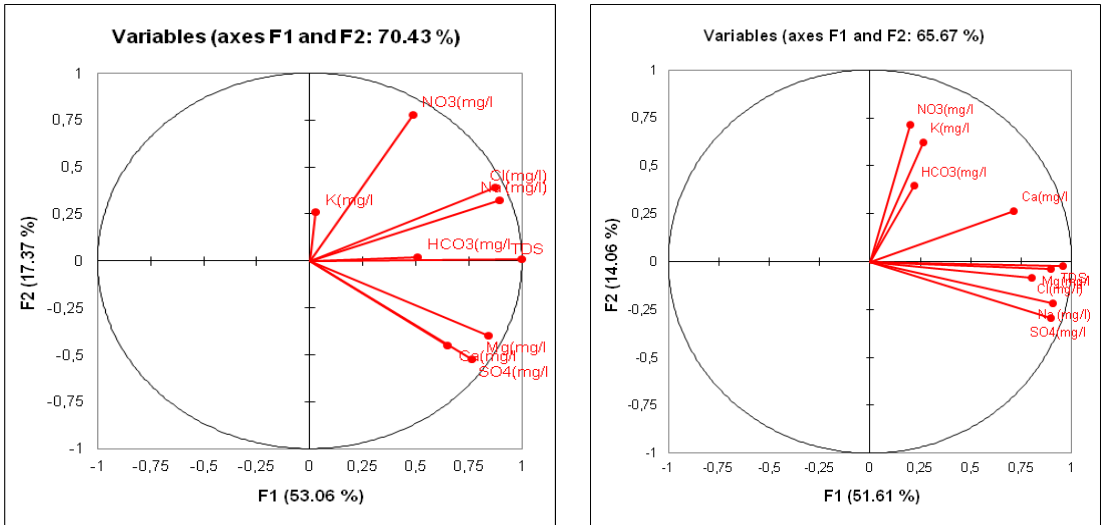


Figure 10.a: correlation circles of samples in the first factorial plane for untransformed data[Bendekkoum, M., 2025]

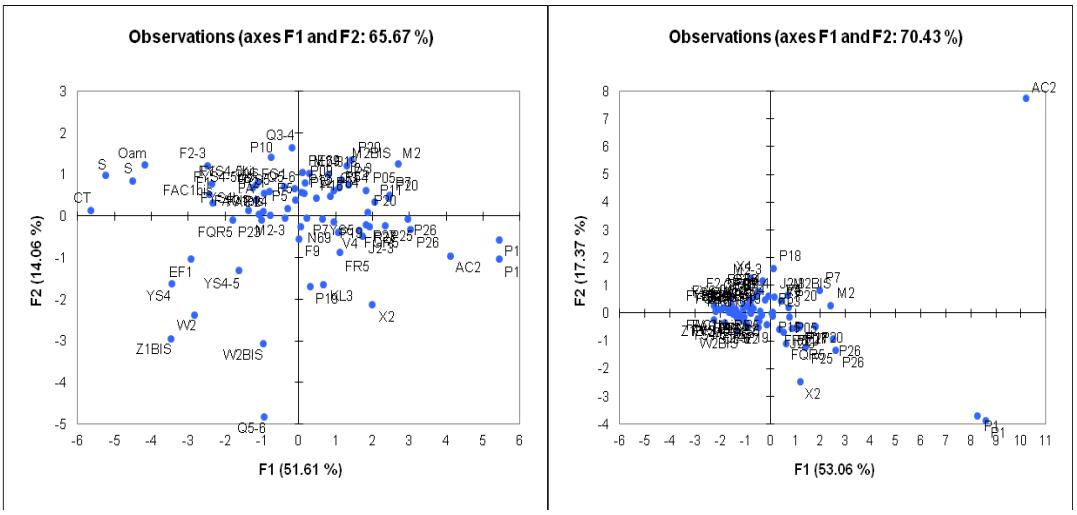


Figure 10.b: distribution of samples in the first factorial plane with a log transformation [Bendekkoum, M., 2025]

3.4.2. *Principal Component Analysis*

Running PCA on the raw dataset already hinted at two distinct processes: TDS correlated broadly with nearly every chemical parameter, while nitrate aligned specifically with sodium and chloride. The log-transformed analysis sharpened this picture considerably — TDS remained tied to almost all major ions, but nitrate now showed no meaningful direct relationship with the rest of the ionic assemblage [Bendekkoum, M., 2025]. That shift is, in our reading, the clearest statistical evidence available in this dataset for nitrate's anthropogenic origin: once the noise from a handful of extreme raw values is removed, nitrate simply does not travel with the geogenic mineralization signal.

Table 5. Correlation matrix of chemical variables computed from the non-transformed dataset [Bendekkoum, M., 2025]

Variables	Ca (mg/l	Mg (mg/l	Na (mg/l)	K (mg/l	SO_4 (mg/l	NO_3 (mg/l	Cl (mg/l)	HCO_3 (mg/l)	TDS (mg/l)
Ca (mg/l)	1	0,646	0,321	0,153	0,664	0,058	0,403	0,222	0,622
Mg (mg/l	0,646	1	0,608	-0,032	0,837	0,111	0,557	0,417	0,823
Na (mg/l)	0,321	0,608	1	-0,049	0,539	0,621	0,943	0,406	0,911
K (mg/l	0,153	-0,032	-0,049	1	-0,092	0,170	0,102	-0,079	0,013
SO_4 (mg/l	0,664	0,837	0,539	-0,092	1	0,001	0,447	0,223	0,773
NO_3 (mg/l	0,058	0,111	0,621	0,170	0,001	1	0,659	0,184	0,471
Cl (mg/l)	0,403	0,557	0,943	0,102	0,447	0,659	1	0,355	0,887
HCO_3 (mg/l	0,222	0,417	0,406	-0,079	0,223	0,184	0,355	1	0,499
TDS (mg/l)	0,622	0,823	0,911	0,013	0,773	0,471	0,887	0,499	1

The PCA results indicate that groundwater chemistry is governed by multiple processes rather than by a single contamination source. The extracted components reflect the combined contribution of mineral dissolution, groundwater-rock interaction, salinity evolution, and surface-derived inputs.

Table 6. Correlation matrix of chemical variables computed from log-transformed data [Bendekkoum, M., 2025]

Variables	Ca (mg/l	Mg (mg/l)	Na (mg/l)	K (mg/l	SO_4 (mg/l)	NO_3 (mg/l)	Cl (mg/l)	HCO_3 (mg/l)	TD (mg/l)
Ca (mg/l	1	0,643	0,423	0,366	0,568	0,186	0,455	0,107	0,635
Mg (mg/l	0,643	1	0,770	0,146	0,818	0,194	0,568	0,162	0,847
Na (mg/l)	0,423	0,770	1	0,078	0,846	0,105	0,767	0,170	0,895
K (mg/l	0,366	0,146	0,078	1	0,123	0,190	0,270	0,046	0,136
SO_4 (mg/l	0,568	0,818	0,846	0,123	1	-0,017	0,675	-0,003	0,833
NO_3 (mg/l	0,186	0,194	0,105	0,190	-0,017	1	0,080	0,147	0,199
Cl (mg/l)	0,455	0,568	0,767	0,270	0,675	0,080	1	0,087	0,711
HCO_3 (mg/l	0,107	0,162	0,170	0,046	-0,003	0,147	0,087	1	0,334
TDS (mg/l)	0,635	0,847	0,895	0,136	0,833	0,199	0,711	0,334	1

Table 7: Correlations between variables and extracted factors (F1-F4).

	F1	F2	F3	F4
Ca (mg/l	0,712	0,262	-0,254	-0,067
Mg (mg/l	0,896	-0,038	0,042	0,123
Na (mg/l)	0,905	-0,218	0,123	0,057
K (mg/l	0,265	0,625	-0,560	-0,384
SO_4 (mg/l	0,897	-0,292	-0,108	0,038
NO_3 (mg/l	0,198	0,714	0,147	0,640
Cl (mg/l)	0,800	-0,083	-0,127	-0,097
HCO_3 (mg/l	0,221	0,395	0,764	-0,455
TDS	0,956	-0,021	0,190	0,004

The PCA performed on log-transformed data provides meaningful insight into the processes governing groundwater chemistry. The correlation circles for the F1/F2 and F1/F3 axes are shown in Figure 11 [Bendekkoum, M., 2025].

The first extracted component groups conductivity, TDS, Na^+, Cl^-, and SO_4^{2-} — a mineralization axis reflecting evaporite dissolution and the progressive chemical evolution of groundwater along its flow path, with the Na^+-

Cl^- pairing pointing specifically to a shared evaporitic source (though irrigation return flow may contribute in shallow sectors) (Chernet et al., 2001).

The relationship between Na^+ and Cl^- suggests a common origin related to evaporitic mineral dissolution. This interpretation aligns with the presence of Triassic evaporitic formations around the basin. However, the contribution of agricultural return flows cannot be excluded, especially in shallow groundwater sectors where irrigation practices influence chemical composition.

The association of sulfate with mineralization parameters reflects the influence of gypsum and anhydrite dissolution. The spatial coincidence between sulfate-rich waters and evaporitic formations supports a dominant geogenic control rather than a purely anthropogenic origin.

The second component isolates nitrate together with variables typical of shallow, surface-connected groundwater — a signature we read as the statistical fingerprint of agricultural and wastewater-driven contamination, operating independently of the basin's natural salinity evolution (Soro et al., 2019; Zerhouni et al., 2018).

This two-component structure is more than a statistical curiosity: it demonstrates that groundwater degradation in the Tébessa aquifer proceeds through two superimposed but distinguishable mechanisms — natural geochemical evolution tied to rock-water interaction, and anthropogenic contamination concentrated in the shallow, vulnerable portions of the system (Ziani et al., 2016).

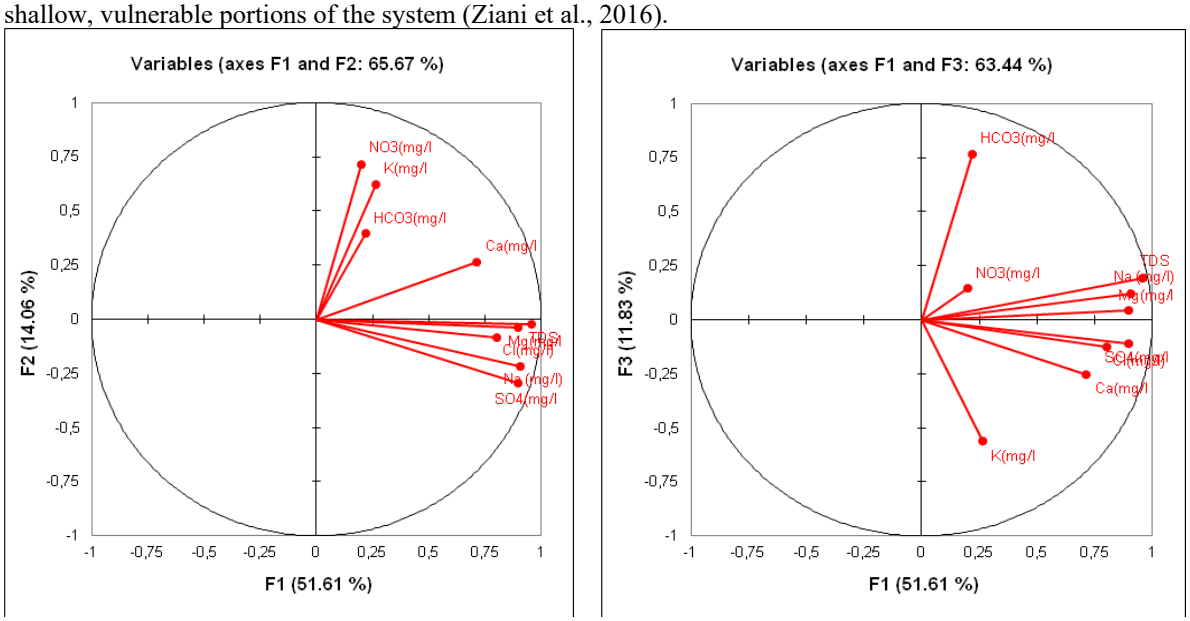


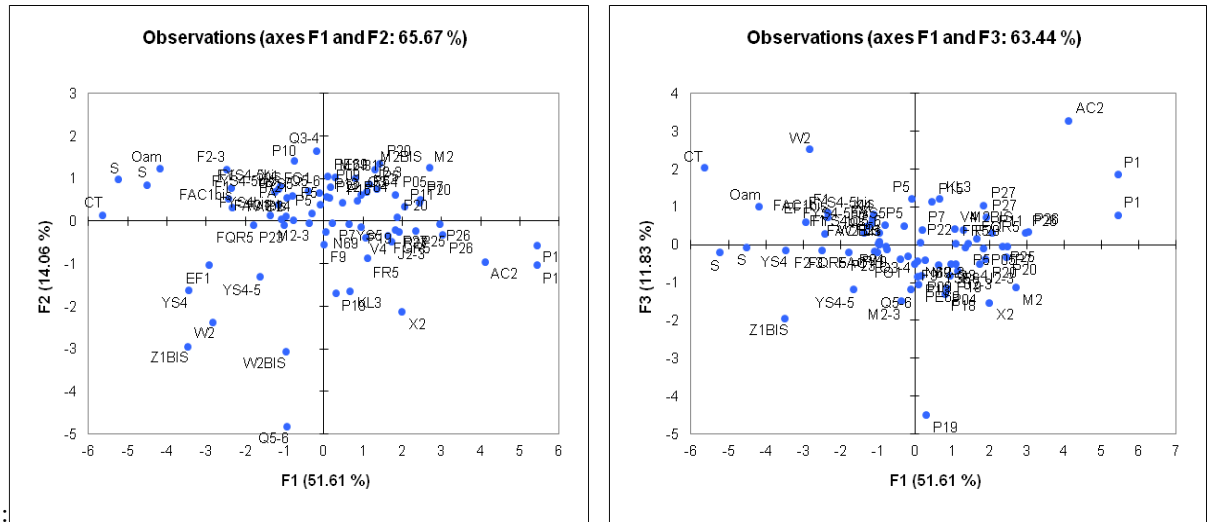
Figure 11. Variable correlation circles plotted along the F1/F2, F1/F3 factorial axes.

3.4.3. Spatial Distribution of Samples and PCA Projection

The distribution graphs of the individuals are clearly visible with the logarithmically transformed data (Figure 12). The most distant water points, which presented extreme values, have become closer to the other points and are now easier to compare with the others [Bendekkoum, M., 2025]

This projection of water samples onto the factorial planes (Figures 12, 13) reveals a clear spatial and chemical differentiation among groundwater points. Samples located along the positive side of F1 correspond to highly mineralized waters, often associated with shallow wells and areas influenced by evaporitic formations and agricultural return flows. In contrast, samples plotting toward the negative side of F1 represent low-mineralized waters, generally derived from deeper boreholes and springs.

The comparison between PCA results obtained from raw and log-transformed data demonstrates that logarithmic transformation significantly improves the structure and interpretability of the data cloud, allowing for a clearer separation of hydrochemical facies and pollution sources.



:

Figure 12. diagram of distribution of water points

The distribution of samples within the factorial space allows the identification of four hydrochemically distinct groups, each representing groundwater influenced by specific geological and hydrogeological conditions (Figure 13).

Group 1 comprises groundwater collected from the Aïn Chabro sector, situated in the western part of the basin (FR5, FKL3, FAC2, FJ2-3 and P1). These samples exhibit comparable hydrochemical characteristics, reflecting similar recharge conditions and groundwater evolution.

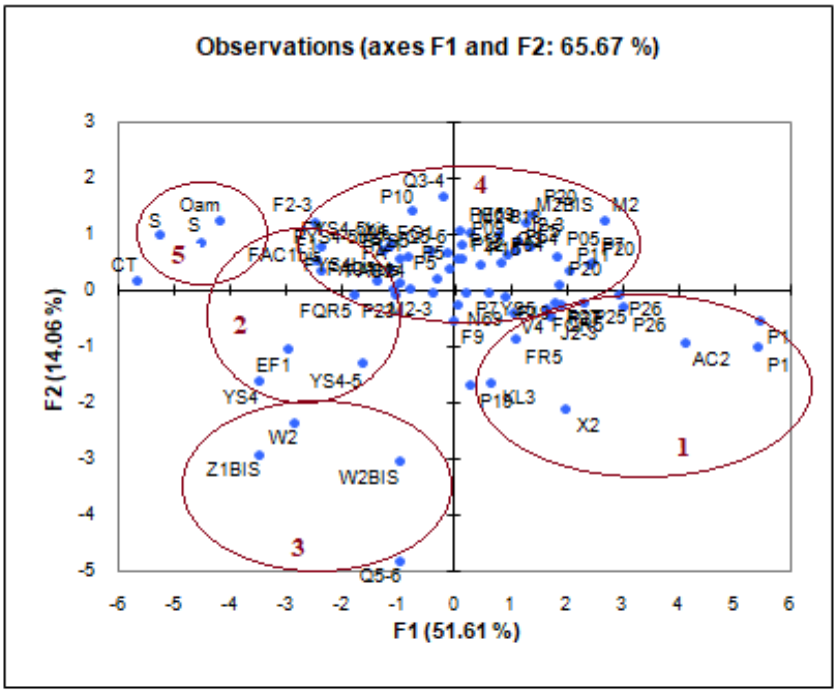
Group 2 includes groundwater from the Bekkaria area in the eastern sector of the basin (FZ1BIS, FW2, FW2BIS and FQ5-6). The chemical similarity observed within this cluster suggests that these sampling points are controlled by common lithological conditions and groundwater flow pathways.

Group 3 groups groundwater collected from the Aïn Chabro–Hammamet zone (FAC1, FQR5, FM2-3, FEG1, FYS4-5, P2 and P5F2-3). Although located within the same alluvial aquifer, these samples represent an intermediate stage of hydrochemical evolution, influenced by both recharge processes and progressive mineralization during groundwater circulation.

Group 4 consists exclusively of surface-water samples, represented by the Youkous Spring and the Youkous River, which remain clearly separated from groundwater because of their distinct hydrological and hydro-chemical characteristics.

Although all groundwater samples originate from the same alluvial aquifer, their distribution within the factorial space demonstrates that hydro-chemical variability is controlled not only by lithological composition but also by groundwater residence time, recharge mechanisms, groundwater depth, and the intensity of human activities. This observation confirms that chemically distinct groundwater bodies may develop within a single

hydrogeological system despite an apparently similar geological framework [Bendekkoum,M,2025]



Legend

Group 1 and 2: Ain Chabro area (western part of the plain);

Group 3: Bekkaria area (eastern part of the plain);

Group 4: Ain chabro- Merdja area located in the centre part of the plain;

Group 5: surface water (Bouakous spring and wadi) western part of the plain.

Figure 13. Groups of water points according to their distribution on the F1/F2 axis[Bendekkoum,M,2025]

3.4.4. Hierarchical Cluster Analysis

Sampling sites were subsequently grouped by chemical resemblance using Hierarchical Cluster Analysis, complemented by a discriminant analysis separating points by type (borehole, well, spring, wadi); this classification is consistent with, and reinforces, the PCA structure described above (Figure 14)

Four hydrochemically distinct groups emerge. One is defined by high salinity and elevated Na^+, Cl^-, SO_4^{2-}, and TDS — a signature of strong geogenic control amplified by evaporation, typical of low-to-moderately mineralized springs and deeper groundwater with shorter exposure to surface contamination. A second groups shallow, nitrate-rich samples where both salinity indicators and nitrate coexist, pointing to sectors where evaporitic dissolution and anthropogenic input act together. A third represents low-mineralized bicarbonate waters typical of active recharge areas. A fourth, where present, captures intermediate conditions arising from mixing between recently recharged and more chemically evolved groundwater — reflecting the dynamic, non-static nature of flow within the alluvial basin (Seghir, Carrillo-Rivera, & Djaafri, 2024).

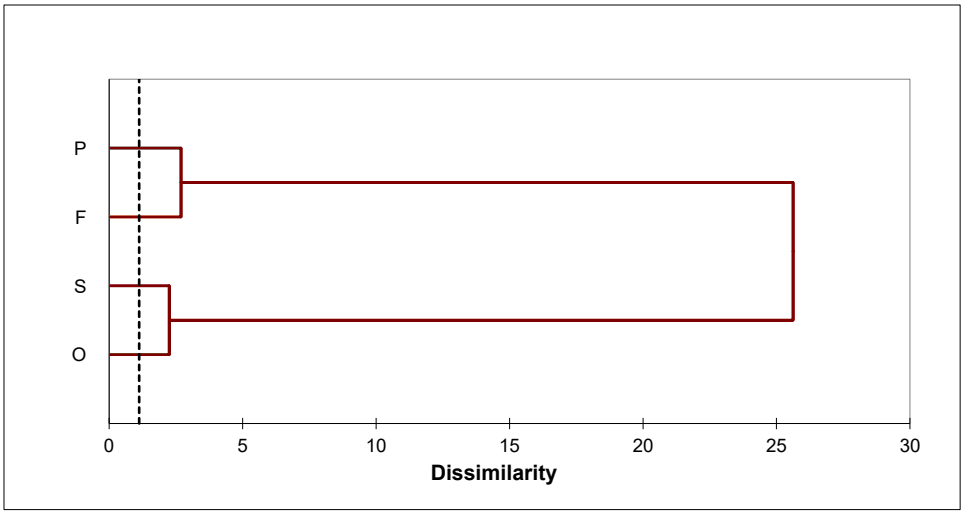


Figure 14. Dendrogram grouping groundwater and surface-water samples by chemical resemblance (P = well; F = borehole; S = spring; O = wadi)

HCA therefore corroborates the spatial heterogeneity already suggested by the hydrochemical and PCA analyses: groundwater quality in the Tébessa aquifer varies systematically with depth, recharge conditions, geological setting, and the intensity of human pressure (Chernet, Travi, & Valles, 2001).

3.5. GIS-based spatial analysis and groundwater vulnerability interpretation

The spatial distribution maps of nitrate and major ions add a further layer of evidence on the aquifer's vulnerability. Nitrate loading is not randomly distributed; it concentrates in sectors combining agricultural activity with shallow groundwater conditions — precisely where contaminant sources and favorable infiltration conditions overlap.

The western and eastern sectors, showing elevated chloride and sulfate, corroborate the geological control exercised by the surrounding evaporitic formations. Importantly, the zones of highest nitrate do not always coincide with the most mineralized waters — confirming that two distinct processes are at work: lithology-driven natural salinization on one hand, and surface-activity-driven nitrate contamination on the other.

Recharge areas adjoining the fractured carbonate uplands play an ambivalent role: they sustain the aquifer through natural infiltration and lateral inflow, yet where they intersect agricultural land, they equally serve as preferential pathways for contaminant transport (Albinet & Margat, 1970; Amharref & Bernoussi, 2007; Amharref et al., 2007). Groundwater vulnerability in this basin therefore results from the combined effect of aquifer permeability, depth to water table, recharge intensity, geology, and land use — which argues for extending protection measures beyond currently contaminated zones to the recharge areas that feed them.

4.Conclusion

This study set out to determine not just *whether* the Tébessa alluvial aquifer is affected by both natural and human processes — that much was already suspected — but *how to separate the two statistically and locate them spatially*. The answer, we argue, lies in treating logarithmic transformation as a deliberate diagnostic step rather than a routine formality: once applied, it reveals that nitrate behaves independently of the basin's geogenic mineralization axis (dominated by carbonate and evaporite dissolution), a separation that raw-data analysis does not achieve with comparable clarity.

Hydrochemically, Ca, Mg, Na,Cl, and SO4 enrichment trace back to water-rock interface and evaporite dissolution, while nitrate follows a distinct pathway rooted in shallow groundwater exposed to agricultural return flow and wastewater infiltration. PCA formalizes this split into two components — one geogenic, one

anthropogenic — and HCA corroborates it by grouping sampling points into hydrochemically coherent clusters tied to depth, recharge conditions, and land use. GIS mapping then locates precisely where this vulnerability peaks: permeable alluvial sectors with shallow water tables lying beneath cultivated land.

We consider the combination of these three elements — validated log-transformation, PCA/HCA-based source discrimination, and GIS-based recharge mapping — to constitute a transferable diagnostic framework, applicable beyond Tébessa to other semi-arid Mediterranean alluvial aquifers facing the same entangled geogenic-anthropogenic pressures (Scanlon et al., 2012; FAO, 2020). Practically, the results argue for extending groundwater protection efforts beyond the currently contaminated zones to include the recharge areas themselves, since these are the pathways through which future contamination will arrive.

Acknowledgments

The authors gratefully acknowledge the editorial team and reviewers of Mediterranean Geoscience Reviews for their constructive comments and efforts in handling this manuscript. Their valuable feedback contributed to improving the quality and clarity of this work.

Funding

The authors declare that this research was carried out entirely without external financial support. No funding, grants, or financial assistance from individuals, institutions, or organizations were received for the design, execution, analysis, or interpretation of the study, nor for the preparation or publication of this manuscript. The work was conducted on the authors' own initiative, and no potential conflicts of interest related to funding sources exist.

Data availability statement

The data utilized in this manuscript were collected specifically for this study through fieldwork and laboratory analysis. This is the first time these data have been employed for publication.Data supporting the findings of this study can be obtained from the corresponding author on justified request.

Declarations

Conflicts of Interest

There is no conflict of interest regarding the publication of this study. All authors have disclosed any financial or personal relationships with other people or organizations that could inappropriately influence or bias the workAll authors have read, understood, and have complied as applicable with the statement on "Ethical responsibilities of Authors" as found in the Instructions for Authors.

Author Contributions

Conceptualization, M.B. and K.S.; methodology, K.S., V.V. and H.A.; validation, K.S., V.V. and H.A.; formal analysis, M.B. and M.A.S.; investigation, M.B. and K.S.; resources, K.S.; data curation, L.B. and V.V.; writing—original draft preparation, M.B. and K.S.; writing—review and editing, K.S.; visualization, K.S.; supervision, K.S., H.A. and V.V.; project administration, K.S. All authors have read and agreed to the published version of the manuscript.

References

Abidi Saad, M., Seghir, K., Touahri, A., Bendekkoum, M., & Bellaoueur, A. (2024). Applying the water quality indices and geographic information system approach to assessing the suitability of groundwater quality for drinking and irrigation purposes in the semi-arid region of Tebessa-Ain Chabro, Northeastern Algeria. Applied Water Science, 14(10), 221.

Aghzar, N., Berdai, H., Bellouti, A., et al. (2002). Pollution nitrique des eaux souterraines au Tadla (Maroc). Revue des sciences de l'eau/Journal of Water Science, 15(2), 459–492.

Albinet, M., & Margat, J. (1970). Cartographie de la vulnérabilité à la pollution des nappes d'eau souterraine. Bulletin BRGM, 2(3), 4.

HADJELA Ali.2023. Water Between Scarcity and Floods, and the Impact of Natural and Human Factors: The Case of the City of Tebessa - Algeria. Int J Innov Stud Sociol Humanities. 2023;8(1): 397-408. DOI: https://doi.org/10.20431/2456-4931.080142

Hadjela A, Annab R, Seghir K.,2021.-Facing floods in the city of Tébessa, in Eastern Algeria, a morphometric and 6. cartographic approach to prioritize intervention in the city's sub-watersheds using GIS and remote sensing. J. Fundam. Appl. Sci, 13(3), 1242-1261. https://www.jfas.info/index.php/JFAS/article/view/1068.

Amharref, M., & Bernoussi, A. (2007). Vulnérabilité et risque de pollution des eaux souterraines. Actes des JSIRAUF, Hanoi, 10.

Amharref, M., Aassine, S., Bernoussi, A., et al. (2007). Cartographie de la vulnérabilité à la pollution des eaux souterraines : Application à la plaine du Gharb (Maroc). Revue des sciences de l'eau/Journal of Water Science, 20(2), 185–199.

Baazi Houria, N., et al. (2020). Impact of agricultural activities on groundwater quality in northeastern Algeria. Water Quality Research Journal, 55(1), 73–84.

Bouchemal, F. (2015). Qualité physico-chimique et paramètres de pollution des eaux souterraines de la région de Biskra. Larhyss Journal, 12(2), 197–212.

Bendekkoum, Mehdi, 2017, Etat de la réserve en eaux souterraines dans la zone de Tébessa Boulhef-Dyr durant la période des hautes eaux du cycle hydrologique 2016/2017,Mémoire de Master, Département des Sciences de la Terre et de l'Univers, Université de Tébessa, 78p.

Mehdi Bendekkoum,Karima Seghir,,Messaoud Abidi Saad, Ali Hadjela,Vincent Valles.2025. Identifying Groundwater Recharge Areas Most Vulnerable to Nitrate Pollution in the Tébessa Region (NE Algeria): A Statistical and Cartographic Study. *Preprint* DOI:10.20944/preprints202512.2281.v1

Carrillo-Rivera, J. J., Cardona, A., & Edmunds, W. M. (2002). Use of extraction regime and knowledge of hydrogeological conditions to control high-fluoride concentrations in abstracted groundwater: Basin of San Luis Potosí, Mexico. Journal of Hydrology, 261, 24–47.

Carrillo-Rivera, J. J., Cardona, A., Huizar-Álvarez, R., & Graniel, E. (2008). Response of the interaction between groundwater and other components of the environment in Mexico. Environmental Geology, 2(476), 303–319.

Chernet, T., Travi, Y., & Valles, V. (2001). Mechanism of degradation of the quality of natural water in the lakes region of the Ethiopian rift valley. Water Resources, 35, 2819–2832.

Collin, J. J. (2004). Les eaux souterraines : Connaissance et gestion. BRGM Éditions, Hermann.

De Marsily, G. (1981). Hydrogéologie quantitative (1st ed.). Masson.

Detay, M. D. (1997). La gestion active des aquifères. Masson.

Djaafri, I. (2025). Contribution of the hydrogeological study of some thermo-mineral springs in the north-east of Algeria [Doctoral dissertation].

Djaafri, I., Seghir K., Valles, V., & Barbiero, L. (2024). Regional hydro-chemistry of hydrothermal springs in northeastern Algeria. Earth, 5(2), 214–227. https://doi.org/10.3390/earth5020011

El Moustaine, R., Chahlaoui, A., El Habib, R. O. U. R., et al. (2013). Étude de la qualité bactériologique et physico-chimique des eaux de certains puits et sources par ACP : Une étude de cas de la région de Meknès (Maroc). Les Technologies de Laboratoire, 8(33).

FAO. (2020). The state of food and agriculture: Overcoming water challenges in agriculture. Food and Agriculture Organization of the United Nations.

Kowalski, W. M., Boudoukha, A., Hemila, M. L., & Pharisat, A. (1997). Les stades d'effondrement du graben de Tébessa. Publications de la Société d'Histoire Naturelle de Montbéliard, 201–215.

Lekrine, Y., Demdoum, A., & Bouaicha, F. (2023). The integration of geochemical and isotopic approaches for thermomineral water characterization: The case of Tebessa (North Eastern Algeria). Acque Sotterranee, 12(2), 77–90.

Messaoud Abidi Saad, Karima Seghir, Abdeldjebar Touahri, Mehdi Bendekkoum, Abdelaziz Bellaoueur (2024). Applying the water quality indices and geographic information system approach to assessing the suitability of groundwater quality for drinking and irrigation purposes in the semi-arid region of Tebessa-Ain Chabro, Northeastern Algeria. Applied Water Science, 14(10), 22. https://doi.org/10.1007/s13201-024-02275-3

Rivett, M. O., et al. (2008). Nitrate attenuation in groundwater: A review of biogeochemical controlling processes. Water Research, 42(16), 4215–4232.

Saad, M. A., Seghir, K., Touahri, A., Bendekkoum, M., Bellaoueur, A., & Pulido-Bosch, A. (2024). Study of the influence of the Tebessa region wastewaters on the hydrochemical quality of the alluvial aquifer groundwater's Tebessa-Ain Chabro (NE Algeria). RBRH, 29, e24.

Scanlon, B. R., et al. (2012). Global groundwater depletion: The facts and the impacts. Nature Climate Change, 2(5), 333–338.

Seghir, K., Carrillo-Rivera, J. J., & Djaafri, I. (2024). Identification of the origin of water salinity in alluvial deposits in the Ain Chabro-Hammamet region, W. Tebessa (North-East Algeria). International Water Forum Conference IWFC 2024, Hammamet, Tunisia.

Seghir, K., (2008c) Vulnérabilité, protection des ressources en eaux, essai de gestion active du sous systèmeaquifère Hammamet Ain chabro (Est Algérien). Thèse de doctorat en hydrogéologie, Université de Annaba, Algérie, 152p.

Seghir, K.; Khérici N. (2008d) contributionto the Study of the Origin of Salinity in Alluvial Aquifer System in the Ain chabro-Hammamet (North-East ALGERIA). American Journal of Environmental Sciences 4 (6): 2008, ISSN 1553-345X,© 2008 Science Publications.DOI:10.3844/ajessp.2008.701.709;http://thescipub.com/issue-ajes/4/6 Am. J. Environ. Sci., 4 (6): 701-709, 2008709

Seghir, K. (2014). La vulnerabilite a la pollution des eaux souterraines de la region Tebessa-Hammamet (Est Algerien). LARHYSS Journal P-ISSN 1112-3680/E-ISSN 2521-9782, (18).

Seghir, K. (2018) Improvement in quantity and quality of groundwater of the Ain chabro-Tebessa alluvial water-table by the application of artificial recharge (East Algeria). 6th Water Forum International Water Forum Conference. Hammamet, TUNISIA, March 19-21, 2018, http://www.isstegb.rnu.tn/IWFC2018/index.html.

Soro, G., Soro, T. D., Adjiri, O., et al. (2019). Application des méthodes statistiques multivariées à l'étude hydrochimique des eaux souterraines de la région des lacs (Côte d'Ivoire). International Journal of Biological and Chemical Sciences, 13(3), 1870–1889.

Souid, F., Agoubi, B., Hamdi, M., et al. (2017). Groundwater chemical and fecal contamination assessment of the Jerba unconfined aquifer, southeast of Tunisia. Arabian Journal of Geosciences, 10(10), 231.

Strebel, O., Duynisveld, W. H. M., & Böttcher, J. (1989). Nitrate pollution of groundwater in Western Europe. Agriculture, Ecosystems & Environment.

Wakida, F. T., & Lerner, D. N. (2005). Non-agricultural sources of groundwater nitrate: A review and case study. Water Research, 39(1), 3–16.

Ward, M. H., et al. (2018). Drinking water nitrate and human health: An updated review. International Journal of Environmental Research and Public Health, 15(7), 1557.

World Health Organization. (2017). Guidelines for drinking-water quality (4th ed.).

World Health Organization. (2023). Guidelines for drinking-water quality (4th ed. + addenda).

Zerhouni, J., Fouzia, R. F., Aboulkacem, A., et al. (2018). Utilisation de l'ACP pour évaluer la pollution des eaux souterraines en milieu rural (ville de Sebaa Ayoune, Maroc). Larhyss Journal, 15(4), 7–20.

577 Ziani, D., Boudoukha, A., Boumazbeur, A., et al. (2016). Investigation of groundwater hydrochemical
578 characteristics using the multivariate statistical analysis in Ain Djacer area, Eastern Algeria. Desalination and
579 Water Treatment, 57(56), 26993–27002.