DOI 10.7343/as-2025-804

ACQUE SOTTERRANEE
Italian Journal of Groundwater
journal homepage: https://www.acquesotterranee.net/

Special Issue:
Hydrogeology in Tunisia

Reassessment of Plio-Quaternary aquifer mineralization (Sidi Mansour plain, Southern Tunisia): a machine learning approach

Rivalutazione della mineralizzazione di un acquifero Plio-Quaternario (Sidi Mansour plain, Southern Tunisia): un approccio Machine Learning

Zohra KRAIEM[a], Kamel ZOUARI[a], Aissa HLEIMI[b], Houda DERBEL FETOUI[c]

[a] Radio-Analysis and Environment Laboratory, the National Engineering School, University of Sfax, Sfax, Tunisia
e-mail : *zohra.kraiem@enis.tn*

[b] Water resources directory, Ministry of Agriculture, Water Resources and Fishing, Tunisia

[c] Faculty of Economic Sciences and Management of Nabeul, university of Carthage, MODILS laboratory

ARTICLE INFO

Ricevuto/*Received*: 8 July 2024
Accettato/*Accepted*: 27 January 2025
Pubblicato online/*Published online*: 30 March 2025

Handling Editor:
Rudy Rossetto

Citation:
Kraiem, Z., Zouari, K., Hleimi, A., Derbel Fetoui, H. (2025). Reassessment of Plio-Quaternary aquifer mineralization (Sidi Mansour plain, Southern Tunisia): a machine learning approach
Acque Sotterranee - *Italian Journal of Groundwater*, 14(1), 09 - 24
https://doi.org/10.7343/as-2025-804

Correspondence to:
Zohra Kraiem :
zohra.kraiem@enis.tn

Keywords: : *major elements, isotopes, mixing process, upward leakage, machine learning; salinity prediction.*

Parole chiave: elementi maggiori, isotopi, miscelazione, filtrazione verso l'alto, machine learning, previsione di salinità.

Riassunto

La comprensione della mineralizzazione delle falde acquifere è fondamentale per una gestione sostenibile delle risorse idriche. L'acquifero plio-quaternario della pianura di Sidi Mansour è stato perciò studiato in dettaglio, per una rivalutazione della eterogenea mineralizzazione delle sue acque. L'analisi geochimica e le analisi degli isotopi stabili (^{18}O, ^{2}H) e radiogenici (^{3}H, ^{14}C) hanno fornito una migliore comprensione dell'idrodinamica e dei processi di mineralizzazione alla base della composizione chimica delle acque. È stata utilizzata una rete neurale artificiale per una valutazione integrata, attraverso la previsione della salinità totale disciolta (TDS) delle acque sotterranee nell'area di Sidi Mansour. A ovest, l'acquifero plio-quaternario ha una facies Na-Ca-SO_4 con elevata salinità a causa della dissoluzione dei minerali, del mescolamento e della filtrazione verso l'alto delle acque più profonde attraverso la faglia di Om Ali (contributo fino al 96%). Ad est, le acque prevalentemente Ca-SO_4 sono meno mineralizzate a causa dell'infiltrazione di precipitazioni meteoriche. La previsione della salinità utilizzando l'apprendimento automatico ha indicato che il modello proposto ha raggiunto un'elevata efficienza. I risultati del modello di rete neurale di back-propagation (10:4:1) hanno indicato un'elevata precisione dell'algoritmo addestrato, come confermato da un test di validazione incrociata (elevata precisione (88,89%), specificità (100%) e R^2 (0,9687)). Questo studio evidenzia l'importanza di questo modello nel prevedere la salinità per la falda acquifera plio-quaternaria in questa regione.

Abstract

Understanding aquifer mineralization is crucial for sustainable water resource management. Thus, the Plio-Quaternary aquifer of the Sidi Mansour plain was studied in detail for reassessment of its heterogeneous mineralization water. Geochemical analysis and stable (^{18}O, ^{2}H) and radiogenic isotope (^{3}H, ^{14}C) analyses have provided a better understanding of the hydrodynamics and mineralization processes underlying the chemical composition. An artificial neural network was used for an integrated groundwater assessment in the Sidi Mansour area via the prediction of total dissolved solids (TDS). To the west, the Plio-Quaternary aquifer has a Na-Ca-SO_4 facies with high salinity because of mineral dissolution, mixing, and upward percolation of deeper waters through the Om Ali fault (contribution up to 96%). To the east, predominantly Ca-SO_4 waters are less mineralized due to the infiltration of recent precipitation. Salinity prediction using machine learning indicated that the proposed model achieved high efficiency. The backpropagation neural network model results (10:4:1) indicated high accuracy of the trained algorithm, as confirmed by a cross-validation test (high accuracy (88.89%), specificity (100%), and R^2 (0.9687)). This study highlights the importance of this model in predicting salinity for the Plio-Quaternary aquifer in this region.

Introduction

Salinization is a real threat to groundwater resources in arid and semiarid areas, leading to limited access to water. It can be caused by several processes, including seawater intrusion (Grassi et al., 2007; Kouzana et al., 2009), upwelling of saltwater from underlying aquifers through fractures (Abid et al., 2010; Gonfiantini et al., 1974) and geogenic dissolution of saltwater deposits or agricultural inputs. Other possible processes may include salinization through the redistribution or concentration of solutes already present in the system, such as the return of irrigation water (Kattan, 2008; Qin et al., 2011), increased exploitation of aquifers (Boughriba et al., 2006, Krimissa et al., 2004;Trabelsi et al., 2005; Zammouri et al., 2007), evaporation (Kattan, 2008) and mixing between different types of water (Ramos-Leal et al., 2007; Rapti-Caputo & Martinelli, 2009). In this context, several researchers have focused on the suitability of groundwater for drinking and irrigation purposes. For example, Ben Brahim et al. (2021) evaluated the groundwater quality for both irrigation and drinking water by using GIS tools. They concluded that the fragile environment of the Kebili area is prone to anthropogenic effects. Mastrocicco (2021) provided a comprehensive review of water resource salinization over 30 years. The author insisted on the effects of climate change, increasing demand for agriculture and demographic growth on the depletion of both surface water and groundwater. The author indicated that the threat to the Italian coast by salinization can be plausibly attributed to the use of "unsuitable waters" and the mobility of trace elements, among other factors. Taherian and Joodavi (2021) concluded that interaction with host rock is the main process of increased salinity in the alluvial aquifer in central Iran. In addition, Dib et al. (2023) used an integrated hydrochemical and statistical approach to study shallow aquifers of the Mio-Plio-Quaternary formations. They proposed the dissolution of evaporite rocks as a possible process behind the salinization of the waters from the eastern high plains of Algeria. In the same context, we intend to study the shallow aquifers in the area, which is situated southwest of Tunisia and is part of the well-known Saharan Atlas domain. It is subjected to cold and marine influences from the Mediterranean to the east and warm air masses from the south. It is characterized by an average annual temperature of 22 °C and evapotranspiration of 1100 mm/year with irregular and low rainfall (130–150 mm/year), reflecting an arid climate. This region is a basin characterized by a synclinal structure, resulting in a flat surface with sandy-clay fillings stretching from east to west; it forms the main body of the Mio-Plio-Quaternary aquifer. These formations exhibit a notable increase in thickness from the basin's edges toward the center with a multilayered aquifer system. This includes the surface aquifer placed in the Plio-Quaternary formations and the deep aquifers belonging to the Miocene, the Upper Cretaceous aquifer (Cenomanian, Turonian and Upper Senonian) and the Lower Cretaceous or Intercalary Continental deposits.

In this area, deep groundwater is poor in quality. Thus, particular interest is given to the Plio-Quaternary aquifer, which is considered the only source of water for livestock and agriculture. These horizons are fed by the infiltration of surface water from rivers (called "wadis") flowing from bordering mountainous areas.

For a deeper understanding of the properties of these resources, a multidisciplinary approach was adopted in this study via the application of hydochemical and isotopic tools. Additionally, we used machine learning to predict salinity. A back propagation neural network (BPNN) was used as a tool to accurately predict the salinity of the shallow waters of the Sidi Mansour plain. A BPNN typically uses the backpropagation algorithm, a supervised learning technique, to train the model. It is a valuable tool for predicting TDS concentrations because of its ability to capture complex relationships and learn from data (Khullar and Singh, 2021; Rashid and Kumari, 2023). This approach was applied to the current dataset because of its simplicity and effectiveness in modeling and predicting various aspects of water (e.g., groundwater level, nitrate concentrations, quality, and isotopic composition as per Lallahem et al., 2005; Zaqoot et al., 2018; Che Nordin et al., 2021; Cemek et al., 2022).

On the other hand, stable isotopes of oxygen and hydrogen are considered good tracers for determining the origins of water and recharge sources and understanding the mechanisms of mixing between different waters (Fontes, 1976; Fritz et al., 1979; Leontiadis et al., 1988; Subyani, 2004).

Characterizing Plio-Quaternary waters requires adequate knowledge of the spatial distribution of these waters and an understanding of the processes that govern their mineralization. As a result, a multidisciplinary study using hydrogeological, geochemical, and isotopic methods and machine learning tools is proposed for the characterization of those groundwater bodies for an accurate assessment of the main mineralization processes from a larger perspective of the sustainable management of these resources. Thus, a comprehensive analysis of those Plio-Quaternary waters was conducted, incorporating multiple aspects (i.e., BPNN, statistics and isotopes).

Materials and methods

Geological, tectonic, and hydrogeological overview

Site description

The Sidi Mansour Basin is located southwestern Tunisia. It stretches 80 km from east to west and 30 to 100 km from north to south. It is defined as a flat plain pinched between Jebel Berda to the north and Sif Laham to the south and opens to the east into a broad alluvial plain. It is bounded to the south by the mountain range beginning in the southeast by Jebel Zemlet el Beidha and ending in the southwest by Jebel Sif Laham, to the northwest by Jebel Berda, to the north by Jebel Chemsi and Jebel Belkheir, to the northeast by the Sebkha Noual and to the east by the coastal plain of Djeffara (Fig. 1).

Geological and tectonic context

Information from geological maps, together with cross-sections of existing boreholes, revealed the presence of geological formations attributed to the Early Cretaceous, Late Cretaceous (i.e., Cenomanian, Turonian, Early Senonian, and Late Senonian), Miocene and Plio-Quaternary (Ben Hadj Ali et al., 1985).

From a lithological perspective, the Early Cretaceous refers to the so-called "continental intercalaire" (CI), geological formations saturated with water separated by thick (10 to 20 m) impermeable layers. It is intercalated between the Jurassic formations at the bottom and the Cenomanian marine formations at the top. The CI-bearing aquifers are detrital formations of clay-sandstone facies that form a homogeneous aquifer (Sokono et al., 2008). The Cenomanian is characterized by a sequence with two limestone and dolomitic limestone members, a marly limestone to marly member in the middle. This set constitutes the Zebbag formation, which continues until the Turonian system. In the Sidi Mansour Basin, the Cenomanian has a variable thickness on both sides of the Gafsa-Elhamma fault. It varies from 300 m at Jebel Sif Laham (to the west of the fault) to 900 m at the northern edge of Jebel Berda. Similarly, the thickness varies between 200 m (Jebel El Haira) and 400 m (Sidi Mansour borehole) east of the fault (Man-1).

The Turonian is characterized by dolomitic limestone facies. Near the center, the Garâat Fatnassa borehole spans 200 m of predominantly calcareous-dolomitic formations with some marl intercalations. The whole area is covered by the early Senonian formation, which is mainly gypseous to marly facies. The late Senonian deposits are mainly carbonates representing the famous Abiod formation. The average thickness reaches 400 m to the west but does not exceed 200 m at the eastern edge. Mio-Plio-Quaternary fillings of the syncline formed continental deposits, consisting mainly of clayey sands, clays, marls, and gypsum. These formations are grouped into a single entity known as the Continental Terminal, which includes both the Beglia and Segui formations (Edmunds et al., 2003). The description of the borehole sections allowed the following subdivisions:

- **Lower series** (Pontian sand to sandstone), less extensive under the Segui; this series is related to the Upper Pontian or Miocene. This deposit is thicker in the western part of Segui (44 m) than in the eastern part (approximately 16 m);
- **Middle series** (Pliocene), a clayey or clayey-sandy bed with coarse intercalations, has a lenticular structure; its thickness and composition change with its position within the geological structure.
- **Upper series** (Quaternary), a clayey deposit that thickens near the Garâats and reaches 40 to 60 m, largely exceeding the average thickness of 10–20 m in the Segui plain.

Structurally, the Sidi Mansour region belongs to the orogenic province of the Saharan Atlas. It is bordered, to the

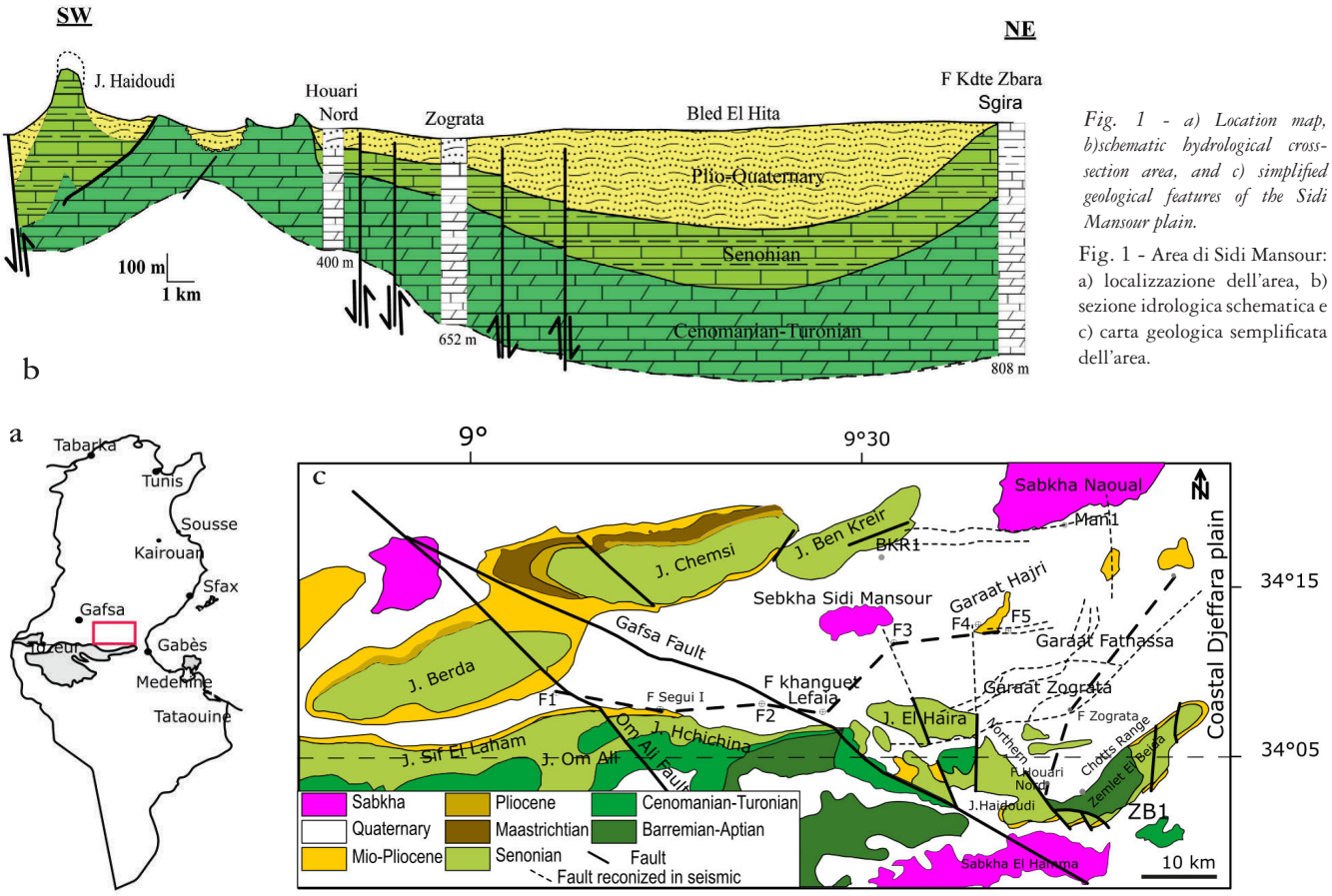

Fig. 1 - a) Location map, b)schematic hydrological cross-section area, and c) simplified geological features of the Sidi Mansour plain.

Fig. 1 - Area di Sidi Mansour: a) localizzazione dell'area, b) sezione idrologica schematica e c) carta geologica semplificata dell'area.

south, by the east-west anticlines of the Chareb range. To the north, the Sidi Mansour plain is bordered by the isolated anticlines of Jebel Berda and Jebel Chemsi. The folding was reinforced by the tectonic effect of the Gafsa-El Hamma fault, whose thickness reaches 1000 m in some places (i.e., Berda, Chemsi, and Orbata); it brought the Early Cretaceous into contact with the Late Cretaceous. The Gafsa-El Hamma fault is the most active structure in the South Atlas, Tunisia, with a 75 km dextral oblique ramp within Triassic evaporites (Ahmadi et al., 2006; Ferjani et al., 1990).

To the west of the study area, the contact of the Hchichina anticline with the Om Ali anticline is marked by the kilometeric Om Ali fault, indicating an apparent dextral strike-slip motion (Bouaziz et al., 2002). This fault brought the Late Cretaceous series into contact with those of the Mio-Pliocene (Bouaziz et al., 2002; Castany, 1954).

Hydrogeological overview

Hydrogeological surveys undertaken in this area revealed several superimposed aquifer horizons. The deep aquifer system is formed by the Continental Intercalaire as a package of superimposed Late Cretaceous aquifers (i.e., Cenomanian, Turonian, and Senonian) and the Miocene or Upper Pontian aquifer. The Turonian-Cenomanian reservoir is lodged in the Zebbag formation and has high salinity (> 17 g/L). The aquifer is characterized by very heterogeneous transmissivity values ranging from 20 × 10^{-3} to 300 × 10^{-3} m^2/s. These values appear to be very high concomitantly with high permeability, probably due to the fissured limestone. Above this formation, the Late Senonian limestones (i.e., Abiod formation) presented an average salinity of 6 g/L. The Miocene aquifer formed the lower part of the Continental Intercalaire as a sandy to coarse-sand aquifer. It is covered by the medium member of the Continental Intercalaire. The well known sandy to clayey layers of the Plio-Quaternary encompass the studied water table. However, these levels presented lenticular passages of gypsum that could be behind the increased salinity in places. Nevertheless, hydrogeological continuity may exist between the Cretaceous and Continental Intercalaire aquifers, particularly near major faults.

Aquifer recharge in Tunisia is severely limited because of the country's arid and semiarid climate. While the northwestern regions receive high amounts of rainfall, averaging 700–1000 mm/year, the southern and central regions, including the Gabès area, experience significantly lower rainfall, often less than 300 mm/year. This arid climate results in minimal infiltration and subsequent groundwater recharge. As a result, the contribution of rainfall to aquifer replenishment is negligible in most parts of the country, especially in the southern arid land. The mean yearly amount of rainfall can be estimated to be between 0 and 300 mm/year (Maroua et al., 2015). According to Jemai et al. (2022), the highest median precipitation reached 191 mm/year in the last four decades. Therefore, the climate in the area is usually considered very arid (Soro et al., 2014). Thus, no significant aquifer recharge from precipitation can be assumed. Consistent with the climatic setting of the area, we did not consider any atmospheric boundary conditions; hence, we did not consider any recharge from rainfall.

The piezometric map established for the Plio-Quaternary aquifer revealed a general flow direction of groundwater from higher reliefs (Jebel Sif Laham, Jebel Hchichina, Jebel El Beida, and Jebel Chemsi) to the Sabkhet Sidi Mansour in the western part. On the other hand, at the eastern edges, the flow took place from Garaat Hajri to Sabkhet Sidi Mansour and Garaat Zograta (Fig. 2a). We can conclude that recharge zones occur within the elevated relief features bordering the region to the west and in the vicinity of the Garâat Hajri area to the east, whereas Sebkhet Sidi Mansour plays the outlet role of the Plio-Quaternary aquifer.

Sampling and analysis

A sampling campaign was carried out in the Sidi Mansour plain, from which 33 water samples were collected. A set of 22 samples came from the Plio-Quaternary aquifer, 7 samples from the Miocene aquifer (5 boreholes), 2 wells located at the Miocene outcrops (P7 and P8), and 4 samples from the Senonian aquifer (3 boreholes and 1 well located at the Senonian outcrop (P5); Fig. 2a). These samples were subjected to chemical and isotopic analyses, the results of which are

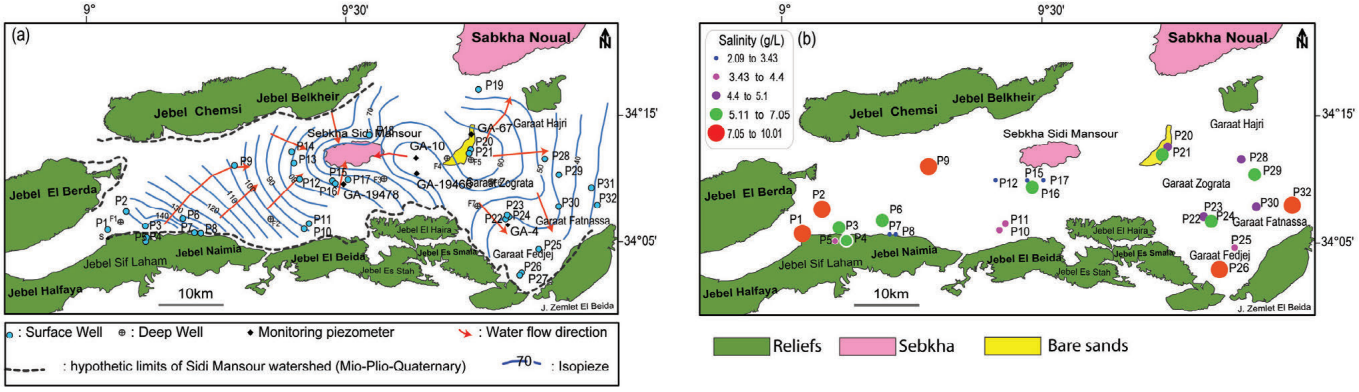

Fig. 2 - Piezometric map (a) and salinity distribution of the PQ water samples from Sidi Mansour area (b).

Fig. 2 - Carta piezometrica (a) e distribuzione della salinità dei campioni d'acqua dell'area (b).

discussed here. The physicochemical parameters (i.e., pH, electrical conductivity, and temperature) were evaluated via an HI 9828 multiparameter portable device (Hanna Instruments, France). The samples for chemical analysis were filtered and stored in plastic polyethylene bottles. Groundwater samples for stable isotope analysis were collected in 30 mL polyethylene bottles with sealing caps. A subset of 15 samples was collected and stored in 500 mL glass bottles for tritium content determination. Both chemical (major elements) and isotopic (^{18}O, ^{2}H, ^{3}H, and ^{14}C activities) analyses were carried out at the Laboratory of Radio-Analysis and Environment, National Engineering School of Sfax, Tunisia. Similarly, major elements (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, Cl^{-} and SO_4^{2-}) were analyzed by liquid chromatography using a chromatograph equipped with IC-PakTM CM/D columns for cations, with EDTA and nitric acid as the eluent, and a Metrohm chromatograph equipped with CI SUPER-SEP columns for anions. The overall detection limit for ions was 0.04 mg/L. The total dissolved solid (TDS) content was measured by a gravimetric method through weighing the residual solids after evaporation of a given amount of the studied water samples. We used a RADWAG type ASR2 Plus precision balance (RADWAG Balances and Scales, La Balance, Tunisia). The total alkalinity (HCO_3) was determined by titration with 0.01 or 0.1 HCl against methyl orange and bromocresol green as indicators. Stable isotopes (^{18}O, ^{2}H) were analyzed by a laser spectrometer (Penna et al. 2010). Carbon activities (^{14}C activities) were measured by liquid scintillation on field-prepared barium carbonate precipitates. The results are expressed as percent modern carbon (PCM) with an analytical uncertainty of 0.3 PCM. Tritium (^{3}H) analyses were performed by electrolyte enrichment followed by liquid scintillation counting (Taylor, 1976).

The contribution of deep aquifers to the recharge of the Plio-Quaternary aquifer can be quantified by calculating an isotope balance in ^{18}O and ^{2}H, which is based on the following equation:

$$\delta 3 = X\delta 1 + (1-X)\,\delta 2 \qquad (1)$$

where X represents the fraction of water from the deep aquifers involved in the recharge of the Plio-Quaternary aquifer, and δ1, δ2, and δ3 are the isotopic contents of ^{18}O and ^{2}H in the deep waters (old pole), the recharge waters (late pole) and the mixture, respectively.

Descriptive statistics and multivariate analysis

Descriptive statistics play a crucial role in summarizing and organizing the characteristics of the collected dataset. After shaping the tabulated data, the first common step involves the systematic running of statistical analysis to describe several features, including the mean, median, minimum, maximum, range, skewness, and kurtosis. The SD and SE were also calculated accordingly. These statistics provide insights into the averages of the values. They help in understanding the central point around which the data tend to cluster. Measures of variability can describe how spread out the values are. Key measures include the range, variance, and standard deviation (Pritha Bhandari, 2019). They reveal the extent of variation within the data, as described earlier. Descriptive statistics served as a basic step for further in-depth and advanced analyses (Tables 1, 2).

Multivariate analyses, such as principal component analysis (PCA) and hierarchical cluster analysis (HCA), are indispensable tools in the realm of statistics. They are two workhorses in data analysis, both of which are unsupervised machine learning methods. While they have different goals, they can be complementary. PCA simplifies complex datasets by identifying a smaller set of features, the principal components, that capture most of the data variation. This approach is useful for visualization and dimensionality reduction (Pham et al., 2023; Yaseen, 2021). HCA, on the other hand, focuses on grouping data points into clusters on the basis of their similarities. It helps uncover natural groupings within the data, which can be used for classification or further exploration. PCA is, therefore, a powerful method for dimensionality reduction and feature extraction.

A scree plot was used to determine the number of significant factors retained in the current multivariate analysis. The biplot offers a combined visualization of both samples and variables resulting from a multivariate analysis technique, such as principal component analysis.

A dendrogram is a tree-like structure, where the bottom section shows individual data points. As branches merge upward, they indicate points that share increasing similarity. The height at which branches join signifies the degree of difference between the merged clusters. In this case, with five clusters, the dendrogram has five main branches, each representing a distinct cluster. The closer the base of a branch is to the bottom, the more similar the data points within that cluster are to each other. Conversely, the higher the point of merging for two branches is, the greater the dissimilarity between the corresponding clusters. This visualization helps researchers understand the natural groupings within their data and identify the optimal number of clusters for further analysis. With a Phenon line of 6.2, we can easily recognize five main groups with close similarity. The distance between two clusters in the dendrogram represents the dissimilarity between those clusters.

ANN as a machine learning tool for TDS prediction

Hydrochemical and isotopic analysis is a valuable tool for understanding the origin and flow paths of groundwater. However, traditional methods can be time-consuming and expensive (Soki et al. 1992; Khalil et al., 2005; Azad et al., 2019; Kouadri et al., 2021; Rashid and Kumari, 2023; Smida et al., 2023; Elshaarawy and Hamed, 2024). Backpropagation neural networks are promising alternatives (Kraiem et al., 2024a,b). By training the network on a dataset of groundwater samples with known isotopic compositions and corresponding environmental factors (such as rainfall data or geological formations), the network can learn to identify isotopic signatures from new, unseen samples (Bhagat et al., 2020; Georgescu et al., 2023). This allows faster and more cost-effective characterization of groundwater sources

and movement, aiding in several tasks, including aquifer management, as was the case in the current study.

A neural network architecture usually describes a system with a relatively simple hidden layer for tasks that do not require extreme complexity. It takes 10 input features, which could represent various data, such as the geochemical properties of the studied water samples (i.e., temperature, pH, HCO_3, Cl, Na, K, Ca, Mg, and EC). These were then processed by a single hidden layer containing 4 nodes. Those nodes act as intermediate processors, combining the weighted influences of the inputs through an activation function. With only one output node, the network produces a single final output (here, we consider the TDS content (mg/L)). This configuration is suitable for problems where the goal is to predict a specific outcome on the basis of the provided input data (Che Nordin et al., 2021; Çimen et al., 2012; Keskin et al., 2015; Kraiem et al., 2024a). The limited hidden layer size makes it efficient for computation while still allowing the network to learn basic relationships between the input features and the desired output.

An ANN was selected for this study because of its ability to capture the nonlinearity and related complexity of the studied groundwater system. Moreover, ANNs have been successfully applied in numerous groundwater studies and have demonstrated superior performance compared with other machine learning methods in many cases.

Results and discussion

Physicochemical parameters

A summary of the saturation indices and stable isotope analyses is reported in Table 1. Salinity (i.e., total dissolved solids) varied from 2.09 to 10.01 g/L for the Plio-Quaternary aquifer. The spatial distribution of salinity indicated significant variation from west to east (Fig. 2b), probably due to the combination of several factors, such as the lithology of the host formation and/or the presence of local sources of contamination. This was not the case in the western part, where tectonic activity (such as the Gafsa or Om Ali faults) might have played a role in the salinization of Plio-Quaternary waters, probably because deep, saltier water seeps through fractures.

In the eastern part, the salinities are lower. These low salinities can be explained by the dilution of rainwater following infiltration at the "garâats" (Hajri, Zograta, and Fatnassa).

Deep water samples from Miocene and Senonian aquifers evolved from a Na-Cl-SO_4 facies (upstream) to a Na-Ca-SO_4 facies (downstream). The Plio-Quaternary aquifer clearly evolves between Ca-Na-SO_4- and Ca-SO_4-dominated waters (Fig. 3). This variability in facies from upstream toward the outlet of the study area further highlights the complex hydrogeochemical processes that can directly affect the salinity of the studied waters, such as the mixing effect, leaching of soluble elements (halite, gypsum, anhydrite, and sodium sulfate or thenardite) and evaporation. The same processes were proposed by Dib et al. (2021), who studied salinization in the Mio-Plio-Quaternary shallow aquifer of eastern Algeria, confirming the current findings and interpretations. Moreover, to elucidate the mechanisms of mineralization in the Sidi Mansour waters, the interrelationships among major elements (i.e., Na^+, Cl^-, Ca^{2+}, and SO_4^{2-}) were investigated.

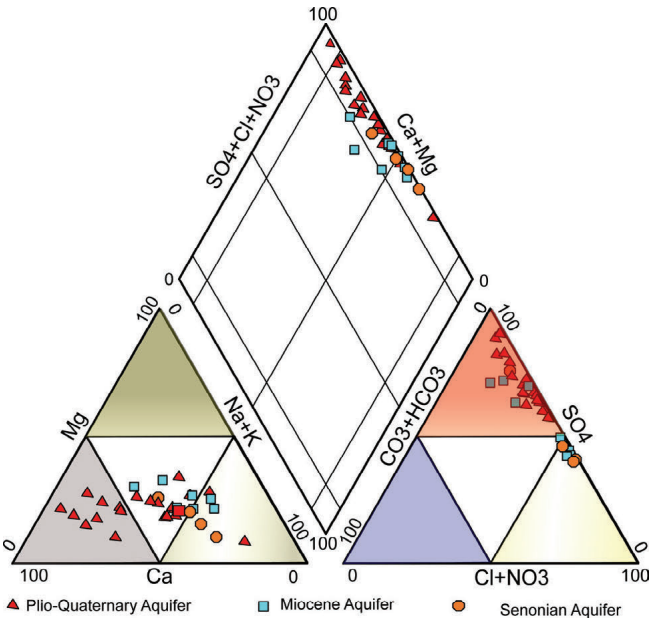

Fig. 3 - Piper diagram of the studied water samples.

Fig. 3 - Diagramma di Piper dei campioni d'acqua studiati.

The variation in Na^+ versus Cl^- had a high coefficient of determination (R^2=0.92), which was probably due to the dissolution of halite (Fig. 4b). This was further confirmed by the relevant halite saturation index (Table 1). However, there is a group of points with an excess of sodium compared to chlorides. This excess can be attributed to the dissolution of other minerals, such as thenardite, whose deposition is generally favored in arid regions and saltwater lakes. Furthermore, SO_4^{2-} was well correlated with Na^+ (R^2=0.75) (Fig. 4c), which revealed an excess of sulfate with respect to sodium, indicating a second origin for sulfate, probably the dissolution of thenardite concomitantly with gypsum. The excess sodium over chlorides can also be explained by base exchange (Ca^{2+}/Na^+) within clay minerals. In fact, on the correlation diagram of Ca^{2+} versus SO_4^{2-}, the majority of the analytical points are positioned below the gypsum dissolution line, reflecting a Ca^{2+} deficit that can be attributed to the precipitation of calcite and/or aragonite (Fig. 4a). The cation exchange capacity of existing clay minerals can also explain such an exchange process.

Fig. 4d shows a scatterplot of Ca^{2+} to Na^+ with slightly different correlations. Water samples from deep aquifers representing the Miocene and Senonian systems (P7, P8, P12, and F2–F7) were more strongly correlated, perhaps because of their cation exchange capacity. In contrast, Plio-Quaternary samples (i.e., P4, P5, P6, P10–12, P15–PP23 and P28–P29) exhibited excessive calcium over sodium; this could

Tab. 1 - Summary of isotopic data and saturation indices of the Sidi Mansour water samples.

Tab. 1 - Sintesi di dati isotopici e indici di saturazionenei campioni d'acqua di Sidi Mansour.

Var.	mean	SD*	median	trimmed	mad	min	max	range	skew	kurtosis	SE**
^{18}O	-5.49	1.63	-5.89	-5.63	1.45	-7.87	-1.31	6.56	0.73	-0.21	0.28
^{2}H	-34.54	10.29	-35.61	-35.35	11.56	-47.84	-11.5	36.34	0.53	-0.82	1.77
SIAnhydrite	-0.29	0.19	-0.24	-0.27	0.1	-0.79	0.18	0.97	-0.69	1.25	0.03
SIAragonite.	0.24	0.41	0.17	0.21	0.41	-0.47	1.38	1.85	0.81	0.71	0.07
SICalcite.	0.39	0.41	0.31	0.36	0.39	-0.3	1.53	1.83	0.83	0.74	0.07
SIDolomite.	0.65	0.93	0.52	0.57	0.78	-0.8	3.2	4	0.94	0.83	0.16
SIGypum.	-0.08	0.17	0	-0.05	0.07	-0.6	0.12	0.72	-1.33	1.13	0.03
SIHalite.	-5.13	0.78	-5.04	-5.07	0.68	-7.05	-3.74	3.31	-0.61	-0.37	0.13
SD*: standard deviation; mad: mean absolute deviation; SE**: standard error											

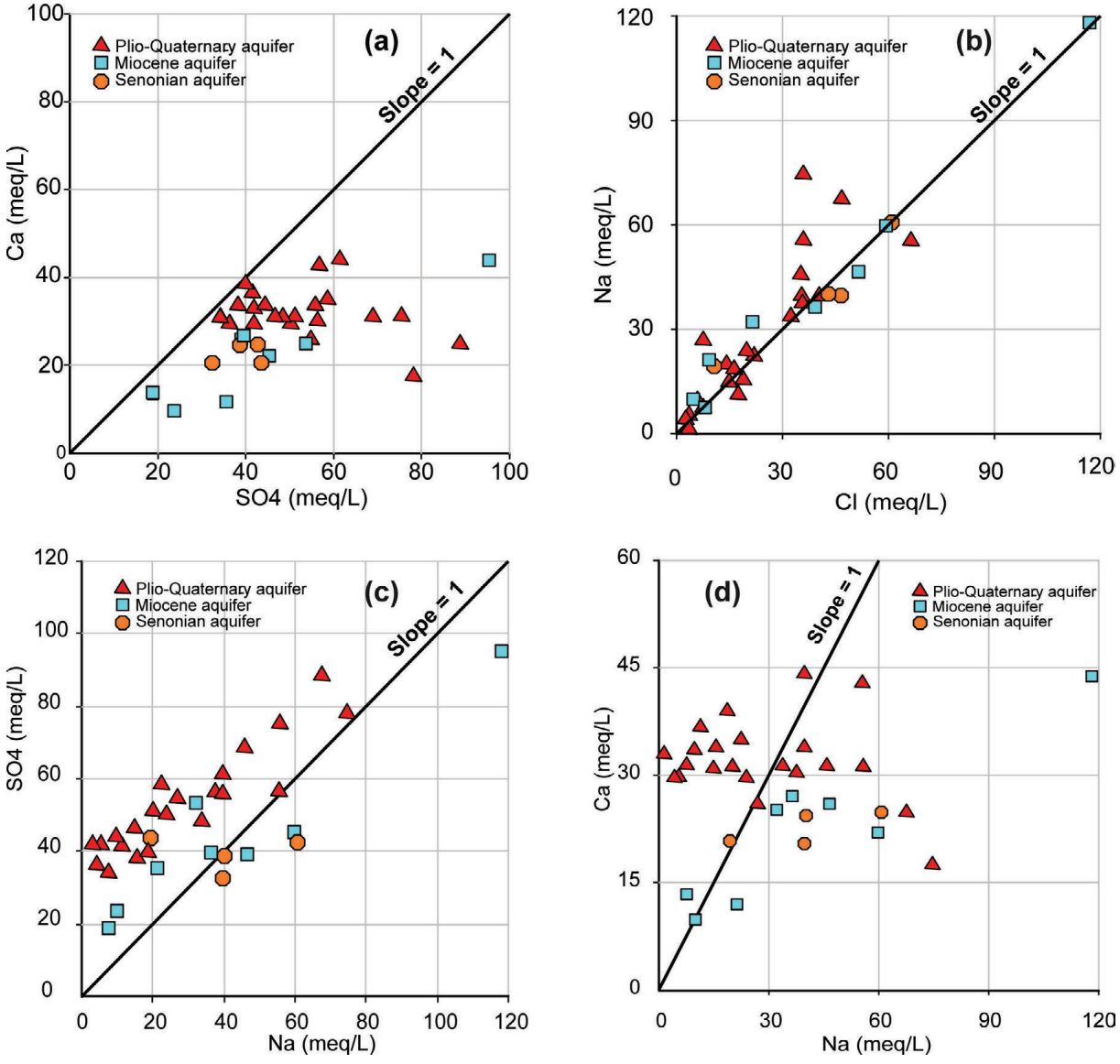

Fig. 4 - Pairs plots of Ca ~SO_4 (a), Na~Cl (b), S04~Na (c) and Ca~Na (d).

Fig. 4 - Grafici a coppie di Ca ~SO_4 (a), Na~Cl (b), SO4~Na (c) and Ca~Na (d).

Tab. 2 - Descriptive statistics of the studied parameters.

Tab. 2 - Statistiche descrittive dei parametri studiati.

vars	n	mean	SD	median	trimmed	mad	min	max	range	skew	kurtosis	SE
Temp	35	20.41	3.26	21.4	20.65	1.78	12	26.3	14.3	-0.81	0.41	0.55
pH	35	7.63	0.35	7.57	7.6	0.28	7.1	8.73	1.63	1.05	1.08	0.06
HCO_3	35	144.36	75.68	128	135.74	72.2	36.6	378	341.4	1.06	1.07	12.79
Cl	35	1133.61	977.85	776.85	987.02	734.11	124.09	4141	4016.91	1.54	2.4	165.29
SO_4	35	2390.07	820.95	2181.4	2327.89	672.21	907.02	4575.6	3668.58	0.8	0.29	138.77
Na	35	832.86	652.29	738.3	745.62	590.93	39	2760	2721	1.31	1.61	110.26
K	35	33.5	24.62	33.32	31.41	25.32	0	115.5	115.5	1.06	1.69	4.16
Ca	35	587.78	179	608	589.63	138.4	196.2	1008	811.8	-0.05	0.02	30.26
Mg	35	220.95	131.42	183.1	200.3	94.17	76	663	587	1.59	2.35	22.21
EC	35	6502.29	3433.37	6760	6234.48	3765.8	1400	15870	14470	0.82	0.69	580.34
TDS	35	5249.57	2701.14	5033	4872.76	2471.49	2090	13080	10990	1.15	1.03	456.58

be attributed to the dissolution of thenardite, gypsum, and halite (Dib et al., 2023; Kraiem et al., 2024a).

Descriptive statistics

The studied variables included the main physicochemical properties of the Sidi Mansour aquifers. In situ measured temperature ranged between 12 and 26.3 °C, with an average of 20.41 °C. All the average values (i.e., mean, median, and 10% trimmed dataset) were very close to each other, indicating very well-distributed data. This was further confirmed by skewness and kurtosis (Table 2).

Multivariate analyses

Given the present dataset, PCA constructed three main axes (dimensions) on the basis of the variables Dim1 (50.2%), Dim2 (15.3%), and Dim 3 (10.3%; Table 3; Fig. 5). TDS, Na^+, Cl^-, Mg^{2+} and SO_4^{2-} significantly contribute to Dim1, which can be plausibly explained by the dissolution of evaporites (halite, gypsum and thenardite and/or epsomite). This is the main mineralization mechanism. Other variables (i.e., HCO_3^- and Ca^{2+}) are more strongly correlated with Dim2, which is probably controlled by the dissolution of carbonates, as confirmed by the saturation indices of calcite (SI calcite) and dolomite (SI dolomite; Table 1). This mechanism is subordinated by cation exchange. We noticed a low contribution of K^+ and temperature to Dim3, which can be explained by sylvite (KCl) dissolution.

PCA-biplot visualization

The biplot displays points representing samples on a two-dimensional scatterplot, where the axes correspond to the most informative principal components. Additionally, arrows representing the variables are superimposed on the plot. The length of the arrows indicates the variable's importance, and their direction shows how they contribute to the variation captured by each principal component. This allows viewers to see how samples relate to each other and how variables influence those relationships, all within a single plot, making it a powerful tool for exploring complex datasets. The biplot revealed several interesting relationships between variables (Fig. 6). For example, TDS (g/L) and sodium are positively correlated, as their arrows point in the same direction. Conversely, other variables, such as pH and Cl, reveal a negative correlation. Additionally, magnesium and calcium are moderately positively correlated.

Tab. 3 - Eigenvalue and variance of different components.

Tab. 3 - Autovalore e varianza di differenti componenti.

Component		Eigenvalue	Variance (%)	Cumulative variance (%)
Dim1	Eigenvalue>1	5.52	50.16	50.16
Dim2		1.68	15.29	65.45
Dim3		1.14	10.33	75.78
Dim4	Eigenvalue<1 (discarded non-significant components)	0.75	6.82	82.6
Dim5		0.64	5.83	88.43
Dim6		0.51	4.63	93.06
Dim7		0.42	3.8	96.86
Dim8		0.18	1.67	98.53
Dim9		0.12	1.09	99.62
Dim10		0.04	0.32	99.94
Dim11		0.01	0.06	100

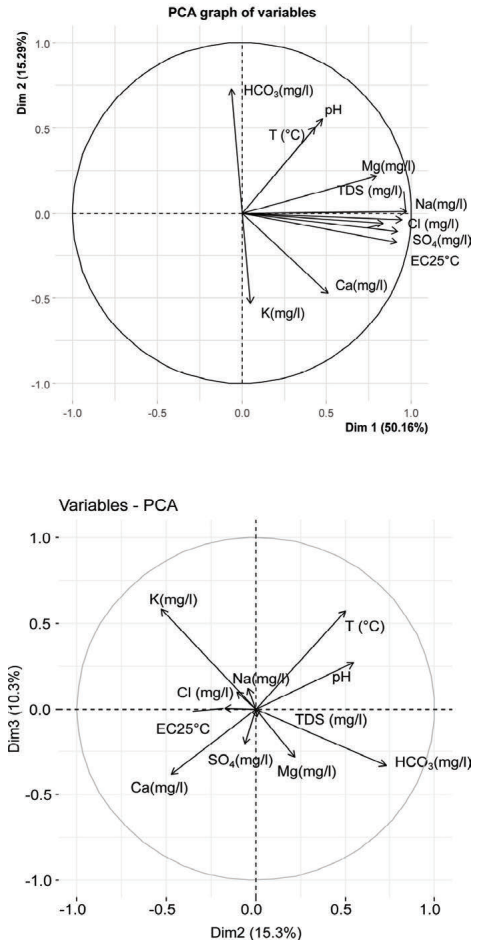

Fig. 5 - . PCA graph for variables. Dim2 vs Dim1 and Dim3 vs Dim2.

Fig. 5 - Grafici PCA per variabili Dim2 vs Dim1 e Dim3 vs Dim2.

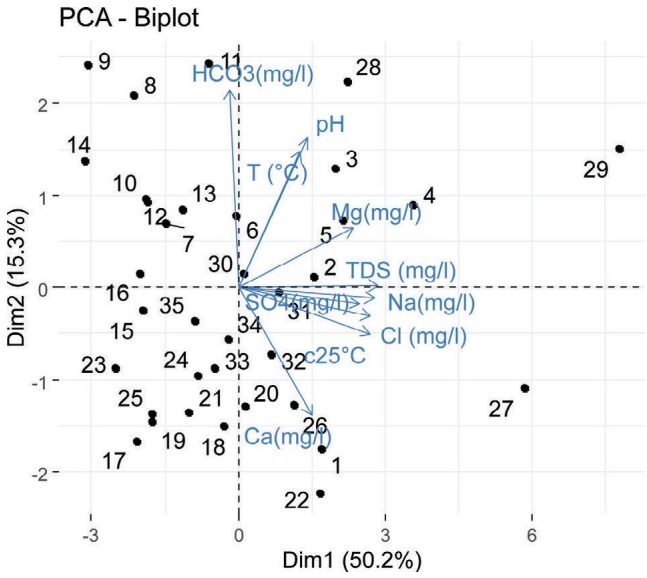

Fig. 6 - PCA biplot for the studied variables and individuals.

Fig. 6 - Biplot PCA per variabili e individui studiati.

Variable contributions

The expected average contribution (EAC) of each factor (i.e., 9.09%) was used. Looking at the data, we can easily recognize that the factors sharply decreased from 50.2% for Dim1 to 15.3% and then to 10.3%. These dimensions were considered more important, as they exceeded the expected average contribution of 9.09%. They explained the most important variance in the data, reaching 75.8%.

Dim1 was influenced, primarily, by six parameters (TDS, Na, Cl, SO_4, Mg, and EC), which exceeded the expected average contribution. This finding supports the role of evaporite mineral dissolution (halite, thenardite, and epsomite) and cation exchange in water mineralization. In contrast, Dim2 was dominated by HCO_3, K, Ca, pH, and temperature, suggesting the influence of carbonate mineral dissolution.

The overall contribution to both Dim1 and Dim2 (Dim1–2) indicated TDS (mg/L) as the greatest contributing variable, followed by Na (mg/L), EC (mS/cm) and Cl- (mg/L); all of them largely exceeded the EAC threshold (i.e., EAC=9.09%). In addition, the values of SO_4 and Mg (mg/L) were almost equal to the EAC values. Therefore, all of these factors significantly contributed to Dim1–2. In contrast, HCO_3, pH, temperature, Ca and K were not considered because of their small contributions to the overall

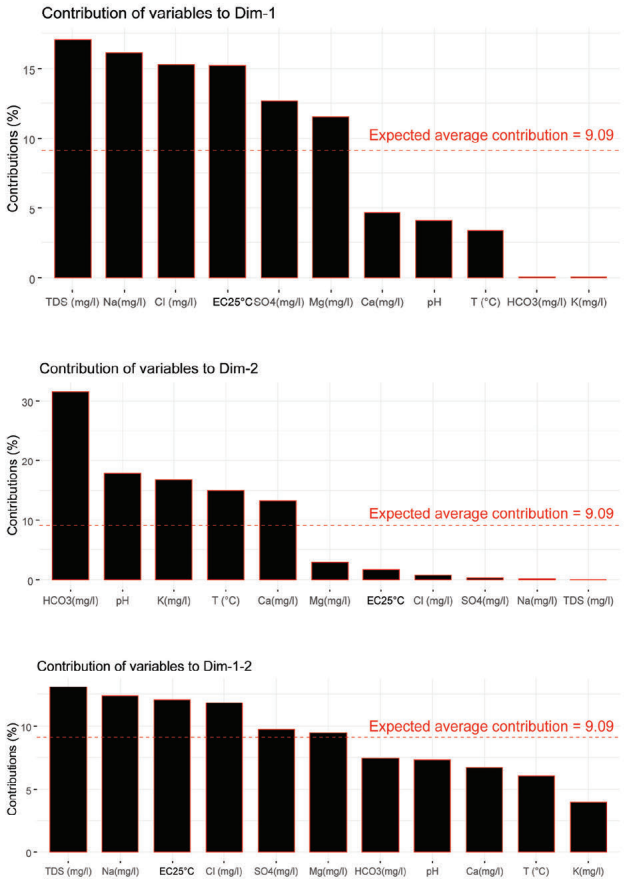

Fig. 7 - Contributions of variables to Dim1, Dim2 and Dim1-2.

Fig. 7 - Contributi delle variabili a Dim1, Dim2 and Dim1-2.

Dim1–2 (Fig. 7). It can be concluded that Dim1 and Dim2 primarily represent mineralization processes in this region, which are influenced by the dissolution of evaporite minerals (halite, thenardite, and epsomite) and cation exchange. In contrast, Dim3 suggests a secondary role for the dissolution of carbonate minerals.

Hierarchical cluster analysis

The dendrogram in Figure 8 displays five clusters, which visually represent the hierarchical relationships between data points. The greater the distance is, the less similar the clusters are. This means that Cluster1, Cluster2, and Cluster3 are very close to each other; they can be grouped into one larger cluster with 18 samples. Similarly, Cluster4 and Cluster5 can be grouped into a new group with 17 samples. Nevertheless, the number of clusters (5 in this case) is determined by cutting the dendrogram to the desired level (Phenon line=6.2 in our case). The level we chose depended on the application (Wickham et al., 2019). This is consistent with the results of Sahal and Kamel (2018), who investigated water quality origins (hydrochemical, mineralogical and isotopic) in the Djeffara aquifer system (southeastern Tunisia). Nevertheless, the decision regarding the overall number of clusters can vary depending on the researcher and the value of the Phenon line. In the case of the studied waters of the Sidi Mansour plain, Cluster1 can be recognized as a group of water layers prone to evaporation; it includes twelve samples (i.e., P32, P2, P28, P26, P33, P34, P30, P31, P22, P18, P20, and P1), most of which belong to the eastern edge of the studied area. The exceptions can be seen for P1, P2, and the subordinates P3, P4 and P5 (identifying Cluster2) of the western edge, which are aligned on the Om Ali fault. They can be influenced by mixing with the underlying water via fractures (i.e., the Om Ali fault). Both Cluster4 and Cluster5 are very close; they also represent waters from the eastern part, except for P6, P7, P8, P9, P10 and P11. Those groups (i.e., Cluster4 and Cluster5) are likely subjected to dissolution and cation exchange processes. This was not the case for samples P6, P7, P8, P9, P10 and P11, which were influenced by direct infiltration of meteoric water due to their low piezometric level (< 10 m). This finding is in perfect agreement with the conclusions of Ben Brahim et al. (2021).

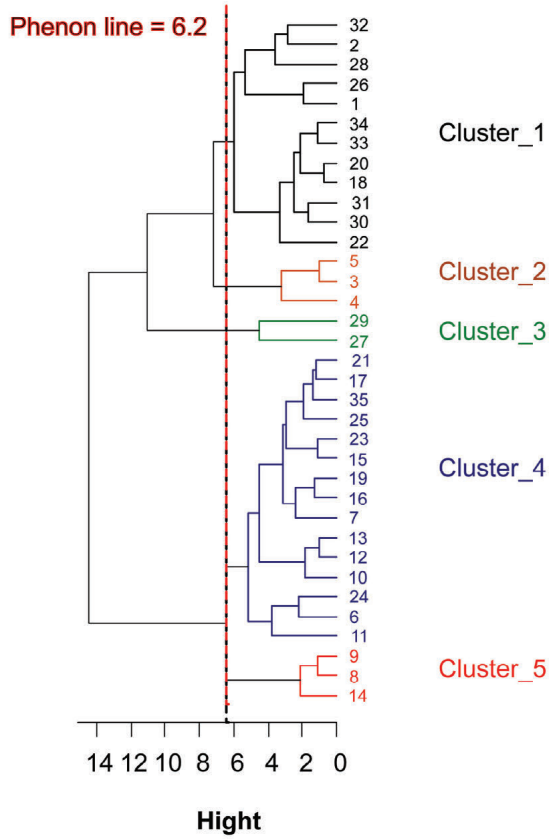

Fig. 8 - Hierarchical cluster analysis of the studied water samples.

Fig. 8 - Cluster analisi gerarchica dei campioni d'acqua studiati.

Isotopic study

For a better explanation of the main mineralization mechanisms underlying the saline character of the Sidi Mansour waters, isotopic analyses were carried out (^{18}O, ^{2}H, ^{3}H, ^{14}C). Isotopic signatures can differentiate water sources, such as groundwater, surface water, or mixed. This is fundamentally important in deciphering the age and mixing behavior of both old and freshwater. The obtained results were plotted on a $\delta^{18}O/\delta^{2}H$ correlation plot (Fig. 10a), which shows the position of the samples with respect to the global meteoric water line (GMWL: $\delta^{2}H=8\delta^{18}O+10$) (Craig 1961) and the local meteoric water line (LMWL: $\delta^{2}H=8\delta^{18}O+13.5$) (Celle-Jeanton et al., 2001; Maliki et al., 2000). Most samples are positioned below the LMWL and around the GMWL, clearly revealing the meteoric origin of those waters. Points F4 and F7, which capture the Miocene, have variable isotopic compositions but are higher than the average Miocene isotopic composition (i.e., -6.87‰ VSMOW and -47.49‰ VSMOW). This enrichment can be explained by the infiltration of recent water at these locations through existing faults. P5, P7, and P8 are depleted in heavy isotopes (P5 has tritium contents of 1.4 TU and is located in the Senonian level), and P7 and P8 were dug in the Miocene layers. Their depletion could be due to the altitude effect. The scatterplot clearly shows the effects of mixing and evaporation on the isotopic composition of groundwater in the study area. We recognize (1) the upper group representing freshwater, (2) the lower right group representing old water, possibly from the Miocene and/or Senonian aquifers, and (3) an intermediate group, perhaps coming from a mixture of freshwater and old water. This isotopic composition confirmed the effects of mixing and evaporation processes (Fig. 9b). This strongly supports the earlier interpretations given by hydrochemical and multivariate statistical analyses.

The Plio-Quaternary waters resulted from the infiltration of rainwater (high ^{14}C activities) but with high evaporation. Water samples from the Senonian and Miocene waters (with mean isotopic compositions of -6.87‰ VSMOW for ^{18}O and -45.8‰ VSMOW for ^{2}H) presented low ^{14}C activities and older apparent ages (≥10000 years B.P.); these samples are considered paleoclimatic waters.

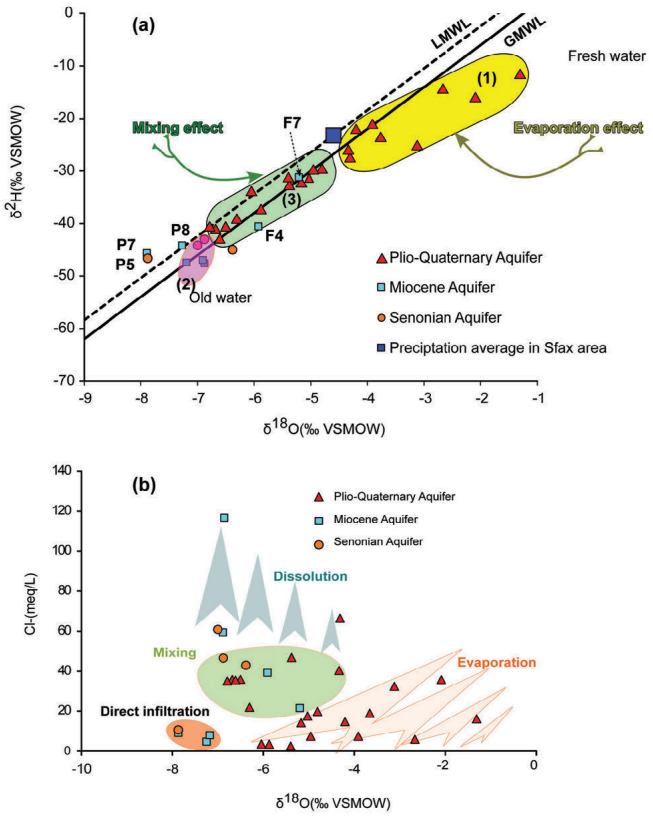

Fig. 9 - δ^2H versus $\delta^{18}O$ (a) and Cl- versus $\delta^{18}O$ (b) for the studied water samples from the Sidi Mansour area.

Fig. 9 - δ^2H vs $\delta^{18}O$ (a) e Cl vs $\delta^{18}O$ (b) per i campioni d'acqua studiati.

The Sidi Mansour region is located to the south of the Sfax area. Therefore, given the absence of a follow-up of precipitation in the study area, the average isotopic composition of precipitation in the Sfax region (-4.59‰ VSMOW for ^{18}O and -23.26‰ VSMOW for ^{2}H) was used as a recent pole (Celle-Jeanton et al., 2001; Maliki et al., 2000). Fig. 9a shows that, compared to precipitation, most of the Plio-Quaternary waters presented low isotope levels but remained higher than the average isotope level of the deeper aquifers. A plausible explanation can be allocated to the contribution of water from deep aquifers. Nevertheless, increased exploitation of Plio-Quaternary aquifers may promote upward leakage through major faults (the Om Ali fault and/or the Gafsa fault), threatening the whole aquifer system. The isotopic contents of the Plio-Quaternary aquifer therefore reflect a mixture of water recharged by the current rainfall and water from upward seepage from deeper aquifer levels (Miocene and Senonian). This mixing effect can be indicated by the distribution diagram of ^{18}O content versus ^{14}C activity (Fig. 10).

Chloride versus ^{18}O (Fig. 9b) provides a summary of the potential processes underlying water mineralization of the Sidi Mansour plain. They may include evaporation, dissolution, and mixing between different water types, possibly due to the effect of upward drainage through major faults (Gafsa fault and/or Om Ali fault; Fig. 11b). However, the exception of P11 would argue for minor hydraulic activity attributable to the Gafsa fault. This is related mainly to the effect of the Om Ali fault (Fig. 11b), indicating the prominent role of this fracture in the recharge of Plio-Quaternary aquifers by ascending drainage from deeper aquifers. This observation was supported by the distribution of ^{18}O contents, which evolved between very depleted waters (to the right of the Om Ali fault) and increasingly enriched waters eastward (Fig. 11a).

To calculate the abovementioned balance, the following assumptions were applied:

1. For the deep aquifer waters (i.e., Senonian and Miocene aquifers), the average isotopic composition was used (-6.87‰ VSMOW for ^{18}O and -45.8‰ VSMOW for ^{2}H);
2. For precipitation, the average precipitation of the Sfax area (-4.59‰ VSMOW for ^{18}O and -23.26‰ VSMOW for ^{2}H) was adopted.

Isotope analyses using $\delta^{18}O$ and $\delta^{2}H$ showed slight discrepancies in the calculated water balances (Fig. 9a). This difference is likely caused by evaporation, which preferentially removes lighter isotopes, leaving the remaining water enriched in ^{18}O rather than ^{2}H. The rate at which deep aquifer waters (Senonian and Miocene) refilled the Plio-Quaternary aquifer varied significantly across the study area, ranging from 9.6% to 96.5% (Fig. 11b). Higher rates of replenishment were

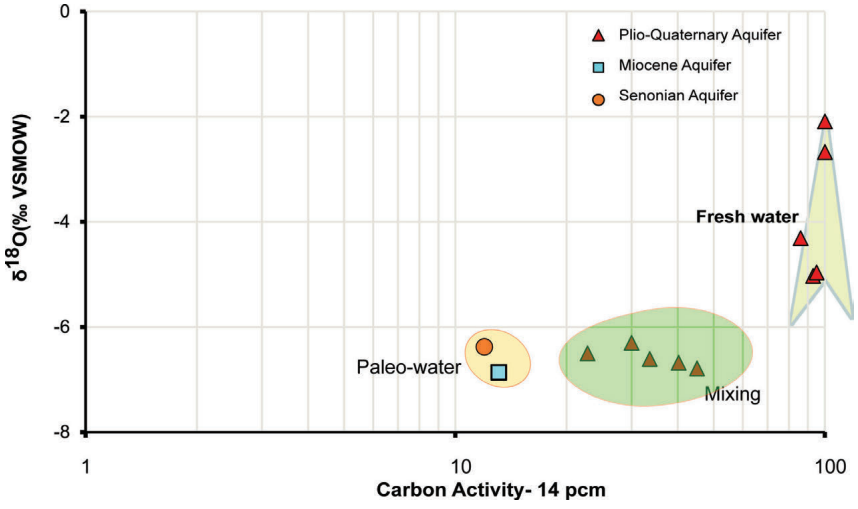

Fig. 10 - $\delta^{18}O$ vs ^{14}C activity distribution diagram.

Fig. 10 - $\delta^{18}O$ vs ^{14}C: diagramma di distribuzione dell'attività.

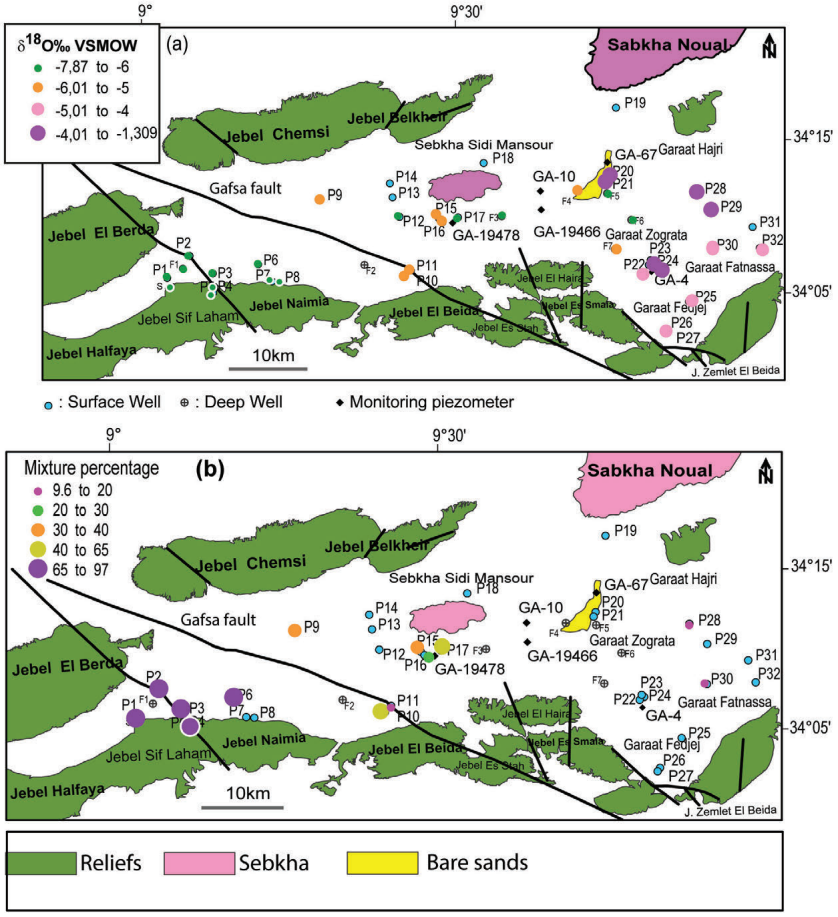

Fig. 11 - Distribution map of $\delta^{18}O$ (a) and mixture percentage (b) within the Sidi Mansour catchment area.

Fig. 11 - Carta della distribuzione di $\delta^{18}O$ entro l'area del bacino di Sidi Mansour (a) e percentuale di miscelamento (b).

found near the western edge, especially close to the Om Ali fault. These findings confirm that the Om Ali fault plays an important role in directing water to recharge the Plio-Quaternary aquifer.

Back propagation neural network (BPNN)

Backpropagation neural networks were used to explore Sidi Mansour water salinization. The data collected were split into 80% for training the model and 20% for testing and cross validation. The BPNN-developed model converged in a swift manner with 59 iterations. It reached the desired threshold of 0.01. The results matrix of the structured 10:4:1 BPNN is given in Table 4. One can easily recognize the positive weights of pH, Cl, Mg and EC at the first node (H1) and pH, HCO_3, Ca, and Mg at H2. Moreover, node H3 was most influenced by EC, Cl, Ca, K, and HCO_3, whereas only HCO_3 and pH had positive effects on node H4. These positive weights are shown as black solid lines with proportional thicknesses (Fig. 12). In contrast, all the negative values shown in Table 4 may exert negative effects on the relevant connection node; they appear as gray solid lines in the BPNN structure. The activation functions connecting the hidden layer nodes (i.e., H1, H2, H3, and H4) to the general output (TDS in mg/L) confirmed the influential weights from bias (B2) and node H1; both were dominated by node H3. The output from H2 was the least influential, whereas H4 had a negative impact on the TDS content.

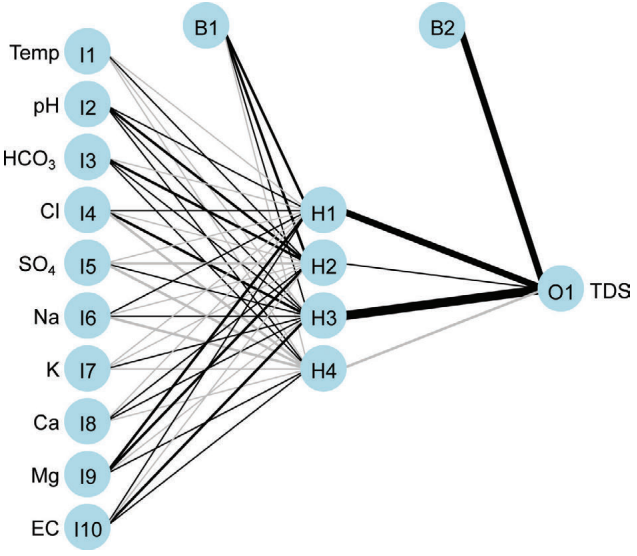

Fig. 12 - Back propagation neural network structure (10:4:1) used to predict TDS (mg/L).

Fig. 12 - Struttura della rete neurale di *back propagation* usata per prevedere TDS (mg/L).

Tab. 4 - Results matrix of the proposed model for 10:4:1 BPNN.

Tab. 4 - Matrice dei risultatidel modello proposto per 10:4:1 BPNN.

	Node 1 (H1)	Node 2 (H2)	Node 3 (H3)	Node 4 (H4)		TDS (mg/L)
Intercept (B1)	1.11	1.72	0.37	-0.68	B2	5.2
Temperature	-0.48	0.07	-0.23	-0.41	H1	4.81
pH	0.23	1.13	0.02	1.1	H2	1.03
HCO_3	-0.3	1.98	0.31	0.99	H3	6.55
Cl	0.87	-0.28	1.33	-1.39	H4	-1.21
SO^4	-0.35	-1.32	0.12	-1.17		
Na	0.52	-0.24	0.71	-1.66		
K	-0.39	-0.21	0.78	-0.12		
Ca	-1.09	0.15	0.91	-0.14		
Mg	1.21	1.71	-0.57	0.06		
EC	0.74	-0.33	1.63	0.25		

Model evaluation

In machine learning, cross-validation is a crucial step for ensuring a model's generalizability and robustness. It tackles the challenge of evaluating a model on unseen data. In the present study, R-squared was used as a metric to evaluate the model fitting robustness (Fig. 13). It represents the proportion of variance in the target variable explained by model. Table 5 presents the performance metrics of the prediction model used for the Sidi Mansour water assessment with the BPNN. The data collected were split into 80% for training the model and 20% for testing and cross validation. The results were evaluated using various statistical performance criteria and revealed that the BPNN model properly predicted the salinity of the Plio-Quaternary water samples. The model achieved high accuracy (88.89%) and could perfectly identify all positive cases (100% sensitivity). However, its ability to distinguish negative cases is only moderate (50% specificity). This means that the model might misclassify some negative instances as positive. The prevalence (77.78%) indicates that the positive class is more common than the negative class. The coefficient of determination (0.9687) and kappa coefficient (60%) suggested a good correlation between the measured and predicted values (Table 5).

The BPNN model results (10:4:1) indicated the high accuracy of the trained algorithm, as confirmed by a cross-validation test. This study highlights the importance of this model in predicting salinity for the Plio-Quaternary aquifer in this region.

Tab. 5 - Confusion matrix based predicted versus measured TDS.

Tab. 5 - Matrice di confusione basata su TDS previsti vs misurati.

Criterion	%
Accuracy	88.89
Sensitivity	100.00
Specificity	50
Positive Predicted Value	87.5
Negative Predicted Value	100.00
Prevalence	77.78
Balanced Accuracy	75
Kappa	60

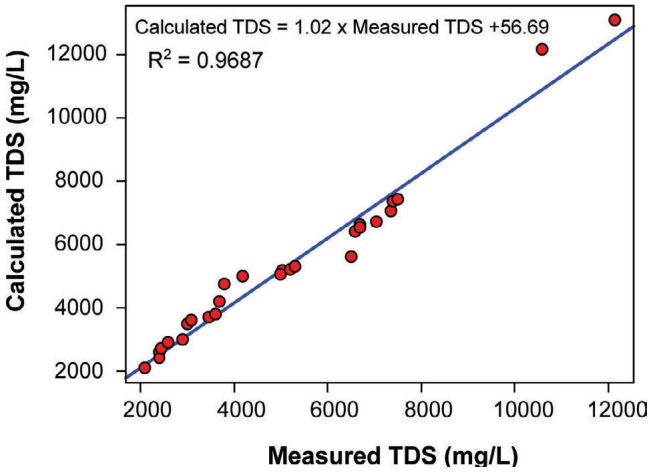

Fig. 13 - Scatter plot showing predicted versus measured TDS (mg/L).

Fig. 13 - Scatter plot illustrante TDS (mg/L) previsti vs misurati.

Analysis methods: advantages and limitations

The application of an integrated approach has contributed to obtaining a clear understanding of the hydrochemical properties of shallow groundwater from southern Tunisia. Multivariate statistical analyses (i.e., PCA, HCA, biplot), ANN, stable isotopes of water (^{18}O, ^{2}H), carbon-14, and tritium offered several advantages in the current hydrogeological study. For example, multivariate statistical analyses may address complex relationships between multiple variables, providing insights into water quality and possible mineralization processes (Rodríguez et al., 2021). Artificial neural networks (ANNs) can model nonlinear relationships, making them valuable for predicting water quality, as is the case for the present water samples. Similar findings have been reported by numerous researchers (Abuzir and Abuzir, 2022; Islam Khan et al., 2022; Uddin et al., 2023). Stable isotopes of water (e.g., deuterium and oxygen-18) provide valuable information on water sources, evaporation, and recharge

processes. However, these advanced techniques have several limitations. PCA can be difficult to interpret and often requires assumptions about how the data are distributed (Ayed et al., 2017). Additionally, training artificial neural networks (ANNs) can be complex, potentially leading to uncertain results. While stable isotopes provide valuable information, their interpretation can be intricate due to complex processes. Moreover, the use of radioisotopes, such as carbon-14 and tritium, is expensive and requires specialized equipment, although they provide crucial information about groundwater age and residence times.

Conclusions

The application of hydrogeochemical, isotopic, and machine learning tools in the characterization of Sidi Mansour waters could help with integrated assessments of hydrodynamic functioning and identification of the main geochemical processes involved in the mineralization of these waters. Several conclusions can be drawn:

- Chemical analyses revealed different types of waters with deep Miocene and Senonian aquifers as Na-Cl-SO_4 facies in the west migrating to Na-Ca-SO_4 facies in the east. The studied Plio-Quaternary aquifer featured Na-Ca-SO_4 facies to the west, but Ca-SO_4-dominated waters to the east.
- Salinity has a heterogeneous spatial distribution due to the combination of several factors, such as the dissolution of evaporites (halite, gypsum, thenardite, and epsomite), cation exchange with clay minerals, evaporation and mixing of water.
- The correlations between several major elements (Na^+, Cl^-, Ca^{2+}, and SO_4^{2-}) revealed the dominant role of evaporite dissolution (mainly halite, gypsum, thenardite and epsomite) in the mineralization of Plio-Quaternary waters. Nevertheless, Miocene and Senonian waters are affected by exchanges with clay minerals.
- The developed BPNN model successfully predicts the overall salinity with high accuracy, sensitivity, and coefficient of determination (R^2=0.9687). On the basis of salinity and other physicochemical parameters, the trained BPNN aimed to identify areas of water scarcity or predict future water availability of the collected water samples. Isotopic analyses (^{18}O, ^{2}H, ^{3}H and ^{14}C) revealed the meteoric origin of the waters in the Sidi Mansour plain.
- The eastern edge of the aquifer was influenced by evaporation, and the western part was influenced by the mixing effect of upward drainage from deeper aquifers, especially at the Om Ali fault, where the contribution can reach 96%.

This study provides insight into water management in southern Tunisia, which requires a multipronged approach that combine traditional techniques with new technologies and a strong emphasis on groundwater conservation for future generations.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the first author used Google AI service in order to generate description and polished text. After using this tool, the author reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgment

We would like to thank the regional commissariat for agricultural development (Water Resources Division) of Kebili and Gabes for their logistic help during the sampling campaigns. Special appreciations go to Professor Rudy Rossetto and the anonymous reviewers who have significantly improved the final version of the submitted manuscript.

Author contributions

Zohra KRAIEM: Conceptualization (lead), Methodology(lead), Formal analysis, Writing - original draft (lead), Writing - review & editing (lead), Supervision (equal), Data Curation (lead), Visualization (equal), Validation (equal).
Kamel ZOUARI: Conceptualization, Methodology, Visualization (equal), Validation (equal).
Aissa HLEIMI: Formal analysis (supporting), Methodology (supporting), Visualization (equal), Validation (equal).
Houda DERBEL FETOUI: Formal analysis (supporting), Writing - review & editing, Resources (supporting), Data Curation (supporting), Software (lead), Visualization (equal), Validation (equal).
All authors have read and agreed to the published version of the manuscript.

Funding source

This research received no external funding.

Competing interest

The authors declare no competing interests.

Additional information

DOI: https://doi.org/10.7343/as-2025-804 **Reprint and permission information** are available writing to acquesotterranee@anipapozzi.it
Publisher's note Associazione Acque Sotterranee remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

Abid, K., Zouari, K. & Abidi, B. (2010). Identification and characterisation of hydrogeological relays of continental intercalaire aquifer of southern Tunisia. Carbonates and Evaporites 25, 65–75. Springer.

Abuzir, S.Y., Abuzir, Y.S., (2022). Machine learning for water quality classification. Water Qual. Res. J. 57, 152–164. https://doi.org/10.2166/wqrj.2022.004

Ahmadi, R., Ouali, J., Mercier, E., Mansy, J.-L., Van-Vliet Lanoë, B., Launeau, P., Rhekhiss, F., et al. (2006). The geomorphologic responses to hinge migration in the fault-related folds in the Southern Tunisian Atlas. J. Struct. Geol. 28(4), 721–728. doi:https://doi.org/10.1016/j.jsg.2006.01.004

Ayed, B., Jmal, I., Sahal, S., Mokadem, N., Saidi, S., Boughariou, E., Bouri, S. (2017). Hydrochemical characterization of groundwater using multivariate statistical analysis: the Maritime Djeffara shallow aquifer (Southeastern Tunisia). Environ. Earth Sci. 76. https://doi.org/10.1007/s12665-017-7168-6

Azad, A., Karami, H., Farzin, S., Mousavi, S.F., Kisi, O. (2019). Modeling river water quality parameters using modified adaptive neuro fuzzy inference system. Water Sci. Eng. 12, 45–54. https://doi.org/10.1016/j.wse.2018.11.001

Ben Brahim, F., Boughariou, E., Bouri, S. (2021). Multicriteria-analysis of deep groundwater quality using WQI and fuzzy logic tool in GIS: A case study of Kebilli region, SW Tunisia. J. African Earth Sci. 180, 104224. https://doi.org/10.1016/j.jafrearsci.2021.104224

Bhagat, S. K., Tung, T. M. & Yaseen, Z. M. (2020) Development of artificial intelligence for modeling wastewater heavy metal removal: State of the art, application assessment and possible future research. J. Clean. Prod. 250, 119473. Elsevier. doi:10.1016/J.JCLEPRO.2019.119473

Bouaziz, S., Barrier, E., Soussi, M., Turki, M. M. & Zouari, H. (2002) Tectonic evolution of the northern African margin in Tunisia from paleostress data and sedimentary record. Tectonophysics 357(1–4), 227–253. Elsevier. doi:10.1016/S0040-1951(02)00370-0

Boughriba, M., Melloul, A., Zarhloule, Y. & Ouardi, A. (2006) Extension spatiale de la salinisation des ressources en eau et modèle conceptuel des sources salées dans la plaine des Triffa (Maroc nord-oriental) *"Spatial extension of salinization in groundwater and conceptual model of the brackish springs in the Triffa plain (northeastern Morocco)"*. Comptes Rendus. Géoscience 338(11), 768–774. Elsevier. doi:10.1016/j.crte.2006.07.007

Castany, G. (1954) Etude géologique de l'Atlas tunisien oriental. Ann. des mines la géologie, Tunis 8, 632,.

Celle-Jeanton, H., Zouari, K., Travi, Y. & Daoud, A. (2001) Caractérisation isotopique des pluies en Tunisie. Essai de typologie dans la région de Sfax *"Isotopic characterisation of the precipitation in Tunisia. Variations of the stable isotope compositions of rainfall events related to the origin of air masses"*. Comptes Rendus l'Académie des Sci. IIA-Earth Planet. Sci. 333(10), 625–631. Elsevier.

Cemek, B., Arslan, H., Küçüktopcu, E., Simsek, H. (2022). Comparative analysis of machine learning techniques for estimating groundwater deuterium and oxygen-18 isotopes. Stoch. Environ. Res. Risk Assess. 36, 4271–4285. https://doi.org/10.1007/s00477-022-02262-7

Che Nordin, N. F., Mohd, N. S., Koting, S., Ismail, Z., Sherif, M. & El-Shafie, A. (2021) Groundwater quality forecasting modelling using artificial intelligence: A review. Groundw. Sustain. Dev. 14, 100643. Elsevier B.V. doi:10.1016/j.gsd.2021.100643

Çimen, Ö., Keskin, S. N. & Yıldırım, H. (2012) Prediction of Swelling Potential and Pressure in Compacted Clay. Arab. J. Sci. Eng. 37(6), 1535–1546. Springer Verlag. doi:10.1007/S13369-012-0268-4

Dib, I., Khedidja, A., Chettah, W., (2023). Multivariate statistical analysis of the alluvial aquifer of Tadjenanet- Chelghoum Laid (Eastern Algeria). Acque Sotter. - Ital. J. Groundw. 45:643, 67–75. https://doi.org/10.7343/as-2023-643

Edmunds, W. M., Guendouz, A. H., Mamou, A., Moulla, A., Shand, P. & Zouari, K. (2003) Groundwater evolution in the Continental Intercalaire aquifer of southern Algeria and Tunisia: trace element and isotopic indicators. Appl. Geochemistry 18(6), 805–822. doi:https://doi.org/10.1016/S0883-2927(02)00189-0

Elshaarawy, M.K., Hamed, A.K. (2024). Predicting discharge coefficient of triangular side orifice using ANN and GEP models. Water Sci. 38, 1–20. https://doi.org/10.1080/23570008.2023.2290301

Ferjani, A. Ben, Burollet, P. F. & Mejri, F. (1990) Petroleum geology of Tunisia. Enterprise tunisienne d'activités pétrolières.

Fontes, J. C. (1976) Isotopes du milieu et cycles des eaux naturelles: quelques aspects. These Dr. d'Etat, Univ. Pierre Marie Curie.

Fritz, P., Silva Hennings, C., Suzuki, O. & Salati, E. (1979) Isotope hydrology in northern Chile. In: Isotope hydrology 1978.

Georgescu, P. L., Moldovanu, S., Iticescu, C., Calmuc, M., Calmuc, V., Topa, C. & Moraru, L. (2023) Assessing and forecasting water quality in the Danube River by using neural network approaches. Sci. Total Environ. 879(January), 162998. The Authors. doi:10.1016/j.scitotenv.2023.162998

Gonfiantini, R., Conrad, G., Fontes, J. C., Sauzay, G. & Payne, B. R. (1974). Etude isotopique de la nappe du Continental Intercalaire et de ses relations avec les autres nappes du Sahara septentrional *"Isotopic study of the continental intercalary nappe and its relations with the other nappes of the northern Sahara"*. Isot. Tech. Groundw. Hydrol. 1, 227–241. IAEA Vienna.

Grassi, S., Cortecci, G. & Squarci, P. (2007). Groundwater resource degradation in coastal plains: The example of the Cecina area (Tuscany – Central Italy). Appl. Geochemistry 22(11), 2273–2289. doi:https://doi.org/10.1016/j.apgeochem.2007.04.025

Hadj Ali, M. Ben, Jedoui, Y., Dali, T., Salem, N. Ben & Memmi, L. (1985). Carte géologique de la Tunisie à 1/500,000 *"Geological map of Tunisia 1/500000"*. Off Nat Mines, Serv Geol Nat Tunisie.

Islam Khan, M.S., Islam, N., Uddin, J., Islam, S., Nasir, M.K. (2022). Water quality prediction and classification based on principal component regression and gradient boosting classifier approach. J. King Saud Univ. - Comput. Inf. Sci. 34, 4773–4781. https://doi.org/10.1016/J.JKSUCI.2021.06.003

Jemai, S., Kallel, A., Agoubi, B., & Abida, H. (2022). Spatial and temporal rainfall variability and its controlling factors under an arid climate condition: case of Gabes Catchment, Southern Tunisia. Environment, Development and Sustainability, 24(4), 5496–5513. https://doi.org/10.1007/s10668-021-01668-7

Kattan, Z. (2008). Estimation of evaporation and irrigation return flow in arid zones using stable isotope ratios and chloride mass-balance analysis: Case of the Euphrates River, Syria. J. Arid Environ. 72(5), 730–747. doi:https://doi.org/10.1016/j.jaridenv.2007.10.011

Keskin, T. E., Düğenci, M. & Kaçaroğlu, F. (2015). Prediction of water pollution sources using artificial neural networks in the study areas of Sivas, Karabük and Bartın (Turkey). Environ. Earth Sci. 73(9), 5333–5347. doi:10.1007/s12665-014-3784-6

Khalil, A., Almasri, M.N., McKee, M., Kaluarachchi, J.J. (2005). Applicability of statistical learning algorithms in groundwater quality modeling. Water Resour. Res. 41, 1–16. https://doi.org/10.1029/2004WR003608

Khullar, S., Singh, N. (2021). Machine learning techniques in river water quality modelling: A research travelogue. Water Sci. Technol. Water Supply 21. https://doi.org/10.2166/ws.2020.277

Kouadri, S., Elbeltagi, A., Islam, A.R.M.T., Kateb, S. (2021). Performance of machine learning methods in predicting water quality index based on irregular data set: application on Illizi region (Algerian southeast). Appl. Water Sci. 11, 1–20. https://doi.org/10.1007/S13201-021-01528-9

Kouzana, L., Mammou, A. Ben & Felfoul, M. S. (2009). Seawater intrusion and associated processes: case of the Korba aquifer (Cap-Bon, Tunisia). Comptes Rendus. Géoscience 341(1), 21–35.

Kraiem, Z., Zouari, K. & Chkir, N. (2024a). Accurate prediction of salinity in Chott Djerid shallow aquifers, southern Tunisia: Machine learning model development. Water Sci. 38(1), 33–47. Taylor & Francis. doi:10.1080/23570008.2023.2294535

Kraiem, Z., Zouari, K. & Trabelsi, R. (2024b). Harnessing machine learning tools for water quality assessment in the Kebili shallow aquifers , Southwestern Tunisia. Acta Geochim. (0123456789). Science Press. doi:10.1007/s11631-024-00689-z

Krimissa, S., Michelot, J.-L., Bouchaou, L., Mudry, J. & Hsissou, Y. (2004). Sur l'origine par altération du substratum schisteux de la minéralisation chlorurée des eaux d'une nappe côtière sous climat semi-aride (Chtouka-Massa, Maroc) *"About the origin of chloride in groundwater from a coastal aquifer under semi-arid climate (Chtouka-Massa, Morocco)"* . Comptes rendus. Géoscience 336(15), 1363–1369.

Lallahem, S., Mania, J., Hani, A., Najjar, Y. (2005). On the use of neural networks to evaluate groundwater levels in fractured media. J. Hydrol. 307, 92–111. https://doi.org/10.1016/J.JHYDROL.2004.10.005

Leontiadis, I. L., Payne, B. R. & Christodoulou, T. (1988). Isotope hydrology of the Aghios Nikolaos area of Crete, Greece. J. Hydrol. 98(1), 121–132. doi:https://doi.org/10.1016/0022-1694(88)90209-0

Maliki, A., Krimissa, M., Michelot, J.-L. & Zouari, K. (2000). Relation entre nappes superficielles et aquifère profond dans le bassin de Sfax (Tunisie) *"Relationship between shallow and deep aquifers in the Sfax basin (Tunisia)*. Comptes Rendus l'Académie des Sci. - Ser. IIA - Earth Planet. Sci. 331(1), 1–6. Elsevier Masson. doi:10.1016/S1251-8050(00)01386-0

Maroua, B.A., Nouiri, I., Tarhouni, J., (2015). Spatio-temporal variation of standardized precipitation index (spi) in tunisia and mapping of degree of drought severity for the period1973-2015. 2nd ICIEM 2016, International Conference on Integrated Environmental Management for Sustainable Development

Mastrocicco, M. (2021). Studies on water resources salinization along the Italian coast: 30 years of work. Acque Sotter. - Ital. J. Groundw. 39:537, 7–13. https://doi.org/10.7343/as-2021-537

Pham, Q. B., Mohammadi, B., Moazenzadeh, R., Heddam, S., Zolá, R. P., Sankaran, A., Gupta, V., et al. (2023). Prediction of lake water-level fluctuations using adaptive neuro-fuzzy inference system hybridized with metaheuristic optimization algorithms. Appl. Water Sci. 13(1), 13. doi:10.1007/s13201-022-01815-z

Pritha Bhandari (2019). Descriptive Statistics|Definitions, Types, Examples. scribbr.com. Retrieved April 8, 2024, from https://www.scribbr.com/statistics/descriptive-statistics/

Qin, D., Qian, Y., Han, L., Wang, Z., Li, C. & Zhao, Z. (2011). Assessing impact of irrigation water on groundwater recharge and quality in arid environment using CFCs, tritium and stable isotopes, in the Zhangye Basin, Northwest China. J. Hydrol. 405(1), 194–208. doi:https://doi.org/10.1016/j.jhydrol.2011.05.023

Ramos-Leal, J. A., Martínez-Ruiz, V. J., Rangel-Mendez, J. R. & Alfaro de la Torre, M. C. (2007). Hydrogeological and mixing process of waters in aquifers in arid regions: a case study in San Luis Potosi Valley, Mexico. Environ. Geol. 53(2), 325–337. doi:10.1007/s00254-007-0648-3

Rapti-Caputo, D. & Martinelli, G. (2009). The geochemical and isotopic composition of aquifer systems in the deltaic region of the Po River plain (northern Italy). Hydrogeol. J. 17(2), 467–480. doi:10.1007/s10040-008-0370-6

Rashid, A., Kumari, S. (2023). Performance evaluation of ANN and ANFIS models for estimating velocity and pressure in water distribution networks. Water Supply 0, 1. https://doi.org/10.2166/WS.2023.224

Rodríguez, R., Pastorini, M., Etcheverry, L., Chreties, C., Fossati, M., Castro, A., Gorgoglione, A., (2021). Water-Quality Data Imputation with a High Percentage of Missing Values: A Machine Learning Approach. Sustainability 13, 6318. https://doi.org/10.3390/su13116318

Sahal, S. & Kamel, S. (2018). Investigation of water quality origins (hydrochemical, mineralogical and isotopic) in the Jeffara of Medenine aquifer system (South-Eastern Tunisia). J. Water Supply Res. Technol. 67(6), 586–606. doi:10.2166/aqua.2018.019

Schiavo, M., Giambastiani, B. M. S., Greggio, N., Colombani, N., and Mastrocicco, M. (2024). Geostatistical assessment of groundwater Arsenic contamination in the Padana Plain. Science for the Total Environment. https://doi.org/10.1016/j.scitotenv.2024.172998

Smida, H., Tarki, M., Gammoudi, N., Dassi, L. (2023). GIS-based multicriteria and artificial neural network (ANN) investigation for the assessment of groundwater vulnerability and pollution hazard in the Braga shallow aquifer (Central Tunisia): A critical review of generic and modified DRASTIC models. J. Contam. Hydrol. 104245. https://doi.org/10.1016/J.JCONHYD.2023.104245

Socki RA, Karlsson HR, Gibson EK Jr. (1992). Extraction technique for the determination of oxygen-18 in water using preevacuated glass vials. Anal Chem. 1992 Apr 1;64(7):829-31. doi: 10.1021/ac00031a026. PMID: 11536503.

Sokono, Y., Diallo, O., Sy, L. B., Baubion, C., Latrech, D. & Mamou, A. (2008). The North-Western Sahara Aquifer System (Algeria, Tunisia, Libya): Concerted management of a transboundary water basin. Retrieved from http://www.ossonline.org/sites/default/files/publications/OSS-SASS-CSn1_En.pdf

Soro, GE; Anouman DL; Goula BI; TA; Srohorou B; Savane I. Caractérisation des séquences desécheresse météorologique à diverses échelles de temps en climat de type Soudanais : Cas de l'extrême NordOuest de la Cote d'Ivoire *"Characterization of metrological drought sequences at various time scales in Sudanese-Type climate: Case of the far north-west of Cote d'Ivoire*. Larhyss Journal 2014 ,18, 107-124.

Subyani, A. M. (2004). Use of chloride-mass balance and environmental isotopes for evaluation of groundwater recharge in the alluvial aquifer, Wadi Tharad, western Saudi Arabia. Environ. Geol. 46(6), 741–749. doi:10.1007/s00254-004-1096-y

Taherian, P., Joodavi, A. (2021). Hydrogeochemical characteristics and source identification of salinity in groundwater resources in an arid plain , northeast of Iran: implication for drinking and irrigation purposes. Acque Sotter. - Ital. J. Groundw. 67 :502, 21–31. https://doi.org/10.7343/as-2021-502

Taylor, C. B. (1976). IAEA isotope hydrology laboratory. Vienna Int. At. Energy Agency. Tech. Proced. note (19).

Trabelsi, R., Kacem, A., Zouari, K. & Rozanski, K. (2009). Quantifying regional groundwater flow between Continental Intercalaire and Djeffara aquifers in southern Tunisia using isotope methods. Environ. Geol. 58(1), 171–183. John Wiley & Sons, Ltd. doi:10.1007/s00254-008-1503-x

Trabelsi, R., Zaïri, M., Smida, H. & Dhia, H. Ben. (2005). Salinisation des nappes côtières: cas de la nappe nord du Sahel de Sfax, Tunisie *"Salinization of coastal aquifers: case of the North Sfax Sahel aquifer, Tunisia"*. Comptes Rendus. Géoscience 337(5), 515–524. doi:10.1016/j.crte.2005.01.010

Uddin, M.G., Nash, S., Rahman, A., Olbert, A.I. (2023). Assessing optimization techniques for improving water quality model. J. Clean. Prod. 385, 135671. https://doi.org/10.1016/j.jclepro.2022.135671

Uddin, M.G., Nash, S., Rahman, A., Olbert, A.I., (2023). Assessing optimization techniques for improving water quality model. J. Clean. Prod. 385, 135671. https://doi.org/10.1016/j.jclepro.2022.135671

Wederni, K., Schiavo, M., Boulbaba, H., Hamed, Y., Bouri, S., and Colombani, N. (2024) SEAWAT modelling scenarios evaluating links among the Southern Gabès (TN) confined aquifer and the Mediterranean Sea. Water-3205793. Accepted.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. 4(43), 1686. doi:10.21105/joss.01686

Yaseen, Z. M. (2021). An insight into machine learning models era in simulating soil, water bodies and adsorption heavy metals: Review, challenges and solutions. Chemosphere 277, 130126. Elsevier Ltd. doi:10.1016/j.chemosphere.2021.130126

Zammouri, M., Siegfried, T., El-Fahem, T., Kriâa, S. & Kinzelbach, W. (2007). Salinization of groundwater in the Nefzawa oases region, Tunisia: results of a regional-scale hydrogeologic approach. Hydrogeol. J. 15(7), 1357–1375. doi:10.1007/s10040-007-0185-x

Zaqoot, H.A., Hamada, M., Miqdad, S. (2018). A Comparative Study Of Ann For Predicting Nitrate Concentration In Groundwater Wells In The Southern Area Of Gaza Strip. Appl. Artif. Intell. 32, 727–744. https://doi.org/10.1080/08839514.2018.1506970