

Article

Hydrochemical Characterization and Predictive Modeling of Groundwater Quality in Karst Aquifers Under Semi-Arid Climate: A Case Study of Ghar Boumaaza, Algeria

Sabrina Guettaia ¹, Abderrezzak Boudjema ¹, Abdessamed Derdour ^{2,3,*}, Abdessalam Laoufi ¹, Hussein Almohamad ⁴, Motrih Al-Mutiry ⁵ and Hazem Ghassan Abdo ⁶

¹ Laboratory n°25 Promotion of Water, Mineral and Soil Resources, Environmental Legislation and Technological Choices, University of Tlemcen, P.O. Box 119, Tlemcen 13000, Algeria; sabrine.guettaia@univ-tlemcen.dz (S.G.); abderrezzak.boudjema@univ-tlemcen.dz (A.B.); abdessalam.laoufi@univ-tlemcen.dz (A.L.)

² Artificial Intelligence Laboratory for Mechanical and Civil Structures, and Soil, University Center of Naama, P.O. Box 66, Naama 45000, Algeria

³ Laboratory for the Sustainable Management of Natural Resources in Arid and Semi Arid Zones, University Center of Naama, P.O. Box 66, Naama 45000, Algeria

⁴ Department of Geography, College of Languages and Human Sciences, Qassim University, Buraydah 51452, Saudi Arabia; h.almohamad@qu.edu.sa

⁵ Department of Geography and Environmental Sustainability, College of Humanities and Social Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia; mkalmutairy@pnu.edu.sa

⁶ National Institute of Oceanography and Applied Geophysics, OGS, Via Treviso 55, 33100 Udine, Italy

* Correspondence: derdour@cuniv-naama.dz

Abstract

Understanding groundwater quality in karst environments is essential, particularly in semi-arid regions where water resources are highly vulnerable to both climatic variability and anthropogenic pressures. The Ghar Boumaaza karst aquifer, located in the semi-arid Tlemcen Mountains of Algeria, represents a critical yet understudied water resource increasingly threatened by climate change and human activity. This study integrates hydrochemical analysis, multivariate statistical techniques, and predictive modeling to assess groundwater quality and characterize the relationship between total dissolved solids (TDSs) and discharge (Q). An analysis of 66 water samples revealed that 96.97% belonged to a Ca^{2+} – HCO_3^- facies, reflecting carbonate rock dissolution, while 3% exhibited a Cl^- – HCO_3^- facies associated with agricultural contamination. A principal component analysis identified carbonate weathering (40.35%) and agricultural leaching (18.67%) as the dominant drivers of mineralization. A third-degree polynomial regression model ($R^2 = 0.953$) effectively captured the nonlinear relationship between TDSs and flow, demonstrating strong predictive capacity. Independent validation ($R^2 = 0.954$) confirmed the model's robustness and reliability. This study provides the first integrated hydrogeochemical assessment of the Ghar Boumaaza system in decades and offers a transferable methodological framework for managing vulnerable karst aquifers under similar climatic and anthropogenic conditions.

Keywords: karst; water; Ghar Boumaaza; Tlemcen; aquifer; hydrochemistry; sustainable management



check for updates

Academic Editor: Xiaomeng Song

Received: 28 June 2025

Revised: 15 July 2025

Accepted: 18 July 2025

Published: 29 July 2025

Citation: Guettaia, S.; Boudjema, A.; Derdour, A.; Laoufi, A.; Almohamad, H.; Al-Mutiry, M.; Abdo, H.G.

Hydrochemical Characterization and Predictive Modeling of Groundwater Quality in Karst Aquifers Under Semi-Arid Climate: A Case Study of Ghar Boumaaza, Algeria.

Sustainability **2025**, *17*, 6883. <https://doi.org/10.3390/su17156883>

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Although karst regions cover only 7 to 12 percent of the Earth's land area, they provide 20 to 25 percent of the world's drinking water needs [1,2]. These aquifers, formed by the

dissolution of rocks, create complex networks of cracks, crevices, and conduits, providing high permeability and storage capacity [3]. However, their heterogeneity and anisotropy, combined with high permeability, make them particularly vulnerable to surface contaminants, with limited self-purification capacity [4]. To do this, the chemical composition of the waters sheltered in these aquifers is influenced by natural processes (hydrogeological conditions, precipitation, and lithology) and human activities [5,6]. Understanding the dynamics of dissolved solids is essential to assess the weathering of karst systems, as it directly influences water quality and weathering processes [7]. Studies show that turbidity and particle transport, as well as mineral saturation, are determinants of this dynamic [8,9]. In addition, elements such as temperature, pH, and saltwater circulation play a crucial role in the chemical dissolution of carbonate rocks, thus affecting their porosity and integrity [7,10]. This relationship is based on the complex interaction between solution chemistry and the weathering of karst systems, highlighting the link between dissolved solids (TDSs) and groundwater flow (Q) [11]. The establishment of a reliable methodology for assessing the amount of dissolved solids as a function of flow variations is crucial for water resource management and decision making in various sectors [12]. Traditional methods of measuring and monitoring concentrations of chemical elements in water have limitations in terms of cost, time, and spatial coverage [13]. To address these challenges, hydrogeochemical studies and multivariate statistical analyses have been adopted to understand and predict dissolved solids concentration as a function of flow rates and to characterize the geochemical and hydrodynamic processes influencing water quality [14]. These approaches are crucial for assessing the environmental threats and vulnerabilities of these systems, especially in the context of pollution and environmental degradation. Recent studies have highlighted the importance of these methods in assessing the vulnerability of karst systems. For example, a study conducted in northern Morocco used vulnerability mapping techniques, such as the VUKA method, to assess the sensitivity of karst aquifers to contamination [15]. These methods are designed to address the unique challenges posed by karst terrains, including those in North Africa. In addition, recent research has highlighted the usefulness of multivariate statistical techniques, such as principal component analysis (PCA), for analyzing the natural and anthropogenic origins of elements in cave systems [16]. The Q- and TDS-based regression curve is an essential tool for studying the relationship between total dissolved solids concentration and flow rate. The construction of this TDS vs. Q curve first involves the collection of flow and dissolved solids data over a significant period of time. Next, it is essential to perform a thorough statistical analysis to determine the nature of the relationship between the two variables. This analysis includes the calculation of the correlation, slope, and y-intercept of the regression curve. The choice of the type of regression model (linear, quadratic, etc.) also requires a justification based on relevant statistical criteria [17]. They are useful for identifying appropriate probability distributions in hydrological studies, especially when the available samples are small and the parental distribution is highly skewed. The integration of these approaches and geological and environmental conditions makes the construction and interpretation of the Q-function TDS regression curve more robust and reliable, allowing a better understanding of the dynamics of dissolved solids in relation to discharge in hydrological systems. This method offers new perspectives for the evaluation and prediction of dissolved solids concentrations as a function of flows, allowing a more efficient and more precise management of water resources. These techniques help to understand water–rock interactions and the impact of climate change on water quality, and to develop more effective conservation strategies to protect these unique and fragile ecosystems, such as Ghar Boumaaza.

The Tlemcen Mountains, home to Algeria's largest underground cave system, called Ghar Boumaaza, stretch for more than 4 km. Despite its importance, Ghar Boumaaza has

not benefited from in-depth scientific research since 1985, with the exception of that of Bensaoula in 2017 [18]. Overgrazing, agriculture, and uncontrolled exploitation threaten its ecological integrity. In addition, this system is very vulnerable to pollution, with more than 66% of its surface classified as a high vulnerability zone according to the R.I.S.K. method. [19].

However, most previous studies were limited to descriptive or point analyses, without integrating robust predictive models or combining hydrochemical, multivariate statistical, and nonlinear modeling approaches to characterize dissolved matter dynamics, leaving Ghar Boumaaza's system largely underdocumented. In this context, the present study proposes an integrated approach combining hydrochemical characterization, multivariate statistical analysis (MCA), and nonlinear predictive modeling (polynomial regression of degree 3) to study the relationship between TDSs and flow in the Ghar Boumaaza karst. In addition, it updates and completes the knowledge on this major karst system, taking into account recent anthropogenic and climatic pressures, and provides a methodological framework that can be transferred to other karst aquifers in the region, thus responding to the need to develop management and protection tools adapted to the specific vulnerability of North African karst systems.

2. Materials and Methods

2.1. Study Area

Located in the mountains of Tlemcen, near the city of Seb Dou, the Ghar Boumaaza karst system is located at the following coordinates: latitude $34^{\circ}37'$ North and longitude $1^{\circ}20'$ West. This remarkable site is located about twenty kilometers from Tlemcen, the main city of the region (Figure 1). In recent decades, the Tlemcen region has experienced strong population growth, leading to increasing pressure on water resources, especially on the Ghar Boumaaza karst system. The population's drinking water supply is based on a few boreholes and springs. However, these infrastructures are frequently exposed to contamination linked to domestic water, particularly in the locality of Dar Maamar [19]. This situation highlights the vulnerability of karst water resources to anthropogenic pollution. The regional climate is semi-arid Mediterranean, with cool winters and hot, dry summers. According to data from the Meffrouch weather station (1976–2022), average annual precipitation varies from 350 to 500 mm, with peaks in winter and spring. These seasonal rains contribute significantly to aquifer recharge and influence flow variability [20]. The hydrodynamic behavior of the system is highly variable. Flow rates at the outlet vary from 0 to 4000 L/s, closely depending on rainfall events. Dye Tracing Experiments [21] provide confirmation of a fast subsurface connectivity, typical of conduit-dominated karst systems. This indicates minimal filtration and short residence times, which increases vulnerability to pollution.

2.2. Geological and Hydrogeological Setting

The underground karst network of Ghar Boumaaza developed within limestone and dolomitic formations in a synclinal structure (the Merchiche syncline) oriented SW–NE. According to Benest [22], these formations are composed, from base to summit, of the following units: the Terni dolomites (DTI-100 m) and the Hariga marl–limestone (MC—mainly limestone over 195 m and sandstone over 52 m, the latter corresponding to the Merchiche sandstones). The sandstone formations feed small, slow-flowing springs, used both for water supply for local populations and for watering livestock. In total, these formations represent 347 m of mainly calcareous deposits, often difficult to dissociate due to dolomitization. They rest on an impermeable bedrock, made up of the marl–limestone of Raourai, which forms the base of the aquifer and through which the galleries of the

Ghar Boumaaza karst network were formed [21]. This complex geological structure is typical of Mediterranean karst systems, where past geological events have shaped the current hydrological features. Recent studies, based on aeromagnetic data, have made it possible to precisely map the geological structures of the Tlemcen region, highlighting the complexity of the Paleozoic basement as well as the sedimentary cover of the Mesozoic and Cenozoic (Figure 2). The karst system is said to be “perched”, the main outlets being the springs of Ain Taga and Ain Hassi El Kelb, located about 500 m downstream of the overflow excrecence called Ghar Boumaaza. The latter provides access to the underground river of Ghar Boumaaza [23]. A route made by Birebent in the 1950s, from a loss on the left bank of the underground river, 1600 m from the entrance to the cave, confirmed that the entire flow came out of the Ain Taga spring and Ain El Kelb after only a few days [21].

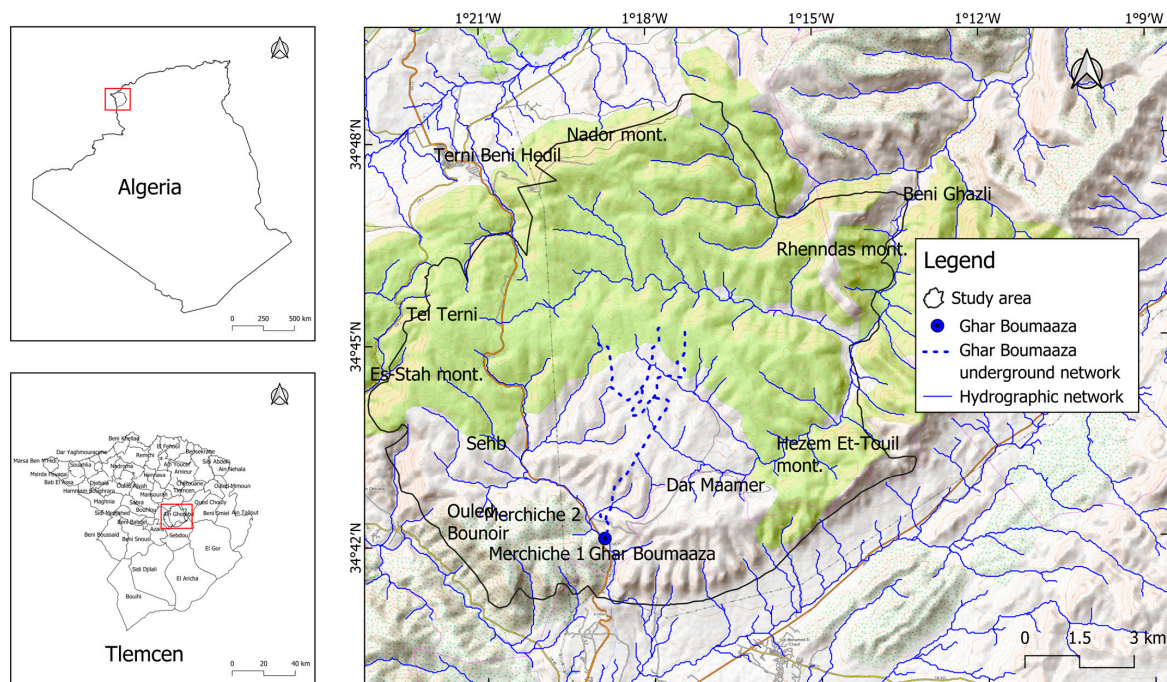


Figure 1. Location of the study area.

2.3. Sampling and Analysis

A total of sixty-seven groundwater samples were collected directly from the Ghar Boumaaza cave system. All water samples were taken at the cave’s main outlet immediately upon the appearance of flow, reflecting the system’s sensitivity to hydrological inputs. Sampling was conducted over a multi-year period, from January 2013 to May 2023 (excluding the COVID-19 pandemic period), covering a wide range of seasonal and interannual conditions.

The sampling was carried out at an irregular frequency, primarily driven by flow availability and weather conditions, particularly rainfall events. On average, sampling occurred approximately every 1.5 to 2 months, with higher frequency (up to every 15 days) during periods of increased hydrological activity, such as the wet season. This allowed for better representation of the dynamic nature of the karst spring system.

It should be noted that the flow regime of the Ghar Boumaaza wadi, characteristic of karst systems in semi-arid climates, is highly variable. It alternates between sudden, high-flow events, often exceeding 4000 L/s following intense rainfall, and long periods of dryness during the summer. This behavior is strongly influenced by the basin’s karstic structure and rapid response to precipitation.

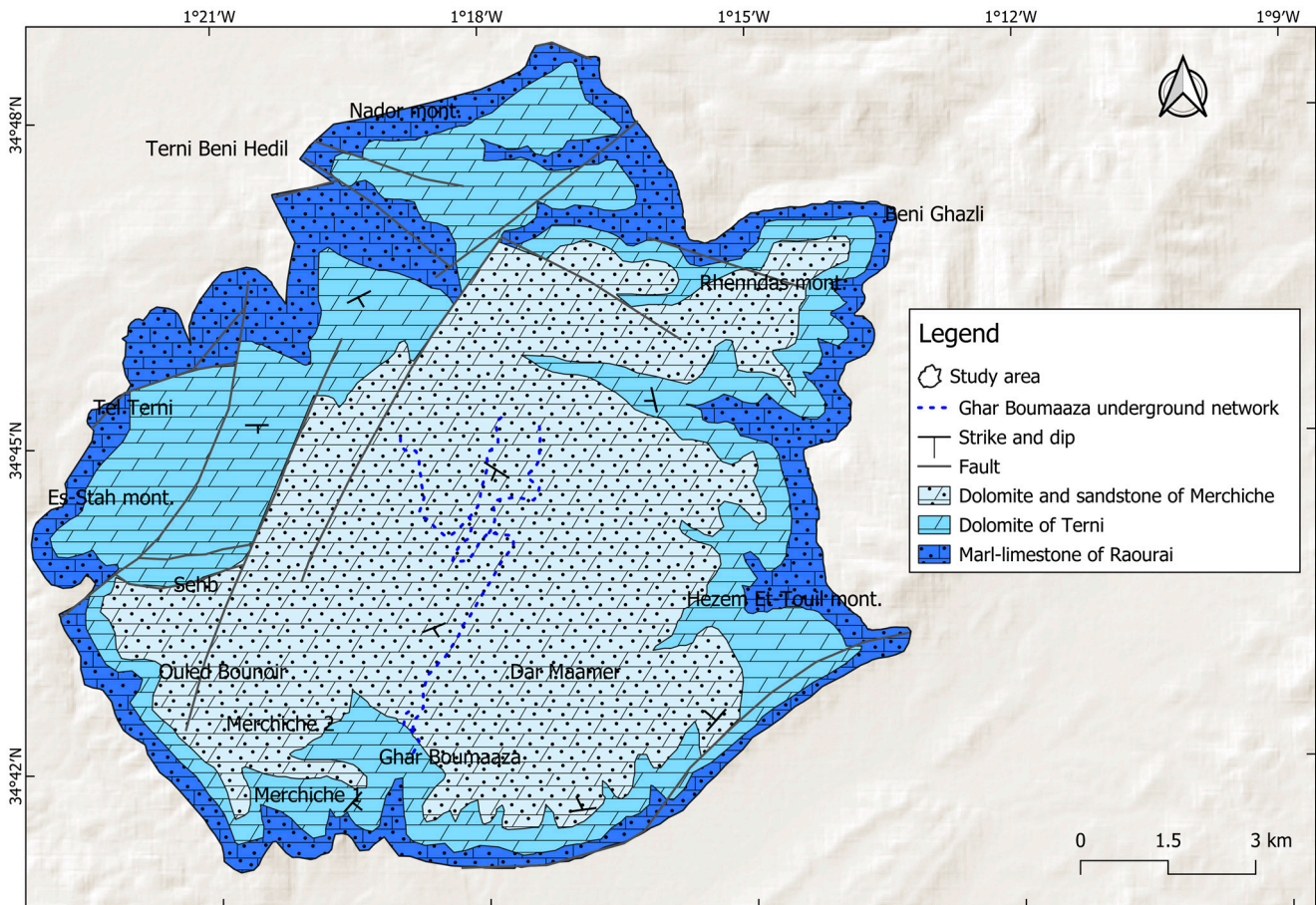


Figure 2. Geological map of the study area.

The sampling campaigns were carried out in coordination with the Directorate of Water Resources of the Wilaya of Tlemcen, under the technical supervision of hydrogeologists from the National Agency for Water Resources (ANRH). Each sample was analyzed for the concentrations of eight major ions: sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), magnesium (Mg^{2+}), bicarbonate (HCO_3^-), chloride (Cl^-), sulfate (SO_4^{2-}), and nitrate (NO_3^-).

Sample collection, retention, and transportation were carried out in accordance with standardized protocols, in strict compliance with quality assurance and quality control (QA/QC) guidelines to minimize contamination and ensure data integrity. Samples were stored in polyethylene cylinders, filtered in situ ($0.45 \mu\text{m}$), and cooled to 4°C prior to analysis, in accordance with APHA [24] recommendations for water sampling procedures.

The chemical analyses were carried out in fully accredited ANRH laboratories, officially authorized to carry out groundwater quality assessments in agricultural and ecologically sensitive areas of Algeria, in compliance with national regulatory frameworks. Once the ion concentration measurements were completed, the analytical accuracy was verified by calculating the ion balance, using Equation (1). The load balance error (IB%), expressed as a percentage, ensures internal consistency between the measured cation and anion concentrations:

$$IB(\%) = \frac{[\sum \text{cations} - \sum \text{anions}]}{\sum \text{cations} + \sum \text{anions}} \times 100 \quad (1)$$

where

- Cations, expressed in milliequivalents per liter (meq/L), include positively charged ions such as calcium, magnesium, sodium, and potassium, contributing to the overall positive charge of the water.
- Anions, also expressed in milliequivalents per liter (meq/L), include negatively charged ions such as bicarbonate, carbonate, chloride, sulfate, and nitrate, which balance the total ionic charge in the water sample.

An acceptable ion balance is typically less than $\pm 5\%$ for high-quality analytical data. If the balance exceeds this range, it may indicate sampling, analysis, or calculation errors [25].

2.4. Chart Analysis

Hydrochemical diagrams such as Piper, Chadha, and the ascending hierarchical classification (HFC) are fundamental tools in hydrogeology to decipher the origin and evolution of groundwater. These graphical representations make it possible to transform complex analytical data into intuitive visualizations, revealing essential information about mineralization processes, the geochemical origin of water, water–rock interactions, and the mechanisms of recharge and evolution of aquifers. Note that the CAH diagram offers a statistical classification of samples, the Piper diagram allows us to characterize the overall chemical composition, and the Gibbs diagram analyzes the control mechanisms of the chemistry, while the Chadha diagram provides a detailed representation of the hydrochemical facies.

2.4.1. Ascending Hierarchical Classification (CAH)

Ascending hierarchical classification (HFC) is an effective statistical technique for the multivariate analysis of hydrochemical data, based on the fundamentals of hierarchical clustering [26]. It facilitates the categorization of water samples with similar chemical properties by converting complex data into a tree structure known as a dendrogram. The main objectives include the identification of homogeneous water families, the elucidation of similarities and differences between samples, the simplification of hydrochemical data, and the identification of common sources or processes of mineralization. The procedure adheres to specific methodological steps: data normalization, calculation of the distance between samples, hierarchical creation of clusters, and graphical display. CAH allows for the fast processing of multiple variables, clear visualization of correlations between samples, and independence from the number of variables. The interpretation of the dendrogram is based on the idea that each branch represents a group of samples, with the height of the link reflecting the degree of similarity, thus facilitating the understanding of the genesis and evolution of groundwater [27].

2.4.2. Piper Diagram

The Piper diagram, developed by Piper [28], is a graphical representation used to illustrate the main chemical elements and the different facies of groundwater. It makes it possible to visualize the evolution of the water, passing from one facies to another, either by analyses spaced out in time or by analyses of samples taken from various locations. This diagram is invaluable for representing various analysis groups. It consists of two triangles and a diamond. The two triangles, one for the cations and the other for the anions, are filled first, followed by the diamond. The values used are expressed in % meq/L.

2.4.3. Chadha Diagram

The Chadha diagram, introduced by Chadha [29], is a valuable tool for hydrochemical analysis. By plotting the sample groups on this diagram, we can describe different types

of water and trace the evolution of hydrochemical processes. This diagram allows us to understand the factors that control groundwater chemistry in the study area. It provides a clear visual representation of the interactions between the different chemical elements in the water. This detailed understanding helps identify changes over time and variations between different locations, providing insight into the underlying processes that affect groundwater quality

2.5. Statistical Analyses

Modern hydrogeochemical analysis effectively combines principal component analysis (PCA) and multiple linear regression (MLR) to decipher the complexities of aquifer systems. This approach, which has been widely validated by recent studies, is particularly suitable for the analysis of datasets including ion concentrations (Ca^{2+} , Mg^{2+} , Na^+ , K^+ , Cl^- , SO_4^{2-} , HCO_3^- , and NO_3^-), conductivity, measured TDSs, and flow rate (Q in l/s) [30]. PCA, by reducing the dimensionality of the data, makes it possible to identify the main factors influencing water quality, and this is performed through the use of a correlation matrix and varimax rotation [30]. This method is particularly effective for identifying and distributing sources of groundwater pollution, a common problem in hydrochemical studies [31]. The MLR, applied to the principal components, then makes it possible to model the dependence of target variables such as the TDSs on the flow, offering a direct interpretation of the coefficients and an easy validation by statistical indicators such as R^2 and RMSE [32]. To ensure the robustness of the model, it is recommended to divide the data into training (70%) and test (30%) subsets, thus allowing a rigorous evaluation of predictive performance by residual analysis [33]. This hybrid methodology, combining PCA and MLR, offers not only a clear interpretation of complex hydrochemical interactions but also a predictive capacity adapted to the specificities of hydrogeochemical data, thus facilitating a more informed and sustainable management of groundwater resources in the face of growing environmental challenges.

3. Results

This section presents the main results of the hydrochemical and statistical analyses carried out on groundwater samples collected from the Ghar Boumaaza karst system. The results are based on graphical interpretations (cluster analysis and Piper and Chadha diagrams), descriptive and multivariate statistical evaluations, and the development of a predictive model relating total dissolved solids (TDSs) to releases (Q). These analyses provide an overview of the dominant hydrochemical facies, spatial and temporal variability, pollution indicators, and the main geochemical processes influencing groundwater quality in the study area.

3.1. Chart Analysis

3.1.1. Cluster Analysis

Applying a cluster analysis to the sampled water data reveals the presence of two distinct classes (Figure 3):

- Class 1: Includes ions associated with saline and evaporitic formations, such as Mg^{2+} , Na^+ , K^+ , Cl^- , and SO_4^{2-} . Also, this class is marked by the presence of a narrow line between K^+ and NO_3^- , testifying to the presence of anthropogenic pollution.
- Class 2: characterized by the presence of bicarbonates (HCO_3^-) and Ca^{2+} .

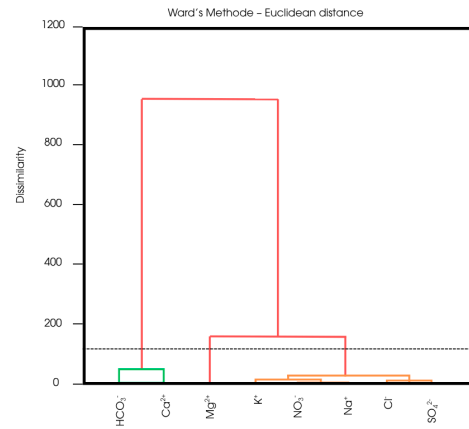


Figure 3. Ascending hierarchical classification (HAC) of the chemical parameters of the waters of the Ghar Boumaaza cave (2013–2023).

3.1.2. Piper Diagram

The Piper diagram is a commonly used tool for categorizing and illustrating the chemical composition of water samples based on their predominant ion concentrations (Figure 4). The report of the analysis results on this diagram indicates that the HCO_3^- - Ca^{2+} - Mg^{2+} facies is predominant, constituting 96.97% of the water samples. The groundwater in the study area contains mainly substantial amounts of calcium, magnesium, and bicarbonate ions. This facies is frequently linked to groundwater which has undergone considerable interaction with carbonate rocks, resulting in the decomposition of calcium and magnesium carbonates. This water is usually present in recharge areas when rainfall seeps into the soil and solubilizes carbonate minerals. The Cl^- - SO_4^{2-} - Ca^{2+} - Na^+ facies is represented by 3% of the samples, presented by samples 13 and 33, indicating a significant impact of sodium and potassium in conjunction with chloride and sulphate, which may indicate that mineralization is due to the presence of evaporitic formations or anthropogenic pollution.

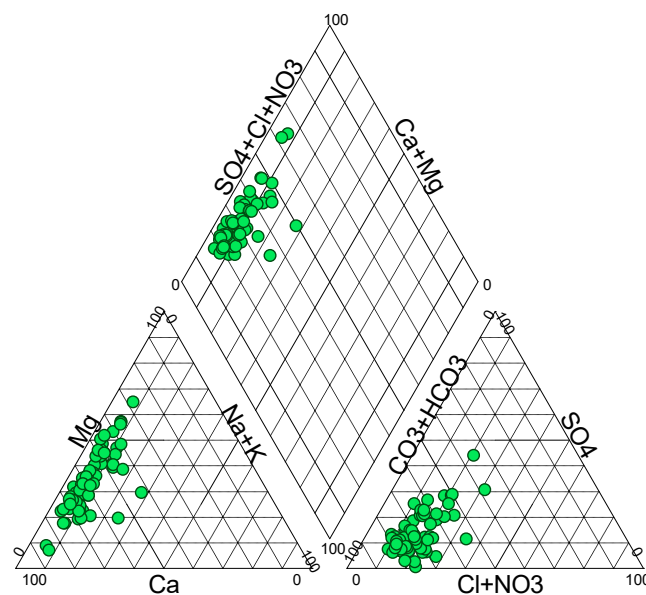


Figure 4. Piper's diagram of the waters of the Ghar Boumaaza cave (2013–2023).

3.1.3. Chadha Diagram

The Chadha diagram is a useful tool for classifying groundwater types and visualizing the dominant hydrochemical processes influencing water chemistry. In this study, the

diagram was applied to assess the hydrochemical facies of groundwater samples from the Ghar Boumaaza cave system over the period 2013–2023 (Figure 5).

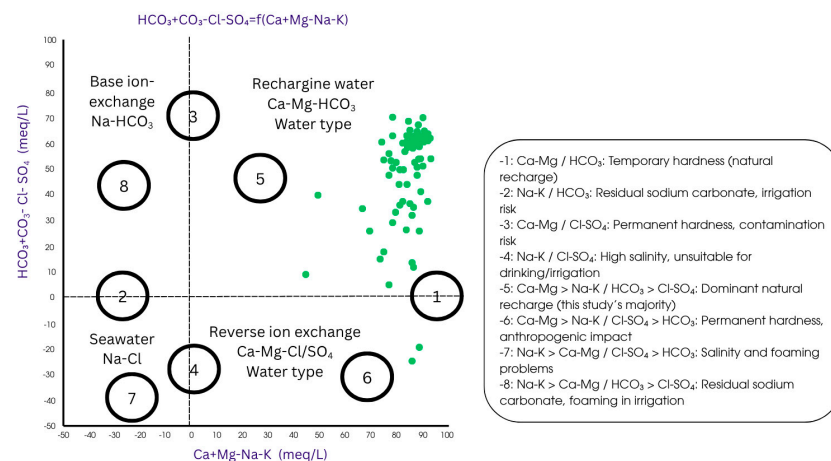


Figure 5. Chadha diagram of groundwater samples from the Ghar Boumaaza cave (2013–2023), illustrating hydrochemical facies and dominant ion relationships.

The majority of samples are plotted in Field 5 of the Chadha diagram, where alkaline earths (Ca^{2+} and Mg^{2+}) exceed alkali metals (Na^+ and K^+), and weak acidic anions (HCO_3^-) exceed strong acidic anions (Cl^- and SO_4^{2-}). This distribution reflects a dominance of carbonate weathering, resulting in Ca^{2+} – Mg^{2+} – HCO_3^- -type waters. These waters are typically characterized by temporary hardness and are commonly associated with active recharge zones, where slightly acidic rainwater reacts with carbonate rocks such as calcite and dolomite. The presence of this facies suggests a system largely controlled by natural geochemical processes and highlights the vulnerability of the recharge area to potential surface contamination.

Two samples, 13 and 33, fall into Field 6, where alkaline earths still dominate over alkali metals, but strong acidic anions exceed weak acidic anions. These samples correspond to Ca^{2+} – Mg^{2+} – Cl^- -type waters, indicating the influence of ion exchange processes or anthropogenic inputs such as agricultural leaching. Such waters are associated with permanent hardness and do not leave residual sodium carbonate, making them distinct from the majority of the sampled waters.

At the end, the Chadha diagram confirms that the primary hydrochemical processes in the Ghar Boumaaza aquifer are driven by carbonate dissolution, with localized deviations pointing to external influences that may require further investigation.

3.2. Statistical Analysis

A descriptive statistical analysis of the chemical elements of the waters of Ghar Boumaaza reveals a temporal variability ranging from low to significant in their distribution. The majority of the coefficients of variation, with the exception of those for Ca^{2+} and HCO_3^- , are close to or slightly above 50% (Table 1), indicating moderate heterogeneity in the concentration of these elements in the study area. This heterogeneity could result from a combination of geological, hydrological, and anthropogenic factors influencing the chemical composition of the waters. On the other hand, bicarbonates and Ca^{2+} and HCO_3^- have a relative stability of their concentrations, with coefficients of variation of less than 30%. This suggests a relatively homogeneous temporal distribution, potentially due to a more uniform source or process governing their concentrations, such as the dissolution of carbonate rock. Outliers were identified and eliminated using statistical methods, including the exclusion of values greater than 1.5 times the interquartile range. It is important to

note that these outliers, especially in flow (Q) and total dissolved solids (TDSs) measurements, can influence the interpretation of the relationship between variables and should be carefully considered.

Table 1. Physicochemical analysis of the waters of the Ghar Boumaaza cave (2013–2023).

Variable	Minimum Concentrations (Min)	Maximum Concentrations (Max)	Coefficient of Variation (CV)
Ca ²⁺	31.0	120.0	29.8
Mg ²⁺	6.0	51.0	43.1
Na ⁺	5.0	35.0	51.1
K ⁺	0.0	4.0	55.4
Cl ⁻	11.0	82.0	43.7
SO ₄ ²⁺	1.0	127.0	61.7
HCO ₃ ⁻	120.0	372.0	25.8
NO ₃ ⁻	7.0	55.0	41.7
Electrical conductivity	280.0	760.0	18.2
Mineralization	216.0	544.0	17.7
TDS	219.8	539.6	17.6
Q (l/s)	19.9	31.8	59.4

3.2.1. Correlation Analysis

The correlation matrix in Figure 6 reveals very strong to significant relationships between the physicochemical parameters of the waters analyzed, thus showing the correlations observed between Ca²⁺/HCO₃⁻ (0.64), Ca²⁺/Cond (0.59), Ca²⁺/Min (0.58), Ca²⁺/TDS (0.69), Ca²⁺/Q (0.71), K⁺/Cl⁻ (0.53), HCO₃⁻/Cond (0.69), HCO₃⁻/Min (0.69), HCO₃⁻/TDS (0.63), HCO₃⁻/Q (0.69), Cond/Min (1.00), Cond/TDS (0.67), Cond/Q (0.67), Min/TDS (0.66), Min/Q (0.67), and TDS/Q (0.97). These suggest that calcium and bicarbonates are the main culprits in the mineralization of the waters of this cave. In addition, the matrix highlights the role of the dissolution/precipitation and ion exchange processes with clay and sulfate minerals, reflected by moderate and non-significant correlations among other parameters.

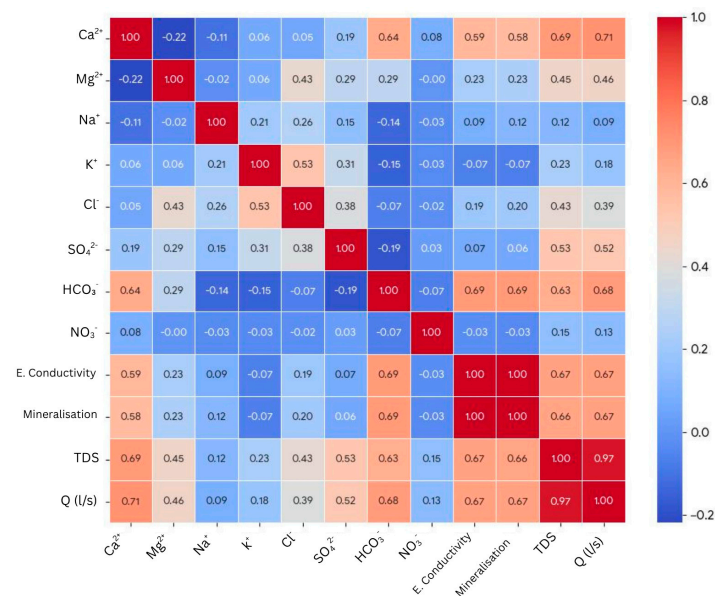


Figure 6. Correlation matrix of the different physicochemical parameters of the waters of the Ghar Boumaaza cave (2013–2023). Values in bold indicate correlation coefficients (Pearson (n)).

The correlation matrix provides a comprehensive view of the dominant geochemical processes and interactions influencing groundwater quality in the study area, with a focus on the complex interaction between the different chemical constituents.

3.2.2. Principal Component Analysis (PCA)

Table 2 of factors from the principal component analysis (PCA) applied to the groundwater of the Ghar Boumaaza karst shows that five factors explain 86.28% of the total variance of the hydrochemical data, the first factor (F1, 40.35%) being mainly associated with natural mineralization by carbonate dissolution (strong contributions of TDSs, flow, conductivity, mineralization, HCO_3^- , and Ca^{2+} ions), while the second factor (F2, 18.67%) reflects the influence of anthropogenic inputs, particularly agricultural, through high contributions of Cl^- , K^+ , and SO_4^{2-} , although the small share of NO_3^- invites this interpretation to be qualified; the following factors (F3 to F5) reflect mixed processes, such as dolomitization (Mg^{2+} , Ca^{2+} , and NO_3^- for F3) and occasional pollution inputs (Na^+ and NO_3^- for F4 and F5), which highlights that water quality is mainly controlled by the dissolution of carbonate rocks but remains significantly influenced by human activities, in particular agriculture and domestic discharges, thus highlighting the need for integrated management and enhanced monitoring to protect this vulnerable aquifer.

Table 2. Eigenvalues and contributions of variables (%) of the waters of the Ghar Boumaaza cave (2013–2023).

Eigenvalues:						
	F1	F2	F3	F4	F5	
Eigenvalue	4.84	2.24	1.19	1.14	0.94	
Variability (%)	40.35	18.67	9.95	9.52	7.79	
Cumulative %	40.35	59.02	68.97	78.49	86.28	
Variable contributions (%):						
	F1	F2	F3	F4	F5	
Ca^{2+}	10.95	4.17	22.51	1.13	5.82	
Mg^{2+}	3.64	5.71	36.66	19.66	0.69	
Na^+	0.2	6.81	0.05	33.47	35.02	
K^+	0.49	20.64	4.12	6.01	8.23	
Cl^-	3.26	22.67	2.21	0.91	0.01	
SO_4^{2+}	2.97	18.65	4.35	3.15	2.52	
HCO_3^-	11.45	8.59	2.35	0.14	0.96	
NO_3^-	0.07	0.04	21.97	23.57	41.49	
Cond	14.82	3.95	1.78	3.77	2.21	
Min	14.72	3.8	2.13	4.74	2.9	
TDS measured	18.56	1.34	1.08	1.38	0.01	
Q l/s	18.87	0.62	0.79	2.07	0.13	

3.2.3. Statistical Approach for the Prediction of Dissolved Solids

The curve representing the relationship between total dissolved solids (TDSs) and flow rate (Q), as shown in Figure 7, shows a generally positive trend. This suggests that as the flow increases, TDS concentrations also tend to increase. A simple linear regression model was applied to the data, and the resulting regression line indicates a moderate positive relationship. The coefficient of determination (R^2) is 0.141, which means that approximately 14.1% of the variability in TDSs can be explained by changes in flow. The corresponding Pearson correlation coefficient (r) is about 0.376, indicating a low to moderate positive linear correlation. However, the dispersion of data points and the relatively low R^2 value

suggest that the relationship is not strictly linear and that other factors may also influence TDS levels.

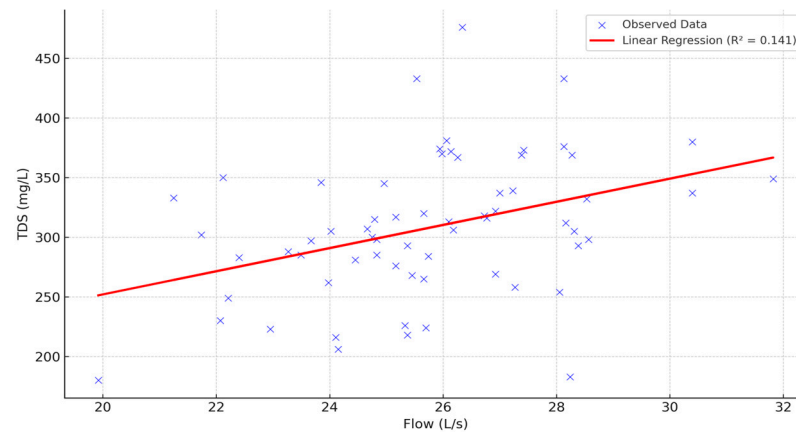


Figure 7. Linear regression between the flow (Q) and the total dissolved solids (TDSs) of the waters of the Ghar Boumaaza cave (2013–2023).

Due to the low linear correlation observed between TDSs and flow, we opted for the use of a polynomial regression model, described by the equation below (Equation (2)). Prior to conducting the analysis, outliers were eliminated using statistical methods, including the exclusion of data greater than 1.5 times the interquartile range. This choice is justified by the ability of polynomial regression to model nonlinear relationships between a dependent variable and one or more independent variables, as mentioned in Dargahi, et al. [34] in their study on models for predicting groundwater quality parameters using multiple linear regression (MLR): a case study from Kermanshah, Iran. Polynomial regression offers great flexibility to fit complex curves:

$$\text{TDS} = \beta_0 + \beta_1Q + \beta_2Q^2 + \dots + \beta_nQ^n + C \quad (2)$$

where TDS is the predicted TDS value and β_0 , β_n , and c are the coefficients to be determined by fitting the model.

Using the Python programming language (version 3.11.4), we fitted the polynomial regression model to the data, and the β_0 , β_n , and c values were determined.

Fitting the data of the function $f(Q) = \text{TDS}$ to the polynomial regression model shows us that the best fit for these data is a third-degree polynomial:

$$\text{Simulated TDS} = 234.4435 + (1.4133 \times Q) - (0.0354 \times Q^2) + (0.0002 \times Q^3) \quad (3)$$

This model offers a good fit for the data, as shown in the graph (Figure 8). The degree 3 polynomial regression curve fits well with the data points, indicating that a nonlinear relationship exists between flow and measured TDSs, with a correlation coefficient of $R^2 = 0.953$.

To validate our results, we compared the measured TDS values with the simulated TDSs using various regression models (Figure 9). Of these, the polynomial regression model showed the best performance, reaching a coefficient of determination (R^2) of 0.954. This high value confirms a high agreement between the simulated and observed data, thus strengthening the reliability and accuracy of the TDS prediction model at the Ghar Boumaaza site.

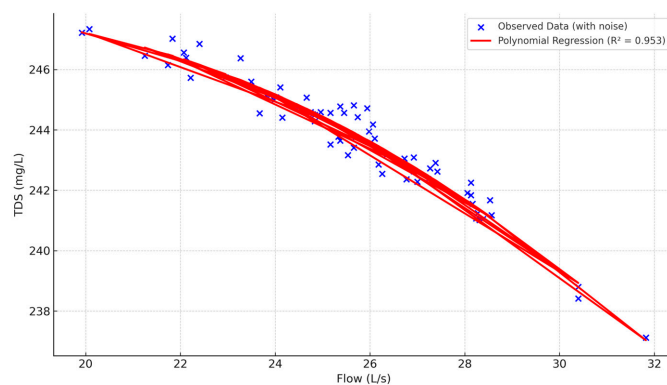


Figure 8. TDS polynomial regression model as a function of the flow (Q) of the waters of the Ghar Boumaaza cave (2013–2023).

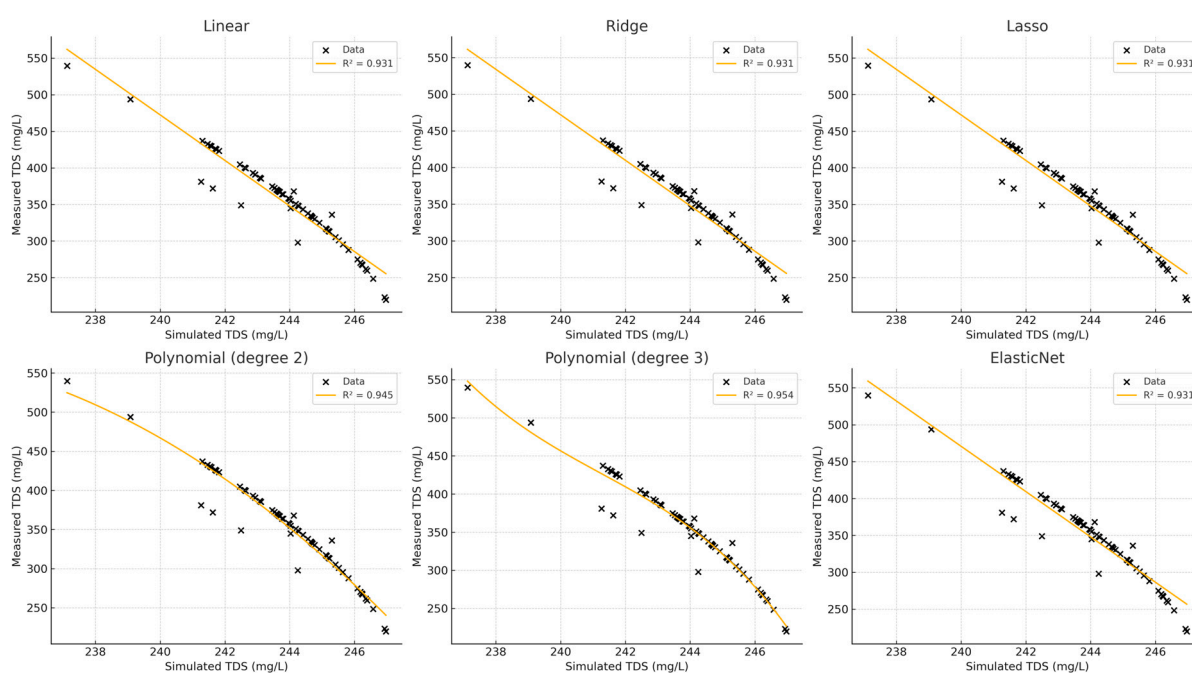


Figure 9. Comparison of regression models for the prediction of TDSs measured from simulated TDSs of the waters of the Ghar Boumaaza cave (2013–2023).

4. Discussion

The hydrochemical and statistical analysis carried out on the groundwater of the Ghar Boumaaza region revealed that 96.97% of the samples belong to the HCO_3^- - Ca^{2+} - Mg^{2+} facies, indicating a strong interaction with carbonate rocks. This is typical of recharge areas where rainfall solubilizes minerals, as found in other karst regions of the world, such as the Unica Springs Study in Slovenia [16], where groundwater chemistry has been strongly influenced by carbonate dissolution. The interaction between rainwater and carbonate rocks in these regions is a well-documented process that significantly shapes the hydrochemistry of karst aquifers [1]. However, the novelty of our study lies in the integration of multivariate statistical techniques, such as principal component analysis (PCA), to identify key influencing factors, which is a more advanced approach compared to previous Algerian studies that have not yet used these modern techniques. On the other hand, 3% of the samples are affected by inverse ion exchanges and exhibit a Cl^- - SO_4^{2-} - Ca^{2+} - Na^+ facies. This result aligns with the findings of Bensaoula and Collignon [19] in the same region of Tlemcen, where anthropogenic contamination has been identified as

a critical problem in the Ghar Boumaaza karst systems. The presence of anthropogenic contaminants in karst systems is well known, especially in Mediterranean karst aquifers, as highlighted by [20], which observed similar anthropogenic impacts on groundwater quality in the Tlemcen Mountains. However, our study goes further by specifically linking these contaminants to agricultural practices and overgrazing, with a focus on the need for sustainable land management practices to reduce pollution risks.

The statistical analysis of this study shows the variability of the chemical elements, with coefficients of variation greater than 50% for the majority of parameters, indicating a moderate heterogeneity influenced by geological and anthropogenic factors. This finding is consistent with the findings of Bouteraa, Mebarki, Bouaicha, Nouaceur and Laignel [26], which noted significant variability in groundwater composition due to both natural geological processes and human activities in the Boumerzoug-El Khroub Valley in northeastern Algeria. On the other hand, the study of Gao, Li, Liu, Sun, Zhao, Lv and Gang [9] in China focused more on assessing the risks of groundwater pollution, showing that geological features and human activities can both lead to significant changes in water quality, which confirms the conclusions of our study regarding the dual influence of natural and anthropogenic factors.

The principal component analysis (PCA) highlights that the first factor (F1, 40.35%) is mainly associated with natural mineralization by the dissolution of carbonates (strong contributions of TDSs, flow, conductivity, mineralization, HCO_3^- , and Ca^{2+} ions), while the second factor (F2, 18.67%) reflects the influence of anthropogenic inputs, particularly agricultural, through high contributions of Cl^- , K^+ , and SO_4^{2-} , although the small share of NO_3^- invites us to qualify this interpretation; the following factors (F3 to F5) reflect mixed processes, such as dolomitization (Mg^{2+} , Ca^{2+} , and NO_3^- for F3) and occasional pollution inputs (Na^+ and NO_3^- for F4 and F5), which highlights that water quality is mainly controlled by the dissolution of carbonate rocks but remains significantly influenced by human activities, in particular agriculture and domestic discharges, thus highlighting the need for integrated management and enhanced monitoring to protect this vulnerable aquifer. These findings are in perfect agreement with the results of research by Leins, Scheller, Çalli, Ravbar, Mayaud, Petrič, Liu and Hartmann [5], which highlighted the important role of agricultural activities in changing groundwater chemistry in karst systems. However, our study adds a new perspective by incorporating multivariate statistical models like PCA and regression analysis to assess how these human activities interact with natural processes to affect water quality. This is a methodological advance compared to previous Algerian studies, such as Bensaoula [23], which mainly used simple geochemical analyses without incorporating modern statistical methods to predict changes in water quality.

Finally, a polynomial regression approach was used to establish a relationship between flow and total dissolved solids (TDSs), revealing a nonlinear correlation with a good fit of the model ($R^2 = 0.9313$). This approach marks a significant advance in the modeling of the behavior of karst aquifers. In comparison, Hartmann, Goldscheider, Wagener, Lange and Weiler [8] used linear regression models in their study of karst aquifers in Europe, but the use of polynomial regression here allows for a more accurate representation of the complex and nonlinear relationship between flow and TDSs. This innovative approach improves our understanding of the dynamics between flows and groundwater mineralization in karst systems, providing a robust tool to predict changes in water quality and improve water resource management.

In conclusion, this study highlights the urgent need for sustainable management strategies that integrate hydrological, geological, ecological, and socio-economic factors to protect the Ghar Boumaaza karst system. The integration of advanced statistical techniques, the first of their kind in the region, provides a better understanding of the factors influ-

encing water quality and provides a reliable tool for future monitoring and management efforts. The use of polynomial regression to model TDSs in relation to flow represents a breakthrough in the management of karst aquifers, offering new insights into the long-term sustainability of water resources in semi-arid regions such as Algeria.

5. Conclusions

This study investigated the hydrochemical characteristics and predictive modeling of groundwater quality in the Ghar Boumaaza karst aquifer, a vulnerable system in the semi-arid Tlemcen region of Algeria. Groundwater samples revealed a dominant HCO_3^- - Ca^{2+} - Mg^{2+} facies (96.97%), indicative of a strong interaction with carbonate formations, while a minority (3%) showed signs of anthropogenic influence, notably from agricultural activities. These results underscore both the natural mineralization processes and emerging pollution risks in the aquifer.

Through the application of principal component analysis (PCA), the main hydrochemical processes were identified, distinguishing between carbonate dissolution and anthropogenic contributions such as nutrient leaching. Additionally, a third-degree polynomial regression model demonstrated a robust nonlinear relationship between discharge (Q) and total dissolved solids (TDSs), with an R^2 of 0.953, providing a powerful predictive tool for water quality under variable hydrological conditions.

Despite these insights, this study acknowledges certain limitations. In particular, this research did not incorporate isotopic analyses, which are essential for tracing recharge sources, identifying pollutant origins, and understanding water–rock interaction times. This omission limits the depth of interpretation concerning groundwater flow dynamics and contamination pathways.

Future work should therefore integrate isotopic tracers (e.g., $\delta^{18}\text{O}$, $\delta^2\text{H}$, $\delta^{15}\text{N}$, and tritium) to provide more direct and reliable insights into hydrological and anthropogenic processes affecting groundwater. Additionally, future research should adopt extended spatial and temporal monitoring frameworks—including key spring outlets such as Ain El Kelb and Ain Taga—and apply multivariate modeling approaches. Validation strategies such as k -fold or leave-one-out cross-validation (LOOCV) are recommended to enhance the robustness and generalizability of predictive models. The incorporation of advanced tools such as machine learning algorithms, hydrochemical mapping, and land use data will further strengthen the decision-making foundation for groundwater protection.

Ultimately, this research offers a transferable methodological framework for assessing water quality in karst aquifers under semi-arid climates. Its findings support the urgent need for sustainable groundwater governance that balances natural recharge dynamics with anthropogenic pressures. By linking science to policy, this study contributes to the development of locally adapted strategies for the protection and resilience of groundwater resources in North Africa and beyond.

Author Contributions: Conceptualization, S.G., A.B., and A.D.; methodology, S.G. and A.B.; software, S.G. and H.G.A.; validation, S.G., A.B., A.D., and A.L.; formal analysis, S.G., A.D., and M.A.-M.; investigation, S.G., A.D., and M.A.-M.; resources, S.G. and A.L.; data curation, S.G. and A.L.; writing—original draft preparation, S.G.; writing—review and editing, A.B., A.L., and H.A.; visualization, A.B. and H.G.A.; supervision, A.D. and A.L.; project administration, A.D. and A.L.; funding acquisition, H.A. All authors have read and agreed to the published version of the manuscript.

Funding: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R241). Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available upon request from the corresponding author [A.D.].

Acknowledgments: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R241). Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ford, D.; Williams, P.D. *Karst Hydrogeology and Geomorphology*; John Wiley & Sons: Hoboken, NJ, USA, 2007.
2. Goldscheider, N.; Haller, L.; Poté, J.; Wildi, W.; Zopfi, J. Characterizing water circulation and contaminant transport in Lake Geneva using bacteriophage tracer experiments and limnological methods. *Environ. Sci. Technol.* **2007**, *41*, 5252–5258. [[CrossRef](#)] [[PubMed](#)]
3. Frisia, S.; Borsato, A. *Karst. Dev. Sedimentol.* **2010**, *61*, 269–318.
4. Nanou, E.-A.; Zagana, E. Groundwater vulnerability to pollution map for karst aquifer protection (Ziria Karst System, Southern Greece). *Geosciences* **2018**, *8*, 125. [[CrossRef](#)]
5. Leins, T.; Scheller, M.; Çalli, K.Ö.; Ravbar, N.; Mayaud, C.; Petrič, M.; Liu, Y.; Hartmann, A. A new process-based approach for defining karst aquifer vulnerability to contamination risks under global changes. *Sci. Total Environ.* **2025**, *966*, 178561. [[CrossRef](#)] [[PubMed](#)]
6. Almaliki, A.H.; Derdour, A.; Ali, E. Air Quality Index (AQI) Prediction in Holy Makkah based on machine learning methods. *Sustainability* **2023**, *15*, 13168. [[CrossRef](#)]
7. Jeelani, G.; Shah, R.A.; Goldscheider, N. Karst geomorphology, cave development and hydrogeology in the Kashmir valley, Western Himalaya, India. *Acta Carsologica* **2018**, *47*, 5–21. [[CrossRef](#)]
8. Hartmann, A.; Goldscheider, N.; Wagener, T.; Lange, J.; Weiler, M. Karst water resources in a changing world: Review of hydrological modeling approaches. *Rev. Geophys.* **2014**, *52*, 218–242. [[CrossRef](#)]
9. Gao, S.; Li, C.; Liu, Y.; Sun, B.; Zhao, Z.; Lv, M.; Gang, S. Hydrogeochemical Characteristics and Evolution Processes of Karst Groundwater Affected by Multiple Influencing Factors in a Karst Spring Basin, Eastern China. *Water* **2023**, *15*, 3899. [[CrossRef](#)]
10. Hosni, A.; Derdour, A.; Nouri, T.; Moussaoui, T.; Zahi, F.; Reghais, A.; Jodar-Abellan, A.; Pardo, M.Á. Cultivating sustainability: A multi-assessment of groundwater quality and irrigation suitability in the arid agricultural district of Dzira (Ksour Mountains, Algeria). *Environ. Monit. Assess.* **2024**, *196*, 886. [[CrossRef](#)] [[PubMed](#)]
11. Stroj, A.; Briški, M.; Oštrič, M. Study of groundwater flow properties in a karst system by coupled analysis of diverse environmental tracers and discharge dynamics. *Water* **2020**, *12*, 2442. [[CrossRef](#)]
12. Lachache, S.; Derdour, A.; Maazouzi, I.; Amroune, A.; Guastaldi, E.; Merzougui, T. Statistical approach of groundwater quality assessment at Naama Region, South-West Algeria. *Larhyss J.* **2023**, *55*, 125–144.
13. Derdour, A.; Abellán, A.J.; Navarro, A.M.; Bailey, R.T. Assessment of Land Degradation and Droughts in an Arid Area Using Drought Indices, Modified Soil-Adjusted Vegetation Index and Landsat Remote Sensing Data. *Cuad. Investig. Geográfica Geogr. Res. Lett.* **2023**, *49*, 65–81. [[CrossRef](#)]
14. Adjovu, G.E.; Stephen, H.; James, D.; Ahmad, S. Measurement of total dissolved solids and total suspended solids in water systems: A review of the issues, conventional, and remote sensing techniques. *Remote Sens.* **2023**, *15*, 3534. [[CrossRef](#)]
15. Leyland, R.; Witthüser, K. VUKA: A modified cop vulnerability mapping method for karst terrains in South Africa. *Q. J. Eng. Geol. Hydrogeol.* **2010**, *43*, 107–116. [[CrossRef](#)]
16. Čuk Đurovič, M.; Petrič, M.; Jemcov, I.; Mulec, J.; Grudnik, Z.M.; Mayaud, C.; Blatnik, M.; Kogovšek, B.; Ravbar, N. Multivariate statistical analysis of hydrochemical and microbiological natural tracers as a tool for understanding karst hydrodynamics (the Unica Springs, SW Slovenia). *Water Resour. Res.* **2022**, *58*, e2021WR031831. [[CrossRef](#)]
17. Chundu, M.L.; Banda, K.; Sichingabula, H.M.; Nyambe, I.A. Integrated water quality assessment of open water bodies using empirical equations and remote sensing techniques in Bangweulu Wetland lakes, Zambia. *J. Great Lakes Res.* **2024**, *50*, 102451. [[CrossRef](#)]
18. Bensaoula, F.; Adjim, M.; Benazzouz, B.K.; Khatir, O.; Bouchama, M.-A.; Harkat, M.; Fellah, H.-S. Vulnérabilité et risque de pollution du système karstique de Ghar Boumaza (Monts de Tlemcen, Nord-Ouest algérien) Application de la méthode RISK. *Karstologia* **2017**, *70*, 53–62. [[CrossRef](#)]
19. Bensaoula, F.; Collignon, B. Vulnerability to Pollution of Karstic Aquifers in the Tafna River Basin and Risk Mitigation Strategies (Northwest Algeria). In *Groundwater in Arid and Semi-Arid Areas: Monitoring, Assessment, Modelling, and Management*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 55–80.
20. Azzaz, H.; Cherchali, M.; Meddi, M.; Houha, B.; Puig, J.M.; Achachi, A. The use of environmental isotopic and hydrochemical tracers to characterize the functioning of karst systems in the Tlemcen Mountains, northwest Algeria. *Hydrogeol. J.* **2008**, *16*, 531–546. [[CrossRef](#)]

21. Collignon, B. Hydrogéologie Appliquée des Aquifères Karstiques des Monts de Tlemcen (Algérie). Ph.D. Thesis, Université d'Avignon, Faculté des Sciences, Avignon, France, 1986. Published by the Atelier national de reproduction des thèses (ANRT), Grenoble, France.
22. Benest, M. Stratigraphie et sédimentologie du Jurassique supérieur des Monts de Tlemcen. *Dr. Diss. Univ. Lyon Fr.* **1984**, *2*, 540.
23. Bensaoula, F. Exportation des carbonates et dissolution spécifique dans le système karstique de Boumaza (Monts de Tlemcen-NO algérien). *Karstologia* **2008**, *52*, 31–38. [[CrossRef](#)]
24. APHA American Public Health Association; American Water Works Association; Water Environment Federation. *Standard Methods for the Examination of Water and Wastewater*; American Public Health Association: Washington, DC, USA, 2017; Volume 23.
25. Nordstrom, D.K.; Ball, J.W.; Donahoe, R.J.; Whittemore, D. Groundwater chemistry and water-rock interactions at Stripa. *Geochim. Cosmochim. Acta* **1989**, *53*, 1727–1740. [[CrossRef](#)]
26. Bouteraa, O.; Mebarki, A.; Bouaicha, F.; Nouaceur, Z.; Laignel, B. Groundwater quality assessment using multivariate analysis, geostatistical modeling, and water quality index (WQI): A case of study in the Boumerzoug-El Khroub valley of Northeast Algeria. *Acta Geochim.* **2019**, *38*, 796–814. [[CrossRef](#)]
27. Zhang, J.; Li, R.; Zhang, X.; Bai, Y.; Cao, P.; Hua, P. Vehicular contribution of PAHs in size dependent road dust: A source apportionment by PCA-MLR, PMF, and Unmix receptor models. *Sci. Total Environ.* **2019**, *649*, 1314–1322. [[CrossRef](#)] [[PubMed](#)]
28. Piper, A.M. A graphic procedure in the geochemical interpretation of water-analyses. *Eos Trans. Am. Geophys. Union* **1944**, *25*, 914–928.
29. Chadha, D. A proposed new diagram for geochemical classification of natural waters and interpretation of chemical data. *Hydrogeol. J.* **1999**, *7*, 431–439. [[CrossRef](#)]
30. Guan, Y.; Lu, H.; Dong, J.; Ge, Y.; Zhang, W.; Deng, Y. A novel groundwater pollution risk assessment method for producing-enterprises sites: Integrating scenario-specific pollution evaluation with Gaussian mixture model clustering. *Process Saf. Environ. Prot.* **2025**, *195*, 106760. [[CrossRef](#)]
31. Li, Q.; Li, P.; Elumalai, V. Identification and apportionment of groundwater pollution sources in the Guanzhong region based on PMF model. *Hum. Ecol. Risk Assess. Int. J.* **2025**, *31*, 30–52. [[CrossRef](#)]
32. Ruan, D.; Bian, J.; Wang, Y.; Wu, J.; Gu, Z. Identification of groundwater pollution sources and health risk assessment in the Songnen Plain based on PCA-APCS-MLR and trapezoidal fuzzy number-Monte Carlo stochastic simulation model. *J. Hydrol.* **2024**, *632*, 130897. [[CrossRef](#)]
33. Chen, K.; Liu, Q.; Jiang, Q.; Hou, X.; Gao, W. Source apportionment of surface water pollution in North Anhui Plain, Eastern China, using APCS-MLR model combined with GIS approach and socioeconomic parameters. *Ecol. Indic.* **2022**, *143*, 109324. [[CrossRef](#)]
34. Dargahi, P.; Nasserli, S.; Hadi, M.; Nodehi, R.N.; Mahvi, A.H. Prediction models for groundwater quality parameters using a multiple linear regression (MLR): A case study of Kermanshah, Iran. *J. Environ. Health Sci. Eng.* **2023**, *21*, 63–71. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.