

Research

Modeling and characterization of shallow aquifer water based on ion concentrations to salinity variations using multivariate statistical approach

Riaz Hossain Khan¹

Received: 8 September 2023 / Accepted: 21 January 2024

Published online: 29 January 2024

© The Author(s) 2024 [OPEN](#)

Abstract

Water chemistry data (1965–2015) from the Chicot and Evangeline aquifers in Nueces and Kleberg counties of Texas were analyzed to assess the hydrogeochemical processes based on variations in ion concentrations affecting groundwater salinity levels in those coastal aquifers. A further aim of this study was to identify a proxy using major-ion concentrations to predict groundwater salinity. Correlation analysis suggests that the hydrogeochemical processes operating in these aquifers differ significantly. In principal component analyses, the first three principal components explain 91% and 94.8% of the total variabilities of the variation of groundwater chemistry in the Evangeline and Chicot aquifers, respectively. The rotated factors (factors 1, 2, and 3) using the varimax rotation method implied that depth control and salinity variations predominantly cause the highest variabilities in water chemistry in the Chicot and Evangeline aquifers in the study area. Results showed depth control on water quality parameters was more pronounced in the Evangeline aquifer compared to the Chicot aquifer. Stepwise elimination of the least significant predictors to identify the proxy for groundwater salinity revealed chloride (Cl) could be the most significant predictor to estimate groundwater salinity variations in both aquifers. However, regression models generated from 75% of the training datasets predicted total dissolved solid (TDS) variations with 78% and 43% accuracies in Chicot and Evangeline aquifers, indicating that Cl can be considered the proxy for the Chicot aquifer only but not suitable for the Evangeline aquifer in the study area.

Keywords Aquifer · Groundwater chemistry · Regression · Salinity intrusion

1 Introduction

Groundwater, one of the most important drinking water resources worldwide, is threatened by contamination from different natural and anthropogenic activities. For instance, saltwater intrusion, the movement of saline water into freshwater aquifers, can contaminate and render groundwater resources unsuitable for drinking, agriculture, and industrial uses and also can negatively affect aquatic biota and organisms [1]. Irreversible ecological changes and economic losses are the far-reaching consequences of the increased salinities in many coastal environments. This type of contamination can be the result of a combination of different factors acting simultaneously or individually, such as groundwater overdraft (i.e., withdrawals exceeding the available resources or safe yield), disposal of oil field

✉ Riaz Hossain Khan, khanriazhossain@yahoo.com; riaz.khan@bracu.ac.bd | ¹BRAC James P Grant School of Public Health, BRAC University, Bir Uttom A K Khandakar Road, Dhaka 1213, Bangladesh.



brines and/or surface infiltration from storm surge events or leaching of salts from soils (i.e., in semi-arid and arid environments) [2].

The study area, which was the focus of this investigation, lies in Nueces County and a small portion of Kleberg County and extends from 27.38° to 27.98° N latitude and from – 97.20° to – 97.94° E longitude (Fig. 1a).

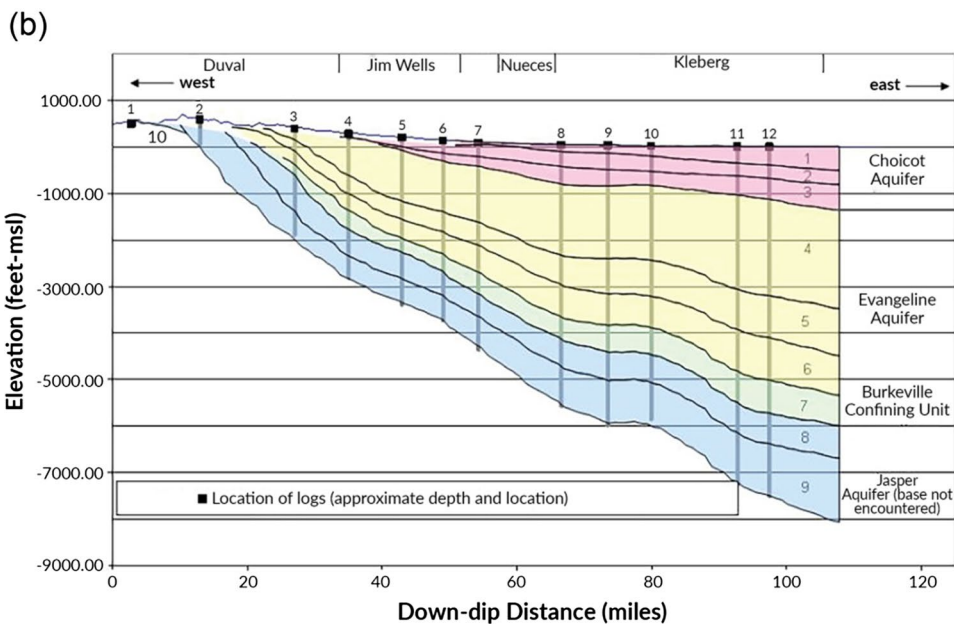
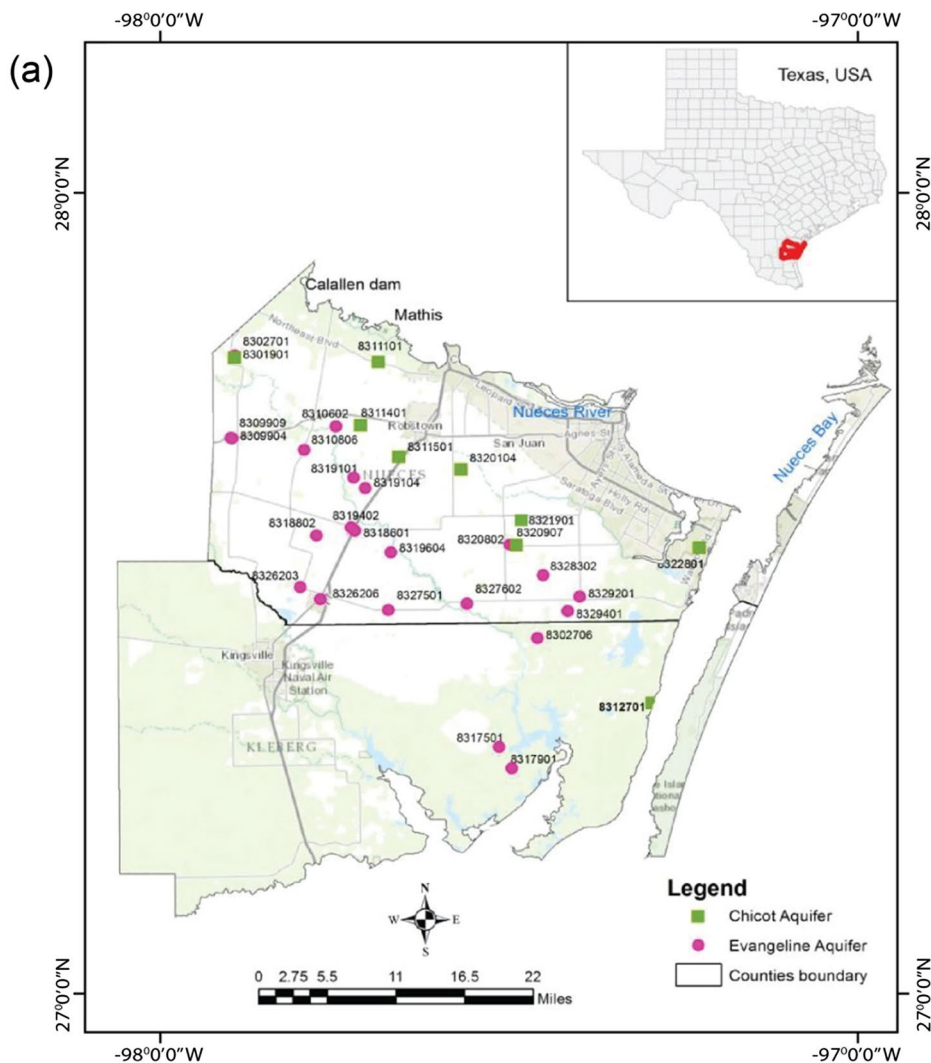
The geographic location of the study area, including its proximity to the Gulf of Mexico, significantly influences its climate [3]. Short and longer-term climate events such as El Niño/Southern Oscillation (ENSO) change precipitation patterns and exacerbate drought conditions in Texas [3]. Under the current climate change conditions, coastal water resources will be at risk due to the combined effects of saltwater intrusion due to the fall of groundwater levels, the contemporary rise of sea levels, and predicted declining rainfall rates in the region [4]. Freshwater discharge to the sea occurs mainly from coastal riverine sources and submarine groundwater discharges (SGD). For instance, in Coastal Florida, southerly winds are more pervasive, and coastal waters are observed to advance inland (i.e., bay waters) during the dry season months when riverine discharge rates are generally extremely low [5].

Cenozoic sediments that underlie the coastal plain of Texas are thousands of meters thick near the coast [6]. The shallowest coastal aquifers are in close contact with surface water drainage features. Haley et al. [7] reported that exploration of petroleum and groundwater coupled with the vertical movements along growth faults might increase land subsidence in the Texas Coastal Bend areas. Therefore, the depth and thickness of the hydrostratigraphic units vary within short distances in the study area. Regionally, the major aquifer system of interest consists of the following stratigraphic units of Cenozoic age in ascending stratigraphic order: (a) Catahoula confining system (restricted), (b) Jasper aquifer, (c) Burkeville confining system, (d) Evangeline aquifer, and (e) Chicot aquifer (Fig. 1b). The Gulf Coast aquifer in the study area consists of interbedded sand, silt and clay materials [8]. Quartz arenite is the predominant sandstone category in the northern part, whereas arkosic sandstone and greywacke sandstone, with a significant proportion of feldspar and rock fragments, are present near the southern part of the Central Gulf Coast Aquifer [9]. Large amounts of orthoclase, plagioclase, and volcanic rock fragments were reported in the Goliad Formation of the Evangeline aquifer towards the south adjacent to the San Patricio-Refugio County line [10]. In the Nueces and San Patricio counties and their close vicinities, the sand fractions of the Chicot and Evangeline aquifers can be as high as 60% [11]. Earlier studies reported that the cation exchange capacity is significantly high in fine-grained materials such as clay and organic matter [12]. The surface soils of Nueces and San Patricio Counties are mostly Clay-rich sediments [11]. Therefore, the groundwater chemistry in the study area is partially controlled by the soil materials near the outcrop areas. Soil RF [13] reported three small salt diapirs within the Miocene-Pliocene sediments in adjoining areas at a depth of 850 and 1000 feet (ft) below the land surface, which are overlain by gypsum-anhydrite, sulfur, and limestone caprock. The limestone layer grades downward into a thin anhydrite deposit. The Gulf Coast hydrostratigraphic units show significant spatial variation in thickness and depth. These disparities can cause hydrogeochemical dissimilarities between the aquifers or even within the same aquifer system. Therefore, factors that may directly and/or indirectly control groundwater composition include aquifer mineralogy and its capacity to buffer ion concentration changes, aquifer hydraulic characteristics, rates and types of chemical reactions, mixing of water from different sources, groundwater residence time, retention of residual connate water, climate conditions and so on [8, 9, 12].

Low permeability zones, which contain naturally high solute concentrations, are in the aquifers toward the southern portion of the south Texas region. High evaporation and low precipitation rates are also responsible for higher dissolved solid concentrations in surface and groundwater in those semi-arid regions [14]. Consequently, abundant good-quality water is not readily available throughout the South Texas coastal area. Previous studies conducted by the U.S. Geological Survey (USGS) and others have indicated a significant increase in the concentration of dissolved minerals in the lower Nueces River between Mathis and the Calallen Saltwater Barrier Dam. The purpose of constructing the saltwater barrier dam was to prevent saltwater intrusion from Nueces Bay into the Nueces River through the surface water, for example, during tropical storm surge. Furthermore, high rates of saline submarine groundwater discharge (SGD) have been reported for Nueces Bay [15]. Various natural and anthropogenic factors simultaneously contribute to high salinity levels in Texas aquifers, common in many semi-arid regions [14]. In such areas, irrigation and intermittent increased precipitation events enhance the leaching of accumulated solutes from soils and shallower groundwater to the deeper aquifers and the surface water [14, 16–18].

Therefore, the main objective of this research was to assess how the hydrogeochemical processes operating in these two aquifers differ based on variations in ion concentrations, ultimately affecting groundwater salinity in the Chicot and Evangeline aquifers in the study area. This provided valuable information to characterize individual aquifer systems in the subsurface environment. A further aim of this study was to develop a proxy method using major-ion concentrations

Fig. 1 **a** Map of the water sampling locations (well IDs shown) in the study area, and **b** vertical litho-stratigraphic cross-section perpendicular to the Gulf Coast [3]



and water type to predict total dissolved solids (TDS), considering an indicator of groundwater salinity using regression analyses.

Times series analysis, multivariate statistics, artificial intelligence techniques, and other statistical methods have been applied for environmental and biological investigations such as hydrology/climatology, biological, bio-medical, and genetic studies, and coastal vulnerability to sea level rise predictions [19–26]. Multivariate statistical analyses were applied in the past to delineate spatial- and depth-variation of water chemistry in the central Canterbury plains [27]. Multiple Regression (MR) analysis was used to predict groundwater level changes with satisfactory results [28]. Missing water chemistry data values often make the modeling tasks unreliable, mainly when a data set is composed of many variables and seemingly small percentages of missing data for different sampling events/observations for different variables cumulatively affect the data set. This may require discarding a large amount of available data from the datasets [29]. Several well-established techniques to address this exist, such as case-wise deletion, standard iterative and non-iterative missing data imputation methods, and modern machine learning techniques, are frequently used. Suitable techniques are chosen depending on the degree of linearity among paired variables, normality of datasets, and percentage of missing data, targeting to reduce the mean square error (MSE) [29].

2 Data and methods

Annual dry period water chemistry data were collected from the Texas Water Development Board (TWDB) website, available at <http://www.twdb.texas.gov/groundwater/data/gwdbbrpt.asp>, and consisted of 121 individual measurements collected between 1965 and 2015 from the Chicot aquifer at depths 35–491 feet, and in Evangeline aquifer at 181–1258 feet. Annual water chemistry data such as major ions, including dissolved silica, and physicochemical parameters were considered for statistical analyses for nine wells from the shallowest Chicot aquifer and twenty-one from the Evangeline aquifer, directly overlain by the Chicot aquifer. In the study area, more frequent and severe droughts were reported after 1977, which affected more than 25% of the region noticeably in 1979, 2006, 2009, and 2011 [30]. This, along with a change in irrigation pattern and recent high groundwater pumping rates, impacted the groundwater level and altered the local groundwater flow direction. Therefore, considering the annual fluctuation of groundwater level in the time series plot, water chemistry data of the Chicot and Evangeline aquifers were split into two sets: pre-1977 and 1977–2015. Four representative samples were chosen for each individual well, two of which were taken from individual sub-groups (pre and post-1977) to keep the dataset balanced to minimize bias in the analysis. This study did not consider wells with excessive missing data (i.e., > 40%) for the water chemistry parameters. Groundwater physicochemical parameters considered for this study were pH, Electric conductivity (EC), and Total Dissolved Solids (TDS). Statistical analyses were made on most of the major ions, including Calcium (Ca), Magnesium (Mg), Sodium (Na), Potassium (K), Bicarbonate (HCO_3), Sulfate (SO_4), Chloride (Cl), Fluoride (F), and Nitrate (NO_3). Missing values for F and NO_3 (less than 10% of the total dataset) were imputed using the linear interpolation technique.

The most efficient way to measure salinity and its units has been a matter of debate in the past few decades. TDS can be used as a salinity indicator but needs to be measured very cautiously to achieve maximum accuracy and precision with highly sensitive instruments. Alternatively, it has been suggested that the electric conductivity measurements be utilized and multiplied by the appropriate correction factor to obtain the required TDS [31]. There is unreliability about the complete accuracy of all available EC measurements in the TWDB database (Source: <http://www.twdb.texas.gov/groundwater/faq/faqgwdb.asp>). However, Pearson correlations showed that the available EC dataset perfectly correlated with the TDS values; hence, EC was excluded from the subsequent statistical analyses.

Correlation, principal component analyses (PCA), and regression analysis for Chicot and Evangeline aquifer samples were conducted separately using 37 and 84 groundwater samples in each subset. No significant temporal differences were noticed in water chemistry for most of the wells in the individual aquifers. The Pearson correlation aims to quantify the strength of the linear relationship between the paired variables. Pearson correlations were used to determine the major ions, which strongly correlated with the observed variation of TDS and among individual ion pairs. Therefore, it provided information to assess the geochemical environment and possible sources of groundwater contamination based on the magnitude and direction of the correlation coefficients. The PCA is an unsupervised method that does not consider any dependent variables. Good correlation strengths among the variables allowed conducting the PCA to assess the major components contributing to the highest variabilities within the dataset. Individual PCAs were run for the available measured variables for the Evangeline and Chicot aquifer. Rotated factors were used to visualize and comprehend the first three principal components using the varimax rotation method. Results showed that the predominant principal

components indicate subsurface stratigraphy where depth relation can be explained for different ion concentrations. The other major principal component indicates the variation of salinity within the groundwater system. PCA results suggest that the depth control and salinity variations predominantly cause the highest variabilities within the datasets for the Chicot and Evangeline aquifers in the study area.

Therefore, in the subsequent stages, regression analysis was conducted to generate models to predict the depth control on aquifer hydrochemistry considering depth as the dependent variable and then the influences of the ion concentrations on the variations in groundwater salinity (where salinity is the dependent variable) within the individual aquifers. Finally, the regression analyses were conducted to develop a proxy method using major-ion concentrations to predict groundwater salinity in the Chicot and Evangeline aquifers through a stepwise process of elimination using the least number of variables within the datasets.

The depth control in aquifer hydrochemistry was assessed using Tukey lines, taking both aquifer and depth as classes. The preliminary multilinear regressions were run for the following variables to model their variations with depth.

$$\text{Model: } TDS, pH, Na, K, Ca, Mg, Cl, F, SO_4, HCO_3, NO_3, Silica = f(\text{depth}) \quad (1)$$

Taking the TDS (salinity indicator) values as a dependent variable, forward and backward multiple regression models were run to determine the relative influences of major ions on the observed TDS values.

$$\text{Model: } TDS = f(Na, Ca, Mg, K, Cl, F, SO_4, HCO_3, NO_3, Silica) \quad (2)$$

Individual data values were treated as outliers if those were three standard deviations above or below the mean values. Variables that showed significant outliers were log-transformed to prepare for the multiple regression models. Minimal and full models were run to identify the significant variables for predictions.

Regression models were calibrated between TDS and one of the most significant multiple regression model-derived predictors by separating 75% of the dataset in a training set and the remaining 25% in a testing set to assess the performance of the models. At first, point estimations were conducted to predict each sample using the following equation,

$$Y = \beta_0 + \beta_1 X \quad (3)$$

where Y is the target variable, X is the predictor variable, β_0 is the intercept parameter, and β_1 is the slope parameter.

Later, to estimate the precision of each prediction, the 95% confidence interval (significance level, $\alpha = 0.025$) was used to derive the upper and lower bounds for every model prediction of the testing set. If the quantile–quantile ($Q-Q$) plot and Shapiro Wilk normality tests did not support the normality of the data distribution for the predictor in the aquifers in both training and testing sets, bootstrapping was used by resampling 5000 times with a 95% confidence interval. Hypothesis tests were carried out to see the differences in mean between those two training and testing datasets. Finally, the accuracy estimation for the model predictions provided clues to assess the suitability of considering the predictor(s) as a proxy for salinity estimations in the Chicot and Evangeline aquifers.

3 Results and discussion

In the Chicot and Evangeline aquifer, post-1977 samples showed significant spatial variations of major ion concentrations based on the Tukey lines and P values of statistical significance ($p < 0.05$). However, individual wells showed no noticeable temporal variations between the pre and post-1977 samples. The distribution of the mean values of the major ion concentrations and associated parameters are shown in Figs. 2 and 3 below.

3.1 Pearson correlations among major ions and TDS

Correlation coefficients of the water chemistry parameters for the Chicot and the Evangeline aquifers are shown at 0.05 significance level (Fig. 4a and b). Groundwater samples from both aquifers showed a strong positive correlation of TDS with Na , SO_4 , and Cl ions. However, an inverse relation of TDS was observed with the HCO_3 and silica concentrations in the Chicot aquifer. Moreover, a relatively weak correlation of Ca and Mg with the groundwater TDS indicates that the ion exchange process replaces Ca , resulting in an immobilization of Ca ions and a concomitant increase of Na concentrations in the Chicot aquifer. These results conform to the findings by Chowdhury et al. [9]. Uddameri et al. [32] stated that the source of calcium in the Chicot aquifer is partly from the overlying caliche deposits. A moderately strong negative

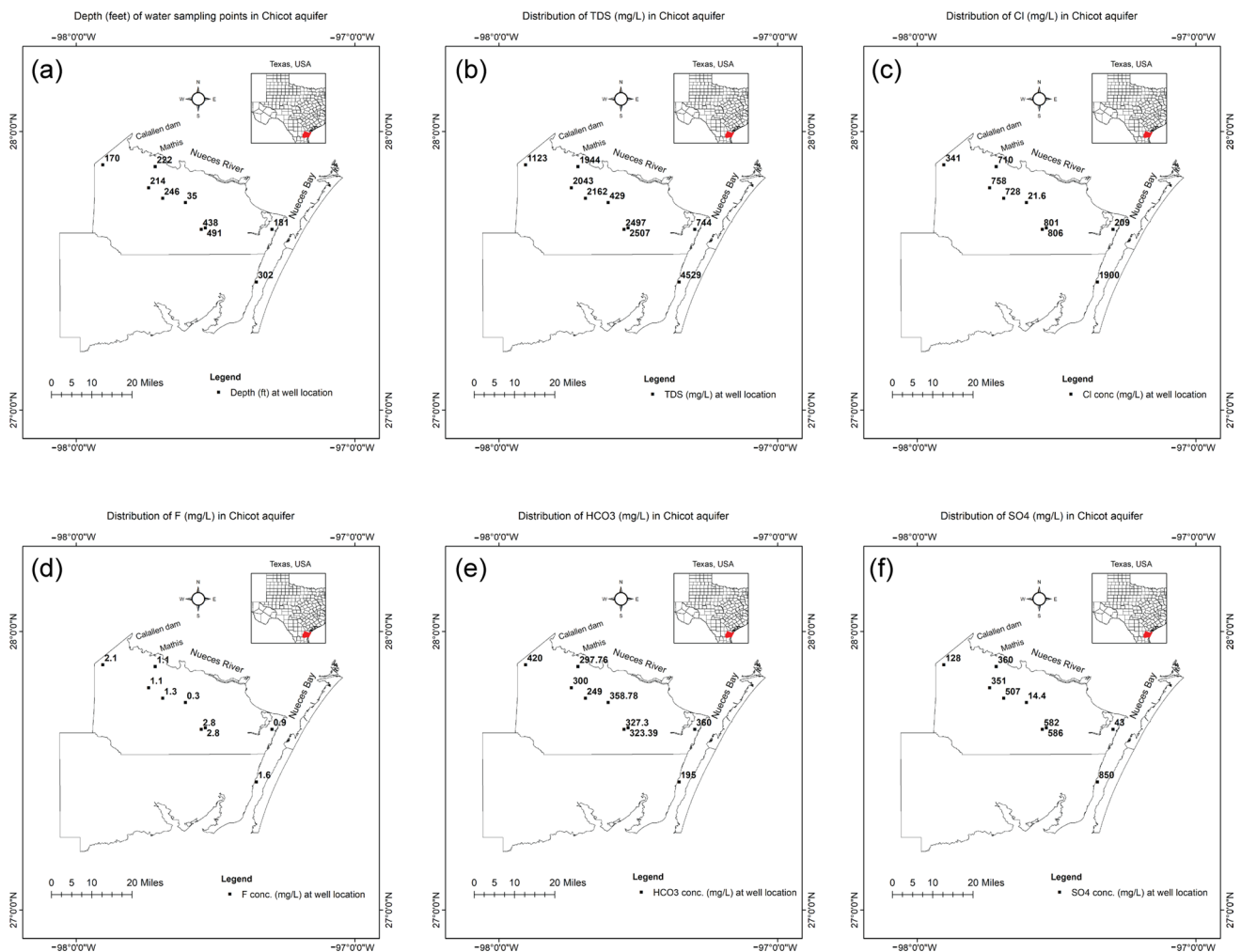


Fig. 2 Spatial distribution of groundwater sampling depth and water chemistry parameters (mean values) in the post-1977 dataset in the Chicot aquifer

correlation between TDS and HCO₃ in the Chicot aquifer might result from secondary carbonate precipitation, which was not noticed in the Evangeline aquifer.

The correlation strength between SO₄ and Cl is significantly higher in the Chicot aquifer (0.85) than in the Evangeline aquifer (0.47). A strong correlation of SO₄ and Cl with TDS in the aquifers indicates the major contributors of groundwater salinity in the study area, consistent with the study findings in the arid Maadher region of Hodna, northern Algeria [33]. In most groundwater systems where Cl is the dominant anion, Na is the predominant cation [34]. The strong correlation between sulfate and chloride might result predominantly from saltwater intrusion, high evaporation, anthropogenic contributions, and sulfate mineral dissolution, consistent with the results from the Suruliyar sub-basin in Tamil Nadu, India [35].

In the Chicot aquifer, groundwater pH showed a weak to moderate inverse relation with dissolved silica and Ca concentrations. However, it was not noticed in the Evangeline aquifer. Although Ca and Mg showed a strong positive correlation, no correlation was observed with the corresponding anion, HCO₃, in the Chicot aquifer. In contrast, Ca and Mg showed a relatively strong positive correlation with HCO₃ and Cl ions in the Evangeline aquifer. The aquifers in the study area are siliciclastic in nature with high sand fractions. Therefore, the possible source of HCO₃ in the groundwater system is the degradation of sedimentary organic matter within the subsurface environment and partially by the chemical reaction of CO₂ and precipitated water near the outcrops [36, 37].

A weak positive correlation was noticed between pH and Na with F ions in both aquifers. Besides, a moderately strong negative correlation (−0.66) was observed between F and Ca in the Chicot aquifer, providing evidence of the minor dissolution of fluoride-containing minerals under a slightly alkaline environment. This result is consistent with a study

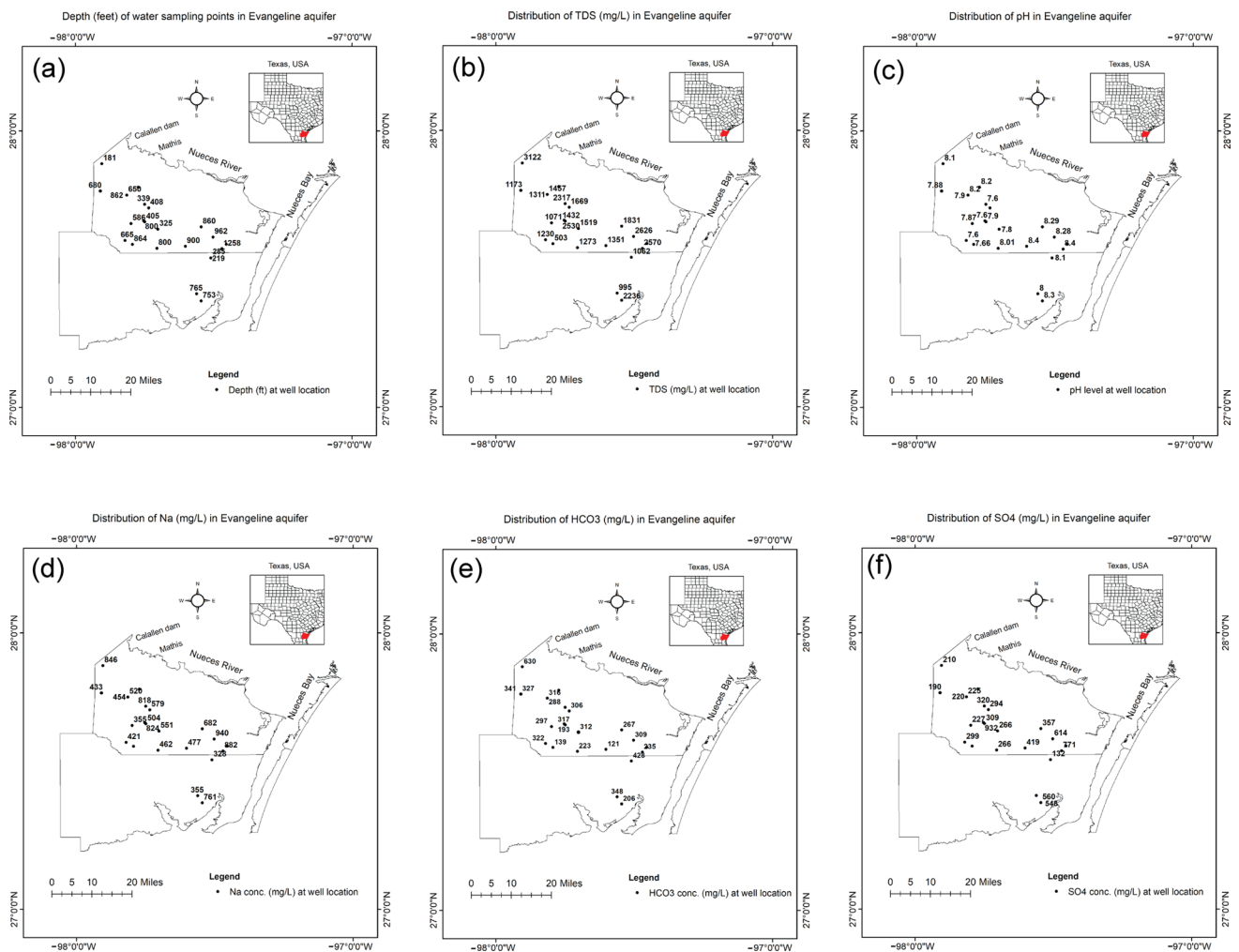


Fig. 3 Spatial distribution of groundwater sampling depth and water chemistry parameters (mean values) in the post-1977 dataset in the Evangeline aquifer

by Uddameri et al. [32]. Like the Chicot aquifer, no significant correlations are noticed for the Ca-F and Mg-F pairs in the Evangeline except HCO₃-F, which shows a weak positive correlation. Some physicochemical properties, such as high pH, Na, K, and Cl concentrations, relatively low total dissolved silica, and Na/Cl molar ratios ≥ 0.84 in four Chicot aquifer samples, indicate the possibility of seawater intrusion. Baker [38] proposed that growth faults have become barriers to fluid flow or conduits for cross-formational flow in some adjoining areas. The mass concentration of Cl/SO₄ and Cl/Na indicates the possibility of mixing water through a cross-formation flow via faults/fractures.

Overall, the unique differences in the relationships among TDS, Na, SO₄, and Cl ions with HCO₃ and dissolved silica concentrations, pH with silica and Ca ions, and bivalent Ca and Mg with HCO₃, Cl, and F ions indicate the geochemical processes at Chicot and Evangeline aquifers are not similar.

3.2 Principal component analyses and rotated factor pattern

PCA was run to reduce the dimensionality of the dataset containing the variables being considered and to extract the common underlying factors contributing to the highest variability within the datasets. The eigenvalues for the correlation matrix were calculated to total 15, of which the first three principal components explained about 91% and 94.8% of the total variabilities within the dataset in the Evangeline and Chicot aquifer, respectively.

In factor 1 versus factor 3 in the Evangeline aquifer, strong positive loadings were noticed for TDS with Na, SO₄, and Cl ions, indicating similarity in attributes of those variables and effects of non-carbonate minerals in the groundwater environment (Fig. 5a). This result is consistent with the correlation analysis results in the previous section. The first

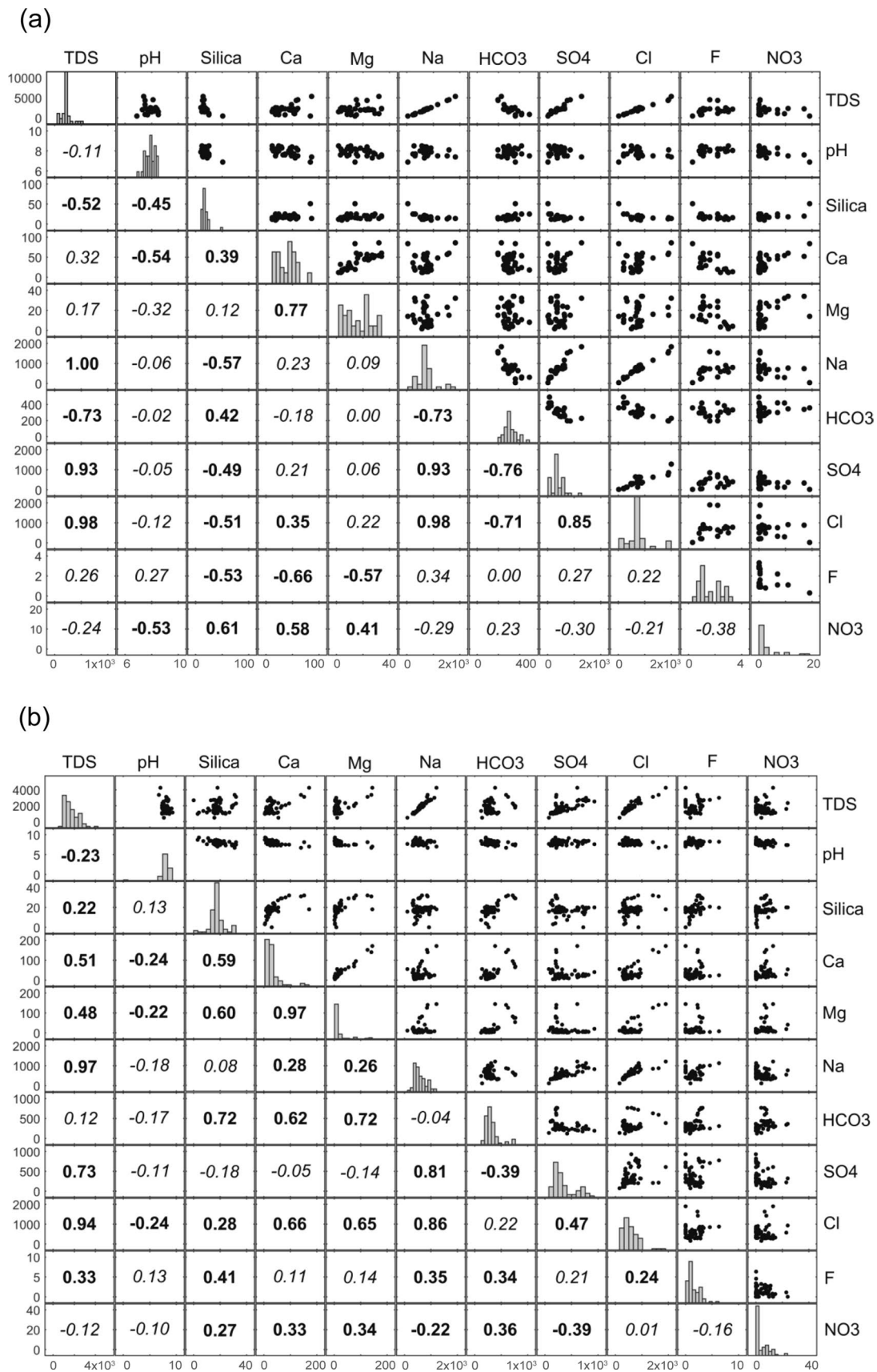


Fig. 4 Pearson Correlation analyses of major ions and physicochemical parameters in the **a** Chicot and **b** Evangeline Aquifers. The dissolved ion concentrations are in milligrams per liter (mg/L), TDS in mg/L, and pH in pH unit

PC suggests that high values of TDS are affected mainly by the given major ions in the Evangeline aquifer. The strong negative association of bicarbonate with the groundwater pH indicates the neutralizing effects of the newly recharged rainwater containing a relatively low pH level.

Almost identical and high loadings for the Na, Cl, and SO_4 in the Chicot aquifer imply common and significant sources of water salinization at a relatively higher depth close to the Nueces Bay area (Fig. 5c). These relations also indicate predominantly non-carbonate hardness in the Chicot aquifer. Chaudhuri and Ale [39] reported that deep regional groundwater circulation resulted in a minor dissolution of fluoride minerals (such as muscovite, biotite, and fluorapatite), ultimately releasing F ions in the deeper part of the reservoir. A slight deviation and relatively weak loadings for F suggest that it contributes less to the groundwater salinity in that aquifer. In contrast, strong negative loadings were noticed for HCO_3 , but relatively weak loadings were seen for NO_3 and pH levels that were not observed in the Evangeline water samples. The strong positive loadings for the Ca and Mg in the Chicot aquifer indicate a possibility of weathering of the carbonate minerals. Multiple sources cause the spatial and depth variations of SO_4 and Cl in the study area, including anthropogenic sources such as agricultural wastewater and industrial effluents and natural sources, including saltwater intrusion and release of SO_4 from oxidation of gypsum-anhydrite minerals in the reservoirs. Along with some of these common sources, the dissolution of albite and some mica causes an increase in Na concentration in the aquifers. Chowdhury et al. [9] reported evidence of recent salinity intrusion based on analyzing distinctive chemical fingerprints and high Br concentrations in several wells in the Nueces and Kleberg counties, consistent with these results.

Factor 1 versus factor 2 and factor 1 versus factor 3 indicate depth relationships in the Evangeline and Chicot aquifer, respectively (Fig. 5b and d). A moderately strong positive loading was noticed for SO_4 in Evangeline but for F concentrations in Chicot aquifer samples. The inverse loadings for NO_3 with water table depth in both aquifers indicate anthropogenic sources of pollution (e.g., agricultural practice and farming) at the recharge areas. In both aquifers, HCO_3 , Ca, Mg, and NO_3 showed negative loadings as a function of progressive depths. The higher concentrations were noticed near the Mathis and Calallen dams towards the north at the shallowest depths in both aquifers. High silica at shallower depth in the Chicot aquifer indicates the dissolution of aluminosilicate minerals (e.g., feldspars and mica) with recent infiltration of rainwater at the shallower portion of the aquifer. Previous studies showed that fluctuations in groundwater level are relatively higher in the Evangeline aquifer (standard deviation 9–16 feet) compared to the Chicot aquifer (standard deviation 2–6 feet) in the study area [37]. The degrees of groundwater level fluctuations at those two aquifers also contribute to the water chemistry variations in the study area (Fig. 5).

Based on these observations, it is apparent that the depth relations to major ion chemistry and major ion contributions to the variations of groundwater salinity level differ noticeably between the Chicot and Evangeline aquifers. Therefore, the PCA results imply that the geochemical environments vary, and the sources of high groundwater salinity differ in these aquifers to some extent. Considering the above, the depth control on groundwater chemistry and its relation to salinity are discussed in detail in the subsequent sections.

3.3 Depth models using linear regression

A high sedimentation rate and the presence of growth faults are responsible for the significant variation of depth and thickness of the major hydro-stratigraphic units that influence the hydrogeochemical characteristics of different aquifers or even within the same aquifer system in the study area [7, 9, 32, 37, 38].

Significant depth relations were observed for the groundwater TDS, Ca, Na, HCO_3 , SO_4 , NO_3 , and F in Chicot and Evangeline aquifers using all available wells in the study area. Unlike the Chicot aquifer, depth variations were also noticed for pH, Mg, and Cl ions in the Evangeline aquifer. Table 1 shows the depth differences in the water chemistry, based on the Tukey lines and p values < 0.05) for the Chicot and Evangeline aquifer water samples. The results showed that depth control for most ions is more prominent in the Evangeline aquifer than in the Chicot aquifer samples.

The depth relation of the concentrations of major ions and TDS in Evangeline aquifer water samples for the available wells at specified depth ranges are shown in Fig. 6 below. Relatively higher fluctuations of groundwater levels were noticed in several wells within high pumping areas and are located close to the Nueces Bay or near the surface outcrops of the aquifer. It affected groundwater chemistry, including TDS, by leaching accumulated solutes from soils, changing the groundwater redox environment, and rock-water chemical reactions within the same depth intervals. Therefore, excluding the groundwater well measurements containing significantly higher fluctuations, the relevant best-fitted depth models for the Evangeline aquifer samples are shown in Fig. 7 below.

The model used to predict the variations of water chemistry in Chicot and Evangeline aquifers as a function of depth is provided in Eq. 1 in the method section. The depth models for TDS, Na, K, HCO_3 , NO_3 , and SO_4 in Evangeline aquifer

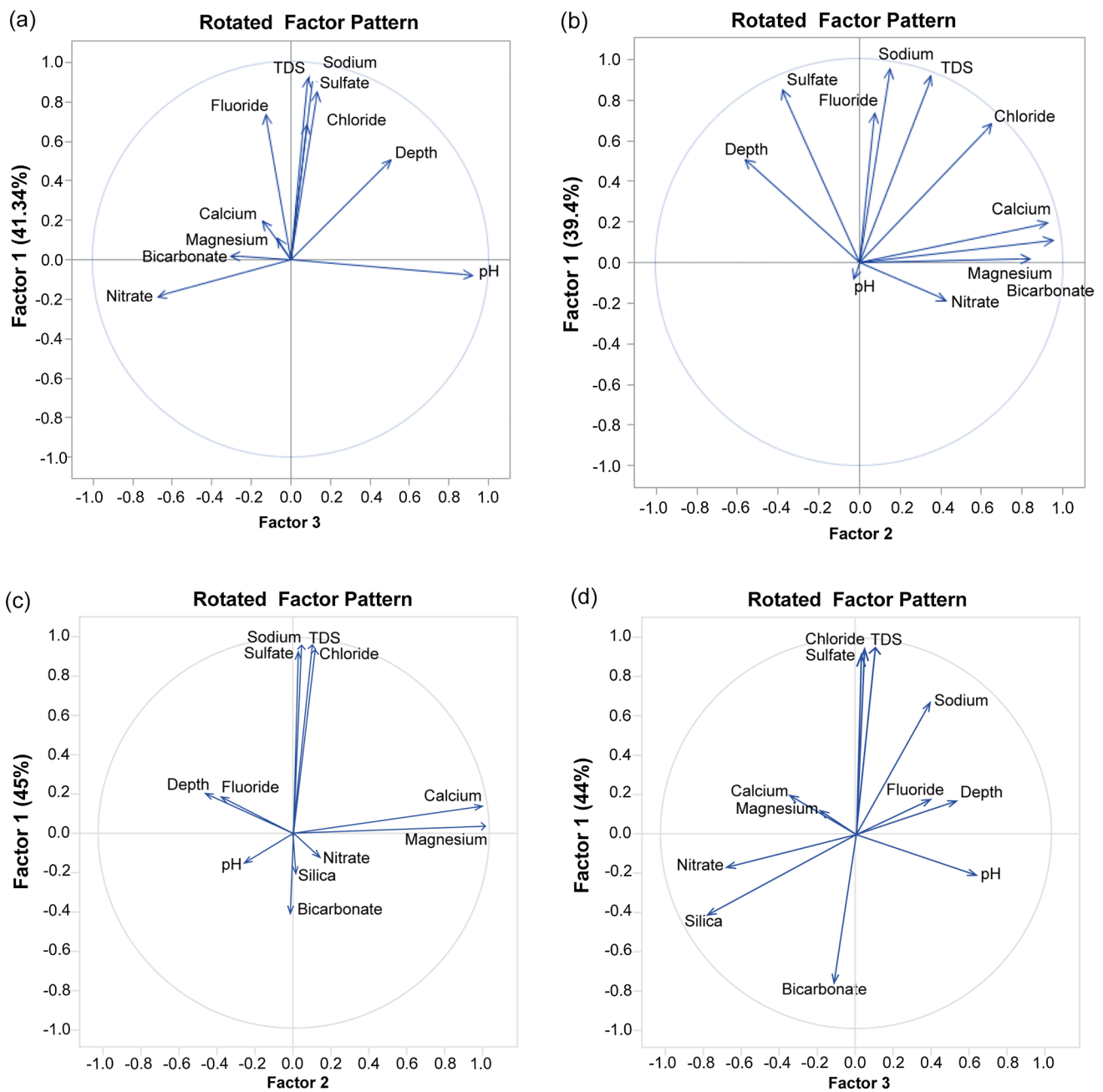


Fig. 5 Rotated factor patterns represent the standardized loading scores for the groundwater dataset in **a, b** Evangeline and **c, d** Chicot aquifers

were found statistically significant (p -value < 0.05) with weak to moderate explanatory power (for SO_4 : $R^2 = 65\%$, HCO_3 : $R^2 = 48\%$, TDS: $R^2 = 57\%$, Na: $R^2 = 52\%$, NO_3 : $R^2 = 58\%$, and K: $R^2 = 48\%$). Models indicate a gradual increase of TDS, Na, and SO_4 and declination of HCO_3 and K at higher depths in the Evangeline aquifer samples (Figs. 6 and 7). As mentioned in the previous section, the high concentrations of SO_4 , Na, and Cl at greater depths in the aquifers might be of both natural and anthropogenic origin. The concentration of K is low in both aquifers because of the relatively slow chemical weathering of K-bearing minerals in the reservoirs. However, the minor dissolution of K-feldspar and illite through the recharge zones increases the pH level and the concentrations of dissolved K in the shallower portion of the aquifer, which resulted in inverse depth relations in the Evangeline aquifer (Fig. 5a).

In the Chicot aquifer, the linear depth model for F was found statistically significant with high explanatory power ($R^2 = 86\%$ and Mean Square Error 0.15). The causes of slightly higher F in the deeper part of the Chicot aquifer might

Table 1 Depth differences of water chemistry during post-1977 in the Chicot and Evangeline Aquifers

Variables	Chicot aquifer depth difference (p value)	Evangeline aquifer depth difference (p value)
pH	0.36	0.014
TDS	0.034	<0.0001
Silica	0.031	–
Ca	–	0.0002
Mg	0.72	<0.0001
Na	0.047	<0.0001
HCO ₃	0.0015	<0.0001
SO ₄	<0.0001	<0.0001
Cl	<0.0001	0.001
NO ₃	0.014	0.001
F	0.0001	–

P-value < 0.05 represents significant differences in concentrations as a function of depth

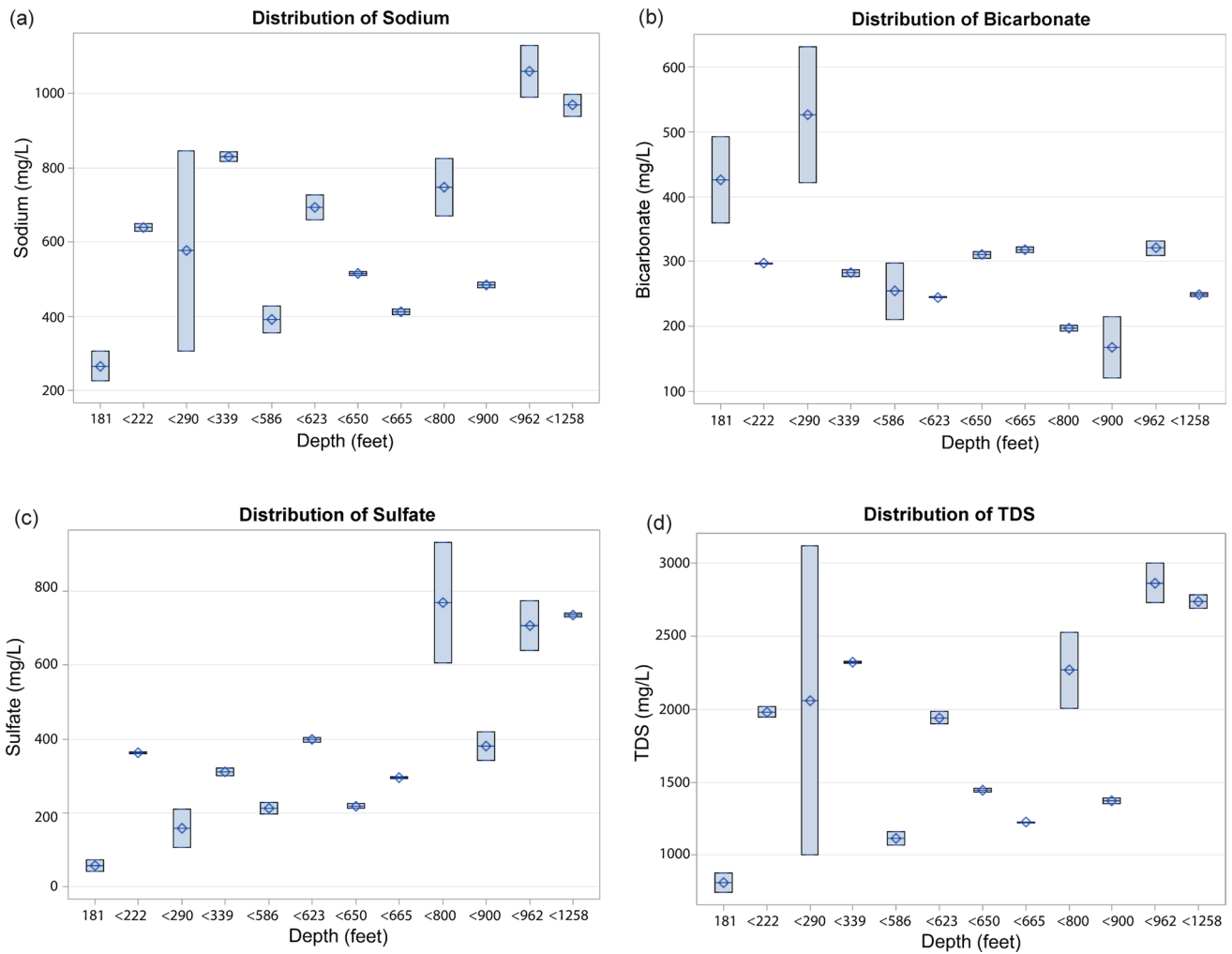
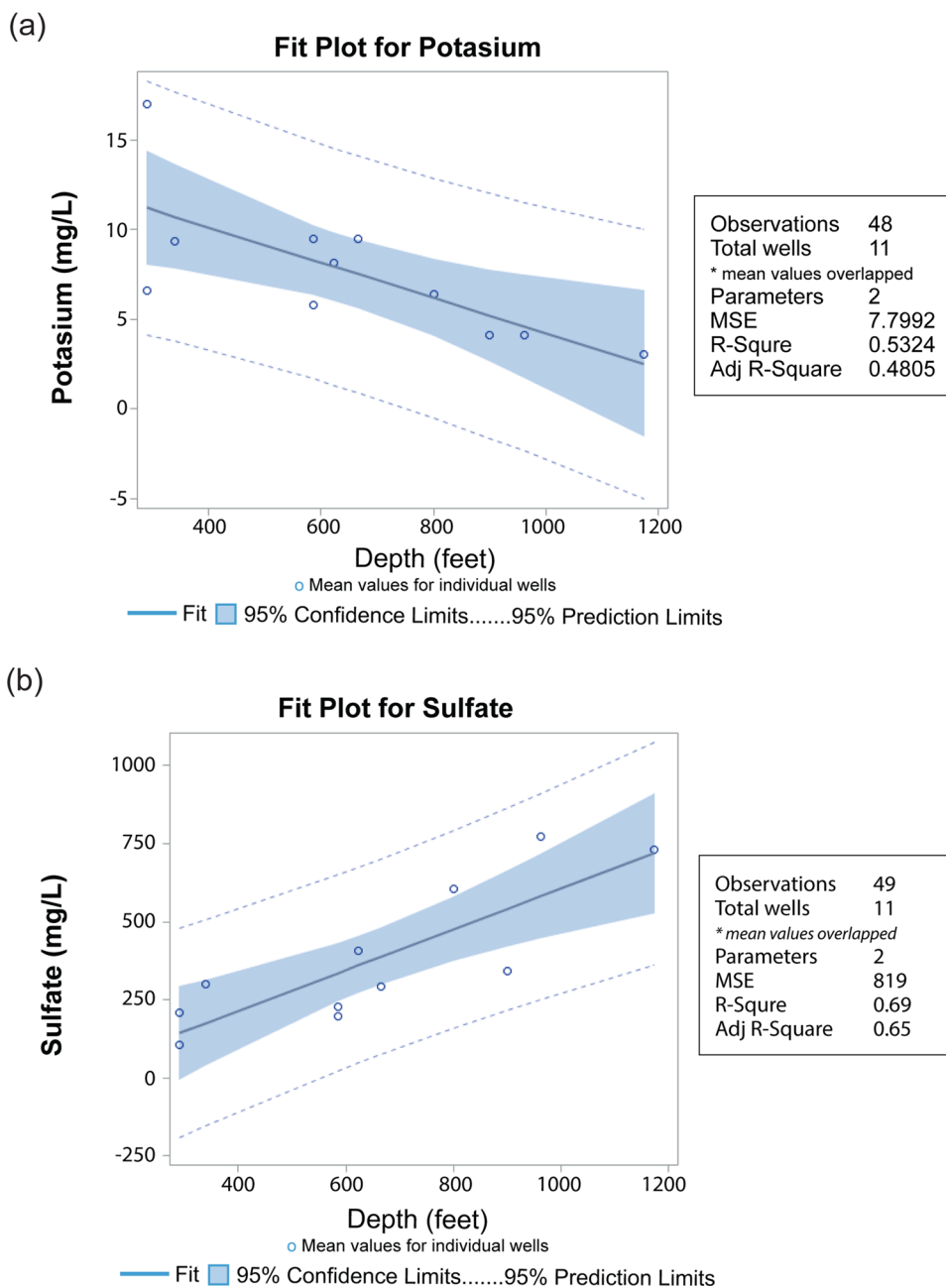


Fig. 6 Depth variations of groundwater chemistry in Evangeline aquifer. A high range in data values indicates multiple measurement results at different depth intervals

Fig. 7 Depth models of groundwater chemistry in Evangeline aquifer



result from the minor dissolution of fluoride-bearing minerals and the release of pore water from compacted clay layers through land subsidence. Groundwater availability models (GAM) based on field and lab-based data showed that land subsidence and compaction ultimately released connate water from the interbedded clay layers into the Chicot aquifer [9, 40, 41]. Furthermore, those studies reported a slight increase in groundwater salinity in areas where no significant mixing of the expelled fresh connate water within the aquifer happened. However, due to the release of water through a slow diffusive mechanism, including an alteration of the natural hydraulic gradient, which ultimately captured more fresh water from the outcrop, no noticeable increase in salinity was observed in some other areas.

On the other hand, ions such as NO_3 , HCO_3 , and dissolved Silica models showed a decreasing trend with progressively greater depths at a high statistical significance. Relatively higher silica at the shallower depths results from chemical reactions of aluminosilicates with rainwater near the outcrop areas. The depth models for the rest of the parameters in the Chicot aquifer were statistically non-significant. The lack of a well-defined depth relationship between TDS and the rest of the major ions in the Chicot aquifer might be due to the variation of aquifer depth

within short intervals and high salinity of groundwater from multiple sources, including salinity intrusion, rock-water interactions, and different anthropogenic sources of contamination.

A decrease of HCO_3 along the progressively greater depths was noticed in these aquifers, which conforms with a study by Gao et al. [42] in sand-gravel confined aquifers in the Songnen Basin of China. The gradually lower concentrations of NO_3 in both aquifers are associated with a relatively reducing environment at progressively greater depths.

Overall, the differences in the well-defined depth relations between the aquifers, more specifically, the lack of model's significance for TDS in Chicot aquifer, and most of the major ions except F, NO_3 , HCO_3 , and dissolved silica imply that the sources of salinity and the geochemical processes within the vertical profile along the groundwater flow paths are dissimilar within those two aquifer systems. Multiple factors simultaneously contribute to the variations in water chemistry between the Chicot and Evangeline aquifers in the study area. The major causes include longer groundwater travel paths from the outcrops, lower fluid flow rates [43] and a hence longer residence time in the Evangeline aquifer, heterogeneity in aquifer mineral composition, hydraulic properties, and retention of connate water influence the geochemical mechanisms of the rock-water interactions. Besides, various near-surface processes such as types of surface soils, proximity to the Nueces Bay, groundwater pumping and land subsidence rates, spatiotemporal variabilities of precipitation and evaporation rates, including anthropogenic pollution sources such as agricultural practice and farming, resulting in variations of water composition between those two aquifers and even within the same aquifer unit.

3.4 Multiple regression models to identify proxy for groundwater salinity

The goal of this regression analysis was to identify the proxies in a stepwise manner to predict the salinity using the least number of predictors with the highest explanatory power and least standard errors. Bie et al. [44] reported that the chemical composition of groundwater in the Chicot and Evangeline aquifers is generally fresh near the outcrop areas. The hydrogeochemical facies in the central part of the aquifers change from Ca- HCO_3 types to Na-Cl- HCO_3 or Na-Cl- SO_4 types along the regional flow path. Data showed that the groundwater facies in both aquifers are Na-K-Cl- SO_4 to Na-Cl- SO_4 type in the study area [37].

3.5 Stepwise elimination of less significant predictors for groundwater salinity in Chicot aquifer

Assuming TDS as a good indicator of groundwater salinity, forward and backward progressions were used to calibrate regression models to predict TDS using the major ions as potential predictors. Linear exhaustive search methods were applied separately for the Chicot and Evangeline aquifers for all samples during the given period.

The initial model started with all major ions to predict the groundwater salinity variations in Chicot and Evangeline aquifers, as provided in Eq. 2 above. The multiple linear regression models showed all variables to be highly significant except K, F, and log-transformed Mg^{2+} . Exhaustive search techniques were applied to make model selections based on explanatory power (R^2) and Akaike information criterion (AICc) values. In the subsequent stages, the choice of the models was based on the statistical significance of the model and individual predictors (p -value < 0.05 means significant), the model's high R^2 value, low AICc value, low standard errors, minimizing multicollinearity effects, and justifying the diagnostic plots of individual models in a stepwise manner. Each selected model is provided below with the model selection criteria mentioned herein.

The multiple regression-derived first model is given below:

$$TDS = Na + HCO_3 + SO_4 + Cl + NO_3 + \log(\text{Silica}) + \log(\text{Ca}) \quad (4)$$

Model p value ($< 2.2e^{-16}$), R^2 : 99%, Residual standard error: 6.7 and AICc value: 240.71).

Following the stepwise procedures, the final selected multiple linear models included only three predictors, HCO_3 , SO_4 , and Cl, as shown in Eq. 5 below. In the final model, the model properties remained the same based on multiple R^2 and P values. Residual standard error increased slightly (19.7), and the Q-Q plot showed that the distribution of residuals was normal. However, the Residual versus Fitted (RVF) plot showed higher values of residual distribution near the center, and the fitted plot was slightly curved.

$$TDS = 1.51SO_4 - 0.92HCO_3 + 1.61Cl + 4.13 \quad (5)$$

Among the ten variables tested, three in Eq. 5 were recognized as significant to explain the variations of TDS values. In this model, the highest weight was estimated for Cl and SO_4 . This preliminary model in Eq. 5 indicates that Cl, SO_4 , and HCO_3 could be the ideal proxies to predict groundwater salinity in the aquifer.

Due to a relatively high variable Influential Factor (VIF) and slight curviness in the scatter plots between SO_4 and Cl , SO_4 was later removed from the model, considering the higher significance level of Cl . A quadratic model was calibrated, which improved the RVF plot line slightly but reduced the model's explanatory power by approximately 3%. The standard error was increased to 27.03, as shown in Eq. 6.

$$TDS = -1.184\text{HCO}_3 + 2.304\text{Cl} - 0.00015\text{Cl}^2 + 829.8 \quad (6)$$

The model explains 96.1% of the variation of TDS (p value- $< 2.2e^{-16}$).

Although the above model 6 was found significant (p-value $< 2.2e^{-16}$), individual p-values for HCO_3 and Cl^2 became insignificant in this model. As a result, the concentration of Cl was seen as the most significant predictor and, hence, the most representative proxy to predict groundwater salinity variations in the Chicot aquifer.

3.6 Stepwise elimination of less significant predictors for groundwater salinity in Evangeline aquifer

All of the procedures mentioned above were also applied to data from the Evangeline aquifer that derived the final selected model, as shown in Eq. 7 below.

$$TDS = 2.2\text{Na} + 0.000479\text{Cl}^2 + 255.3 \quad (7)$$

This model explains 97.95% of the variation of TDS (p value- $< 2.2e^{-16}$).

Model results showed that the concentrations of Na and Cl were the most appropriate proxies to predict groundwater salinity variations in the Evangeline aquifer.

3.6.1 Linear models in Chicot and Evangeline aquifers using proxy variable(s)

The linear models were prepared using 75% of the measurements in each aquifer randomly chosen to avoid sample bias and considered as the training set. To assess the appropriateness of the proxy variables, the remaining 25% of the datasets were treated as testing sets to estimate the accuracies of the predictions of salinity level based on the linear model using only the proxy variables from the previous section.

For the chloride (predictor) distribution in the training dataset, the Q-Q plot showed a non-normal distribution of data, and the Shapiro-Wilk Normality test gave a p-value = 0.0005 < 0.01 , so both tests suggested the data were non-normal. Data for the testing set also showed a clear curve, and the Shapiro test showed the p-value = 0.0215, just above 0.01. Therefore, the t-test was found inappropriate, and bootstrap resampling was conducted to assess mean differences in Cl concentration between training and testing datasets. Hypothesis tests showed no significant difference between the chloride concentrations in the training and testing sets; hence, the model predictions were considered valid using those randomized datasets at the 90% confidence interval.

3.6.1.1 Models with the proxy variable in Chicot Aquifer For the Chicot aquifer, TDS showed a linear relationship to chloride concentration, as provided in Eq. 8. Boxplots of TDS and Cl were also similar in distribution. Results showed the model is statistically highly significant (p-value: $< 2.2e^{-16}$). The predictor (Cl) and the intercept are also highly significant based on P values. The diagnostic plot showed slight heteroscedasticity in the RVF plot. The Q-Q plot showed that the data distribution is non-normal, and a high Cook's distance close to 2 for a few data values indicates the presence of outliers in chloride concentration.

$$TDS = 368.29 + 2.27\text{Cl} \quad (8)$$

This model can explain 97.8% of the total variation of TDS. It has a standard error of 167.8.

Log-transformed values were applied to reduce the effect of outliers, but they significantly increased the intercept and the coefficient of $\log(\text{Cl})$. The model became:

$$TDS = -4363.6 + 2337.6\log(\text{Cl}) \quad (9)$$

Both the intercept and the predictor were significant, but the residual standard error became 671.2, and the model's explanatory power was also reduced to 65.1%.

Finally, a quadratic model was developed to improve the model diagnostics, such as the RVF plot. In this case, Cl , the square term of Cl , and the intercept of both were observed to be significant for predictions. The final model was found

to be highly significant (p -value: $1.01e^{-15}$), but the explanatory power of that model was reduced from 97.8 to 93.7%, and the residual standard error increased from 167.8 to 291. Furthermore, the model could not improve the diagnostic plots. Considering the above, the linear model in Eq. 8 was finally chosen for the groundwater salinity prediction to assess the efficiency of the proxy variable (Cl).

3.6.1.2 Models with the proxy variable in Evangeline aquifer TDS also showed a strong linear relationship with chloride concentration for the Evangeline aquifer. The model in Eq. 10 was highly significant based on the p value (p -value: $< 2.2e^{-16}$). Both the intercept and the predictor (Cl) were also significant. The predictor explained 92.7% of the total variation of TDS. The residual standard error was 167.5. Curviness was observed in the RVF plot, and the Q–Q plot showed the data to be normal enough, and no outlier effect was noticed.

$$TDS = 572.52 + 1.92Cl \quad (10)$$

The quadratic model in Eq. 11 was developed with the expectation of improving the model's performance.

$$TDS = 388.1 + 2.487Cl - 0.0003Cl^2 \quad (11)$$

This quadratic model was also highly significant (p -value: $< 2.2e^{-16}$).

In this case, Cl, the square term of Cl, and the intercept all were significant. This model increased its explanatory power from 92.7% to 93.5%, and the residual standard error decreased from 167.5 to 160.2, but the model could not improve the diagnostic plots. Considering all the above, the quadratic model in Eq. 11 was a better choice to assess the accuracy of predictions.

3.6.2 Salinity prediction interval and prediction accuracy estimations using proxy variables

In this section, to assess the appropriateness of the proxy variables to present groundwater salinity, the prediction performance of the testing sets was evaluated using Eq. 8 and Eq. 11 (as provided in the previous section) for the Chicot and Evangeline aquifers, respectively.

Results showed that 16 of the 30 test case predictions were accurate within a 95% confidence interval. For the Chicot aquifer, the model accurately predicted 78% of salinity variations using the proxy variable and underestimated TDS measurements by up to (–) 20 mg/L in the rest 22% of the testing dataset (Fig. 8a). On the other hand, in the Evangeline aquifer, the model could accurately predict only 43% and underestimated in the range of up to (–) 250 mg/L of the testing datasets (Fig. 8b). The model predictions in the Evangeline aquifer salinity level were highly influenced by the presence of significant outliers that ultimately affected the accuracy of predictions. The accuracy of the model predictions is summarized for both aquifers in Table 2 below.

One of the limitations of these models for potential implementation is that they underestimated actual measured values for a significant number of predictions in the Evangeline aquifer. The significant spatial and depth variability of water chemistry, specifically Cl and TDS values, affected the choice to include Cl as a proxy for this aquifer system.

4 Conclusions

A high variability of depth and thickness in the Chicot and Evangeline aquifers considerably influences groundwater chemistry in the study area. The correlations for different ion pairs between the Chicot and Evangeline aquifers showed some unique signatures, particularly the relationships among the different physicochemical parameters and major ion concentrations, indicating that the geochemical processes at those aquifers are not entirely similar. Na, SO_4 , and Cl concentrations account for the highest contributions to the elevated salinity level in both aquifers. The highest values of those parameters were noticed towards the southeast, close to Nueces Bay, indicating the possibility of saltwater intrusion in the Chicot aquifer. In the Evangeline aquifer, the maximum TDS, Na, Cl, and SO_4 were noticed at a greater depth within a relatively higher alkaline environment towards the southeast direction close to the head of the Nueces River.

Along with salinity intrusion, high evaporation, and groundwater withdrawal in the study area, rock-water interactions predominantly dissolution and precipitations of gypsum-anhydrite minerals, aluminosilicate minerals (e.g., feldspars and mica), illites, and carbonates influence the level of concentrations of Na, SO_4 , K, and Ca ions. Anthropogenic sources partially contribute to elevated SO_4 and Cl levels. High NO_3 and silica were restricted towards the shallower portion of

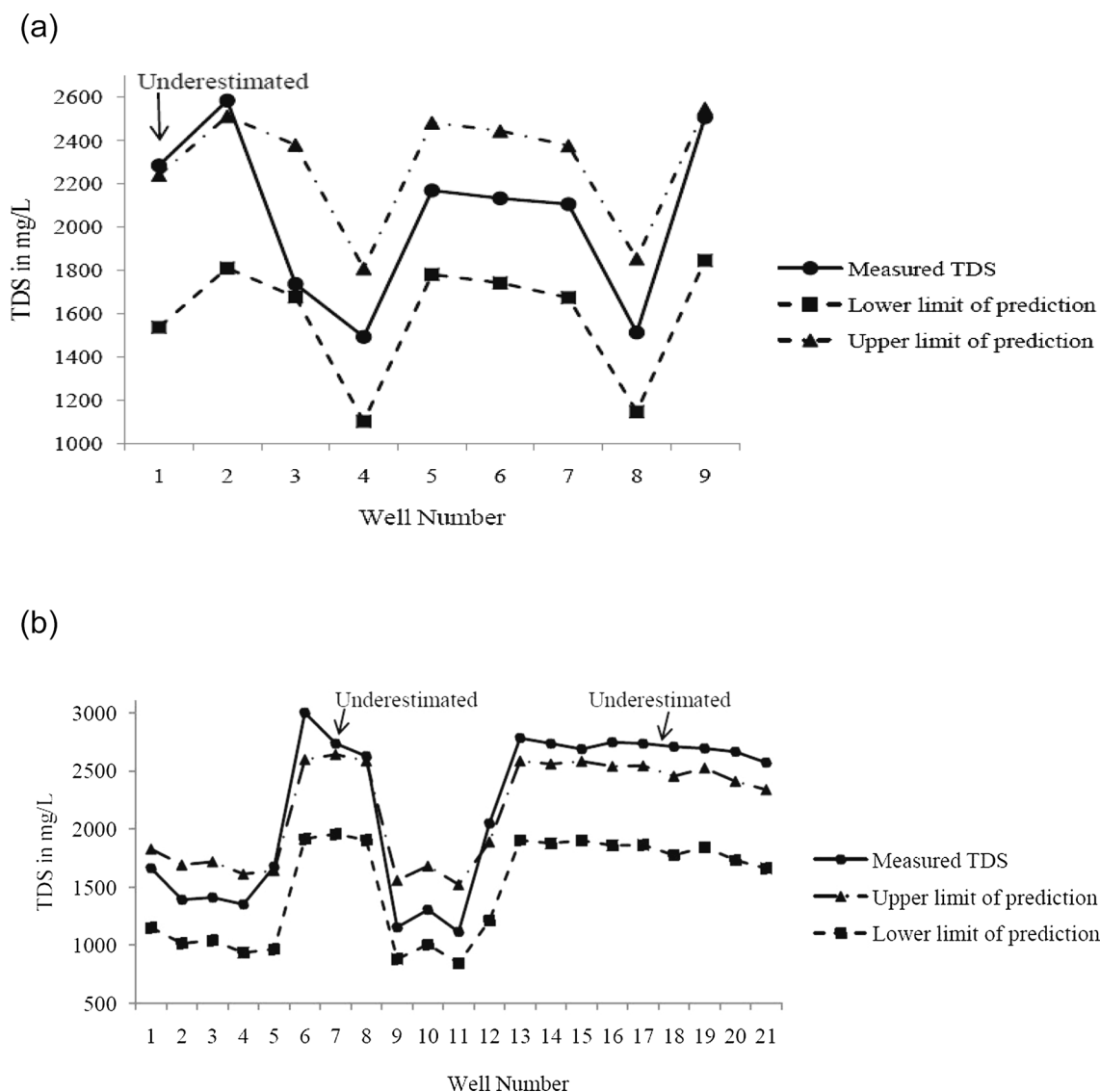


Fig. 8 Comparison of model predicted TDS values (upper and lower limits) with the measured TDS in the testing dataset for the **a** Chicot aquifer and **b** Evangeline Aquifer (at 95% confidence interval)

Table 2 Accuracy of the model predictions in Chicot and Evangeline aquifers

	Accurate prediction	Underestimated (below 95% confidence interval)	Overestimated (above 95% confidence interval)
Chicot Aquifer	7/9	2/9	0
Evangeline Aquifer	9/21	12/21	0

the aquifers and noticeably influenced by the anthropogenic pollution sources (e.g., agricultural practice and farming), pumping rates, and groundwater fluctuations, including the dissolution of aluminosilicate minerals with newly recharged rainwater near the outcrops. Although fluoride does not contribute much to groundwater salinity, the minor dissolution of fluoride minerals and water expulsion from compacted shale during land subsidence slightly increased the concentration at the deeper part of the Chicot aquifer. Unlike the Evangeline aquifer, TDS and the other predominant major ions, except HCO_3 , do not show any well-defined depth relations along the groundwater flow paths, which indicates

that the sources of salinity and the geochemical processes within the vertical profile are dissimilar in the Chicot and the Evangeline aquifer systems.

Stepwise elimination using regression analysis revealed that Cl could be the possible proxy to estimate groundwater salinity variations in both aquifers. The accuracy of the model predictions for the Chicot and Evangeline aquifers was 78% and 43%, respectively, at a 95% confidence interval. Therefore, it would be valid to consider Cl as the proxy variable for the Chicot aquifer but not appropriate for the Evangeline aquifer based on the model results. The availability of seasonal water chemistry data at a higher spatial and temporal resolution might reduce the outlier effects, yielding better model accuracy and precision for the Chicot and Evangeline aquifers in the study area.

Acknowledgements The author is thankful to the Texas Water Development Board for making the groundwater chemistry and groundwater level data available for public use. The author sincerely thanks the anonymous reviewers for their valuable comments that improved the quality of this manuscript.

Author contributions RHK contributed to the study's conception and design, performed data analysis, and prepared the first draft of the manuscript.

Funding This research received no specific grant from public, commercial, or not-for-profit funding agencies.

Data and code availability Data is available for public use from the TWDB website. Statistical R and SAS packages were used to conduct the analysis. The codes and datasets generated during and/or analyzed during the current study are not publicly available but are available from the corresponding author at reasonable request.

Declarations

Ethics approval and consent to participate The author has read, understood, and complied as applicable with the "Ethical responsibilities of Authors" statement as found in the Instructions for Authors.

Institutional review board statement Not applicable.

Informed consent Not applicable.

Competing interests The author has no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. McCleskey RB, Cravotta CA, Miller MP, Tillman F, Stackelberg P, Knierim KJ, Wise DR. Salinity and total dissolved solids measurements for natural waters: an overview and a new salinity method based on specific conductance and water type. *Appl Geochem*. 2023. <https://doi.org/10.1016/j.apgeochem.2023.105684>.
2. Todd DK. *Groundwater hydrology*, 3rd ed. John Wiley and Sons Inc; 1980.
3. Texas Water Development Board, Water for Texas 2012 State Water Plan. Chapter 4, climate of Texas; 2012, pp. 1–9. https://www.twdb.texas.gov/publications/state_water_plan/2012/2012_SWP.pdf
4. Union of Concerned scientists and Ecological Society of America. Fact sheet is based on the findings of Confronting Climate Change in the Gulf Coast Region; 2001, pp. 1–2. www.ucsusa.org
5. Langevin C, Swain E, Wolfeer M. Simulation of integrated SW/GW flow and salinity for a coastal wetland and adjacent estuary, U.S. Geological Survey greater Everglades priority ecosystem sciences program, Open file report, 2005, p. 214–233.
6. Ferrill DA, Sims DW, Waiting DJ, Morris AP, Franklin NM, Schultz AL. Structural framework of the Edwards Aquifer recharge zone in South Central Texas. *Geol Soc Am Bull*. 2004;116:407–18.
7. Haley M, Ahmed M, Gebremichael E, Murgulet D, Starek M. Land subsidence in the Texas Coastal bend: locations, rates, triggers, and consequences. *Remote Sens*. 2022. <https://doi.org/10.3390/rs14010192>.
8. Scanlon BR, Reedy R, Strassberg G, Huang Y, Senay G. Estimation of groundwater recharge to the gulf coast aquifer in Texas, USA. Final contract report to Texas Water Development Board; 2011. p. 9–30.
9. Chowdhury AH, Boghici R, Hopkins J. Hydrochemistry, salinity distribution, and trace constituents, implications for salinity sources, geochemical evolution, and flow system characterization, Gulf coast aquifer, Texas, Chapter 5, TWDB Report 365; 2006. p. 3–39.
10. Hoel HD. Goliad formation of the south Texas Gulf Coastal Plain—Regional genetic stratigraphy and uranium mineralization, Masters' thesis, The University of Texas at Austin; 1982.

11. Waterstone Environmental Hydrology and Engineering Inc., Groundwater Availability of the Central Gulf Coast Aquifer: Numerical Simulations to 2050 Central Gulf Coast, Texas; 2003. p.114.
12. Langmuir D. Aqueous environmental geochemistry. Prentice Hall; 1997. p. 60–365 (ISBN 0-02-367412-1).
13. Soil, R.F., Upper Tertiary and Quaternary depositional systems, central Coastal Plain, Texas—Regional geology of the Coastal aquifer and potential liquid-waste repositories: Texas Bureau of Economic Geology Report of Investigations 108, Austin; 1981. p.89.
14. Hudak PF. Sulfate and chloride concentrations in Texas aquifers. *Environ Int.* 2000. [https://doi.org/10.1016/S0160-4120\(00\)00078-7](https://doi.org/10.1016/S0160-4120(00)00078-7).
15. Breier JA, Edmonds HN. High ²²⁶Ra and ²²⁸Ra activities in Nueces Bay, Texas indicate large submarine saline discharges. *J Mar Chem.* 2007;103:131–45.
16. Bauer P, Held RJ, Zimmermann S, Linn F, Kinzelbach W. Coupled flow and salinity transport modelling in semi-arid environments: the Shashe River Valley, Botswana. *J Hydrol.* 2006;316:163–83.
17. Costelloe JF, Grayson RB, McMahon TA, Argent RM. Spatial and temporal variability of water salinity in an ephemeral arid-zone river, central Australia. *J Hydrol Process.* 2005;19:3147–66.
18. Subyani AM. Hydrochemical identification and salinity problem of ground-water in Wadi Yalamlam basin, western Saudi Arabia. *J Arid Environ.* 2005;60:53–66.
19. Bickel S, Bogojeska J, Lengauer T, Scheffer T. Multi-task learning for HIV therapy screening. In: Proceedings of the 25th international conference on machine learning, Helsinki, Finland; 2008.
20. Foll M, Gaggiotti O. A genome -scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics Soc Am.* 2008. <https://doi.org/10.1534/genetics.108.092221>.
21. Ganoulis J. Risk analysis of water pollution. Second revised and expanded edition. Chapter 1: Water resources-quantity and quality. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim; 2009. ISBN: 978-3-527-32173-5
22. Gutierrez BT, Plant NG, Thieler ER. A Bayesian network to predict coastal vulnerability to sea level rise. *J Geophys Res.* 2011;116:1–15.
23. Montagna PA. GEM: a general experimental method for the design and analysis of Ecological experiments. Harte Research Institute for Gulf of Mexico studies, Texas A&M University-Corpus Christi; 2012. p. 1–49.
24. Pathak H, Limaye SN. Study of seasonal variation in groundwater quality of Sagar city (India) by principal component analysis. *E-J Chem.* 2011;8:2000–9.
25. Sahoo S, Jha MK. Groundwater level prediction using multiple linear regression and artificial neural techniques: A comparative assessment. *J Hydrol.* 2013;21:1865–87.
26. Vahid G. Modeling of groundwater level using statistical methods and GIS: a case study: Amol-Babol Plain, Iran. *Int J Water Resour Environ Sci.* 2013;2:53–9.
27. Hanson C, Abraham P. Depth and spatial variation in groundwater chemistry-Central Canterbury Plain. Environment Canterbury Regional Council. Report R09/39, 2009. ISBN 978-1-86937-969-8
28. Maata S. Predicting groundwater levels using linear regression and neural networks. CS 229 Final Report; 2011. p. 1–5.
29. Richman MB, Trafalis TB, Adrianto I. Missing data imputation through machine learning algorithms. *Artificial intelligence methods in the environmental sciences*; 2009. p. 154–169.
30. Drought History for Texas’ 10 Regions. South Central Climate Science Center in Norman, Oklahoma, 2018. <https://southcentralclimate.org/wp-content/uploads/2018/12/Drought-History-for-Texas-10-Regions-pdf>
31. Walton NRG. Electrical conductivity and total dissolved solids—what is their precise relationship? *Desalination.* 1989;72:275–92.
32. Uddameri V, Honnunar V, Hernandez EA. Assessment of groundwater water quality in central and southern Gulf Coast aquifer, TX using principal component analysis. *Environ Earth Sci.* 2014;71:2653–71. <https://doi.org/10.1007/s12665-013-2896-8>.
33. Selmane T, Dougha M, Hasbaia M, et al. Hydrogeochemical processes and multivariate analysis for groundwater quality in the arid Maadher region of Hodna, northern Algeria. *Acta Geochim.* 2022;41:893–909. <https://doi.org/10.1007/s11631-022-00553-y>.
34. Hem JD. Study and interpretation of the chemical characteristics of natural water. 3rd ed. Washington: USGS WSP2254; 1989.
35. Chandran S, Selvan P, Dhanasekarapandian M, et al. Hydrogeochemical characteristics of surface and groundwater: suitability for human consumption and irrigated agriculture purposes in Suruliyar sub basin, South India. *Environ Geochem Health.* 2022. <https://doi.org/10.1007/s10653-021-01145-0>.
36. Capuano RM, Jones CR. Cation exchange in groundwater chemical evolution and prediction of Paleo-groundwater flow: a natural system study. *Water Resour Res.* 2020. <https://doi.org/10.1029/2019WR026318>.
37. Khan RH, Smith-Engle J, Tissot P, Murgulet D. Temporal spatial and depth variations of ground water chemistry: an indicator of hydro-geochemical evolution in shallow coastal aquifers, South Texas, Gulf Coast Association of Geological Societies Transactions; 2016. vol. 66. p. 313–29.
38. Baker ET. Stratigraphic and hydrogeologic framework of part of the Coastal Plain of Texas: Texas Department of Water Resources Report 236, Austin; 1979. p. 1–41.
39. Chaudhuri S, Ale S. Evaluation of long-term (1960–2010) groundwater fluoride contamination in Texas. *J Environ Qual.* 2014;43:1404–16. <https://doi.org/10.2134/jeq2013.04.0133>.
40. Kasmarek MC, Robinson JL. Hydrogeology and simulation of Groundwater flow and land surface subsidence in the Northern part of the Gulf Coast Aquifer System. U. S. Geological Survey Scientific Investigations Report. 2004, 2004-5102, p. 103. <https://doi.org/10.3133/sir20045102>
41. Gabrysck, R. K. Case history no. 9.12—The Houston-Galveston region, Texas, USA. In Poland JF, editor. Guidebook to studies of land subsidence due to groundwater withdrawal: UNESCO studies and reports in hydrology; 1984. vol. 20. p. 253–62.
42. Gao Z, Guo H, Qiao W, Ke T, Zhu Z, Cao Y, Su X, Wan L. Abundant Fe (III) oxide-bound arsenic and depleted Mn oxides facilitate arsenic enrichment in groundwater from a sand-gravel confined aquifer. *J Geophys Res Biogeosci.* 2022;127(8):e2022JG006942.
43. Chowdhury AH, Wade S, Mace RE, Ridgeway C. Groundwater availability model of the central gulf coast aquifer system: numerical simulations through 1999, Texas. TWDB report; 2004. p. 108.
44. Bie H, Yang C, Liu P. Assessing groundwater monitoring strategy for leakage detection in the Texas Gulf Coast Aquifer (USA) at a hypothetical CO2 storage site: a reactive transport modeling approach. *Hydrogeol J.* 2019. <https://doi.org/10.1007/s10040-018-1887-y>.