Assessment of groundwater water quality in central and southern Gulf Coast aquifer, TX using principal component analysis

**Venkatesh Uddameri · Vivekanand Honnungar ·
E. Annette Hernandez**

Abstract Principal component analysis has been applied for source identification and to assess factors affecting concentration variations. In particular, this study utilizes principal component analysis (PCA) to understand groundwater geochemical characteristics in the central and southern portions of the Gulf Coast aquifer in Texas. PCA, along with exploratory data analysis and correlation analysis is applied to a spatially extensive multivariate dataset in an exploratory mode to conceptualize the geochemical evolution of groundwater. A general trend was observed in all formations of the target aquifers with over 75 % of the observed variance explained by the first four factors identified by the PCA. The first factor consisted of older water subjected to weathering reactions and was named the ionic strength index. The second factor, named the alkalinity index explained greater variance in the younger formations rather than in the older formations. The third group represented younger waters entering the aquifers from the land surface and was labeled the recharge index. The fourth group which varied between aquifers was either the hardness index or the acidity index depending on whether it represented the influences of carbonate minerals or parameters affecting the dissolution of fluoride minerals, respectively. The PCA approach was also extended to the well scale to determine and identify the geographic influences on geochemical evolution. It was found that wells located in outcrop areas and near rivers and streams had a larger influence on the factors suggesting the importance of surface water–groundwater interactions.

Introduction

Groundwater plays an important role in shaping the economic and ecological makeup of arid and semiarid regions. Aquifers are often seen as accessible, cost-effective and reliable sources of water supplies. However, aquifers are slowly replenished and given their benefits, the potential for overexploitation exists (Zekster et al. 2005). Sustainable groundwater management requires that aquifer resources be available to future generations as they are to current users. Furthermore, it is important that current generations have equitable access to this resource (Uddameri 2005). There is a growing realization that local-scale aquifer management is vital for striking a balance between using groundwater for economic development but without compromising the ecosystem services it provides (Uddameri and Kuchanur 2007).

The physical, chemical and biological characteristics of groundwater define its intended use. Furthermore, the quality of groundwater limits its availability. Aquifer resources are particularly susceptible to pollution and deterioration due to activities carried out at the land surface (Connell and van den Daele 2003; Gogu and Dassargues 2000; NRC 1993; Uddameri and Honnungar 2007). Aquifers are active geochemical zones, where a variety of chemical reactions such as hydrolysis, ion exchange, dissolution and dissociation occur (Freeze and Cherry 1979). In addition, they are open systems that exchange water,

V. Uddameri · E. A. Hernandez (✉)
Department of Civil and Environmental Engineering, Texas
Tech University, Box 41023, Lubbock, TX 79409, USA
e-mail: annette.hernandez@ttu.edu

V. Honnungar
The Energy and Resources Institute (TERI), Bengaluru, India

gases and other colloidal and dissolved constituents with the atmosphere, soil, surface water bodies and other aquifers to which they are connected. Therefore, groundwater quality is affected by a variety of anthropogenic as well as natural processes. Consequently, understanding groundwater quality is crucial to properly managing this resource.

Monitoring groundwater quality is an expensive proposition as water quality can be defined using many different parameters. In addition, monitoring is challenging because most existing wells tend to be on privately owned property which may restrict their accessibility. As such, groundwater monitoring often tends to be ad hoc and, typically, not collected unless there is an evidence of contamination (Uddameri 2007). Furthermore, regional-scale monitoring is often limited to the measurement of a few bulk parameters such as total dissolved solids (TDS) and specific conductance (Cartwright et al. 2004).

The total groundwater in an aquifer is a net result of accumulation that has taken place over a long geologic time. The residence time in the aquifer affects the nature and extent of the chemical reactions to which the groundwater is subjected. Moreover, aquifers are intrinsically heterogeneous and the movement of water and the associated geochemical interactions along the flowpaths has an impact on the solution chemistry. Generally speaking, groundwater at any given location is a mixture of waters that have different origins and followed various flowpaths and as such are affected by diverse geochemical controls. Therefore, the information pertaining to the origin and evolution of groundwater is embodied in the concentration of solutes (Bicalho et al. 2012; Stuart et al. 2010). The spatial variability in the concentrations of a given solute and the relative abundance of distinct solutes within an aquifer implicitly contain information related to the origin of the groundwater and the potential reactions that have occurred (or are occurring) to control the composition of the solutes. Hydrochemical facies represent distinct zones or compositional categories that can be used to group water based on measured quality characteristics (Dalton and Upchurch 1978; Todd and Mays 2005). The groupings provide basic insights into the origin and evolution of water. These insights, in turn, provide the fundamental understanding required to develop and guide monitoring activities and facilitate scientifically credible groundwater quality management.

Principal component analysis (PCA) is a statistical data reduction technique that can be used to aggregate the effects of a large number of variables into a small subset of factors (Hamilton 1992). Dalton and Upchurch (1978) used a variant of PCA (called factor analysis) for interpretation of hydrochemical facies. They indicate that factor analysis is more convenient than traditional graphical techniques because they do not require closed systems nor are they limited to a select number of ions. Lawrence and Upchurch

(1982) used factor analysis to separate chemicals that reflect aerially significant recharge processes. The analysis by Usunoff and Guzman–Guzman (1989) suggests principal component can be beneficial during the initial stages of hydrochemical studies. They used PCA-based factor analysis both in the R-mode to identify sources and in the Q-mode to assess correlations among various sampling points. Melloul and Collin (1992) highlight the ability of PCA to combine geochemical and physical factors during analysis and recommend it as a complementary technique that enhances traditional geochemical methods such as piper plots and other graphical techniques (Hem 1970). Grande et al. (1996) used factor analysis to study agricultural contamination in the Ayamonte-Huelva aquifer system in Spain. Powers et al. (1997) used PCA for identification of source areas with gasoline and coal tar at a manufacturing gas plant site. Suk and Lee (1999) used factor analysis in conjunction with cluster analysis to characterize hydrochemical system in Inchon, Korea.

More recently, Duffy and Brandes (2001) used PCA to classify 116 chemical compounds into three groups based on their chemical properties as well as their environmental responses. They used PCA-derived groups to identify potential sources at a contaminated site that has no documented disposal history. McGuire et al. (2005) used a combination of factor analysis and clustering to study redox changes due to recharge events. Lucas and Jauzein (2008) employed PCA for source identification and to assess factors affecting concentration variations in various chlorinated solvents to support site remediation activities. Hildebrandt et al. (2008) also used PCA to elucidate pesticide contaminant patterns and their seasonal trends in surface and groundwater systems in Spain. Mor et al. (2009) used PCA-based factor analysis in conjunction with correlation analysis to appraise salinity and fluoride levels in India. Fernandes et al. (2010) applied PCA to identify water–rock interaction and anthropogenic activities in the Essaouira aquifer, Morocco. Bakari et al. (2012) used hierarchical cluster analysis in conjunction with factor analysis to classify groundwater and identify the major factors influencing its quality in Tanzania.

Building along these lines of inquiry, the primary goal of this study is to utilize PCA to understand geochemical characteristics of the groundwater in the central and southern portions of the Gulf Coast aquifer in Texas. More specifically, PCA is used in an exploratory mode on a spatially extensive multivariate dataset consisting of major ions, groundwater levels and important bulk chemical measures to delineate regional-scale sources and controls on groundwater in distinct aquifer formations of the Gulf Coast aquifer. The study is one of the first attempts to understand groundwater evolution characteristics in different formations of the Gulf Coast aquifer in South Texas.

Methodology

Hydrogeochemical information is often tabulated to contain values of various parameters collected at each location or time. Each column represents a hydrogeochemical parameter of interest (e.g., water level, total dissolved solids (TDS), etc.) and each row (or record) represents a measurement set (observation vector) in space or time. Principal component analysis is a multivariate statistical technique that transforms a collection of correlated parameters (e.g., water levels, TDS and other water quality parameters) into a group of uncorrelated (orthogonal) variables called principal components.

Principal component analysis can be carried out using either the correlation matrix or the variance–covariance matrix (Borgognone et al. 2001; Hamilton 1992). Correlation coefficients are calculated by computing the variance–covariance upon standardization of the data (i.e., subtraction of the mean from the original value and dividing by its standard deviation). This standardization makes all the measurements of parameters commensurate and eliminates bias arising from different measurement scales. Therefore, the correlation matrix is commonly used in hydrologic and environmental literature for carrying out PCA (Duffy and Brandes 2001; Güler et al. 2002; McCuen 1993) as the measured hydrogeochemical parameters of interest typically exhibit wide variability.

Principal component analysis succinctly summarizes the information contained in a series of multivariate observations using eigenvalues and eigenvectors. The eigenvalues represent the total amount of variance in the original dataset. As the variance of each standardized variable is equal to unity, the total variance (which the eigenvalues explain) is equal to the number of variables. The first principal component has the highest eigenvalue, the second principal component has the second highest eigenvalue and so on. By construction, the principal components are not correlated with each other. Typically, the first few factors explain a significant (>80 %) amount of variance and are generally sufficient to represent the data. When PCA is carried out using the correlation matrix, the Kaiser rule indicates that factors having eigenvalues greater than 1 are of most significance (McCuen 1993). Scree plots map each component and the corresponding eigenvalue and are a useful visual aid to eliminate subjectivity associated with the Kaiser rule, particularly when the eigenvalues of components are in the vicinity of one. A subset of components selected to represent the data are generally referred to as factors.

Each parameter in the original dataset is correlated to each principal component (or factor). The factor loadings represent the correlation between the parameter and the factor. As such, the factor loadings range between ±1 with the negative sign indicating a negative relationship between the parameter and the principal component. A parameter can be assumed to be associated with a particular factor if the absolute value of its factor loading is close to 1. This information can be used to interpret what the factor represents. Furthermore, as the first factor explains more variance than the second and so on, parameters that load significantly onto the first factor can be considered to be more important to explain the total variance in the dataset than the second and so on. Thus, PCA enables the categorization of parameters into manageable groups. Communalities (C_{ij}) measure the cumulative fraction of variance of variable i in j principal components and in conjunction with eigenvalues can also be helpful in selecting a subset of factors. As a rule of thumb, the selected number of factors must yield a communality value of 0.6 or higher for each parameter, although this limit can be relaxed when the number of variables is large (McCuen 1993).

In many instances, a parameter may have similar associations with more than one factor and many diverse sets of parameters may also have similar associations with a given factor. Both these conditions make the interpretation of principal components difficult. Factor rotation seeks more interpretable factors by polarizing the factor loadings such that each parameter loads strongly on only one factor and near zero on the other factors. The varimax factor rotation scheme preserves the orthogonal structure of the principal components. An oblique (promax) rotation on the other hand permits correlation among factors and as such leads to even greater polarization and thus seeks a more simple structure and interpretable components. Oblique rotation is more realistic if the factors (principal components) are viewed as unmeasured variables that underlie the measured data (Hamilton 1992). There are several commercial software programs that implement principal component analysis. Two standard software, MATLAB (Mathworks, Inc.) and STATISTICA (Statsoft Inc.) were utilized in this study.

Study area

The central and the southern portions of the Gulf coast aquifer in Texas were the focus of this study (Fig. 1). The northern portion of the study area falls under the jurisdiction of the Groundwater Management Area 15 (GMA 15) while the southern portion falls under the Groundwater Management Area 16 (GMA 16). The geology of the Gulf Coast aquifer is characterized by deposition of sedimentary facies under a fluvial-deltaic to shallow marine environments during the Miocene and Pleistocene periods. Repeated coastal water incursions and subsidence have led to
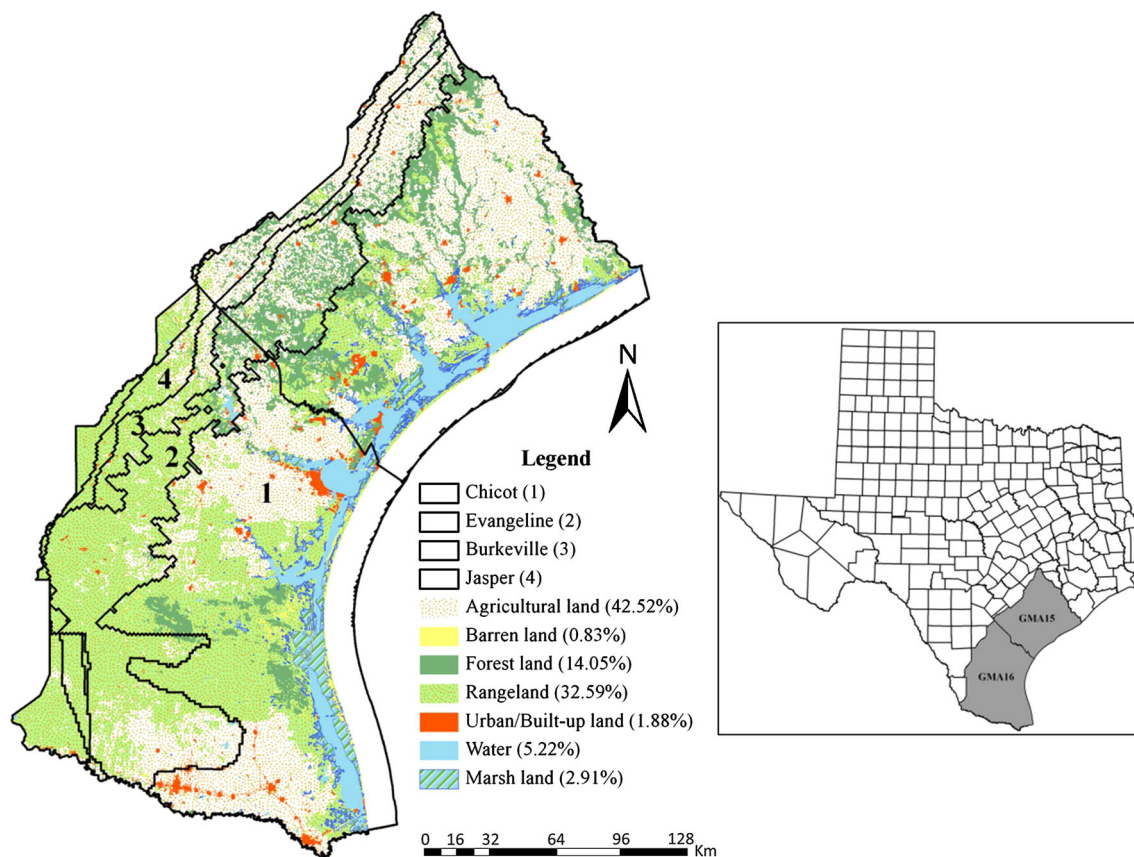
Fig. 1 Study area with aquifers and land use

cyclic deposition of discontinuous beds of sand, silt, clay and gravel (Galloway et al. 2000). While complex and controversial, the aquifer is generally accepted to be comprised of four water-bearing stratigraphic units: (1) the Jasper aquifer; (2) the Burkeville confining system; (3) the Evangeline aquifer and (4) the Chicot aquifer (Baker 1979). All these formations generally slope eastward with increasing thicknesses.

The Jasper aquifer is the deepest confined water-bearing unit in the Gulf Coast aquifer with a small outcrop area along the western sections of the study area. The aquifer is comprised of interbedded sand, silt and clay sediments with intermixed volcano-clastic and tuffaceous material (Hosman 1996). The Burkeville confining system predominantly consists of terrigenous clastic sediments (silt and clay). However, it is important to note that the formation is comprised of many individual sand layers which contain fresh to slightly saline water (Baker 1979). Both Jasper and Burkeville formations are made up of Miocene aged sediments and treated as a single unit in this study. The Evangeline formation is mainly comprised of Pliocene-aged sediments and consists of a greater percentage of coarse-grained sediments including cobbles, clay balls and wood fragments at the base of the formation (Hosman

1996). The upper sections are comprised of fine grained sands that are cemented with calcium carbonate and referred to as caliche (Chowdhury and Turco 2006). The Evangeline aquifer is also locally referred to as Goliad sand formation. This aquifer formation is one of the most prolific formations in the study area and used widely for water supply. The Chicot aquifer mainly consists of depositions of Pleistocene-aged sediments with Lissie formation and Beaumont clay being the two major subdivisions. The Holocene era alluvium deposits along the major river basins are also included in this formation. The distinction between Chicot and Evangeline aquifer formations is difficult to make in many locations and is a subject of controversy among geologists (Chowdhury and Turco 2006). For the purpose of this study, the outcrop areas and the base of these aquifer formations were derived from the stratigraphic structure developed by Baker (1979) and are incorporated in the central and southern Gulf Coast aquifer groundwater availability models (Chowdhury and Mace 2004; TWDB 2003).

The mineralogical components of Miocene-Pliocene sandstones (Burkeville and Jasper formations) are poorly known (Chowdhury et al. 2006). Calcium carbonate deposits (Caliche) are predominant in Evangeline aquifer
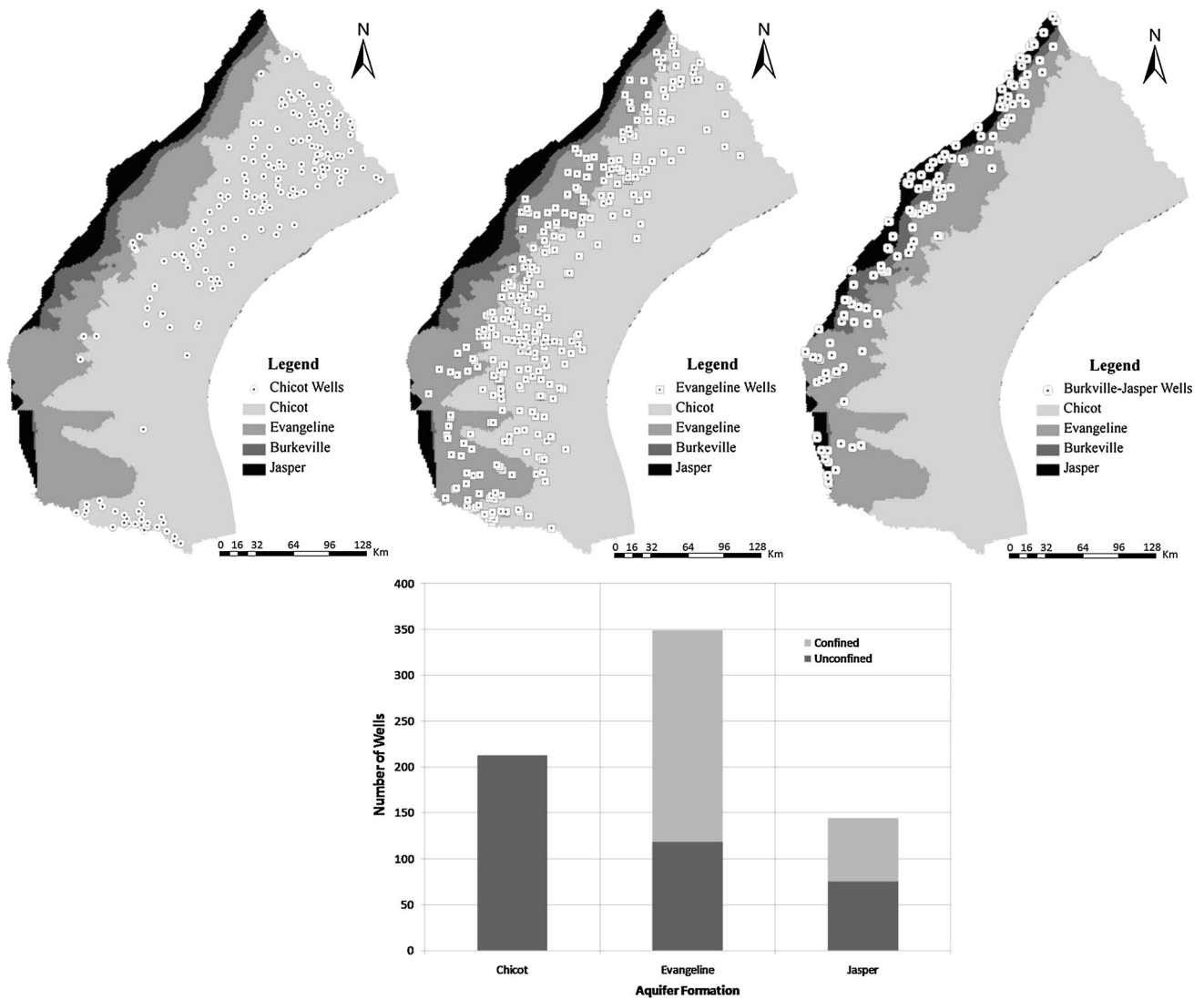
Fig. 2 Wells in each of the aquifers with fraction in confined and unconfined portions

formation (Sellards et al. 1932). Furthermore, the Evangeline formation is known to contain significant amounts of orthoclase and plagioclase feldspars and volcanic rock fragments (Hoel 1982). The clays are mostly comprised of montmorillonite with minor amounts of illite, chlorite and kaolinite (Gabrysch and Bonnet 1975). The hydrologic characteristics vary considerably in the study area. The annual average rainfall is close to 50.8 cm (20 inches) in the southern and western portions and increases to over 101.6 cm (40 inches) in the northern portions. The study area is predominantly overlain by rangelands in the south and agricultural lands in the north. Significant water deficits are projected in the south along the USA–Mexico border (TWDB 2007) and several large-scale groundwater development projects are being contemplated to meet these demands and supply water to the urban corridors of San Antonio and Houston (Uddameri and Kuchanur 2007).

The data used in this study were obtained from the comprehensive statewide groundwater monitoring program carried out by the Texas Water Development Board and tabulated in their groundwater database. In addition to annual water level and periodic water quality measurements, the database also contains data collected by cooperating agencies. Sample collection and analysis follow well-documented standard protocols with adherence to comprehensive quality control/quality assurance procedures (Chowdhury et al. 2006; Hopkins 2010). As the focus here was to evaluate ambient conditions, long-term average values (1980–2010) of groundwater levels and 15 other chemical constituents were extracted. The parameters considered here include groundwater levels, pH, total alkalinity, major cations including (Na^+, K^+, Ca^{2+}, Mg^{2+}), major and minor anions (SO_4^{2-}, HCO_3^-, Cl^-, NO_3^-, F^-), Silica (SiO_2) and bulk parameters (total dissolved solids

(TDS), hardness, total alkalinity and specific conductance). The dataset comprised of a total of 706 wells in the entire study area. Figure 2 depicts the wells in each aquifer formation and the proportion that lie in unconfined (outcrop) and confined conditions. As Chicot formation is the upper most layer, all wells exist in unconfined conditions. The number of confined to unconfined wells in the Evangeline formation was 66 and 34 %, respectively. The Evangeline formation is used more extensively particularly in the southeastern sections of the study area due to its greater sandcontent and better water-bearing characteristics than the overlying Chicot formation in that part of the study area (e.g., Mason 1963). The confined and unconfined wells in the Burkeville–Jasper formation were close to 50 %.

Results and discussion

Summary statistics and correlation measures

The summary statistics for Chicot, Evangeline and Jasper formations are presented in Tables 1, 2 and 3. The mean values of TDS sodium, potassium and chloride increase with depth in the Gulf Coast aquifer (i.e., from Chicot to Jasper) indicating the natural evolution of groundwater (Freeze and Cherry 1979). The Gulf Coast aquifer formations are well buffered with respect to pH and the water is generally near neutral or slightly alkaline. The total

hardness of the water decreases with depth most notably due to a decrease in the magnesium concentrations. However, there is no noticeable trend in the calcium concentrations between different aquifer formations indicating that there could be multiple sources operating within the aquifer. Most parameters, with the exception of bicarbonate, pH and alkalinity exhibit considerable positive skewness (extended right-tailed distributions) in the Chicot and Evangeline formation. Interestingly, bicarbonate, pH and alkalinity exhibit considerable skewness in the Jasper formation. As the bicarbonate concentrations in the aquifers are largely controlled by carbonate exchange from the vadose zone, the result indicates that the aquifer recharge to the Jasper aquifer is more variable than that in Chicot and Evangeline formations. This result is borne out by the fact that precipitation patterns in the western sections (Jasper outcrop areas) tend to be more erratic and significantly less than that along the eastern sections of the study area (Norwine et al. 2007). Most parameters exhibit considerable variability with median values substantially different from the mean. With the exceptions of bicarbonate, alkalinity and silica, the coefficient of variation for all other parameters is quite large (>0.75) and in many instances greater than unity indicating that the variability within each aquifer formation is at least as great or often times significantly greater than that between the formations. The intra-formation variability is most pronounced in the Chicot formation which is to be expected because it is

Table 1 Summary statistics for wells in the Chicot aquifer

	Mean	Median	Mode	SD	Kurtosis	Skewness	Range	Minimum	Maximum
Static water level	10.24	7.01	1.22	19.25	3.02	1.44	110.83	−26.52	84.31
Calcium	88.87	72.33	26.00	105.71	50.88	6.18	1,086.85	3.15	1,090.00
Magnesium	28.55	18.00	23.00	41.41	38.34	5.47	373.50	0.50	374.00
Sodium	268.51	172.00	215.00	300.17	9.50	2.69	2,028.70	9.30	2,038.00
Potassium	4.37	3.00	2.00	3.86	14.52	3.21	27.00	1.00	28.00
Bicarbonate	347.52	342.05	350.24	93.50	1.29	−0.04	635.19	37.83	673.02
Sulfate	141.56	26.00	10.00	312.29	22.73	4.39	2,188.80	1.00	2,189.80
Chloride	357.49	211.60	182.00	509.04	22.23	4.08	4,041.52	8.48	4,050.00
Fluoride	0.74	0.57	0.40	0.61	5.87	2.20	3.51	0.10	3.61
Nitrate	3.50	0.34	0.09	10.48	28.66	5.04	84.92	0.01	84.93
pH	7.50	7.49	8.00	0.48	0.03	−0.09	3.31	5.71	9.02
TDS	1,089.10	784.50	957.00	1,178.65	18.50	3.85	8,692.00	82.00	8,774.00
Total hardness	339.98	265.00	52.00	426.32	47.89	6.09	4,124.00	15.00	4,139.00
Silica	28.33	27.00	16.00	13.58	2.95	1.44	82.10	2.90	85.00
Total alkalinity	285.64	283.00	287.00	76.76	1.29	−0.05	520.50	31.00	551.50
Specific conductance	1,930.23	1,349.00	2,624.00	2,012.80	14.45	3.40	14,142.40	0.00	14,142.40

Depth: water table depth from MSL in meters; Calcium, Magnesium, Sodium, Potassium, Bicarbonate, Sulfate, Chloride, Fluoride: dissolved in mg/L; Nitrate: Nitrate nitrogen in mg/L; pH: Standard units (field measurement); TDS: solids dissolved sum of constituents; Total hardness: total hardness; Silica: mg/L; Total alkalinity: dissolved (analyzed in lab); Specific conductance: μmhos/cm at 25 °C (field measurement). All parameters were averaged for the time period between 1980 and 2006

Table 2 Summary statistics for wells in the Evangeline aquifer

	Mean	Median	Mode	SD	Kurtosis	Skewness	Range	Minimum	Maximum
Static water level	27.64	23.96	20.73	34.07	1.41	0.87	219.34	−55.67	163.66
Calcium	81.25	58.50	113.00	76.86	12.25	2.83	598.23	4.27	602.50
Magnesium	24.29	15.20	20.00	26.87	14.34	3.06	237.20	0.80	238.00
Sodium	330.78	251.83	192.00	328.90	28.82	4.05	3,481.04	4.96	3,486.00
Potassium	7.61	6.51	2.00	6.11	4.93	1.92	38.44	0.56	39.00
Bicarbonate	308.60	306.37	360.00	78.52	3.29	0.57	623.35	72.25	695.60
Sulfate	204.21	94.00	18.00	314.39	14.28	3.39	2,193.00	0.50	2,193.50
Chloride	400.68	251.00	251.00	468.39	31.35	4.35	5,045.41	4.59	5,050.00
Fluoride	0.97	0.70	0.70	1.30	195.83	12.34	21.91	0.09	22.00
Nitrate	13.47	3.01	0.09	41.29	106.42	9.28	559.96	0.04	560.00
pH	7.63	7.70	8.00	0.48	−0.65	−0.21	2.40	6.30	8.70
TDS	1,246.44	904.00	766.00	1,088.97	20.72	3.62	10,376.53	187.47	10,564.00
Total hardness	303.64	227.00	271.00	291.21	13.16	2.93	2,474.00	20.00	2,494.00
Silica	32.47	25.74	20.00	19.05	1.59	1.49	93.35	3.00	96.35
Total alkalinity	254.41	252.50	295.00	64.70	3.16	0.58	503.90	66.10	570.00
Specific conductance	2,074.61	1,587.50	1,380.00	1,778.35	45.75	4.93	21,666.50	341.50	22,008.00

Depth: water table depth from MSL in meters; Calcium, Magnesium, Sodium, Potassium, Bicarbonate, Sulfate, Chloride, Fluoride: dissolved in mg/L; Nitrate: Nitrate nitrogen in mg/L; pH: Standard units (field measurement); TDS: solids dissolved sum of constituents; Total hardness: total hardness; Silica: mg/L; Total alkalinity: dissolved (analyzed in lab); Specific conductance: μmhos/cm at 25 °C (field measurement). All parameters were averaged for the time period between 1980 and 2006

completely unconfined and as such subject to interactions with the atmosphere as well as overlaying surface hydrologic entities (i.e., rivers, streams and the coast).

The Pearson product moment correlation coefficients for various analyte pairs are presented in Tables 4, 5 and 6 for the three different aquifer formations. For the purposes of this analysis, the labels–strong (bold), moderate (bold and italics) and weak (italics) refer to correlation values >0.75, 0.5–0.75 and 0.3–0.5, respectively (Chen-Wuing et al. 2003). The bulk parameter TDS exhibits a strong correlation with sodium in all three formations indicating the importance of halite on the overall ionic composition in the aquifer. As to be expected, the correlation of TDS with calcium and magnesium is less pronounced in the deeper formations as much of calcium is immobilized due to ion exchange. There is a moderate correlation (~0.6) between potassium and sodium in the Chicot and Evangeline formations pointing again to the possibility of ion exchange reactions. As the potassium ion has a smaller ionic radius, it is preferentially sorbed onto the clay minerals relative to sodium which in turn reduces the covariation between the two compounds (Lorite-Herrera et al. 2008). The high correlation between calcium and magnesium indicates that the carbonate minerals—calcite and dolomite likely occur in the same strata (Freeze and Cherry 1979). The correlation between calcium and sulfate is moderate in Chicot but weak in Evangeline and Burkeville–Jasper formations demonstrating the role of gypsum dissolution in the aquifer. Interestingly, none of the water quality parameters considered

exhibits strong correlations with static water level elevations which are indicative of the highly heterogeneous characteristic of the aquifer formations. Silica has a moderate positive (~0.5) correlation and nitrate exhibits a weak(0.3–0.4) correlation with the water table elevations in the Chicot and Evangeline formations signifying the possibility of intermingling of freshly recharged waters from the land surface and the vadose zone.

Hydrogeochemical classification of groundwater

Piper plots were employed to classify groundwater of the Gulf Coast aquifer units based on their major ion chemistry and are presented in Figs. 3, 4 and 5. These plots indicate that sodium and calcium are the dominant cations relative to magnesium. On the other hand, bicarbonate and chloride are more prominent than sulfate. The major water types include $Ca–HCO_3$; $Na–HCO_3$; $Ca–Na–HCO_3–Cl$. $Ca–HCO_3$ is noted to be more common in the recharge areas (along the western sections) which evolve into mixed $Ca–Na–HCO_3–Cl$; $Na–HCO_3$ or $Na–Cl–HCO_3$ or $Na–Cl–SO_4$ along the coast (Chowdhury et al. 2006). The Na/Cl molar ratio exhibits considerable variability in all the units (see Figs. 6, 7) and is most pronounced in the Jasper aquifer. These results indicate that a variety of processes affect these ions including halite dissolution, mixing, evaporation, ion exchange and weathering (Cartwright et al. 2004).

Table 3 Summary statistics for wells in the Burkeville–Jasper unit

	Mean	Median	Mode	SD	Kurtosis	Skewness	Range	Minimum	Maximum
Static water level	88.94	74.95	146.91	49.47	0.56	1.04	241.75	−11.41	230.34
Calcium	87.12	72.72	146.00	78.30	2.36	1.39	402.00	1.00	403.00
Magnesium	14.61	10.08	3.90	16.49	8.56	2.39	109.13	0.20	109.33
Sodium	343.50	260.00	330.00	345.32	9.11	2.72	2,038.50	21.50	2,060.00
Potassium	12.16	8.08	5.00	11.63	5.58	2.18	66.95	1.00	67.95
Bicarbonate	326.52	324.01	390.51	107.07	10.67	2.02	938.70	58.58	997.28
Sulfate	145.82	78.76	31.00	238.82	35.24	5.13	2,139.00	1.00	2,140.00
Chloride	436.60	258.00	132.00	505.05	9.02	2.69	2,924.25	10.75	2,935.00
Fluoride	0.92	0.72	0.50	0.65	3.61	1.68	3.83	0.10	3.93
Nitrate	13.10	3.91	0.04	20.91	7.22	2.49	114.49	0.02	114.51
pH	7.64	7.61	8.00	0.78	50.20	5.60	8.22	6.65	14.86
TDS	1,259.31	865.00	2,316.00	1,042.81	7.02	2.53	5,383.00	337.00	5,720.00
Total hardness	278.02	231.75	790.00	253.30	3.60	1.62	1,455.83	4.50	1,460.33
Silica	47.99	42.30	34.00	26.50	−0.05	0.87	117.50	3.00	120.50
Total alkalinity	270.33	267.00	214.00	88.75	10.09	1.98	761.21	56.00	817.21
Specific Conductance	2,113.48	1,525.38	1,450.00	1,646.86	7.18	2.43	9,786.20	545.80	10,332.00

Depth: water table depth from MSL in meters; Calcium, Magnesium, Sodium, Potassium, Bicarbonate, Sulfate, Chloride, Fluoride: dissolved in mg/L; Nitrate: Nitrate nitrogen in mg/L; pH: standard units (field measurement); TDS: solids dissolved sum of constituents; Total hardness: total hardness; Silica: mg/L; Total alkalinity: dissolved (analyzed in lab); Specific conductance: μmhos/cm at 25 °C (field measurement). All parameters were averaged for the time period between 1980 and 2006

Aquifer-scale principal component analysis

Principal component analysis was carried out separately for Chicot, Jasper and Burkeville–Jasper aquifers. The number of significant components (factors) in each case was ascertained from the eigenvalues plot presented in Fig. 8. As can be seen, the first four eigenvalues were typically greater than or equal to one. These factors explain more than 75 % of the variability (see Fig. 8 inset). The communalities of the variables were also inspected and the cumulative value considering the first four factors was generally high (≥0.75) for most parameters. However, the cumulative variance explained for nitrate and fluoride was moderate (∼0.5) in Evangeline and Jasper aquifers. The depth to water table had a lower communality (∼0.5) in the Jasper aquifer as well. Given the relatively large number of parameters and data records, the amount of variance explained by the first four factors was deemed sufficient and selected for further evaluation.

The varimax rotated orthogonal factor loadings of the variables on the first four factors extracted in Chicot, Evangeline and Burkeville–Jasper aquifers are presented in Tables 7, 8 and 9. As varimax preserves orthogonality, the factors are independent of each other. Oblique rotation was also carried out using a promax rotation scheme. The results of the promax rotation did not vary much from the varimax rotation. This result indicated that the promax rotated factors were approximately orthogonal as well. As

such, further analysis and interpretation were carried out using the varimax rotated factor.

Chicot aquifer

The Chicot aquifer is the youngest of all aquifer units considered here. The first factor presented in Table 7 depicts that all major cations–calcium, magnesium, sodium and potassium as well as major non-carbonate anions–sulfate and chloride load strongly on Factor 1. Not surprisingly, the bulk parameters—TDS and specific conductivity—are strongly correlated with this factor as well. Furthermore, hardness is also correlated to this factor as expected given the strong relationship of both calcium and magnesium. As sulfates and chlorides load strongly on this factor and not carbonates or bicarbonates, it can be inferred that the hardness in this aquifer is largely due to non-carbonate hardness. The high loadings of non-carbonate mineral forming ions—Na, K, SO_4 and Cl suggest that the dissolution of carbonate minerals (calcium and magnesium) are enhanced due to ionic strength effects (Freeze and Cherry 1979) and explain why the specific conductance (a measure of ionic strength) as well as both Ca and Mg load strongly on this factor as well. Given these conditions, the factor is labeled ionic strength index in this study.

The second factor is strongly correlated to bicarbonate and alkalinity and as such represents the interactions with carbonate minerals in the aquifer. Fluoride has a weak

Table 4 Pearson correlation coefficients between various analytes for wells in the Chicot aquifer

	Static water level	Ca	Mg	Na	K	HCO₃⁻	SO₄²⁻	Cl⁻	F⁻	NO₃⁻	pH	TDS	Total hardness	SiO₂	Total alkalinity	Specific conductance
Static water level	1.00															
Ca	0.25	1.00														
Mg	0.14	**0.90**	1.00													
Na	−0.05	0.45	0.67	1.00												
K	0.09	0.53	0.64	0.62	1.00											
HCO₃⁻	−0.25	−0.02	0.10	0.32	0.14	1.00										
SO₄²⁻	0.11	0.65	**0.81**	**0.80**	0.60	0.15	1.00									
Cl⁻	0.06	**0.76**	**0.87**	**0.88**	0.66	0.12	0.75	1.00								
F⁻	−0.12	−0.11	0.09	0.53	0.20	0.35	0.27	0.30	1.00							
NO₃⁻	0.44	0.18	0.24	0.15	0.18	0.06	0.21	0.16	0.14	1.00						
pH	−0.29	−0.33	−0.22	0.10	−0.05	0.04	−0.16	−0.04	0.38	−0.05	1.00					
TDS	0.07	0.74	**0.89**	**0.93**	0.68	0.22	**0.89**	**0.96**	0.35	0.21	−0.08	1.00				
Total hardness	0.21	**0.98**	**0.96**	0.55	0.59	0.03	0.73	**0.82**	−0.03	0.21	−0.30	**0.82**	1.00			
SiO₂	0.53	0.30	0.25	−0.04	0.22	0.06	0.18	0.05	−0.06	0.49	−0.27	0.11	0.29	1.00		
Total alkalinity	−0.26	−0.02	0.10	0.32	0.13	**1.00**	0.15	0.12	0.36	0.06	0.05	0.22	0.02	0.06	1.00	
Specific conductance	0.06	0.68	**0.84**	**0.94**	0.69	0.24	**0.85**	**0.95**	0.38	0.19	−0.01	**0.98**	**0.76**	0.09	0.24	1.00

Table 5 Pearson correlation coefficients between various analytes for wells in the Evangeline aquifer

|	Static water level	Ca	Mg	Na	K	HCO₃⁻	SO₄²⁻	Cl⁻	F⁻	NO₃⁻	pH	TDS	Total hardness	SiO₂	Total alkalinity	Specific conductance
Static water level	1.00															
Ca	0.43	1.00														
Mg	0.28	**0.83**	1.00													
Na	−0.11	0.30	0.54	1.00												
K	0.18	0.43	0.65	0.57	1.00											
HCO₃⁻	0.15	0.05	0.09	−0.14	0.04	1.00										
SO₄²⁻	−0.15	0.34	0.50	**0.79**	0.38	−0.32	1.00									
Cl⁻	0.09	0.59	**0.77**	**0.90**	0.67	−0.07	0.62	1.00								
F⁻	0.01	−0.02	0.13	0.30	0.18	0.02	0.24	0.23	1.00							
NO₃⁻	0.32	0.38	0.25	0.04	0.29	0.24	−0.03	0.15	−0.01	1.00						
pH	−0.37	−0.41	−0.29	0.16	−0.06	−0.15	0.07	−0.03	0.12	−0.08	1.00					
TDS	0.03	0.56	0.74	**0.95**	0.64	−0.11	**0.82**	**0.95**	0.26	0.15	0.00	1.00				
Total hardness	0.39	**0.98**	**0.93**	0.40	0.54	0.06	0.42	0.68	0.03	0.35	−0.38	0.65	1.00			
SiO₂	0.59	0.44	0.33	−0.10	0.16	0.30	−0.13	0.09	0.00	0.35	−0.29	0.05	0.41	1.00		
Total alkalinity	0.13	0.03	0.07	−0.13	0.03	**1.00**	−0.32	−0.07	0.03	0.23	−0.11	−0.11	0.05	0.29	1.00	
Specific conductance	0.04	0.46	0.61	**0.88**	0.61	−0.10	0.67	**0.88**	0.28	0.14	0.10	**0.89**	0.53	0.07	−0.10	1.00

Table 6 Pearson correlation coefficients between various analytes for wells in the Burkeville–Jasper unit

	Static water level	Ca	Mg	Na	K	HCO$_3^-$	SO$_4^{2-}$	Cl$^-$	F$^-$	NO$_3^-$	pH	TDS	Total hardness	SiO$_2$	Total alkalinity	Specific conductance
Static water level	1.00															
Ca	−0.10	1.00														
Mg	−0.03	**0.79**	1.00													
Na	0.15	0.11	0.14	1.00												
K	0.07	0.29	0.33	0.72	1.00											
HCO$_3^-$	−0.02	−0.37	−0.35	−0.10	−0.15	1.00										
SO$_4^{2-}$	0.06	0.40	0.29	0.66	0.59	−0.25	1.00									
Cl$^-$	0.11	0.39	0.43	**0.92**	0.73	−0.29	0.56	1.00								
F$^-$	0.15	−0.11	0.09	0.57	0.26	0.17	0.19	0.50	1.00							
NO$_3^-$	0.09	0.34	0.27	0.24	0.28	−0.18	0.29	0.29	0.07	1.00						
pH	0.00	−0.37	−0.27	0.06	−0.12	0.07	−0.08	−0.06	0.10	−0.17	1.00					
TDS	0.11	0.40	0.39	**0.95**	**0.77**	−0.22	**0.76**	**0.96**	0.48	0.34	−0.07	1.00				
Total hardness	−0.09	**0.99**	**0.88**	0.13	0.32	−0.38	0.39	0.42	−0.06	0.33	−0.36	0.42	1.00			
SiO$_2$	0.07	0.28	0.23	0.15	0.41	0.02	0.29	0.17	0.05	0.37	−0.28	0.26	0.28	1.00		
Total alkalinity	−0.02	−0.39	−0.37	−0.09	−0.17	**0.99**	−0.25	−0.29	0.21	−0.19	0.12	−0.22	−0.40	−0.02	1.00	
Specific conductance	0.08	0.41	0.40	**0.89**	0.70	−0.29	**0.73**	**0.91**	0.38	0.34	−0.02	**0.94**	0.42	0.18	−0.28	1.00

loading on this factor and can be explained by the fact that the dissolution of fluoride is more favored under alkaline conditions. Therefore, this factor is labeled alkalinity index in this study and represents the carbonate mineral effects on groundwater. As the second eigenvalue explains less variance than the first, the effects of carbonate minerals is less pronounced than that of the non-carbonate minerals in this aquifer. As such, the carbonate hardness is less important than the non-carbonate hardness in the aquifer.

Depth, nitrate and silica are associated with the third factor. The water table depth and nitrate concentrations both point toward recharge of freshwater and a concomitant pollutant load associated with anthropogenic activities such as agriculture and cattle raising. The silica contributions arise from the dissolution of aluminosilicate minerals (feldspars and mica). The dissolution of aluminosilicates is usually brought forth by waters that are acidic (or contain dissolved CO_2) which once again points toward freshly recharged water. This factor is, therefore, labeled recharge index here and represents the component of groundwater mixture that has recently entered the aquifer. From a mineralogical perspective, this factor represents the alterations of feldspars to clay minerals.

The fourth factor has strong correlations with fluoride and pH. Sodium has a weaker positive correlation and calcium has a weak negative correlation with this factor. The factor represents the dissolution of fluoride containing minerals (e.g., fluorite) in the aquifer. Fluorite concentrations are noted to be strongly affected by mineralogy in this aquifer (Hudak 1999). In the presence of excessive sodium bicarbonates, the dissociation of fluoride minerals is high and can be expressed by the following reactions (Saxena and Ahmed 2001):

$$CaF_2 + 2NaHCO_3 \rightarrow CaCO_3 + 2Na + 2F + H_2O + CO_2 \tag{1}$$
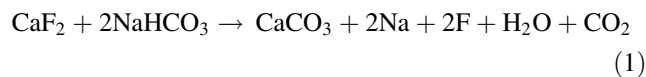
The above reaction indicates the formation of dissolved fluoride and sodium both of which have positive correlation with the factor and the precipitation of calcium through the formation of calcium carbonate which explains its negative correlation. The formation of carbon dioxide and its subsequent dissolution leading to the formation of bicarbonate has an effect on the pH. This factor which explains 6 % of the total variance is labeled acidity index due to its relationship with pH. From a mineralogical standpoint, this factor points to the reactions of fluorite in the aquifer.

Evangeline aquifer

The first factor in the Evangeline aquifer has high loadings corresponding to sodium, chloride, sulfate and magnesium and moderate loadings with respect to calcium and
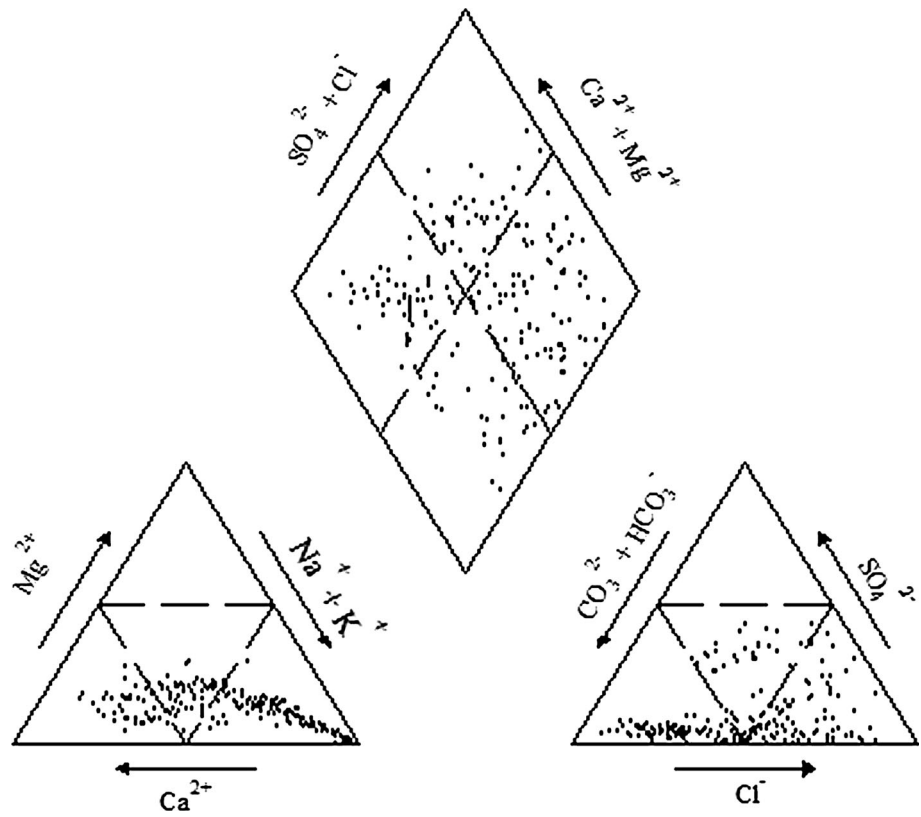
Fig. 3 Piper plot for the Chicot
aquifer

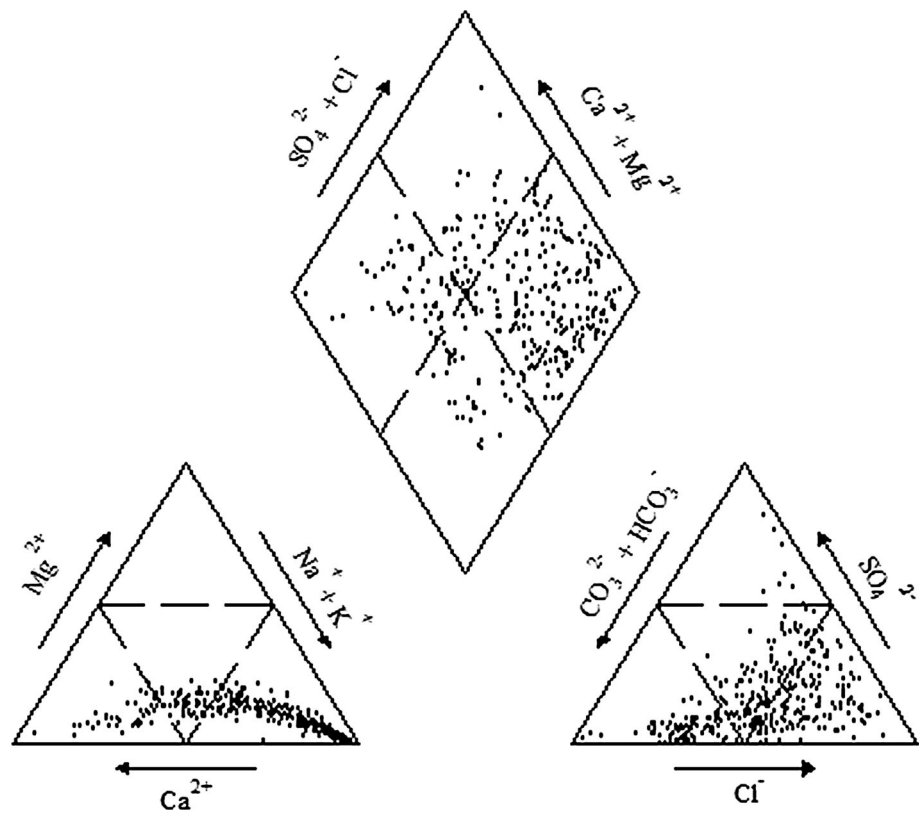**Fig. 4** Piper plot for the
Evangeline aquifer

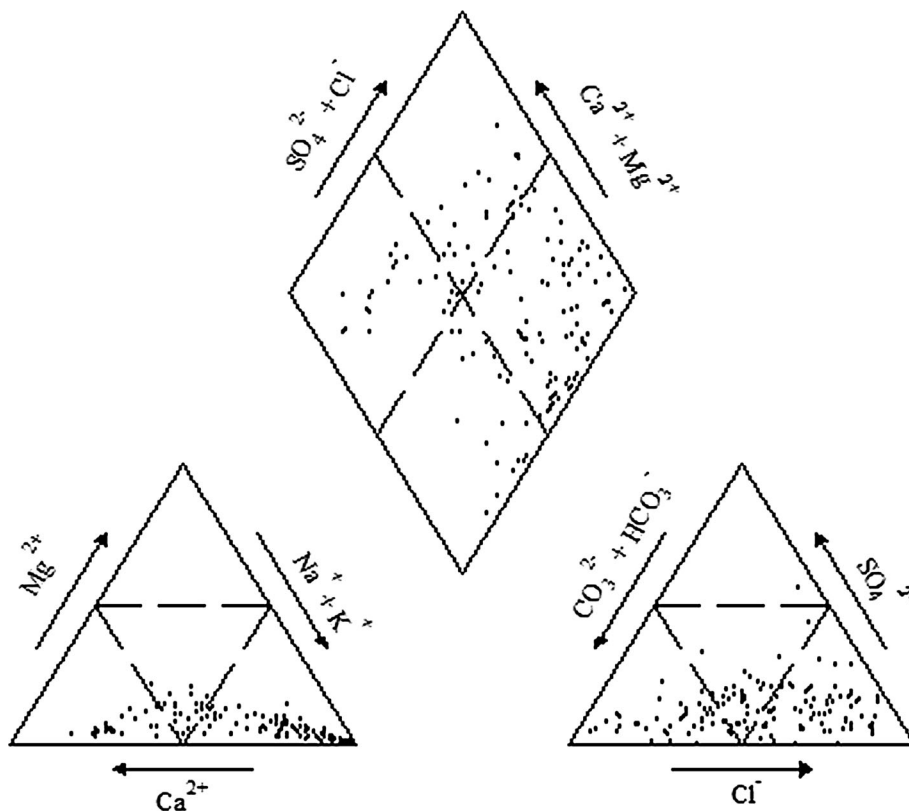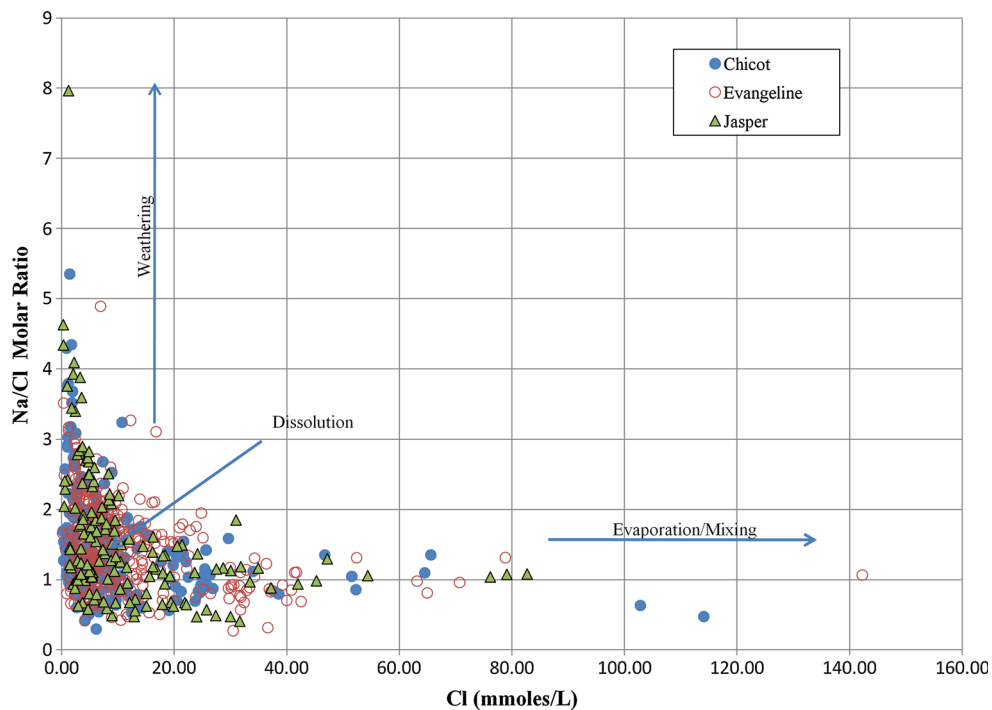**Fig. 5** Piper plot for the Burkeville–Jasper unit

Fig. 6 Variability of Na/Cl molar ratio and associated processes

potassium. The bulk parameters, TDS and specific conductance, also correlate strongly with this factor. This factor predominantly represents the effects of non-

carbonate minerals and the moderate loadings of hardness represent the non-carbonate hardness in the system. The concentrations of sodium, sulfate and chloride in the

Evangeline aquifer are considerably higher than that in the Chicot aquifer. As such, the ionic strength of the solution (as measured using specific conductance or total dissolved solids) is also high. Therefore, the common ion effect that facilitates the dissolution of calcium and magnesium from carbonate minerals is to be expected. However, given that groundwater in the Evangeline aquifer is at a greater depth, the amount of calcium and magnesium to be found in the strata are lesser than that in the Chicot aquifer due to ion exchange reactions which retard the mobility of calcium and magnesium in favor of smaller sodium and potassium ions. Principal component analysis results indicate that while the dissolution of carbonate minerals cannot be ruled

Fig. 7 Variability and outliers of the Na/Cl molar ratio for each aquifer

out, their effects are somewhat diminished due to limited availability of such minerals. The factor is, as such, labeled ionic strength index and seeks to explain the effects of non-carbonate minerals.

The second factor is strongly correlated to groundwater levels and silica and moderately correlated with nitrate and calcium. The factor also exhibits a weak correlation with pH and a moderate correlation with hardness. These effects point toward the notion that the factor is attempting to group waters that have recently interacted with the land surface. As stated previously, a significant portion of the Evangeline aquifer is overlain by caliche (calcium carbonate) deposits (Chowdhury and Turco 2006) and explains the moderate loading of calcium. Furthermore, waters passing through these deposits can accumulate carbon dioxide which affects the pH. The interactions of these acidic waters with mineral matter, particularly the large deposits of feldspars in the aquifer (Hoel 1982) cause the leaching of silica and helps explain the noted strong correlation. Therefore, the hardness associated with this factor is the carbonate hardness. The carbonate hardness is also inversely linked to pH as it helps neutralize the acid in waters (Vesilind and Morgan 2004). This relationship is borne out in the factor loadings as hardness exhibits a negative correlation while pH is positively correlated. The strong correlations of non-lithological parameters i.e., groundwater levels and nitrate also point toward recharging waters. As such this factor is labeled recharge index in this study.

The interpretation of the third factor is relatively straightforward as it explains the interactions of carbonate
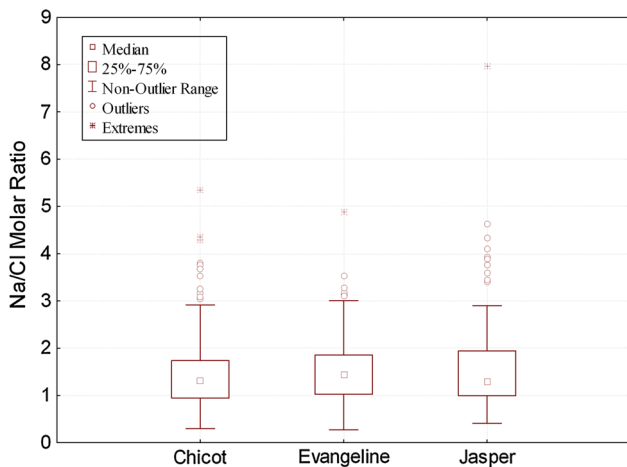
Fig. 8 Eigenvalue plot of significant components with Kaiser rule cutoff for each aquifer
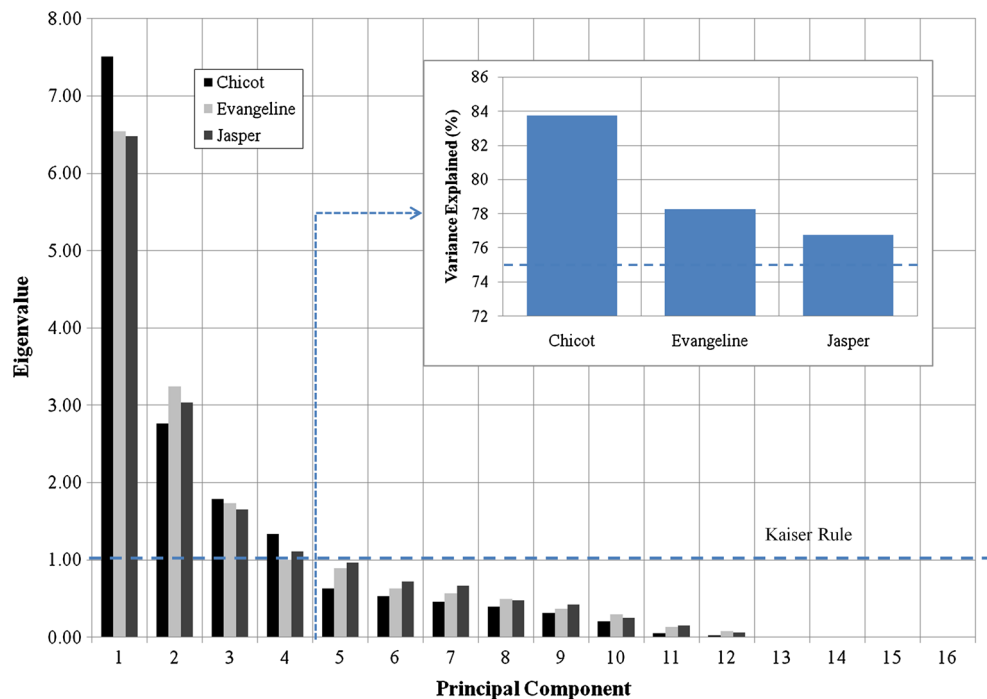
Table 7 Factor loadings for wells in the Chicot aquifer

	Factor 1	Factor 2	Factor 3	Factor 4
Static water level	−0.06	−0.27	**−0.78**	−0.16
Calcium	**−0.83**	−0.08	−0.16	*−0.42*
Magnesium	**−0.93**	0.02	−0.14	−0.20
Sodium	**−0.85**	0.22	0.04	*0.38*
Potassium	**−0.72**	0.07	−0.14	0.06
Bicarbonate	−0.10	**0.98**	0.02	0.07
Sulfate	**−0.87**	0.08	−0.11	0.04
Chloride	**−0.96**	0.02	0.01	0.10
Fluoride	−0.24	*0.31*	−0.06	**0.77**
Nitrate	−0.14	0.05	**−0.81**	0.18
pH	0.12	−0.06	0.19	**0.77**
TDS	**−0.98**	0.12	−0.05	0.10
Total hardness	**−0.89**	−0.04	−0.15	*−0.34*
Silica	−0.09	0.12	**−0.82**	−0.24
Total alkalinity	−0.10	**0.98**	0.03	0.08
Specific conductance	**−0.95**	0.14	−0.04	0.18

Table 9 Factor loadings for wells in the Burkeville–Jasper unit

	Factor 1	Factor 2	Factor 3	Factor 4
Static water level	0.11	*0.46*	−0.21	**−0.45**
Calcium	0.20	**−0.86**	−0.26	−0.14
Magnesium	0.26	**−0.77**	−0.25	−0.09
Sodium	**0.98**	0.09	−0.03	−0.07
Potassium	**0.75**	−0.17	−0.02	−0.28
Bicarbonate	−0.11	0.18	**0.95**	0.02
Sulfate	*0.68*	−0.20	−0.18	−0.24
Chloride	**0.93**	−0.17	−0.18	−0.04
Fluoride	*0.63*	0.16	0.22	0.01
Nitrate	0.20	−0.19	−0.13	**−0.67**
pH	0.15	**0.73**	−0.03	0.22
TDS	**0.96**	−0.18	−0.12	−0.14
Total hardness	0.23	**−0.88**	−0.27	−0.14
Silica	0.13	−0.27	0.16	**−0.77**
Total alkalinity	−0.10	0.21	**0.94**	0.04
Specific conductance	**0.91**	−0.16	−0.22	−0.08

Table 8 Factor loadings for wells in the Evangeline aquifer

	Factor 1	Factor 2	Factor 3	Factor 4
Static water level	−0.03	**−0.84**	−0.02	−0.11
Calcium	*0.59*	*−0.54*	−0.02	*−0.49*
Magnesium	**0.79**	−0.36	0.08	*−0.33*
Sodium	**0.92**	0.15	−0.05	0.23
Potassium	*0.69*	−0.27	0.08	0.06
Bicarbonate	−0.07	−0.16	**0.98**	−0.03
Sulfate	**0.79**	0.17	−0.28	0.05
Chloride	**0.95**	−0.09	−0.01	0.00
Fluoride	0.26	−0.19	0.01	*0.68*
Nitrate	0.13	**−0.59**	0.20	0.08
pH	0.00	*0.34*	−0.06	*0.65*
TDS	**0.98**	−0.03	−0.05	0.04
Total hardness	**0.69**	−0.50	0.02	*−0.45*
Silica	0.01	**−0.79**	0.18	−0.10
Total alkalinity	−0.07	−0.14	**0.98**	−0.01
Specific conductance	**0.90**	−0.05	−0.04	0.17

minerals in the aquifer and is labeled alkalinity index. However, fluoride and pH have moderately strong loadings (>0.65). Also, calcium, magnesium and hardness have weak to moderate loadings on this factor (−0.3 to −0.49). Again cations and hardness are inversely related while pH and fluoride are directly correlated to this factor. Therefore, the factor is labeled as acidity index and is noted to explain a little over 6 % of the total observed variance. As was the case in Chicot aquifer, this factor groups parameters affecting the dissolution of fluoride minerals.

Burkeville–Jasper aquifer units

The first factor in the Burkeville–Jasper unit has strong loadings with sodium, potassium and chloride. Sulfate and fluoride have moderate loadings with this factor. The strong loadings of sodium and chloride indicate that the ionic strength is predominantly influenced by brines. This result is to be expected because the waters in this unit are significantly older than the overlaying Evangeline and Chicot aquifer. The moderate loadings of fluoride along with high concentrations of sodium and chloride and an insignificant correlation with pH and calcium indicate that the fluoride sources are likely attributable to the leakage of brines more so than pH-controlled mineral reactions. As seen from Fig. 6, weathering (i.e., rock–water interactions) is likely more dominant than evaporation and mixing in explaining these ions. The factor is, therefore, labeled ionic strength index and is controlled by non-carbonate minerals in the aquifer unit.

The second factor is largely defined by calcium and magnesium ions. As to be expected, hardness has a strong correlation to this factor along with pH. These loadings suggest that the factor is influenced by carbonate mineralogy (i.e., the dissolution of calcite and dolomite) which significantly affects the pH. These geological factors point toward the exchange of water from the overlying Evangeline aquifer or areas which, in turn, are overlain by calcium carbonate deposits (caliche). This reasoning is supported by the fact that the factor has almost moderate correlation with groundwater levels. This factor is labeled hardness index here and from a mineralogical standpoint represents the influences of carbonate minerals particularly calcite and dolomite.

Fig. 9 Spatial location of wells having an impact on each PCA factor for the Chicot aquifer
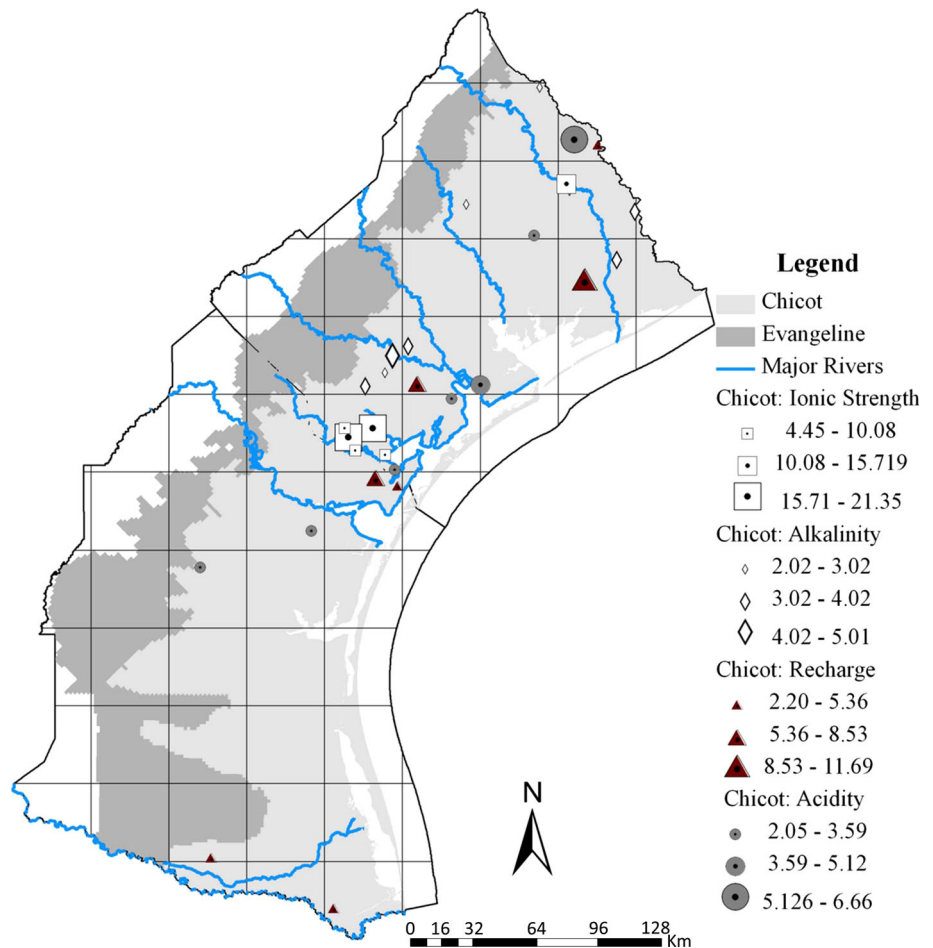
The third factor separates the effects of bicarbonate and alkalinity and as such is labeled alkalinity index. The fourth factor has a strong correlation with silica and moderate correlations with nitrate and static groundwater elevations and as such potentially represents recharging water from the surface. The factor, as before, is interpreted to represent processes affecting the groundwater that has recently migrated into the aquifer unit from the land and as such is labeled recharge index.

Well-scale PCA analysis

The aquifer-scale PCA evaluation enables the grouping of water quality parameters according to their origin and evolution characteristics. As PCA is based on geochemical records obtained at individual wells, the contribution of measurements at each well on the factors can also be ascertained. The measurements at all wells cumulatively explain the observed variation in the data. As such, the contribution of the measurement set at a particular well to the amount of variance explained by the factor (i.e., eigenvalue) can be determined. In most instances, a small number of wells contribute significantly toward a factor.

The percent contribution of each well toward each factor was computed, and those wells having at least 2 % contribution toward a factor were retained for further analysis. Figures 9, 10 and 11 depict the spatial location of wells that have sufficient impact on each factor for the three aquifer units under consideration. An attempt was also made to evaluate whether regional-scale geographic data provided any additional clues with regards to the established factors. The influencing wells did not exhibit correlations with surface soil texture (STATSGO) and land use land cover (LULC) datasets. This result is to be expected because the resolution of these datasets is much coarser (in the order of hundreds of meters) than the radius of influence of the well (in the order of tens of meters).

However, it is evident from Figs. 9, 10 and 11 that most of the influencing wells for the recharge index are in the outcrop regions of their respective aquifers. In addition, the depths to the static water levels are lowest for this factor in all the aquifers providing a strong evidence for interactions with the surface (Table 10). The correlation of factors with bulk (or field measurable) parameters was also evaluated at these influencing wells. As can be seen from Table 10, the correlation between factors and bulk parameters is not

Fig. 10 Spatial location of wells having an impact on each PCA factor for the Evangeline aquifer
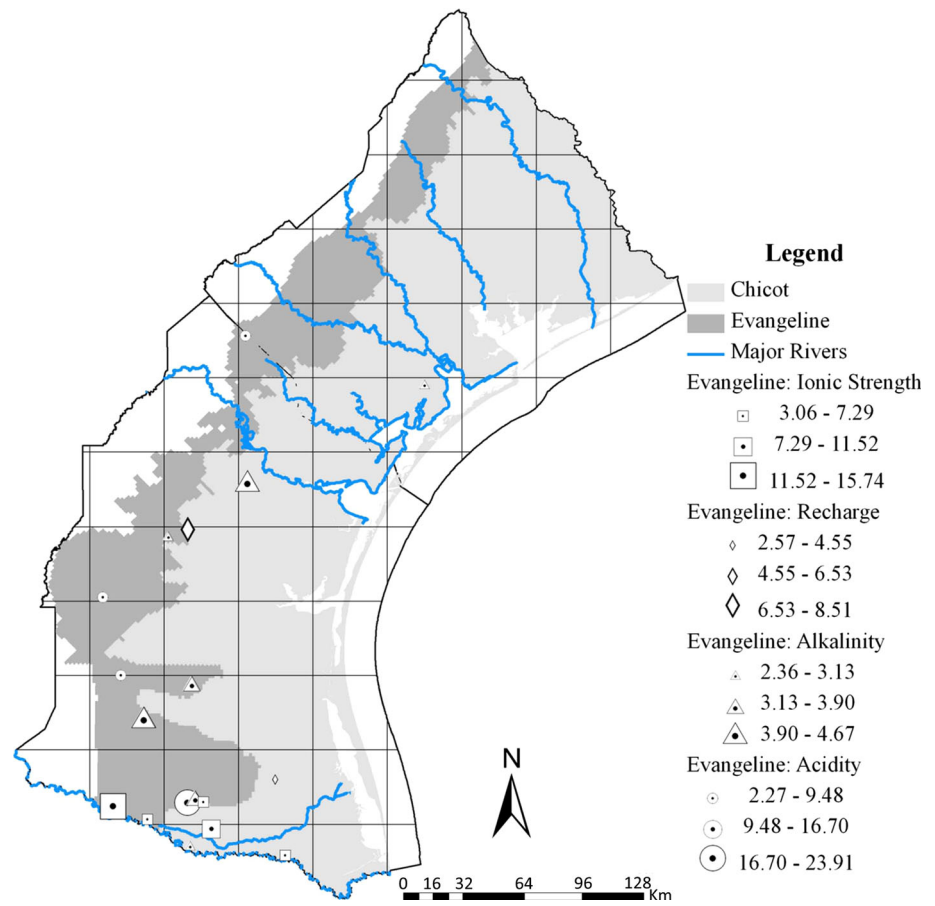
straightforward indicating that the bulk measures are affected by mixing of waters from different origins and heterogeneous geochemical reactions. Sampling of bulk parameters while inexpensive may not provide insights in the origin and evolution of groundwater in the aquifer. Clearly, there is a need for detailed multivariate geochemical information to separate out the potential sources of groundwater. From a monitoring perspective, the influencing wells depicted in Figs. 9, 10 and 11 represent sampling locations with the highest priority. The proximity of many influencing wells to major surface water bodies points toward the prominent role of the surface water–groundwater interactions in defining the solution chemistry of the aquifer. Additional sampling and investigations in and around the influencing wells are undoubtedly necessary to identify the reasons for their impact.

Summary and conclusions

Principal component analysis (PCA) was utilized to identify the potential source and evolution of groundwater in the various aquifer units of the Gulf Coast aquifer in South Texas using a large spatially extensive multivariate dataset. The results of the study indicated certain general trends in

all the aquifers. Over 75 % of the observed variance was explained by the first four factors in the three aquifers considered here. The first factor, named ionic strength index, generally consisted of the influence of sodium, potassium, sulfate and chloride and represented older water in the aquifers that have been subjected to cation exchange and weathering reactions. Bulk parameters, total dissolved solids (TDS) and specific conductance had high correlations with this factor. Principal component analysis also grouped bicarbonate and total alkalinity into a separate group. This group, designated as the alkalinity index, explained a greater amount of variance in the younger Chicot aquifer than in the older Evangeline and Burkeville–Jasper aquifers. Furthermore, PCA clustered nonlithological parameters namely static water levels and nitrate into a separate group. This group represented waters that entered the aquifer more recently from the land surface and was labeled the recharge index. The dissolution of fluorite was seen to have an influence on the pH in the Chicot and Evangeline aquifers. The calcium concentrations in the Evangeline and Burkeville–Jasper aquifers were greatly controlled by the overlying caliche deposits. The dissolution of calcium carbonate affected the pH in the Burkeville–Jasper unit. Carbonate hardness is more significant in Evangeline, Burkeville–Jasper aquifers while

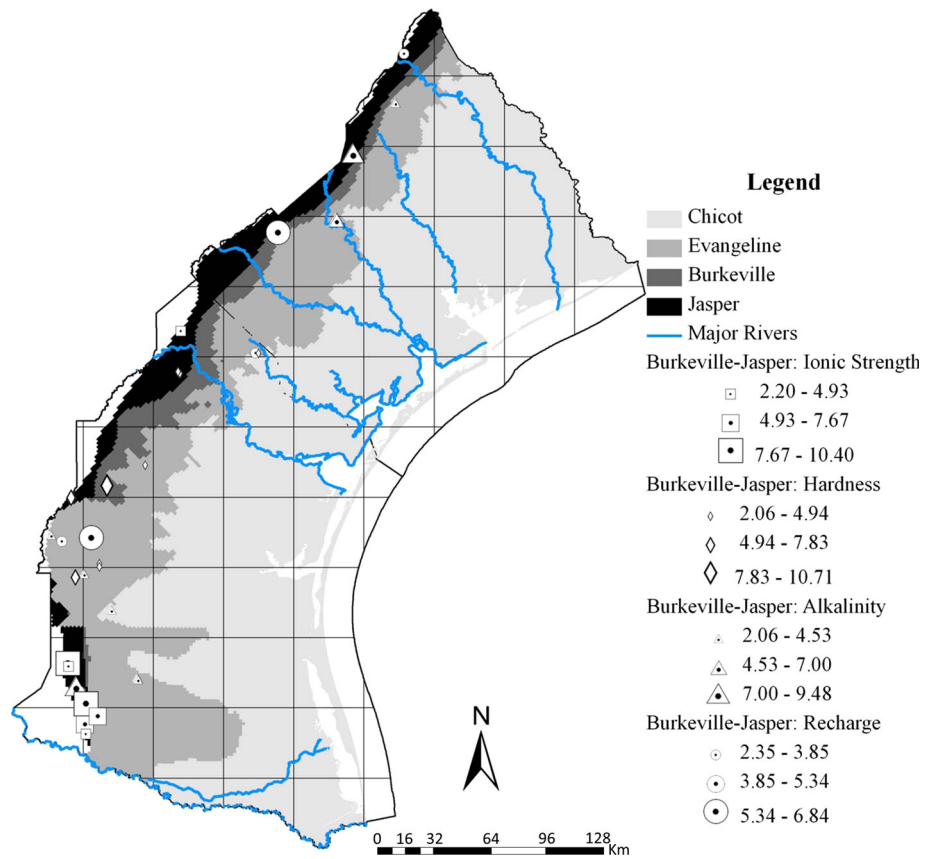**Fig. 11** Spatial location of wells having an impact on each PCA factor for the Burkeville–Jasper unit

Table 10 Major ion chemistry in the Gulf Coast aquifer units

	Na/Cl	Ca + Mg–SO$_4$–0.5HCO$_3$	Static water level	Excess sodium	Ratio
Chicot					
Ionic strength	1.39	−1.79	−38.56	3.04	−1.70
Alkalinity	1.57	−2.04	−42.18	1.15	−0.57
Recharge	1.15	−5.90	−18.21	2.79	−0.47
Acidity	1.19	−2.18	−26.34	0.45	−0.20
Evangeline					
Ionic strength	1.03	−1.92	−30.18	3.55	−1.85
Recharge	1.93	−1.44	−5.87	13.71	−9.54
Alkalinity	0.95	−1.97	−42.90	−1.03	0.52
Acidity	1.24	−2.12	−29.26	1.26	−0.59
Burkeville–Jasper					
Ionic strength	1.24	−4.53	−128.10	7.91	−1.75
Recharge	1.13	0.67	−107.20	−1.98	−2.97
Alkalinity	2.15	−3.42	−123.85	5.71	−1.67
Acidity	1.61	−2.09	−103.34	4.38	−2.09

non-carbonate hardness has a greater influence in the Chicot aquifer. Principal component analysis can also help identify wells that have the greatest influence on each of the factors. These influencing wells point toward areas that need additional monitoring and further evaluation. The spatial locations of the influencing wells highlight the importance of surface water–groundwater interactions in defining the solution chemistry of the aquifer. Well-scale PCA also highlighted the necessity of multivariate sampling to understand the evolution of groundwater. In conclusion, PCA in conjunction with other statistical techniques such as exploratory data analysis (EDA) and

correlation analysis is advantageous for conceptualizing the geochemical evolution of groundwater. Principal component analysis is also useful to identify critical wells where additional monitoring is warranted and in conjunction with GIS can be beneficial to evaluate the influence of geographic factors on the geochemical behavior of the aquifer.

References

Bakari SS, Aagaard P, Vogt RD, Ruden F, Johansen I, Vuai SA (2012) Delineation of groundwater provenance in a coastal aquifer using statistical and isotopic methods Southeast Tanzania. Environ Earth Sci 66(3):889–902

Baker ET (1979) Stratigraphic and hydrogeologic framework of part of the coastal plain of Texas. Texas Department of Water Resources, Austin

Bicalho CC, Batiot-Guilhe C, Seidel JL, Exter S, Jourde H (2012) Hydrodynamical changes and their consequences on groundwater hydrochemistry induced by three decades of intense exploitation in a Mediterranean Karst system. Environ Earth Sci 65(8):2311–2319

Borgognone MG, Bussi J, Hough G (2001) Principal component analysis in sensory analysis: covariance or correlation matrix? Food Qual Prefer 12:323–326

Cartwright I, Weaver TR, Fulton S, Nichol C, Reid M, Cheng X (2004) Hydrogeochemical and isotopic constraints on the origins of dryland salinity, Murray Basin, Victoria, Australia. Appl Geochem 19:1233–1254

Chen-Wuing L, Kao-Hung L, Yi-Ming K (2003) Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. Sci Total Environ 313:77–89

Chowdhury AH, Mace RE (2004) Geochemical evolution of the groundwater in the Gulf Coast aquifer of South Texas in Proceedings of Groundwater flow understanding from local to the regional scales. In: XXXIII Congress International Association of Hydrogeologists and 7th Asociación Latinoamericana de Hidrologia Subterránea para el Desarrollo, Zacatecas City, Mexico, p 4

Chowdhury AH, Turco MJ (2006) Geology of the Gulf Coast Aquifer, Texas. In: Mace RE, Davidson SC, Angle ES, Mullican WF (eds) Aquifers of the Gulf Coast of Texas. Texas Water Development Board Report 365, Austin, pp 23–50

Chowdhury AH, Boghici R, Hopkins J (2006) Hydrochemistry, salinity distribution, and trace constituents: implications for salinity sources, geochemical evolution, and flow systems characterization, Gulf Coast Aquifer, Texas. In: Mace RE, Davidson SC, Angle ES, Mullican WF (eds) Aquifers of the Gulf Coast of Texas. Texas Water Development Board Report 365, pp 81–128

Connell LD, van den Daele G (2003) A quantitative approach to aquifer vulnerability mapping. J Hydrol 276:71–88

Dalton MG, Upchurch SB (1978) Interpretation of hydrochemical facies by factor analysis. Ground Water 16(4):228–233

Duffy CJ, Brandes D (2001) Dimension reduction and source identification for multispecies groundwater contamination. Contam Hydrol 48:151–165

Fernandes PG, Carreira PM, Bahir M (2010) Mass balance simulation and principal components analysis applied to groundwater resources: Essaouira basin (Morocco). Environ Earth Sci 59(7):1475–1484

Freeze RA, Cherry JA (1979) Groundwater. Prentice Hall, New Jersey

Gabrysch RK, Bonnet CW (1975) Land-surface subsidence in the Houston–Galveston region. Texas Water Development Board, Austin

Galloway WE, Ganey-Curry PE, Li X, Buffler RT (2000) Cenozoic depositional history of the Gulf of Mexico Basin. Am Assoc Pet Geol Bull 84:1743–1774

Gogu RC, Dassargues A (2000) Current trends and future challenges in groundwater vulnerability assessment using overlay and index methods. Environ Geol 39(6):549–559

Grande JA, Gonzalez A, Beltran R, Sanchez-Rodas D (1996) Application of factor analysis to the study of contamination in the aquifer system of Ayamonte-Huelva (Spain). Ground Water 34(1):155–161

Güler C, Thyne GD, McCray JE, Turner KA (2002) Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeol J 10(4):455–474

Hamilton LC (1992) Regression with graphics: a second course in applied statistics. Brooks/Cole Publishing Co., Pacific Grove

Hem JD (1970) Study and interpretation of the chemical characteristics of natural water. 3rd edn. US Geological Survey. Water-Supply Paper, Austin

Hildebrandt A, Guillamon M, Lacorte S, Tauler R, Barcelo D (2008) Impact of pesticides used in agriculture and vineyards to surface and groundwater quality (North Spain). Water Res 42(13):3315–3326

Hoel HD (1982) Goliad formation of the south Texas Gulf Coastal Plain: regional genetic stratigraphy and uranium mineralization. The University of Texas, Austin

Hopkins J (2010) Groundwater monitoring activities of the Texas Water Development Board. http://www.twdb.state.tx.us/GwRD/HEMON/GMSA.asp. Accessed 30th March 2010

Hosman RL (1996) Regional stratigraphy and subsurface geology of Cenozoic deposits, Gulf Coastal Plain, South- Central United States-Regional aquifer system analysis-Gulf Coastal Plain. US Geol Surv Prof Pap 1416:G1–G35

Hudak PF (1999) Chloride and nitrate distributions in the Hickory aquifer, Central Texas, USA. Environ Int 25(4):393–401

Lawrence FW, Upchurch SB (1982) Identification of recharge areas using geochemical factor analysis. Ground Water 20(6):680–687

Lorite-Herrera M, Jimenez-Espinosa R, Jimenez-Millan J, Hiscock KM (2008) Integrated hydrochemical assessment of the quaternary alluvial aquifer of the Guadalquivir River, southern Spain. Appl Geochem 23(8):2040–2054

Lucas L, Jauzein M (2008) Use of principal component analysis to profile temporal and spatial variations of chlorinated solvent concentration in groundwater. Environ Pollut 151(1):205–212

Mason C (1963) Water resources of Refugio County. Texas Water Commission, Austin

McCuen RH (1993) Microcomputer applications in statistical hydrology. Prentice Hall, New Jersey

McGuire JT, Long DT, Hyndman DW (2005) Analysis of recharge - induced geochemical change in a contaminated Aquifer. Ground Water 43(4):518–530

Melloul A, Collin M (1992) The principal components statistical method as a complementary approach to geochemical methods in quality factor identification: application to the Coastal Plain aquifer of Israel. J Hydrol 140:49–73

Mor S, Singh S, Yadav P, Rani V, Rani P, Sheoran M, Singh G, Ravindra K (2009) Appraisal of salinity and fluoride in a semi-arid region of India using statistical and multivariate techniques. Environ Geochem Health 31(6):643–655

Norwine J, Harriss R, Yu J, Tebaldi C, Bingham R (2007) The problematic climate of South Texas 1900–2100. In: Norwine J, John K (eds) The changing climate of South Texas 1900–2100: problems and prospects, impacts and implications. pp 15–55

NRC (1993) Groundwater vulnerability assessment, contamination potential under conditions of uncertainty. National Academy Press, Washington DC

Powers SE, Villaume JF, Ripp JA (1997) Multivariate analyses to improve understanding of NAPL pollutant sources. Ground Water Monitor Rev 130–140

Saxena VK, Ahmed S (2001) Dissolution of fluoride in groundwater: a water-reaction study. Environ Geol 40:1084–1087

Sellards EH, Adkins WS, Plummer FB (1932) The geology of Texas, Bulletin 3232. The University of Texas at Austin, Bureau of Economic Geology, Austin

Stuart M, Maurice L, Sapiano M, Micallef Sultana M, Gooddy D, Chilton P (2010) Groundwater residence time and movement in the Maltese islands: a geochemical approach. Appl Geochem 25(5):609–620

Suk H, Lee K–K (1999) Characterization of a ground water hydrochemical system through multivariate analysis: clustering into ground water zones. Ground Water 37(3):358–366

Todd DK, Mays LW (2005) Groundwater Hydrol. Willy & Sons Inc., New York

TWDB (2003) Groundwater availability of the central Gulf Coast aquifer: Numerical simulations to 2050, Central Gulf Coast. Texas Water Development Board, Austin

TWDB (2007) Water for Texas. Texas Water Development Board, Austin

Uddameri V (2005) Sustainability and groundwater management. Clean Technol Environ Policy 7(4):231–232

Uddameri V (2007) Systems analysis for sustainable aquifer management in coastal semi-arid South Texas. Environ Geol 51:883–884

Uddameri V, Honnungar V (2007) Combining rough sets and GIS techniques to assess aquifer vulnerability characteristics in the semi-arid South Texas. Environ Geol 51:931–939

Uddameri V, Kuchanur M (2007) Simulation-optimization approach to assess groundwater availability in Refugio County, TX. Environl Geol 51(6):921–929

Usunoff EJ, Guzman–Guzman A (1989) Multivariate analysis in hydrochemistry: an example of the use of factor and correspondence analyses. Ground Water 27(1):27–34

Vesilind PA, Morgan SM (2004) Introduction in environmental engineering, 2nd edn. Brooks/Cole, Belmont

Zekster S, Loáiciga HA, Wolf JT (2005) Environmental impacts of groundwater overdraft: selected case studies from the southwestern United States. Environ Geol 47(3):396–404