communications earth & environment

Article

A Nature Portfolio journal

https://doi.org/10.1038/s43247-025-02456-1

Groundwater salinization patterns in the Yucatan Peninsula reveal contamination and vulnerability of the karst aquifer

Christian Narvaez-Montoya [1], Rogelio Mondragón Bonilla [2], Nico Goldscheider [3] & Jürgen Mahlknecht [1]

The Yucatan Peninsula is one of the largest coastal and lowland karst regions worldwide. This groundwater-dependent region is highly vulnerable to contaminants that spread easily due to the karst environment. Here, the spatio-temporal patterns of major ions in 1528 water samples sourced from a government institution suggest the main factors triggering salinization in the aquifer system. The hydrogeochemical analysis, supported by dimensional reduction and network-based clustering, linked one-third of the samples to contamination outbreaks from seawater intrusion, extensive gypsum dissolution in the south, and nitrate pollution in the ubiquitous carbonate aquifer matrix. Temporal variations of water quality indicated changes in regional recharge trends and increasing human impact in recent decades. Moreover, ~23% of water samples from human-use sources exceeded acceptable sulfate and nitrate limits for drinking water purposes. The study underscores the need for continuous water quality monitoring and enhanced regional knowledge to support management plans.

The Yucatan Peninsula is one of the world's largest coastal and lowland karst regions, spanning ~160,000 km^2 across southeastern Mexico, northern Guatemala, and Belize[1–3] (Fig. 1a). The area with limited surface run-off has a tropical climate with precipitation and temperature around 1000 mm year^{-1} and 25 °C, respectively[4,5] (Fig. 1b; Supplementary Fig. 1). The peninsula's karst environment sustains a range of groundwater-dependent ecosystems, including biodiverse tropical forests and wetlands. It also features the world's longest underwater cave system and encompasses 23 Ramsar Convention-protected sites[6,7]. The region's historical significance is further underscored by the ancient Maya civilization, which incorporated the underwater system into its mythology and faced its decline, in part, due to prolonged droughts before the arrival of the Spanish colonizers[8,9].

Karst environments result from the chemical dissolution of primarily carbonate rocks, increasing groundwater salinization and feeding back the complex underground conduit systems[3,10]. In the Yucatan Peninsula, groundwater flows through a range of pathways, from ubiquitous small-scale fractures up to regional fault-fractured zones (Fig. 1b, c). The conduit system also includes large caves and thousands of cenotes, which are natural sinkholes resulting from the collapse of carbonate rocks, exposing groundwater to the atmosphere[2]. Cenotes are abundant in the coastal plain at elevations below 40 m.a.s.l. (meters above sea level)[11] (Fig. 1c), in which Cenozoic sequences of limestone and dolomite are predominant to more than 250 m depth[12–14] (Fig. 1b; Supplementary Fig. 2). In contrast, the stratigraphy of the Elevated Interior Region (EIR)[15], ranging from 40 to 300 m.a.s.l., remains poorly studied. However, the dissolution of gypsum layers up to 35 m thick that outcrop within the carbonate environment has been linked to the formation of enclosed basins with flat floors known as karst poljes (Fig. 1b, c), as well as large subsurface conduits, low permeability due to residual sediments, and perched water bodies[16–19]. It is suggested that this gypsum member underlies carbonate succession to more than 200 m in the northern part of the peninsula[17] (Fig. 1b; Supplementary Fig. 2).

The Yucatan karst features high contamination vulnerability due to the easy spreading of substances through the multi-scale flow paths[3,10]. Groundwater in the coastal plain, with hydraulic heads less than 5 m.a.s.l, is constrained by an extensive seawater wedge. Salinity profiles from specific sites in the zone indicate seawater underlying the aquifer at less than 90 meters depth with a mixing interface of up to 40 m[20–23]. Samples collected from UNAM 2 and UNAM 5 sites to more than 200 m.b.s.l. (meters below sea level) were associated with groundwater beneath the diffusion zone[24] (Fig. 1b). Moreover, electromagnetic measurements suggest that the halocline can be modeled with the Ghyben–Herzberg principle at the regional scale[19], which posits that the depth of the seawater-groundwater interface below sea level is 40 times the groundwater hydraulic head above sea level (Fig. 1c). Conversely, high concentrations of Ca^{2+} and SO_4^{2-} in the south

[1]Escuela de Ingenieria y Ciencias, Tecnologico de Monterrey, Monterrey, Mexico. [2]Dirección General de Geografía y Medio Ambiente, Instituto Nacional de Estadística y Geografía, Aguascalientes, Mexico. [3]Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany. e-mail: jurgen@tec.mx

have been linked to the dissolution of the gypsum layers within the poorly studied Paleocene Icaiche formation in the EIR[17] (Fig. 1b). Hydraulic heads up to 250 m.a.s.l. in the EIR indicate a regional recharge flowing towards the coastal plain[5] (Fig. 1c).

Natural salinization and inadequate public service coverage complicate aquifer management and water supply for over 5 million inhabitants[25]. In 2020, around 121,000 people lacked access to potable water, while 291,000 had no sewage service[26]. For many, the aquifer is a crucial water source

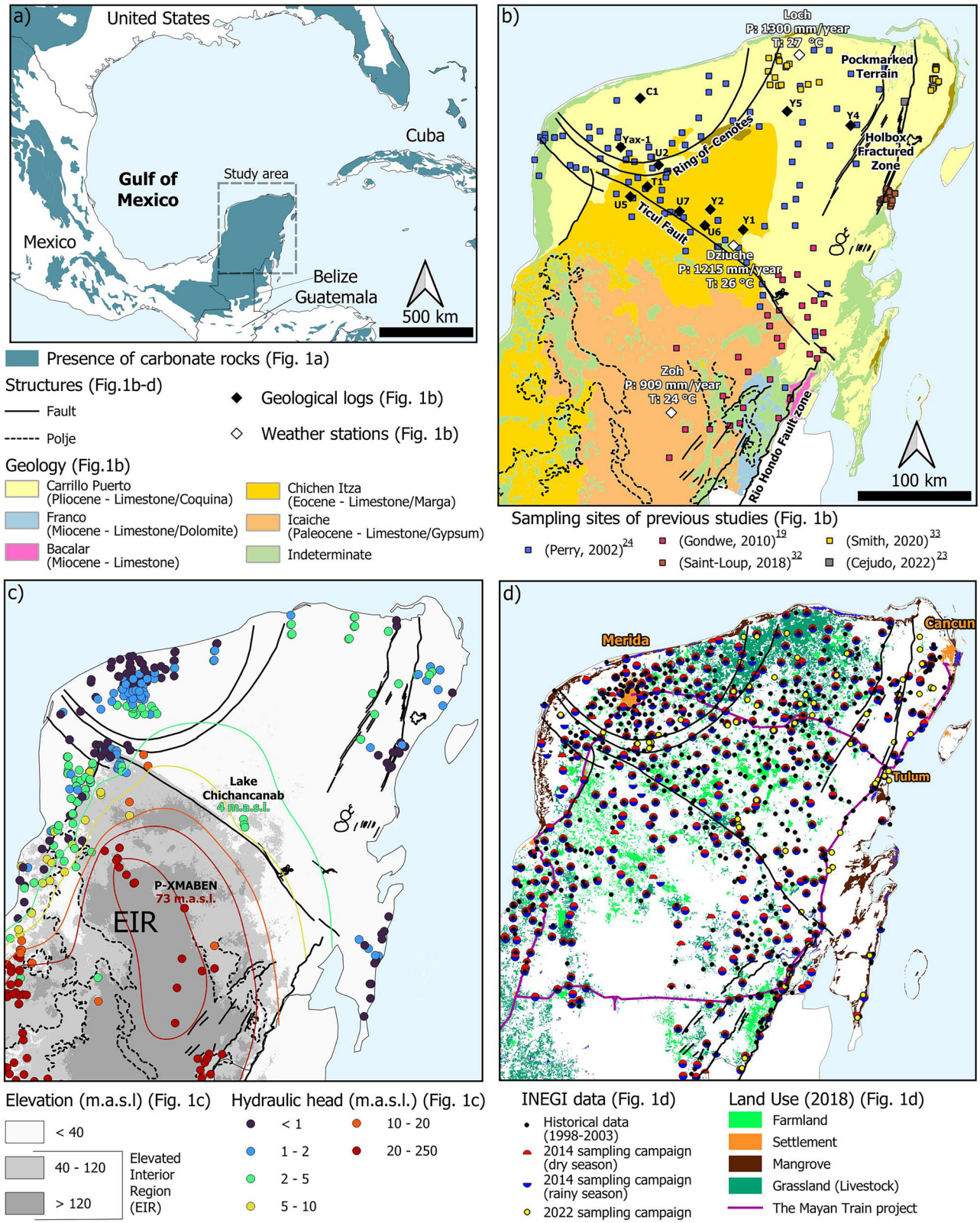

Fig. 1 | Yucatan Peninsula. a Presence of carbonate rocks in the Gulf of Mexico[3] and location of the Yucatan Peninsula. **b** Mexican geology[78], water sampling sites of previous studies[19,23,24,32,33], climatological data[4] (Supplementary Fig. 1) and geological logs[14] (Supplementary Fig. 2). **c** Elevation and hydraulic head records[79,80]. Contour lines, achieved through kriging interpolation, provide a rough idea of the hydraulic head distribution. **d** Land use in the Mexican territory[57] and INEGI water sampling sites[36].

through cenotes and dug wells, and it also serves as a disposal site for household sewage. Even those with reported access to water services remain at risk, as groundwater from tube wells may not meet quality standards[25,27,28]. Additionally, 21% of collected wastewater is discharged untreated into the aquifer[26]. Although the aquifer is considered subexploited[5], pumping for public supply and irrigation can induce localized seawater upconing (Fig. 1d). These pressures are further exacerbated by over 26 million tourists annually[29]. The Mayan Train project[30], a railway network that aims to boost regional development (Fig. 1d), is expected to increase the population to 8 million and tourist numbers to 36 million by 2035 following the anticipated project completion in 2025[29]. This development may intensify anthropogenic pressures on the peninsula.

Addressing water quality deterioration in the Yucatan Peninsula requires an enhanced understanding of the hydrogeological system[2,31]. However, the multiscale variabilities in hydrogeochemical conditions modified by different regimes and compositions of flows in the karst environment make a proper description of processes challenging[10]. Furthermore, the lack of regional water quality data and insufficient knowledge of the geological settings in the EIR have limited the understanding of the Yucatan system. In local studies, researchers have utilized water quality data to assess pollution and water quality concerns, highlighting major ions (hydrogeochemistry) as valuable tracers[19,23,24,32–35]. Despite these contributions, a comprehensive regional-level water quality analysis remains lacking, limiting the understanding of the complex dynamics of the vast karst aquifer system. Recently, the Mexican National Institute of Statistics and Geography (INEGI) began publishing extensive water quality records for the Peninsula, including major ion concentrations. Utilizing pattern recognition in this data, coupled with comparison to previous studies, offers the potential to clarify regional hydrogeological processes associated with salinization and inform future water management plans.

Here, the spatio-temporal hydrogeochemical patterns of the Yucatan Peninsula are recognized by analyzing 1528 water quality samples provided by the INEGI[36]. For this, network-based clustering was used to group similar water quality samples, dimensional reduction to explore salinization trends, traditional hydrogeochemical criteria to assess the resulting groups, and inferential statistics to suggest changes in dominant processes over time. The results are compared with major ion datasets from smaller-scale studies within the peninsula that mark local salinization patterns[19,23,24,32,33]. Although the karst environment poses challenges in the identification of hydrogeological processes due to the inherent heterogeneity, the large amount of data analyzed enables the identification of specific trends. This study marks a comprehensive assessment of water quality across the entire peninsula, highlighting one-third of the samples as contamination outbreaks linked to seawater intrusion, gypsum dissolution from the Paleocene Icaiche formation within the EIR, and nitrate pollution. Most of the remaining samples represent the influence of the ubiquitous carbonate dissolution interacting with multiple processes. Additionally, temporal variations of major ions indicated changes in regional recharge trends and increasing human impact in recent decades in the southern part of the peninsula.

Results and discussion

Pattern recognition

Unsupervised techniques were applied to explore patterns within the INEGI dataset[36,37], allowing for the clustering of similar water samples and exploring distinctive trends within the multivariate space. Fourteen clusters of water samples were defined through network-based clustering. With this, samples are connected to the nearest ones, making it easy to relate the samples and the groups (Fig. 2a; Supplementary Fig. 3). Regarding PCA,-four linear environmental trends explain 89% of the INEGI dataset (Fig. 2b–g). Clusters C9, C10, C11, C12, and C14, accounting for 561 samples, mark the greatest impact on the dataset variance. Meanwhile, C1, C2 and C4, accounting for 578 samples, are strongly related and have little effect on the variance. C3, C5, C6, and C7 are tiny clusters without an appreciable pattern. For simplicity, clusters C1 to C7 were grouped for further analysis.

The first principal component (PC1), which accounts for 39% of the explained variance, is influenced by all the major ions, except K^+ and HCO_3^- (Supplementary Table 1). Samples' scores on this component have a high correlation with TDS ($r = 0.98$), highlighting that this direction represents a general salinization process[38]. PC2, explaining 24% of the variance, is characterized by the moderate influence of SO_4^{2-}, Ca^{2+}, and HCO_3^-. This component is associated with karst rock minerals dissolution, as carbonate and sulfate minerals contribute to these ions and are ubiquitous in the peninsula[17,19,24]. Moreover, the most scattered samples plotted on the first two PCs indicate two distinctive trends triggering salinization: an increase in Na^+ and Cl^- suggests seawater-groundwater mixing[32,39,40], and an increase in Ca^{2+} and SO_4^{2-} indicates gypsum dissolution[19,24,41] (Fig. 2b, c). PC3 has a high influence of NO_3^- explaining 15% of the variance (Fig. 2d, e). The PC associated with NO_3^- usually indicates anthropogenic pollution from sewage, manure, or fertilizer leaching, as documented in several hydrogeochemical studies[42–46]. Finally, PC4, accounting for 11% of the variance, exhibits a strong influence of HCO_3^-.

From a hydrogeochemistry perspective, the relative contribution of major ions in the Piper diagram indicates groundwater types as calcium-bicarbonate, calcium-sulfate and sodium-chloride (Figs. 3a, b; 4a). Seawater fraction values, regarding the ideal mixing between rainwater and seawater, reach up to 15% with a median of 7.6% (Figs. 3c, d; 4a, b). Since anthropogenic sources and halite dissolution may contribute Cl^-, caution must be taken in evaluating this indicator[47]. Furthermore, previous studies have highlighted the milliequivalent ratio $100{\cdot}SO_4^{2-}/Cl^-$ as a valuable indicator for major ion source differentiation within the Yucatan Peninsula[19,24] (Figs. 3e, f; 4a, c). Samples with a ratio close to 10.3 (seawater ratio) suggest seawater-groundwater mixing, while ratios >100 indicate evaporite layers dissolution. Median saturation indices indicate most of the groundwater is in equilibrium with calcite ($SI_{Calcite} = 0.16$) and dolomite ($SI_{Dolomite} = 0.05$), and undersaturated regarding halite ($SI_{Halite} = -1.83$) and gypsum ($SI_{Gypsum} = -6.6$). The spatial distribution of halite and gypsum has similar patterns to the seawater fraction and $100{\cdot}SO_4^{2-}/Cl^-$ ratio, respectively (Supplementary Fig. 5).

Based on the multivariate and spatial patterns of hydrogeochemistry, the predominant hydrogeological and anthropological processes, along with their associated triggering factors, were inferred (Fig. 4a, e; Table 1). Owing to the karstic nature of the Yucatan Peninsula, water samples are the result of a complex mixture of overlapping processes, resulting in a heterogeneous water quality pattern (Fig. 3). In this analysis, the water quality clusters were associated with processes that contributed the most to the variance of the INEGI dataset. This does not mean that samples are exclusive to individual processes or unaffected by other processes outside the scope of this study, such as biochemical reactions, dedolomization, cation exchange, among others[10]. Furthermore, the discussion is limited to the post-Cretaceous carbonate successions in the northern part. Although water sample extraction depths are unknown, wells are drilled up to 100 m in the peninsula[48], while the limestone and dolomite successions are deeper than 200 m[14] (Figs. 1b; 4e). On the other hand, the stratigraphy south of the Ticul fault is unknown, apart from the outcrops of carbonate and gypsum rocks of the Paleocene Icaiche formation.

Seawater intrusion outbreaks, represented by the linked C9 and C10 samples, are predominant in the coastal zones; gypsum dissolution outbreaks, represented by the linked C12 and C14 samples, are within and around the EIR; whereas nitrate pollution outbreaks (C11) are ubiquitous (Figs. 2a; 4a). Although (C1–C7) reflect mainly calcium-bicarbonate facies, explaining the excess of ions relative to the ideal mixture (Fig. 3c). The lack of a higher correlation between the excess of Ca^{2+} + excess Mg^{2+} and excess of HCO_3^- ($r = 0.53$) denotes the influence of sources other than carbonate dissolution and seawater intrusion. The contribution of gypsum dissolution, nitrate pollution, and seawater-groundwater mixing at a lesser intensity than the contamination outbreak clusters partially explains the scattered hydrogeochemical pattern. C13 reflects a strong relation to C9, C12 and C1 (Fig. 2a). This may refer to a specific pattern indicating recharge from the EIR with

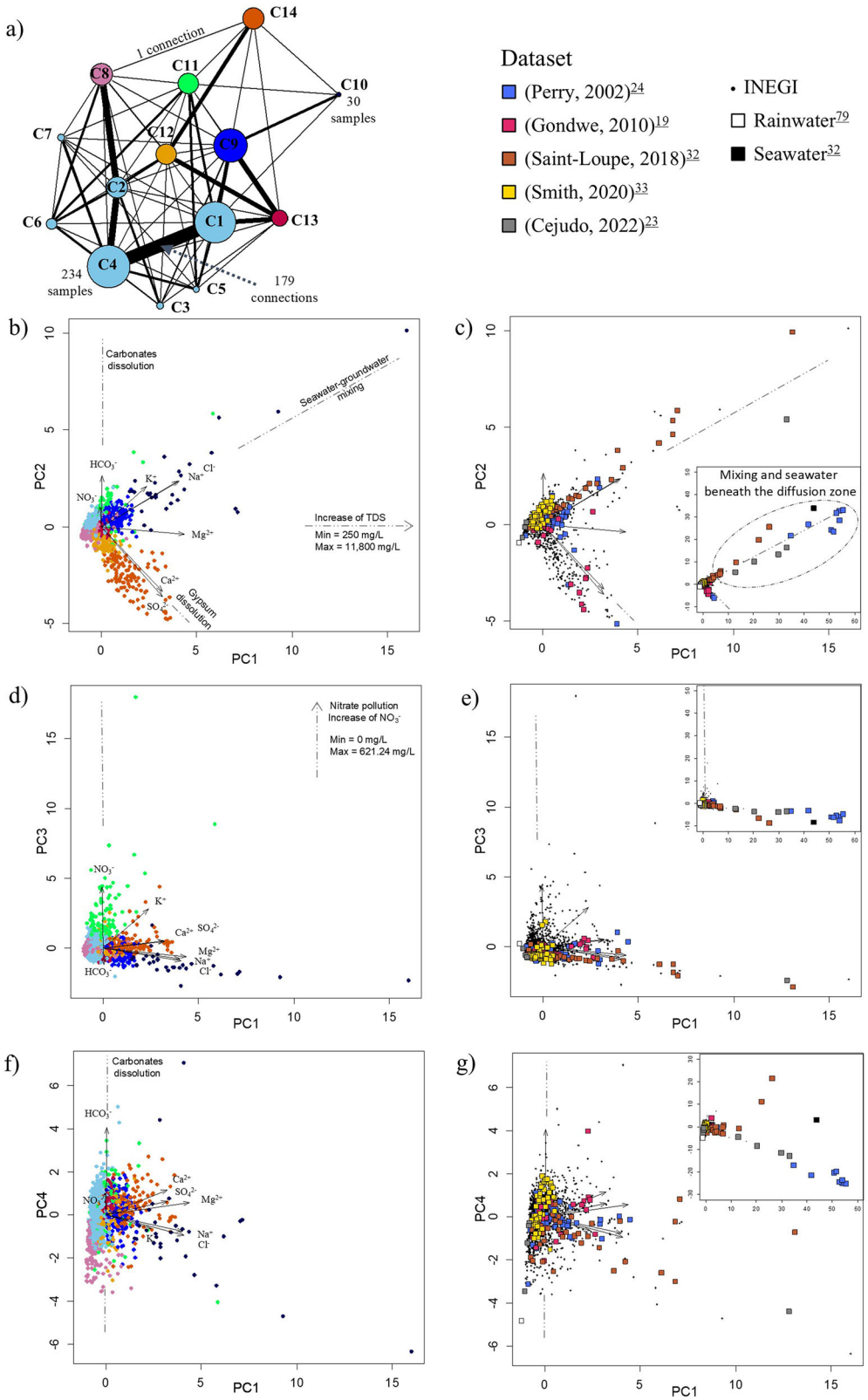

Fig. 2 | Water quality multivariate pattern of the Yucatan Peninsula. a Simplified Edge Betweenness on the Nearest Neighbor Graph of the INEGI dataset, considering nodes as the clusters defined by the community detection algorithm: nodes size represents the number of samples within the cluster and edge size represents the number of linked samples. The full representation of the network is shown in Supplementary Fig. 3. **b**, **d**, **f** Principal components biplots of the INEGI dataset, illustrating the cluster membership of the samples. Loadings and correlation coefficients associated with the PCA are included in Supplementary Table 1. **c**, **e**, **g** Principal components biplots of INEGI dataset, showing the projection of previous studies' datasets onto the INEGI data multivariate space.

calcium-sulfate facies type to the lower-elevation area with predominantly carbonate rocks through large fault-fractured zones which serve as preferential flow channels (e.g., Ticul, Rio Hondo, Holbox-fractured zones) (Fig. 4a). Another interesting group is C8, which is more related to C1, with samples tending to a lower water-rock interaction but with a higher dispersed hydrogeochemical pattern (Figs. 2a; 4a; 3).

The spatial distribution of variables suggests the influence of main processes in individual samples. Seawater fractions with values higher than

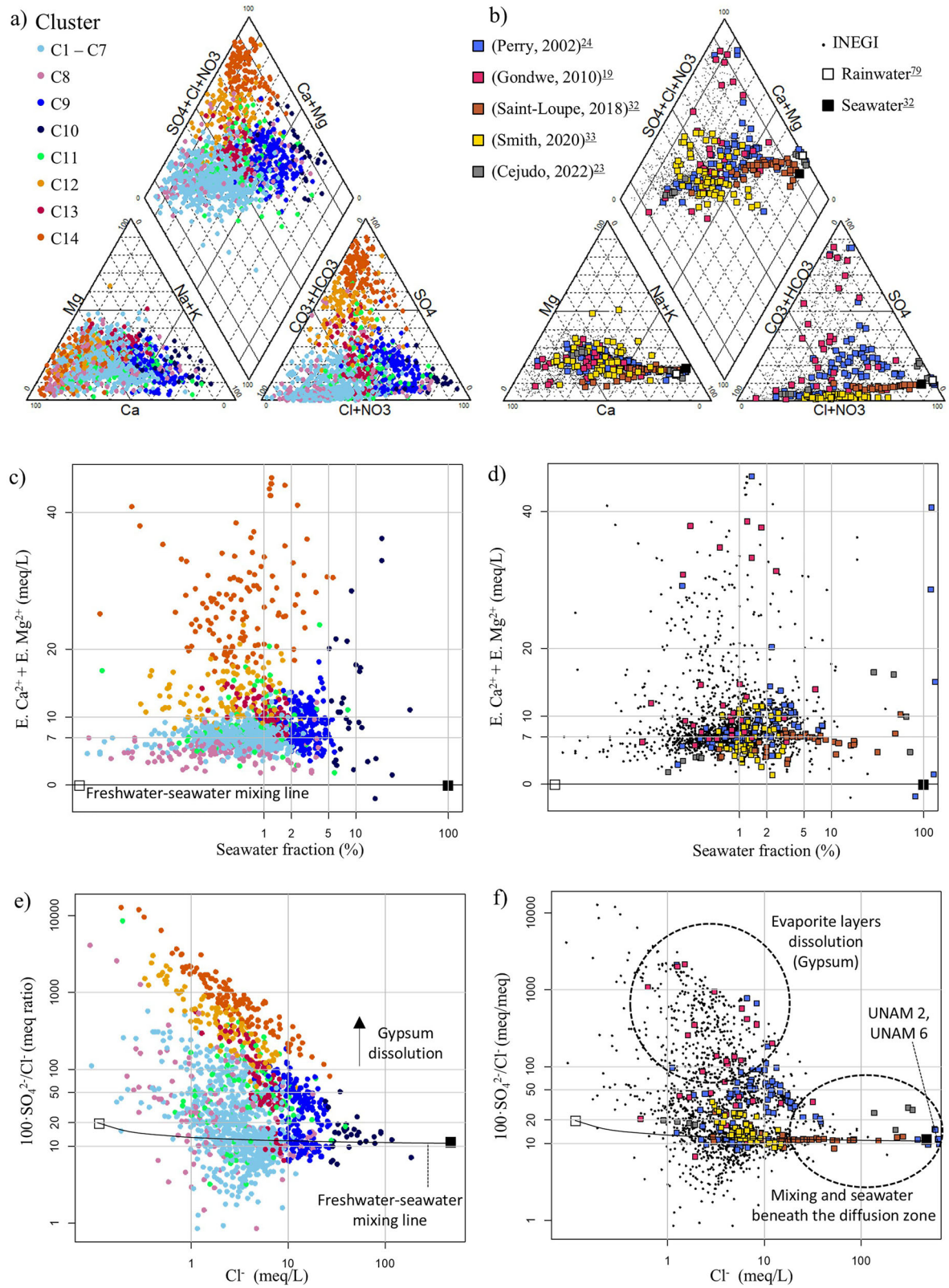

Fig. 3 | Ion relationships. a Piper diagram for INEGI dataset clusters. **b** Piper diagram for the previous studies. **c** Seawater fraction vs excess (E.) of cations regarding freshwater-seawater mixing for INEGI dataset clusters. **d** Seawater fraction vs excess (E.) of cations regarding freshwater-seawater mixing for previous studies. **e** Cl^- vs $100{\cdot}SO_4^{2-}/Cl^-$ ratio for INEGI dataset clusters. **f** Cl^- vs $100{\cdot}SO_4^{2-}/Cl^-$ for previous studies.

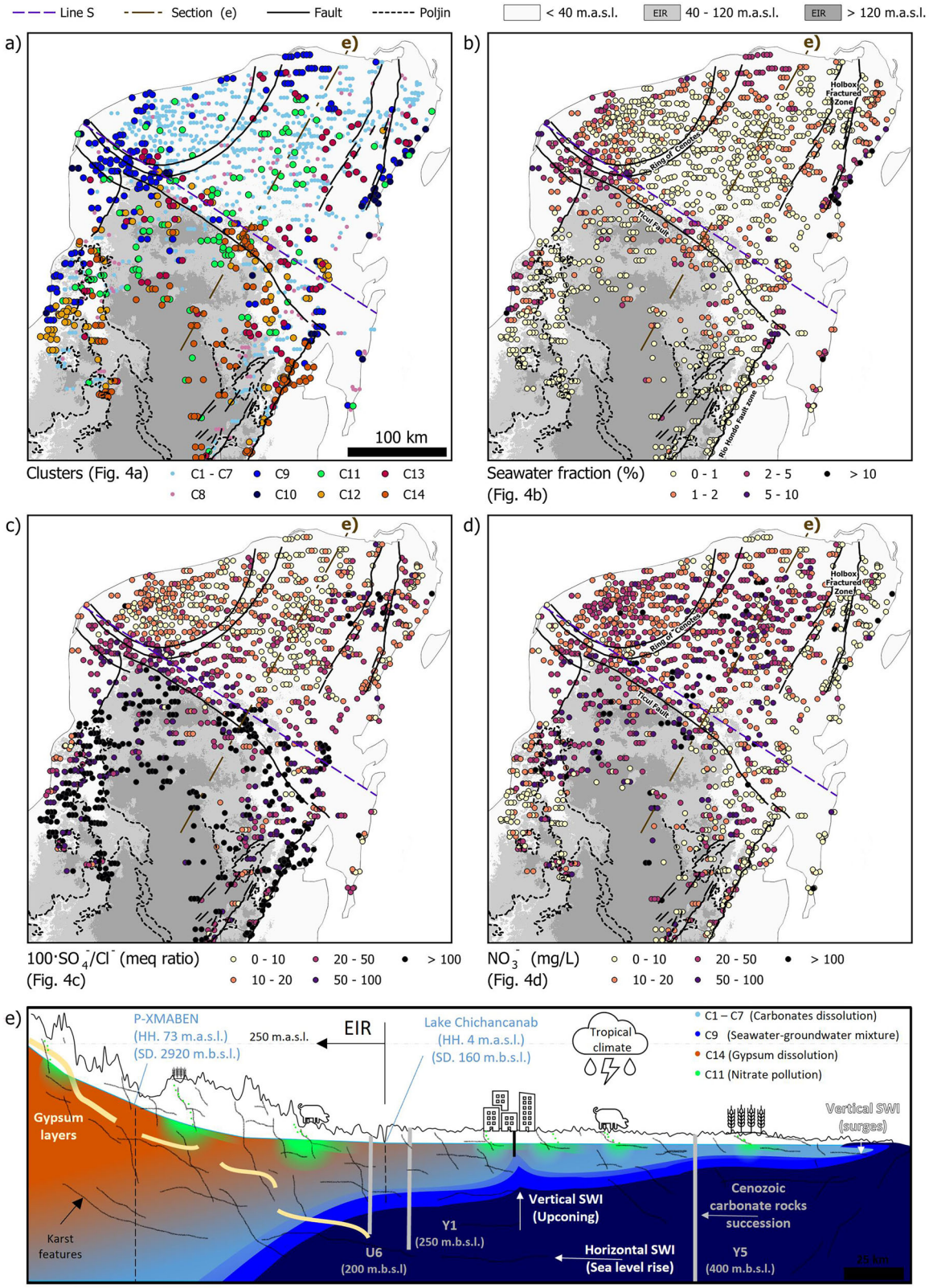

Fig. 4 | Groundwater salinization patterns of the Yucatan Peninsula. **a** Occurrence of clusters. Occurrence of clusters by sampling campaign can be seen in Supplementary Fig 4. **b** Seawater fraction. **c** $100{\cdot}SO_4^{2-}/Cl^-$ ratio. **d** NO_3^- concentrations. **e** Conceptual sections of the aquifer system in the Yucatan Peninsula, highlighting the processes in Table 1. HH. refers to hydraulic head, and SD. refers to seawater depth. A representation of the seawater wedge was constructed with the hydraulic heads' records in Lake Chichancanab and P-XMABEN sites (Fig. 1c), and the Ghyben–Herzberg principle[2]. Line S delineates the division of samples for hypothesis testing by region (northern and southern) in Table 2.

1% are located in zones with low topography and low hydraulic heads (Figs. 4b; 1c), consistent with the samples related to seawater-groundwater mixture (Fig. 4a, e). However, samples close to the P-XMABEN site, with a recorded hydraulic head of 73 m.a.s.l. and a presumed depth to the seawater interface below 2900 m.b.s.l, may not reflect modern seawater intrusion through the actual water mining technology. The predominance of gypsum dissolution outbreaks (Fig. 4a, e) in the EIR suggests that residual halite from the evaporite layers could explain these high Cl^- values in some water samples[49]. On the other hand, anthropogenic inputs, marked by the occurrence of nitrate pollution outbreaks (Fig. 4a, d, e), could also contribute to Cl^- release as a local factor[47].

Regarding the $100{\cdot}SO_4^{2-}/Cl^-$ distribution, values greater than 100 occur mainly within and around the EIR with some high values in the Holbox Fractured Zone (Fig. 4c). This predominance of sulfate over chloride suggests that the gypsum dissolution is regionally associated with the EIR and the Icaiche formation, as stipulated in previous studies[16,17,19,49]. Furthermore, ratios greater than 50 on the coastal plain could suggest the regional recharge from the EIR, and values higher than 100 in the Holbox Fractured Zone may indicate preferential subsurface flow as marked by water samples in C13 (Figs. 3a, e; 4a, c). A buried gypsum layer of 35 m in U6 located beneath more than 200 m.b.s.l. of the carbonate rocks succession has been correlated with the Icaiche formation, which plunges to the north[14,17]. Due to this, gypsum dissolution may not occur locally in the exploitable portion of the aquifer further north of the Ticul fault, but much further below the seawater wedge[14,17] (Fig. 4a, e).

The INEGI dataset reveals the Yucatan Peninsula has been heavily affected by nitrate pollution regionally, with concentrations reaching up to 621.2 mg L^{-1} (Fig. 4d). Sixty percent of the water samples exceed 13.8 mg L^{-1}, the threshold used to define the natural nitrate background of NO_3^- in the peninsula[50]. Pollution is as scattered as the anthropogenic activities in a karst system (Figs. 1d; 4e), in which heterogeneous hydrological features facilitate the infiltration, mixing and transport[10,31] (Fig. 4e). Anthropogenic contamination influences deviations in the water quality pattern affecting the mineral dissolution and groundwater-seawater mixing relationships[47,51], contributing to the heterogeneous pattern of water quality in the Yucatan Peninsula (Fig. 3). Other anthropogenic contaminants recorded on the peninsula that may accompany NO_3^- leaching include bacteria, viruses, fecal sterols, polycyclic aromatic hydrocarbons, pesticides, microplastics and other emerging contaminants[11,28,52,53].

Temporal variation

To elucidate long-term and seasonal changes in water quality, differences in variables associated with the main drivers of salinization were tested by contrasting the historical (1998–2003) to the 2014 data and the 2014 rainy season to the 2014 dry season data (Table 2). These comparisons were made in the northern and southern regions of the Peninsula, divided by an imaginary line S parallel to the Ticul fault (Fig. 4). This division was based on the observed geochemical and hydrogeological differences. The northern region is characterized by elevations less than 40 m.a.s.l. with recorded hydraulic heads around 5 m.a.s.l., and the influence of Eocene-Pliocene carbonate rocks, predominantly showing water samples associated with carbonate dissolution (C1–C7) (Figs. 1b, c; 3a; 4a, e). Meanwhile, the southern region recorded hydraulic heads up to 250 m.a.s.l. within the EIR and exhibited most of the outbreaks of gypsum dissolution (C12 and C14) (Figs. 1b, c; 3a; 4a, e). The $100{\cdot}SO_4^{2-}/Cl^-$ ratio spatial distribution highlights the regional differentiation (Figs. 3e; 4c).

Significant long-term and seasonal calcite and dolomite saturation indices variations were observed for the northern region and long-term variations for the southern region. These changes can be explained by the fact that carbonate dissolution/precipitation kinetics and dilution/concentration vary with the hydrogeological conditions at different time scales in a carbonate karst system[10]. For instance, influenced by seasonal and interannual meteoric recharge trends in the peninsula (Fig. 1b; Supplementary Fig. 1), mixing with other sources like seawater[54], or matrix porosity/permeability and water composition alterations due to the groundwater flow[10]. Results from a local flow model suggest there is a short residence time in the north of the peninsula with flow velocities in the order of one kilometer per month[55], supporting temporal hydrogeological changes. Non-seasonal significant variations in carbonates saturation indices are suggested in the southern part of the peninsula, possibly associated with the less dynamic system related with the low permeability of the Icaiche formation[17,19].

Although variations in SI_{gypsum} are not suggested in the southern region, where gypsum dissolution occurs, a significant long-term change is inferred for the northern region. One explanation for this could be changes in the regional recharge patterns with calcium-sulfate groundwater type from the EIR (Figs. 3a, e; 4a, c, e). Other explanations can come from the seawater-groundwater mixture, increasing the SI_{gypsum} value. However, non-variation in the seawater fraction indicated no changes relative to the seawater-groundwater mixing in all the comparisons, assuming seawater as the major Cl^- contributor in water quality. Meanwhile, there was no significant long-term change in NO_3^- in the northern region, with both medians above the background level of 13.6 mg L^{-1} (20.34 and 19.02 mg L^{-1}), but a significant increase was observed in the southern region with medians of 14.99 and 20.88 mg L^{-1}. The expansion of anthropogenic activities in the southern region in the early 2000s[17,56,57] in contrast to the already-established development in the northern region before the 1990s could explain this difference (Supplementary Fig. 6). This is because high sulfate content has constrained the settlement within the EIR, although the Mexican government's efforts in the last decades have tried to expand the anthropogenic activities in the zone[17].

Implications for water management. Water quality poses substantial challenges for water management in the Yucatan Peninsula. Although 85% of the samples in the INEGI dataset came from sources categorized for domestic, potable, and urban uses[36], many of these failed to meet WHO quality recommendations[58]. Specifically, TDS exceeded 1000 mg L^{-1} in 42% of the samples, a level beyond which water may taste unpleasant. The primary factors that trigger the high salinization values in groundwater are seawater-groundwater mixing and gypsum dissolution (Fig. 2b; Table 1). With a projected sea level rise of 20 cm on the peninsula by 2100[59], horizontal seawater intrusion could advance the seawater wedge hundreds of meters or even kilometers inland[22,60], reducing the thickness of the usable aquifer and increasing the vulnerability of upcoming seawater intrusion (Fig. 4e). Moreover, climate change may disrupt hydrogeological patterns in the karst environment, affecting the rate of carbonate and gypsum dissolution. In turn, ~11% of samples exceeded the recommended 500 mg L^{-1} for SO_4^{2-}, primarily due to gypsum dissolution from the Paleocene limestone-gypsum Icaiche formation. This issue has constrained human settlements and delayed agricultural development in the southern region[17]. The release of high SO_4^{2-} concentrations in the complex network of perched aquifers and preferential conduits in the zone could abruptly affect local water quality, necessitating ongoing monitoring of drinking water sources[16].

Regarding NO_3^-, 13% of the samples exceeded the 50 mg L^{-1} limit. This contamination mirrors the distribution of human activities on the peninsula (Figs. 1d; 4d). Sewage discharge and fertilizer leaching mainly affect the upper part of the water table, where some populations are supplied with water from cenotes, springs, and dug wells without further treatment, making these sources particularly vulnerable to anthropogenic contamination[11,28,52,53]. The national water authority recognizes the contamination issues and potential exacerbating factors[25], including climate change and population growth driven by the Maya Train project. Projections indicate that these drivers could strain water resources beyond 100% of their current availability by 2050 while also increasing anthropogenic pollution. The hydrogeochemistry spatial pattern of this study can inform the authority's future water quality sampling campaigns, focusing on monitoring seawater intrusion, nitrate pollution, and gypsum dissolution across the Yucatan peninsula. In particular, monitoring in areas with the highest occurrence of contamination outbreaks, like the intersection between the

Table 1 | Overview of clusters' pattern recognition

Clusters	Predominant processes	Triggering factors	Supporting observations
C1–C7	Carbonate minerals dissolution	Dissolution of ubiquitous carbonate rocks (limestone, dolomite, marga and coquina) from the Paleocene to Pliocene (Cenozoic) geology (Figs. 1b; 4a, e).	• Major occurrence of samples in the northern part of the peninsula, where logs show Mesozoic sedimentary sequences of carbonate rocks at more than 150 meters deep[14] (Fig. 1b, Supplementary Fig. 2, Fig. 4a, e). • Increase in HCO_3^- without a dramatic increase in salinization (Fig. 2f). • Samples are mainly characterized by calcium-bicarbonate facies in the Piper diagram (Fig. 3a).
C9 and C10	Seawater-groundwater mixing	Seawater intrusion[75] (Fig. 4a, e): • Upconing due to groundwater abstraction (vertical seawater intrusion). • Surge inundation (vertical seawater intrusion). • Sea level rise (horizontal seawater intrusion).	• Samples spreading along Na^+ and Cl^- trend in PC1 vs PC2 plot (Fig. 2b), with high correlation between Na^+ and Cl^- ($r = 0.96$). • Samples have the highest seawater fraction and Cl^- values, with a median of 2.89% and 13.75 meq/L (487.62 mg L^{-1}), respectively (Fig. 3c). • 97% of the samples have $100{\cdot}SO_4^{2-}/Cl^- < 100$, with the highest concentrations of Cl^- close to the seawater ratio (Fig. 3e). • Similar multivariate pattern to samples from previous studies identified as seawater-groundwater mixtures[23,32] and seawater beneath the diffusion zone in the case of UNAM 2 and UNAM 5, which are samples taken to more than 200 m.b.s.l.[24] (Figs. 1b; 2b, c; 3e, f). • Most samples are concentrated close to the coast (Fig. 4a). • Large seawater wedge beneath the coastal plain due to the lowland and consequent low hydraulic heads. This facilitates the upconing of the groundwater-seawater mixture (Fig. 4e). Considering a hydraulic head of 4 m.a.s.l. in Lake Chincancanab[24], the seawater-groundwater interface could be at 160 m.b.s.l. there according to the Ghyben–Herzberg principle[2]. Meanwhile, the diffusion thicknesses could be tens of meters[21–23], and for each meter of drawdown when pumping, the interface can ascend 40 m according to the same principle. All this makes the groundwater-seawater mixture accessible to wells up to 100 meters deep in the peninsula[48]. On the other hand, variations in sea level could affect the seawater wedge position due to the low hydraulic gradient[50]. Furthermore, the coastal lowland makes the peninsula vulnerable to seawater surges from tropical storms[76], common in the Gulf of Mexico[77].
C12 and C14	Gypsum dissolution	Dissolution of gypsum layers from the Paleocene limestone-gypsum Icaiche formation within and around the EIR[17] (Figs. 1b; 4a, e).	• Samples spreading along Ca^{2+} and SO_4^{2-} trend in PC1 vs PC2 plot (Fig. 2b), with high correlation between Ca^{2+} and SO_4^{2-} ($r = 0.93$). Furthermore, there is a high correlation between the excess of Ca^{2+} and the excess of SO_4^{2-} regarding the conservative seawater-groundwater mixture ($r = 0.93$), indicating that high concentrations of these ions have a different source from seawater (Fig. 3c). • Samples are mainly characterized by calcium-sulfate facies in the Piper diagram (Fig. 3a). • 96% of the samples have $100{\cdot}SO_4^{2-}/Cl^- > 100$, indicating evaporite layers dissolution[24] (Fig. 3e, f). • Most samples are concentrated inside or close to the EIR that encloses the outbreaks of gypsum layers with thicknesses of up to 35 m of the Paleocene limestone-gypsum Icaiche formation[17,49] (Figs. 1b; 4a, e). • The samples are the closest to gypsum saturation. C14 has a median SI_{Gypsum} value of −0.3 and C12 has a median value of −0.86. Thus, C12 reflects groundwater affected by gypsum dissolution without achieving saturation[19].
C5, C6 and C11	Nitrate pollution	Leaching of fertilizers, manure and sewage from anthropogenic activities: Farming, livestock and urban settlements (Figs. 1d; 4a, e).	• Samples in C11 are the most influenced in PC3 with a median NO_3^- value of 94.92 mg L^{-1} (Fig. 2d). Water quality in the Yucatan peninsula with concentrations less than 13.8 mg L^{-1} has been indicated as non-polluted groundwater[50]. • Samples are distributed across the peninsula, particularly between the Ticul Fault, the Ring of Cenotes and the Holbox Fractured Zone. Farmlands, cultivated grasslands for livestock and urban settlements are well-known drivers of NO_3^- pollution in the study area[2,33,34,50] (Figs. 1d; 4a, e).

Predominant processes, triggering factors, and supporting observations.

Ring of Cenotes and the Ticul fault, sites close to lake Chichancanab, the central EIR, and in the center of the northern region (Fig. 4a).

Further research could provide information to enhance the understanding of the Yucatan Peninsula and improve the interpretation of this and previous local studies. Although the pattern within and around the EIR supports gypsum dissolution acting regionally, deep geological logs in the southern region could clarify the stratigraphy of this zone and increase the knowledge of the Icaiche and deeper Cretaceous formations. On the other hand, a regional groundwater flow model would be beneficial to understanding recharge trends, contamination pathways, and the diffuse interface

Table 2 | Temporal variations in water quality

Region	Comparison	Test summary	No. of samples	$SI_{Calcite}$	$SI_{Dolomite}$	SI_{Gypsum}	NO_3^-	Seawater fraction
Northern	Long-term	*p* value	-	<0.005	<0.005	<0.005	0.95	0.23
		2014 median	324	0.66	1.05	−2.11	19.96 mg L^{-1}	0.84%
		(1998–2003) median	411	−0.12	−0.45	−2.27	18.61 mg L^{-1}	0.82%
	Seasonal	*p* value	-	<0.005	<0.005	0.12	0.40	0.45
		2014 dry median	157	0.61	0.89	−2.08	20.34 mg L^{-1}	0.86%
		2014 rainy median	167	1.35	2.37	−2.14	19.02 mg L^{-1}	0.83%
Southern	Long-term	*p* value	-	<0.005	<0.005	0.96	<0.005	0.16
		2014 median	370	0.31	0.30	−1.17	20.98 mg L^{-1}	0.59%
		(1998– 2003) median	295	0.13	−0.14	−1.19	14.99 mg L^{-1}	0.68%
	Seasonal	*p* value	-	0.35	0.37	0.63	0.12	0.32
		2014 dry median	181	0.33	0.31	−1.18	22.80 mg L^{-1}	0.70%
		2014 rainy median	189	0.31	0.26	−1.16	18.96 mg L^{-1}	0.53%

Two-tailed Wilcoxon rank sum test to determine whether there is a significant difference between the distribution of variables in two periods. *p* values less than 0.005 were considered statistically significant. The long-term test refers to the comparison between the 2014 and the historical data (1998–2003), and the seasonal test compares data between the 2014 dry and rainy seasons.

between groundwater and the seawater wedge. With this, the aquifer hydrodynamics could be simulated under different pumping allocation, vertical recharge, and sea level rise scenarios, providing more concrete information about optimal land use and protection areas. Substantial challenges for modeling include the collection of reliable spatio-temporal hydraulic and hydrological data and emulating the multiscale karst heterogeneous properties. Hydraulic heads for calibration in the extensive lowland require at least decimetric precision in the measurement regarding the sea-level datum, since 1 m of error could roughly suggest 40 m of error in the vertical position of the seawater-groundwater interface, according to the Ghyben–Herzberg principle.

Conclusions

The comprehensive analysis of spatio-temporal water quality patterns in the Yucatan Peninsula, supported by unsupervised learning techniques, provides valuable insights into regional hydrogeochemical processes and contamination vulnerabilities. Examination of major ions reveals a heterogeneous and complex water quality profile driven by distinct processes. Key factors triggering salinization include the seawater-groundwater mixing driven by seawater intrusion, with the coastal lowland being the most vulnerable to upconing due to the proximity of the seawater wedge and potential seawater surges. On the other hand, gypsum dissolution acts regionally at the south of the peninsula within the Elevated Interior Region. Furthermore, regional recharge to the coastal plain with calcium-sulfate groundwater type is suggested, influencing water quality in the predominantly carbonate aquifer. Additionally, the study highlights a widespread nitrate contamination across the Yucatán Peninsula attributed to inadequate waste management and the high vulnerability of the karst environment, which facilitates rapid infiltration and mixing.

The findings emphasize the importance of spatio-temporal monitoring of major ions to ensure safe water sources for human consumption in the Yucatan Peninsula. Approximately 23% of samples from water sources designated for human use exceeded acceptable limits for the SO_4^{2-} or NO_3^-, posing health risks. With anticipated population and tourism growth, anthropogenic pollution is expected to intensify, while climate change may disrupt hydrogeological patterns in the karst environment, affecting the rate of carbonate and gypsum dissolution. Furthermore, sea level rise could introduce the large seawater wedge into the aquifer due to the low hydraulic heads, further contributing to salinization. The study suggests future research directions to better understand aquifer behavior, including increasing knowledge of seawater-groundwater mixing diffusion zones at the regional scale through a regional flow model and advancing hydrogeological comprehension in the south of the peninsula. These efforts are essential for informed water resource management and sustainable development in the region.

Methods

Data and quality

The dataset analyzed in this study comprises recently published and unpublished reports of 1528 water quality samples collected between 1998 and 2021, focusing on major ions as measurement parameters (Ca^{2+}, Mg^{2+}, Na^+, K^+, SO_4^{2-}, HCO_3^-, Cl^-, and NO_3^-)[36] (Fig. 1d). The dataset includes 706 water samples from historical records (1998–2003), 356 samples at the end of the 2014 rainy season (October and November), 338 at the end of the 2014 dry season (April and May), and 128 samples at the end of the 2021 dry season (April and May). Of these, 78% were collected from tube wells after purging, while the remainder were from cenotes, springs and dug wells. Most of the samples from the wells, mainly intended for water supply, were pumped without certainty about the screening intervals. According to the Mexican water authority, wells have depths of up to 100 m[48]. The samples from the other sources were taken close to the phreatic surface, except for 63 samples from the 2021 campaign that were taken from depths up to 20 m.

Major ions were measured according to the INEGI guidelines for water quality analysis[61]. To ensure the reliability of the analysis, 117 out of the initial 1645 samples were excluded. These exclusions were due to exceeding the 5% threshold for the absolute Charge Balance Error, related to major ions measurement accuracy[62] as well as those exhibiting inconsistencies in concentration values or sampling sites outside the study area. Total Dissolved Solids (TDS) were calculated as the sum of the major ions for subsequent analysis. CO_3^{2-} was excluded from the multivariate analyses due to null values in 64% of the samples; in turn, it constitutes only 0.7% of TDS on average. The dataset statistical description is provided in Supplementary Table 1 and Supplementary Table 2.

Statistical analysis

Statistical analyses were conducted using the R programming environment utilizing the stats package[37,63]. This included descriptive statistics and the two-tailed Wilcoxon rank-sum tests[64] for assessing temporal differences between non-parametric distributions. *p* values less than 0.005 were considered statistically significant.

Multivariate analysis

Clustering of water samples was achieved through the Edge Betweenness[65] community detection algorithm applied to an undirected and unweighted

Nearest Neighbor Graph[66] of the INEGI dataset[36,37]. First, the Euclidean distance matrix of the standardized dataset was computed, and the K shortest distances for each sample were selected as the network's connections to simplify the topology of the multivariate data. The Edge Betweenness community detection algorithm was then applied to the graph network to identify clusters (communities) based on the resulting structure. For this analysis, $K = 8$ was chosen to minimize the number of clusters and maximize the modularity, which quantifies the strength of a network's division into communities. Fewer clusters simplify the analysis, while higher modularity values indicate better-defined separation. Supplementary Table 3 shows the variation in clusters' numbers and modularity concerning the K values. The R packages igraph[67] and cccd[68] were used to construct the network and perform community detection, and the Kamada Kawai layout was utilized for network visualization[67].

Principal components analysis (PCA)[69] was conducted on the INEGI dataset for multivariate data exploration and dimensional reduction[36,37]. Data from previous studies[19,23,24,32,33] were projected onto the PCA results by multiplying each standardized dataset by the PCA weight matrix. The means and standard deviations of the INEGI dataset were used for standardization. PCA was conducted using the R package psych[70].

Hydrogeochemical analysis

Water samples were plotted into the Piper diagram to observe the relative dominance of major ions and infer the main hydrogeochemical facies[71]. Seawater fraction and excess of major ions were calculated regarding the conservative mixing between freshwater and seawater with chloride as the conservative tracer[72]. For this, the chemical compositions of rainwater[73] and seawater[32] in the Yucatan Peninsula were used as end-members. The $100{\cdot}SO_4^{2-}/Cl^-$ ratio was also calculated, as it has been highlighted in previous studies as a valuable tracer in the study area.

Saturation indices for minerals of interest were calculated to assess trends in mineral dissolution or precipitation. The indices were determined by the formula $SI = Log(IAP/K)$, where IAP represents the ion activity product and K denotes the equilibrium constant for the specific mineral. The calculations were performed using the USGS PHREEQC 2.7[74] program, with default equilibrium constants set as follows: $Log_{10}(K_{Calcite}) = \mathbf{-8.48}$, $Log_{10}(K_{Dolomite}) = \mathbf{-17.09}$, and $Log_{10}(K_{Gypsum}) = \mathbf{-4.58}$.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The INEGI dataset and the statistical summary of the variables analyzed by cluster are publicly available at https://doi.org/10.5281/zenodo.15604505[36].

Code availability

The R Markdown related to the statistical and multivariate analysis is publicly available at https://doi.org/10.5281/zenodo.15547586[37].

Received: 5 November 2024; Accepted: 6 June 2025;
Published online: 16 June 2025

References

1. Platt, N. H. & Wright, V. P. Flooding of a carbonate platform: the Sian Ka'an Wetlands, Yucatán, Mexico—a model for the formation and evolution of palustrine carbonate factories around the modern Caribbean Sea and in the depositional record. *Depos. Rec.* **9**, 99–151 (2023).
2. Bauer-Gottwein, P. et al. Review: the Yucatán Peninsula karst aquifer, Mexico. *Hydrogeol. J.* **19**, 507–524 (2011).
3. Goldscheider, N. et al. Global distribution of carbonate rocks and karst water resources. *Hydrogeol. J.* **28**, 1661–1677 (2020).
4. CONAGUA. Información Estadística Climatológica [Mexican Climatological Statistical Information]. Available at https://smn.conagua.gob.mx/es/climatologia/informacion-climatologica/informacion-estadistica-climatologica (2025).
5. CONAGUA. *Actualización De La Disponibilidad Media Anual De Agua En El Acuífero Península De Yucatán (3105), Estado De Yucatán [Update on the Average Annual Availability of Water in the Yucatan Peninsula Aquifer (3105), State of Yucatan]*. Available at https://sigagis.conagua.gob.mx/gas1/Edos_Acuiferos_18/yucatan/DR_3105.pdf (2024).
6. Vidal, L., Vallarino, A., Benítez, I. & Correa, J. Implementation of the Ramsar strategic plan in coastal wetlands of the Península de Yucatán: regulations and normativity. *Lat. Am. J. Aquat. Res.* **43**, 873–887 (2017).
7. Arroyave, J., Martinez, C. M., Martínez-Oriol, F. H., Sosa, E. & Alter, S. E. Regional-scale aquifer hydrogeology as a driver of phylogeographic structure in the Neotropical catfish Rhamdia guatemalensis (Siluriformes: Heptapteridae) from cenotes of the Yucatán Peninsula, Mexico. *Freshw. Biol.* **66**, 332–348 (2021).
8. Kennett, D. J. et al. Drought-induced civil conflict among the ancient Maya. *Nat. Commun.* **13**, 3911 (2022).
9. Melo Zurita, M., de, L. & Munro, P. G. Voluminous territorialisation: historical contestations over the Yucatan Peninsula's subterranean waterscape. *Geoforum* **102**, 38–47 (2019).
10. Çallı, K. Ö. et al. Karst water resources in a changing world: review of solute transport modeling approaches. *Rev. Geophys.* **63**, e2023RG000811 (2025).
11. Moreno-Pérez, P. A., Hernández-Téllez, M. & Bautista-Gálvez, A. In danger one of the largest aquifers in the world, the great mayan aquifer, based on monitoring the cenotes of the Yucatan Peninsula. *Arch. Environ. Contam Toxicol.* **81**, 189–198 (2021).
12. Whalen, M. T. et al. Winding down the Chicxulub impact: The transition between impact and normal marine sedimentation near ground zero. *Mar. Geol.* **430**, 106368 (2020).
13. Rebolledo-Vieyra, M. & Urrutia-Fucugauchi, J. Magnetostratigraphy of the Cretaceous/Tertiary boundary and early Paleocene sedimentary sequence from the Chicxulub Impact Crater. *Earth Planets Space* **58**, 1309–1314 (2006).
14. Guzmán-Hidalgo, E. et al. Seismic stratigraphic evidence of a pre-impact basin in the Yucatán Platform: morphology of the Chicxulub crater and K/Pg boundary deposits. *Mar. Geol.* **441**, 106594 (2021).
15. Dunning, N. P., Beach, T. P. & Luzzadder-Beach, S. Kax and kol: collapse and resilience in lowland Maya civilization. *Proc. Natl. Acad. Sci. USA* **109**, 3652–3657 (2012).
16. Leal-Bautista, R. M. et al. Karst drainage of lake Chakanbacán and its relation to Icaiche Formation. *Bol Soc. Geol. Mex.* **73**, A021020 (2021).
17. Perry, E., Velazquez-Oliman, G., Leal-Bautista, R. M. & Dunning, N. The Icaiche Formation: major contributor to the stratigraphy, hydrogeochemistry and geomorphology of the northern Yucatán peninsula, Mexico. *Bol. Soc. Geol. Mex.* **71**, 741–760 (2019).
18. Perry, E. C., Leal-Bautista, R. M., Velázquez-Olimán, G., Sánchez-Sánchez, J. A. & Wagner, N. Aspects of the hydrogeology of southern campeche and quintana roo, Mexico. *Bol. Soc. Geol. Mex.* **73**, A011020 (2021).
19. Gondwe, B. R. N. et al. Hydrogeology of the south-eastern Yucatan Peninsula: new insights from water level measurements, geochemistry, geophysics and remote sensing. *J. Hydrol.***389**, 1–17 (2010).
20. Brankovits, D. et al. Methane- and dissolved organic carbon-fueled microbial loop supports a tropical subterranean estuary ecosystem. *Nat. Commun.* **8**, 1835 (2017).
21. Zamora-Luria, J. C., Perera-Burgos, J. A., González-Calderón, A., Marin Stillman, L. E. & Leal-Bautista, R. M. Control of fracture networks on a coastal karstic aquifer: a case study from northeastern Yucatán Peninsula (Mexico). *Hydrogeol. J.* **28**, 2765–2777 (2020).
22. Canul-Macario, C., Salles, P., Hernández-Espriú, A. & Pacheco-Castro, R. Numerical modelling of the saline interface in coastal

karstic aquifers within a conceptual model uncertainty framework. *Hydrogeol. J.* **29**, 2347–2362 (2021).
23. Cejudo, E., Ortega-Almazán, P. J., Ortega-Camacho, D. & Acosta-González, G. Hydrochemistry and water isotopes of a deep sinkhole in north Quintana Roo, Mexico. *J. South Am. Earth Sci.* **116**, 103846 (2022).
24. Perry, E., Velazquez-Oliman, G. & Marin, L. The Hydrogeochemistry of the Karst Aquifer System of the Northern Yucatan Peninsula, Mexico. *Int Geol. Rev.* **44**, 191–221 (2002).
25. CONAGUA. Programa Hídrico Regional 2021-2024 - Región Hidrológico-Administrativa XII Península de Yucatán [Regional Water Program 2021-2024 - Hydrological-Administrative Region XII Yucatan Peninsula]. Available at https://files.conagua.gob.mx/conagua/generico/PNH/PHR_2021-2024_RHA_XII_Península_de_Yucatán.pdf (2022).
26. CONAGUA. *Situación Del Subsector Agua Potable, Alcantarillado y Saneamiento—Edición 2023 [Situation of the Drinking Water, Sewage and Sanitation Subsector—2023 Edition]*. Available at https://www.gob.mx/cms/uploads/attachment/file/876087/Edici_n_c_2023.pdf (2023).
27. Guardiola, J., González-Gómez, F. & Grajales, ÁL. Is access to water as good as the data claim? Case study of Yucatan. *Int. J. Water Resour. Dev.* **26**, 219–233 (2010).
28. Polanco Rodríguez, A. G. et al. Organochlorine pesticides in the drinking water of Merida and its Metropolitan Zone, a Karst Region. *Urban Water J.* **19**, 40–50 (2022).
29. ONU-Habitat. Sistema De Evaluación Exante De Los Impactos Generados Por La Estrategia De Desarrollo Integral De La Región Sureste [Exante Evaluation System of the Impacts Generated by the Comprehensive Development Strategy of the Southeast Region]. Available at https://publicacionesonuhabitat.org/onuhabitatmexico/Sistema-Evaluacion-Exante.pdf (2020).
30. Ortega, R. P. & Jaber, I. G. A controversial train heads for the Maya forest. *Science* **375**, 250–251 (2022).
31. Goldscheider, N. A holistic approach to groundwater protection and ecosystem services in karst terrains. *Carbonates Evaporites* **34**, 1241–1249 (2019).
32. Saint-Loup, R., Felix, T., Maqueda, A., Schiller, A. & Renard, P. A survey of groundwater quality in Tulum region, Yucatan Peninsula, Mexico. *Environ. Earth Sci.* **77**, 644 (2018).
33. Smith, D. N. I. et al. A multi-approach assessment of land use effects on groundwater quality in a karstic aquifer. *Heliyon* **6**, e03970 (2020).
34. Pacheco Castro, R., Pacheco Ávila, J., Ye, M. & Cabrera Sansores, A. Groundwater quality: analysis of its temporal and spatial variability in a karst aquifer. *Groundwater* **56**, 62–72 (2018).
35. Pérez-Ceballos, R. et al. Regional hydrogeochemical evolution of groundwater in the ring of cenotes, Yucatán (Mexico): An Inverse Modelling Approach. *Water* **13**, 614 (2021).
36. Narvaez-Montoya, C., Mondragon-Bonilla, R., Goldscheider, N. & Mahlknecht, J. Hydrogeochemistry of the large Yucatan Peninsula karst aquifer system (V2). Available at https://doi.org/10.5281/zenodo.15604505 (2025).
37. Narvaez-Montoya, C., Mondragon-Bonilla, R., Goldscheider, N. & Mahlknecht, J. Network-based clustering, PCA and hypothesis testing for multivariate water quality data. Available at https://doi.org/10.5281/zenodo.15547585 (2025).
38. Narvaez-Montoya, C., Mahlknecht, J., Torres-Martínez, J. A., Mora, A. & Bertrand, G. Seawater intrusion pattern recognition supported by unsupervised learning: a systematic review and application. *Sci. Total Environ.* **864**, 160933 (2023).
39. Chae, G.-T., Yun, S.-T., Yun, S.-M., Kim, K.-H. & So, C.-S. Seawater–freshwater mixing and resulting calcite dissolution: an example from a coastal alluvial aquifer in eastern South Korea. *Hydrol. Sci. J.* **57**, 1672–1683 (2012).
40. Terzić, J., Peh, Z. & Marković, T. Hydrochemical properties of transition zone between fresh groundwater and seawater in karst environment of the Adriatic islands, Croatia. *Environ. Earth Sci.* **59**, 1629–1642 (2010).
41. Gaofeng, Z., Yonghong, S., Chunlin, H., Qi, F. & Zhiguang, L. Hydrogeochemical processes in the groundwater environment of Heihe River Basin, northwest China. *Environ. Earth Sci.* **60**, 139–153 (2010).
42. Kanellopoulos, C. & Argyraki, A. Multivariate statistical assessment of groundwater in cases with ultramafic rocks and anthropogenic activities influence. *Appl. Geochem.* **141**, 105292 (2022).
43. Mao, H. et al. Deciphering spatial pattern of groundwater chemistry and nitrogen pollution in Poyang Lake Basin (eastern China) using self-organizing map and multivariate statistics. *J. Clean. Prod.* **329**, 129697 (2021).
44. Fang, Y. et al. Assessment of the hydrodynamics role for groundwater quality using an integration of GIS, water quality index and multivariate statistical techniques. *J. Environ. Manag.* **273**, 111185 (2020).
45. Belkhiri, L. & Narany, T. S. Using multivariate statistical analysis, geostatistical techniques and structural equation modeling to identify spatial variability of groundwater quality. *Water Resour. Manag.* **29**, 2073–2089 (2015).
46. Sangadi, P., Kuppan, C. & Ravinathan, P. Effect of hydro-geochemical processes and saltwater intrusion on groundwater quality and irrigational suitability assessed by geo-statistical techniques in coastal region of eastern Andhra Pradesh, India. *Mar. Pollut. Bull.* **175**, 113390 (2022).
47. Fidelibus, M. D. et al. A chloride threshold to identify the onset of seawater/saltwater intrusion and a novel categorization of groundwater in coastal aquifers. *J. Hydrol.* **653**, 132775 (2025).
48. CONAGUA. Actualización De La Disponibilidad Media Anual De Agua En El Acuífero Península De Yucatán (3105), Estado De Yucatán [Update on the Average Annual Availability of Water in the Yucatan Peninsula Aquifer (3105), State of Yucatan]. Available at https://www.gob.mx/cms/uploads/attachment/file/103392/DR_3105.pdf (2015).
49. Perry, E., Paytan, A., Pedersen, B. & Velazquez-Oliman, G. Groundwater geochemistry of the Yucatan Peninsula, Mexico: constraints on stratigraphy and hydrogeology. *J. Hydrol.* **367**, 27–40 (2009).
50. Rojas Fabro, A. Y., Pacheco Ávila, J. G., Esteller Alberich, M. V., Cabrera Sansores, S. A. & Camargo-Valero, M. A. Spatial distribution of nitrate health risk associated with groundwater use as drinking water in Merida, Mexico. *Appl. Geogr.* **65**, 49–57 (2015).
51. Pacheco, F. A. L., Landim, P. M. B. & Szocs, T. Anthropogenic impacts on mineral weathering: A statistical perspective. *Appl. Geochem.* **36**, 34–48 (2013).
52. Mendoza-Olea, I. J. et al. Contaminación por microplásticos en el acuífero kárstico de la península de Yucatán. *Ecosistemas Recur. Agropecu*. **9**, e3360 (2022).
53. Martínez-Casales, Y. et al. Caffeine and paraxanthine as tracers of anthropogenic wastewater in coastal Lagoons in Yucatan, Mexico. *Bull. Environ. Contam Toxicol.* **108**, 182–189 (2022).
54. Sabarathinam, C., Bhandary, H. & Ali, A. Strategies to characterize the geochemical interrelationship between coastal saline groundwater and seawater. *Environ. Earth Sci.* **80**, 642 (2021).
55. Martínez-Salvador, C., Moreno-Gómez, M. & Liedl, R. Estimating pollutant residence time and NO3 concentrations in the Yucatan Karst Aquifer; Considerations for an Integrated Karst Aquifer Vulnerability Methodology. *Water* **11**, 1431 (2019).
56. INEGI. Uso del Suelo y Vegetación Serie II Conjunto Nacional—Decada de los 90s [Land Use and Land Cover Series II National Set - the 90's]. Available at https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=702825007021 (2001).
57. INEGI. Uso del Suelo y Vegetación Serie VII Conjunto Nacional -2018 [Land Use and Land Cover Series VII National Set - 2018]. Available at

https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=889463842781 (2021).
58. World Health Organization. *Guidelines for Drinking-Water Quality*. (Gutenberg, 2011).
59. Boretti, A. A realistic expectation of sea level rise in the Mexican Caribbean. *J. Ocean Eng. Sci.* **4**, 379–386 (2019).
60. Ferguson, G. & Gleeson, T. Vulnerability of coastal aquifers to groundwater use and climate change. *Nat. Clim. Chang* **2**, 342–345 (2012).
61. INEGI. Los análisis físicos y químicos en la cartografía hidrológica del INEGI [Physical and chemical analyses in INEGI's hydrological mapping]. Available at https://www.inegi.org.mx/contenidos/temas/mapas/hidrologia/metadatos/normhidro.pdf (2020).
62. Fritz, S. J. A survey of charge-balance errors on published analyses of potable ground and surface waters. *Groundwater* **32**, 539–546 (1994).
63. Dalgaard, P. *Introductory Statistics with R* (Springer New York, 2008).
64. Krzywinski, M. & Altman, N. Nonparametric tests. *Nat. Methods* **11**, 467–468 (2014).
65. Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, 026113 (2004).
66. Marchette, D. J. *Random Graphs for Statistical Pattern Recognition* (Wiley, 2004).
67. Csárdi, G. et al. Package 'igraph': network analysis and visualization. Version 2.0.2. Available at https://cran.r-project.org/web/packages/igraph/igraph.pdf (2024).
68. Marchette, D. J. Package 'CCCD': Class Cover Catch Digraphs. Version 1.6. Available at https://cran.r-project.org/web/packages/cccd/cccd.pdf (2022).
69. Fraino, P. E. Using principal component analysis to explore multivariable relationships. *Nat. Rev. Earth Environ.* **4**, 294–294 (2023).
70. Revelle, W. Package 'psych': procedures for psychological, psychometric, and personality research. Version 2.5.3. Avaliable at https://cran.r-project.org/web/packages/psych/psych.pdf (2025).
71. Kuniansky, E. L., Taylor, C. J., Williams, J. H. & Paillet, F. *Introduction to Karst Aquifers*. https://doi.org/10.21083/978-1-77470-040-2 (The Groundwater Project, 2024).
72. Cook, P. *Introduction to Isotopes and Environmental Tracers as Indicators of Groundwater Flow*. https://doi.org/10.21083/978-1-7770541-8-2 (The Groundwater Project, 2020).
73. Cerón, R. M. B. et al. Rainwater chemical composition at the end of the mid-summer drought in the Caribbean shore of the Yucatan Peninsula. *Atmos. Environ.* **36**, 2367–2374 (2002).
74. USGS. *User's Guide to PHREEQC (Version 2): A Computer Program for Speciation, Batch-Reaction, One-Dimensional Transport, and Inverse Geochemical Calculations*. Available at https://pubs.usgs.gov/publication/wri994259 (1999).
75. Kretschmer, D. V. et al. A perceptual model of drivers and limiters of coastal groundwater dynamics. *Hydrol. Process.* **39**, e70058 (2025).
76. Smith, C. G., Cable, J. E. & Martin, J. B. Episodic high intensity mixing events in a subterranean estuary: effects of tropical cyclones. *Limnol. Oceanogr.* **53**, 666–674 (2008).
77. Yang, Z. et al. A modeling study of coastal inundation induced by storm surge, sea-level rise, and subsidence in the Gulf of Mexico. *Nat. Hazards* **71**, 1771–1794 (2014).
78. SGM. Carta Geológico-Minera, Estados de: Campeche, Quintana Roo y Yucatán. Escala 1:250,000 [Geological-Mining Chart, States of: Campeche, Quintana Roo and Yucatan. Scale 1:250,000]. Available at https://www.sgm.gob.mx/GeoInfoMexGobMx/# (2007).
79. Ortiz, G., Gonzales, L. & Guitérrez, C. Estudio hidrogeológico en el acuífero Xpujil en el estado de Campeche [Hydrogeological study of the Xpujil aquifer in the state of Campeche]. Available at http://repositorio.imta.mx/handle/20.500.12013/2199 (2019).
80. CONAGUA. GeoVisor de Consulta de Mediciones Piezométricas [GeoViewer for consulting piezometric measurements]. Available at https://sigagis.conagua.gob.mx/rp20/ (2024).

Acknowledgements

We extend our gratitude to the Consejo Nacional de Ciencia y Tecnología (CONAHCYT) (CVU: 1014283) and Tecnologico de Monterrey for providing scholarship and tuition support to C.N.-M. during the Ph.D. degree program. Additionally, we thank the Mexican National Institute of Statistics and Geography—Instituto Nacional de Estadística y Geografía (INEGI) for sharing unpublished data from the historical period (1998–2003). Finally, we thank the German Academic Exchange Service (DAAD) for awarding C.N.-M. the Short-Term-Grants 2024 scholarship.

Author contributions

Christian Narvaez-Montoya: Initial idea and conceptualization, Data curation, Data analysis, Investigation, Writing-Original draft preparation, Visualization, Funding acquisition. Rogelio Mondragón Bonilla: Data curation, Validation, Reviewing and editing. Nico Goldscheider: Validation, Reviewing and editing. Jürgen Mahlknecht: Supervision, Funding acquisition, Project administration, Reviewing and editing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-025-02456-1.

Correspondence and requests for materials should be addressed to Jürgen Mahlknecht.

Peer review information *Communications Earth & Environment* thanks PJ Moore and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Patricia Spellman, Somaparna Ghosh. A peer review file is available.