Journal of Groundwater Science and Engineering
ISSN 2305-7068
Indexed by ESCI CABI CSA
Scopus GeoRef AJ CNKI

New insights into soda water in shallow groundwater of the North China Plain

Bo Gao, Jiang-tao He, Bao-nan He, Yan-jia Chu, Zhen Chen

Citation:
Gao B, He JT, He BN, Chu Y-J, Chen Z. 2026. New insights into soda water in shallow groundwater of the North China Plain. Journal of Groundwater Science and Engineering, 14(4): 458-477.

View online: https://doi.org/10.26599/JGSE.2026.9280103

Articles you may be interested in

Optimizing groundwater recharge plan in North China Plain to repair shallow groundwater depression zone, China
Journal of Groundwater Science and Engineering. 2023, 11(2): 133-145 https://doi.org/10.26599/JGSE.2023.9280012

ANN-based prediction model for single-hole water inflow from piedmont to inland plain areas of Hebei Province, North China Plain
Journal of Groundwater Science and Engineering. 2025, 13(4): 434-448 https://doi.org/10.26599/JGSE.2025.9280064

Analysis of driving factors for land subsidence in typical cities of the North China Plain based on geodetector technology
Journal of Groundwater Science and Engineering. 2025, 13(1): 74-89 https://doi.org/10.26599/JGSE.2025.9280040

Effectiveness of groundwater extraction in Beijing since the ingauration of the first phase of the South-to-North Water Diversion Project, China
Journal of Groundwater Science and Engineering. 2023, 11(3): 294-307 https://doi.org/10.26599/JGSE.2023.9280024

Determining safe yield and mapping water level zoning in groundwater resources of the Neishabour Plain
Journal of Groundwater Science and Engineering. 2023, 11(1): 47-54 https://doi.org/10.26599/JGSE.2023.9280005

Assessment of water level threshold for groundwater restoration and over-exploitation remediation the Beijing-Tianjin-Hebei Plain
Journal of Groundwater Science and Engineering. 2022, 10(2): 113-127 https://doi.org/10.19637/j.cnki.2305-7068.2022.02.002

Journal of Groundwater Science and Engineering 14(2026) 458–477

Journal homepage: http://gwse.iheg.org.cn

Research Article

New insights into soda water in shallow groundwater of the North China Plain

Bo Gao[1,2], Jiang-tao He[1,2*], Bao-nan He[1,2], Yan-jia Chu[1,2], Zhen Chen[1,2]

[1] *Key Laboratory of Groundwater Conservation of Ministry of Water Resources (in preparation)*, *China University of Geosciences (Beijing)*, *Beijing* 100083, *China*.
[2] *School of Water resources and Environment*, *China University of Geosciences (Beijing)*, *Beijing* 100083, *China*.

Abstract: Soda water in shallow aquifers represents a unique hydrochemical type, often enriched in arsenic (As), Fluorine (F), Iodine (I), and other components, while also acting as a critical driver of soil salinization. However, existing studies have failed to effectively distinguish between salinization (characterized by soluble salt accumulation) and alkalization (characterized by soda-alkali enrichment). The "new insights" of this study do not rely on new data but derive from an in-depth excavation and interpretation of the 2006–2009 National Groundwater Pollution Survey dataset—the only authoritative background dataset covering the entire North China Plain. Focusing on shallow groundwater in the North China Plain, this study refines the identification criteria for soda water based on existing concepts, analyzes its spatial distribution characteristics, delineates typical zones, and conducts a preliminary investigation into the genetic differences across regions. Results show that when using the criterion—"HCO_3^- + CO_3^{2-} as dominant anions with [(HCO_3^- + CO_3^{2-}) - (Ca^{2+} + Mg^{2+})] > 0 meq%"—combined with hydrochemical cluster analysis, soda water is primarily concentrated in two zones: The mountain-front discharge zone (Area A) and the runoff-ancient Yellow River channel zone (Area B). These two zones account for 88.48% of all soda water samples and exhibit distinct hydrochemical features. In Area A, groundwater has a simple anion composition dominated by HCO_3^-, a median Total Dissolved Solids (TDS) content of 501.15 mg/L, and elevated concentrations of F^- and NO_3^-. In contrast, Area B is characterized by diverse anions (HCO_3^-, SO_4^{2-}, and Cl^-), a higher median TDS (863.56 mg/L), and enrichment of reductive components including As, F^-, I^-, Fe, and Mn. Genetic analysis reveals that soda water in Area A forms through the combined effects of mineral weathering, dissolution, and calcite-dolomite precipitation. In contrast, groundwater in Area B evolves under calcite-dolomite precipitation controlled by evaporative concentration, with further modifications by microbial geochemical processes and agricultural activities. This study clarifies the spatial distribution patterns and genetic mechanisms of soda water in the North China Plain, laying a foundation for further research on its formation processes.

Keywords: Soda water; Zoning characteristics; SOM-KM clustering; North China Plain; Preliminary genetic analysis

Received: 16 Sep 2025/ Accepted: 16 Mar 2026/ Published: 12 Jun 2026

Introduction

In hydrogeology, soda water typically refers to groundwater naturally enriched in sodium bicarbonate ($NaHCO_3$), characterized by an alkaline pH and moderate Total Dissolved Solids (TDS). Such waters exhibit a global distribution with notable regional variability and diverse genetic mechanisms. They occur not only in tectonically active

*Corresponding author: Jiang-tao He, *E-mail address*: jthe@cugb.edu.cn

DOI: 10.26599/JGSE.2026.9280103

Gao B, He JT, He BN, et al. 2026. New insights into soda water in shallow groundwater of the North China Plain. Journal of Groundwater Science and Engineering, 14(4): 458-477.

2305-7068/© 2026 Journal of Groundwater Science and Engineering Editorial Office This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0)

zones or closed sedimentary basins but also predominate in plains, plateaus, and piedmont fault zones. While certain soda waters hold potential for therapeutic, potable, and industrial applications (Mansouri et al., 2024), their distinctive hydrochemical properties generally exert dual effects on environmental systems—particularly with respect to soil salinization and groundwater ecosystem security (Ma et al., 2025). Consequently, systematic investigations into their formation mechanisms and eco-environmental impacts remain imperative.

Existing studies demonstrate that the formation of soda water is closely associated with specific geological structures, rock-mineral compositions, and hydrogeochemical processes (Talib et al., 2019; Xu et al., 2019; Si et al., 2016; Grobe et al., 2002). For instance, in volcanic zones, sedimentary basins, or strata rich in sodium-bearing minerals (e.g., albite), prolonged water-rock interactions —including dissolution and ion exchange—during groundwater circulation may lead to the enrichment of sodium ions (Na^+) and bicarbonate ions (HCO_3^-) (Han et al., 2024; Qian et al., 1995). Furthermore, climatic conditions and anthropogenic activities can also exert significant influences on the formation and evolution of soda water.

Based on current research, the formation of soda water can be categorized into four genetic models (Table 1): (1) deep groundwater circulation coupled with sulfate reduction, (2) evaporation-induced salt accumulation, (3) marine sediment leaching coupled with ion exchange, and (4) crystalline rock weathering under water-rock equilibrium conditions (Wang et al., 2021). These models exhibit distinct hydrogeological conditions and controlling factors. In particular, soda water in shallow groundwater of plain basins is typically attributed to the evaporation-induced salt accumulation model. Such waters predominantly occur in arid to semi-arid plain regions, where alternating sandy and clay layers in shallow aquifers facilitate salt accumulation driven by intense evaporation. This process generates moderately alkaline, high-TDS waters, often characterized by notable fluoride anomalies and pronounced vertical salinity stratification (Lei et al., 2022; Irén et al., 1997).

However, shallow groundwater in the North China Plain exhibits highly complex hydrochemical characteristics and associated geochemical processes. Consequently, soda water in this region is widely distributed and likely originates from multiple genetic mechanisms rather than a single dominant mode. Furthermore, previous studies on soil salinization—a process closely linked to soda water dynamics—have often failed to distinguish between salinization (characterized by soluble salt accumulation) and alkalization (characterized by soda enrichment). This oversight disregards their fundamental differences in mineral dissolution mechanisms, hydrochemical indicators, and ecological impacts (Xiong, 1979; Chen et al., 2022; Wu, 1999; Huang et al., 2025; Zhi et al., 2023).

This issue inevitably leads to a biased understanding of spatial distribution patterns and environmental effects, thereby hindering the accurate interpretation of soda-alkalization evolution processes. Against this background, this study focuses on soda water in the shallow groundwater of the North China Plain, aiming to achieve the following objectives:

(1) Building upon existing conceptual frame-

Table 1 Synthesis of genetic mechanistic frameworks for soda water formation

Genetic model	Key controls	Hydrogeological settings	References
Deep groundwater circulation coupled with sulfate reduction	Calcium carbonate precipitation	Fault zones, tectonic activity, sluggish hydrodynamic conditions Sulfate reduction ($SO_4^{2-} \rightarrow HCO_3^-$ conversion)	(Cooper et al., 2006) (Chae et al., 2006) (Lepokurova, 2020)
Evaporation-induced salt accumulation	Calcium carbonate precipitation	Sand-clay interbedded layers, depression-dominated topography High temperature, low humidity, slow groundwater flow Evaporative enrichment of salts	(Lei et al., 2022) (Shvartsev and Wang, 2006) (Deocampo, 2010)
Marine sediment leaching coupled with ion exchange	Ion exchange	Connate paleo-seawater in marine sedimentary formations Na^+-Ca^{2+} cation exchange	(Irén et al., 1997)
Crystalline rock weathering under water-rock equilibrium conditions	Weathering of aluminosilicate minerals	Aluminosilicate mineral dissolution, organic matter decomposition Seasonal water table fluctuations Slow hydrodynamic circulation	(Azaria et al., 2023) (Dutova, 2020) (Lepokurova and Shvartsev, 2019) (Hanor et al., 2023)

works, we propose criteria to distinguish salinization from alkalization, establish definitive identification standards for soda water, and conduct a comparative analysis of hydrochemical characteristics between soda and non-soda water;

(2) By integrating groundwater system zoning with geomorphological features, supplemented by Self-Organizing Map-k-Means (SOM-KM) clustering analysis, we identify typical soda water distribution zones and systematically characterize their spatial patterns and hydrochemical signatures;

(3) Through hydrogeochemical modeling and multivariate statistical analysis, we preliminarily elucidate regional differences in soda water genesis, providing a foundation for future investigations into its formation mechanisms and eco-environmental impacts.

1 Materials and methods

1.1 Study area

The North China Plain is bounded by Bohai Bay to the east and adjoins the Cenozoic Taihang Mountain tectonic uplift zone to the west. It is adjacent to the Mesozoic Yanshan fault zone in the north and demarcated by the southern margin of the Yellow River alluvial fan to the south. As part of the Huang-Huai-Hai Plain, this region covers approximately 140,000 km^2, encompassing parts of Beijing, Tianjin, Shandong Province, Hebei Province, and Henan Province. Climatically, the North China Plain is located in a semi-arid zone characterized by a continental monsoon climate. The mean annual temperature ranges from 4°C to 14°C. Influenced by monsoon circulation, precipitation exhibits a southeast-to-northwest decreasing gradient, with annual totals declining from 1,200 mm to 400 mm. The mean annual evaporation varies between 1,000 mm and 2,000 mm (Wang et al., 2015).

During the Neogene and Quaternary periods, the North China Plain was situated within the Cenozoic fault-depression zone of the North China Platform, with marginal mountains uplifting and the central area subsiding. Based on the Quaternary geological landforms, groundwater hydrodynamic characteristics, and groundwater chemical properties of the North China Plain, groundwater in this region exhibits distinct zonal regularity in horizontal distribution. A complete and unified groundwater system has formed from the piedmont to the coastal areas, and the Quaternary aquifer system of the North China Plain is divided into three zones: the piedmont alluvial-proluvial inclined plain (recharge area), the central alluvial-lacustrine plain (runoff area), and the eastern alluvial-marine coastal plain (discharge area) (Fig. 1a) (Zhang et al., 2009; Xing et al., 2012).

The piedmont alluvial-proluvial plain is formed by the deposition of weathered clastic materials from the Yanshan Mountains and Taihang Mountains, distributed in a zonal pattern along the mountain foothills. The central alluvial-lacustrine plain is composed of sediments transported by the Yellow River, Luan River, Haihe River, and their tributaries, with a layered distribution in the vertical direction. The eastern alluvial-marine coastal plain is mainly formed by the connection of river deltas and coastal depressions via marine sandbars. The central and eastern parts of the plain have experienced relatively intense weathering and

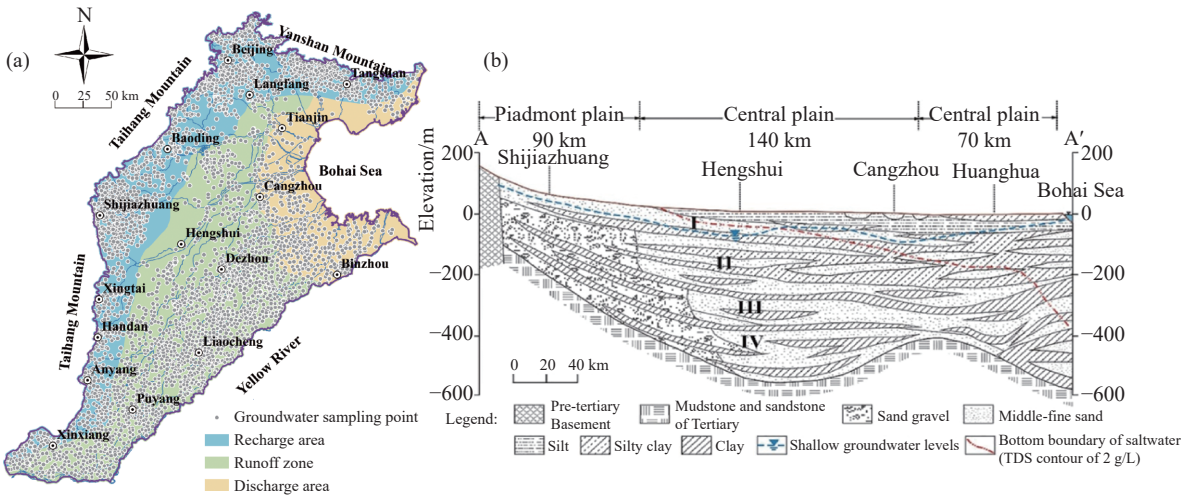


Fig. 1 Zoning of recharge, runoff and discharge areas (a) and hydrogeological section (b) of the North China Plain (Han et al., 2021)

erosion, and paleochannels have been cut off and straightened, resulting in relatively complex geomorphology (Liu et al., 2022).

Based on lithological characteristics, geological age, aquifer distribution, and hydrodynamic conditions, the Quaternary aquifers in the North China Plain can be vertically divided into four aquifer groups, as shown in Fig. 1b. From top to bottom, these aquifer groups correspond sequentially to the Holocene strata (Q_4), Upper Pleistocene strata (Q_3), Middle Pleistocene strata (Q_2), and Lower Pleistocene strata (Q_1). Among them, Aquifer Group I is an unconfined aquifer with a thickness of approximately 60 m; its sediment composition transitions from gravel in the piedmont area to fine sand in the coastal area. Aquifer Group II is a shallow confined aquifer, mainly composed of gravel, medium sand, and fine sand, with its bottom interface at a depth of approximately 120–210 m. Aquifer Group III is a relatively confined aquifer: This unit consists mainly of gravel in the recharge area and medium-fine sand in the runoff and discharge areas. Aquifer Group IV is a deep aquifer, with its bottom interface at a depth of approximately 350–550 m, and is composed of cemented sandy gravel and sand layers (Meng, 2011; Chen et al., 2003). Since groundwater in Aquifer Groups I and II of the piedmont plain has been exploited in a mixed manner, these two groups are collectively referred to as shallow groundwater (Xing et al., 2012).

In this study, based on the characteristics of the hydrochemical analysis data used and the regional hydrogeological conditions, the study area is divided into shallow and deep groundwater, with 100 m set as the boundary. Shallow groundwater is selected as the research object.

1.2 Data acquisition and preprocessing

The data used in this study were obtained from the National Groundwater Pollution Survey Comprehensive Research.

The dataset utilized in this study was obtained from the National Groundwater Pollution Survey conducted during 2006–2009, which is an authoritative, systematic, and comprehensive groundwater quality dataset in China, featuring wide spatial coverage and a complete set of hydrochemical indicators. It is currently the most intact dataset with reliable sampling and testing standards for the North China Plain. Notably, although a certain degree of anthropogenic interference already existed in the region during 2006–2009, the formation of soda water is dominated by natural hydrogeochemical processes (e.g., reductive dissolution, competitive adsorption, evaporative concentration), which are relatively stable and less affected by short-term anthropogenic activities. Given that the core geochemical principles governing soda water formation are thermodynamically stable, the findings from this historical dataset retain strong explanatory power for understanding the current hydrochemical characteristics of soda water in the region.

Sampling and testing were primarily conducted during the summer months (June to August) from 2006 to 2009. All procedures for sampling site selection, sample collection, laboratory testing, and quality control strictly adhered to the unified technical requirements and standard methods specified in the DD2008-1 Specifications for Evaluation of Groundwater Pollution Geological Survey (China Geological Survey, 2008). The tested components included all mandatory indicators prescribed by DD2008-1. The relative error of all measured parameters was controlled within ±5%. To ensure the usability and reliability of the data, data preprocessing was performed. First, samples where both well depth and sampling depth exceeded 100 m were excluded. Subsequently, abnormal data were removed through charge balance verification (ionic balance error, EB < 5%) and carbonate equilibrium checks. Ultimately, 3,608 valid datasets were retained (Fig. 1a) for subsequent hydrochemical characteristic analysis.

1.3 Research methods

The data processing methods adopted in this study mainly include the Self-Organizing Map-K-Means (SOM-KM) method, Principal Component Analysis (PCA), and the Gibbs hydrogeochemical diagram method. A brief description is provided as follows:

The Self-Organizing Map-K-means (SOM-KM) method combines two clustering techniques: SOM for dimensionality reduction and K-means for subsequent clustering of SOM outputs. SOM is an unsupervised learning algorithm that projects high-dimensional data into a low-dimensional space while preserving topological relationships, and it is widely applied in data reduction and unsupervised classification (Mao et al., 2021). K-means clustering is a classical unsupervised learning method that efficiently partitions samples into predefined clusters through iterative computation (Haaf and Barthel, 2018). In this study, the coupled SOM-KM model was implemented on the MATLAB platform for clustering and visualization. Fifteen

key hydrochemical parameters were selected for analyzing the spatial variability of groundwater characteristics, including Total Dissolved Solids (TDS), pH, Ca^{2+}, Mg^{2+}, K^{+}, Na^{+}, Cl^{-}, SO_4^{2-}, HCO_3^{-}, NO_3^{-}, Fe, Mn^{2+}, As, F^{-}, and I^{-}.

Principal Component Analysis (PCA) examines interdependencies among variables by reducing the dimensionality of complex relational data and classifying datasets based on common factors (Safia and Abderrahmane, 2018). Its primary advantage lies in interpreting the original datasets through extracted key variables while minimizing information loss. Although PCA is fundamentally an exploratory descriptive method, this study specifically applies it to identify the dominant factors controlling groundwater chemical evolution. Using IBM SPSS Statistics 27 software, 16 hydrochemical parameters were systematically analyzed: pH, Total Dissolved Solids (TDS), Total Hardness (TH), Ca^{2+}, Mg^{2+}, K^{+}, Na^{+}, Cl^{-}, SO_4^{2-}, HCO_3^{-}, Fe, Mn^{2+}, As, F^{-}, I^{-}, and NO_3^{-}.

The Gibbs hydrogeochemical diagram method involves constructing a binary ratio plot of $Na^{+}/(Na^{+}+Ca^{2+})$ versus $Cl^{-}/(Cl^{-}+HCO_3^{-})$. This graphical approach quantitatively identifies the controlling effects of three major end-members on water chemistry: Atmospheric precipitation input, rock weathering, and evaporation-crystallization processes (Wang et al., 2018). Using Origin 2024 software, standardized Gibbs diagrams were generated. The spatial distribution patterns of sample points within the diagram domains were analyzed to determine the relative contributions of different hydrogeochemical processes to the formation of regional groundwater. These results provide geochemical mechanistic evidence for the dominant factors extracted by PCA.

2 Results and discussion

2.1 Soda water identification and characteristics

2.1.1 The identification of soda water and hydrochemical characteristics

In previous studies, the Russian chemist Shvartsev defined soda water as groundwater that satisfies the condition $Na^{+}/(Cl^{-}+SO_4^{2-})>1$ when concentrations are expressed in milliequivalents (Wang et al., 2009). Emil Boros (Boros and Kolpakova, 2018) classified groundwater using the criterion that the milliequivalent percentage of total cations or total anions exceeds 25%. He specified that soda water must meet two conditions simultaneously: (1) Na^{+} accounts for > 25% of total cations and is the dominant cation (ranks first), and (2) (HCO_3^{-} + CO_3^{2-}) accounts for > 25% of total anions and is the dominant anion (ranks first). However, as a transitional water type between leached water and typical continental saline-alkali water, defining soda groundwater solely based on milliequivalent concentrations often results in overly stringent data exclusion. The method that considers the milliequivalent percentages of both total cations and total anions is primarily applicable to inland saline surface water, and significant errors occur when it is applied to deep groundwater. Therefore, to improve identification accuracy, this study establishes the following refined criteria for soda water:

(1) [(HCO_3^{-} + CO_3^{2-}) - (Ca^{2+} + Mg^{2+})] > 0 meq%, and the value of this difference can be used to characterize the degree of "soda nature" (or "soda degree") of the groundwater.

(2) The meq% of (HCO_3^{-} + CO_3^{2-}) ranks first among all anions.

Groundwater is identified as soda water only when both of the above conditions are satisfied. Conversely, non-soda water refers to all shallow groundwater samples in the study area that do not satisfy both of the aforementioned identification criteria for soda water. This binary classification (soda water and non-soda water) covers all sampled groundwater without overlap, providing a clear and comprehensive framework for the subsequent comparative analysis of their spatial distribution characteristics and genetic mechanisms.

Since the soda water identification criteria inherently impose hydrochemical constraints, significant differences exist between soda water and non-soda water. To compare these differences, hydrochemical mapping of shallow groundwater in the North China Plain was conducted. Piper trilinear diagrams (Figs. 2a-2b) show that the anionic compositions of groundwater are predominantly HCO_3^{-}-type, $HCO_3^{-}\cdot Cl^{-}$-type, $HCO_3^{-}\cdot SO_4^{2-}$-type, and Cl^{-}-type, while the cationic compositions are mainly $Na^{+}\cdot Mg^{2+}$-type, $Ca^{2+}\cdot Mg^{2+}$-type, Na^{+}-type, and Ca^{2+}-type. In contrast, soda water exhibits a distinct HCO_3^{-}-Na^{+} signature, which is determined by the first identification criterion: [(HCO_3^{-} + CO_3^{2-}) - (Ca^{2+} + Mg^{2+})] > 0 meq%. Further analysis (Fig. 2c) demonstrates that when the soda degree index exceeds 0, groundwater typically shows HCO_3^{-}-Na^{+} characteristics, with occasional occurrences of SO_4^{2-}-Na^{+} or Cl^{-}-Na^{+} types. These variations primarily depend on whether (HCO_3^{-} + CO_3^{2-}) dominates the anion composition.

By constructing multi-parameter correlation diagrams, including those of TDS, soda degree,

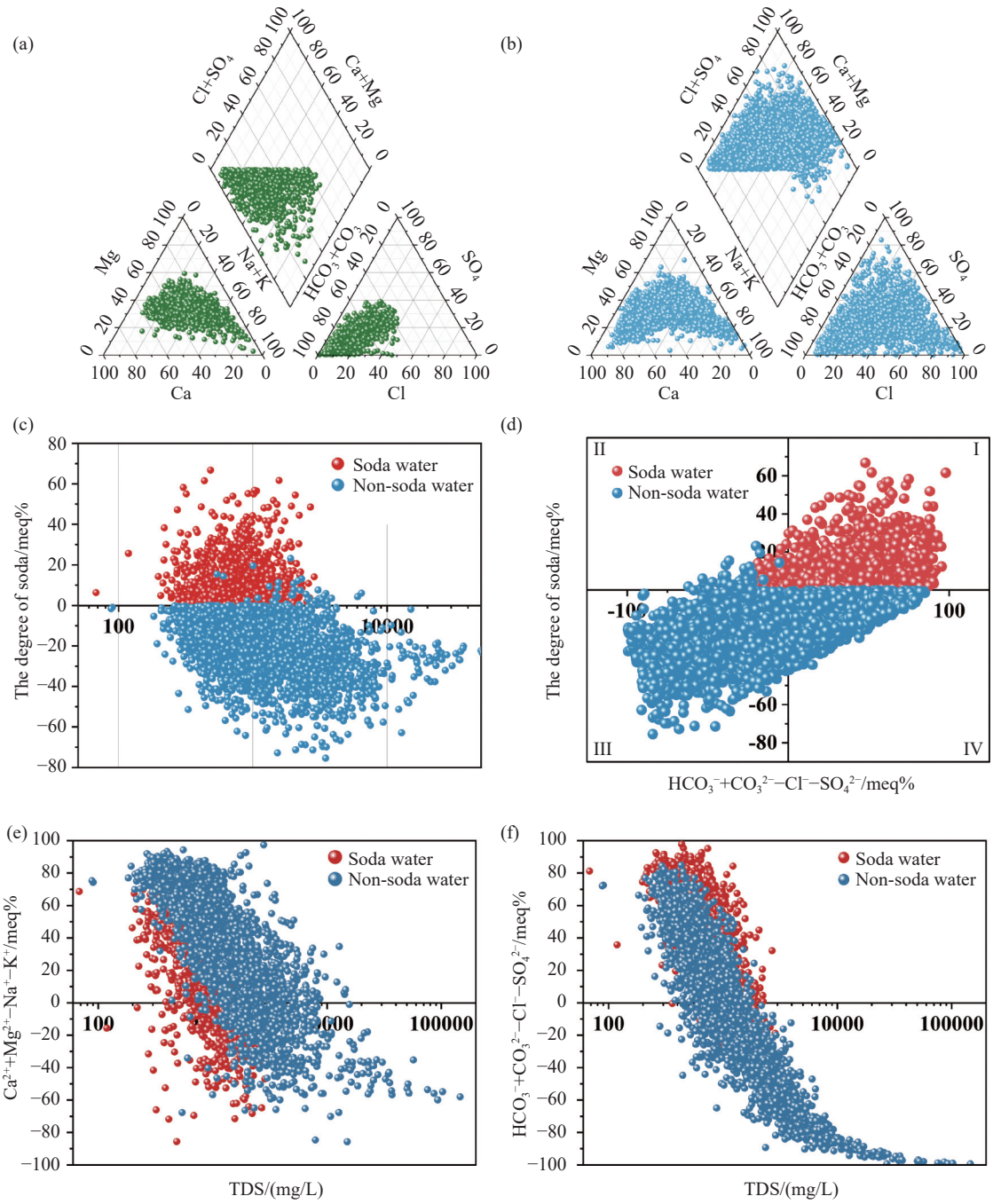


Fig. 2 Hydrogeochemical signatures of soda water vs. Non-soda water

Notes: a- Piper trilinear diagram for soda water; b- Piper trilinear diagram for non-soda water; c- Relationship between groundwater and TDS under different classification criteria; d- Relationship between soda degree and anion difference under different TDS contents; e- Relationship between major anion differences and TDS; f- Relationship between major cation differences and TDS.

and anion/cation internal differences (Figs. 2c-2f), it is found that the TDS of shallow soda water is mainly distributed between 0.2 g/L and 2.0 g/L, with a dominant concentration of approximately 0.8 g/L. This distribution is attributed to the second constraint condition. Since TDS exhibits a significant negative correlation with both anion internal difference and cation internal difference, it implies that as TDS increases, calcite and dolomite in groundwater gradually tend to reach saturation. The precipitation process of carbonate minerals inhibits the continuous increase in the concentrations of Ca^{2+} and Mg^{2+}. Under the same conditions, a higher Na^{+} concentration in the water body is more conducive to the formation of soda water. This also indicates that the identification criteria for soda water can effectively distinguish between sodification and salinization processes.

2.1.2 The associated components of soda water and non-soda water

To further clarify the differences in hydrochemical characteristics between soda water and non-soda water, statistical analysis was conducted on the data distribution of hydrochemical components and hydrochemical types across different regions, with details shown in Fig. 3. The analysis indicates that these two types of water bodies exhibit differences in hydrochemical parameters and environmental indicators. Through statistical summary, the comparative analysis involves 6 functional categories of core parameters (a total of 9 specific indicators), which are sorted out as follows:

(1) In terms of acid-base properties, the pH value of soda water is significantly higher than that of non-soda water (median: Soda water [e.g., 7.65], non-soda water [e.g., 7.48]), exhibiting typical weakly alkaline characteristics;

(2) Although the Sodium Adsorption Ratio (SAR) shows no significant difference between the two groups, the fluctuation range of SAR values for soda water is significantly narrower than that for non-soda water (range: Soda water [e.g., 0.26–19.42], non-soda water [e.g., 0.1–121.8]), which reflects the spatial stability of soda water hydrochemical composition;

(3) In terms of the redox environment, the redox potential (Eh) of soda water is significantly lower than that of non-soda water (median: Soda water [e.g., 0 mV], non-soda water [e.g., 89.65 mV]), indicating that it tends to be in a more reducing hydrogeochemical environment;

(4) Regarding the difference in characteristic component concentrations, the enrichment levels of As, F^-, and I^- in soda water are higher than those in non-soda water (F^- median: Soda water [e.g., 0.65 mg/L], non-soda water [e.g., 0.40 mg/L]), among which the difference in F^- concentration between the two groups is the most significant;

(5) In terms of pollutant indicators, the concentrations of NO_3^- and Mn in soda water are significantly lower than those in non-soda water (NO_3^- median: Soda water [e.g., 1.24 mg/L] vs. non-soda water [e.g., 5.24 mg/L]); Mn median: Soda water [e.g., 0.15 mg/L] vs. non-soda water [e.g., 0.21 mg/L]), and their fluctuation ranges are narrower than those of non-soda water, showing more stable low-pollution characteristics;

(6) There is no statistical difference in iron (Fe)

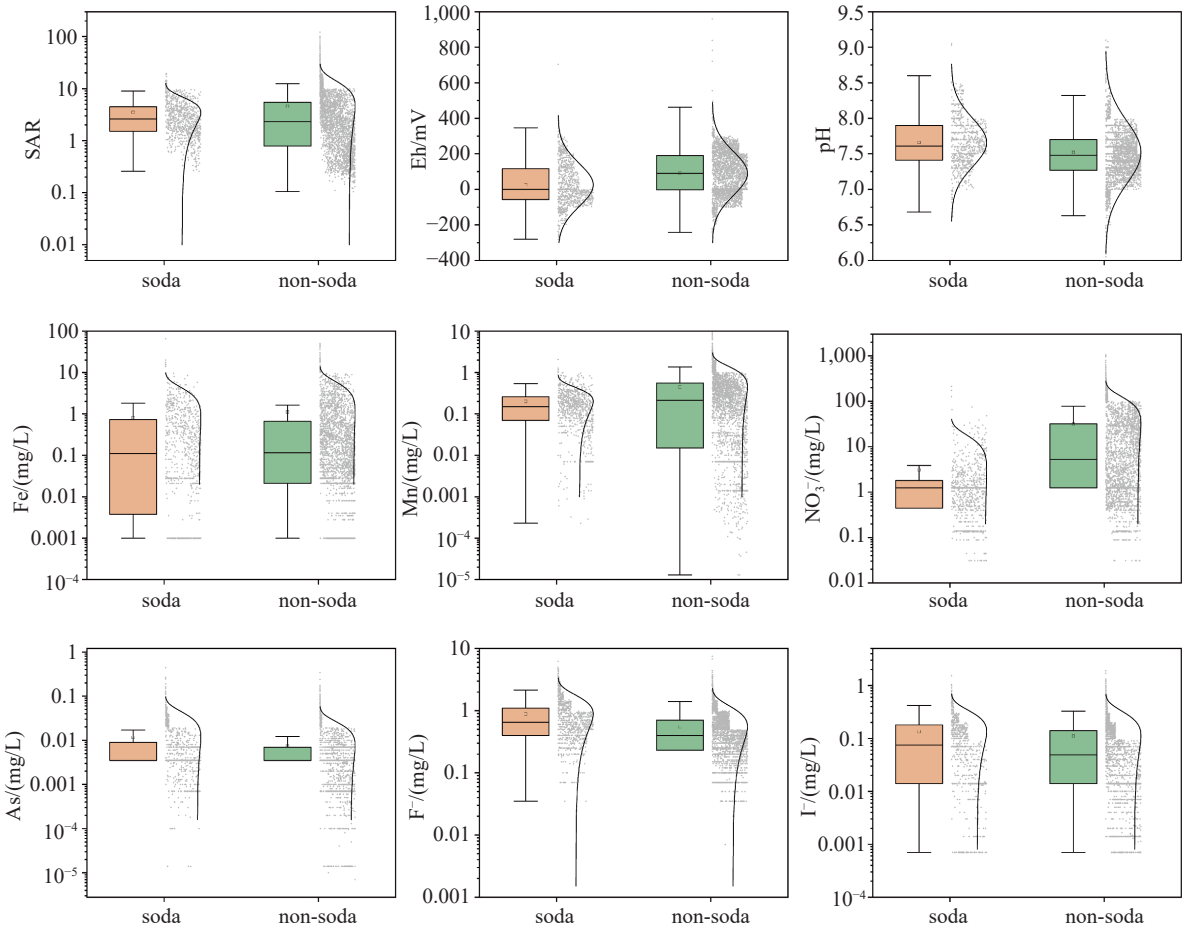


Fig. 3 Box plots of associated components in soda water vs. non-soda water

content between the two groups, but the fluctuation range of Fe concentration in soda water is significantly wider than that in non-soda water.

In conclusion, soda water has weakly alkaline and low-pollution properties, and is conducive to the associated occurrence of arsenic, fluoride, and iodide ions.

All the above comparative analyses between soda water and non-soda water are intended to highlight the unique hydrochemical characteristics, stable distribution range, and specific environmental conditions of soda water. These findings provide a critical foundation for the subsequent core research content of this manuscript: The analysis of the spatial distribution characteristics and genetic mechanisms of soda water in the North China Plain.

2.2 Typical distribution areas division and characteristics of soda water

2.2.1 Spatial distribution characteristics and zoning of soda water

Based on the aforementioned identification criteria, 868 soda water samples were identified from the shallow aquifers in the North China Plain, accounting for 24.06% of the total samples. The spatial distribution of soda water is shown in Fig. 4. In terms of spatial distribution, soda water exhibits an obvious zonal distribution pattern. It is mainly distributed in the mountain-front discharge zone of the Taihang Mountains and Yanshan Mountains in the North China Plain, as well as in the runoff area of the Yellow River paleochannel.

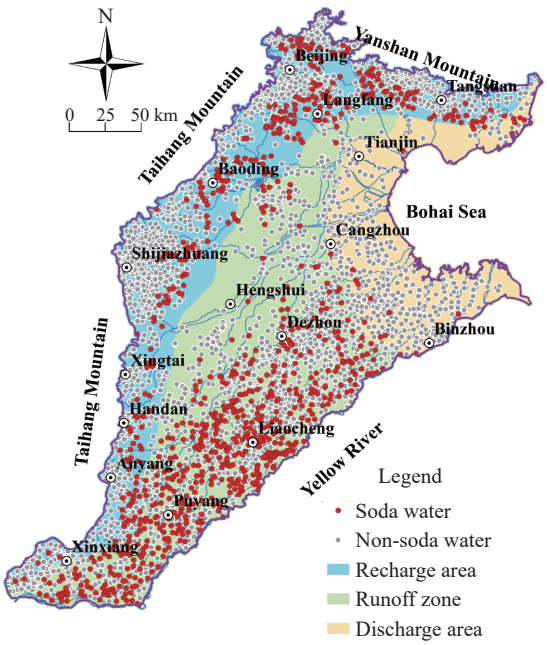


Fig. 4 Spatial distribution of soda water vs. non-soda water in unconfined aquifers across the north China Plain

To achieve more accurate zoning of soda groundwater, this study integrated the groundwater system classification (Fig. 5a) (Zhang and Fei, 2009) and geomorphological division (Fig. 5b) (Hou and Liu, 2020) of the North China Plain. The Self-Organizing Map-K-Means hybrid algorithm (SOM-KM) was applied for hydrochemical data clustering, with the results shown in Fig. 5d.

The concentrations of TDS, Na^+, Cl^- and SO_4^{2-} show a significant increasing trend along the diagonal from top-left to bottom-right. High pH values are concentrated in the top-right region. The concentration of Ca^{2+} increases from top-right to bottom-left, with notable enrichment in the left area. The concentrations of Mg^{2+}, HCO_3^-, and I^- decrease vertically from top to bottom. K^+, Fe, As, and NO_3^- maintain generally low background values but exhibit localized anomalies in the bottom-left corner. Mn^{2+} and F^- show horizontal gradients, with concentrations increasing from left to right.

Based on the prototype vectors obtained from the Self-Organizing Map (SOM) neural network training, the K-Means algorithm was further applied for clustering (Fig. 5e). The results show that soda water in the study area can be divided into two typical clustering units: Cluster 1 samples are mainly distributed in the upper-middle part of the feature space, while Cluster 2 samples are located in the lower part of the feature space (Table 2).

Combined with spatial distribution information for visual analysis (Fig. 5c), different clustering types exhibit clear distribution patterns in the hydrogeological system: Cluster 1 is concentrated in the piedmont alluvial-proluvial fan area, with some parts extending in a zonal pattern along the ancient Yellow River channel; Cluster 2 is enriched in the main regional groundwater runoff channels and the transition zone of the lower reaches of the ancient Yellow River system.

In summary, by integrating the groundwater system zoning, geomorphological division, and SOM-KM clustering results of soda water, this study finally merged regions with similar hydrochemical characteristics and hydrogeological backgrounds. Two typical study areas were delineated: (A) the mountain-front discharge zone and (B) the runoff-ancient Yellow River channel zone (Fig. 5f). This provides a basis for further exploring the zoning characteristics and formation mechanisms of soda water.

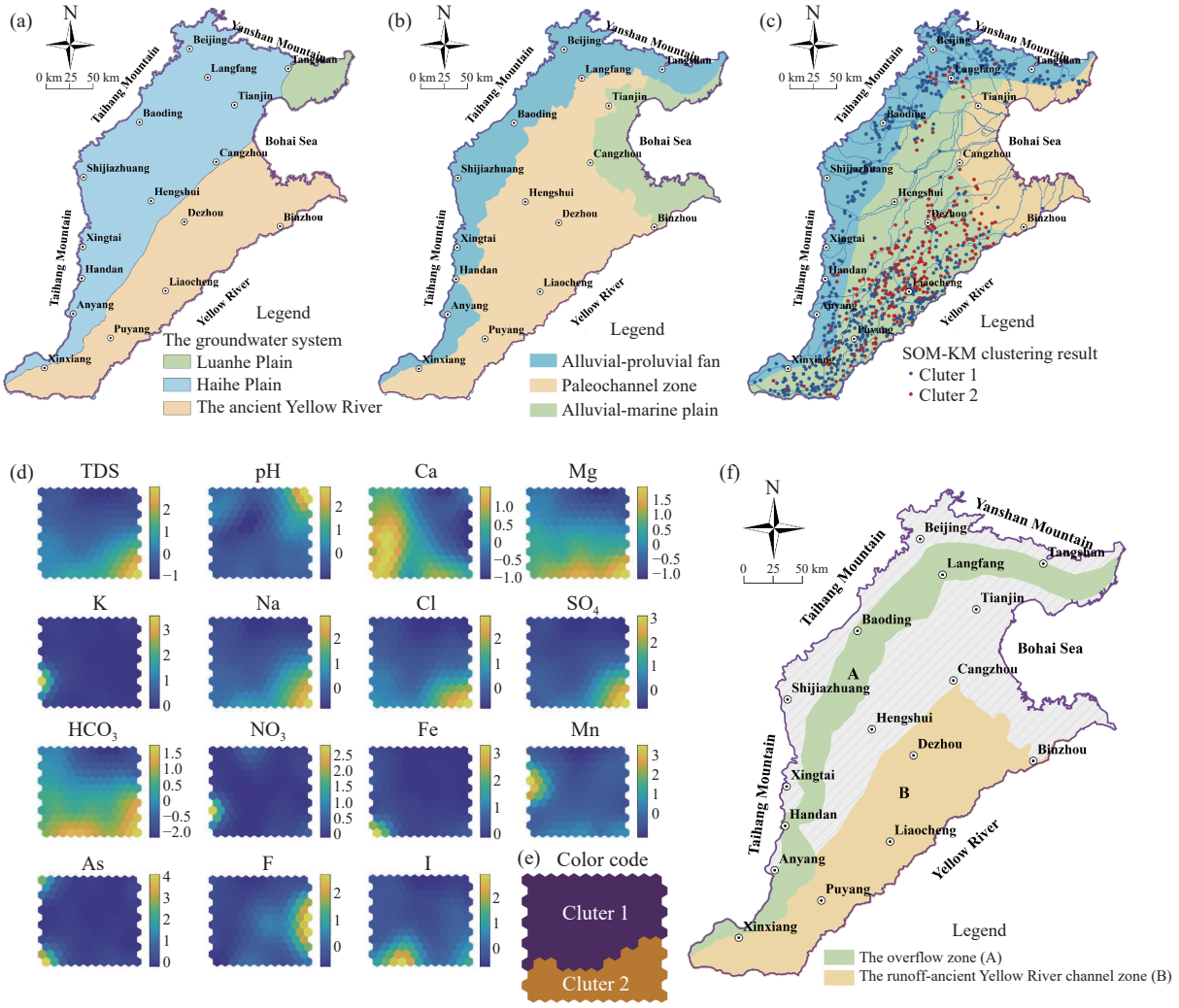


Fig. 5 Typification of soda water distribution in the North China Plain

Notes: a- Groundwater system zoning map of the North China Plain (Zhang and Fei, 2009); b- Landform classification map of the North China Plain (Hou and Liu, 2010); c- Spatial distribution map of SOM-KM clustering results; d- Plan view visualization result map; e- Clustering results of KM neurons; f- Typical regional division map.

Table 2 Breakdown analysis of spatial distribution patterns for soda water hydrochemical clusters

Category	Proportion	Spatial distribution	Component concentrations	Hydrochemical types	Summary
Cluster 1	70.74%	Mainly distributed in piedmont alluvial-proluvial fans and near the Yellow River paleochannel	TDS and main ion components have low concentrations; among them, pH, Ca^{2+}, and Mn^{2+} concentrations are at medium levels	HCO_3-Na HCO_3-Ca·Mg HCO_3-Na·Ca HCO_3-Na·Mg	Low TDS Low Fe High Mn
Cluster 2	29.26%	Mainly distributed near the groundwater system in the lower reaches of the ancient Yellow River	TDS and main ion components have relatively high concentrations; among them, Mg^{2+}, HCO_3^-, Fe, and I^- concentrations are at relatively high levels	HCO_3·Cl-Na HCO_3-Na·Mg HCO_3·Cl-Na·Mg HCO_3·SO_4-Na·Mg	High TDS High Fe High I^-

2.2.2 Characteristics of the hydrochemicalzone in soda water

The hydrochemical zoning results from SOM-KM clustering show that in the mountain-front discharge zone, Cluster I samples account for 91.87% and Cluster II samples account for 8.13%. In the runoff-ancient Yellow River channel zone, Cluster I samples account for 59.20% and Cluster II samples account for 40.80%. To compare the

main hydrochemical characteristics of the two zones, statistical analysis was conducted on the data distribution of hydrochemical components and hydrochemical types in each zone. Details are shown in Table 3, Fig. 6, and Fig. 7.

Zone A is the transitional area from the recharge zone to the runoff zone in the groundwater system of the Luanhe-Haihe Plain. Therefore, it is referred to as the mountain-front discharge zone. Its hydrogeological characteristics are characterized by a transition from a steep to a gentle terrain slope, a change in aquifer structure from single-layered to multi-layered, variations in water exchange intensity, and a slowdown in hydrodynamic conditions (Zhang et al., 2009). The hydrochemical characteristics are as follows: the concentrations of major cations and anions are low; HCO_3^- accounts for a high proportion of anions; and among other ion components, only NO_3^- exhibits a high concentration. Therefore, this zone can be summarized as a low-TDS, oxidizing zone dominated by relatively simple HCO_3^--type anions.

Table 3 Comparative evaluation of key hydrochemical parameters in Soda water from unconfined aquifers: Zone A vs. zone B across the North China Plain

Hydrochemical index	The mountain-front discharge zone (A)				The runoff-ancient Yellow River channel zone (B)			
	Max	Min	Average	Median	Max	Min	Average	Median
The degree of soda (meq%)	58.23	0.06	14.28	11.22	55.14	0.02	13.96	11.15
TDS (mg/L)	2,069	68	584.78	501.15	2,863.71	326	989.87	863.56
pH	9.05	6.82	7.68	7.63	9	6.62	7.65	7.60
Eh	346.10	−245	69.31	57.70	704	−281	−13.44	−20.05
Proportion of soda water	28.34%				60.14%			

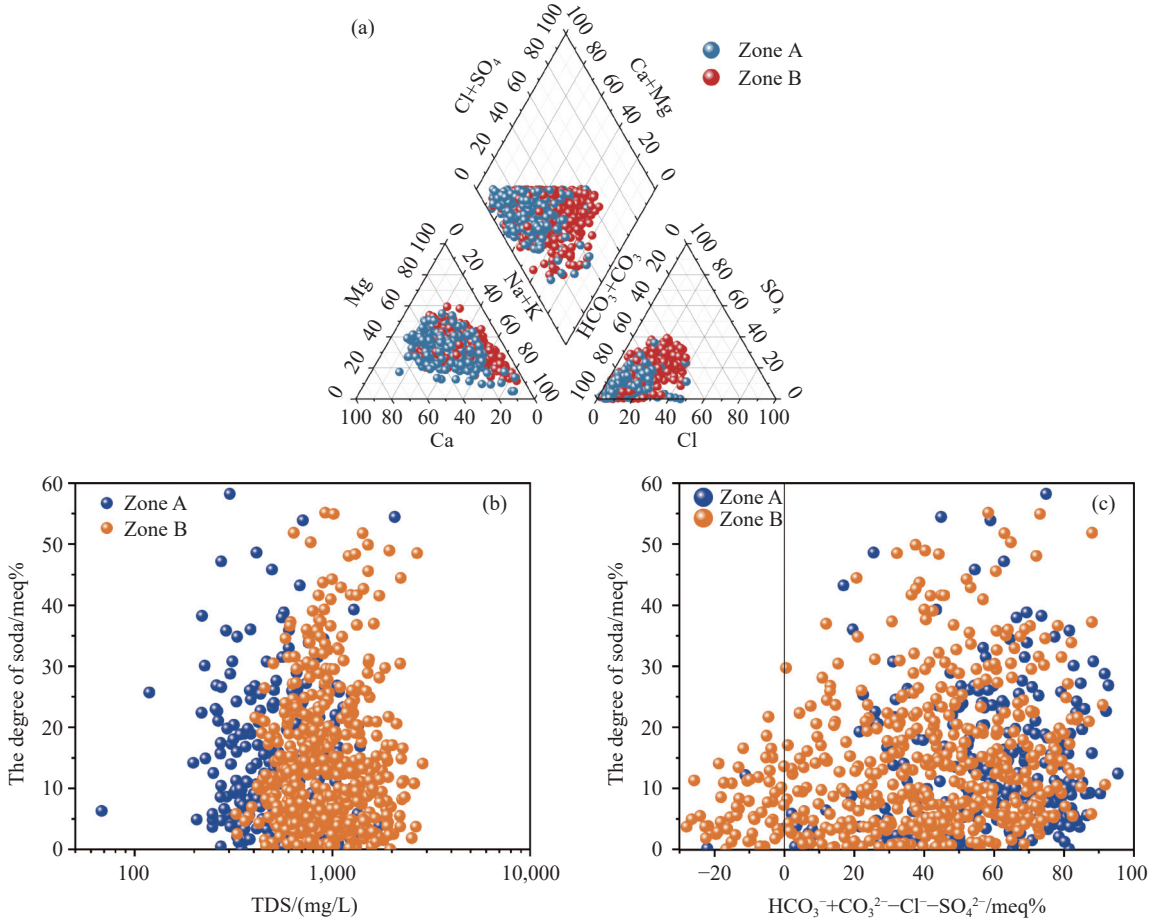


Fig. 6 Intercomparison of soda water hydrochemistry in Zone A vs. Zone B

Notes: a - Distribution map of shallow groundwater hydrochemical types in each zone; b - Relationship map between sodification degree and TDS; c - Relationship map between sodification degree and proportion of major anions.

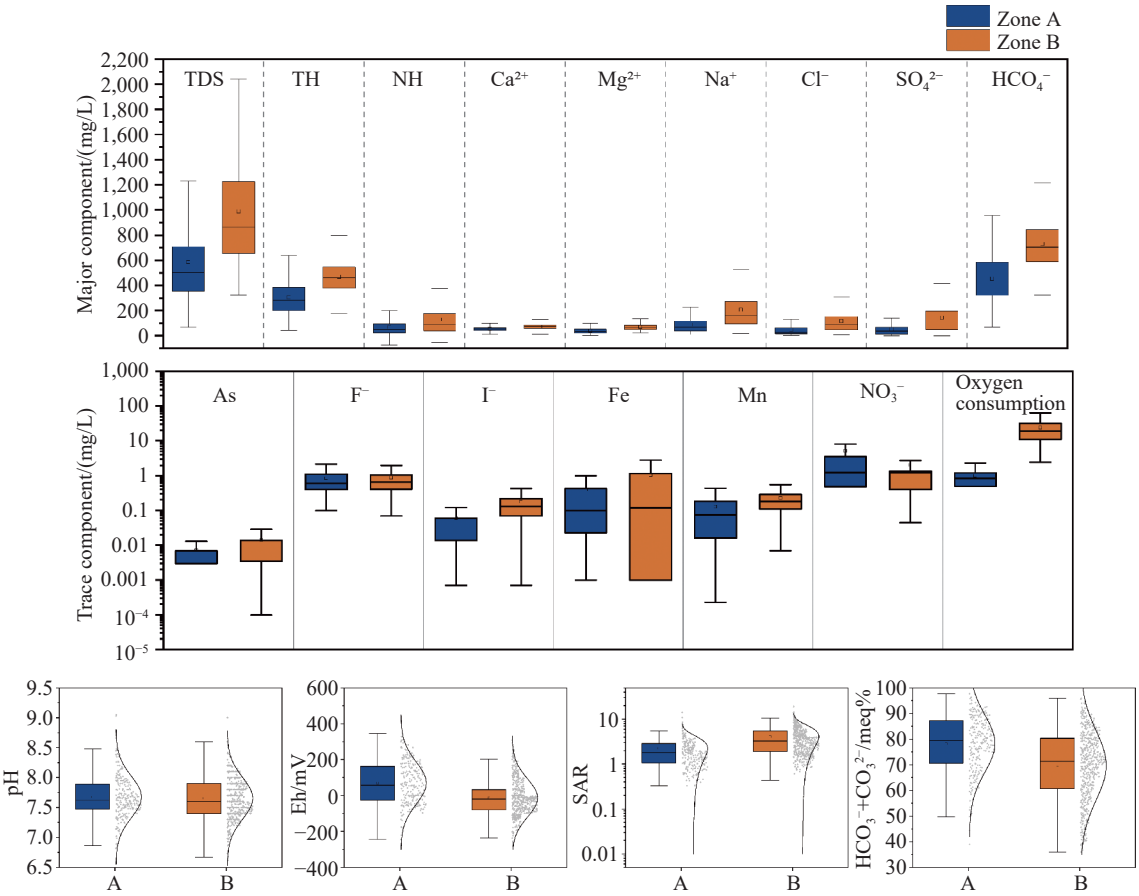


Fig. 7 Box plots of major ions and associated components in soda water: Zone A vs. Zone B

Zone B is the runoff area of the groundwater system in the lower reaches of the ancient Yellow River. Its hydrogeological characteristics are dominated by the Yellow River alluvial deposits, with sluggish hydrodynamic conditions (Zhang et al., 2009). The hydrochemical characteristics exhibit complex anionic compositions, including HCO_3^-, SO_4^{2-}, and Cl^-. The concentrations of major cations and anions are high. Among other ion components, As, I^-, Fe, Mn^{2+}, Total Hardness (TH), negative hardness, oxygen consumption, and Sodium Adsorption Ratio (SAR) also show high values. Therefore, this zone can be summarized as a high-TDS, weakly oxidizing zone with complex anions. The variations in redox potential (Eh) and TDS values between the two zones indicate that they are subject to distinct hydrogeological conditions. In addition, among the shallow groundwater in the North China Plain, soda water accounts for 28.34% of the total water samples in zone A and 60.14% of the total water samples in zone B. Soda water in zones A and B together accounts for 88.48% of the total soda water samples in the North China Plain, confirming the distinct zonal distribution of soda water.

Statistical analysis shows that the hydrochemical characteristics of the overflow zone (zone A) and the runoff-ancient Yellow River channel zone (zone B) share certain common features: Both have weakly alkaline pH and higher F^- concentrations compared to non-soda water, and TDS exerts a significant influence on ion composition. Although both zones are classified as HCO_3^--Na^+ type, their hydrochemical characteristics differ markedly due to variations in hydrodynamic conditions, water exchange intensity, and geological settings. Therefore, further analysis will focus on these two distinct zones: The overflow zone (zone A) and the runoff-ancient Yellow River channel zone (zone B).

2.3 The causes of soda water

2.3.1 The mountain-front discharge zone

Sixteen indicators, including pH, TDS, Total Hardness (TH), Ca^{2+}, Mg^{2+}, K^+, Na^+, Cl^-, SO_4^{2-}, HCO_3^-, Fe, Mn^{2+}, As, F^-, I^-, and NO_3^-, were selected for Principal Component Analysis (PCA) in the overflow zone. The results showed that the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity significance value for the 246 soda

water samples in this zone were 0.62 and < 0.05, respectively (note: Bartlett's test does not yield a value of 0; "0" typically refers to a significance level < 0.05, which is the standard expression for statistical significance), indicating a moderate correlation among variables and suitability for PCA (Ravish et al., 2020). A total of five principal components were extracted from the PCA of hydrochemical indicators in zone A (Table 4).

The results of the Rotated Factor Loading Matrix (Table 4) show that F1, with a variance contribution rate of 38.92%, is the dominant factor. The factor loadings of Na^+ and HCO_3^- in F1 reach 0.95 and 0.74, respectively, which are significantly higher than those in other principal components. This indicates that F1 is not only the core factor controlling regional water quality variation, but also has a significant positive correlation with the formation process of soda water. In F1, highly positive correlations are observed among TDS, Mg^{2+}, Na^+, Cl^-, SO_4^{2-}, HCO_3^-, and F^-, covering five of the eight major conventional groundwater ions. This suggests that the hydrochemical composition of groundwater is influenced by natural hydrogeochemical processes, and it is inferred that the groundwater chemistry in this zone is mainly controlled by mineral leaching processes.

F2 has a variance contribution rate of 13.03%. The factor loading of HCO_3^- ranks second, reflecting its secondary influence. In F2, pH has the highest loading coefficient, followed by Total Hardness (TH), Ca^{2+}, Mg^{2+}, and HCO_3^-. This indicates that carbonate rock dissolution also affects the formation of hydrochemical indicators. Atmospheric precipitation and CO_2 participate in the carbonate equilibrium system through water-rock interactions, facilitating the stable presence of HCO_3^- in a weakly alkaline environment (Wang and Zhang, 1998). Therefore, F2 is inferred to represent the dynamic carbonate equilibrium in this zone.

Notably, although F3 has a relatively low variance contribution rate, the loading value of HCO_3^- remains statistically significant, suggesting that HCO_3^- may participate in water-rock interactions through auxiliary mechanisms. In F3, HCO_3^- exhibits significant positive correlations with Fe and Mn^{2+}, indicating that this component may represent the activation of metal elements in a reducing hydrogeochemical environment. This process is closely driven by microbial metabolic activities: anaerobic microorganisms in the aquifer use organic matter as an electron donor, and iron-manganese oxides/hydroxides as electron acceptors for dissimilatory iron reduction. This reductive dissolution process not only releases adsorbed

Table 4 Principal component analysis results and factor loading matrix

	F1	F2	F3	F4	F5
pH	−0.01	**−0.81**	−0.02	−0.03	−0.10
TDS	**0.85**	0.37	0.17	0.12	−0.05
TH	0.49	**0.75**	0.34	0.09	−0.05
Ca^{2+}	0.13	**0.68**	0.47	0.28	0.01
Mg^{2+}	**0.61**	**0.63**	0.22	−0.05	0.10
K^+	0.02	0.09	−0.06	**0.72**	0.15
Na^+	**0.95**	0.04	0.05	0.00	0.01
Cl^-	**0.78**	0.12	0.23	0.13	0.06
SO_4^{2-}	**0.75**	0.18	0.02	0.10	−0.04
HCO_3^-	**0.74**	**0.54**	**0.21**	−0.04	−0.04
Fe	0.03	0.21	**0.69**	−0.07	0.06
Mn^{2+}	0.26	0.09	**0.82**	−0.09	−0.08
As	0.03	0.05	0.00	−0.02	**0.96**
F^-	**0.69**	−0.36	−0.02	−0.15	0.13
I^-	0.49	0.32	−0.33	−0.36	−0.17
NO_3^-	0.09	0.04	−0.06	**0.70**	−0.19
Eigenvalue	6.23	2.08	1.13	1.12	1.02
Variance contribution rate (%)	38.92	13.03	7.92	7.01	6.40
Cumulative variance contribution rate (%)	38.92	51.94	59.86	66.87	73.27

Fe and Mn^{2+} into groundwater but also synchronously generates HCO_3^- through the decomposition of organic matter, which explains the significant positive correlation between HCO_3^- and Fe/Mn^{2+} in F3. The weak reducing environment in the mountain-front discharge zone provides a suitable habitat for these functional microorganisms, confirming their participation in the hydrochemical evolution of soda water. Driven by dissimilatory iron reduction involving organic matter, the dissolution of iron-manganese oxides/hydroxides releases Fe and Mn^{2+}, accompanied by the synchronous generation of HCO_3^-.

In summary, the variance contribution rates of F1, F2, and F3 to groundwater hydrochemical indicators are 38.92%, 13.03%, and 7.92%, respectively. It is concluded that soda water in the study area is primarily influenced by mineral leaching and enrichment processes, while also being affected by regional carbonate equilibrium and the dissolution of iron-manganese oxides driven by dissimilatory iron reduction under reducing conditions.

The Saturation Index (SI) reflects the dissolution and equilibrium state of minerals. When SI > 0, the mineral is in a supersaturated state; when SI < 0, the mineral is in an undersaturated (dissolved) state; when SI = 0 or close to 0 (0 ± 0.5), the mineral is in an equilibrium state (Tang et al., 2012). The saturation indices of various mineral phases in the aquifer of the study area were calculated using PHREEQC software. The relationship between the saturation indices of mineral phases and TDS was further plotted.

As shown in Fig. 8a, when TDS exceeds 500 mg/L, most dolomite is in a supersaturated state, while calcite is in a nearly saturated or supersaturated state. With increasing TDS, fluorite, gypsum, and halite remain in an undersaturated (dissolved) state but show a trend toward equilibrium. A Gibbs diagram was also plotted. As shown in Fig. 9, the hydrochemical composition of soda water in Zone A is mainly controlled by mineral weathering processes, and its hydrochemical types are primarily concentrated in the medium-to-low TDS range.

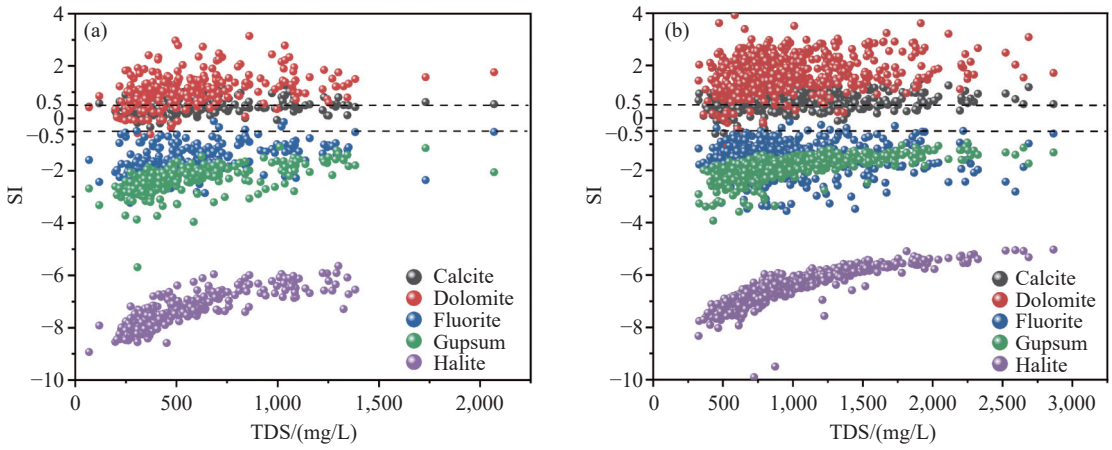


Fig. 8 Relationship between mineral Saturation Indices (SI) and TDS in soda water: Zone A (a) and Zone B (b)

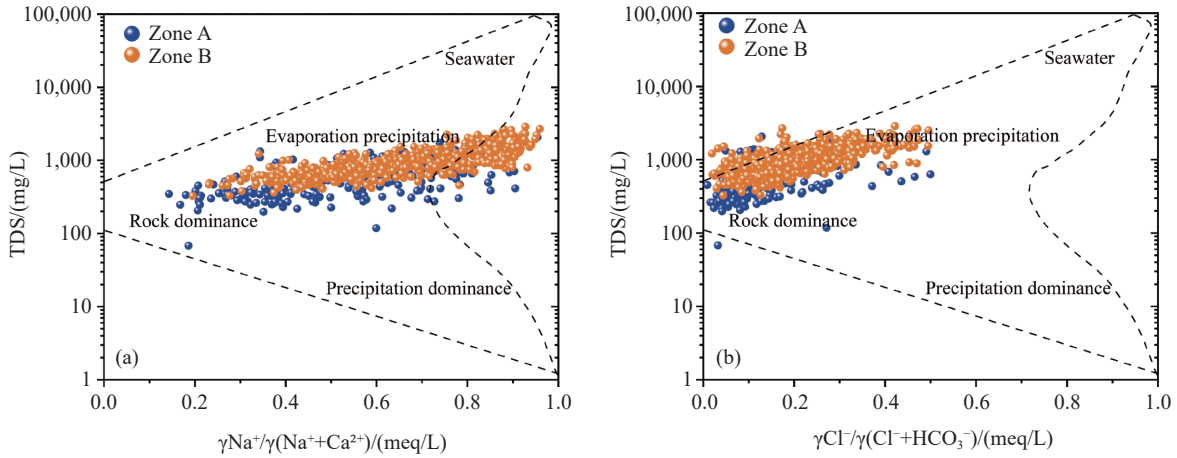


Fig. 9 Gibbs diagram analysis of dominant hydrogeochemical processes: Zone A (a) groundwater vs. Zone B (b)

2.3.2 The runoff-ancient Yellow River channel zone

Sixteen indicators, including pH, TDS, Total Hardness (TH), Ca^{2+}, Mg^{2+}, K^{+}, Na^{+}, Cl^{-}, SO_4^{2-}, HCO_3^{-}, Fe, Mn^{2+}, As, F^{-}, I^{-}, and NO_3^{-}, were selected for Principal Component Analysis (PCA) in the runoff-ancient Yellow River channel zone. The results showed that the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test significance value for the 522 soda water samples in this zone were 0.61 and < 0.05, respectively (note: Bartlett's test of sphericity does not yield a value of 0; "0" typically refers to a significance level < 0.05, which is the standard expression for statistical significance), indicating a certain degree of correlation among variables and suitability for PCA. A total of five principal components were extracted from the PCA of hydrochemical indicators in zone B (Table 5).

The results of the Rotated Factor Loading Matrix (Table 5) indicate that F1, with a variance contribution rate of 34.30%, is the dominant factor. The factor loadings of Na^{+} and HCO_3^{-} in F1 are significantly higher than those of other variables. This characteristic shows that F1 has a strong correlation with the formation process of soda water and plays a crucial role in explaining its origin.

In F1, TDS, Total Hardness (TH), Mg^{2+}, Na^{+}, Cl^{-}, SO_4^{2-}, and HCO_3^{-} exhibit significant positive correlations, covering five of the eight major conventional groundwater ions. As a comprehensive parameter representing total dissolved solids, TDS has a significantly higher loading coefficient than other components. Hydrogeological analysis shows that the Yellow River alluvial plain develops a shallow unconfined aquifer; combined with the strong recharge from precipitation in the flat terrain, the groundwater system is dominated by an evaporative discharge hydrogeochemical evolution model. It is inferred that F1 represents the influence of evaporative concentration in the study area.

Notably, although F3 and F5 have relatively low variance contribution rates, their HCO_3^{-} loading values are similar in magnitude but opposite in direction. This suggests that they may respectively represent differences in the occurrence forms of alkaline substances and ionic synergistic effects in the groundwater system of zone B, this finding is statistically significant for explaining the complexity of the soda water formation mechanisms in this region.

In F3, HCO_3^{-} exhibits significant positive correlations with Mn^{2+} and I^{-}, which may represent unique microbial geochemical processes in the downstream alluvial plain of the ancient Yellow River. Zone B's interbedded clay-sand sedimen-

Table 5 Principal component analysis results and factor loading matrix

	F1	F2	F3	F4	F5
pH	−0.06	**−0.61**	0.34	0.15	0.15
TDS	**0.98**	−0.02	0.07	0.04	0.04
TH	**0.62**	**0.66**	0.28	0.07	0.07
Ca^{2+}	−0.04	**0.87**	0.11	0.18	0.18
Mg^{2+}	**0.77**	0.32	0.27	−0.02	−0.02
K^{+}	−0.08	−0.04	0.18	−0.11	−0.11
Na^{+}	**0.94**	−0.25	−0.00	0.01	0.01
Cl^{-}	**0.86**	0.08	−0.05	0.08	0.08
SO_4^{2-}	**0.88**	−0.11	−0.06	−0.05	−0.05
HCO_3^{-}	**0.79**	0.03	**0.28**	0.04	0.04
Fe	0.11	0.03	0.19	**0.71**	**0.71**
Mn^{2+}	0.08	0.16	**0.72**	−0.04	−0.04
As	−0.05	0.14	−0.15	**0.79**	**0.79**
F^{-}	0.31	**-0.61**	−0.19	−0.26	−0.26
I^{-}	0.42	−0.12	**0.57**	0.15	0.15
NO_3^{-}	0.25	0.05	−0.09	0.21	0.21
Eigenvalue	5.49	2.40	1.22	1.17	1.17
Variance contribution rate (%)	34.30	14.98	7.61	7.31	7.31
Cumulative variance contribution rate (%)	34.30	49.28	56.89	64.20	64.20

tary structure results in sluggish groundwater flow and a strongly reducing environment, which is conducive to the accumulation of organic matter and the proliferation of anaerobic functional microorganisms. Specifically: ① Iron-manganese reducing bacteria use organic matter as an electron donor to drive the reductive dissolution of iron-manganese oxides/hydroxides, releasing Mn^{2+} and adsorbed I^- into groundwater; ② During microbial metabolism, CO_2 is produced and reacts with water to form HCO_3^-, leading to a synchronous increase in HCO_3^- concentration. This microbial-driven coupling process directly explains the significant positive correlation between HCO_3^-, Mn^{2+}, and I^- in F3, and is consistent with Zone B's enrichment of reduced components observed in hydrochemical analysis. Similar microbial geochemical processes have been verified in other ancient river channel aquifers with similar sedimentary and hydrogeological conditions (Zhang et al., 2013). In this region, Quaternary soils and clayey or silty clay deposits formed by Yellow River alluviation, together with anthropogenic inputs, lead to the formation of organic-rich surface soils. The decomposition of organic matter is accompanied by CO_2 production.

In F5, HCO_3^- exhibits negative correlations with K^+ and NO_3^-, suggesting the influence of agricultural activities on soda water. K^+ is an essential component of fertilizers in agricultural production; the application of potassium fertilizer can improve soil fertility and promote plant growth and development (Li et al., 2021). NO_3^- is a typical chemical component in groundwater, mainly originating from the mineralization of soil organic nitrogen, atmospheric deposition (rainfall), and human activities such as domestic sewage discharge and agricultural fertilization (Matiatos, 2016; Liu et al., 2015).

In summary, the variance contribution rates of F1, F3, and F5 to groundwater hydrochemical indicators are 34.30%, 7.61%, and 6.83%, respectively. Soda water in this zone is mainly influenced by evaporative concentration, as well as microbial geochemical processes and agricultural activities.

As shown in Fig. 8b, TDS is generally >500 mg/L, and most calcite and dolomite are in a supersaturated state. With increasing TDS, fluorite, gypsum, and halite remain in an undersaturated state but show a trend toward equilibrium. Further analysis using the Gibbs diagram (Fig. 9) shows that sample points in Zone B are concentrated in the high-value zone where $Na^+/(Na^+ + Ca^{2+}) > 0.6$, close to the evaporation-crystallization trend line. This further indicates that the zone is strongly influenced by evaporative concentration.

2.3.3 Comparative analysis

A comparative analysis of the spatial distribution, hydrochemical composition, and genetic mechanisms of soda water in the mountain-front discharge zone and the runoff-ancient Yellow River channel zone reveals significant differences in their occurrence environments, geochemical processes, and ecological effects.

From the perspective of hydrochemical characteristics, soda water in zone A is characterized by low TDS and a simple anionic composition. The associated components show significant enrichment of F^- and NO_3^-, indicating formation in an open hydrogeochemical system. The enrichment of NO_3^- in zone A is closely related to agricultural activities in the region: The mountain-front discharge zone is dominated by dryland farming, with extensive application of nitrogen fertilizers. Urea and ammonium bicarbonate are hydrolyzed by soil microorganisms to produce NH_4^+, which is further nitrified to NO_3^- under aerobic conditions; nitrate-based fertilizers directly increase soil NO_3^- content. The open hydrogeochemical system in zone A facilitates the leaching of NO_3^- into groundwater by precipitation, resulting in the significant enrichment of NO_3^- in soda water. Statistical data from local agricultural yearbooks show that the average nitrogen fertilizer application rate in zone A is 150–200 kg/ha·a, which is significantly higher than that in zone B, further confirming the direct correlation between agricultural nitrogen input and NO_3^- enrichment in zone A soda water. Principal component analysis shows that F1 in Zone A is dominated by Na^+, HCO_3^-, and TDS, supporting a synergistic mechanism of mineral leaching and calcite precipitation. In the Gibbs diagram, samples from zone A are clustered in the rock weathering endmember, demonstrating that hydrochemical evolution is primarily controlled by water-rock interactions.

In contrast, soda water in zone B exhibits high TDS and a complex anionic composition. Reduced components such as As and I^- are significantly enriched, suggesting formation in a relatively closed reducing environment. F1 in zone B is dominated by TDS, Cl^-, and SO_4^{2-}; combined with the Gibbs diagram showing concentrated distribution in the evaporation endmember, this indicates that evaporative concentration is the dominant process. Additionally, the co-enrichment of Fe and Mn in zone B, along with negative redox potential (Eh) values, confirms the critical influence of

dissimilatory iron reduction-driven metal oxide dissolution processes on hydrochemical evolution.

The genetic differences between soda water in the two zones reflect the spatial heterogeneity of regional hydrogeological conditions. Zone A, located in the mountain-front discharge zone transitioning from piedmont alluvial-proluvial fans to the plain, is primarily shaped by mineral weathering, dissolution, and leaching processes. During hydrochemical evolution, TDS gradually increases; upon reaching the mountain-front discharge zone, hydrodynamic conditions change significantly. Calcite and dolomite gradually reach saturation and precipitate, restricting further enrichment of calcium (Ca^{2+}) and magnesium (Mg^{2+}). Meanwhile, Na^{+} becomes relatively enriched, resulting in the formation of soda water. The slightly positive Saturation Indices (SIs) of calcite and dolomite indicate that precipitation is a key hydrochemical process controlling soda water formation.

In contrast, zone B is situated in an interbedded clay-sand structure formed within the ancient Yellow River sedimentary system. This structure results in sluggish groundwater hydrodynamic conditions, and strong evaporative concentration serves as the dominant driving force of hydrochemical evolution. It causes a significant increase in TDS, leading to the supersaturation and precipitation of calcite and dolomite, which limits further enrichment of calcium (Ca^{2+}) and magnesium (Mg^{2+}) ions in the water. Meanwhile, Na^{+} becomes increasingly enriched due to continuous evaporation, ultimately forming soda water. Therefore, intense evaporative concentration and the resulting carbonate precipitation are the main controlling mechanisms for soda water formation in Zone B.

2.4 Research limitations and future prospects

2.4.1 Research limitations

This study has made progress in clarifying the identification criteria, zonal distribution, and genetic mechanisms of shallow soda water in the North China Plain, but there are still several limitations that need to be acknowledged:

(1) Timeliness of data: The dataset used in this study was collected during 2006–2009. Although it can reflect the natural background geochemical characteristics of soda water in that period, it fails to capture the dynamic changes in soda water induced by subsequent large-scale anthropogenic disturbances and climate change. Therefore, the conclusions cannot fully reflect the current status of soda water under the combined impacts of natural processes and human activities.

(2) Insufficient coverage of deep groundwater: This study only focuses on shallow groundwater with a sampling depth of less than 100 m and does not involve deep groundwater systems. Deep and shallow groundwater are affected by distinct hydrodynamic conditions and geochemical processes, so the findings of this study cannot be generalized to deep soda water, resulting in limitations in the comprehensiveness of regional soda water research.

(3) Insufficient evidence for microbial mechanisms: Although this study infers that microbial geochemical processes may participate in the modification of soda water components in the runoff-ancient Yellow River paleochannel zone (zone B), there is a lack of direct evidence from microbial molecular experiments. The specific role and mechanism of microorganisms in soda water formation and evolution remain to be further verified.

(4) Insufficient analysis of ecological impacts: Insufficient ecological impact analysis: The study only notes soda water as a driver of soil salinization but lacks quantification of its impacts on soil physicochemical properties. Due to the lack of extensive systematic soil data across the North China Plain (scattered local data are insufficient for reliable analysis) and constraints in time and cost, research on soda water's impacts on crop growth is also absent, leading to incomplete understanding of its regional ecological effects.

2.4.2 Future prospects

To address the above limitations and further deepen the research on regional soda water, the following future research directions are proposed:

(1) Expand the research scope to deep groundwater: Conduct systematic sampling and hydrochemical analysis of deep groundwater in the study area, compare the differences in hydrochemical characteristics, spatial distribution patterns, and genetic mechanisms between deep and shallow soda water, and construct a comprehensive framework for understanding regional soda water systems.

(2) Supplement microbial molecular experimental studies: In typical soda water distribution zones (especially zone B), carry out high-throughput sequencing of microbial communities and functional gene detection, clarify the composition of dominant microbial groups and their functional characteristics, and reveal the role and mechanism of microbial processes in soda water formation and evolution.

(3) Carry out long-term monitoring and dynamic evolution research: Establish long-term monitoring sites in key soda water zones, integrate post-2015 recent groundwater monitoring data with the 2006–2009 historical background data of this study, focus on analyzing the impacts of climate change and human activities on soda water evolution, and clarify the driving mechanisms of anthropogenic disturbances on groundwater sodification.

(4) Strengthen ecological impact research: Collect extensive and systematic soil data in the study area to quantify the effects of soda water on soil physicochemical properties; carry out targeted field surveys or controlled experiments to explore soda water's impacts on crop growth, thereby improving the understanding of regional soda water ecological effects and providing more comprehensive support for ecological protection and agricultural management.

3 Conclusions

This study focuses on soda water in shallow groundwater of the North China Plain. Based on clarifying the identification criteria for soda water, it identifies and analyzes the spatial distribution characteristics of soda water, delineates typical distribution areas, and preliminarily discusses differences in the genetic mechanisms of soda water across different zones. The main research conclusions are as follows:

(1) The identification criteria for soda water are further clarified: $[(HCO_3^- + CO_3^{2-}) - (Ca^{2+} + Mg^{2+})] > 0$ meq% and the meq% of $(HCO_3^- + CO_3^{2-})$ ranks first among anionic equivalents. According to these criteria, shallow soda water samples in the North China Plain account for 24.06% of the total samples. The TDS (total dissolved solids) of shallow soda water is mainly distributed between 0.2 g/L and 2 g/L, with the dominant concentration around 0.8 g/L, which can effectively distinguish it from salinized water.

(2) The zoning results of shallow groundwater soda water in the North China Plain show that soda water has obvious zonal distribution characteristics, which can be divided into the discharge zone (zone A) and the runoff-ancient Yellow River paleochannel zone (zone B). In Zone A, the dominant anion is HCO_3^-, with a median TDS of 501.15 mg/L, accompanied by components such as F^- and NO_3^-. In contrast, the anion types in Zone B are complex; in addition to HCO_3^-, they also include SO_4^{2-} and Cl^-, with a median TDS of 863.56 mg/L, and are enriched in reduced components such as As, F^-, I^-, Fe, and Mn.

(3) Genetic analysis shows that Zone A is a key area where groundwater transitions from a single-layered aquifer structure to a multi-layered aquifer system, accompanied by changes in hydrodynamic conditions. The soda water in this zone is mainly controlled by the combined processes of mineral weathering and leaching, as well as calcite and dolomite precipitation. In contrast, the soda water in Zone B is mainly formed under the condition of enhanced evaporative concentration, and the precipitation of calcite and dolomite serves as its dominant controlling process.

The three core findings of this study—an accurate dual-index identification system for soda water, systematic zonal distribution across the North China Plain, and differentiated genetic mechanisms—all originate from the scientific interpretation of the 2006–2009 historical dataset. This study clarifies that soda water formation in the region is mainly driven by natural hydrochemical processes with relative stability and provides an important reference baseline for evaluating anthropogenic impacts on groundwater sodification.

Acknowledgements

This research was funded by the National Land and Resources Survey Project (1212011121170): "Comprehensive Investigation and Assessment of Groundwater Pollution in China".

References

Azaria SL, Julianna S, Emil B. 2023. Review of chemical properties of inland soda and saline waters in East Africa (rift valley region). Journal of Hydrology: Regional Studies, 46: 101323. DOI：10.1016/j.ejrh.2023.101323.

Boros E, Kolpakova M. 2018. A review of the defining chemical properties of soda lakes and pans: An assessment on a large geographic scale of Eurasian inland saline surface waters. PLoS ONE, 13(8): 1−20. DOI：10.1371/journal.pone.0202205.

Chae GT, Yun ST. 2006. Hydrogeochemistry of sodium-bicarbonate type bedrock groundwater in the Pocheon spa area, South Korea: Water-rock interaction and hydrologic mixing. Journal of Hydrology, 321(1): 326−343. DOI：10.1016/j.jhydrol.2005.08.006.

China Geological Survey. 2008. Specification for

geological investigation and assessment of groundwater pollution: DD 2008-01. Beijing: China Geological Survey. (in Chinese)

Chen P, Wang W, Wang GM, et al. 2022. Distribution characteristics of soil salinization in the east coastal plain of Cangzhou. North China Geology, 45(3): 36−43. (in Chinese) DOI：10.19948/j.12-1471/P.2022.03.06.

Chen ZY, Qi JX, Xu JM, et al. 2003. Paleoclimatic interpretation of the past 30 ka from isotopic studies of the deep confined aquifer of the North China plain. Applied Geochemistry, 18(7): 997−1009. DOI：10.1016/S0883-2927(02)00206-8.

Cooper M, Tweed SO, Ahearne D, et al. 2006. Controls on chemistry during fracture-hosted flow of cold CO_2-bearing mineral waters, Daylesford, Victoria, Australia: Implications for resource protection. Applied Geochemistry, 21(2): 289−304. DOI：10.1016/j.apgeochem.2005.09.011.

Deocampo DM. 2010. The geochemistry of continental carbonates. Developments In Sedimentology, 62(1): 1−60. DOI：10.1016/s0070-4571(09)06201-3.

Dutova E. 2020. Geochemistry of fresh groundwater in the Altai-Sayan folded area and adjacent areas of the West Siberian plate. Applied Geochemistry, 120: 104673. DOI：10.1016/j.apgeochem.2020.104673.

Grobe M, Machel HJ. 2002. Saline groundwater in the Münsterland Cretaceous Basin, Germany: Clues to its origin and evolution. Marine and Petroleum Geology, 19(3): 307−322. DOI：10.1016/S0264-8172(02)00019-3.

Haaf E, Barthel R. 2018. An inter-comparison of similarity-based methods for organisation and classification of groundwater hydrographs. Journal of Hydrology, 559: 222−237. DOI：10.1016/j.jhydrol.2018.02.035.

Han DM, Currell MJ, Guo HM. 2021. Controls on distributions of sulphate, fluoride, and salinity in aquitard porewater from the North China Plain: Long-term implications for groundwater quality. Journal of Hydrology, 603: 126828. DOI：10.1016/j.jhydrol.2021.126828.

Hanor JS, Wendeborn FC, et al. 2023. Origin of sodium bicarbonate groundwaters, Southern Hills Aquifer System, USA by silicate hydrolysis. Applied Geochemistry, 148: 105512. DOI：10.1016/j.apgeochem.2022.105512.

Han SB, Zhou YZ, Zheng Y, et al. 2024. Formation mechanism and source apportionment of hydrochemical components in groundwater in the Yinchuan Plain. Environmental Science, 45(8): 4577−4588. (in Chinese) DOI：10.13227/j.hjkx.202308211.

Hou CT, Liu XD. 2010. Atlas of hydro-geo-environmental conditions in the North China Plain. Beijing: Geological Publishing House. (in Chinese)

Huang YH, Yang J, Yu X, et al. 2025. Hydrogeochemical analysis and paleo-hydrogeological modeling of shallow groundwater salinization processes in North China Plain. Journal of Hydrology, 651: 132616. DOI：10.2139/ssrn.4720574.

Irén Varsányi, Matray JM, Lajos Ó Kovács. 1997. Geochemistry of formation waters in the Pannonian Basin (southeast Hungary). Chemical Geology, 140(1): 89−106. DOI：10.1016/S0009-2541(97)00045-4.

Lei M, Zhou JL, Liang X, et al. 2022. Hydrochemical characteristics of pore water and genesis of soda water in the middle of the northern piedmont of Tianshan Mountain, Xinjiang. Earth Science, 47(2): 674−688. (in Chinese) DOI：10.3799/dqkx.2021.027.

Lepokurova OE, Shvartsev SL. 2019. Geochemistry of soda waters of the Chulym-Yenisei Artesian Basin (West Siberia). Russian Geology and Geophysics, 60(5): 558−569. DOI：10.15372/RGG2019044.

Lepokurova OE. 2020. Sodium-bicarbonate groundwaters in southeastern West Siberia, Russia: Compositions, types, and formation conditions. Applied Geochemistry, 116: 104579. DOI：10.1016/j.apgeochem.2020.104579.

Li J, Zou SZ, Zhao Y, et al. 2021. Major ionic characteristics and factors of karst groundwater at Huixian Karst Wetland, China. Environmental Science, 42(04): 1750−1760. (in Chinese) DOI：10.13227/j.hjkx.202009027.

Liu F, Song XF, Yang LH, et al. 2015. The role of anthropogenic and natural factors in shaping the geochemical evolution of groundwater in the Subei Lake Basin, Ordos energy base, Northwestern China. Science of the Total

Environment, 538: 327−340. DOI：10.1016/j.scitotenv.2015.08.057.

Liu HY, Liu MH, Zhang WM, et al. 2022. Distribution and fractionation of rare earth elements in high fluoride groundwater from the North China Plain. Earth Science Frontiers, 29(03): 129−144. (in Chinese) DOI：10.13745/j.esf.sf.2021.7.24.

Ma M, Zuo Z, Han YD, et al. 2025. Origin of surface substrate for soil salinization and alkalization in the Songnen Plain. Natural Resources and Remote Sensing, 37(2): 128−139. (in Chinese). DOI：10.6046/zrzyyg.2023330.

Mansouri, Katharina, Greupner, et al. 2024. Acid-base balance in healthy adults: Beneficial effects of bicarbonate and sodium-rich mineral water in a randomized controlled trial: The bicarbo water study. Journal of Nutrition and Metabolism, 2024(1): 1−14. DOI：10.1155/2024/3905500.

Mao HR, Wang GC, Rao Z, et al. 2021. Deciphering spatial pattern of groundwater chemistry and nitrogen pollution in Poyang Lake Basin (Eastern China) using self-organizing map and multivariate statistics. Journal of Cleaner Production, 329: 129697. DOI：10.1016/j.jclepro.2021.129697.

Matiatos I. 2016. Nitrate source identification in groundwater of multiple land-use areas by combining isotopes and multivariate statistical analysis: A case study of Asopos basin (Central Greece). Science of the Total Environment, 541: 802−814. DOI：10.1016/j.scitotenv.2015.09.134.

Meng SH. 2011. Research of groundwater vulnerability assessment and pollution prevention and treatment regionalization in North China Plain. M. S. thesis. Beijing: Chinese Academy of Geological Sciences. (in Chinese)

Qian TW, Cao YQ, Liu SX. 1995. Discussion on the thermodynamic model of shallow soda water formation in Qian'an Area, Jilin Province. Hydrogeology and Engineering Geology, 22(6): 22−26. (in Chinese) DOI：10.16030/j.cnki.issn.1000-3665.1995.06.010.

Ravish S, Setia B, Deswal S. 2020. Groundwater quality analysis of northeastern haryana using multivariate statistical techniques. Journal of the Geological Society of India, 95(4): 407−416. DOI：10.1007/s12594-020-1450-z.

Safia K, Abderrahmane B. 2018. Multivariate statistical characterization of groundwater quality in Fesdis, East of Algeria. Journal of Water and Land Development, 37(1): 65−74. DOI：10.2478/jwld-2018-0026.

Shvartsev SL, Wang YX. 2006. Geochemistry of sodic waters in the Datong intermountain basin, Shanxi Province, Northwestern China. Geochemistry International, 44(10): 1015−1026. DOI：10.1134/S0016702906100065.

Si CS, Zhang RH, Yao GS, et al. 2016. Tectonism and hydrocarbon preservation conditions of Qianbei depression and its margin. Journal of China University of Mining and Technology, 45(5): 1010−1021. (in Chinese) DOI：10.13247/j.cnki.jcumt.000493.

Tang H, Chen J, Qian H. 2012. Application of saturation index to research of water-rock interaction and its sensitivity analysis. Journal of Water Resources and Hydraulic Engineering, 23(06): 180−183. (in Chinese)

Talib MA, Tang Z, Shahab A, et al. 2019. Hydrogeochemical characterization and suitability assessment of groundwater: A case study in Central Sindh, Pakistan. International Journal of Environmental Research and Public Health, 16(5): 886. DOI：10.3390/ijerph16050886.

Wang JC, Zhang Y, Wen JL, et al. 2015. Temporal and spatial changing features of climate in North China Plain. Geoscience, 29(2): 299−306. (in Chinese) DOI：10.3969/j.issn.1000-8527.2015.02.012.

Wang YX, Shvartsev SL, So CL. 2009. Genesis of arsenic/fluoride-enriched soda water: A case study at Datong, Northern China. Applied Geochemistry, 24(4): 641−649. DOI：10.1016/j.apgeochem.2008.12.015.

Wang YX, Li JX, Ma T, et al. 2021. Genesis of geogenic contaminated groundwater: As, F and I. Critical Reviews in Environmental Science and Technology, 51(24): 1−39. DOI：10.1080/10643389.2020.1807452.

Wang ZR, Tian X, Wu X. 2018. Hydrochemical characteristics and quality assessment of shallow groundwater and CBM co-produced water in the Shizhuangnan block, Qinshui Basin, China. Environmental Earth Sciences, 77(3): 57. DOI：10.1007/s12665-017-7212-6.

Wang ZH, Zhang SJ. 1998. Discovery and characteristics of high-mineralization sodium bicarbonate-type water in the karamay oilfield. Petroleum Geology and Experiment, 20(1): 39−43. (in Chinese) DOI：10.11781/sysydz199801039.

Wu C. 1999. The impact of river channel changes in the North China Plain on soils and soil salinization. Geography and Territorial Research, 15(4): 70−75. (in Chinese)

Xing LN, Guo HM, Wei L, et al. 2012. Evolution feature and gensis of fluoride groundwater in shallow aquifer from North China Plain. Journal of Earth Science and Environment, 34(04): 57−67. (in Chinese) DOI：10.3969/j.issn.1672-6561.2012.04.008.

Xiong Y. 1979. The ecosystem of the Huang-Huai-Hai Plain. Soils, 02: 41−44. (in Chinese)

Xu PP, Feng WW, Qian H, et al. 2019. Hydrogeochemical characterization and irrigation quality assessment of shallow groundwater in the Central-Western Guanzhong Basin, China. International Journal of Environmental Research and Public Health, 16(9): 1492. DOI：10.3390/ijerph16091492.

Zhang LP, Xie XJ, Li JX, et al. 2013. Hydrochemical and geochemical investigations on high arsenic groundwater from Datong Basin, Northern China. Asian Journal of Ecotoxicology, 8(2): 215−221. DOI：10.7524/AJE.1673-5897.20121223002.

Zhang ZJ, Fei YH, Chen ZY. 2009. Investigation and evaluation of sustainable utilization of groundwater in the North China Plain. Beijing: Geological Publishing House. (in Chinese)

Zhi CS, Bill XH, Chang WB, et al. 2023. Enrichment mechanism of fluoride and iodine in saline groundwater in the lower flood plain of the Yellow River, Northern China. Journal of Hydrology, 621: 129529. DOI：10.1016/j.jhydrol.2023.129529.