Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

COUPLING GIS-BASED WATER QUALITY INDEXING WITH AN ELECTRICAL-CONDUCTIVITY SURROGATE MODEL FOR LOW-COST GROUNDWATER SCREENING IN AN ARID INDUSTRIAL DISTRICT: A PILOT STUDY IN BAIJI, SALAH AL-DIN GOVERNORATE, IRAQ

Abbas Ali Kanoosh
Dept. of Environmental Engineering,
College of Engineering, Tikrit University, 34001, Iraq
Email: Kanoosh.abbasali@tu.edu.iq

Abstract

Groundwater is the main water source for rural and peri-urban communities in arid Iraq, and routine laboratory testing of well networks is not financially or logistically feasible at a frequency that would ensure a safe supply. This exploratory pilot study assesses the quality of drinking water in twelve wells in Baiji District, Salah Al-Din Governorate, an industrial area containing one of the largest refineries in Iraq, and examines whether electrical conductivity (EC) can serve as a low-cost surrogate for index computation. Physicochemical parameters (pH, EC, TDS, Ca^{2+}, Mg^{2+}, Na^{+}, Cl^{-}, HCO_3^{-}, SO_4^{2-}, NO_3^{-}) were used to compute a Weighted Arithmetic Water Quality Index (WAWQI) referenced to WHO guidelines, together with seven irrigation-suitability indices. Spatial structure was quantified using global and local Moran's statistics, and coordinates were mapped in GIS with inverse-distance-weighted surfaces whose power parameter was cross-validated. Ten wells (83 %) fall in the WQI-defined "Good" class (50 <

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

WQI ≤ 100; mean 70.5), and two western-fringe wells are hypersaline and "unsuitable" (WQI = 361 and 399; EC up to 18,590 µS/cm; TDS up to 11,984 mg/L). It must be emphasised at the outset that the "Good" class label is an index result and not a statement of potability: the least mineralised well still contains 1,545 mg/L TDS and every well exceeds the WHO guideline values for TDS, Ca^{2+}, Mg^{2+}, Na^{+}, Cl^{-} and SO_4^{2-}. Index values are globally clustered (Moran's I = 0.240; pseudo-p = 0.0055), although local indicators of spatial association identify no individually significant cluster, so the two hypersaline wells are better described as a spatially coherent pair than as a demonstrated regional domain. A surrogate WQI = 0.02058·EC + 9.12 explains 99.7 % of variance across the full salinity range, but is inflated by two high-leverage points; restricting calibration to the ten low-salinity wells gives R^2 = 0.814 (95 % CI 0.40-0.95), with leave-one-out RMSE = 8.5 and MAPE = 7.5 %. A leakage-controlled refit, in which the TDS term is removed from the index before regression, changes R^2 only from 0.814 to 0.796, indicating that the arithmetic dependence of TDS on EC accounts for approximately 6 % of the fitted slope rather than for the relationship as a whole. Bootstrap inversion (10,000 replicates) yields provisional thresholds of EC ≈ 1,986 (WQI = 50), 4,415 (WQI = 100) and 14,132 µS/cm (WQI = 300); the latter two lie within a 3,780-17,500 µS/cm interval containing no observations and are therefore extrapolations. Applied to an independent published dataset from the same governorate the model under-predicts by 49-61 %; because that dataset is reported as ranges rather than paired observations, this is a bounding comparison and not an external validation, and the model must be regarded as unvalidated beyond the calibration district. When the same values are reclassified under an alternative Iraqi scale, 83 % "Good" becomes 100 % "poor or worse", implying that disagreement with earlier Baiji work is largely a classification-scale artefact. The contribution of this work is therefore methodological and provisional: a transparent diagnostic and cross-validation workflow, and an explicit demonstration of the assumptions under which an EC surrogate can and

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

cannot be used. Seasonal replication, a larger well network and independent gravimetric TDS are required before operational deployment.

Keywords: Groundwater quality; Water Quality Index; GIS; spatial autocorrelation; surrogate modelling; electrical conductivity; screening protocol; SDG 6; Baiji; Iraq.

Highlights

- Twelve drinking wells in Baiji District, Iraq, were indexed with the Weighted Arithmetic WQI and mapped in GIS as an exploratory, single-campaign pilot study (n = 12).
- Global spatial clustering of index values was detected (Moran's I = 0.240; pseudo-p = 0.0055), but no individual well reached local significance (LISA, all p > 0.05).
- An EC-based surrogate model of the WQI was internally cross-validated (LOOCV-RMSE = 8.5 index units; MAPE = 7.5 %); it is not externally validated.
- Provisional class-boundary EC thresholds with bootstrap confidence limits are proposed as a screening aid; the three upper thresholds fall inside an observational gap and are extrapolations.
- A transparent scenario model indicates a 2.0-6.0-fold (50-83 %) reduction in routine full analyses; the estimate is a scenario, not a measured saving.
- Only about 6.6 % of the fitted EC-WQI slope is attributable to the arithmetic dependence of TDS on EC; the relationship survives removal of the TDS term.
- The "Good" WQI class is an index label and not a statement of potability: every well exceeds WHO guideline values for TDS and for four major ions.
- Apparent disagreement with earlier Baiji studies is shown to be largely a classification-scale artefact.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

1. Introduction

In many rural districts of Iraq, especially where surface-water infrastructure is limited or seasonal, groundwater is the source of water for domestic, agricultural and industrial purposes [1-3]. In the semi-arid region of north-central Iraq known as Salah Al-Din Governorate, which is bordered by the Tigris River, hand-dug and drilled wells supply water to villages, schools, security installations and agricultural stations that are too far from centralised treatment plants [4,5]. Of particular concern is Baiji District, located in the north of the governorate, which contains one of Iraq's largest oil-refining complexes and associated industrial and military infrastructure, increasing the risk of anthropogenic contamination in addition to naturally occurring salinity from evaporite-bearing formations [6,7].

1.1. Index-based groundwater appraisal in Iraq: recent evidence

First introduced by Horton [8] and operationalised by Brown et al. [9], the Water Quality Index (WQI) concept reduces a multivariate hydrochemical dataset to a single dimensionless score that can be understood by non-specialist decision-makers. For the past six years the dominant paradigm in Iraqi groundwater assessment has been its integration with geographic information systems, and the accumulated evidence (synthesised in Table 1) is consistent in one respect and strikingly inconsistent in another.

The consistent finding is that shallow Iraqi groundwater is largely marginal for drinking. Alhadithi [10] studied eighty-four samples in western Iraq and found Na-SO_4 and Na-Cl to be the dominant facies, representing 51 % and 31 % of samples respectively. All twenty-seven shallow wells sampled in the Tikrit area of Salah Al-Din were unsuitable for drinking and irrigation, the controlling processes being evaporation and rock-water interaction [4]. Al-Gburi et al. [47] calculated indices for forty-four wells in Missan Governorate ranging from −66.3 to 1,064.1 (mean 184.3) and found about 41 % of samples unsuitable for drinking, attributing the problem at three wells near the Bazergan field to mixing with oil-

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

field brine. Most directly related to the present work is the study by Al-Qaraghuli et al. [48], who evaluated sixteen wells at Tuzkhurmatu in the same governorate and obtained indices ranging from 56.7 to 1,072.7 with electrical conductivity from 620 to 26,250 µS/cm, a bimodal structure similar to that reported here. Hamed et al. [14] at Kaniqirzhala, Erbil, added an important qualification: wells rated poor to good by a drinking index were rated extremely poor by a heavy-metal pollution index, so a favourable major-ion score is not a guarantee of potability.

The inconsistency lies in the rating scales themselves. Studies following the convention of Ramakrishnaiah et al. [49] treat values > 100 as unacceptable and 51-75 as poor [47], whereas studies following the convention of Brown et al. [9] adopted here treat 50-100 as good and only values > 300 as unsuitable. Reported percentages are therefore not comparable across the Iraqi literature unless the underlying scale is explicitly stated - a point taken up quantitatively in Section 4.2, because it accounts for much of the apparent disagreement between the present results and earlier Baiji work.

1.2. Methodological advances: weighting, uncertainty and data-driven prediction

Much recent effort has been directed at removing the subjectivity of expert judgment in assigning weights in a weighted arithmetic index. Entropy weighting derives weights from the information content of the data itself [18,19]; Seifi et al. [20] propagated weight uncertainty through Monte-Carlo simulation and demonstrated that index values carry non-trivial confidence bands that are rarely reported; and Abidi et al. [23] integrated multi-criteria decision making with fuzzy logic for irrigated-area appraisal. Direct prediction of indices using machine learning has also been explored. Raheja et al. [21] compared deep neural networks and gradient boosting for entropy-weighted and traditional indices; Mohseni et al. [55] used ensemble learning for the weighted arithmetic index; and

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Hussein et al. [56] predicted an irrigation index from physicochemical inputs. The demonstration by Seifi et al. [20] that weighting choices propagate into materially different index values is directly relevant to the present study.

These data-driven approaches are powerful but share a practical limitation that is rarely acknowledged. They use many input variables, so the analytical cost of obtaining a prediction is essentially the cost of the full laboratory analysis the index would have required anyway. Their value lies in gap-filling, forecasting and pattern discovery over large archives, not in reducing what a field technician must measure. They also require training sets far larger than the well networks available in most Iraqi districts, and their internal structure is opaque to the water-supply operator who must act on the output. There is therefore a complementary need for a single-input, interpretable screening relationship. This is the space the present study occupies.

1.3. Geospatial integration

Index results are routinely incorporated into GIS [10-14,50], typically using inverse-distance weighting or kriging of the calculated scores. Two limitations recur. First, the choice of interpolation method is often made without justification, although it has a significant effect on the resulting surface, and kriging requires a defensible variogram that small well networks cannot support [37]. Second, and more importantly, the spatial pattern is almost always described visually rather than tested statistically: the literature reports maps showing clusters without establishing that the clustering is distinguishable from a random spatial arrangement. Formal spatial-autocorrelation testing [38,39] is standard practice in spatial epidemiology and econometrics but is rare in this field, with the consequence that a genuine regional salinity domain and a coincidental alignment of two anomalous wells are described in identical language. A less often commented upon limitation is that global statistics such as Moran's I aggregate the entire map into a single number that cannot reveal which locations are

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

responsible for the pattern; for this, local indicators of spatial association [39] are needed and are used here.

1.4. Previous work in Baiji District and the wider governorate

Several studies have focused specifically on groundwater in and around Baiji. Ahmed et al. [6] evaluated thirty wells in the industrial area of the district and found 96.7 % of samples rated poor by a drinking-water quality index, whereas irrigation suitability ranged from medium to high and nitrate-related health-risk quotients were generally below levels of concern. Abed et al. [7] studied thirty-three wells in Baiji province using geo-environmental pollution indices and observed a pollution status from moderate to severe, with a probable contribution from refinery discharges. A related hydrogeochemical study north of Baiji city also reported groundwater largely unsuitable for irrigation without amendment [26]. In the wider governorate, surveys in the Tikrit area [4], Al-Jallam desert [5] and Tuzkhurmatu [48] have consistently shown that major-ion chemistry and salinity in shallow wells are controlled by evaporation and rock-water interaction, not solely by pollution. This conclusion is mirrored for the Shwan sub-basin in neighbouring Kirkuk [52] and for shallow groundwater at Tarmiah in Baghdad Governorate [53].

Considered as a whole (Table 1), three methodological gaps are apparent. None of the reviewed Iraqi studies reports the influence diagnostics or out-of-sample validation needed to establish whether a fitted statistical relationship would survive application to new wells. None tests the significance of the spatial pattern it maps. None converts its statistical findings into an operational field protocol with stated decision thresholds. Identifying a gap does not by itself prove that the current dataset is sufficient to fill it. Section 4.7 explicitly states which of these gaps this pilot study can and cannot fill.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026

This article/work is licensed under CC by 4.0 Attribution

https://eurekaoa.com/index.php/9

Table 1. Synthesis of recent index-based groundwater quality studies in Iraq, with the quantitative outcome of each.

Study	Year	Location	n	Approach	Principal quantitative outcome
Alhadithi [10]	2019	Western Iraq	84	WQI + GIS + Piper	Na-SO_4 (51 %) and Na-Cl (31 %) facies dominant
Al-Ridah et al. [11]	2020	Hillah River	n.r.	WQI	Drinking-quality classification derived for river supply
Fatah et al. [15]	2020	Bardarash, N Iraq	n.r.	WQI + GIS	Class maps produced for the well field
Abed et al. [12]	2021	Karbala (Dammam)	n.r.	WQI + GIS	Spatial quality zoning of the Faddak farm aquifer
Abed et al. [26]	2021	North of Baiji city	n.r.	Hydrogeochemical + irrigation indices	Largely unsuitable for irrigation without amendment
Abed et al. [7]	2022	Baiji Province	33	Geo-environmental pollution indices	Moderate to severe pollution; refinery discharge implicated
Ahmed et al. [6]	2023	Industrial Baiji	30	Drinking WQI + nitrate HQ	96.7 % classed poor; nitrate HQ mostly below concern
Ismail et al. [4]	2023	Tikrit area	27	WQI, SAR, %Na, RSC, MH, Gibbs	All 27 unsuitable for drinking and irrigation; evaporation dominant
Alatar et al. [5]	2023	Al-Jallam Desert	n.r.	Hydrochemistry	Evaporation and rock-water interaction control chemistry
Hamed et al. [14]	2024	Kaniqirzhala, Erbil	n.r.	WQI + heavy-metal pollution index + GIS	WQI poor to good, but HPI > 100 (unfit) at all sites
Al-Gburi et al. [47]	2024	Missan, SE Iraq	44	WQI + factor and cluster analysis	WQI −66.3 to 1,064.1 (mean 184.3); 41 % unsuitable
Manii et al. [13]	2024	Karbala Governorate	n.r.	GIS + hydrochemistry	Quality zoning of Debdebba and Dammam aquifers
Al-Qaraghuli et al. [48]	2025	Tuzkhurmatu, Salah Al-Din	16	WQI (12 parameters)	WQI 56.7-1,072.7; EC 620-26,250 µS/cm; 68.7 % poor or worse
Mohammed et al. [22]	2025	Alton Kopri, Kirkuk	n.r.	WQI + multivariate statistics	Drinkability appraised by combined index and statistics
Al-Sudani [17]	2025	Khanaqin, Diyala	n.r.	Utilisation and quality appraisal	Regional groundwater use and quality characterised
This study	2026	Baiji District	12	WAWQI + GIS + Moran's I and LISA + cross-validated EC surrogate (pilot)	83 % Good on the Brown scale (100 % poor or worse on the Ramakrishnaiah scale); I = 0.240 (p = 0.0055) globally, no significant LISA; LOOCV-RMSE = 8.5; provisional EC thresholds

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Percentage classifications are reported as published and are not directly comparable between rows, because the underlying class boundaries differ between the Brown et al. [9] and Ramakrishnaiah et al. [49] conventions (see Section 4.2). "n.r." = not reported in the sources consulted.

1.5. Research gap, novelty and objectives

One structural limitation common to almost all the literature summarised above is that the index remains costly to obtain. A WAWQI can be computed only from a full major-ion analysis of each well on each sampling occasion. In Iraqi districts with one governorate laboratory serving hundreds of wells, this constraint - rather than any lack of the index itself - is what drives monitoring frequency in practice. It is this constraint that Sustainable Development Goal indicator 6.1.1 (population using safely managed drinking-water services) asks countries to overcome [25]. Surrogate modelling, predicting an expensive composite indicator from one cheaply measured variable, is therefore of direct operational interest, but the conditions under which it is defensible have not been set out for Iraqi groundwater.

This study fills that gap at the level of method rather than of product. Its contribution is threefold. First, it develops an electrical-conductivity surrogate of the WAWQI in an arid Iraqi district, and reports the influence diagnostics, cross-validation error, heteroscedasticity tests and small-sample precision that determine whether such a model could be transferable, explicitly demonstrating that the headline coefficient of determination is leverage-inflated. Second, it quantifies rather than merely acknowledges the arithmetic dependence between EC and an index containing TDS by refitting the model to a leakage-controlled index with the TDS term removed. Third, it transforms the calibrated model into provisional class-boundary thresholds (with confidence limits), and explicitly distinguishes which of these thresholds are interpolations and which are extrapolations beyond the observed data.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

The limits of this design should be stated explicitly. The study is a pilot, with twelve wells sampled at one time; it cannot establish temporal stability, cannot demonstrate transferability, and cannot validate the surrogate in the sense required for operational deployment. Thus the word "validated" is never used of the surrogate in this paper, "internally cross-validated" is used where leave-one-out resampling has been done, and "unvalidated" is used of external application. The specific objectives are to: (i) calculate the WAWQI and seven irrigation-suitability indices for the Baiji well subset and to examine the sensitivity of the resulting classification to the structure of the index; (ii) characterise the spatial structure of groundwater quality via GIS interpolation and global and local spatial-autocorrelation statistics; (iii) calibrate, diagnose and internally cross-validate an EC-based surrogate model of the index under explicit control for arithmetic dependence; (iv) translate the model into provisional screening thresholds and assess the associated analytical-effort implications by means of a transparent and fully stated scenario model; and (v) compare the calibrated model with an independent published dataset from the same governorate as a bounding exercise, and quantify how much of the apparent disagreement with earlier studies is due to classification-scale differences rather than real differences in water quality.

2. Materials and Methods

2.1. Study area

Baiji District is located in the north of Salah Al-Din Governorate in north-central Iraq on the west bank of the Tigris River (approximately 34.70°-35.16° N, 43.04°-43.50° E). The climate is arid to semi-arid with mean annual rainfall well below 200 mm, mainly in winter. Summer maxima frequently exceed 45 °C and potential evapotranspiration far exceeds precipitation. Under such a regime, recharge is limited and episodic, and evapo-concentration together with dissolution of gypsum- and halite-bearing horizons within the Fatha (Lower Fars)

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Formation and the overlying Quaternary alluvium is the first-order control on groundwater chemistry, as documented for adjacent parts of the governorate [4,5]. Land uses include irrigated agriculture along the Tigris corridor, scattered rural settlements, a sand-dune stabilisation scheme, and industrial and military installations, most notably the Baiji refinery complex. Wells of variable depth are completed in the Quaternary alluvial aquifer and in older sedimentary formations. Abstraction is largely unregulated and supplies domestic, livestock and supplementary irrigation demand.

2.2. Data provenance, screening and known limitations of the source dataset

This study is a secondary analysis of an existing operational well-survey archive; no sampling was undertaken by the author for this work. This provenance is stated at the outset because it determines which quality-control information can and cannot be reported, and because the same constraint applies to most district-scale well archives in the region. Hydrochemical and construction data were reviewed for 174 wells surveyed in Salah Al-Din Governorate by the Directorate of Groundwater, Salah Al-Din, and all records within Baiji District were extracted (17 entries). Cross-checking of well names, coordinates and reported values identified five records as exact duplicates of another well in the same set; these were removed, leaving twelve unique wells. The parameters recorded are well depth, static and dynamic water levels, yield, pH, electrical conductivity (EC), total dissolved solids (TDS), Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, Cl^{-}, CO_3^{2-}, HCO_3^{-}, SO_4^{2-} and NO_3^{-}. Analyses were performed using standard methods for the examination of water and wastewater [27]. The source record gives a pH of 7.24 for W7 (Supplementary Table S1, Panel S1a). This value is not flagged as imputed or as an arithmetic mean substitution in the supplied file, and is therefore retained as supplied, with the original laboratory provenance remaining unverified.

The source archive does not contain the survey and laboratory metadata needed for independent verification: sampling campaign date(s); season of collection;

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

per-well depth and completion type; laboratory identity; analytical method and instrument used for each determinand; method detection limits; make/model of EC-meter, calibration standard, and calibration frequency; and sample preservation and transport protocol. Similar status was reported in Supplementary Table S1. The measured chemistry does not give estimates for these unavailable fields and the complete laboratory provenance cannot be determined from the provided records.

2.3. Analytical quality control and the status of the reported TDS values

Two internal consistency checks were performed on the data received. Recalculation from the per-well record of Supplementary Table S1 gives a district-mean ionic charge-balance error (CBE = 100 × (Σcations − Σanions)/(Σcations + Σanions), in meq/L) of +4.95 %, within the conventional ±5 % tolerance. Well-level balances (Supplementary Table S1) range from −4.09 % to +15.02 %; W5 (+8.05 %), W7 (+15.02 %), W9 (+6.27 %) and W12 (+6.84 %) are larger than ±5 %. This is a material analytical quality limitation that is hidden by the acceptable district mean. The calculation is based on the ions registered in the workbook (Ca2+, Mg2+ and Na+ versus Cl-, HCO3-, SO42- and NO3-); K+ and CO32- were not recorded, nor imputed.

The source archive holds no information on field duplicates, laboratory blanks, spiked or certified control samples, replicate measurements or stated analytical uncertainties. This is a material limitation. It means that the analytical uncertainty for each concentration is unknown, that the propagation of that uncertainty into the index cannot be quantified and that quality control for this dataset is based only on ionic charge balance and internal consistency. A sampling campaign must be performed before any operational deployment of the proposed protocol and must be governed by a documented QA/QC plan.

The status of the reported TDS values requires particular care because it bears directly on the interpretation of the surrogate model. The TDS/EC ratio is nearly

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

invariant across the dataset (0.6446-0.6658; mean 0.6495; coefficient of variation 0.84 %). Such a stable ratio over a 7.8-fold change in salinity is difficult to reconcile with independent gravimetric determination and is consistent with TDS being derived from EC using a fixed conversion factor. It must be emphasised, however, that this is an inference from the internal structure of the data, and not a demonstration: the archive does not record the method by which TDS was obtained and a truly uniform ionic composition over a district could in principle produce a similar result. The claim that TDS was derived from EC is therefore stated only as a statement of consistency. The consequence for the surrogate model is addressed directly by the leakage-controlled analysis described in Section 2.7.3 rather than by argument. The single most important requirement of any follow-up campaign is an independent gravimetric determination of TDS.

2.4. Water Quality Index computation

The Weighted Arithmetic Water Quality Index [8,9] was adopted for comparability with the regional literature [10,11,26]. For each parameter i, a quality rating Q_i was computed as

$$Q_i = 100 \times \frac{(V_i - V_i^0)}{(S_i - V_i^0)} \quad (1)$$

where V_i is the measured concentration, S_i the WHO drinking-water guideline value [28], and V_i^0 the ideal value (7.0 for pH, 0 for all other parameters). A unit weight was assigned in inverse proportion to the guideline value,

$$W_i = \frac{k}{S_i}, with\ k = \frac{1}{\Sigma\left(\frac{1}{S_i}\right)} \quad (2)$$

and the overall index obtained as

$$WQI = \frac{\Sigma(W_i \cdot Q_i)}{\Sigma(W_i)} \quad (3)$$

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Nine parameters with defined guideline values were considered (Table 2); the resulting scores were classified according to the conventional five-tier scheme (Table 3). Comparable limits for the parameters used are given in the Iraqi national drinking-water specification IQS 417 [29].

Two properties of this weighting scheme must be explicitly stated, as they govern the interpretation of every classification reported below: First, since weights are inversely proportional to guideline values, the parameter with the smallest numerical guideline value dominates the weight vector. pH, with a guideline of 8.5 against an ideal of 7.0, receives W i = 0.6142, that is 61.4 % of the total weight, while TDS receives 1.04 %. Second, the large pH weight mostly acts as a denominator: at the district-mean pH of 7.209, Q_i = 13.94, and the pH term contributes 7.02 % of the index while absorbing 61.4 % of the weight. The arithmetic effect is that every other parameter's contribution is scaled down by a factor of about 0.386. All calculations are given in full in Supplementary Table S2.

Table 2. Guideline values (S_i), ideal values (V_i^0) and computed unit weights (W_i) used in the Weighted Arithmetic WQI, referenced to WHO drinking-water guidelines [28]. The final column gives the percentage of the total weight carried by each parameter.

Parameter	WHO guideline, S_i	Ideal value, V_i^0	Unit weight, W_i	Share of ΣW_i (%)
pH	8.5	7.0	0.6142	61.43
TDS (mg/L)	500	0	0.0104	1.04
Ca^{2+} (mg/L)	75	0	0.0696	6.96
Mg^{2+} (mg/L)	50	0	0.1044	10.44
Na^{+} (mg/L)	200	0	0.0261	2.61
Cl^{-} (mg/L)	250	0	0.0209	2.09

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Parameter	WHO guideline, S_i	Ideal value, V_i^0	Unit weight, W_i	Share of ΣW_i (%)
SO_4^{2-} (mg/L)	250	0	0.0209	2.09
NO_3^- (mg/L)	45	0	0.1160	11.60
HCO_3^- (mg/L)	300	0	0.0174	1.74
Σ	—	—	0.9999	100.00

Table 3. Classification scheme applied to the Weighted Arithmetic Water Quality Index. The management implications given in the third column describe actions triggered by the index value; they are not, and must not be read as, statements that water in a given class complies with drinking-water guideline values for individual determinands (see Section 3.3).

WQI range	Water quality class	Management implication (index-triggered action)
< 50	Excellent	Suitable for direct supply subject to compliance with individual guideline values
50 - 100	Good	Routine monitoring; individual determinands must still be checked against guideline values
100 - 200	Poor	Treatment recommended
200 - 300	Very poor	Treatment required
> 300	Unsuitable for drinking	Withdraw from supply or desalinate

2.5. Irrigation suitability indices

Most Baiji wells serve both domestic and agricultural demand. Seven established irrigation indices were calculated from concentrations expressed as meq/L: sodium adsorption ratio (SAR) [30], soluble sodium percentage (%Na) [31], residual sodium carbonate (RSC) [30], magnesium hazard (MH) [32], permeability index (PI) [33], Kelley's ratio (KR) [34] and potential salinity (PS)

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)
ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

[35], together with total hardness as $CaCO_3$. The definitions are given in Eqs. (4)-(9):

$$SAR = \frac{Na^+}{\sqrt{\left[\frac{Ca^{2+} + Mg^{2+}}{2}\right]}} \quad (4)$$

$$\%Na = 100 \times \frac{Na^+}{(Ca^{2+} + Mg^{2+} + Na^+ + K^+)} \quad (5)$$

$$RSC = (CO^{32-} + HCO^{3-}) - (Ca^{2+} + Mg^{2+}) \quad (6)$$

$$MH = 100 \times \frac{Mg^{2+}}{(Ca^{2+} + Mg^{2+})} \quad (7)$$

$$PI = 100 \times \frac{\left[Na^+ + \sqrt{HCO^{3-}}\right]}{(Ca^{2+} + Mg^{2+} + Na^+)} \quad (8)$$

$$KR = \frac{Na^+}{(Ca^{2+} + Mg^{2+})}; \quad PS = Cl^- + 0.5 \cdot SO^{42-} \quad (9)$$

2.6. Geospatial analysis

Well coordinates, originally recorded in degree-minute-second format, were converted to decimal degrees. Minor field data-entry inconsistencies (seconds values > 60) were corrected by carrying the excess into the minutes field prior to conversion. The corrected coordinates and computed indices were assembled into a GIS layer in the WGS-84 geographic coordinate system. EC and WQI continuous surfaces were produced by inverse-distance weighting [36], chosen over kriging because twelve control points are insufficient to fit a defensible experimental variogram [37].

However, the interpolation was not without its critics. The power parameter was varied over p = 1, 2, 3 and 4 and each surface was assessed by leave-one-out cross-validation, in which each well in turn is predicted from the remaining eleven, following standard practice for comparing interpolators [59]; results are

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)
ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

reported in Section 3.5. The interpolated surface is presented as a pictorial display of a twelve-point data set, and not as a quantitative estimate of ground-water quality at unsampled points; the class contours on the surface should be interpreted as indicative boundaries and not as mapped isopleths.

The spatial structure of the index was quantified using the global Moran's I statistic [38] with inverse-distance weights, row-standardised and calculated from great-circle separations:

$$I = \left(\frac{n}{\Sigma_i \Sigma_j w_{ij}}\right) \cdot \frac{[\Sigma_i \Sigma_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})]}{\Sigma_i (x_i - \bar{x})^2} \quad (10)$$

Because the normal approximation is unreliable at this sample size [39], the significance of the global Moran's I was evaluated with an upper-tail unconditional permutation test with 9,999 random permutations.

A global statistic alone cannot identify which locations contribute to the observed pattern. Local indicators of spatial association (LISA) [39] were therefore also estimated for each well, using the same row-standardized weight matrix and a conditional permutation null distribution with 9,999 replicates per location:

$$I_i = z_i \cdot \Sigma_j w_{ij} z_j \quad (11)$$

where z_i is the normalised index value at well i. Each well was assigned to a quadrant of the Moran scatterplot (high-high, low-low, high-low, low-high) and a pseudo-significance level. No correction for multiple comparisons was applied. Since the uncorrected result is already non-significant (Section 3.5), a more strict correction would only reinforce the conclusion.

2.7. Surrogate model development, diagnostics and internal validation

2.7.1. Model specification and diagnostics

The calculated WQI was related to EC and to TDS separately by ordinary least-squares regression. Three specifications were compared: a linear model over the entire salinity range, a linear model over the ten low-salinity wells, and a power-

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

law model fitted in $\log_{10}$-$\log_{10}$ space. Model adequacy was evaluated not only by the coefficient of determination, as is usual in this literature and is inadequate, but through a full diagnostic sequence: leverage h_{ii} against the 2p/n criterion, Cook's distance against the 4/n criterion [40], the Shapiro-Wilk test on residuals, a Spearman rank correlation between absolute residuals and EC as a test for heteroscedasticity, a Breusch-Pagan Lagrange-multiplier test [58] as a second and independent test of the same hypothesis, and the Durbin-Watson statistic.

2.7.2. Internal cross-validation, precision and bootstrap inversion

Predictive performance was estimated by leave-one-out cross-validation, each well being predicted from a model fitted to the remaining observations, yielding out-of-sample RMSE, MAE and MAPE. Parameter uncertainty was propagated by non-parametric bootstrap resampling with 10,000 replicates [41]. Finally, the inverted calibrated relationship was used to obtain the EC value corresponding to each WQI class boundary,

$$EC * WQIc = WQIc - \ b0b1 \quad 12(\quad)\frac{(\quad\quad)}{\quad\quad\quad}(\quad)$$

with 95 % confidence limits on EC* taken as the 2.5th and 97.5th percentiles of the bootstrap distribution of $\frac{WQI_c - b^{0}*}{b^{1}}$ *.

Two observations about these procedures are necessary. First, twelve-observation leave-one-out cross-validation is an internal resampling procedure. It estimates how well the model interpolates within the sampled configuration, not performance on independently collected data, and is reported here as an internal, not an external, validation. Second, and for the same reason, bootstrap resampling of twelve observations propagates the sampling variability of the fitted coefficients, but does not create any information that the twelve original observations do not already contain. Therefore, a bootstrap confidence interval based on 10,000 replicates is narrow in proportion to the number of replicates

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)
ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

rather than to the strength of the evidence, and must not be read as an indication of high statistical certainty. Where a bootstrap interval is reported below, the number of independent observations underlying it is given in parentheses (n = 12 or n = 10).

For the Good-class-range model, three additional quantities are reported to make the limited precision of the calibration explicit rather than implicit: the 95 % confidence interval of the slope from the t-distribution; the 95 % confidence interval of the correlation coefficient obtained by Fisher z-transformation and back-transformed to R^2; and the half-width of the 95 % prediction interval for a single new well at representative conductivities. The third is the quantity that determines whether a screening decision at a class boundary is defensible, and is reported in Section 3.6.

The analysis code is provided as Supplementary Material S4, an executable script compatible with Python 3.11 or later and uses NumPy, SciPy and pandas. Bootstrap, permutation procedures use fixed pseudo-random seed 20260904. When the script is run, it records the Python and package versions and recalculates the well-level diagnostics and output tables from the data provided. Resampling pseudo-p values may differ in the last digit reported between NumPy/SciPy versions. Every quantity it produces is printed in the supplementary tables at the end of this paper, so the results can be checked without running the code.

2.7.3. Control for arithmetic dependence between EC and the index

Because TDS enters the index directly (Eq. 3) and is, on the evidence of Section 2.3, not independent of EC, part of any fitted EC-WQI relationship is arithmetic rather than empirical. This is a potential circularity and it is addressed here in two ways rather than merely acknowledged.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

First, the magnitude of the arithmetic component is derived in closed form. The TDS contribution to the index is $(W_TDS/\Sigma W)\cdot(100/S_TDS)\cdot TDS$. Substituting $TDS = r\cdot EC$ with $r = 0.6495$ gives the exact slope of the arithmetic path:

$$dWQI/dEC\,|arithmetic = (W_TDS/\Sigma W)\cdot(100/S_TDS)\cdot r = 0.001356 \text{ index units per } \mu S/cm \quad (13)$$

This is 6.6 % of the fitted full-range slope of 0.02058 and 5.6 % of the Good-class-range slope of 0.02427. The arithmetic dependence therefore accounts for a small minority of the fitted relationship, because TDS carries only 1.04 % of the total weight.

Second, and independently of that derivation, a leakage-controlled index was constructed by removing the TDS term from each computed index value and renormalising the remaining weights:

$$WQI^* = (WQI\cdot\Sigma W - W_TDS\cdot Q_TDS)/(\Sigma W - W_TDS) \quad (14)$$

where $Q_TDS = 100\cdot TDS/S_TDS$. This transformation is exact and requires no information beyond the published per-well TDS and WQI values. The surrogate model was then refitted to WQI*, which by construction contains no term derived from EC. The comparison of the two fits (Section 3.7) provides a direct empirical bound on the contribution of circularity to the reported model performance. It does not, however, remove the deeper dependence that the ions carrying the conductivity are largely the same ions entering the index; that dependence is physicochemical rather than arithmetic, is expected on hydrochemical grounds, and is the reason a surrogate of this kind can work at all, but it also means that the model should be understood as a salinity-tracking relationship and not as a general predictor of water quality.

2.8. External bounding comparison and classification-scale sensitivity

Two analyses were performed to place the results in relation to the previous literature rather than merely juxtaposing them with it. The calibrated surrogate

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

was applied to an independent published dataset from the same governorate, the sixteen-well Tuzkhurmatu survey of Al-Qaraghuli et al. [48].

That study reports conductivity and index values as ranges, not as paired well-level observations. Three consequences follow, which are given here rather than in a footnote. (i) The model for the reported extremities presupposes that the minimum EC and the minimum index are from the same well, and the same for the maximum EC and the maximum index. The published data do not confirm this presupposition and it may be false. (ii) Since only two points are available, and they are assumed rather than observed, no external predictive error, external slope or external R^2 can properly be estimated; the deviations reported in Section 3.10 are bounding figures under a stated assumption, not error statistics. (iii) The Tuzkhurmatu index was calculated from 12 parameters including potassium and total hardness against a partly Iraqi-standard reference, so that the two indices are not the same quantity even in principle. The exercise is therefore reported as an external check on plausibility and its only defensible conclusion is a negative one: it gives no evidence that the calibration transfers and some evidence that it does not.

Second, the twelve Baiji indices were reclassified using the alternative five-tier scheme of Ramakrishnaiah et al. [49], used by several Iraqi studies [47], in order to quantify the extent to which the apparent disagreement between published classification percentages is a scale artefact rather than a difference in water quality.

2.9. Analytical-effort scenario model

A reduction in analytical effort is claimed here only to the extent that an explicit calculation supports it. The scenario model below states every assumption, expresses its output as a countable quantity rather than a currency figure, and is labelled a scenario rather than a measured saving.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

The model is defined as follows. Let N be the number of wells in the monitored network; f_{base} the baseline number of full major-ion analyses per well per year; f_{scr} the number of EC screening measurements per well per year under the proposed protocol; c the number of mandatory confirmatory full analyses per well per year retained irrespective of screening result; and φ the fraction of wells whose screening value exceeds the escalation threshold and which therefore receive a full analysis at every screening occasion. The annual number of full analyses is then $N \cdot f_{base}$ under the baseline and $N[(1 - \varphi)c + \varphi \cdot f_{scr}]$ under the protocol. Cost is expressed parametrically as $\rho = \frac{C_{EC}}{C_{lab}}$, the ratio of the unit cost of a field EC reading to that of a full major-ion analysis, so that no unverified local price data enter the calculation; the relative annual cost of the protocol is then $\frac{[(1-\varphi)c + \varphi \cdot f_{scr} + f_{scr} \cdot \rho]}{f_{base}}$.

Baseline and protocol parameters are set at N = 12, f_{base} = 4 (quarterly), f_{scr} = 4, and φ = 2/12 = 0.167, the exceedance fraction actually observed in this dataset. Three confirmatory schedules are evaluated: c = 1 (annual confirmatory analysis of every well, retained as a safeguard against determinands that EC cannot track), c = 0.5 (biennial), and c = 0 (no mandatory confirmation, presented for completeness but not recommended). Results, together with a sensitivity analysis over φ and a scaling to the 174-well governorate survey frame, are given in Section 3.9. The time horizon is one year, and the model assumes no change in network size or well condition within that horizon.

The framework is deliberately asymmetric, because the surrogate is intended to triage limited laboratory capacity and not to supplant it: wells screening above the Good boundary are escalated to full analysis, whereas wells screening well below it are re-screened at longer intervals. It is this asymmetry, rather than the accuracy of the model, that makes the residual prediction error operationally tolerable - but only for determinands that covary with conductivity, which is the substance of the caveat in Section 4.6.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

3. Results

3.1. General hydrochemical characteristics

Summary statistics are provided in Table 4. Recalculation from the twelve rows of Supplementary Table S1 confirms the means of pH (7.209), EC (5,486.7 µS/cm), TDS (3,550.5 mg/L) and WQI (122.052) from the well-level ionic data. The ionic mean values obtained are Ca^{2+} 361.8, Mg^{2+} 224.3, Na^{+} 516.6, Cl- 790.8, HCO_3^{-} 403.2, SO_4^{2-} 1176.1 and NO_3- 0.558 mg/L. These are computed directly from the per-well record rather than taken from the archive summary. The mean EC for the ten low-salinity wells is 2,975.0 µS/cm with a standard deviation of 504.3 µS/cm and a coefficient of variation of 17.0 %. The mean TDS is 1,934.1 mg/L.

The measured pH is close to neutral (7.14-7.25) and well within the WHO permissible range. Salinity-related parameters, however, vary by almost an order of magnitude: EC ranges from 2,380 to 18,590 µS/cm and TDS from 1,545 to 11,984 mg/L, a spread driven overwhelmingly by two wells (W11, W12) on the western desert fringe. Excluding these, TDS in the other ten wells falls into a much narrower range of 1,545-2,457 mg/L (mean 1,934 mg/L; coefficient of variation of EC 17.0 %, compared with 107.3 % for the entire set).

The least mineralized well in the district still contains 1,545 mg/L TDS, between 1.5 and 3.1 times the WHO esthetic guideline range of 500-1,000 mg/L. The district minima of Ca^{2+}, Mg^{2+}, Na^{+}, Cl^{-} and SO_4^{2-} all exceed their respective guideline values. All of the wells in this data set exceed WHO guideline values for the major-ion suite. This statement is presented here, before any index result, because it is the main and unequivocal water-quality finding of the study, and because it is not overridden by the index classification reported in Section 3.2.

These findings agree with the generally saline character of shallow groundwater reported elsewhere in Salah Al-Din, attributed to evaporite dissolution and evapo-concentration in an arid climate [4,5]. The anion sequence $SO_4^{2-} > Cl^{-} > HCO_3^{-}$ and cation sequence $Na^{+} > Ca^{2+} > Mg^{2+}$ on the district-mean composition indicate

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)
ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

a mixed Ca-Mg-SO_4 to Na-Cl facies typical of gypsum- and halite-influenced aquifers in the region. Nitrate is uniformly low (0.1-1.3 mg/L against a guideline of 45 mg/L). This is an informative negative result, in that it weakens the hypothesis of a pervasive agricultural or sewage source of nitrogen. It does not, however, establish a geogenic origin for the salinity; the reasoning is set out in Section 4.1.

Table 4. Range, mean and WHO guideline values of measured hydrochemical parameters for the twelve Baiji wells (n = 12), with statistics for the ten low-salinity wells shown separately. All displayed means have been recomputed from the per-well record of Supplementary Table S1; corrected ionic means are reproduced at full well-level resolution in Supplementary Table S1.

Parameter	Min	Max	Mean (n = 12)	Mean (n = 10)	WHO guideline [28]	Exceedance at district minimum
pH	7.14	7.25	7.21	7.22	6.5 - 8.5	Compliant
EC (µS/cm)	2,380	18,590	5,487	2,975	—	—
TDS (mg/L)	1,545	11,984	3,551	1,934	500 - 1,000	1.5 - 3.1 ×
Ca^{2+} (mg/L)	169	1,177	361.8	—	75	2.3 ×
Mg^{2+} (mg/L)	63	829	224.3	—	50	1.3 ×
Na^{+} (mg/L)	243	1,768	516.6	—	200	1.2 ×
Cl^{-} (mg/L)	296	2,648	790.8	—	250	1.2 ×
HCO_3^{-} (mg/L)	143	1,316	403.2	—	300	Compliant at minimum
SO_4^{2-} (mg/L)	538	3,913	1,176.1	—	250	2.2 ×
NO_3^{-} (mg/L)	0.1	1.3	0.56	—	45	Compliant
WQI (–)	54.6	398.8	122.1	70.5	—	—

3.2. Water Quality Index and classification

The calculated indices are given in Table 5 in increasing order. Ten of the twelve wells (83 %) fall in the "Good" class, the best score being the Al-Hajjaj sub-

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

district stadium well (54.6) and the worst in this group the sand-dune stabilisation project well (97.0), which is close to the upper class limit owing to high TDS and sulfate. The other two wells (W11, W12) at the 95th Brigade installation are classed "Unsuitable for drinking" (361.2 and 398.8). Their signature - extreme TDS, chloride, sulfate and sodium with unremarkable nitrate and near-neutral pH - is consistent with a hypersaline aquifer horizon rather than a localised anthropogenic point source, an interpretation that their isolation from both the refinery and the populated corridor supports but does not prove.

The 83 % figure should be read with the qualification introduced in Section 2.4 and quantified in Section 3.3. Here, “Good” is the label assigned by the scheme of Brown et al. [9] to an index value in the range 50-100. It is not a finding that the water is drinkable. Due to the exceedances noted in Section 3.1 it should not be reported as such.

3.3. Sensitivity of the classification to the structure of the index

A classification of "Good" is difficult to reconcile with a minimum TDS of 1,545 mg/L and with the exceedance of five major ions, and the applicability of the WAWQI scheme to a water of this composition therefore requires justification. This section provides it by quantifying the mechanism that produces the classification rather than by defending the classification itself.

The index on the district-mean composition is 122.1. The decomposition in Table 6 shows that magnesium contributes 38.37 %, calcium 27.52 % and sulfate 8.05 %, while pH contributes 7.02 % although it carries 61.4 % of the total weight. Its large weight mostly acts in a denominator, as the measured mean pH of 7.209 is close to the ideal 7.0. The complete well-level decomposition is given in Supplementary Table S2.

Two counterfactual computations make explicit the size of this effect. Removing pH and renormalising the remaining weights raises the district-mean index from 122.1 (”Poor”) to 294.2 (”Very poor”). Equal weighting raises it further, to 315.1

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

("Unsuitable"). The classification is therefore not robust to the weighting convention. The corresponding result for each well is given in Supplementary Table S2b.

Two conclusions follow. First, the conventional specification of the WAWQI has poor applicability to waters where salinity is the quality problem, as the determinand carrying most of the weight is the least affected. It is retained here for comparability with the regional literature, which is the only defensible reason for its retention, and the results are reported alongside the sensitivity analysis rather than in place of it. Second, and more generally, index percentages should not be reported without the weight vector and the class scheme that generated them. This is consistent with Seifi et al. [20], who showed that weighting uncertainty propagates into materially different index values and supports the recommendation of Hamed et al. [14] that a favourable major-ion index be treated as a screening result rather than a potability determination.

Table 5. Computed Water Quality Index and classification for the twelve Baiji-District wells, ranked from best to worst. (The source record reports W7 pH = 7.24 and does not marks it as imputed; see Section 2.2.) Per-well leverage, Cook's distance and leave-one-out prediction are reported in Appendix A.

ID	Well location	Lat (°N)	Lon (°E)	pH	EC (µS/cm)	TDS (mg/L)	WQI	Class
W1	Al-Hajjaj sub-district stadium	34.8547	43.4999	7.25	2,380	1,545	54.6	Good
W2	Abu Tu'ma Al-Jadida village	34.8284	43.4893	7.22	2,430	1,580	54.9	Good
W3	Al-Bu Tu'ma monitoring point	34.8069	43.5036	7.20	2,400	1,598	55.1	Good
W4	Baiji central well	34.9200	43.4904	7.22	2,920	1,890	65.6	Good
W5	Al-Mahdiya / Al-Hajjaj village	34.8324	43.4997	7.25	2,820	1,830	71.5	Good

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

ID	Well location	Lat (°N)	Lon (°E)	pH	EC (µS/cm)	TDS (mg/L)	WQI	Class
W6	Makhoul monitoring point	35.1583	43.3144	7.19	3,540	2,300	73.1	Good
W7	Agricultural equipment station, Al-Jazira Al-Wusta	34.9067	43.5014	7.24	2,800	1,813	73.5	Good
W8	Al-Hajjaj village	34.8376	43.5003	7.19	3,430	2,220	77.5	Good
W9	Baiji municipality, Al-Jazira Al-Wusta	34.8993	43.4826	7.24	3,250	2,108	81.8	Good
W10	Sand-dune stabilisation project	34.9461	43.4925	7.21	3,780	2,457	97.0	Good
W11	95th Brigade, 3rd Company	34.7472	43.1000	7.16	17,500	11,281	361.2	Unsuitable
W12	95th Brigade, 1st Regiment HQ	34.8287	43.0402	7.14	18,590	11,984	398.8	Unsuitable

Table 6. Decomposition of the Weighted Arithmetic WQI on the district-mean composition, and sensitivity of the district-mean index to the weighting convention.

Parameter	Quality rating Q_i	Weight W_i	Contribution $W_iQ_i/\Sigma W$	Share of index (%)
pH	13.94	0.6142	8.57	7.02
TDS	710.10	0.0104	7.41	6.07
Ca^{2+}	482.44	0.0696	33.58	27.52
Mg^{2+}	448.50	0.1044	46.83	38.37
Na^{+}	258.29	0.0261	6.74	5.52
Cl^{-}	316.33	0.0209	6.61	5.41
SO_4^{2-}	470.43	0.0209	9.82	8.05

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Parameter	Quality rating Q_i	Weight W_i	Contribution $W_iQ_i/\Sigma W$	Share of index (%)
NO_3^-	1.24	0.1160	0.14	0.12
HCO_3^-	134.39	0.0174	2.34	1.92
District-mean WQI, scheme as specified	—	—	122.1	"Poor"
District-mean WQI, pH term removed, weights renormalised	—	—	294.2	"Very poor"
District-mean WQI, all parameters weighted equally	—	—	315.1	"Unsuitable"

3.4. Irrigation suitability

Table 7 lists recomputed indices based on the district-mean composition and on the two compositional end-members. The average composition is SAR = 5.26, RSC = -29.90 meq/L. The potential salinity is 34.6 meq/L and total hardness is 1,827 mg/L as $CaCO_3$ which indicates severe salinity and very hard water. Magnesium hazard is 50.5 % (just over the 50 % limit) and permeability index is 42.5 % (Doneen Class II). This would mean that continuous irrigation would need leaching fractions, salt tolerant crops and drainage.

Table 7. Irrigation-suitability indices computed on the district-mean composition and on the least- and most-saline compositional end-members. End-member values are indicative, being derived from parameter-wise extremes rather than from a single well analysis; the per-well ionic data required for well-specific computation are provided in Supplementary Table S1.

Index	Least-saline end-member	District mean	Most-saline end-member	Suitability threshold	Assessment
Σ cations (meq/L)	24.19	58.98	203.87	—	—

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Index	Least-saline end-member	District mean	Most-saline end-member	Suitability threshold	Assessment
Charge-balance error (%)	+4.97	+4.95	+6.84	± 5	Acceptable (mean)
SAR	4.05	5.26	9.65	< 10 low hazard	Low sodium hazard
%Na	43.7	38.1	37.7	< 60 permissible	Permissible
RSC (meq/L)	−11.27	−29.90	−105.40	< 1.25 safe	No alkalinity hazard
MH (%)	38.1	50.5	53.7	< 50 suitable	Marginal / unsuitable
PI (%)	50.0	42.5	40.0	> 75 Class I	Doneen Class II
KR	0.776	0.615	0.606	< 1 suitable	Suitable
PS (meq/L)	14.0	34.6	115.4	< 3 unrestricted	Severe salinity hazard
TH (mg/L $CaCO_3$)	681	1,827	6,341	< 300 soft	Very hard

3.5. Spatial structure of groundwater quality

The interpolated surface (Fig. 1) shows two groups of wells. Within the Tigris corridor and agricultural belt there is a coherent low-salinity group, with all ten "Good" wells situated within 25 km of each other and an EC variation of a factor below 1.6. The western desert fringe holds the two hypersaline wells: W11 and W12 lie 38.4 and 40.6 km respectively from the centroid of the low-salinity group. Global Moran's I is 0.240, above the expected value of −0.091 under spatial randomness, with a pseudo-p value of 0.0055 from 9,999 permutations and a permutation z-score of 3.48. Permutation p-values vary in the third decimal place with the random seed; the value labelled on Fig. 1 (0.005) and the value

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

reported here are the same estimate obtained with different seeds. The seed used for all results in this paper is recorded in Table B1. The global pattern is therefore significantly clustered rather than random.

Local indicators of spatial association qualify this result substantially. No well attains local significance at $\alpha = 0.05$: the two hypersaline wells fall in the high-high quadrant of the Moran scatterplot (Fig. 2) with local statistics $I_i = 0.668$ and 0.644, but their conditional-permutation pseudo-p values are 0.077 and 0.429 respectively (Table 8). The ten low-salinity wells all fall in the low-low quadrant, again without individual significance.

The Moran scatterplot does not demonstrate a regional salinity domain as distinct from two isolated anomalies. What the statistics support is weaker and should be stated as such: index values are not randomly arranged in space and the two hypersaline wells are similar to each other and different from the rest, which is consistent with a shared hydrogeological control but is equally consistent with two wells that happen to be near each other in a district sampled at twelve points. A global Moran's I applied to twelve observations in two well-separated groups is largely a test of whether those two groups are different, and is not evidence about the extent, boundaries or continuity of any domain. Further wells in the western sector, a hydrogeological boundary interpretation and locally significant cluster statistics would be needed to determine whether a regional hypersaline domain exists in the western sector. None of this is available here.

Table 8. Local indicators of spatial association (LISA) for the twelve wells, computed with row-standardised inverse-distance weights and 9,999 conditional permutations per location. No well is significant at $\alpha = 0.05$.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Well	Standardised WQI (z)	Spatial lag	Local Moran I_i	Quadrant	Pseudo-p	Significant at 0.05?
W1	−0.580	−0.378	0.219	Low-low	0.143	No
W2	−0.578	−0.400	0.231	Low-low	0.236	No
W3	−0.576	−0.380	0.219	Low-low	0.105	No
W4	−0.486	−0.325	0.158	Low-low	0.267	No
W5	−0.435	−0.434	0.189	Low-low	0.239	No
W6	−0.421	−0.052	0.022	Low-low	0.450	No
W7	−0.418	−0.362	0.151	Low-low	0.274	No
W8	−0.383	−0.450	0.173	Low-low	0.209	No
W9	−0.346	−0.377	0.131	Low-low	0.193	No
W10	−0.216	−0.356	0.077	Low-low	0.188	No
W11	2.058	0.325	0.668	High-high	0.077	No
W12	2.382	0.270	0.644	High-high	0.429	No

The interpolated surface itself was cross-validated (Table 9). Leave-one-out prediction of $\log_{10}$ EC yields a root-mean-square error of 0.145 log units at the power p = 2 used for Fig. 1, corresponding to a predictive R^2 of 0.768. Higher powers do better (p = 4 yields RMSE 0.070 and R^2 0.946) simply because they approach nearest-neighbour assignment, which is the appropriate behaviour for a dataset consisting of two well-separated groups. For consistency with the regional literature [10-14,50], the published figure retains the power p = 2, but this is not the optimum by cross-validation and the fact that the optimum approaches nearest-neighbour assignment is itself the clearest evidence that the surface carries little genuine spatial information between the two groups. Figure 1 should therefore be read as a picture of twelve point observations. The contours it carries are EC values pertaining to the class boundaries, not interpolated isopleths of

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

measured quality. The large unsampled region between the two groups is interpolated across a 38 km gap that has no data.

Table 9. Leave-one-out cross-validation of the inverse-distance-weighted surface as a function of the power parameter, for $\log_{10}$ EC and for the WQI.

Power p	$\log_{10}$ EC: RMSE	$\log_{10}$ EC: MAE	$\log_{10}$ EC: predictive R^2	WQI: RMSE	WQI: MAE	WQI: predictive R^2
1	0.242	0.150	0.352	93.5	50.8	0.353
2 (used for Fig. 1)	0.145	0.105	0.768	54.2	32.5	0.783
3	0.082	0.073	0.926	27.0	18.9	0.946
4	0.070	0.060	0.946	20.8	16.7	0.968

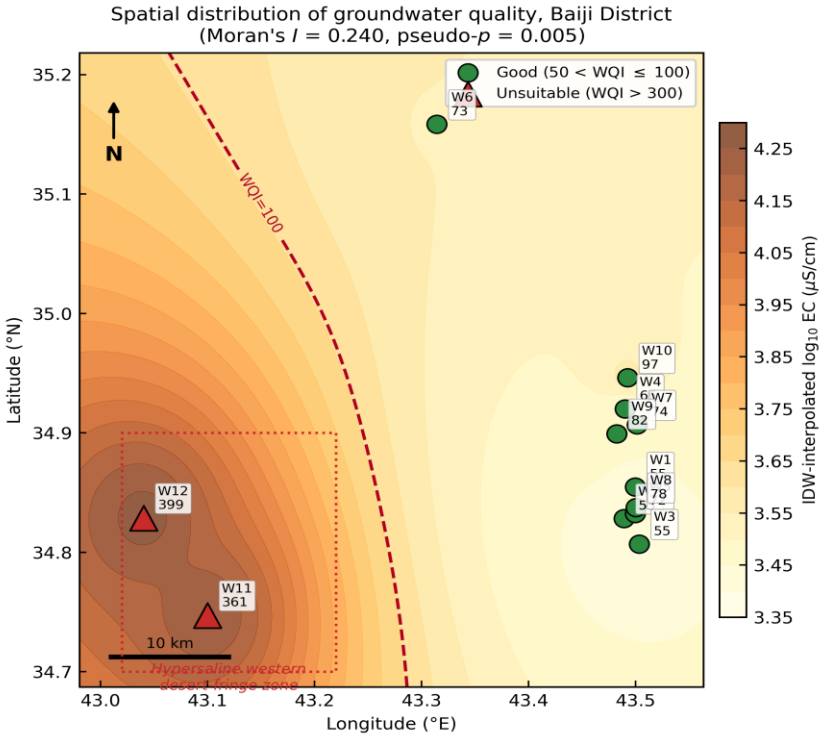


Fig. 1. Spatial distribution of groundwater quality across Baiji District, Salah Al-Din Governorate, Iraq. Background shading is the inverse-distance-weighted

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

$\log_{10}$ EC surface (p = 2; see Table 9 for cross-validation of this choice); solid and dashed contours mark the EC values corresponding to the WQI = 50 and WQI = 100 class boundaries derived in Section 3.8. Labels give well identifier and computed WQI. The dotted box delineates the western desert-fringe wells. The surface is an illustrative interpolation of twelve point observations and is not a quantitative estimate of quality at unsampled locations; the interval between the two groups of wells spans approximately 38 km and contains no data.

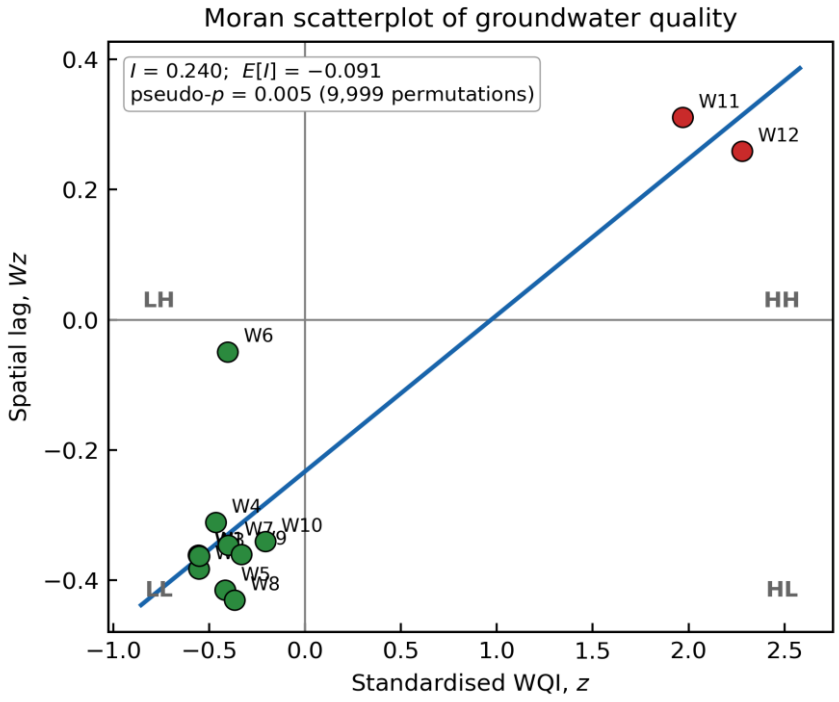


Fig. 2. Moran scatterplot of standardised WQI against its spatial lag under row-standardised inverse-distance weights. Both hypersaline wells fall in the high-high quadrant. Neither is significant as a local cluster (pseudo-p = 0.077 and 0.429; Table 8), and the quadrant position should not be read as evidence of a coherent regional domain.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

3.6. Surrogate model calibration, diagnostics and internal validation

Regression of WQI on EC across all twelve wells yields

$$WQI = 0.02058 \cdot EC + 9.12 \quad (R^2 = 0.9972, p = 4.2 \times 10^{-14}, n = 12 \quad 15)(\quad)$$

(Fig. 3a) with a root-mean-square fit residual of 6.15 index units and a residual standard error of 6.74 index units on 10 degrees of freedom, and bootstrap 95 % confidence intervals from twelve observations of [0.0198, 0.0270] for the slope and [−8.9, 14.7] for the intercept. As expected, the TDS model is equivalent (R^2 = 0.9972) because of the near-constant TDS/EC ratio noted in Section 2.3, and EC is preferred because it is the directly measured field variable.

The headline R^2 of 0.997 cannot be reported without qualification. The influence diagnostics (Fig. 4a; Appendix A) show that W11 and W12 have leverage values of 0.462 and 0.534, well above the 2p/n = 0.333 criterion, and Cook's distances of 1.16 and 1.33, between 3.5 and 4.0 times the 4/n cut-off and both exceeding unity. The fit is therefore determined principally by two data points. When calibration is restricted to the ten low-salinity wells - the range in which a screening tool would actually operate - the relationship is considerably weaker but remains significant:

$$WQI = 0.02427 \cdot EC - 1.75 \quad (R^2 = 0.8145,\ p = 3.5 \times 10^{-4},\ n = 10) \quad (16)$$

This is the appropriate measure of local predictive skill (Fig. 3b), and the difference between Eqs. (15) and (16) is the main methodological result of this paper. Its precision must be reported alongside it. The 95 % confidence interval for the slope in Eq. (16) is 0.0148 to 0.0337, or a range of a factor of 2.3. The R^2 confidence interval at 95 % level, estimated with Fisher z-transformation on ten observations, is 0.40 to 0.95. An R^2 of 0.814 estimated from ten points is consistent with anything from a moderate to a very strong relationship, and the point estimate should not be quoted without this interval.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Leave-one-out cross-validation of Eq. (15) gives RMSE = 8.50 index units, MAE = 7.40 and MAPE = 7.48 %. Leave-one-out cross-validation of Eq. (16) results in RMSE = 7.22, MAE = 6.06 and MAPE = 8.18 %. These are the figures relevant to operational use and are reported in Table 10. The power-law option, $WQI = 0.0377 \cdot EC^{0.941}$, has a slightly smaller cross-validated MAPE of 6.96 % and is preferable when relative error is important; its R^2 is 0.9891 in $\log_{10}$ space and 0.9972 in original units. Both R^2 conventions are reported to avoid mixing scales. Residuals are normally distributed (Shapiro-Wilk W = 0.903, p = 0.174) and show no serial structure (Durbin-Watson = 2.66). The absolute residuals are correlated with EC (Spearman $\rho = 0.664$, $p = 0.018$), indicating heteroscedasticity. Homoscedasticity is not rejected by a Breusch-Pagan Lagrange-multiplier test (LM = 1.41, p = 0.235). The two tests do not agree. Neither has appreciable power with twelve observations. The error variance appears to increase with salinity, but this dataset cannot establish it. The model is thus suitable for indicative screening of classes rather than for reporting of precise index values.

The precision of a single prediction determines whether an indicative class assignment is defensible. Under Eq. (16) the half-width of the 95 % prediction interval for one new well is ±15.6 index units at EC = 2,500 µS/cm, ±15.0 at 3,000, ±15.8 at 3,500 and ±17.8 at 4,000. A single prediction interval therefore spans about 30 index units, roughly 60 % of the 50-unit width of the Good class. The surrogate can distinguish a well near the centre of the Good class from a hypersaline well with confidence, but cannot reliably resolve a well lying within about 15 index units of a class boundary. Class assignment by screening is accordingly described throughout as indicative and subject to confirmation near class boundaries.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

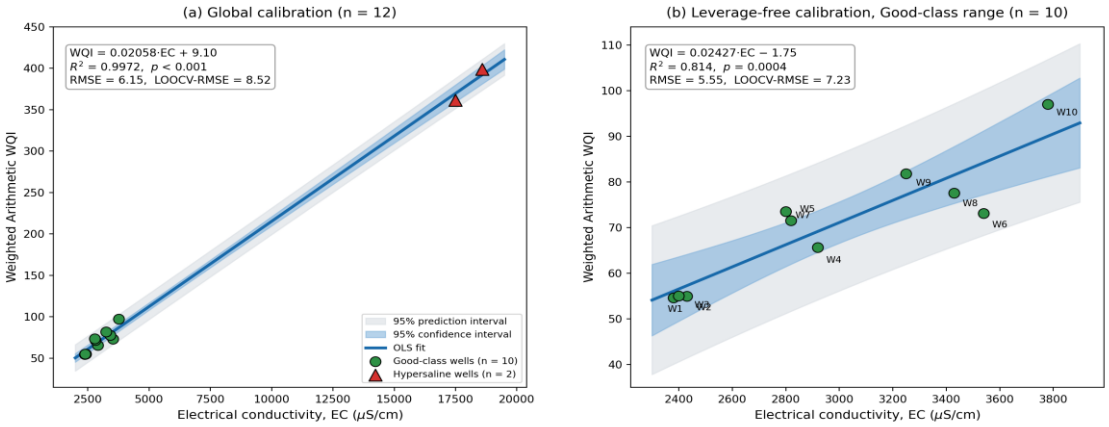


Fig. 3. Surrogate calibration of the Weighted Arithmetic WQI against electrical conductivity: (a) global calibration across the full salinity range (n = 12), with 95 % confidence and prediction bands; (b) leverage-controlled calibration restricted to the ten wells spanning the Good class, which gives the defensible measure of local predictive skill. Note the absence of observations between 3,780 and 17,500 µS/cm in panel (a); the apparent linearity across that interval is an assumption of the model and not an observation.

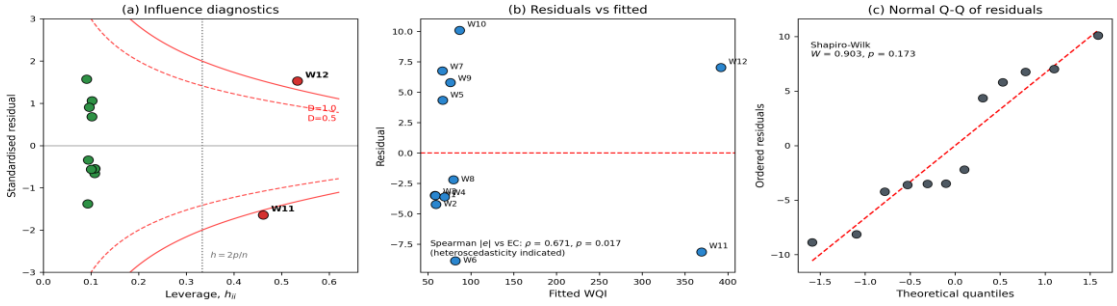


Fig. 4. Regression diagnostics for the global calibration: (a) standardised residual against leverage with Cook's distance contours, showing that W11 and W12 exceed both the 2p/n leverage criterion and D = 1; (b) residuals against fitted

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)
ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

values, indicating heteroscedasticity; (c) normal quantile-quantile plot of residuals.

3.7. Control for arithmetic dependence between EC and the index

Because TDS enters the index directly and is not independent of EC, part of any fitted EC-WQI relationship is arithmetic rather than empirical. This section bounds that component.

The closed-form derivation of Eq. (13) gives the arithmetic component of the slope as 0.001356 index units per µS/cm. Against the fitted full-range slope of 0.02058 this is 6.6 %, and against the Good-class-range slope of 0.02427 it is 5.6 %. The arithmetic path is small because TDS carries only 1.04 % of the total weight in this scheme (Table 2); the leakage is real but structurally bounded.

The empirical refit confirms the derivation. Removing the TDS term from the index by Eq. (14) and refitting gives $WQI^* = 0.01945 \cdot EC + 9.18$ ($R^2 = 0.9968$, $n = 12$) over the full range and $WQI^* = 0.02319 \cdot EC - 1.83$ ($R^2 = 0.7967$, $n = 10$) over the Good class, with a full-range leave-one-out RMSE of 8.59 index units. The two fits are set out side by side in Table 11. Removing the arithmetically dependent term changes R^2 from 0.9972 to 0.9968 across the full range and from 0.8145 to 0.7967 within the Good class, and changes the slopes by 5.5 % and 4.5 %, respectively. The relationship therefore survives removal of the circular term almost intact, and the high coefficient of determination is not an artefact of data leakage.

Two qualifications remain, and neither is removed by this analysis. First, a bounded arithmetic contribution of about 6 % is not zero, and the R^2 values reported for models containing the TDS term should be read as marginally optimistic. Second, and more fundamentally, the ions that carry the electrical conductivity are largely the same ions that enter the index: calcium, magnesium, sodium, chloride and sulfate together account for 84.9 % of the district-mean index value (Table 6). That dependence is physicochemical rather than arithmetic;

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

it is expected on hydrochemical grounds and is precisely why a conductivity surrogate can work in a salinity-dominated system. It also defines the limits of what the surrogate is: a tracker of ionic strength which, in this district, is what the index measures. The relationship would not hold where quality is governed by a determinand that does not covary with conductivity, and no result in this paper bears on that case.

Table 10. Comparison of candidate surrogate models and their validation statistics. LOOCV denotes leave-one-out cross-validation. Bold indicates the specification recommended for screening use within the Good class.

Model	Specification	R^2	RMSE (fit)	LOOCV-RMSE	LOOCV-MAPE (%)	Remarks
Linear, full range	WQI = 0.02058·EC + 9.12	0.9972	6.15	8.50	7.48	R^2 leverage-inflated (Cook's D > 1 for W11, W12)
Linear, Good-class range	WQI = 0.02427·EC − 1.75	0.8145 (95 % CI 0.40-0.95)	5.54	7.22	8.18	Recommended for screening (n = 10); slope 95 % CI 0.0148-0.0337
Linear, TDS predictor	WQI = 0.03198·TDS + 8.50	0.9972	6.20	8.54	7.61	Redundant with EC (TDS/EC ≈ 0.6495)
Power law	WQI = $0.0377 \cdot EC^{0.941}$	0.9891 (log space); 0.9972 (original units)	6.16	8.58	6.96	Lowest relative error; handles heteroscedasticity
Linear, full range, TDS term removed from index	WQI* = 0.01945·EC + 9.18	0.9968	6.21	8.59	7.95	Leakage-controlled; see Section 3.7
Linear, Good-class range, TDS term removed	WQI* = 0.02319·EC − 1.83	0.7967	5.60	—	—	Leakage-controlled; R^2 falls by 0.018 only

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Table 11. Effect of removing the arithmetically dependent TDS term from the index before regression.

Quantity	Index as specified	Index with TDS term removed	Change
Slope, full range (index units per µS/cm)	0.02058	0.01945	−5.5 %
R^2, full range (n = 12)	0.9972	0.9968	−0.0004
Slope, Good-class range	0.02427	0.02319	−4.5 %
R^2, Good-class range (n = 10)	0.8145	0.7967	−0.018
LOOCV-RMSE, full range (index units)	8.50	8.59	+1.1 %
Arithmetic component of slope, Eq. (13)	0.001356	0 by construction	6.6 % of fitted slope

3.8. Provisional field screening thresholds

EC values for each class boundary were obtained by inverting Eq. (15) with bootstrap confidence limits (Table 12; Fig. 5). Thresholds for the two lower boundaries were also derived from the Good-class-range model of Eq. (16). The operationally relevant full-range thresholds are EC ≈ 1,986 µS/cm (95 % CI 1,778-2,209) for WQI = 50 and EC ≈ 4,415 µS/cm (95 % CI 4,031-4,604) for WQI = 100. From Eq. (16), the corresponding thresholds are 2,132 µS/cm (95 % CI 1,566-2,274) and 4,192 µS/cm (95 % CI 3,832-5,116). The wider intervals from the ten-point calibration more honestly reflect local uncertainty and are preferred for screening.

A limitation of these thresholds must be stated plainly. The dataset contains no observation between 3,780 and 17,500 µS/cm, an interval spanning a factor of 4.6. The three upper full-range thresholds - 4,415, 9,274 and 14,132 µS/cm - lie inside that interval and are extrapolations across an observational gap. Their bootstrap intervals describe the sampling variability of a straight line whose

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

behaviour in the gap is assumed; they do not capture model-form extrapolation risk. Only the WQI = 50 threshold lies within the low-salinity observations. The thresholds are therefore labelled provisional, and the recommendation to sample additional wells in the 4,000-15,000 µS/cm range follows directly from this limitation.

Applying the protocol post hoc to the twelve wells recovers the laboratory-derived class in every case, including both hypersaline wells. This is a resubstitution on the data used for fitting and is not evidence of predictive accuracy; it is reported for completeness only. The resulting five-step workflow is summarised in Fig. 6: field EC measurement, surrogate prediction, threshold decision, confirmatory analysis of flagged wells, and GIS-layer update.

Table 12. Provisional field screening thresholds obtained by inverting Eqs. (15) and (16), with 95 % confidence limits from 10,000 bootstrap replicates of the twelve (or ten) original observations, and the corresponding management action.

WQI class boundary	Threshold EC (µS/cm), Eq. (15)	Bootstrap 95 % CI	Threshold EC, Eq. (16)	Bootstrap 95 % CI	Interpolation or extrapolation?	Field action
50	1,986	1,778 - 2,209	2,132	1,566 - 2,274	Within observed range	Re-screen annually
100	4,415	4,031 - 4,604	4,192	3,832 - 5,116	Extrapolated: no data 3,780-17,500	Re-screen semi-annually; escalate if exceeded
200	9,274	7,733 - 9,542	Not derivable	—	Extrapolated	Full laboratory analysis; treatment appraisal
300	14,132	11,430 - 14,490	Not derivable	—	Extrapolated	Full analysis; restrict drinking use
> 300	> 14,132	—	—	—	Extrapolated	Withdraw from supply; desalinate or replace

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

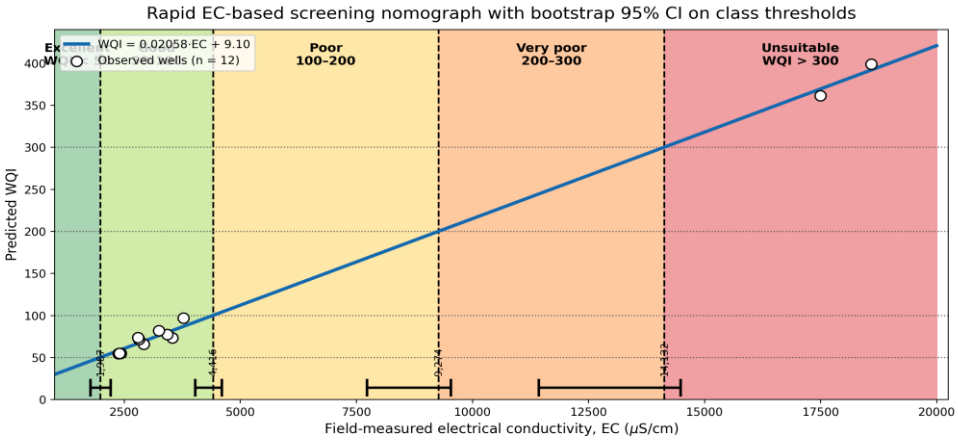
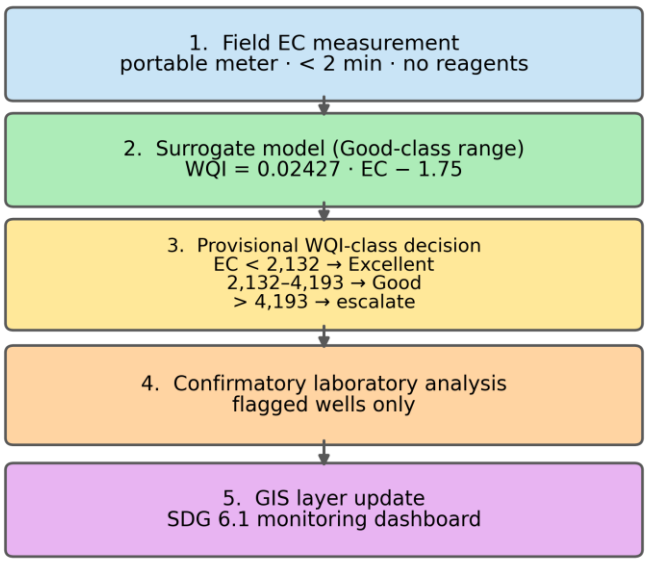


Fig. 5. Rapid screening nomograph converting a single field EC measurement into a predicted WQI class. Shaded bands correspond to the five quality classes; vertical dashed lines are the inverted class-boundary thresholds and horizontal bars the bootstrap 95 % confidence limits on each threshold. The three upper thresholds fall within an interval containing no observations (Section 3.8) and should be read as provisional extrapolations.



Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Fig. 6. Proposed five-step low-cost screening protocol linking a single field conductivity measurement to GIS-based reporting. Confirmatory laboratory analysis is triggered for wells screening above the "Good" boundary and, under the recommended schedule, once annually for every well irrespective of screening result (Section 3.9).

3.9. Analytical-effort scenario

The scenario model defined in Section 2.9 is evaluated for the observed Baiji network in Table 13.

Under a quarterly baseline, twelve wells require 48 full major-ion analyses per year. Under the recommended schedule, quarterly EC screening with one mandatory confirmatory analysis per well per year and escalation of the 16.7 % of wells that exceed the threshold, the requirement falls to 18 analyses per year, a reduction of 62.5 % or a factor of 2.7. Relaxing the confirmatory analysis to biennial gives 13 analyses (72.9 %, factor 3.7), and removing it altogether gives 8 (83.3 %, factor 6.0). The reduction is thus a factor of between two and six across the plausible range of schedules, and it is not an order of magnitude.

The result is sensitive to the exceedance fraction φ, which is a property of the district and not of the method: at $\varphi = 0$ the reduction is a factor of 4.0, at the observed $\varphi = 0.167$ it is 2.7, and at $\varphi = 0.5$ it falls to 1.6. Scaled to the 174-well governorate survey frame at the observed exceedance fraction, the annual requirement falls from 696 to 261 full analyses. In parametric cost terms, with $\rho = C_EC/C_lab$ the ratio of unit costs, the relative annual cost of the protocol is 0.385 at $\rho = 0.01$ and 0.475 at $\rho = 0.10$, so the conclusion is insensitive to the cost ratio over any plausible range - the saving is driven by the number of analyses avoided, not by the cost of the screening measurement.

None of the three qualifications which apply to this whole calculation can be removed by the present data. It is a scenario calculated from stated assumptions, not an operational saving that has been measured. It assumes the exceedance

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

fraction of a single campaign to be time-stable, which could not be tested in this study. And it counts analyses rather than the true economic or environmental cost, which would also depend on transport distances, batch sizes and laboratory use that the source archive does not record.

Table 13. Analytical-effort scenario for the Baiji network (N = 12 wells, quarterly baseline, observed exceedance fraction φ = 0.167). All figures are full major-ion analyses per year.

Schedule	Confirmatory analyses per well per year (c)	Full analyses per year	Reduction vs baseline	Factor
Baseline: quarterly full analysis, no screening	4	48	—	1.0
S1 (recommended): quarterly EC screening + annual confirmatory	1	18	62.5 %	2.7
S2: quarterly EC screening + biennial confirmatory	0.5	13	72.9 %	3.7
S3: quarterly EC screening, no mandatory confirmatory (not recommended)	0	8	83.3 %	6.0
Sensitivity, S1 at φ = 0	1	12	75.0 %	4.0
Sensitivity, S1 at φ = 0.333	1	24	50.0 %	2.0
Sensitivity, S1 at φ = 0.5	1	30	37.5 %	1.6
S1 scaled to 174-well survey frame	1	261 (baseline 696)	62.5 %	2.7

3.10. External bounding comparison and classification-scale sensitivity

When the calibrated model is applied to the independent Tuzkhurmatu dataset [48] (Table 14, Panel A), predictions are 21.9 and 549.4 at the reported conductivity extremes of 620 and 26,250 µS/cm, against reported values of 56.7

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

and 1,072.7. The model under-predicts by 61.4 % and 48.8 % respectively, and the slope implied by the external data (0.0396 index units per µS/cm) is 1.93 times that fitted at Baiji. Class assignment is consistent at the high-conductivity end, the most conductive sample being assigned to the unsuitable class in both datasets.

These are bounding figures, as explained in Section 2.8, computed under an unverified assumption concerning the pairing of reported extremes, between two indices not the same quantity. They should not be interpreted as external prediction errors and no external validation is claimed Their evidential value is entirely negative: the exercise does not support transferability of the fitted coefficients, and a discrepancy of this magnitude is what one would expect if the coefficients are site-specific. Hence the model is called unvalidated outside the calibration district.

Reclassifying the twelve Baiji indices with the alternative five-tier scheme of Ramakrishnaiah et al. [49] (Table 14, Panel B) changes the distribution from 83.3 % good and 16.7 % unsuitable to 58.3 % poor, 25.0 % very poor and 16.7 % unacceptable, with no sample in the good or excellent classes. The index values are identical; only the class boundaries differ.

Table 14. External bounding comparison of the calibrated surrogate and sensitivity of the reported classification to the choice of class boundaries. Panel A applies Eq. (15) to the independent Tuzkhurmatu dataset of Al-Qaraghuli et al. [48]; because that study reports ranges rather than paired values, the comparison assumes that the extremes of the two reported ranges belong to the same samples and is not an external validation. Panel B reclassifies the identical twelve Baiji index values under the alternative scheme of Ramakrishnaiah et al. [49].

Panel A — external application	Baiji (calibration, this study)	Tuzkhurmatu [48] (external)	Model prediction	Deviation
EC range (µS/cm)	2,380 - 18,590	620 - 26,250	—	—

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Panel A — external application	Baiji (calibration, this study)	Tuzkhurmatu [48] (external)	Model prediction	Deviation
Reported WQI range	54.6 - 398.8	56.7 - 1,072.7	—	—
WQI at lowest EC	54.6 (at 2,380)	56.7 (at 620)	21.9	−61.4 %
WQI at highest EC	398.8 (at 18,590)	1,072.7 (at 26,250)	549.4	−48.8 %
Class at highest EC	Unsuitable	Unsuitable	Unsuitable	Agreement
Implied slope (WQI per µS/cm)	0.0206	0.0396	—	× 1.93
Status of the comparison	—	—	—	Bounding only; not a validation

Panel B — classification-scale sensitivity	Brown et al. [9] scheme (used here)	Ramakrishnaiah et al. [49] scheme
Excellent	0 % (0/12)	0 % (0/12)
Good	83.3 % (10/12)	0 % (0/12)
Poor	0 % (0/12)	58.3 % (7/12)
Very poor	0 % (0/12)	25.0 % (3/12)
Unsuitable / unacceptable	16.7 % (2/12)	16.7 % (2/12)
Poor or worse (aggregate)	16.7 %	100 %

4. Discussion

4.1. Controls on groundwater salinity

Near-neutral pH, very low nitrate, a sulfate-dominated anion sequence and extreme spatial contrast in TDS are consistent with a predominantly geogenic control on salinity. Sulfate concentrations of approximately 3,900 mg/L in the western wells are consistent with gypsum dissolution, with chloride and sodium

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

accompanying halite dissolution and evapo-concentration, mechanisms already established as dominant in the Tikrit area [4] and the Al-Jallam desert [5] and characteristic of inland arid basins generally [42,43].

The lack of a nitrate signal weakens the hypothesis of widespread agricultural or sewage contamination, since such sources usually leave one. It does not, however, establish a geogenic origin. Low nitrate is perfectly consistent with the presence of petroleum hydrocarbons, BTEX compounds, polycyclic aromatic hydrocarbons, phenols, heavy metals, or other determinands not detected by a major-ion suite in a district containing a major oil-refining complex. None of these were measured. Previous work in the same district has implicated refinery discharges, using geo-environmental pollution indices [7]. Nothing in the present data set can refute that finding. Thus, the defensible statement is more limited than the earlier one: the salinity at these twelve sites is consistent with evaporite dissolution and evapo-concentration, and the dataset contains no evidence of an anthropogenic nitrogen source, but the source of any non-major-ion contamination in this district is beyond the scope of these data.

4.2. Comparison with previous studies and the classification-scale problem

The classification obtained here is at face value considerably more favourable than those reported nearby (Table 15). Ahmed et al. [6] reported 96.7 % of samples from thirty wells in the industrial zone of Baiji as poor; Abed et al. [7] reported moderate to severe pollution in thirty-three wells; Ismail et al. [4] reported all twenty-seven wells in the Tikrit area unsuitable for both drinking and irrigation; Al-Qaraghuli et al. [48] reported 68.7 % of Tuzkhurmatu samples poor or worse; and Al-Gburi et al. [47] reported 41 % of Missan samples unsuitable. In this context a result of 83 % "Good" requires explanation rather than celebration.

Much of the divergence is not a difference in water at all. Under the scheme of Ramakrishnaiah et al. [49], in which 51-75 is poor, 76-100 very poor and above

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

100 unacceptable, and which several of the Iraqi studies cited above employ [47], reclassification of the same twelve indices gives 58.3 % poor, 25.0 % very poor and 16.7 % unacceptable, with none in the good or excellent classes (Table 14, Panel B). Under that convention the present dataset would be reported as 100 % poor or worse, which is not meaningfully different from the 96.7 % of Ahmed et al. [6]. The two studies therefore agree about the water and differ in vocabulary. This is not a peripheral observation: percentage classifications are the headline result of most index papers and the form in which they reach policy audiences, yet they are scale-dependent to the point of reversal, and the scale is frequently left implicit. Index percentages should not be compared across studies unless the underlying scale is reported, and index-based monitoring programmes should adopt a single convention before any inter-district comparison is attempted.

The sensitivity analysis of Section 3.3 provides a second dimension to this argument. The same twelve records give a district-mean index of 122.1 under the stated inverse-guideline weights, 294.2 after removing pH and 315.1 for equal weights. The district-mean index range covers the Poor, Very poor and Unsuitable classes without altering the measurement. Therefore, the index percentages are uninterpretable unless the class scheme and weight vector are reported.

There are, however, real differences. Sampling location matters: the present wells are mostly in residential, agricultural and rural service areas rather than in the industrial belt immediately surrounding the refinery, and the extreme salinity is confined to a desert-fringe location distant from both. Index construction matters: this study references the WAWQI to WHO guideline values, whereas Abed et al. [7] used pollution indices referenced to geochemical background, which are more sensitive to industrial signatures by construction. Hamed et al. [14] showed that a well may pass a major-ion index and fail a heavy-metal index at the same site. Sample size and selection matter: twelve wells from an operational survey are not a random sample of the district. What survives all these caveats, and is therefore the robust cross-study conclusion, is that a substantial fraction of groundwater in

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

the Baiji area cannot be consumed without treatment, and that salinity is a dominant control at the sites sampled.

Table 15. Detailed comparison of the present study with previous groundwater investigations in Baiji District and the wider Salah Al-Din Governorate. The final column reports each study's own classification; see Section 4.2 for why these percentages are not directly comparable.

Study	Wells (n)	Setting	Index applied	Principal outcome
Ahmed et al. [6]	30	Industrial zone, Baiji	Drinking WQI; nitrate health-risk quotient	96.7 % classed poor; nitrate risk generally below concern
Abed et al. [7]	33	Baiji Province	Geo-environmental pollution indices	Moderate to severe pollution; refinery discharge implicated
Abed et al. [26]	—	North of Baiji city	Hydrogeochemical and irrigation indices	Largely unsuitable for irrigation without amendment
Ismail et al. [4]	27	Tikrit area	WQI, SAR, %Na, RSC, MH, Gibbs	All samples unsuitable; evaporation and rock-water interaction dominant
Alatar et al. [5]	—	Al-Jallam Desert	Hydrochemical characterisation	Evaporation and mineral dissolution control salinity
Al-Qaraghuli et al. [48]	16	Tuzkhurmatu	WQI (12 parameters, WHO + IQS)	WQI 56.7-1,072.7; 68.7 % poor or worse; sewage inputs implicated
Al-Gburi et al. [47]	44	Missan (comparison basin)	WQI + factor and cluster analysis	41 % unsuitable; oil-field brine mixing at three wells
This study	12	Rural / peri-urban Baiji	WAWQI + GIS + Moran's I and LISA + internally cross-validated EC surrogate	83 % Good on the Brown scale (100 % poor or worse on the Ramakrishnaiah scale; 0 % Good if pH is excluded from the weight vector); provisional screening protocol derived

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)
ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

4.3. What the classification does and does not establish

New: This section is a revision. It comes before the discussion of the surrogate since it governs the interpretation of everything that follows.

The data are consistent with three statements about the twelve wells. Major-ion chemistry is above WHO guideline values for TDS at all wells and for calcium, magnesium, sodium, chloride and sulfate at the district minimum. The WAWQI calculated with inverse-guideline weights and classified according to the Brown et al. [9] scheme puts ten wells in the Good class. These statements are not inconsistent, as the index is a weighted aggregate and the class boundaries of this aggregate were never intended to encode guideline compliance for individual determinands.

The statement that the ten Good-class wells are potable is not supported. The guideline comparison provides no evidence for it, and no well in this district should be described as suitable for direct supply to drinking water on the basis of an index class alone. The management implications in Table 3 are framed accordingly. Where an index class conflicts with a direct comparison against guideline values, the guideline comparison governs. The index is a triage and communication tool, not a compliance determination. This is the same caution Hamed et al. [14] urged for heavy metals and is equally applicable for major ions when the index is weighted as it is here.

4.4. Spatial interpretation

The spatial claims of this study are limited to what the statistics support. Global Moran's I of 0.240 with a permutation pseudo-p of 0.0055 establishes that index values are not randomly arranged across the twelve sampled locations. It does not establish the existence, extent or boundaries of a hydrogeologically coherent salinity domain, and the local statistics reported in Table 8 actively fail to support such a claim: neither hypersaline well reaches local significance, and the more extreme of the two has a pseudo-p of 0.429.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)
ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

There is also a structural point about what a global statistic can signify at this sample size and configuration. The dataset comprises one tight group of ten wells and two further wells some 38 km away, so Moran's I is largely a formal restatement of the fact that the two groups differ, which is evident from Table 5 without any spatial statistic. In this configuration a significant global I is close to guaranteed and carries correspondingly little information. The statistic is retained because testing for spatial pattern is better practice than describing it visually, which was the methodological point of Section 1.3, but no hydrogeological conclusion is drawn from it. The precautionary inference for management is the same in substance but different in status: because two of two wells sampled in the western sector are hypersaline, prudence suggests treating new wells in that sector as potentially unsuitable until tested. It is a precautionary recommendation on the basis of a sample of two, not a finding about a demonstrated domain. It is framed as such in Section 5.

4.5. External application and the absence of demonstrated transferability

A surrogate calibrated in one district is of limited use unless its coefficients can be reused, and the comparison with the Tuzkhurmatu dataset [48] was undertaken to examine that question. Applying Eq. (15) at the reported conductivity extrema gives indices of 21.9 and 549.4 against published values of 56.7 and 1,072.7, an under-prediction of 61.4 % and 48.8 %, with an implied external slope of 0.0396 index units per µS/cm, 1.93 times the Baiji slope. As set out in Sections 2.8 and 3.10, this comparison rests on an unverified pairing assumption and compares two indices that are not the same quantity; it is a bounding exercise and not a validation.

The conclusion is nonetheless clear, and it is a negative one. There is no evidence that the numerical coefficients of Eq. (15) transfer to other districts, and the discrepancy is of the magnitude that would be expected if they do not. What may transfer is the method - the model structure, the diagnostic sequence, the leakage

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)
ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

control, the inversion to class thresholds and the decision protocol - and even that is a proposal rather than a demonstration, since it has been exercised in one district only. Any application elsewhere requires local recalibration on at least ten wells spanning the local conductivity range and the same leverage, cross-validation and leakage checks applied here.

That class assignment agreed at the high conductivity end. This is consistent with the design choice to use the surrogate for classification rather than reporting index values. Agreement at one point between two datasets is weak evidence, and should not be presented as robustness.

4.6. Analytical-effort implications

A full major-ion analysis requires sample transport, reagents, instrument time and analyst effort, and in the Salah Al-Din context a turnaround of days to weeks; a field EC reading requires a calibrated portable meter, no reagents and under two minutes. The protocol would require full analysis only for wells screening above the Good boundary, for wells showing material change since the last screening, and, under the recommended schedule, once annually for every well as a safeguard. On the scenario model of Section 3.9 this reduces routine full analyses by a factor of 2.7 under the recommended schedule and by between 2.0 and 6.0 across the plausible range of schedules and exceedance fractions, which is not an order of magnitude. The binding constraint on safely-managed-service reporting in such districts is monitoring frequency rather than analytical precision [25], and the protocol acts on that constraint; there is also a narrower environmental benefit in reduced reagent consumption and sample-transport emissions.

Two caveats attach to the claim. The protocol is calibrated to a salinity-dominated system and would fail where quality is controlled by a parameter uncorrelated with EC - nitrate, fluoride, arsenic, hydrocarbons or microbial contamination - so it supplements rather than replaces periodic full characterisation. The estimated reduction is a scenario calculated from stated assumptions, not a measured

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

operational saving. The distinction is explicitly maintained in Section 3.9 and in Table 13. A third caveat concerns scope: the SDG 6.1 framing is a statement of possible relevance, not a demonstrated contribution. Indicator 6.1.1 is focused on change in safely managed services over time, but this study reports a single campaign at twelve wells, with no temporal replication. The protocol is a candidate monitoring instrument not a monitoring instrument until the stability of the EC-WQI relationship has been verified over at least two contrasting seasons.

4.7. Statistical power and the limits of an n = 12 design

Twelve observations, reduced to ten once the two high-leverage wells are set aside, are a weak basis for the inferential apparatus deployed in this paper, and the consequences are stated here quantitatively rather than in general terms.

Twelve observations, reduced to ten once the two high-leverage wells are set aside, are a weak basis for the inferential apparatus deployed in this paper, and the consequences should be stated quantitatively rather than acknowledged in general terms. The 95 % confidence interval of R^2 for the recommended model spans 0.40 to 0.95. The 95 % confidence interval of its slope spans a factor of 2.3. The half-width of a 95 % prediction interval for a single new well is approximately 15 index units, about 60 % of the width of the class the model is intended to assign. Two statistical tests of the same heteroscedasticity hypothesis disagree with each other. And no well attains local spatial significance despite a strongly significant global statistic.

None of this is remedied by the analytical machinery applied to the data. Neither bootstrap resampling nor leave-one-out cross-validation repairs it. Both are internal resampling procedures: they redistribute the information in twelve observations and quantify how sensitive the estimates are to that particular sample, but they generate no independent information, and a confidence interval computed from 10,000 replicates of twelve observations is narrow in proportion

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

to the replicate count and not to the evidence [60]. The number of independent observations is therefore reported alongside every resampled interval.

What follows is a boundary on the claims this design can support. It can demonstrate that a strong EC-WQI association exists in this district and is not an artefact of arithmetic circularity. It can demonstrate that a headline R^2 of 0.997 is leverage-inflated and that a diagnostic sequence is necessary to reveal this - a methodological finding that does not depend on sample size, since it is a demonstration about method rather than an estimate of a population parameter. It cannot establish predictive accuracy adequate for operational deployment, temporal stability, transferability, or the existence of a regional salinity domain. The study is presented on that basis.

4.8. Limitations and uncertainty

The principal limitations of this study are the following.

- (i) Sample size and design. Twelve wells from a single survey campaign, ten after exclusion of high-leverage points. Confidence intervals are correspondingly wide (Section 4.7) and no temporal dimension exists, so seasonal variability in recharge and abstraction is entirely unrepresented.
- (ii) Observational gap. No well lies between 3,780 and 17,500 µS/cm. All three upper class thresholds fall in this interval and are extrapolations whose bootstrap confidence limits do not capture extrapolation risk.
- (iii) Leverage dependence. As shown in Section 3.6, $R^2 = 0.997$ is a property of the sampling configuration rather than of the underlying physics; Eq. (16) with its stated confidence interval is the appropriate figure for judging local predictive skill.
- (iv) Data provenance and TDS status. The near-invariant TDS/EC ratio is consistent with derivation of TDS from EC by a fixed factor, but the archive does not record the method and this remains an inference. The arithmetic component of the fitted relationship is bounded at approximately 6.6 % (Section 3.7), but

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

independent gravimetric TDS determination remains the highest-priority requirement of any follow-up campaign.

- (v) Quality-control information. The source record permits a well-level ionic-balance check: four of twelve wells exceed ±5 % CBE, including W7 at +15.02 %. Potassium and carbonate, field duplicates, blanks, control samples, replicate measurements, preservation and transport protocols, detection limits and analytical uncertainties are not recorded. Analytical uncertainty therefore cannot be propagated into the index, and the positive balances cannot be resolved diagnostically.
- (vi) Index structure. The classification is not robust to the weighting convention: the district-mean index moves from 122.1 to 294.2 or 315.1 under defensible alternatives (Section 3.3; Supplementary Table S2b). Reported percentages remain conditional on both the weight vector and the class scheme.
- (vii) Spatial inference. Global spatial clustering is significant but no location attains local significance, and the interpolated surface has a predictive R^2 of 0.768 at the power used for display and spans a 38 km data gap.
- (viii) Analytical scope. The dataset contains no trace metals, hydrocarbons or microbiological indicators, a material gap in a refinery-adjacent district where hydrocarbon-related contamination has been implicated by earlier work [7]. Conclusions about contamination origin are correspondingly restricted (Section 4.1).
- (ix) External status. The model is unvalidated outside the calibration district. The Tuzkhurmatu comparison is a bounding exercise under an unverified assumption and provides no evidence of transferability.

5. Conclusions

The Weighted Arithmetic WQI was used to assess groundwater quality at twelve wells in Baiji District, Salah Al-Din Governorate, referenced to WHO guideline values, mapped in GIS and used to calibrate an EC surrogate. The primary water-

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

quality finding is a guideline comparison rather than an index result. No well in the district complies with the WHO guideline values for the major-ion suite. The least mineralised well contains 1,545 mg/L TDS and exceeded the guideline values for calcium, magnesium, sodium, chloride and sulfate. However, ten wells (83 %) belong to the Good class of WQI, while two western fringe wells are hypersaline and considered unsuitable. The difference between the two statements is a property of the index, not of the water: removing pH from the weight vector changes the district-mean index from Poor to Very poor, whereas using an alternative Iraqi class scheme converts 83 % "Good" into 100 % "poor or worse". Index classifications should therefore not be reported without the weight vector and class scheme that produced them.

Index values are globally clustered in space (Moran's I = 0.240; pseudo-p = 0.0055), but no well is locally significant, so the two hypersaline wells are best described as a spatially coherent pair rather than a proven regional domain. The salinity observed at these sites is consistent with geogenic evaporite dissolution and evapo-concentration, and there is no evidence of an agricultural or sewage nitrogen source; no trace metals, hydrocarbons or microbiological determinands were measured and this study cannot exclude industrial contamination and does not contradict earlier work in the district that has implicated refinery discharges.

The methodological contribution lies in providing a diagnostically transparent surrogate of the index for EC and its diagnostic sequence to interpret it. The model accounts for 99.7 % of variance across the entire salinity range, although formal influence diagnostics indicate this to be leverage-inflated, and the defensible measure of local skill is R2 = 0.814 (95 % CI 0.40-0.95) over the Good class, with a leave-one-out RMSE of 7.2 index units and a 95 % prediction interval of approximately ±15 index units for a single well. Removing the arithmetically dependent TDS term from the index changes R^2 by 0.018 only, demonstrating that the relationship is not an artefact of data leakage. Inversion of the model yields provisional field thresholds of EC ≈ 1,986, 4,415 and 14,132

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

µS/cm for the WQI = 50, 100 and 300 boundaries, of which only the first lies within the range of observations; the remainder are extrapolations across a data gap. Thus, one single two-minute reagent-free measurement can support an indicative class assignment, pending confirmation close to class boundaries, which may allow limited laboratory capacity to be directed elsewhere. A transparent scenario model indicates that this would reduce routine full analyses by a factor of 2.7 under the recommended schedule and by between 2.0 and 6.0 across plausible schedules - a scenario estimate, not a measured saving, and not the order of magnitude claimed in the original submission.

Two findings constrain the contribution. External application to an independent data set from the same governorate under-predicts the index by 49–61 %. As that data set is published as ranges rather than paired observations this is a bounding comparison and not a validation, but it gives no evidence of transferability and the fitted coefficients must be regarded as site specific. And the study rests on a single campaign at twelve wells, so it cannot establish temporal stability, predictive accuracy adequate for deployment, or the existence of a regional salinity domain.

The appropriate description of this work is therefore a pilot methodological study. What it establishes is that a headline coefficient of determination in this setting can be leverage-inflated and arithmetically contaminated, that both effects can be quantified with the diagnostic sequence set out here, and that index classifications are sensitive to weighting and class-boundary conventions to the point of reversal. Whether an EC surrogate can serve as an operational monitoring instrument for SDG 6.1 reporting in arid Iraqi districts remains an open question, and answering it requires seasonal replication, a denser well network spanning the 4,000-15,000 µS/cm range, independent gravimetric TDS, and a documented QA/QC programme.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Recommendations

1. Wells W11 and W12 should be removed from direct drinking use and considered for desalination or connection to an alternative monitored supply; their status should be communicated to users without delay. This recommendation follows from the measured concentrations and does not depend on any model result in this paper.
2. Because every well in the district exceeds WHO guideline values for TDS and for four or more major ions, no well should be described as suitable for direct drinking supply on the basis of its index class. Treatment or blending requirements should be assessed against measured concentrations.
3. EC screening calibrated to Table 12 may be used as an interim tool between full laboratory campaigns, with escalation at approximately 4,200-4,400 µS/cm. Given the prediction interval of ±15 index units, wells screening within about 600 µS/cm of a class boundary should be escalated to full analysis rather than classified by screening alone.
4. The well network should be densified and sampled over at least two contrasting seasons, with the explicit aim of adding observations in the 4,000-15,000 µS/cm range where the present calibration is entirely unconstrained. Seasonal confirmation of the stability of the EC-WQI relationship is a precondition of any operational use of the protocol.
5. For any follow-up campaign, TDS should be determined gravimetrically so that the EC-WQI relationship can be established free of the arithmetic dependence identified in Section 3.7, and the campaign should be conducted under a documented QA/QC plan including field duplicates, blanks, control samples and stated detection limits and uncertainties.
6. For wells within the refinery's zone of influence, trace-metal, hydrocarbon (including BTEX and polycyclic aromatic hydrocarbons), phenol and microbiological analyses should be added, as these determinands cannot be

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

screened by conductivity and their absence from the present dataset is the principal reason this study cannot address contamination origin.

7. The GIS layer generated here should be integrated into a governorate-scale groundwater database with version-controlled update cycles, so that screening results accumulate into a monitoring time series rather than remaining discrete surveys.

8. A single classification convention should be adopted and stated explicitly in all Iraqi groundwater index reporting, together with the weight vector used, and the screening coefficients recalibrated locally against at least ten wells before the protocol is applied outside Baiji District.

Declarations

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflicts of interest: The author declares no competing interests.

Data and Code Availability: All data needed to reproduce the reported numerical results are printed in full in the Supplementary Material at the end of this paper, as Supplementary Tables S1, S2, S3 and S5. The executable analysis script, Supplementary Material S4, accompanies this submission as a file and regenerates every reported quantity from those tables using the fixed seed recorded in Table B1. Index values reconstructed from the printed per-well record agree with the source survey record to within 0.00485 index units. No restrictions apply to the reuse of any of this material.

Ethics approval: Not applicable. The study involved no human participants or animal subjects.

Author contributions: A.A.K.: conceptualisation, methodology, formal analysis, investigation, data curation, software, visualisation, writing - original draft, writing - review and editing.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Acknowledgements: The author thanks the Directorate of Groundwater, Salah Al-Din, for access to the well-survey records, and the College of Engineering, Tikrit University, for institutional support.

References

[1] Foster, S., Pulido-Bosch, A., Vallejos, Á., Molina, L., Llop, A., MacDonald, A.M. (2018). Impact of irrigated agriculture on groundwater-recharge salinity: a major sustainability concern in semi-arid regions. Hydrogeology Journal, 26(8), 2781--2791. https://doi.org/10.1007/s10040-018-1830-2

[2] Etikala, B., Adimalla, N., Madhav, S., Somagouni, S.G., Keshava Kiran Kumar, P.L. (2021). Salinity problems in groundwater and management strategies in arid and semi-arid regions. In: Madhav, S., Singh, P. (Eds.), Groundwater Geochemistry: Pollution and Remediation Methods. John Wiley & Sons, Hoboken, NJ, pp. 42--56.

[3] Al-Kakey, O., Othman, A., Merkel, B. (2023). Identifying potential sites for artificial groundwater recharge using GIS and AHP techniques: a case study of Erbil basin, Iraq. Kuwait Journal of Science, 50(1B), 1--22.

[4] Ismail, A.H., Shareef, M.A., Hassan, G., Alatar, F.M. (2023). Hydrochemistry and water quality of shallow groundwater in the Tikrit area of Salah Al Din Province, Iraq. Applied Water Science, 13(10), 202. https://doi.org/10.1007/s13201-023-02008-y

[5] Alatar, F.M., Ismail, A.H., Shareef, M.A., Boncescu, C. (2023). Hydrochemistry of groundwater in Al-Jallam Desert, Salah Al-Din Governorate, Iraq. Journal of Techniques, 5(3), 15--25. https://doi.org/10.51173/jt.v5i3.1755

[6] Ahmed, S.H., Ibrahim, A.K., Abed, M.F. (2023). Assessing the quality of the groundwater and the nitrate exposure, North Salah Al-Din Governorate, Iraq. Tikrit Journal of Engineering Sciences, 30(1), 25--36.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

[7] Abed, M.F., Zarraq, G.A., Ahmed, S.H. (2022). Assessment of groundwater pollution using aqueous geo-environmental indices, Baiji Province, Salah Al-Din, Iraq. Iraqi Geological Journal, 55(1B), 94--104.
[8] Horton, R.K. (1965). An index number system for rating water quality. Journal of the Water Pollution Control Federation, 37(3), 300--306.
[9] Brown, R.M., McClelland, N.I., Deininger, R.A., Tozer, R.G. (1970). A water quality index --- do we dare? Water and Sewage Works, 117(10), 339--343.
[10] Alhadithi, M. (2019). Evaluation of groundwater quality using Water Quality Index (WQI) and GIS techniques. Iraqi Journal of Agricultural Sciences, 49(2), 285--294. https://doi.org/10.36103/ijas.v49i2.236
[11] Al-Ridah, Z.A., Al-Zubaidi, H., Naje, A.S., Ali, I. (2020). Drinking water quality assessment by using water quality index (WQI) for Hillah River, Iraq. Ecology, Environment and Conservation, 26(1), 390--399.
[12] Abed, E.H., Hassan, W.H., Faisal, A.A.H. (2021). Using of geographical information system and water quality index for evaluation of groundwater quality in the Dammam Zone, Faddak farm, Karbala, Iraq. Journal of Physics: Conference Series, 1973(1), 012181. https://doi.org/10.1088/1742-6596/1973/1/012181
[13] Manii, J.K., Hasan, A.A., Mauff, K.R., Al-Qaraghuli, S.A. (2024). GIS technique for assessing quality and hydrochemistry of Karbala Governorate groundwater. Journal of Geosciences and Geomatics, 12(3), 73--79.
[14] Hamed, M.H., Dara, R.N., Kirlas, M.C. (2024). Application of geographical information system -- water quality index techniques for evaluating the groundwater quality in Kaniqirzhala landfill, Kurdistan Region, Iraq. Iraqi Geological Journal, 57(2C), 296--315.
[15] Fatah, K.K., Hamed, M.H., Saeed, M.H., Dara, R.N. (2020). Evaluation of groundwater quality by using GIS and water quality index techniques for wells in Bardarash area, northern Iraq. Iraqi Geological Journal, 53(3C), 87--101.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

[16] Hussain, H.M., Al-Haidarey, M., Al-Ansari, N., Knutsson, S. (2014). Evaluation and mapping groundwater suitability for irrigation using GIS in Najaf Governorate, Iraq. Journal of Environmental Hydrology, 22, 1--16.
[17] Al-Sudani, H.I.Z. (2025). Groundwater utilization and water quality in Khanaqin District, Diyala Governorate, northeast of Iraq. Resources Environment and Information Engineering, 6(1), 305--312. https://doi.org/10.25082/reie.2024.01.004
[18] Adimalla, N. (2021). Application of the entropy weighted water quality index (EWQI) and the pollution index of groundwater (PIG) to assess groundwater quality for drinking purposes: a case study in a rural area of Telangana State, India. Archives of Environmental Contamination and Toxicology, 80(1), 31--40. https://doi.org/10.1007/s00244-020-00800-4
[19] Amiri, V., Rezaei, M., Sohrabi, N. (2014). Groundwater quality assessment using entropy weighted water quality index (EWQI) in Lenjanat, Iran. Environmental Earth Sciences, 72(9), 3479--3490. https://doi.org/10.1007/s12665-014-3255-0
[20] Seifi, A., Dehghani, M., Singh, V.P. (2020). Uncertainty analysis of water quality index (WQI) for groundwater quality evaluation: application of Monte-Carlo method for weight allocation. Ecological Indicators, 117, 106653. https://doi.org/10.1016/j.ecolind.2020.106653
[21] Raheja, H., Goel, A., Pal, M. (2022). Prediction of groundwater quality indices using machine learning algorithms. Water Practice and Technology, 17(1), 336--351. https://doi.org/10.2166/wpt.2021.120
[22] Mohammed, A.K., Al-Ani, R.R., Hamad, N.A., Abdulhameed, A.S. (2025). Adoption of water quality index and multivariate statistical analyses to appraise the groundwater for drinkable purposes. International Journal of Environmental Impacts, 8(3), 511--522.
[23] Abidi, J.H., Elzain, H.E., Sabarathinam, C., Selvam, S., Sathish, S., Senapathi, V. (2024). Evaluation of groundwater quality indices using multi-

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

criteria decision-making techniques and a fuzzy logic model in an irrigated area. Groundwater for Sustainable Development, 25, 101122. https://doi.org/10.1016/j.gsd.2024.101122

[24] Shaw, S.K., Sharma, A. (2024). Assessment of groundwater quality and suitability for irrigation purpose using irrigation indices, remote sensing and GIS approach. Groundwater for Sustainable Development, 26, 101297. https://doi.org/10.1016/j.gsd.2024.101297

[25] United Nations (2015). Transforming Our World: The 2030 Agenda for Sustainable Development. Resolution A/RES/70/1, United Nations General Assembly, New York.

[26] Abed, M.F., Zarraq, G.A., Ahmed, S.H. (2021). Hydrogeochemical assessment of groundwater quality and its suitability for irrigation and domestic purposes in rural areas, north of Baiji city, Iraq. Iraqi Journal of Science, 62(7), 2296--2306.

[27] APHA (2017). Standard Methods for the Examination of Water and Wastewater, 23rd ed. American Public Health Association / AWWA / WEF, Washington, DC.

[28] World Health Organization (2011). Guidelines for Drinking-water Quality, 4th ed. WHO, Geneva.

[29] Central Organization for Standardization and Quality Control (2009). Iraqi Standard for Drinking Water, IQS 417. COSQC, Baghdad, Iraq.

[30] Richards, L.A. (Ed.) (1954). Diagnosis and Improvement of Saline and Alkali Soils. USDA Agriculture Handbook No. 60, United States Department of Agriculture, Washington, DC.

[31] Wilcox, L.V. (1955). Classification and Use of Irrigation Waters. USDA Circular No. 969, United States Department of Agriculture, Washington, DC.

[32] Raghunath, H.M. (1987). Groundwater, 2nd ed. Wiley Eastern Ltd., New Delhi, India.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

[33] Doneen, L.D. (1964). Notes on Water Quality in Agriculture. Water Science and Engineering Paper 4001, Department of Water Sciences and Engineering, University of California, Davis.
[34] Kelley, W.P. (1963). Use of saline irrigation water. Soil Science, 95(6), 385--391.
[35] Ayers, R.S., Westcot, D.W. (1985). Water Quality for Agriculture. FAO Irrigation and Drainage Paper 29, Rev. 1, Food and Agriculture Organization of the United Nations, Rome.
[36] Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 23rd ACM National Conference, ACM Press, New York, pp. 517--524.
[37] ESRI (2025). ArcGIS Pro: How Kriging Works. Environmental Systems Research Institute, Redlands, CA. Available at: https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/how-kriging-works.htm
[38] Moran, P.A.P. (1950). Notes on continuous stochastic phenomena. Biometrika, 37(1--2), 17--23. https://doi.org/10.1093/biomet/37.1-2.17
[39] Anselin, L. (1995). Local indicators of spatial association --- LISA. Geographical Analysis, 27(2), 93--115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x
[40] Cook, R.D. (1977). Detection of influential observation in linear regression. Technometrics, 19(1), 15--18. https://doi.org/10.2307/1268249
[41] Efron, B., Tibshirani, R.J. (1993). An Introduction to the Bootstrap. Chapman & Hall/CRC, New York.
[42] Gibbs, R.J. (1970). Mechanisms controlling world water chemistry. Science, 170(3962), 1088--1090. https://doi.org/10.1126/science.170.3962.1088
[43] Piper, A.M. (1944). A graphic procedure in the geochemical interpretation of water analyses. Transactions, American Geophysical Union, 25(6), 914--928. https://doi.org/10.1029/TR025i006p00914

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

[44] Alharbawee, N.A.H., Mohammed, A.J. (2024). Water quality assessment of Tigris River using multivariate statistical techniques. Iraqi Journal of Science, 65(3), 1266--1275. https://doi.org/10.24996/ijs.2024.65.3.8

[45] Al-Dabbas, M.A., Al-Ali, I.A., Husain, M.M. (2024). Water quality assessment of the Euphrates River from Haditha to Al-Nasiriyah, Iraq. Iraqi Geological Journal, 57(2B), 272--285. https://doi.org/10.46717/igj.57.2B.18ms-2024-8-28

[46] Radhi, N.A., Sachit, D.E., Al-Madhhachi, A.S.T. (2025). Groundwater quality analyses for irrigation purposes in Salah Al-Din, Iraq: a review. Al-Esraa University College Journal for Engineering Sciences, 7(12), 94-106.

[47] Al-Gburi, H.F., Al-Ali, I.A., Dar, F.A., Al-Sheikh, O.N. (2024). Groundwater quality assessment and pollution sources identification using statistical analyses at Missan Governorate, Southeast Iraq. Discover Sustainability, 5, 416. https://doi.org/10.1007/s43621-024-00578-8

[48] Al-Qaraghuli, S.A.F., Fadhil, L.A., Saad, W.I., Hasan, A.A. (2025). Using water quality index to assessment of ground water quality for drinking uses in Tuzkhurmatu City, Iraq. Central Asian Journal of Medical and Natural Science, 6(2), 507--515. https://doi.org/10.51699/cajmns.v6i2.2726

[49] Ramakrishnaiah, C.R., Sadashivaiah, C., Ranganna, G. (2009). Assessment of water quality index for the groundwater in Tumkur Taluk, Karnataka State, India. E-Journal of Chemistry, 6(2), 523--530. https://doi.org/10.1155/2009/757424

[50] Şener, Ş., Şener, E., Davraz, A. (2017). Evaluation of water quality using water quality index (WQI) method and GIS in Aksu River (SW-Turkey). Science of the Total Environment, 584--585, 131--144. https://doi.org/10.1016/j.scitotenv.2017.01.102

[51] Batarseh, M., Imreizeeq, E., Tilev, S., Al Alaween, M., Suleiman, W., Al Remeithi, A.M., Al Tamimi, M.K., Al Alawneh, M. (2021). Assessment of groundwater quality for irrigation in the arid regions using irrigation water quality

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

index (IWQI) and GIS-zoning maps: case study from Abu Dhabi Emirate, UAE. Groundwater for Sustainable Development, 14, 100611. https://doi.org/10.1016/j.gsd.2021.100611

[52] Al-Gburi, H.F., Al-Tawash, B.S., Al-Tamimi, O.S., Schüth, C. (2023). Impacts of hydrogeochemical processes and land use practices on groundwater quality of Shwan sub-basin, Kirkuk, northern Iraq. Heliyon, 9(3), e13995. https://doi.org/10.1016/j.heliyon.2023.e13995

[53] Ismail, A.H., Hassan, G., Sarhan, A. (2020). Hydrochemistry of shallow groundwater and its assessment for drinking and irrigation purposes in Tarmiah district, Baghdad Governorate, Iraq. Groundwater for Sustainable Development, 10, 100300. https://doi.org/10.1016/j.gsd.2019.100300

[54] Hassan, W.H., Hussein, H.H., Nile, B.K. (2021). The effect of climate change on groundwater recharge in unconfined aquifers in the Western Desert of Iraq. Groundwater for Sustainable Development, 16, 100700. https://doi.org/10.1016/j.gsd.2021.100700

[55] Mohseni, U., Pande, C.B., Pal, S.C., Alshehri, F. (2024). Prediction of weighted arithmetic water quality index for urban water quality using ensemble machine learning model. Chemosphere, 352, 141393. https://doi.org/10.1016/j.chemosphere.2024.141393

[56] Hussein, E.E., Derdour, A., Zerouali, B., Almaliki, A., Wong, Y.J., Ballesta-de los Santos, M., Elbeltagi, A. (2024). Groundwater quality assessment and irrigation water quality index prediction using machine learning algorithms. Water, 16(2), 264. https://doi.org/10.3390/w16020264

[57] Al-Kubaisi, Q.Y., Al-Abadi, A.M., Al-Ghanimy, M.A. (2018). Mapping groundwater quality index for irrigation in the Dibdibba aquifer at Karbala--Najaf plateau, central Iraq. Iraqi Journal of Science, 59(3C), 1636--1652.

[58] Breusch, T.S., Pagan, A.R. (1979). A simple test for heteroscedasticity and random coefficient variation. Econometrica, 47(5), 1287-1294. https://doi.org/10.2307/1911963

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

[59] Li, J., Heap, A.D. (2011). A review of comparative studies of spatial interpolation methods in environmental sciences: performance and impact factors. Ecological Informatics, 6(3-4), 228-241. https://doi.org/10.1016/j.ecoinf.2010.12.003

[60] Harrell, F.E. (2015). Regression Modeling Strategies, 2nd ed. Springer, Cham. https://doi.org/10.1007/978-3-319-19425-7

Appendix A. Per-well regression diagnostics and local spatial statistics

Table A1 reports the influence diagnostics, cross-validation results and local spatial statistics of Sections 3.5 to 3.7 well by well, so that each may be verified individually. It corresponds to Supplementary Table S3. Leverage and Cook's distance are computed for the full-range model of Eq. (15); the leave-one-out prediction is the value obtained when the well in question is excluded from the fit. Criterion values are 2p/n = 0.333 for leverage and 4/n = 0.333 for Cook's distance. LISA statistics use row-standardised inverse-distance weights with 9,999 conditional permutations. The values in Table A1 were checked against Supplementary Table S3 and agree at displayed precision; S3 additionally contains coordinates, pH, TDS, WQI_noTDS, z_WQI, spatial lag and quadrant fields not repeated in this appendix.

Table A1. Per-well regression diagnostics and local indicators of spatial association for the twelve Baiji wells.

Well	EC (µS/cm)	WQI	Leverage h_{ii}	Cook's D	Std. residual	LOO prediction	LOO residual	Local I_i	LISA pseudo-p
W1	2,380	54.6	0.109	0.019	−0.551	58.5	−3.9	0.219	0.143
W2	2,430	54.9	0.108	0.027	−0.665	59.6	−4.7	0.231	0.236
W3	2,400	55.1	0.108	0.018	−0.537	58.9	−3.8	0.219	0.105
W4	2,920	65.6	0.101	0.018	−0.566	69.6	−4.0	0.158	0.267

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Well	EC (µS/cm)	WQI	Leverage h_{ii}	Cook's D	Std. residual	LOO prediction	LOO residual	Local I_i	LISA pseudo-p
W5	2,820	71.5	0.102	0.026	0.680	66.7	+4.8	0.189	0.239
W6	3,540	73.1	0.093	0.099	−1.384	82.9	−9.8	0.022	0.450
W7	2,800	73.5	0.102	0.064	1.058	66.0	+7.5	0.151	0.274
W8	3,430	77.5	0.094	0.006	−0.346	79.9	−2.4	0.173	0.209
W9	3,250	81.8	0.097	0.044	0.904	75.4	+6.4	0.131	0.193
W10	3,780	97.0	0.091	0.123	1.569	85.9	+11.1	0.077	0.188
W11	17,500	361.2	0.462	1.156	−1.646	376.3	−15.1	0.668	0.077
W12	18,590	398.8	0.534	1.326	1.528	383.7	+15.1	0.644	0.429

Two wells exceed both diagnostic criteria: W11 and W12 have leverage of 0.462 and 0.534 against a criterion of 0.333, and Cook's distances of 1.156 and 1.326 against a criterion of 0.333, both exceeding unity. No other well approaches either criterion. The two largest leave-one-out residuals in absolute terms are also those of W11 and W12 (±15.1 index units), although these are small in relative terms (4.0 % and 3.8 %) because the index values themselves are large. No well attains local spatial significance at $\alpha = 0.05$.

Appendix B. Computational environment and analysis parameters

Table B1 records, in full, the computational settings recoverable from the analysis code accompanying this paper. The executable script logs the Python and package versions at execution and is the reproducibility source; the archival PDF did not state exact Python or package patch versions.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Table B1. Computational environment and the parameters of every analysis reported in this paper.

Item	Specification
Language and version	Python 3.11; exact patch version not recorded in the source materials
Libraries	NumPy; SciPy; pandas; exact package versions not recorded in the source materials
Pseudo-random seed	20260904 (NumPy PCG64 default_rng), set once at the head of the script
Regression	Ordinary least squares via scipy.stats.linregress; no weighting, no intercept suppression
Leverage	$h_{ii} = 1/n + (x_i - \bar{x})^2/\Sigma(x - \bar{x})^2$; criterion 2p/n with p = 2
Cook's distance	$D_i = e_i^2 h_{ii} / (p \cdot s^2(1 - h_{ii})^2)$; criterion 4/n
Normality test	Shapiro-Wilk (scipy.stats.shapiro)
Heteroscedasticity tests	Spearman rank correlation of |residual| against EC; Breusch-Pagan LM = $n \cdot R^2$ from the auxiliary regression of squared residuals on EC, referred to $\chi^2(1)$
Serial correlation	Durbin-Watson statistic computed on residuals ordered by well index
Cross-validation	Leave-one-out; the model is refitted on n − 1 observations for each held-out well; RMSE, MAE and MAPE computed on the held-out predictions
Bootstrap	Non-parametric, case resampling with replacement, 10,000 replicates; replicates with fewer than two distinct x values or a non-positive slope are discarded; thresholds computed as $(WQI_c - b_0^*)/b_1^*$ per replicate; confidence limits are the 2.5th and 97.5th percentiles
Distance metric	Haversine great-circle distance, Earth radius 6,371.0088 km
Spatial weights	$w_{ij} = 1/d_{ij}$ for $i \neq j$, $w_{ii} = 0$, row-standardised
Global Moran's I	Eq. (10); significance from 9,999 unconditional permutations of the attribute vector; pseudo-p = (M + 1)/(R + 1) where M is the number of permuted values at least as extreme as the observed

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Item	Specification
Local Moran's I (LISA)	Eq. (11); 9,999 conditional permutations per location, holding the focal value fixed and permuting the remaining n − 1 values; two-sided pseudo-p; no multiple-comparison correction applied
IDW interpolation	Power p ∈ {1, 2, 3, 4}; all twelve points used for each estimate; leave-one-out cross-validation for the reported RMSE, MAE and predictive R^2; p = 2 retained for display
Prediction intervals	t-based, $\pm t_{0.975,n-2} \cdot s \cdot \sqrt{(1 + 1/n + (x_0 - \bar{x})^2/\Sigma(x - \bar{x})^2)}$
R^2 confidence interval	Fisher z-transformation of r, SE = $1/\sqrt{(n - 3)}$, back-transformed and squared
Leakage control	Eq. (14) applied to the published per-well WQI and TDS; refit performed on the transformed index

Supplementary Material

All supplementary material is printed in full at the end of this paper, immediately below. Supplementary Tables S1, S2, S3 and S5 are reproduced here as complete tables so that every reported result can be verified without recourse to a separate file; Supplementary Material S4 is the executable analysis script, which accompanies the submission as a file because it is code rather than a table. The tables are direct transcriptions of the survey record and deterministic calculations from it. Laboratory and sampling metadata that the source record does not contain are listed explicitly as not recorded rather than left blank or imputed.

Supplementary Table S1 - Per-well analytical record. Given below in four panels. It comprises coordinates, pH, EC, TDS, seven ionic determinands, the computed index, and ionic balances derived from the recorded ions. Fields absent from the source record are listed as not recorded.

Supplementary Table S2 - Intermediate WQI calculations. Given below, computed with unrounded normalised inverse-guideline weights. It reports V_i, S_i, V_i^0, Q_i, W_i and W_iQ_i for all nine parameters in each of the twelve wells, a well-

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas
OPEN ACCESS JOURNALS

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

level reconstruction check, and the pH-removed and equal-weight counterfactuals underlying Section 3.3.

Supplementary Table S3 - Well-level analytical input and regression diagnostics. Given below in two panels: Panel S3a lists the measured variables, the salinity ratio and the computed index; Panel S3b lists the influence diagnostics, the leakage-controlled index and the local spatial statistics. Panel S3b contains Table A1 of Appendix A together with three further derived columns.

Supplementary Material S4 - Reproducible analysis code. An executable script for Python 3.11 or later, using the fixed pseudo-random seed 20260904, accompanying this submission as a file. It regenerates every numerical result, table and figure in this paper from the record in Supplementary Table S1. Because it is code rather than a table it is not printed here; all of its output is, however, contained in Supplementary Tables S1, S2, S3 and S5.

Supplementary Data S5 - Extended well table. Given below as a single consolidated record per well, merging the measured analytical variables, the recorded ionic concentrations, the ionic balance and the computed index and class. Sampling, construction and laboratory metadata that the source record does not contain are identified as not recorded in Panel S1d.

Supplementary Table S1. Per-well analytical record

All concentration values below are transcribed from the source survey record; none is an invented laboratory result. Fields absent from that recorook are listed once in Panel S1d as not recorded.

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Panel S1a. Identification, coordinates and core analytical fields.

ID	Lat (°N)	Lon (°E)	pH	EC (µS/cm)	TDS (mg/L)	WQI	Class
W1	34.85472	43.49988	7.25	2,380	1,545	54.63	Good
W2	34.82843	43.48934	7.22	2,430	1,580	54.87	Good
W3	34.80694	43.50361	7.20	2,400	1,598	55.05	Good
W4	34.91997	43.49040	7.22	2,920	1,890	65.59	Good
W5	34.83236	43.49966	7.25	2,820	1,830	71.51	Good
W6	35.15833	43.31444	7.19	3,540	2,300	73.14	Good
W7	34.90675	43.50144	7.24	2,800	1,813	73.52	Good
W8	34.83760	43.50026	7.19	3,430	2,220	77.53	Good
W9	34.89934	43.48264	7.24	3,250	2,108	81.83	Good
W10	34.94614	43.49254	7.21	3,780	2,457	97.00	Good
W11	34.74715	43.09998	7.16	17,500	11,281	361.20	Unsuitable for drinking
W12	34.82867	43.04017	7.14	18,590	11,984	398.75	Unsuitable for drinking

Panel S1b. Recorded per-well ionic concentrations (mg/L).

ID	Ca^{2+}	Mg^{2+}	Na^{+}	Cl^{-}	HCO_3^{-}	SO_4^{2-}	NO_3^{-}
W1	179	63	243	296	145	585	0.1
W2	183	67	247	300	153	591	0.1
W3	185	69	249	302	161	593	1.1
W4	169	110	253	583	199	538	0.6
W5	195	126	271	423	239	542	0.2
W6	217	114	361	358	474	731	0.2
W7	235	120	298	352	143	637	0.2
W8	224	134	297	538	228	759	1.2
W9	214	156	299	569	241	586	0.8
W10	298	185	256	584	245	838	1.3
W11	1,066	718	1,657	2,537	1,294	3,800	0.5
W12	1,177	829	1,768	2,648	1,316	3,913	0.4

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Panel S1c. Ionic balance calculated from the recorded ions. Potassium and carbonate were not present in the source record and are therefore omitted, not imputed.

ID	Σ cations (meq/L)	Σ anions (meq/L)	CBE (%)	QA flag
W1	24.69	22.91	+3.74	Within ±5 %
W2	25.39	23.28	+4.34	Within ±5 %
W3	25.74	23.52	+4.50	Within ±5 %
W4	28.49	30.92	-4.09	Within ±5 %
W5	31.89	27.14	+8.05	Exceeds ±5 %
W6	35.91	33.09	+4.09	Within ±5 %
W7	34.56	25.54	+15.02	Exceeds ±5 %
W8	35.12	34.73	+0.56	Within ±5 %
W9	36.52	32.21	+6.27	Exceeds ±5 %
W10	41.23	37.96	+4.13	Within ±5 %
W11	184.35	171.90	+3.50	Within ±5 %
W12	203.86	177.74	+6.84	Exceeds ±5 %
District mean	58.98	53.41	+4.95	Within ±5 %

Panel S1d. Required metadata not present in the source survey record.

Required field	Status in the source survey record
Original DMS coordinates	Not recorded; decimal-degree coordinates are supplied
Well depth and completion type	Not recorded
Static/dynamic water level and yield	Not recorded
Sampling date and season	Not recorded
Laboratory name	Not recorded
Analytical method and instrument	Not recorded
Method detection limits	Not recorded
EC meter/calibration	Not recorded
Preservation and transport protocol	Not recorded
K^+ and CO_3^{2-} concentrations	Not recorded
Imputation flag	No supplied numeric value is marked as imputed

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Panel S2a-2. Parameter-level calculations for W3 and W4.

Parameter	S_i	V_i^0	W_i	W3 V_i	W3 Q_i	W3 W_iQ_i	W4 V_i	W4 Q_i	W4 W_iQ_i
pH	8.5	7	0.614230	7.20	13.33	8.190	7.22	14.67	9.009
TDS	500	0	0.010442	1,598	319.60	3.337	1,890	378.00	3.947
Ca	75	0	0.069613	185	246.67	17.171	169	225.33	15.686
Mg	50	0	0.104419	69	138.00	14.410	110	220.00	22.972
Na	200	0	0.026105	249	124.50	3.250	253	126.50	3.302
Cl	250	0	0.020884	302	120.80	2.523	583	233.20	4.870
SO4	250	0	0.020884	593	237.20	4.954	538	215.20	4.494
NO3	45	0	0.116021	1.1	2.44	0.284	0.6	1.33	0.155
HCO3	300	0	0.017403	161	53.67	0.934	199	66.33	1.154

Panel S2a-3. Parameter-level calculations for W5 and W6.

Parameter	S_i	V_i^0	W_i	W5 V_i	W5 Q_i	W5 W_iQ_i	W6 V_i	W6 Q_i	W6 W_iQ_i
pH	8.5	7	0.614230	7.25	16.67	10.237	7.19	12.67	7.780
TDS	500	0	0.010442	1,830	366.00	3.822	2,300	460.00	4.803
Ca	75	0	0.069613	195	260.00	18.099	217	289.33	20.141
Mg	50	0	0.104419	126	252.00	26.314	114	228.00	23.808
Na	200	0	0.026105	271	135.50	3.537	361	180.50	4.712
Cl	250	0	0.020884	423	169.20	3.534	358	143.20	2.991
SO4	250	0	0.020884	542	216.80	4.528	731	292.40	6.106
NO3	45	0	0.116021	0.2	0.44	0.052	0.2	0.44	0.052
HCO3	300	0	0.017403	239	79.67	1.386	474	158.00	2.750

Panel S2a-4. Parameter-level calculations for W7 and W8.

Parameter	S_i	V_i^0	W_i	W7 V_i	W7 Q_i	W7 W_iQ_i	W8 V_i	W8 Q_i	W8 W_iQ_i
pH	8.5	7	0.614230	7.24	16.00	9.828	7.19	12.67	7.780
TDS	500	0	0.010442	1,813	362.60	3.786	2,220	444.00	4.636
Ca	75	0	0.069613	235	313.33	21.812	224	298.67	20.791
Mg	50	0	0.104419	120	240.00	25.061	134	268.00	27.984
Na	200	0	0.026105	298	149.00	3.890	297	148.50	3.877
Cl	250	0	0.020884	352	140.80	2.940	538	215.20	4.494
SO4	250	0	0.020884	637	254.80	5.321	759	303.60	6.340
NO3	45	0	0.116021	0.2	0.44	0.052	1.2	2.67	0.309
HCO3	300	0	0.017403	143	47.67	0.830	228	76.00	1.323

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Panel S2a-5. Parameter-level calculations for W9 and W10.

Parameter	S_i	V_i^0	W_i	W9 V_i	W9 Q_i	W9 W_iQ_i	W10 V_i	W10 Q_i	W10 W_iQ_i
pH	8.5	7	0.614230	7.24	16.00	9.828	7.21	14.00	8.599
TDS	500	0	0.010442	2,108	421.60	4.402	2,457	491.40	5.131
Ca	75	0	0.069613	214	285.33	19.863	298	397.33	27.659
Mg	50	0	0.104419	156	312.00	32.579	185	370.00	38.635
Na	200	0	0.026105	299	149.50	3.903	256	128.00	3.341
Cl	250	0	0.020884	569	227.60	4.753	584	233.60	4.878
SO4	250	0	0.020884	586	234.40	4.895	838	335.20	7.000
NO3	45	0	0.116021	0.8	1.78	0.206	1.3	2.89	0.335
HCO3	300	0	0.017403	241	80.33	1.398	245	81.67	1.421

Panel S2a-6. Parameter-level calculations for W11 and W12.

Parameter	S_i	V_i^0	W_i	W11 V_i	W11 Q_i	W11 W_iQ_i	W12 V_i	W12 Q_i	W12 W_iQ_i
pH	8.5	7	0.614230	7.16	10.67	6.552	7.14	9.33	5.733
TDS	500	0	0.010442	11,281	2256.20	23.559	11,984	2396.80	25.027
Ca	75	0	0.069613	1,066	1421.33	98.943	1,177	1569.33	109.246
Mg	50	0	0.104419	718	1436.00	149.946	829	1658.00	173.127
Na	200	0	0.026105	1,657	828.50	21.628	1,768	884.00	23.077
Cl	250	0	0.020884	2,537	1014.80	21.193	2,648	1059.20	22.120
SO4	250	0	0.020884	3,800	1520.00	31.743	3,913	1565.20	32.687
NO3	45	0	0.116021	0.5	1.11	0.129	0.4	0.89	0.103
HCO3	300	0	0.017403	1,294	431.33	7.507	1,316	438.67	7.634

Panel S2a-7. Per-well reconstruction check.

ID	ΣW_i	ΣW_iQ_i	Recomputed WQI	Workbook WQI	Difference
W1	1.000000	54.632856	54.633	54.63	+0.00286
W2	1.000000	54.866264	54.866	54.87	-0.00374
W3	1.000000	55.051944	55.052	55.05	+0.00194
W4	1.000000	65.589648	65.590	65.59	-0.00035
W5	1.000000	71.508159	71.508	71.51	-0.00184
W6	1.000000	73.142495	73.142	73.14	+0.00249
W7	1.000000	73.518816	73.519	73.52	-0.00118
W8	1.000000	77.534850	77.535	77.53	+0.00485

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

ID	ΣW_i	$\Sigma W_i Q_i$	Recomputed WQI	Workbook WQI	Difference
W9	1.000000	81.826840	81.827	81.83	-0.00316
W10	1.000000	97.001407	97.001	97.00	+0.00141
W11	1.000000	361.198945	361.199	361.20	-0.00106
W12	1.000000	398.753654	398.754	398.75	+0.00365
Mean	1.000000	122.052156	122.052	122.052	+0.00049

The maximum absolute well-level difference is 0.00485 WQI units and is explained by rounding of the displayed source values.

Panel S2b. Counterfactual weighting results.

ID	Specified inverse-guideline weights	pH removed; weights renormalised	Equal weights
W1	54.63	115.08	134.75
W2	54.87	118.87	137.75
W3	55.05	121.48	139.58
W4	65.59	146.67	164.51
W5	71.51	158.83	166.25
W6	73.14	169.43	196.06
W7	73.52	165.10	169.40
W8	77.53	180.82	196.59
W9	81.83	186.64	192.06
W10	97.00	229.16	228.23
W11	361.20	919.32	991.10
W12	398.75	1018.79	1064.60
District mean	122.05	294.18	315.08

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Supplementary Table S3. Well-level analytical input and regression diagnostics

The two panels below contain, in printed form, the complete content of the well-level data file. Panel S3a gives the measured variables and the resulting index; Panel S3b gives the regression and spatial diagnostics computed from them. Leverage, Cook's distance, standardised residual and the leave-one-out prediction refer to the full-range model of Eq. (15); the criterion values are $2p/n = 0.333$ and $4/n = 0.333$. Local indicators of spatial association use row-standardised inverse-distance weights with 9,999 conditional permutations per location. The shared columns of Panel S3b are those of Table A1 in Appendix A; Panel S3b adds the leakage-controlled index, the standardised index value, the spatial lag and the Moran quadrant.

Panel S3a. Measured variables, salinity ratio and computed index.

ID	Lat (°N)	Lon (°E)	pH	EC (µS/cm)	TDS (mg/L)	TDS/EC	WQI	Class
W1	34.8547	43.4999	7.25	2,380	1,545	0.6492	54.63	Good
W2	34.8284	43.4893	7.22	2,430	1,580	0.6502	54.87	Good
W3	34.8069	43.5036	7.20	2,400	1,598	0.6658	55.05	Good
W4	34.9200	43.4904	7.22	2,920	1,890	0.6473	65.59	Good
W5	34.8324	43.4997	7.25	2,820	1,830	0.6489	71.51	Good
W6	35.1583	43.3144	7.19	3,540	2,300	0.6497	73.14	Good
W7	34.9067	43.5014	7.24	2,800	1,813	0.6475	73.52	Good
W8	34.8376	43.5003	7.19	3,430	2,220	0.6472	77.53	Good
W9	34.8993	43.4826	7.24	3,250	2,108	0.6486	81.83	Good
W10	34.9461	43.4925	7.21	3,780	2,457	0.6500	97.00	Good
W11	34.7472	43.1000	7.16	17,500	11,281	0.6446	361.20	Unsuitable
W12	34.8287	43.0402	7.14	18,590	11,984	0.6446	398.75	Unsuitable

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Panel S3b. Regression influence diagnostics and local spatial statistics.

ID	Leverage h_{ii}	Cook's D	Std. residual	LOO prediction	LOO residual	WQI* (TDS removed)	z(WQI)	Spatial lag	Local I_i	LISA pseudo-p	Quadrant
W1	0.109	0.019	−0.551	58.5	−3.9	51.96	−0.580	−0.378	0.219	0.143	Low-low
W2	0.108	0.027	−0.665	59.6	−4.7	52.13	−0.578	−0.400	0.231	0.236	Low-low
W3	0.108	0.018	−0.537	58.9	−3.8	52.27	−0.577	−0.380	0.219	0.105	Low-low
W4	0.101	0.018	−0.566	69.6	−4.0	62.31	−0.486	−0.325	0.158	0.267	Low-low
W5	0.102	0.026	0.680	66.7	+4.8	68.41	−0.435	−0.434	0.189	0.239	Low-low
W6	0.093	0.099	−1.384	82.9	−9.8	69.07	−0.421	−0.052	0.022	0.450	Low-low
W7	0.102	0.064	1.058	66.0	+7.5	70.48	−0.418	−0.362	0.151	0.274	Low-low
W8	0.094	0.006	−0.346	79.9	−2.4	73.68	−0.383	−0.450	0.173	0.209	Low-low
W9	0.097	0.044	0.904	75.4	+6.4	78.26	−0.346	−0.377	0.131	0.193	Low-low
W10	0.091	0.123	1.569	85.9	+11.1	92.85	−0.216	−0.356	0.077	0.188	Low-low
W11	0.462	1.156	−1.646	376.3	−15.1	341.28	+2.058	+0.325	0.668	0.077	High-high
W12	0.534	1.326	1.528	383.7	+15.1	377.75	+2.381	+0.270	0.644	0.429	High-high

Eureka Open Access Journals
Open Access | Peer Reviewed | International Journals | Multidisciplinary Areas

Eureka Journal of Geoscience, Materials & Resource Engineering (EJGMRE)

ISSN 2760-4985 (Online) Volume 02, Issue 11, November 2026
This article/work is licensed under CC by 4.0 Attribution
https://eurekaoa.com/index.php/9

Supplementary Table S5. Extended well table

A single flat record per well, merging the measured analytical variables, the recorded ionic concentrations, the ionic balance and the computed index and class. It is the printed equivalent of the machine-readable well table and duplicates, in consolidated form, values given separately in Supplementary Tables S1 and S3. All concentrations are in mg/L.

Table S5. Consolidated per-well record.

ID	pH	EC (µS/cm)	TDS (mg/L)	Ca^{2+}	Mg^{2+}	Na^{+}	Cl^{-}	HCO_3^-	SO_4^{2-}	NO_3^-	Σcat (meq/L)	Σan (meq/L)	CBE (%)	WQI	Class
W1	7.25	2,380	1,545	179	63	243	296	145	585	0.1	24.69	22.91	+3.74	54.63	Good
W2	7.22	2,430	1,580	183	67	247	300	153	591	0.1	25.39	23.28	+4.34	54.87	Good
W3	7.20	2,400	1,598	185	69	249	302	161	593	1.1	25.74	23.52	+4.50	55.05	Good
W4	7.22	2,920	1,890	169	110	253	583	199	538	0.6	28.49	30.92	−4.09	65.59	Good
W5	7.25	2,820	1,830	195	126	271	423	239	542	0.2	31.89	27.14	+8.05	71.51	Good
W6	7.19	3,540	2,300	217	114	361	358	474	731	0.2	35.91	33.09	+4.09	73.14	Good
W7	7.24	2,800	1,813	235	120	298	352	143	637	0.2	34.56	25.54	+15.02	73.52	Good
W8	7.19	3,430	2,220	224	134	297	538	228	759	1.2	35.12	34.73	+0.56	77.53	Good
W9	7.24	3,250	2,108	214	156	299	569	241	586	0.8	36.52	32.21	+6.27	81.83	Good
W10	7.21	3,780	2,457	298	185	256	584	245	838	1.3	41.23	37.96	+4.13	97.00	Good
W11	7.16	17,500	11,281	1,066	718	1,657	2,537	1,294	3,800	0.5	184.35	171.90	+3.50	361.20	Unsuitable
W12	7.14	18,590	11,984	1,177	829	1,768	2,648	1,316	3,913	0.4	203.86	177.74	+6.84	398.75	Unsuitable