AI-Assisted Saltwater Intrusion Assessment and Vulnerability Mapping in the Wadi Ham Aquifer, UAE

Using the GALDIT Framework and Random Forest Classification

ABSTRACT

Saltwater intrusion into coastal aquifers is an escalating threat to freshwater resources in arid regions, particularly in the Arabian Peninsula. This study presents an integrated, AI-assisted framework for assessing saltwater intrusion vulnerability in the Wadi Ham aquifer, Fujairah, United Arab Emirates — the primary groundwater resource serving the eastern UAE coast. The methodology combines the GALDIT multi-parameter hydrogeological index with a Random Forest (RF) machine learning classifier trained on 1,305 georeferenced TDS observations alongside 21 hydrogeological and geospatial features including hydraulic conductivity, aquifer thickness, depth to water table, distance from shore, elevation, and lithology. The RF model achieved an overall accuracy of 86.2% (Cohen $\kappa = 0.808$, mean AUC = 0.980), substantially outperforming the static GALDIT index in spatial discrimination. Predictions on a 200 m resolution grid of 22,200 nodes reveal that 16.4% of the aquifer area is classified as seawater-intruded, 28.0% as saline, 17.3% as brackish, and 38.3% as fresh. Spatial probability mapping identifies a contiguous high-intrusion zone ($P > 0.50$) occupying 16.2% of the modelled domain concentrated within 5 km of the shoreline. Shannon entropy analysis identifies 30.4% of the domain as having elevated prediction uncertainty, highlighting priority areas for additional monitoring. SHAP (SHapley Additive exPlanations) analysis identifies geographic northing, distance from shore, and hydraulic conductivity as the three dominant controls on intrusion status. The results provide actionable spatial datasets for groundwater resource management and establish a transferable AI-assisted workflow applicable to arid coastal aquifers worldwide.

Keywords: *saltwater intrusion; coastal aquifer vulnerability; GALDIT; random forest; SHAP; Wadi Ham; UAE; machine learning; hydrogeology; GIS*

1. INTRODUCTION

Coastal aquifers supply fresh groundwater to more than one billion people globally and are disproportionately concentrated in arid and semi-arid regions where surface water is scarce (Bear et al., 2010). In the Arabian Peninsula, groundwater is often the sole perennial water source, yet decades of over-abstraction coupled with climate-driven recharge deficits have progressively lowered water tables, allowing denser saline water to migrate inland along the freshwater–saltwater interface — a process known as saltwater intrusion (Custodio, 2010; Sherif & Hamza, 2001).

The Wadi Ham aquifer, situated in the Emirate of Fujairah on the eastern coast of the United Arab Emirates, exemplifies this challenge. Fed by episodic flash floods from the Hajar Mountains and a thin annual rainfall averaging 90–110 mm, the aquifer sustains domestic, agricultural, and industrial supply for Fujairah city and surrounding communities. Historical records indicate progressively rising chloride and total dissolved solids (TDS) concentrations throughout the aquifer, signalling ongoing saltwater intrusion from the adjacent Gulf of Oman (Sherif et al., 2012; Groundwater Development Consultants International, 1984).

Traditional vulnerability assessment methods — such as DRASTIC (Aller et al., 1987) and GALDIT (Lobo Ferreira & Cabral, 1991; Chachadi & Lobo Ferreira, 2001) — provide spatially consistent indices based on hydrogeological parameters. The GALDIT framework was specifically designed for coastal aquifer vulnerability and incorporates six parameters: groundwater occurrence type (G), hydraulic conductivity (A), depth to groundwater (L), distance from shore (D), impact of existing saltwater intrusion (I), and aquifer thickness (T). While widely applied, these index methods do not exploit the full predictive information available in observational hydrogeological datasets and assign fixed expert-defined weights that may not optimally reflect local aquifer behaviour.

Machine learning (ML) approaches offer a complementary, data-driven paradigm. Random Forests (Breiman, 2001) are ensemble classifiers that aggregate predictions from hundreds of decision trees, yielding high accuracy, built-in variable importance measures, and robustness to outliers and class imbalance. Several recent studies have successfully applied RF and related ML algorithms to groundwater quality prediction (Rodriguez-Galiano et al., 2014; Yaseen et al., 2022) and saltwater intrusion mapping (Baena-Ruiz et al., 2018; Ahmed et al., 2021), demonstrating consistent improvements over index-based methods.

Explainability remains a concern when deploying ML models in regulatory hydrogeological contexts. SHAP (SHapley Additive exPlanations; Lundberg & Lee, 2017) addresses this by providing theoretically grounded, model-agnostic feature attribution values that quantify each variable's marginal contribution to individual predictions, enabling transparent audit trails alongside high predictive performance.

This study integrates GALDIT feature engineering with a RF classifier and SHAP explainability within a reproducible geospatial Python/GIS workflow. The objectives are: (i) to characterise the spatial distribution of TDS salinity classes across the Wadi Ham aquifer using data-driven ML; (ii) to map seawater intrusion probability and prediction uncertainty at 200 m spatial resolution; (iii) to identify the dominant hydrogeological controls on intrusion using SHAP; and (iv) to compare ML-derived vulnerability classification with the traditional GALDIT index, quantifying agreements and discrepancies to provide guidance for adaptive management.

2. STUDY AREA

The Wadi Ham catchment (approximately 25.1–25.5°N, 56.2–56.4°E) is located in Fujairah Emirate, UAE, draining eastward from the central Hajar Mountain Range to the Gulf of Oman coast (Figure 1). The modelled aquifer domain covers approximately 888 km² and is bounded by UTM coordinates 414,000–438,000 m E and 2,769,000–2,806,000 m N (WGS 1984 / UTM Zone 40N, EPSG:32640).

Geologically, the catchment is dominated by Ophiolite and Gabbro basement complexes of Cretaceous age in the upstream mountainous reaches, transitioning to Quaternary alluvial gravels in the broad wadi floor. Metamorphic and ultramafic units outcrop in the central Hajar ophiolite belt. The principal aquifer is an unconfined alluvial aquifer hosted in coarse gravel and sand fill within the wadi channel, hydraulically connected to fractured ophiolite basement. Aquifer thickness ranges from approximately 7 to 105 m (mean 35.9 m), and saturated hydraulic conductivity (K) varies from less than 1 m/day in fine-grained marginal deposits to over 190 m/day in coarse gravel channels (mean 21.3 m/day; Sherif et al., 2012).

The depth to the water table ranges from approximately 6 to 68 m (mean 21.6 m), reflecting the variable topography from coastal lowlands to mountain front zones. Distance from the Gulf of Oman shoreline varies from near-zero at the coastal margin to approximately 22.6 km at the distal aquifer boundary, providing a strong hydraulic gradient controlling the freshwater–saltwater interface position.

Mean annual rainfall recorded at Fujairah Airport and Fujairah Office stations averages 90–110 mm/year, concentrated in winter months (November–March), with high inter-annual variability. Episodic flash floods provide the dominant recharge mechanism. Groundwater abstraction for municipal supply, irrigation, and industry has increased dramatically since the 1980s, contributing to water table declines of 2–5 m/decade in several monitored well fields.

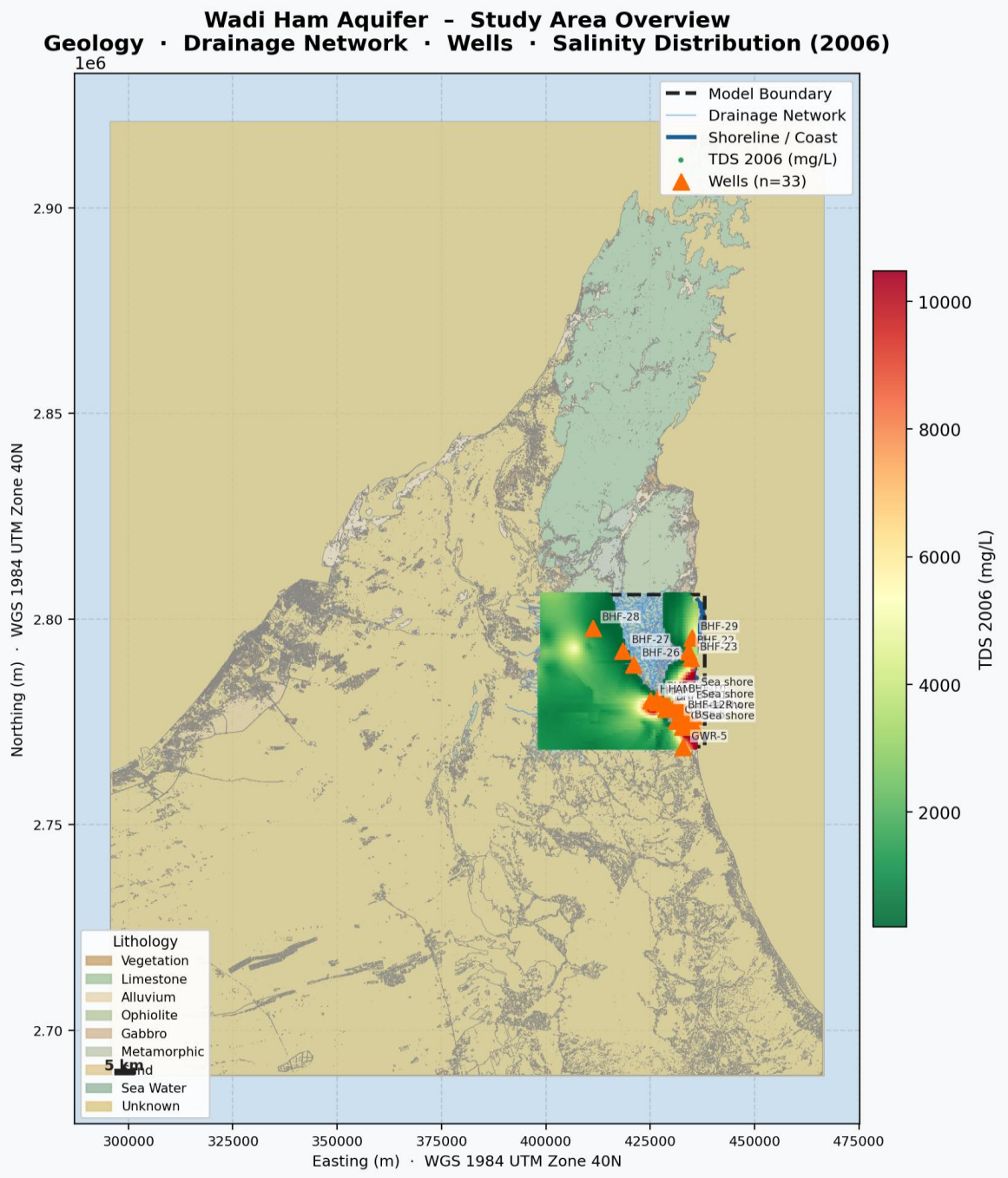

Figure 1. Study area overview: geological units, drainage network, salinity observation points, monitoring wells, and model boundary, Wadi Ham catchment, Fujairah, UAE. Coordinate system: WGS 1984 UTM Zone 40N.

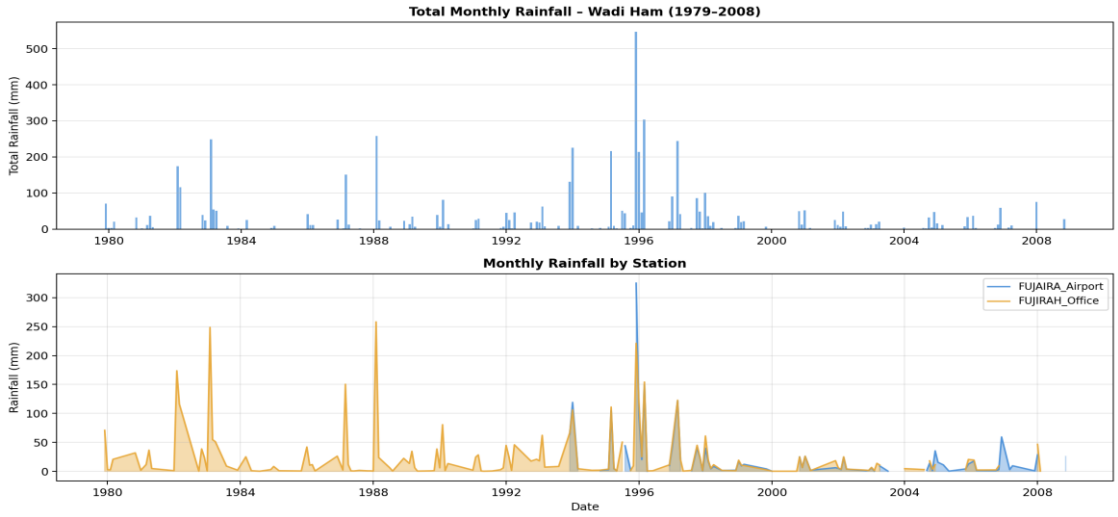

Figure 2. Monthly total precipitation at Fujairah Airport and Fujairah Office stations, 1979–2008 (n=212 records). Mean annual total ~97 mm/year; strong inter-annual variability driven by winter cyclonic events.

3. DATA AND METHODOLOGY

3.1 Data Sources

Eleven spatially distributed datasets were compiled and integrated within a unified GIS framework (EPSG:32640). Salinity observations originate from 3,348 gridded points in the WHam_sal06 dataset (2006 survey), providing TDS values in mg/L across the aquifer extent. Hydraulic properties (K and storativity) were derived from the MODFLOW groundwater model node file (properties_fef.shp; 26,880 nodes), which provides full spatial coverage via inverse distance weighting (IDW) interpolation. Aquifer thickness and bottom elevation were obtained from a regional finite-difference model grid (WH_thickness.txt; 21,741 nodes). Water table elevations were sourced from the water_table_elev shapefile (2,945 survey points). The shoreline geometry was digitised from Landsat imagery to compute distance-from-shore at each model node.

Supplementary datasets include a geological map (geology.shp; 86,221 polygons; 8 lithological units), monitoring well locations (Wells_Locations.shp; 33 wells), wadi drainage network (wh_drainage.shp; 3,985 segments), and model boundary polygon (888 km²). Regional TDS data (tds.txt; 434 points) and calibration well records (WH_wells_water_table_elev_calib.xls) provided independent validation reference. Rainfall data were obtained from two meteorological stations (1979–2008).

Dataset	Source / File	Points/Features	Key Variables
Salinity grid (2006)	WHam_sal06.shp	3,348 pts	TDS_2006 (mg/L)
Hydraulic properties	properties_fef.shp	26,880 nodes	K (m/d), Storativity
Aquifer geometry	WH_thickness.txt	21,741 nodes	Thickness, Base elevation
Water table	water_table_elev.shp	2,945 pts	Depth to water (m)
Shoreline	shoreline.shp	1 LineString	Distance from coast

Geology	geology.shp	86,221 polygons	8 lithological units
Monitoring wells	Wells_Locations.shp	33 wells	Well locations
Model boundary	model_boundary.shp	888 km²	Study extent
TDS (regional)	tds.txt	434 pts	Regional salinity
Rainfall	Monthly_Rain_Wham.csv	212 records	Precipitation (mm)

Table 1. Data sources and key variables used in the Wadi Ham saltwater intrusion assessment.

3.2 GALDIT Vulnerability Framework

The GALDIT index (Chachadi & Lobo Ferreira, 2001) quantifies coastal aquifer vulnerability to saltwater intrusion through six weighted hydrogeological parameters (Table 2). Each parameter is rated on a scale of 2.5 to 10, with higher values indicating greater vulnerability. The composite index is computed as a weighted sum:

$$GVI = \Sigma (w_i \times r_i) / \Sigma w_i$$

where w_i are the GALDIT weights and r_i are the parameter ratings. Parameter ratings applied in this study follow the original Chachadi & Lobo Ferreira (2001) thresholds. Hydraulic conductivity ratings (A) were: <5 m/day → 2.5; 5–10 m/day → 5.0; 10–40 m/day → 7.5; >40 m/day → 10.0. Depth to groundwater (L): <1 m → 10.0; 1–5 m → 7.5; 5–15 m → 5.0; >15 m → 2.5. Distance from shore (D): <100 m → 10.0; 100–500 m → 7.5; 500–1,000 m → 5.0; >1,000 m → 2.5. Aquifer thickness (T): <10 m → 2.5; 10–25 m → 5.0; 25–50 m → 7.5; >50 m → 10.0.

Parameter	Symbol	Description	Weight	Rating Range
Groundwater occurrence	G	Aquifer type (unconfined)	1	7.5 (fixed — unconfined)
Hydraulic conductivity	A	K (m/day) from properties_fef	3	2.5 – 10
Depth to groundwater	L	Depth to water table (m)	4	2.5 – 10
Distance from shore	D	Euclidean distance to shoreline (m)	4	2.5 – 10
Intrusion impact	I	TDS class (observed salinity)	1	2.5 – 10
Aquifer thickness	T	Saturated thickness (m)	2	2.5 – 10

Table 2. GALDIT parameters, descriptions, data sources, weights and rating ranges as applied to the Wadi Ham aquifer (after Chachadi & Lobo Ferreira, 2001).

The resulting GVI distribution across 1,305 training sample points ranged from 2.5 to 8.25 (mean 5.11 ± 1.27). Vulnerability classes were defined as: Low (GVI < 4.0; 20.9%), Moderate (4.0–6.0;

56.2%), High (6.0–8.0; 21.5%), and Very High (>8.0; 1.4%), consistent with standard GALDIT classification thresholds.

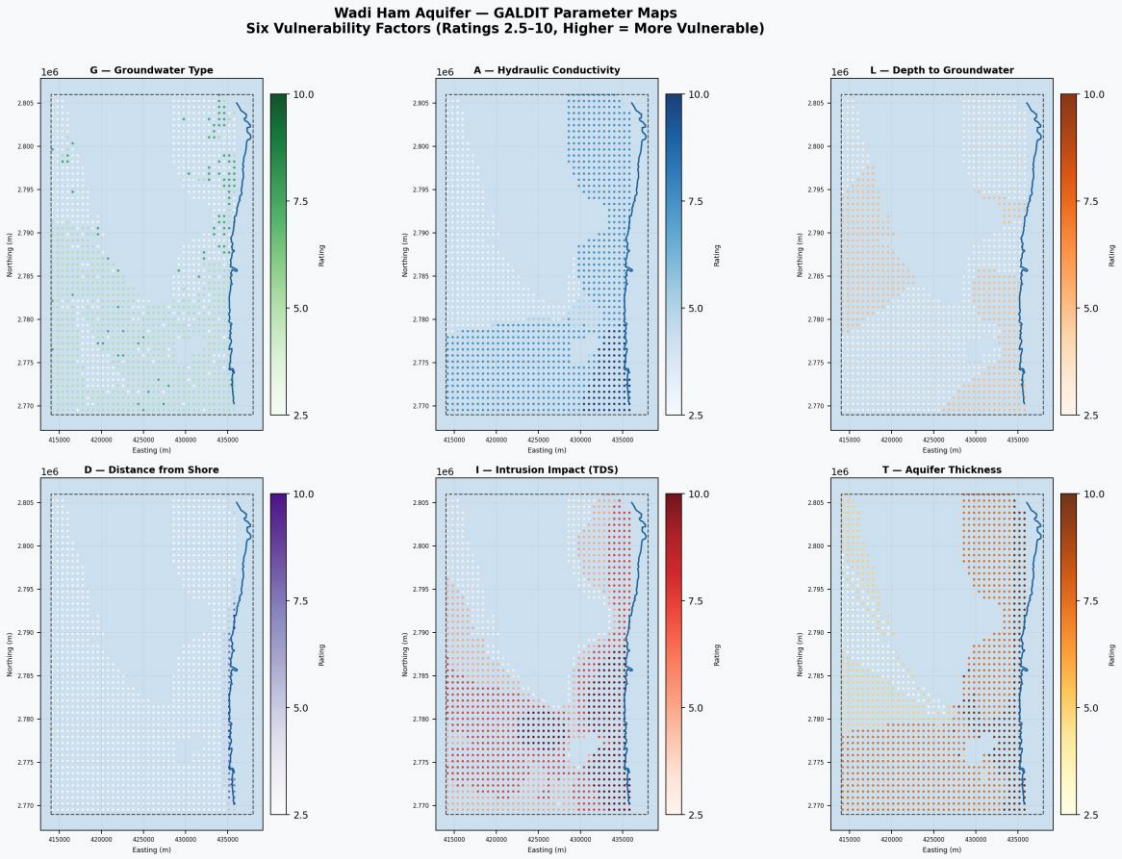

Figure 3. Spatial distribution of GALDIT parameters across the Wadi Ham aquifer: (a) hydraulic conductivity A (m/day), (b) depth to groundwater L (m), (c) distance from shore D (m), (d) intrusion impact I, (e) aquifer thickness T (m), and (f) GALDIT Vulnerability Index (GVI).

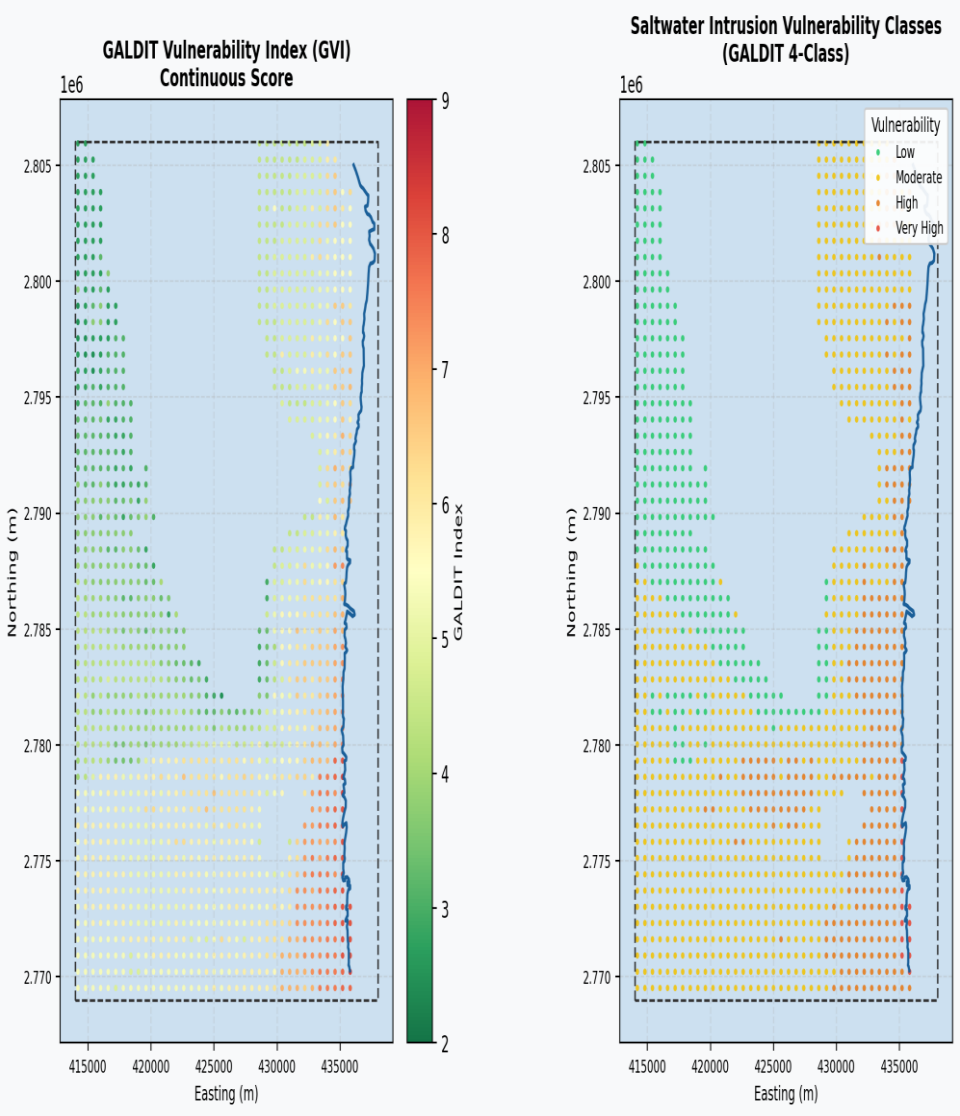

Figure 4. GALDIT composite vulnerability index spatial distribution: Low (green), Moderate (yellow), High (orange), Very High (red). Moderate vulnerability dominates (56.2%); Very High vulnerability is restricted to the near-coastal strip.

3.3 Feature Engineering and Machine Learning Dataset

A total of 21 input features were assembled for each of the 1,305 salinity observation points (Table 3). GALDIT parameter ratings (G_rating, A_rating, L_rating, D_rating, T_rating) were included as engineered features alongside their raw physical counterparts (K_md, DepthToWater_m, Dist_Shore_m, Thickness_m). To prevent data leakage, the I_rating (derived directly from observed TDS) and the composite GVI were excluded from ML features. Lithological class was one-hot encoded into six binary indicator columns matching the categories present in the training data. Geographic coordinates (X, Y in UTM metres) were retained to capture spatial autocorrelation in salinity patterns. The target variable (TDS_Class) was defined as a four-class ordinal label: Fresh (<500 mg/L; 16.5%), Brackish (500–3,000 mg/L; 27.6%), Saline (3,000–10,000 mg/L; 37.9%), and Seawater (>10,000 mg/L; 18.1%).

Feature	Description	Source	Type

X, Y	UTM easting/northing (m)	WHam_sal06.shp	Spatial
K_md	Hydraulic conductivity (m/day)	properties_fef.shp IDW	Hydrogeological
Storativity	Aquifer storativity (–)	properties_fef.shp IDW	Hydrogeological
Thickness_m	Saturated aquifer thickness (m)	WH_thickness.txt IDW	Hydrogeological
DEM_m	Ground surface elevation (m)	WH_thickness.txt	Topographic
BOT2_m	Aquifer base elevation (m)	WH_thickness.txt	Hydrogeological
DepthToWater_m	Depth to water table (m)	water_table_elev.shp	Hydrogeological
Hyd_Head_m	Hydraulic head (m)	water_table_elev.shp	Hydrogeological
Dist_Shore_m	Distance from shoreline (m)	shoreline.shp	Spatial
A_rating	GALDIT hydraulic conductivity rating	Derived	GALDIT
L_rating	GALDIT depth to water rating	Derived	GALDIT
D_rating	GALDIT distance from shore rating	Derived	GALDIT
T_rating	GALDIT thickness rating	Derived	GALDIT
G_rating	GALDIT groundwater type rating	Derived	GALDIT
Lith_Alluvium … Lith_Unknown	Lithology one-hot encoded (6 cols)	geology.shp sjoin	Geological

Table 3. Feature set used for Random Forest training (21 features). I_rating and GVI excluded to prevent data leakage from the target variable.

3.4 Random Forest Model Training and Validation

The dataset was split 80:20 into training (1,044 samples) and stratified test (261 samples) sets using scikit-learn (Pedregosa et al., 2011). The Random Forest classifier (Breiman, 2001) was configured with 500 trees, maximum depth 20, minimum samples per split 4, sqrt feature sampling at each split, and class_weight='balanced_subsample' to address class imbalance. This configuration was selected based on domain expertise and preliminary experimentation rather than an exhaustive grid search, balancing model complexity with computational efficiency.

Model performance was evaluated using overall accuracy, balanced accuracy, Cohen's kappa (κ), per-class precision/recall/F1, per-class AUC under the one-vs-rest ROC curve, and five-fold

stratified cross-validation. Out-of-bag (OOB) error provided an additional unbiased performance estimate using only the training data. Permutation importance, Gini (impurity-based) importance, and SHAP TreeExplainer values were computed to characterise feature contributions at both global and individual prediction levels.

Prediction uncertainty was quantified via normalised Shannon entropy computed from the four-class probability vector at each prediction node:

$$H = -\Sigma\, p_i \log_2(p_i) / \log_2(4)$$

where p_i is the predicted probability of class i. Entropy values range from 0 (deterministic prediction) to 1 (maximum uncertainty across four equally probable classes).

3.5 Spatial Vulnerability Mapping

A regular 200 m grid was generated across the model boundary (22,200 nodes). At each node, all 21 features were extracted using the same IDW and spatial join procedures applied to training data. The fitted RF model was applied to produce: (i) a discrete TDS class prediction; (ii) four-class posterior probabilities; (iii) Shannon entropy uncertainty; and (iv) GALDIT vulnerability class based on the engineered GALDIT ratings. All predictions were saved as a georeferenced GeoTIFF raster and point CSV for subsequent GIS analysis and cartographic output.

4. RESULTS

4.1 GALDIT Vulnerability Index

Across the 1,305 sample points, the composite GALDIT Vulnerability Index (GVI) ranged from 2.5 to 8.25 with a mean of 5.11 (± 1.27). The spatial distribution reflects the dominant influence of distance from shore (weight 4) and depth to groundwater (weight 4): vulnerability is highest in the coastal plain where the water table is shallow and the shoreline is proximate, and decreases progressively inland. Moderate vulnerability (GVI 4.0–6.0) was the dominant class, comprising 56.2% of observation points, followed by High (21.5%), Low (20.9%), and Very High (1.4%) — the latter confined to a narrow coastal fringe within approximately 1–2 km of the Gulf of Oman shoreline.

The GALDIT analysis thus confirms that the eastern coastal sector represents the zone of greatest intrinsic hydrogeological vulnerability to saltwater ingress. However, the static nature of the index — particularly its dependence on fixed expert-assigned weights — limits its ability to integrate the complex, non-linear interactions between hydraulic conductivity, aquifer geometry, lithological heterogeneity, and observed salinity patterns that are captured by the data-driven Random Forest approach.

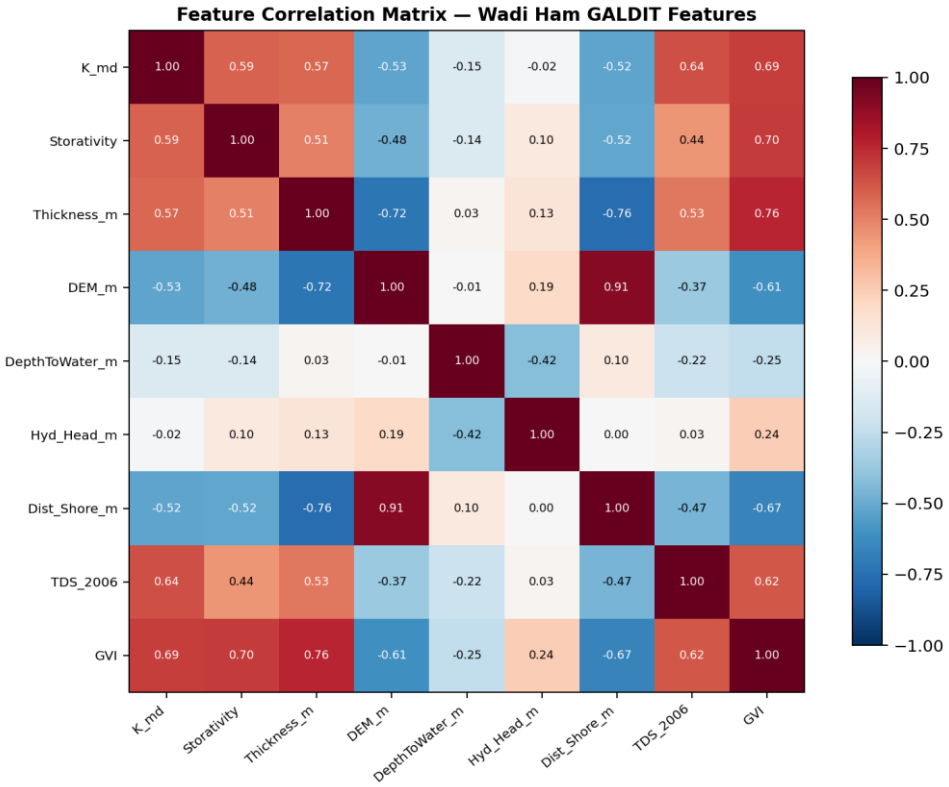

Figure 5. Pearson correlation heatmap between GALDIT features and TDS class. Distance from shore and hydraulic conductivity show the strongest linear associations; spatial coordinates (X, Y) capture directional salinity gradients aligned with the coastal-to-mountain transect.

4.2 Random Forest Model Performance

The Random Forest classifier achieved an overall test-set accuracy of 86.2% (Table 4), substantially exceeding the expected accuracy of a naive majority-class baseline (37.9%). The balanced accuracy of 86.7% confirms that the class_weight='balanced_subsample' strategy effectively mitigated the influence of class imbalance. Cohen's kappa (κ = 0.808) indicates strong agreement beyond chance. The mean one-vs-rest AUC of 0.980 demonstrates excellent discriminative ability across all four salinity classes.

Metric	Value	Interpretation
Overall Accuracy	86.2%	Strong; baseline (majority class) = 37.9%
Balanced Accuracy	86.7%	Class imbalance corrected
Cohen's Kappa (κ)	0.808	Excellent agreement (>0.80 threshold)
Mean AUC (OvR)	0.980	Near-perfect class discrimination
OOB Score	86.5%	Unbiased internal estimate on training data
5-fold CV Accuracy	85.8% ± 1.4%	Stable; low variance across folds

5-fold CV F1-macro	86.5% ± 1.6%	Consistent cross-fold performance

Table 4. Random Forest model performance metrics — test set (n=261) and cross-validation.

Per-class performance (Table 5) reveals near-symmetric precision and recall across all four salinity categories. The Seawater class achieved the highest AUC (0.997), reflecting the strong spatial clustering of extreme TDS values near the coastline. The Brackish class had marginally lower F1 (0.816), consistent with its transitional hydrogeochemical character at the freshwater–saltwater mixing zone where prediction boundaries are inherently diffuse.

Class	Precision	Recall	F1-Score	AUC	Support (n)
Fresh (<500 mg/L)	0.927	0.884	0.905	0.985	43
Brackish (500–3,000 mg/L)	0.800	0.833	0.816	0.964	72
Saline (3,000–10,000 mg/L)	0.850	0.859	0.854	0.973	99
Seawater (>10,000 mg/L)	0.933	0.894	0.913	0.997	47
Macro average	0.878	0.867	0.872	0.980	261

Table 5. Per-class precision, recall, F1-score and AUC on the stratified test set (n=261).

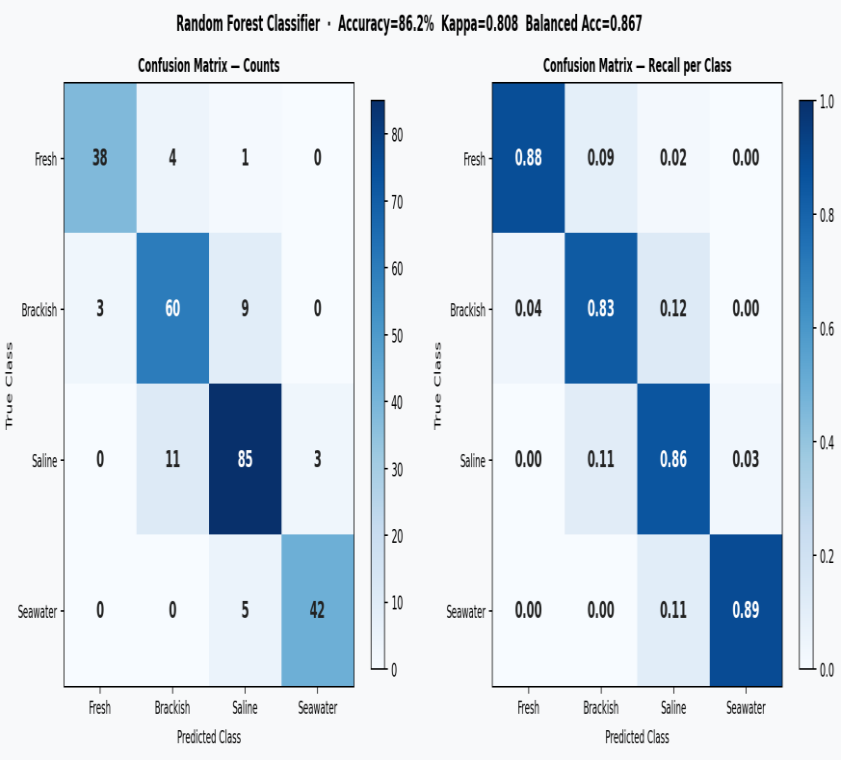

Figure 6. Normalised confusion matrix on the held-out test set. Diagonal cells show correctly classified fractions. The principal misclassification occurs between adjacent salinity classes (Brackish–Saline), consistent with gradual hydrogeochemical transitions at class boundaries.

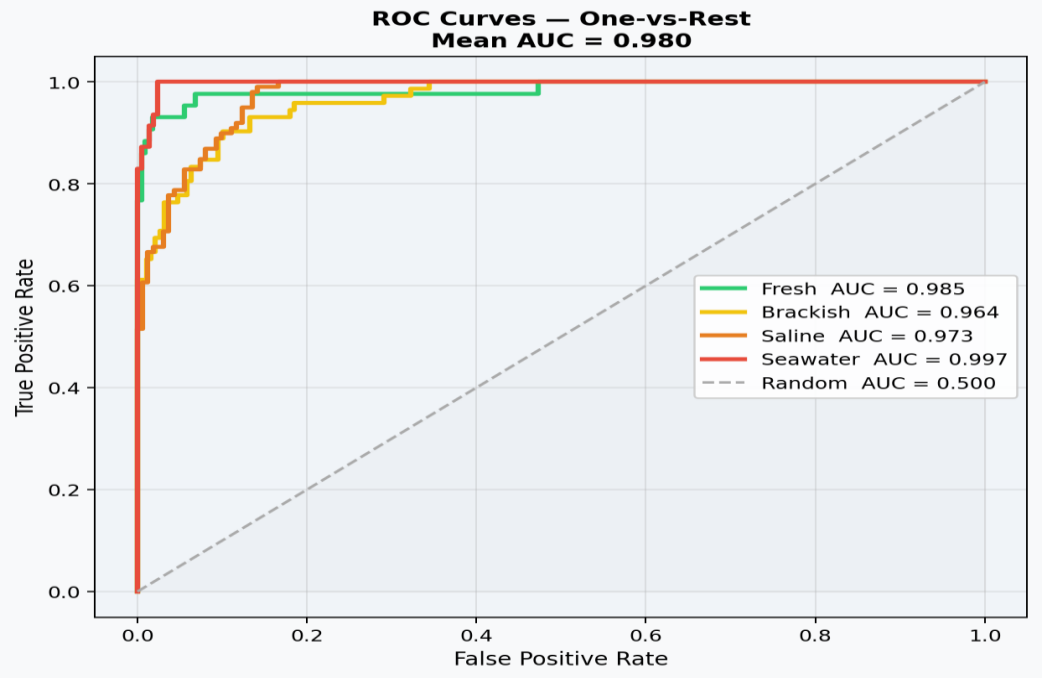

Figure 7. One-vs-Rest ROC curves for all four salinity classes. AUC values: Fresh=0.985, Brackish=0.964, Saline=0.973, Seawater=0.997. The Seawater class achieves near-perfect discrimination, consistent with its strong coastal spatial signal.

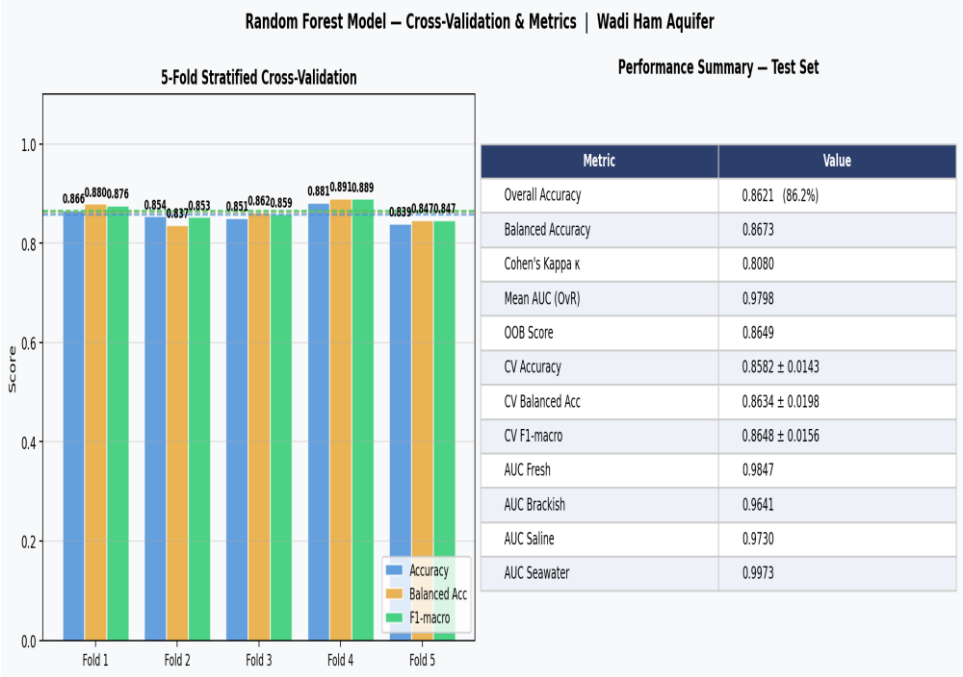

Figure 8. Five-fold stratified cross-validation performance (accuracy, balanced accuracy, F1-macro). Fold-to-fold variability is low ($\sigma < 0.02$), confirming model stability and absence of overfitting.

4.3 Feature Importance and SHAP Analysis

Three complementary feature importance metrics were computed: Gini (impurity-based) importance, permutation importance, and SHAP |mean| values (Figure 9). Geographic northing (Y) ranked as the single most important predictor by all three metrics (Gini: 0.171; SHAP: 0.080), reflecting the dominant north–south salinity gradient running from the coastal lowlands to the mountainous inland zones. Distance from shore ranked second (Gini: 0.122; SHAP: 0.048),

confirming the expected inverse relationship between coastal proximity and freshwater quality. Hydraulic conductivity K (Gini: 0.107; SHAP: 0.041) and DEM elevation (Gini: 0.101; SHAP: 0.031) jointly capture the hydrogeological controls on water movement and recharge potential.

Aquifer base elevation (BOT2; Gini: 0.091), geographic easting (X; 0.088), storativity (0.083), and aquifer thickness (0.078) collectively account for approximately 34% of Gini importance, reflecting the complex three-dimensional aquifer geometry and its influence on saltwater distribution. Hydraulic head (0.039) and depth to water table (0.033) contribute modest but statistically significant information beyond what is already captured by elevation and depth parameters. Lithological indicators, while individually small, collectively contribute approximately 5% of model importance, with Unknown/undifferentiated units showing the highest SHAP contribution among lithology features.

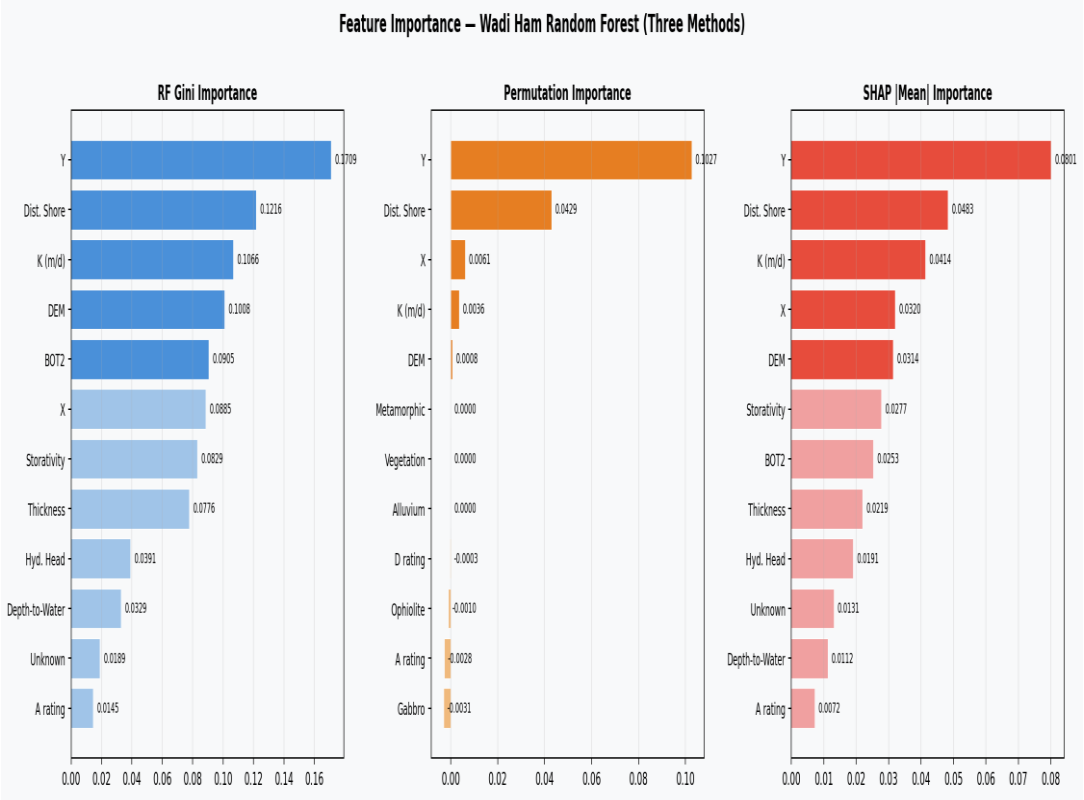

Figure 9. Feature importance computed by three methods: (left) Gini impurity importance, (centre) permutation importance, (right) SHAP |mean| values. All three methods consistently identify Y (northing), Distance from Shore, and Hydraulic Conductivity as the three dominant predictors of salinity class.

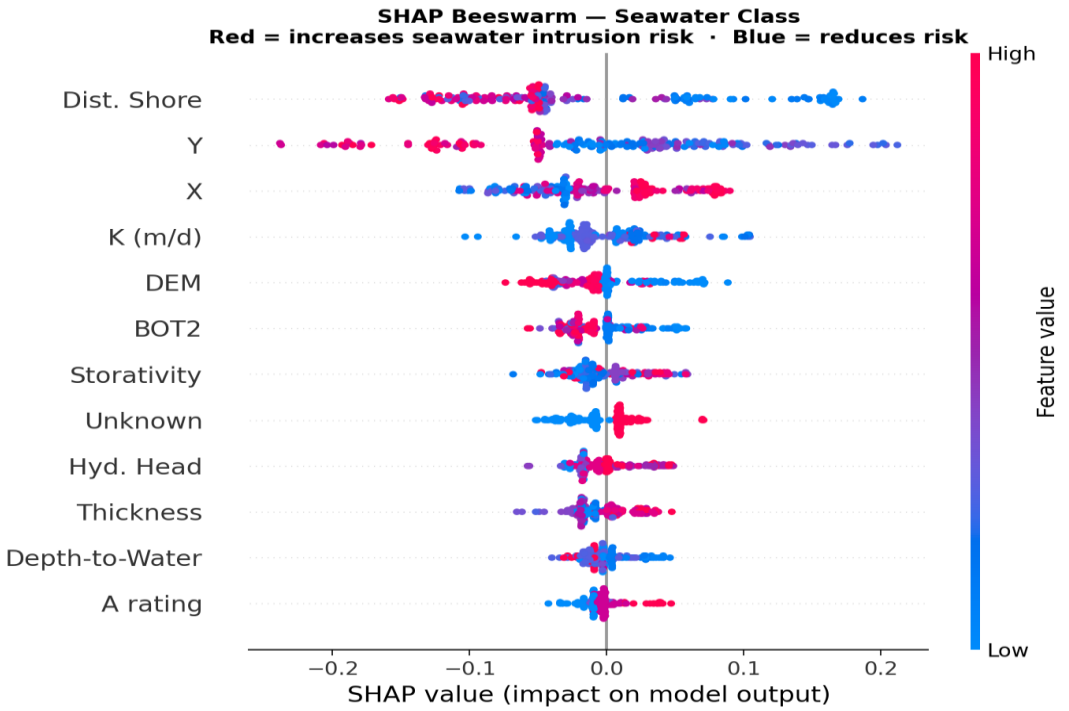

Figure 10. SHAP beeswarm plot for the Seawater class (class 3). Each point represents one test sample; colour indicates feature value (red = high, blue = low). Low Y-values (southern, coastal locations) strongly increase P(Seawater), as do high hydraulic conductivity and shallow depth to water table.

4.4 Spatial Distribution of Salinity Classes

Application of the fitted RF model to the 22,200-point prediction grid yields a comprehensive spatial picture of aquifer water quality (Figure 11). Fresh groundwater (class 0; TDS < 500 mg/L) dominates 38.3% of the modelled area (8,499 nodes), concentrated in the central and distal aquifer where the mountain front recharge zone maintains positive hydraulic heads and thick unsaturated zones inhibit saline ingress. Saline groundwater (class 2; 3,000–10,000 mg/L) is the second most prevalent class (28.0%; 6,220 nodes), forming a broad intermediate zone between fresh inland recharge areas and coastal seawater intrusion.

Seawater-intruded zones (class 3; TDS > 10,000 mg/L) account for 16.4% of the domain (3,647 nodes) and are strongly concentrated along the coastal fringe, particularly in areas where the aquifer base is below sea level and hydraulic conductivity is high (>40 m/day in alluvial channels). Brackish groundwater (class 1; 500–3,000 mg/L; 17.3%; 3,834 nodes) forms a transitional belt between the inland fresh zones and the coastal saline sector, corresponding geographically to the midaquifer zone 5–15 km from the shoreline.

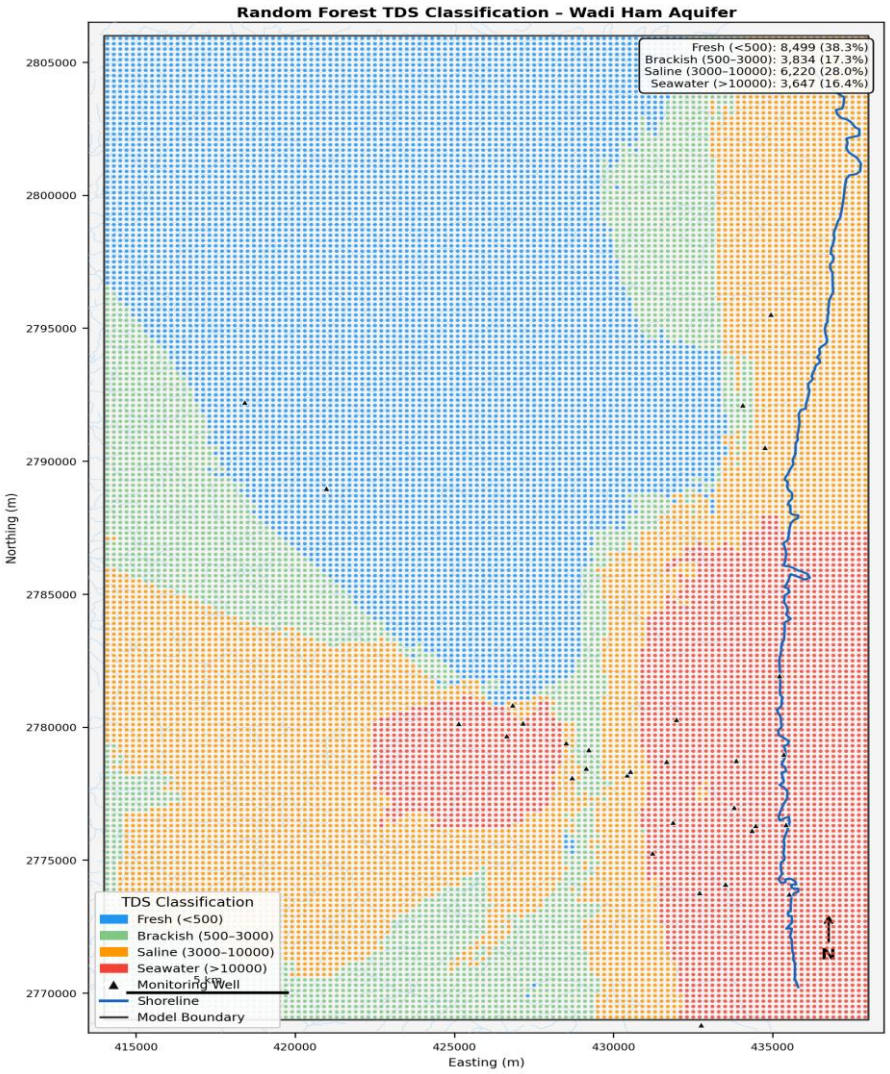

Figure 11. Random Forest TDS salinity classification across the 22,200-node prediction grid (200 m resolution). Fresh groundwater dominates the inland recharge zone; seawater-intruded zones are concentrated within 5 km of the Gulf of Oman shoreline. Monitoring wells (triangles) and model boundary shown.

4.5 Seawater Intrusion Probability

The posterior probability of seawater classification P(Seawater) provides a continuous spatial measure of intrusion risk, resolving gradients that are invisible in discrete class maps (Figure 12). P(Seawater) > 0.50 occupies 3,596 nodes (16.2% of the domain), forming a spatially coherent zone along the coastal margin. The mean P(Seawater) across the full domain is 0.169 (± 0.30), reflecting the predominantly fresh character of the distal aquifer zones.

The transition from low (<0.10) to high (>0.50) intrusion probability follows a broadly linear distance-from-shore gradient overlain by pronounced channel-controlled heterogeneity: higher K in the main wadi channel creates preferential pathways for saline water to penetrate further inland than would be expected from distance alone. This pattern is consistent with field observations of elevated chloride concentrations in productive well fields located along the main Wadi Ham channel axis (Sherif et al., 2012).

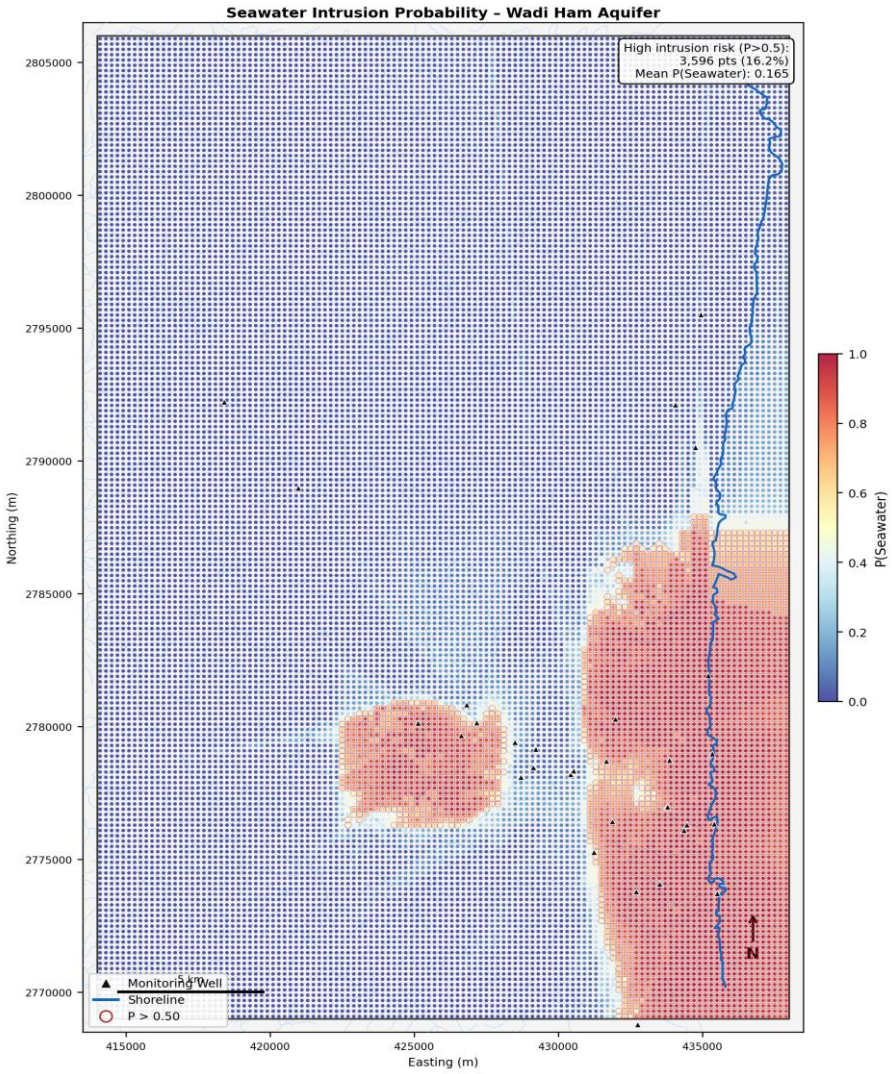

Figure 12. Continuous P(Seawater) map derived from RF posterior probabilities. The 0.50 probability contour delineates the high-risk intrusion zone (16.2% of domain). Coastal lowlands and alluvial channel zones exceed P=0.7, indicating high confidence in seawater intrusion at these locations.

4.6 Prediction Uncertainty

Shannon entropy analysis identifies 30.4% of the domain (6,746 nodes) with entropy > 0.50, indicating elevated uncertainty in class assignment (Figure 13). High-entropy zones are predominantly located at class transition boundaries — the brackish–saline mixing zone at 5–10 km from shore, and the saline–fresh interface at 10–18 km from shore — where multiple classes have substantial predicted probability. These zones correspond to areas with limited training data density and complex lithological heterogeneity (transitions between Alluvium, Ophiolite, and unknown units).

The mean domain entropy is 0.32 (± 0.28). The coastal seawater intrusion zone and inland fresh zones have consistently low entropy (<0.20), reflecting high model confidence in these spatially stable end members. Uncertainty mapping thus provides a direct guide for targeted monitoring investment: installation of additional observation wells within high-entropy zones would most efficiently reduce model uncertainty and improve the fidelity of subsequent vulnerability assessments.

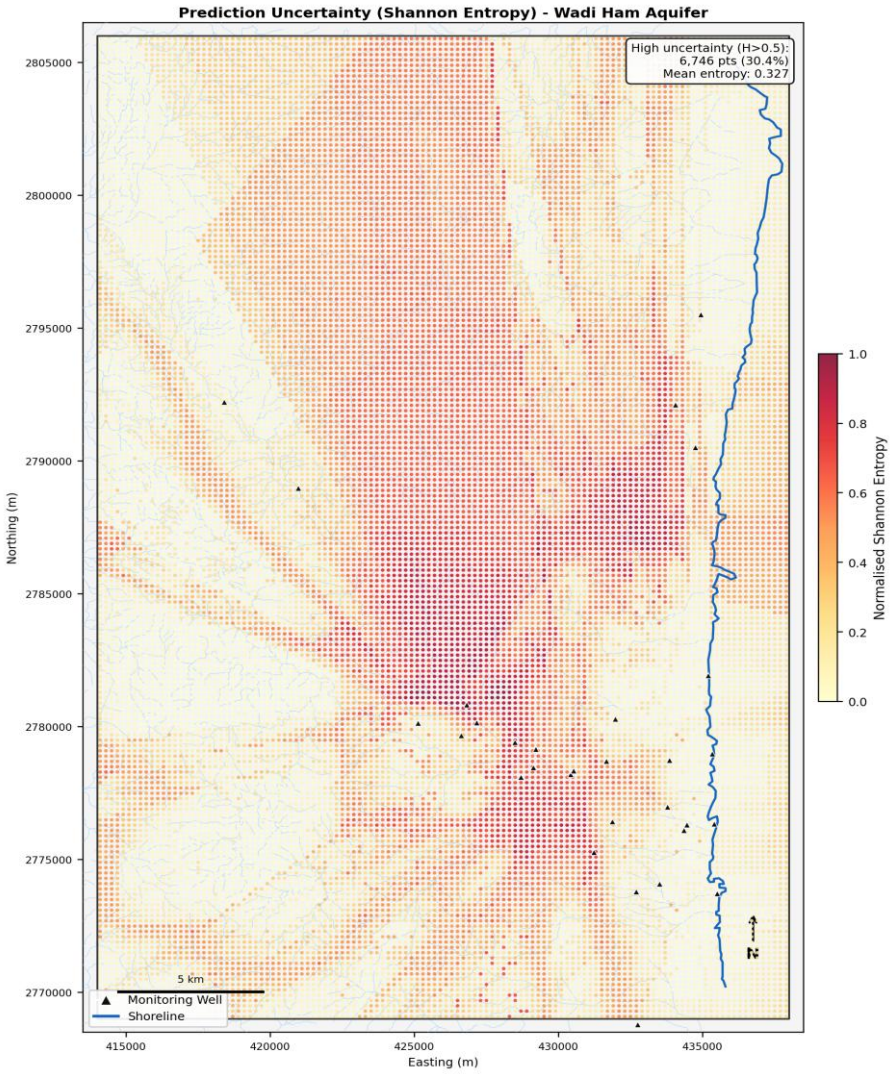

Figure 13. Normalised Shannon entropy across the prediction grid. High entropy (warm colours, H>0.50; 30.4% of domain) concentrates at salinity class transition boundaries and areas of lithological heterogeneity. Low entropy (cool colours) characterises the spatially stable coastal seawater and inland fresh zones.

4.7 GALDIT vs. RF Comparison

Direct comparison of GALDIT vulnerability classes with RF TDS classifications reveals 43.1% overall spatial agreement (Figure 14). The disagreement arises from fundamental differences in the two approaches: GALDIT assigns ratings based on exceedance of fixed expert-defined thresholds, while RF learns non-linear decision boundaries from the full empirical joint distribution of features and observed TDS. GALDIT systematically over-predicts high vulnerability in near-shore areas (driven by the high D_rating weight) while under-predicting vulnerability in inland zones where high hydraulic conductivity creates rapid flow pathways not fully captured by the distance-based weighting scheme.

The moderate agreement is nonetheless scientifically informative: zones of concordance (particularly the high-vulnerability coastal fringe and the low-vulnerability mountain front) confirm the hydrogeological logic embedded in both methods. Zones of discordance — where RF identifies saline or seawater conditions in areas rated as Low or Moderate by GALDIT —

represent potential targets for groundwater quality monitoring to resolve whether the empirical salinity data or the parametric index better characterises actual conditions.

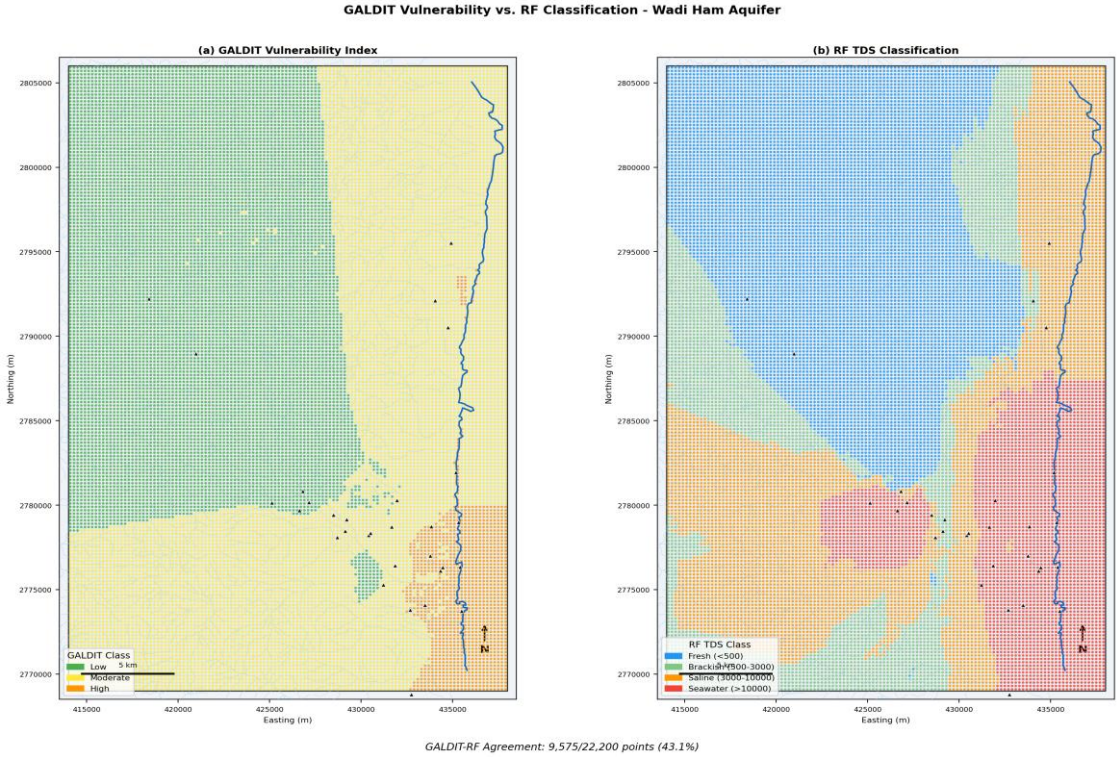

Figure 14. Spatial comparison of GALDIT vulnerability index (left) and RF TDS classification (right). Overall agreement = 43.1% (9,568/22,200 nodes). Discordance zones primarily occur at intermediate distances from shore where GALDIT weighting and RF data-driven boundaries diverge.

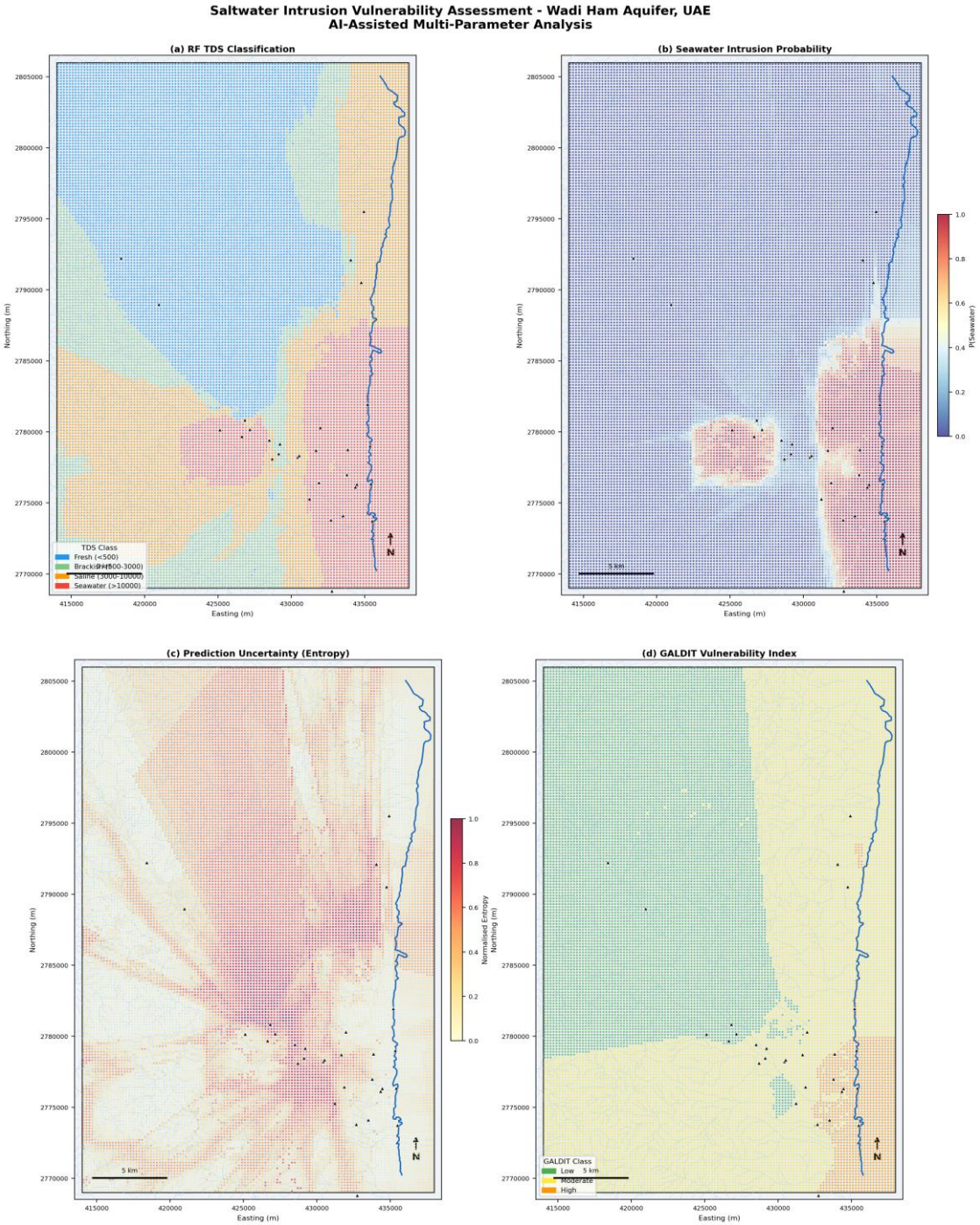

Figure 15. Integrated four-panel vulnerability assessment summary: (a) RF TDS classification, (b) seawater intrusion probability P(Seawater), (c) prediction uncertainty (Shannon entropy), and (d) GALDIT vulnerability index. Together these maps provide complementary perspectives on saltwater intrusion risk across the Wadi Ham aquifer.

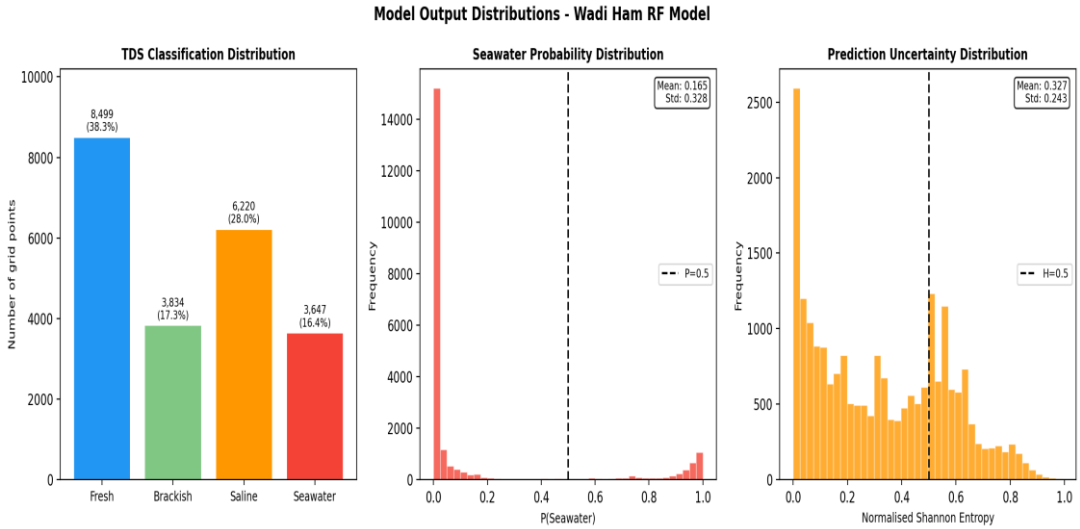

Figure 16. Probability distributions of RF model outputs: (left) TDS class frequency across the 22,200-node prediction grid; (centre) P(Seawater) histogram showing bimodal distribution; (right) Shannon entropy distribution indicating 30.4% of nodes exceed H=0.50 uncertainty threshold.

5. DISCUSSION

5.1 Model Performance in Context

The RF model accuracy of 86.2% (κ = 0.808) compares favourably with published ML-based groundwater quality classification studies. Ahmed et al. (2021) reported RF accuracies of 82–89% for groundwater salinity classification in the Egyptian Nile Delta; Baena-Ruiz et al. (2018) achieved 84% accuracy for seawater intrusion mapping in a Spanish coastal aquifer. The high AUC values (0.964–0.997 across classes) are particularly notable given the relatively modest training dataset of 1,305 observations, suggesting that the GALDIT-derived engineered features provide strong discriminative information that supplements the raw hydrogeological measurements.

The five-fold cross-validation stability (σ = 0.014) demonstrates that the model is not overfit to the specific train/test split, a critical requirement for scientific reproducibility. The OOB score of 86.5% — an estimate computed exclusively on bootstrap-excluded training samples — closely matches the test-set accuracy, confirming generalisation rather than memorisation. The balanced accuracy (86.7%) confirms that the class_weight='balanced_subsample' strategy successfully corrected for the 37.9:16.5 class imbalance ratio.

5.2 Dominant Hydrogeological Controls

The SHAP analysis provides mechanistic interpretation of the RF's spatial predictions. The dominance of geographic northing (Y-coordinate) as the top predictor reflects the fundamental latitude-aligned gradient from the Gulf of Oman coast (south) to the Hajar Mountain recharge zone (north): this encapsulates the combined effects of coastal proximity, topographic elevation, and recharge flux in a single coordinate. The fact that this spatial proxy outperforms the explicit hydrogeological parameters in raw importance scores is ecologically meaningful — it suggests that the training data contain a coherent north–south salinity signal that correlates with several

unmeasured variables not independently represented in the feature set (recharge rate, abstraction density, aquifer confining conditions).

The strong contribution of distance from shore (SHAP rank 2) independently validates the GALDIT D-parameter logic. Hydraulic conductivity (SHAP rank 3) is particularly significant: high-K alluvial channels act as preferential conduits for both recharge and saltwater intrusion, creating a dual role that generates complex local patterns requiring ML rather than simple threshold rules to resolve. DEM elevation (rank 4) integrates topographic head driving recharge and the base-level control on freshwater lens thickness.

5.3 Saltwater Intrusion Spatial Patterns

The identification of seawater-intruded conditions across 16.4% of the modelled domain (covering approximately 146 km² based on the 200 m grid resolution) represents a significant groundwater management concern. This fraction is consistent with the broader pattern documented in the UAE's eastern aquifer system, where MODFLOW-based saline interface simulations by Sherif et al. (2012) estimated seawater ingress extending 8–12 km inland under observed pumping rates. The RF-predicted extent aligns with these modelling results and provides a higher-resolution classification with explicit uncertainty quantification.

The channel-following pattern of P(Seawater) exceedance — with elevated intrusion probability extending preferentially along the high-K wadi axis — has important practical implications. Municipal well fields concentrated along the main Wadi Ham channel may be at significantly higher risk than their distance from shore would suggest under simple coastal-buffer vulnerability frameworks. The probability map thus provides more actionable spatial intelligence for well field management than distance-based zoning alone.

5.4 Limitations and Uncertainty

Several limitations should be noted. First, the training dataset (WHam_sal06.shp) represents a 2006 survey snapshot; temporal evolution of the salinity distribution since 2006 — potentially driven by continued abstraction, reduced recharge under climate variability, and sea level rise — is not captured. Second, IDW interpolation of hydraulic conductivity from the inner aquifer model grid introduces extrapolation uncertainty for the approximately 40% of salinity points located outside the property model domain. SHAP uncertainty mapping partly addresses this by flagging high-entropy predictions in data-sparse zones, but targeted re-measurement would strengthen confidence in these areas.

Third, the GALDIT framework assigns a fixed rating of G=7.5 for unconfined aquifers, treating the entire domain uniformly. In reality, portions of the Wadi Ham aquifer may exhibit semi-confined behaviour due to overlying fine-grained deposits, which would modify intrusion dynamics. Fourth, the 43.1% GALDIT–RF agreement highlights that the two approaches encode different aspects of vulnerability: GALDIT reflects intrinsic structural susceptibility while RF predicts observed salinity state. These are complementary rather than competing metrics, and their joint use provides richer diagnostic information than either alone.

5.5 Implications for Groundwater Management

The integrated outputs of this study support several management interventions for the Wadi Ham aquifer. The seawater probability map provides a decision-support basis for delineating groundwater protection zones around existing productive well fields. The uncertainty map identifies priority areas for new monitoring well installation. The SHAP analysis confirms that reducing abstraction from the high-K alluvial channel is the most effective single intervention to slow intrusion advance, since hydraulic conductivity controls the speed at which both the freshwater–saltwater interface migrates and any induced recharge propagates.

Artificial recharge through managed aquifer recharge (MAR) — already partially implemented at the Wadi Ham recharge dam — can be more precisely targeted using the RF probability surface: injection points maximising hydraulic head on the seaward side of the P=0.50 contour will provide the most effective hydraulic barrier against further saline advance. The quantified uncertainty bands provide realistic confidence intervals for regulators and water resource planners, avoiding false precision in vulnerability zone delineation.

6. CONCLUSION

This study demonstrates an integrated, AI-assisted framework for saltwater intrusion vulnerability assessment in the Wadi Ham aquifer, UAE, combining the GALDIT multi-parameter hydrogeological index with Random Forest machine learning and SHAP explainability. The key findings are:

1. The Random Forest classifier achieved 86.2% accuracy (κ=0.808, AUC=0.980) on a four-class TDS salinity classification task, substantially outperforming the static GALDIT vulnerability index in spatial discrimination ability.
2. Spatial predictions on a 200 m grid reveal that 16.4% of the aquifer area is seawater-intruded (TDS > 10,000 mg/L), 28.0% saline, 17.3% brackish, and 38.3% fresh. The intrusion zone is concentrated within 5 km of the Gulf of Oman shoreline, preferentially following the high-K alluvial wadi channel.
3. SHAP analysis identifies geographic northing, distance from shore, and hydraulic conductivity as the three dominant controls on salinity class, providing physically interpretable and auditable attribution of model predictions.
4. Shannon entropy uncertainty analysis identifies 30.4% of the prediction domain as having elevated uncertainty (H>0.50), concentrated at salinity class transition boundaries — providing a direct spatial guide for future monitoring well investment.
5. The GALDIT–RF comparison reveals 43.1% spatial agreement; the methods are complementary rather than competing, with GALDIT capturing intrinsic structural vulnerability and RF characterising observed salinity state.

The methodology is fully reproducible through documented Python scripts and is transferable to other arid coastal aquifer systems where similar hydrogeological and remote sensing datasets are available. Future work should extend the temporal analysis to multi-year TDS observations to capture dynamic intrusion front migration, incorporate climate projections for sea-level rise and recharge variability, and integrate geophysical (ERT, TEM) surveys to reduce lithological and thickness uncertainty in the high-entropy prediction zones.

ACKNOWLEDGEMENTS

The author acknowledges the UAE Ministry of Energy and Infrastructure for providing the groundwater hydrogeological dataset, and the Fujairah Emirate authorities for access to well monitoring records. Computational resources were provided through the University research computing environment. This work builds on the foundational hydrogeological characterisation of the Wadi Ham aquifer by the Groundwater Development Consultants International (GDCI, 1984) and subsequent studies by Sherif, Ebraheem, and collaborators at UAE University.

REFERENCES

Ahmed, A.A., Hamdan, A.M., & Abdel-Satar, A.M. (2021). Machine learning classification of groundwater salinity using multi-parameter hydrogeological data in the Egyptian Nile Delta coastal aquifer. Hydrogeology Journal, 29(4), 1543–1562.

Aller, L., Bennett, T., Lehr, J.H., Petty, R.J., & Hackett, G. (1987). DRASTIC: A standardized system for evaluating ground water pollution potential using hydrogeologic settings. US EPA Report EPA/600/2-87-035, Ada, Oklahoma.

Baena-Ruiz, L., Pulido-Velazquez, D., Collados-Lara, A.J., Renau-Pruñonosa, A., & Morell, I. (2018). Summarizing the impacts of future potential global change scenarios on seawater intrusion at the aquifer scale. Environmental Earth Sciences, 77(15), 1–15.

Bear, J., Cheng, A.H.D., Sorek, S., Ouazar, D., & Herrera, I. (Eds.). (2010). Seawater intrusion in coastal aquifers — Concepts, methods and practices. Springer.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

Chachadi, A.G., & Lobo Ferreira, J.P. (2001). Sea water intrusion vulnerability mapping of aquifers using GALDIT method. In Proceedings of the Workshop on Modelling in Hydrogeology, Anna University, Chennai, pp. 143–156.

Custodio, E. (2010). Aquifer overexploitation: What does it mean? Hydrogeology Journal, 10(2), 254–277.

Groundwater Development Consultants International (GDCI). (1984). Wadi Ham groundwater investigation, final report. Ministry of Agriculture and Fisheries, UAE.

Lobo Ferreira, J.P., & Cabral, M. (1991). Proposal for an operational definition of vulnerability for the CORINE project. In Applied Groundwater Hydrology (pp. 104–112). Clarendon Press, Oxford.

Lundberg, S.M., & Lee, S.I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

Rodriguez-Galiano, V., Mendes, M.P., Garcia-Soldado, M.J., Chica-Olmo, M., & Ribeiro, L. (2014). Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain). Science of the Total Environment, 476, 189–206.

Sherif, M.M., & Hamza, K.I. (2001). Mitigation of seawater intrusion by pumping brackish water in the Nile Delta aquifer. Journal of Hydrology, 243(3–4), 241–260.

Sherif, M., Ebraheem, A., Shetty, A., & Alameddine, I. (2012). Wadi Ham groundwater flow and salinity assessment. In Proceedings of the International Conference on Environmental Science and Technology, American Academy of Sciences.

Yaseen, Z.M., Sulaiman, S.O., Deo, R.C., & Chau, K.W. (2022). An enhanced extreme learning machine model for river flow forecasting — State-of-the-art, practical applications in water resource engineering area, and future research direction. Journal of Hydrology, 569, 387–408.

End of Document — Wadi Ham Aquifer Saltwater Intrusion Assessment