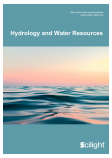
Hydrology and Water Resources
https://www.sciltp.com/journals/hwr



Article

AI-Driven Analysis of Saltwater Intrusion Vulnerability

Vahid Nourani *, Elnaz Bayat Khajeh, Sana Maleki, Nardin Jabbarian Paknezhad and Elnaz Sharghi

Center of Excellence in Hydroinformatics, Faculty of Civil Engineering, University of Tabriz, Tabriz 5166-61647, Iran
* Correspondence: nourani@tabrizu.ac.ir

How To Cite: Nourani, V.; Khajeh, E.B.; Maleki, S.; et al. AI-Driven Analysis of Saltwater Intrusion Vulnerability. *Hydrology and Water Resources* **2026**, *1*(1), 4. https://doi.org/10.53941/hwr.2026.100004

Received: 4 August 2025
Revised: 30 December 2025
Accepted: 31 December 2025
Published: 9 January 2026

Abstract: Evaluating groundwater salinity is crucial, particularly in arid and semi-arid regions where access to fresh water is essential for various needs. In this study, the GALDIT method, which uses expert judgment to score six key parameters to assess vulnerability to seawater intrusion (SWI), was utilized to evaluate the susceptibility of the Salmas and Urmia aquifers. However, depending on expert judgment may lead to subjectivity and potential bias in assessing vulnerability. To address these limitations, different artificial intelligence-based models were applied to enhance model performance. In the Urmia aquifer, the vulnerability index obtained by GALDIT varied from 3.8 to 6.8, and in the Salmas aquifer, it ranged from 4.3 to 7.8. The difference in the GALDIT vulnerability index ranges primarily reflects distinct hydrogeological conditions in each area. A qualitative comparison between the GALDIT index maps and the observed Electrical Conductivity (EC) distribution highlighted notable spatial mismatches, suggesting that the GALDIT index alone may not fully capture the spatial variability of salinity levels. To assess vulnerability using AI-based models, parameters of GALDIT index were used as input, while observed EC values, as effective salinity indicators, were used as the model outputs. The results showed that all AI-based models outperformed the conventional, expert-based GALDIT index in predicting aquifer vulnerability. While the GALDIT model demonstrated limited correspondence with observed salinity patterns, the AI-based models—especially BO-ELM and Integrated Ensemble—provided significantly improved predictions. These models achieved the highest accuracy in both Urmia and Salmas Plains, with DC values approaching 0.99–1.0 and substantially reduced RMSE, effectively capturing the complex spatial variability of salinity in the aquifers. Although the proposed framework is demonstrated using the Urmia and Salmas aquifers, the methodology is transferable and can be applied to other coastal aquifers with similar hydrogeological settings through site-specific recalibration, provided sufficient data are available.

Keywords: seawater intrusion; vulnerability assessment; GALDIT; AI-based models; coastal aquifers

1. Introduction

Groundwater resources stand as a prevalent source of fresh water, and the gradual rise in the salt content of many primary aquifers utilized for water supply in coastal zones, particularly in arid or semi-arid regions, presents clear indications of declining water quality [1,2]. So, groundwater vulnerability assessment is widely acknowledged as a key method for evaluating aquifer susceptibility to contamination and informing effective groundwater management strategies [3]. The concept of vulnerability is based on the understanding that some

areas of an aquifer system are naturally less shielded by geological materials and are therefore more prone to contamination [4,5].

In essence, the process of groundwater vulnerability assessments involves simplifying complex hydrogeological data into a user-friendly map. This map can then be readily utilized by decision-makers, policy developers, environmental stewards, and the public [6]. Overall, vulnerability assessment techniques can be divided into three primary categories: overlay or index-based approaches, statistical methods, and process-based or numerical methodologies [7]. Overlay or index-based methods are widely used due to their relative simplicity and ease of implementation, especially in data-scarce regions [8,9]. Several widely applied methodologies are used to delineate zones vulnerable to groundwater contamination, including DRASTIC, GALDIT, SINTACS, and the Aquifer Vulnerability Index (AVI) [10–13]. Among the numerous indexing techniques, GALDIT adopts an approach for assessing coastal aquifer vulnerability [14]. Utilizing numerical calculations, this method evaluates the likelihood of Seawater Intrusion (SWI) occurrence using six parameters of Groundwater occurrence (G), Aquifer hydraulic conductivity (A), Height of groundwater above sea Level (L), Distance from the shore (D), Impact of existing SWI status (I) and aquifer Thickness (T) [15]. GALDIT has been widely adopted by many researchers around the world for performing vulnerability assessments in various regions [16–21]. In many studies, modified versions of GALDIT have undergone validation using SWI indicators and statistical techniques. These validations have mostly resulted in improvements to the traditional GALDIT framework, making it more effective and dependable for evaluating and forecasting the occurrence of SWI [22]. For instance, a novel AI-based modeling strategy was developed to enhance the GALDIT method using a two-level learning process involving Artificial Neural Network (ANN), Sugeno Fuzzy Logic (SFL), Neuro-Fuzzy (NF), and Support Vector Machine (SVM) models [23]. The results showed that the methodology enhances SWI vulnerability simulation and demonstrates improved accuracy, as evaluated using statistical metrics such as Root Mean Square Error (RMSE) and the coefficient of determination (R^2). Resampling techniques along with machine learning models were applied to enhance the GALDIT framework for groundwater vulnerability mapping, which helped improve prediction accuracy by tackling the problems of subjectivity and limited data sets [24]. A study integrating Adaptive Neuro-Fuzzy Inference System (ANFIS), SVM, and ANN developed a Supervised Committee Machine Artificial Intelligence (SCMAI) model to predict groundwater vulnerability to SWI in a coastal aquifer [25]. For improving GALDIT method, a study used set of boosting and tree-based algorithms and integrated them with two resampling techniques, Bootstrap Aggregating and Disjoint Aggregating. This study demonstrated that combining ML algorithms with resampling methods can markedly enhance the predictive performance of GALDIT, underscoring the importance of integrated and hybrid approaches in developing more reliable groundwater vulnerability assessments [24]. In another study, Bayesian Model Averaging (BMA) has applied to address the subjectivity of index-based model. By combining prediction from multiple first-level AI models into a probabilistic ensemble, BMA quantified both within-model and between-model variability. The results showed that Bayesian integration improves the stability and reliability of vulnerability outputs [26]. A Z-number adaptation of Fuzzy Logic (FL) and conventional FL addressed the complexities and uncertainties in examining groundwater susceptibility to pollution, surpassing the conventional GALDIT method and leading to improved accuracy in vulnerability predictions [21].

In this research, the GALDIT approach has been employed to evaluate the vulnerability of the Urmia and Salmas plains, considering their coastal aquifer nature where SWI is a critical concern. While GALDIT provides valuable insights, its reliance on expert judgment introduces subjectivity and potential biases in the weighting and rating process. To enhance the precision of vulnerability assessments, AI-based approaches are employed to enable dynamic modeling that can accommodate and adapt to evolving environmental conditions, thereby improving predictive resilience [27]. Unlike GALDIT and other unsupervised index-based methods, introducing a target variable enables supervision and potential improvement of the method. AI-based models, renowned for their self-learning and self-adaptive capabilities, are utilized in modeling nonlinear hydrological processes owing to their black box nature, allowing them to adapt to changing conditions [28,29].

In this study, the vulnerability assessment of the coastal aquifers in northwest of Iran in the Urmia Lake catchment area (Urmia and Salmas) to saltwater intrusion was initially conducted using the GALDIT method in ArcMap software to prepare vulnerability and parameter zoning maps. Although the GALDIT model is one of the most widely used models for assessing the vulnerability of coastal aquifers, as mentioned before, its unsupervised nature and reliance on expert opinion for parameter ranking and weighting can introduce bias into the vulnerability assessment. To improve modeling performance and achieve more reliable results of vulnerability, artificial intelligence-based models (ANFIS, Support Vector Regression (SVR), Ensemble model integrating FFNN, SVR and the Extreme Learning Machine (ELM) integrated with Bayesian Optimization (BO), (BO-ELM)) were utilized in this study. To this end, the parameters of GALDIT are used as inputs. The G parameter in the GALDIT method

represents “Groundwater occurrence”, and describes the aquifer type (confined, unconfined, or semi-confined) and its inherent level of natural protection against SWI [11,30]. In this study, G was excluded from the modeling because it shows no spatial or temporal variability across the region. Electric Conductivity (EC) is selected as the target variable because of its effectiveness as an indicator of SWI in coastal aquifers. However, to assess aquifer susceptibility across regions and conditions, it is feasible to include additional variables as model outputs and targets. This flexibility further supports the applicability of the proposed framework to other regions, subject to the availability of suitable local data.

This study is among the first to explore the use of several AI-based models within the GALDIT framework for SWI vulnerability assessment. First DFFNN model was used to examine whether increasing the depth of feedforward architectures leads the model to learn more complex pattern within the GALDIT parameters and better reproduce observed salinity. Second, an ensemble model was developed by combination of three different algorithms (FFNN, ANFIS, SVR) to leverage their complementary strength. Third, the ELM was further enhanced through BO, this allows the structure of the model and hyperparameters to be selected automatically rather than trial and error. So, this approach allows us to systematically compare shallow learning, deep learning, integration of shallow learning models and hybrid optimization-learning model that addresses the subjectivity of the conventional GALDIT method. Moreover, this study aims to produce a more reliable depiction of aquifer vulnerability to SWI.

2. Materials and Methods

2.1. Study Area

Understanding groundwater vulnerability in coastal aquifers is vital due to the reliance on groundwater for agricultural and domestic needs in arid and semi-arid regions. The Salmas plain is in the northern part of West Azerbaijan province, within the catchment area of Lake Urmia to the west (See Figure 1a). Encompassing an area between 44°38′31″ E and 45°05′51″ E, and between 38°04′13″ N and 38°20′58″ N, the study area spans approximately 550 km^2 at the heart of the Zoolachay watershed. With elevations ranging from 1221 m to 3211 m above sea level, the average elevation of the Salmas plain stands at 1341 m. The primary river nourishing the Salmas plain is the Zoolachay. The average temperature is 9.8 °C, while the average annual rainfall in the region amounts to 263.8 mm [26]. The Salmas aquifer system is predominantly unconfined, although local semi-confined conditions occur where fine-grained layers are interbedded with coarser sediments. Groundwater recharge primarily originates from infiltration of precipitation over the surrounding uplands, runoff and stream infiltration from rivers such as the Zola and Shapur Rivers, and subsurface inflow from adjacent mountain aquifers. The alluvial deposits that constitute the aquifer, in combination with local rivers, play a crucial role in facilitating recharge across the plain. Groundwater flow within the Salmas Plain generally occurs from west to east, with occasional short local gradients from north to south, ultimately discharging into the northern margin of Lake Urmia. The region is characterized by intensive agricultural activities, with irrigation serving as the principal economic driver [26]. However, excessive groundwater extraction, particularly in the downstream and coastal zones of the plain, has disrupted the balance between freshwater and saline interfaces, leading to the development of SWI and declining groundwater levels [31]. So, it is really important to thoroughly evaluate how vulnerable and sustainable this aquifer system is to manage groundwater properly in the future.

Geologically, the Salmas Plain is characterized by extensive Quaternary alluvial deposits overlying Miocene volcanic and sedimentary units [32], as illustrated in Figure 2, which depicts the distribution of lithological units, faults, and structural features across the plain. The alluvial deposits mainly comprise layers of gravel, sand, silt, and clay with varying thickness, reaching up to 200 m in the central region of the plain. The upper layers (coarse sediments) show high permeability and constitute the primary aquifer, whereas the fine-grained strata (silt and clay) in the lower areas occasionally generate semi-confined conditions. The surrounding highlands are mainly composed of Tertiary volcanic rocks notably andesite, basalt, and tuff, interbedded with marl and limestone formations, which play a significant role in supplying recharge to the alluvial aquifer through runoff infiltration and subsurface flow [32]. Several fault systems crossing the plain (including the Salmas Fault Zone) influence both the hydrogeological behavior and the groundwater flow regime. These faults act as preferential flow paths in some areas and as hydraulic barriers in others, depending on their lithological characteristics and degree of fracturing.

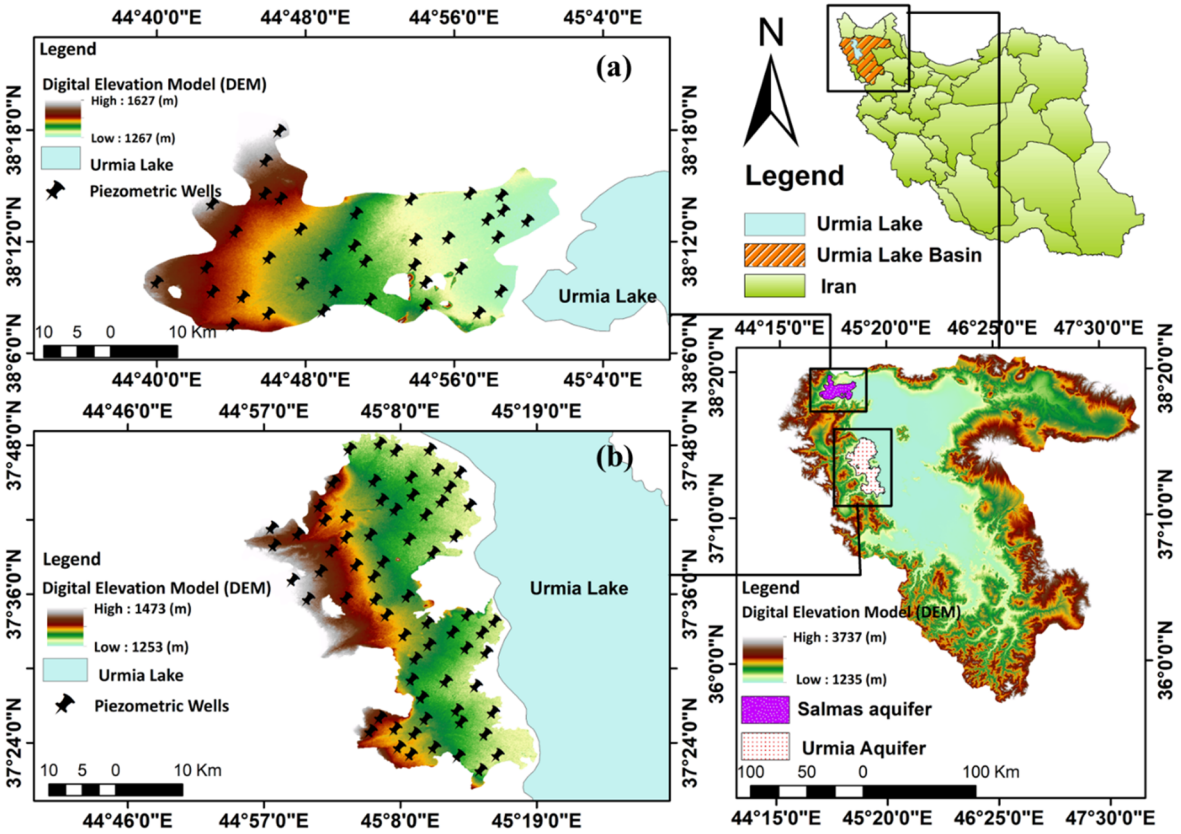


Figure 1. The Digital Elevation Model (DEM) and position of piezometers: (**a**) The Salmas Aquifer, West Azerbaijan, Iran, (**b**) The Urmia Aquifer, West Azerbaijan, Iran.

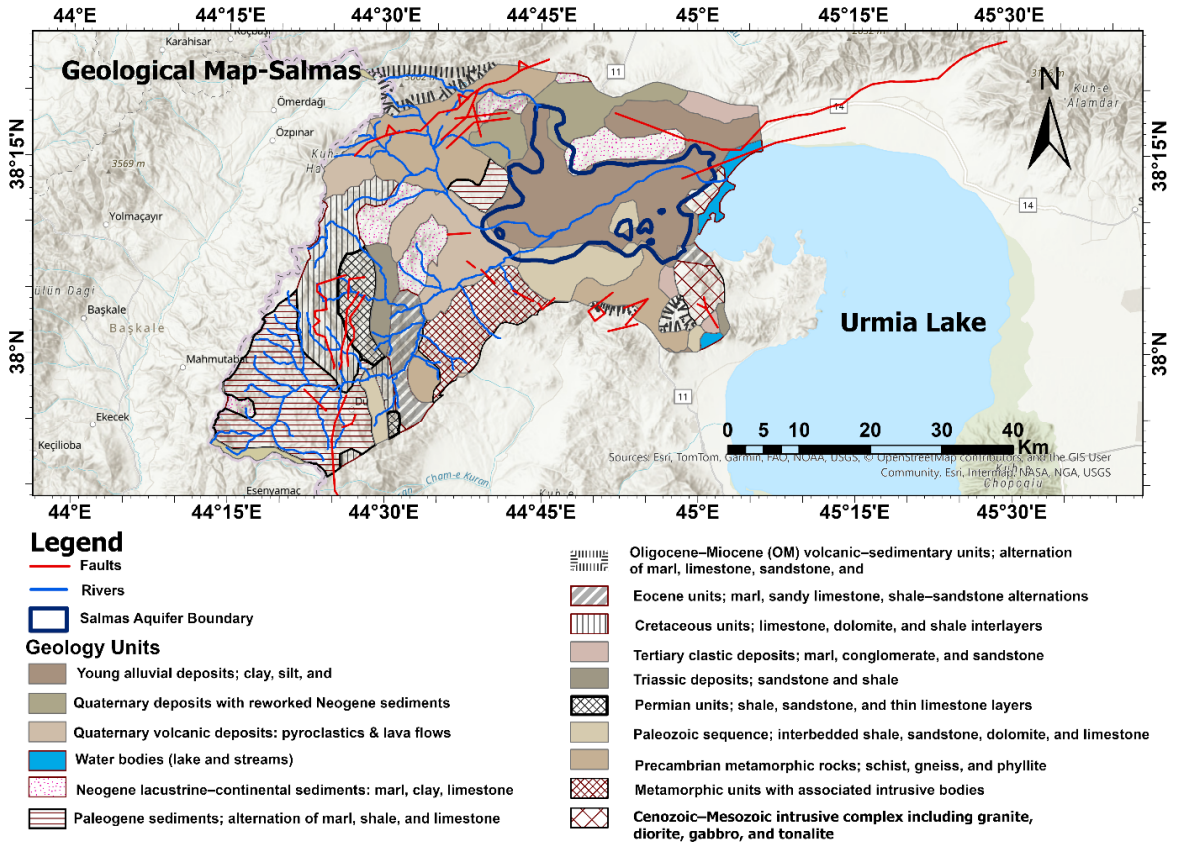


Figure 2. Geological map of Salmas Plain.

The Urmia Plain is situated in the central region of West Azerbaijan Province, to the west of Lake Urmia, Iran. The plain encompasses an unconfined aquifer system extending between 44°49′20″ and 45°19′14″ E longitudes and 37°06′39″ and 37°50′41″ N latitudes, covering an area of approximately 1479 km^2 (see Figure 1b). While the western part of the plain consists of rugged mountainous terrain, the eastern portion gradually declines in elevation toward Lake Urmia, forming a natural groundwater discharge zone. The area experiences a cold semi-arid climate, with an average annual precipitation of about 329 mm and a mean annual temperature of nearly 13 °C. From a geological view, the Urmia Plain is underlain by thick Quaternary alluvial deposits that rest unconformably on Tertiary volcanic and sedimentary formations (see Figure 3 for the geological map of the region). The alluvial sediments—composed primarily of sand, gravel, silt, and clay—reach thicknesses of up to 250 m in the central portions of the plain [33]. The upper coarse-grained units constitute the main unconfined aquifer, while interbedded fine-grained layers locally create semi-confined conditions. The basement complex beneath the alluvial sequence consists mainly of andesitic and basaltic volcanic rocks, tuff, and marl, which crop out in the surrounding highlands and serve as recharge boundaries for the aquifer system.

The hydrogeological structure of the plain is influenced by several fault systems trending northwest–southeast, which control both groundwater flow paths and hydraulic connectivity. Groundwater generally flows from the elevated western highlands toward the east, eventually discharging into Lake Urmia. Recharge to the aquifer occurs through multiple processes, including infiltration of precipitation on the highlands, infiltration of surface runoff from ephemeral and perennial streams (such as the Nazlou and Shahar Rivers), and percolation of excess irrigation water in agricultural zones.

Groundwater is the primary water resource for agricultural, domestic, and industrial use within the Urmia Plain. However, prolonged over-extraction has led to declining water levels and deteriorating water quality, particularly along the eastern and southeastern margins adjacent to Lake Urmia. The EC of groundwater increases eastward, ranging from 700 µS/cm in recharge zones to over 2500 µS/cm near the lake, indicating progressive salinization caused by lake-groundwater interactions and evaporation effects.

Given the high degree of exploitation of the aquifer and the sensitivity of its hydrogeological system to changes in recharge and abstraction, assessing its vulnerability and sustainable management potential is essential for long-term water resource planning.

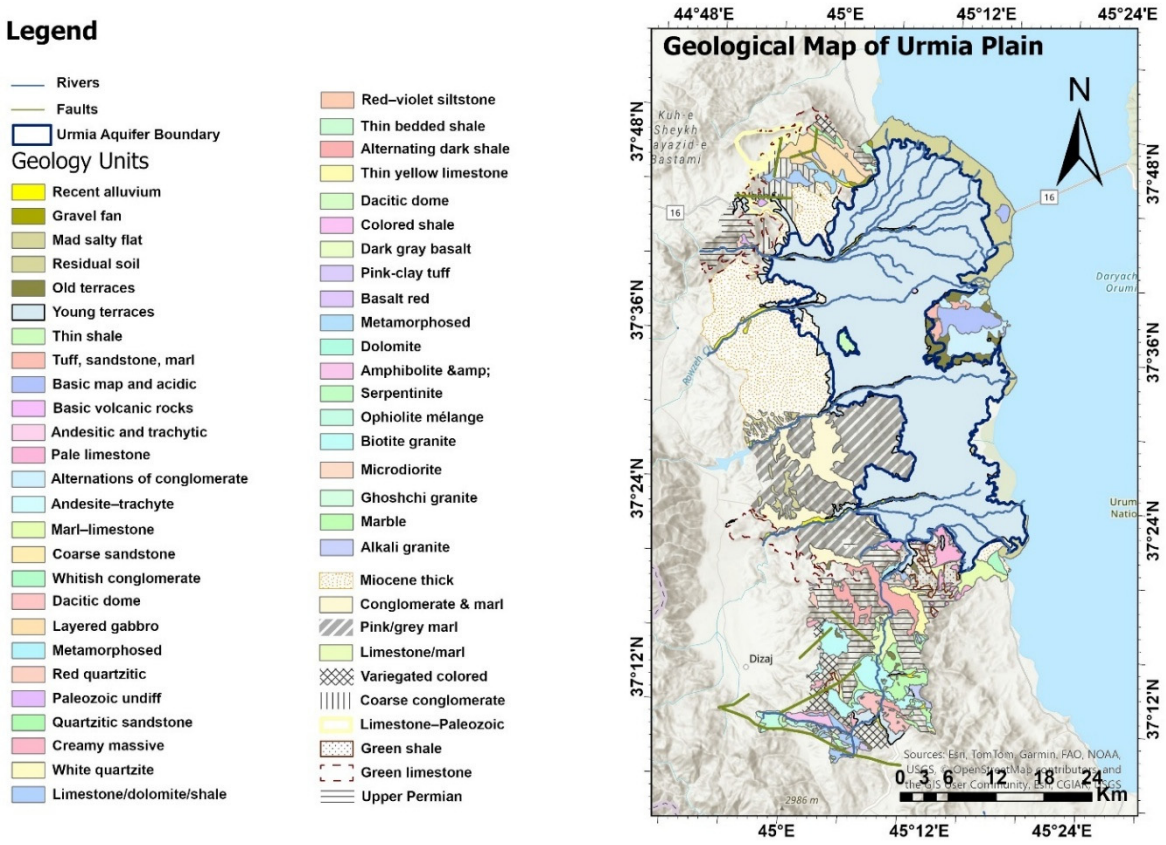


Figure 3. Geological Map of Urmia Plain.

2.2. Data Preparation

All the data used in this study were collected by the West Azerbaijan Water Organization. The dataset contains information about the hydrogeological features and water quality of both the Salmas and Urmia aquifers. For the Urmia aquifer, 65 piezometric wells were used to measure groundwater levels, and water quality samples were collected from 59 observation wells in September 2022 [31]. For the Salmas aquifer, groundwater levels were monitored through 37 piezometric wells, and water quality measurements were collected from 36 sampling wells in the same period.

The water salinity analysis included electrical conductivity (EC), chloride (Cl^-), and bicarbonate (HCO_3^-) concentrations, which were used to derive the GALDIT "Impact of SWI" (I) parameter based on the $Cl^-/(HCO_3^- + CO_3^{2-})$ ratio. Aquifer characteristics, such as aquifer type and thickness, were determined using well logs distributed across the study area. The D parameter represents the perpendicular distance from each location (e.g., observation well point) to the current shoreline of Lake Urmia. This was calculated using the ArcGIS buffer tool, with zones categorized according to the standard GALDIT rating scale [25].

Groundwater level elevations (L) were calculated by integrating piezometric well data with the reference elevation of Lake Urmia. All spatial data layers, including aquifer thickness, groundwater level, and distance from the lake (shoreline), were processed in ArcGIS to construct maps for the GALDIT method. The point data were spatially interpolated using the Inverse Distance Weighting (IDW) method to create continuous surfaces for use as model inputs, following the procedures described in references [9,24].

2.3. Proposed Methodology

This research endeavors to enhance the accuracy of vulnerability assessment by integrating advanced AI-based methodologies into the GALDIT framework. GALDIT, although a robust method, relies heavily on expert knowledge in rating and weighting parameters and is subject to variation across different geographic regions due to differences in aquifer characteristics and area morphology. Recognizing this limitation, AI and ML-based models are proposed to address these challenges. FFNN and DFFNN are advantageous because they are not restricted by theoretical assumptions or residual analysis, allowing for adaptability across various data types. Additionally, FFNNs are widely used in environmental modeling due to their ability to capture complex, nonlinear relationships that are often difficult to represent using traditional statistical regression techniques, such as linear or polynomial regression. While FFNNs, like other machine learning models, generally require high-salinity data, they have shown comparative resilience to moderate noise or missing patterns. However, it is acknowledged that no model—regardless of its complexity—can perform well under persistently poor or insufficient data conditions. Therefore, ensuring adequate data salinity and quantity remains essential for reliable model predictions. Although FFNN is a robust model, DFFNN—with its multiple hidden layers and greater depth—can capture more intricate patterns and interactions compared to a simple FFNN, thereby enhancing the predictive capabilities for assessing SWI.

To further improve neural network learning performance, ANFIS and SVR were incorporated as complementary models. These approaches enhance the ability of framework to maintain strong generalization capacity, capture complex nonlinear behavior, and, importantly, manage data uncertainty, which are inherent characteristics of water quality datasets [34]. Moreover, to complement neural network learning, additional models including ANFIS, SVR were incorporated, each offering unique strengths in handling uncertainty, nonlinear boundaries, and rapid generalization. To integrate these capabilities, an ensemble model combining FFNN, ANFIS, and SVR was developed to leverage the complementary advantages of multiple algorithms, providing a more robust and balanced representation of salinity dynamics. Then, the hybrid BO-ELM model was further adopted to introduce an optimization-driven structure, allowing BO to refine the ELM parameters and improve model stability.

The proposed methodology for this study was divided into four steps (See Figure 4). Firstly, data were prepared and gathered from the West Azerbaijan Water Organization. Secondly, map layers of the GALDIT method were created using GIS for both Salmas and Urmia plains aquifer. Before beginning the third step a data preprocessing workflow was implemented to prepare a sufficient dataset for training AI-based models. Since the number of observed EC samples and GALDIT parameter measurements was limited and not spatially uniform across the study areas, each parameter layer was first interpolated (using inverse distance weighting). To reduce spatial autocorrelation and obtain representative values at a consistent resolution, all interpolated rasters were resampled to a coarser grid, ensuring an equal number of cells across the study domains. After resampling, the Raster-to-Point tool was applied to extract the value of every raster cell into a point feature with an explicit geographic location. These point datasets were exported as attribute tables to Excel and assembled into a structured

database, where each row represents a single spatial sample and each column corresponds to a GALDIT input parameter or the target EC value.

This approach was necessary because machine learning algorithms require a sufficiently large number of inputs–output pairs for effective training. Using only the observed EC measurements, would not provide enough samples to capture the spatial variability of aquifer properties. Therefore, interpolated raster layers were used to generate training datasets, a practice widely adopted in groundwater modeling and data-driven aquifer assessments. Notably, the machine learning inputs used in this study correspond to the original GIS-derived GALDIT parameter values rather than the expert-assigned ratings.

In the third step, AI-based models (FFNN, DFFNN, ANFIS, SVR, integrated ensemble, and BO-ELM) were employed to improve the GALDIT method, addressing the unsupervised nature of index-based methods. Despite the capabilities of FFNN, the need for a deeper understanding of aquifer features was recognized, leading to the utilization of DFFNN. In which, the inputs included GALDIT parameters, excluding the G parameter due to its lack of spatial and temporal variability, which did not contribute to the FFNN and DFFNN learning processes. EC was selected as the output variable, as it serves as a reliable indicator of salinity levels within the aquifer and offers critical insights into potential seawater contamination. Lastly in the fourth step, the AI-based models were evaluated using RMSE and DC to assess their performance in simulating EC values.

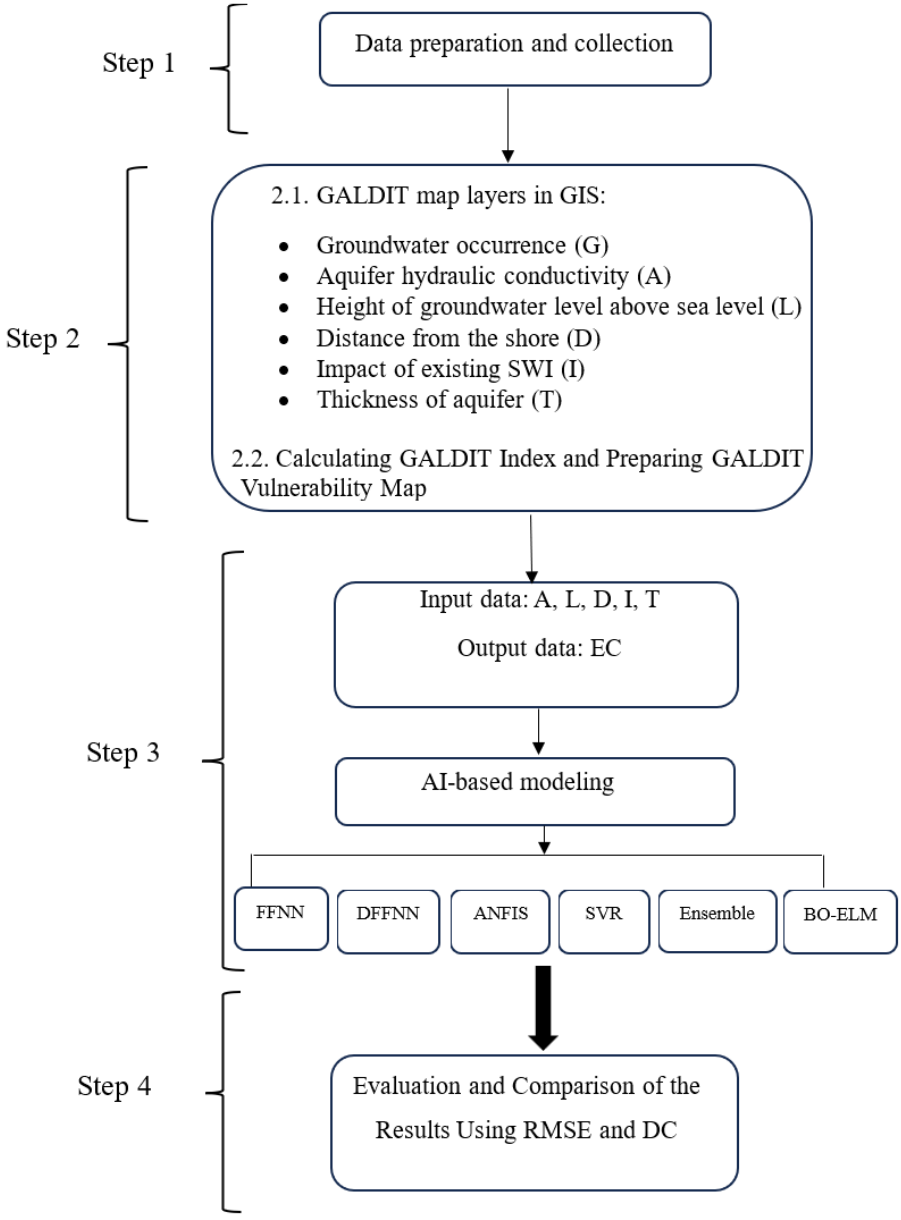


Figure 4. Flowchart of the proposed methodology.

2.4. Model Evaluation Criteria

To evaluate how well the trained networks performed, statistical measures like the Determination Coefficient (DC) and Root Mean Square Error (RMSE) were used. RMSE quantifies the average magnitude of prediction errors, highlighting large deviations, while DC measures how well the model explains the variance in observed data. RMSE values typically range from 0 to infinity, with lower values indicating better performance. Specifically, an RMSE closer to 0 signifies high predictive accuracy. On the other hand, DC values range from 0 to 1, with higher values meaning a better fit. Using these metrics ensures a comprehensive evaluation, that balances prediction accuracy with overall model reliability. The statistical criteria mentioned earlier are often used to assess how well models perform, offering a strong foundation that is calculated through the formulas listed below. The correlation coefficient (CC), which measures the strength and direction of the linear relationship between observed and predicted values, was also used to evaluate model performance (Equation (3)).

$$\mathrm{RMSE} = \sqrt{\frac{\sum(X - \mathrm{C})^2}{\mathrm{N}}} \tag{1}$$

$$\mathrm{DC} = 1 - \frac{\sum(X - \mathrm{C})^2}{\sum(X - \bar{X})^2} \tag{2}$$

$$\mathrm{CC} = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}} \tag{3}$$

where X_i and Y_i represent the observed and model-predicted EC values, respectively; $\bar{X}$ and $\bar{Y}$ denote their corresponding mean values, and N is the total number of samples.

2.5. GALDIT Method

The GALDIT method is a technique utilized to diagnose the susceptibility of groundwater to saltwater intrusion through an index-based method. Chachadi et al. proposed a methodology for mapping aquifer vulnerability, particularly focusing on factors influencing SWI [13]. Which include: type of aquifer; unconfined, confined, or semi-confined (G), the ability of an aquifer to transmit water through its porous media (A), the altitude of the water table measured above sea level (L), measured distance of the observation wells perpendicularly inland from the coast (D), the impact of existing status of SWI (I), and saturated thickness (T) [26]. A numerical ranking system has been developed utilizing these factors to evaluate the extent of SWI, comprising three main components: weights, ranges, and ratings. The range of parameter variables is divided into four ratings, each customized for the specific area being studied, considering the available hydrogeologic and morphologic data. Each rating is given an importance score of 2.5, 5, 7.5, or 10, indicating its significance relative to SWI. Higher importance ratings signify a greater influence on SWI. The GALDIT equation (Equation (4)) includes the set of weights and rating values, which are used to calculate the GALDIT index by multiplying the weights with the rating values for each data point [35].

$$GALDIT_{Index} = \frac{\sum_{i=1}^{6}(W_i * R_i)}{\sum_{i=1}^{6} W_i} \tag{4}$$

where W_i stands for the weight given to the indicator parameter, and R_i shows the importance level of the i^{th} indicator parameter. The GALDIT index ranges from 2.5 to 10 [30]. This classification system enables the evaluation and description of how vulnerable aquifers are to seawater intrusion, based on the GALDIT index values [30]. Presents the weight assigned to each parameter, the range of variables, and the corresponding importance rating associated with each parameter. This information is important for performing a detailed vulnerability assessment with the GALDIT model [15].

2.6. Feed Forward Neural Networks (FFNN)

FFNNs are computational systems engineered to process large-scale data, mirroring the neural structures of the human brain. They are designed to recognize complex patterns and relationships within data. These models are particularly effective at handling non-linear relationships making them powerful tools for various applications [36]. The FFNN involves three key components: input layer, hidden layers, and output layer. The input layer receives raw data, which is then processed by one or more hidden layers. Each hidden layer consists of neurons

that apply activation functions to weighted inputs, transforming the data through non-linear mappings. These transformations enable the network to capture complex patterns and representations. The output layer produces the final prediction or classification based on the processed information from the hidden layers [37,38]. The process begins with identifying suitable input variables, followed by choosing the type of neural network to be used. Next, the data is preprocessed and organized for optimal performance. The structure of the network is then designed, after which criteria for evaluating the performance of the model are established. The model undergoes a training phase to optimize the weights, and lastly, it is validated to ensure accuracy and reliability [28].

In this study, the FFNN is trained using a feedforward architecture combined with a back-propagation algorithm. This approach has been effectively applied to model various aspects of hydrological processes within FFNN frameworks [39]. The output of an FFNN is computed using the following formula:

$$\hat{y}_j = f_j\left[\sum_{h=1}^{m} w_{jh} \times f_j\left(\sum_{i=1}^{n} w_{hi}x_i + w_{hb}\right) + w_{jb}\right] \tag{5}$$

where, *i*, *h*, *j*, *b*, and *w* represent neurons of the input, hidden, and output layers, denoting the bias and weight provided by the neuron, respectively. The activation functions for the hidden and output layers are denoted as f_h and f_j, respectively. Additionally, x_i, n, and m represent the input value, total number of inputs, and number of neurons in the hidden layer, respectively. Finally, y and $\hat{y}_j$ signify the observed and calculated target values, respectively.

2.7. Deep Feed Forward Neural Network (DFFNN)

Multi-Layer Perceptron (MLP) networks, among the most widely used neural network models, are nowadays employed to develop the FFNNs. The MLP architecture comprises three main layers: the input layer, one or more hidden layers, and the output layer [40]. DFFNN is a type of FFNNs that consists of multiple layers of interconnected nodes, or neurons [36]. Unlike conventional FFNN, DFFNNs have multiple hidden layers (≥ 2) between the input and output layers, allowing for the extraction of complex features from the input data [41]. Within the architecture of a DFFNN, data propagate from the input layer through these hidden layers to the output layer, lacking any feedback or loops, thereby providing it with enhanced depth and speed [42,43].

Several studies have classified DFFNN as a deep learning method due to its multi-layered structure [41,44,45]. The architecture of a DFFNN enables each layer to progressively learn and represent abstract features of the input data, with the final layer generating the output of the network. This learning process involves adjusting the weights of connections between layers to minimize a loss function, measuring the disparity between predicted and actual outputs. Through extensive calibration on substantial datasets, DFFNNs can yield highly accurate models of intricate phenomena [46].

2.8. Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS is a hybrid intelligent model that integrates the neural learning capability of artificial neural networks with the fuzzy reasoning structure of fuzzy logic. It is based on a first-order Sugeno-type fuzzy inference system where relationships are nonlinear, uncertain, or gradually varying [47]. The model consists of five computational layers that perform fuzzification, rule activation, normalization, defuzzification, and output computation [48]. In the fuzzification layer, input variables are converted into fuzzy membership degrees using functions such as Gaussian or bell-shaped curves. The rule layer then generates fuzzy if–then rules, while the normalization and consequent layers compute the weighted contributions of each rule. Parameter optimization is performed through a hybrid learning approach combining least-squares estimation with gradient descent. For a two-input, first-order Sugeno ANFIS (inputs: $x$and y), the fuzzy rules are formulated as:

Rule 1: If x is $A_1$and y is B_1, then

$$f_1 = p_1x + q_1y + r_1 \tag{6}$$

Rule 2: If x is A_2 and y is B_2, then

$$f_2 = p_2x + q_2y + r_2 \tag{7}$$

Membership functions for antecedents are commonly defined using Gaussian functions:

$$\mu_{A_i}(x) = \exp\left[-\frac{(x-c_i)^2}{2\sigma_i^2}\right], \mu_{B_i}(y) = \exp\left[-\frac{(y-d_i)^2}{2\gamma_i^2}\right] \tag{8}$$

The firing strengths of the fuzzy rules are:

$$w_i = \mu_{A_i}(x)\,\mu_{B_i}(y) \tag{9}$$

Normalized firing strengths:

$$\bar{w}_i = \frac{w_i}{w_1 + w_2} \tag{10}$$

Final output of the ANFIS model:

$$f = \bar{w}_1 f_1 + \bar{w}_2 f_2 \tag{11}$$

2.9. Support Vector Regression (SVR)

Support Vector Regression (SVR) is a kernel-based machine learning method derived from the Support Vector Machine (SVM) framework and is widely used for nonlinear regression. Unlike traditional regression models that minimize the squared error, SVR adopts the principle of structural risk minimization to balance model complexity and prediction accuracy, thereby improved generalization in noisy or irregular datasets [49,50]. SVR maps input data into a high-dimensional feature space using kernel functions, where linear relationships can effectively approximate nonlinear interactions observed in the original input space [51]. Commonly applied kernels include the radial basis function (RBF), which is particularly suited for environmental modeling due to its flexibility in capturing complex nonlinear patterns.

Given a training dataset (x_i, d_i) for $i = 1,2,\ldots,N$, the general SVR function can be written as:

$$f(x) = w^T \phi(x) + b \tag{12}$$

where $\phi(x)$ is the nonlinear mapping from the input space to the feature space, w is the weight vector, and b is the bias term. The SVR optimization problem aims to minimize:

$$\frac{1}{2} \parallel w \parallel^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i^*) \tag{13}$$

subject to:

$$d_i - w^T \phi(x_i) - b \le \varepsilon + \xi_i \tag{14}$$

$$w^T \phi(x_i) + b - d_i \le \varepsilon + \xi_i^* \tag{15}$$

where C is the regularization constant controlling the trade-off between model flatness and approximation error, ε is the insensitive loss margin, and ξ_i, ξ_i^* are slack variables [50]. The dual formulation yields the final SVR function:

$$f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*)\, k(x_i, x) + b \tag{16}$$

where α_i and α_i^* are Lagrange multipliers, and $k(x_i, x)$ is the kernel function.

The Gaussian radial basis function (RBF), one of the most frequently used kernels, is given by:

$$k(x_i, x_j) = \exp\left(-\gamma \parallel x_i - x_j \parallel^2\right) \tag{17}$$

where γ is the kernel width parameter.

2.10. Ensemble

It is widely accepted that combining the outputs of multiple forecasting models can significantly improve time-series prediction accuracy. A major advantage of ensemble prediction is that the final estimate becomes less dependent on a single model behavior, producing more stable and reliable results. Numerous theoretical and empirical studies have shown that aggregated predictions often outperform individual models, and that combining several relatively simple approaches can sometimes be safer and more robust than relying solely on a single complex model [52]. In this study, the predictions of three machine learning models (FFNN, SVR, and ANFIS) were combined.

2.11. Extreme Learning Machine (ELM)

The Extreme Learning Machine (ELM) is a single-hidden-layer feedforward neural network designed for fast and efficient learning. Unlike traditional neural networks which update all weights iteratively, ELM randomly

initializes the hidden-layer parameters and analytically determines the output weights using a least-squares solution [53]. This structure significantly reduces training time while maintaining strong approximation capabilities for nonlinear functions, making ELM particularly suitable for large-scale or computationally demanding environmental datasets [54]. The architecture consists of an input layer, a single hidden layer with randomly generated nodes, and an output layer where optimal weights are computed in closed form.

Given training data (x_i, d_i) for $i = 1,2,\ldots,N$, the ELM model can be expressed as:

$$f(x) = \sum_{j=1}^{L} \beta_j \, h_j(x) \tag{18}$$

where L is the number of hidden neurons, $h_j(x)$ are the activation outputs of the hidden nodes, and β_j are the output weights. The hidden-layer output matrix H is computed as:

$$H_{ij} = g(w_j \cdot x_i + b_j) \tag{19}$$

where $g(\cdot)$ is the activation function, $w_j$represents randomly assigned input weights, and b_j denotes the bias of the j-th hidden neuron.

The output weights β are obtained using:

$$\beta = H^{\dagger} D \tag{20}$$

where $H^{\dagger}$ is the Moore–Penrose pseudoinverse of H, and D is the target vector [53].

2.12. Bayesian Optimization (BO)

Bayesian Optimization (BO) is a probabilistic, model-based optimization technique designed for efficiently searching high-dimensional or computationally expensive parameter spaces. Instead of evaluating every possible parameter configuration, BO constructs a surrogate model—typically a Gaussian Process (GP)—that approximates the underlying objective function and quantifies uncertainty across the search domain [55,56]. Using this surrogate, BO selects new candidate points through an acquisition function that balances exploration of uncertain regions with exploitation of areas expected to yield improved performance. This strategy allows BO to identify near-optimal hyperparameters with significantly fewer evaluations compared to exhaustive or grid-based searches. The general optimization objective can be expressed as:

$$\theta^{*} = \arg\min_{\theta} \; \mathcal{L}(\theta) \tag{21}$$

where θ represents the model hyperparameters and $\mathcal{L}(\theta)$ is the loss function evaluated on validation data. Bayesian Optimization models the loss function as a Gaussian Process:

$$\mathcal{L}(\theta) \sim GP(m(\theta), k(\theta, \theta')) \tag{22}$$

where $m(\theta)$ is the mean function and $k(\theta, \theta')$ is the covariance kernel representing similarity between points. The next hyperparameter candidate is chosen by maximizing the acquisition function:

$$\theta_{\text{next}} = \arg\max_{\theta} \; \alpha(\theta \mid GP) \tag{23}$$

3. Results and Discussion

3.1. Outcomes Derived from the GALDIT Approach

This study utilized the GALDIT framework to evaluate the risk of SWI affecting the aquifers in the Salmas and Urmia Plains, by integrating six critical parameters. The GALDIT index, while valuable, relies on subjective expert judgment and lacks the ability to adapt to varying aquifer characteristics dynamically. FFNN and DFFNN models were introduced to mitigate these limitations by employing their advanced pattern recognition capabilities. The parameters of the GALDIT framework served as inputs for these models; the G parameter was excluded because it exhibits negligible temporal and geographic variance. EC was selected as the output variable to represent salinity levels, offering critical insights into SWI. In this study, EC values were derived from salinity measurements of groundwater samples collected from wells spread throughout the aquifer during September. While alternative targets such as nitrate or sulfate levels could be used in different settings, EC was most appropriate for evaluating the vulnerability of the aquifer in the study area to SWI. Considering the strong correlation between salinity and EC in water, higher salinity generally results in increased EC [21]. Therefore, EC

can be a valuable parameter for assessing the susceptibility of coastal aquifers to saltwater contamination. The GALDIT method assesses six key parameters to determine the vulnerability of aquifers to SWI:

i Groundwater Occurrence (G)

G refers to the type and presence of groundwater in the aquifer, which determines the aquifer type [30], as shown in Table 1. The type of aquifer can be inferred from well logs, geological surveys, and aquifer tests [24]. Well logs provide detailed information about subsurface materials and their properties, while geological surveys identify geological formations and their extents [57]. In this study, the data indicate that aquifers in the Urmia and Salmas Plains are unconfined, with a rating of 7.5 as depicted in Table 2 and Figure 5a for Salmas, and Figure 6a for Urmia aquifers.

Table 1. Assessment guidelines for GALDIT parameters adapted from [15,21].

Parameter	Weight	Standard GALDIT Factor Variables Range	Classes	Importance Rating
Groundwater occurrence (G)	1	Confined aquifer	Confined aquifer	10
		Unconfined aquifer	Unconfined aquifer	7.5
		Leaky confined aquifer	Leaky confined aquifer	5
		Bounded aquifer	Bounded aquifer	2.5
Aquifer hydraulic conductivity (A) [m/day]	3	>40	High	10
		10–40	Medium	7.5
		5–10	Low	5
		<5	Very Low	2.5
Height of groundwater level above sea level (L) [m]	4	<1	High	10
		1–1.5	Medium	7.5
		1.5–2	Low	5
		>2	Very Low	2.5
Distance from the shore (D) [m]	4	<500	Very Small	10
		500–750	Small	7.5
		750–1000	Medium	5
		>1000	Far	2.5
Impact of existing status of SWI (I)	1	>2 (Ratio of (Cl^-/HCO_3))	High	10
		1.5–2	Medium	7.5
		1–1.5	Low	5
		<1	Very Low	2.5
Saturated aquifer thickness (T) [m]	2	>10	Large	10
		7.5–10	Medium	7.5
		5–7.5	Small	5
		<5	Very Small	2.5

Table 2. GALDIT Parameter Ratings and Associated Percentages for Salmas and Urmia Aquifers.

GALDIT Parameters (Units)	Salmas Aquifer Rates	Percentage of Salmas Aquifer Rating	Urmia Aquifer Rates	Percentage of Urmia Aquifer Ratings
G	7.5	100%	7.5	100%
A (md^{-1})	5	0.05%	2.5	0.8
	7.5	7.8%	5	4.2
	10	92.1%	7.5	91.1
			10	3.9
L (m)	2.5	95%	2.5	100
	5	0.6%		
	7.5	0.5%		
	10	3.9%		
D (m)	2.5	100	2.5	98.55
			5	1.2
			7.5	0.2
			10	
I	2.5	79%	2.5	92.6
	5	6.9%		
	7.5	4.6%	5	7.4
	10	3.9%		
T (m)	10	100%	10	100%

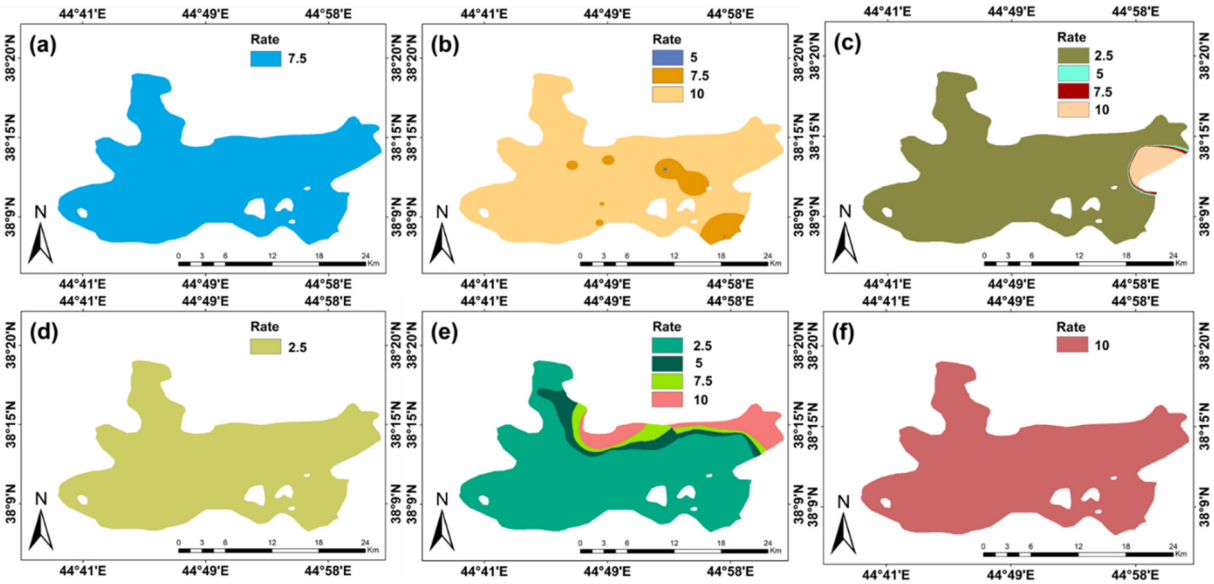


Figure 5. Variation and corresponding ratings of GALDIT index parameters for the Salmas aquifer: (**a**) Groundwater occurrence, (**b**) Aquifer hydraulic conductivity, (**c**) Groundwater level, (**d**) Distance from the shoreline, (**e**) Impact of existing SWI, (**f**) Aquifer thickness.

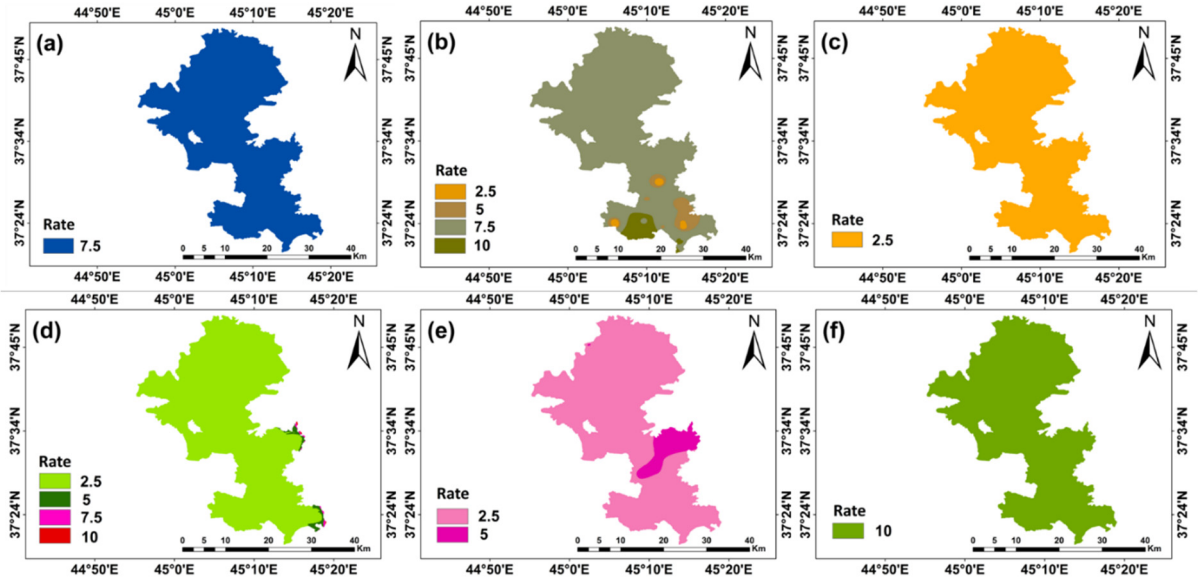


Figure 6. Distribution and associated ratings of GALDIT index parameters for the Urmia aquifer: (**a**) Groundwater occurrence, (**b**) Aquifer hydraulic conductivity, (**c**) Groundwater level, (**d**) Distance from the shoreline, (**e**) Impact of existing SWI, (**f**) Aquifer thickness.

ii Aquifer Hydraulic Conductivity (A)

"A" refers to the ability of the aquifer material to transmit water [58]. Hydraulic conductivity is typically measured in meters per day (m/day) and is determined through various methods, including aquifer pumping tests. Pumping tests involve extracting water from a well at a known rate and measuring the resulting drawdown in the water table [59]. This data allows for the calculation of the transmissivity of the aquifer, from which hydraulic conductivity can be derived using the aquifer thickness [21]. The rates and percentages of A for both aquifers are shown in Table 2. Figure 5b illustrates the ratings for the Salmas aquifer, which is between 5 and 10, while Figure 6b depicts those for the Urmia aquifer between 2.5 and 10. Salmas plain hydraulic conductivity ranges from 0.8 to 96 and Urmia aquifer Hydraulic conductivity range from 0.03 to 46. In the eastern part of the Urmia aquifer, there are salty flats and salt flats, whereas moving westward reveals gravel plains. Therefore, hydraulic conductivity increases as we move toward western parts of the aquifer.

iii Groundwater Level (L)

Parameter L represents the elevation of the water table above sea level. This parameter is crucial because it controls the pressure and hydraulic gradient in the freshwater-saltwater transition zone, thereby influencing the potential for SWI [21]. Groundwater levels are typically measured using piezometers or observation wells, which provide precise data on the depth of the water table. To clarify the basis for the "L" parameter, this study adopts the standard GALDIT definition, in which groundwater level is measured as elevation above mean sea level [31]. In the Salmas and Urmia aquifers, groundwater levels range from −5 to 224 m, reflecting the complex terrain of the region and elevation profile. Although these values place groundwater well above the lake elevation (~1271 m a.s.l.), saltwater intrusion remains a valid concern due to lateral salinity pressure from Lake Urmia and extensive groundwater withdrawal. Overextraction lowers hydraulic head locally, potentially reversing gradients near the lake interface and allowing brackish water to migrate inland, particularly in shallow alluvial zones with hydraulic connectivity. Therefore, SWI vulnerability is not solely determined by absolute elevation but also by groundwater pumping, aquifer structure, and lake-aquifer interactions. The groundwater level ratings percentages for each aquifer are provided (see Table 2). The GIS layer for L in the Salmas aquifer, which ranges from −5 to 224 m with ratings between 2.5 and 10, is shown in Figure 5c. For the Urmia aquifer, the groundwater levels range from 5 to 88 m, all having a rating of 10 (See Figure 6c).

iv Distance from shoreline (D)

The parameter D represents the distance of the aquifer from the shoreline. This is a significant factor as it influences the extent of SWI, with aquifers closer to the shoreline being more susceptible to contamination by seawater. The distance from the shoreline is typically measured using spatial analysis tools in GIS software, which calculates the Euclidean distance between the aquifer locations and the nearest shoreline [24]. The aquifers are closest to Lake Urmia in the eastern parts. Aquifers closer to the lake are more vulnerable due to their proximity to saltwater sources. The rating values for D range from 2.5 to 10, with closer distances receiving higher vulnerability ratings (See Table 2). The GIS layer depicting D for the Salmas aquifer, rated 2.5, is shown in Figure 5d, while those for the Urmia aquifer, with ratings between 2.5 and 10, are illustrated in Figure 6d.

v Impact of the existing status of SWI (I)

This parameter evaluates the current degree of SWI by examining chloride and bicarbonate ion concentrations in groundwater [30]. The ratio of these ions provides an indication of salinity levels, with higher chloride concentrations signifying greater intrusion. The percentage of ratings for both aquifers is presented in Table 2. The ratings for the Salmas aquifer range from 2.5 to 10, as shown in Figure 5e, while the ratings for the Urmia aquifer are 2.5 and 5 (See Figure 6e.)

vi Aquifer thickness (T)

Aquifer thickness refers to the vertical extent of the saturated zone, measured in meters. Thicker aquifers have a larger volume of water that can be affected by SWI [21]. The Salmas aquifer has a thickness ranging from 10 to 80 m, whereas the Urmia aquifer has a thickness ranging from 40 to 122 m. The rating for both aquifers is 2.5, as depicted in Figures 5f and 6f for the Salmas and Urmia aquifers, respectively.

Raster files with a 30 m resolution were created for each GALDIT variable. The final vulnerability maps for the Salmas and Urmia aquifers (Figure 7a,b) were produced by overlaying the six layers, each weighted according to the standards outlined in Table 1. Different classification techniques, including natural breaks, equal intervals, quantiles, and geometric intervals—have been explored by Huan et al. to identify categories of groundwater vulnerability [60]. Based on the guidance provided by Huan et al., Nakhaei et al., and Barzegar et al. this study applied geometric interval classification to divide GALDIT index values into four classes: low, moderate, high, and very high vulnerability [24,60,61]. The data for the Salmas Plain (Figure 7a) shows that, according to the GALDIT vulnerability map, 7.8% of the area is classified as low vulnerability, 68.6% as moderate vulnerability, 18.9% as high vulnerability, and 4.7% as very high vulnerability. For the Urmia aquifer, 38.1% of the region exhibits low vulnerability, 57.5% shows moderate vulnerability, 4.2% falls under high vulnerability, and 0.2% is identified as very high vulnerability.

In the Salmas Plain aquifer, the high vulnerability in the northeastern part is primarily due to excessive groundwater pumping, which lowers the groundwater level, thereby increasing the hydraulic gradient and the potential for SWI. Additionally, the reduction in groundwater levels results in a lower hydraulic head, facilitating the inland movement of seawater and exacerbating vulnerability. The proximity of the region to Lake Urmia further contributes to its very high vulnerability (D), as the shorter distance to the saltwater source means that any reduction in groundwater levels or increase in hydraulic conductivity directly impacts the rate and extent of SWI. Furthermore, the presence of the Zola River, which flows from the western part of the plain (originating in Turkey)

to the east, plays a significant role. The flow dynamics of the river can influence the hydraulic gradient and groundwater recharge patterns, potentially increasing vulnerability in areas adjacent to its course.

In the Urmia Plain aquifer, the exceptionally high vulnerability areas near Lake Urmia are due to the proximity of pollution sources to the water table, which significantly increases the likelihood of groundwater contamination. Areas closer to the lake exhibit higher vulnerability, since pollutants from the lake have greater potential to infiltrate the surrounding groundwater [62]. The presence of gypsum and salt deposits near the lake significantly increases vulnerability. These deposits dissolve easily, raising the salinity of the groundwater. This directly impacts groundwater quality and salinity, thereby increases the vulnerability rating of these areas. Groundwater in the Urmia aquifer flows from west to east towards Lake Urmia, carrying potential contaminants towards the lake [31]. This flow pattern contributes to the higher vulnerability near the lake as pollutants are more likely to be transported into these areas. In the southwestern parts of the Urmia Plain, the alluvial thickness ranges from 150 to 200 m [31]. According to the GALDIT framework, thickness (T) is another significant factor influencing vulnerability. Thicker alluvial deposits in these regions may facilitate greater pollutant transport and infiltration, thereby increasing vulnerability. The presence of major rivers such as Barandoz Chai in the southern part of the aquifer, Nazlu and Roze Chay in the northern part also impacts vulnerability. These rivers can induce recharge carrying pollutants into the aquifer. This river-induced recharge, especially in areas close to these water bodies, elevates vulnerability.

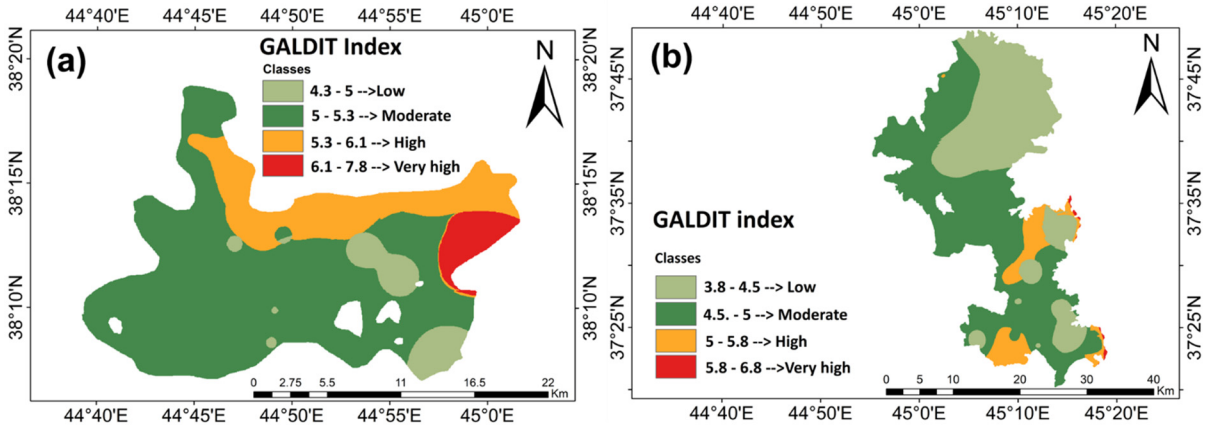
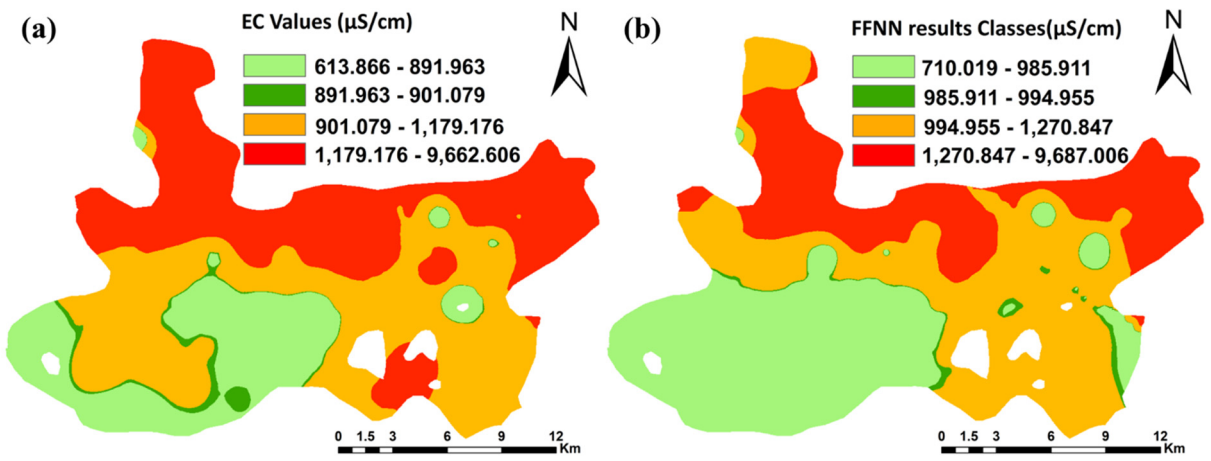


Figure 7. (**a**) GALDIT vulnerability map of the Salmas aquifer and (**b**) Urmia Aquifers GALDIT index map.

The comparison between the GALDIT vulnerability maps (Figure 7) and the observed and modeled EC distributions (Figures 8 and 9) clearly illustrates how differently these approaches represent groundwater conditions. The GALDIT index provides a broad, generalized view of vulnerability based on hydrogeological parameters, whereas the EC maps—both the observed values and the predictions generated by the AI models—display much more detailed and locally variable patterns. These spatial contrasts show that the unsupervised, expert-based GALDIT method cannot reflect the fine-scale salinity variations captured in the field. In contrast, the AI-based predictions align far more closely with the observed EC distribution, indicating a stronger ability to present actual salinity patterns throughout the aquifer.



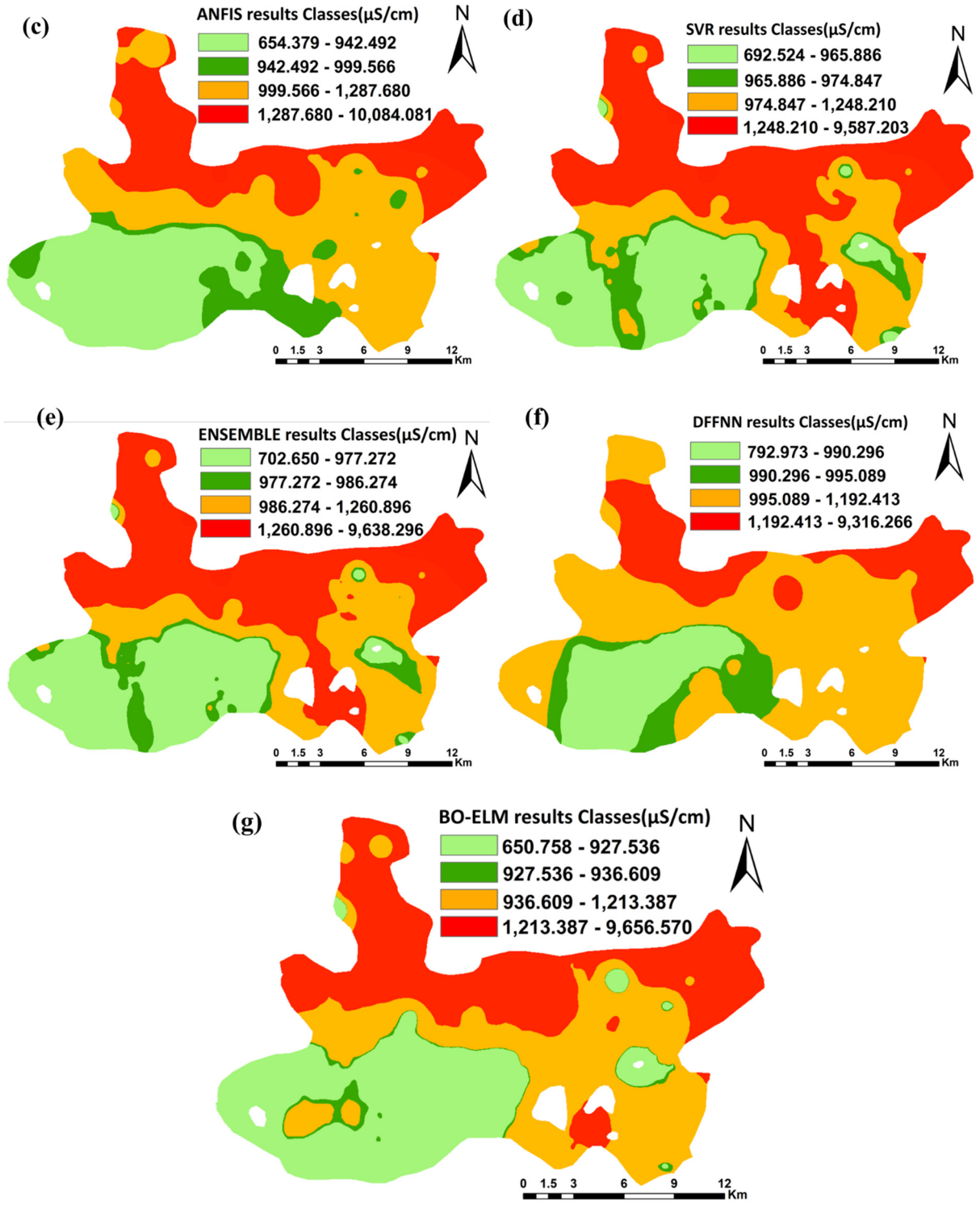
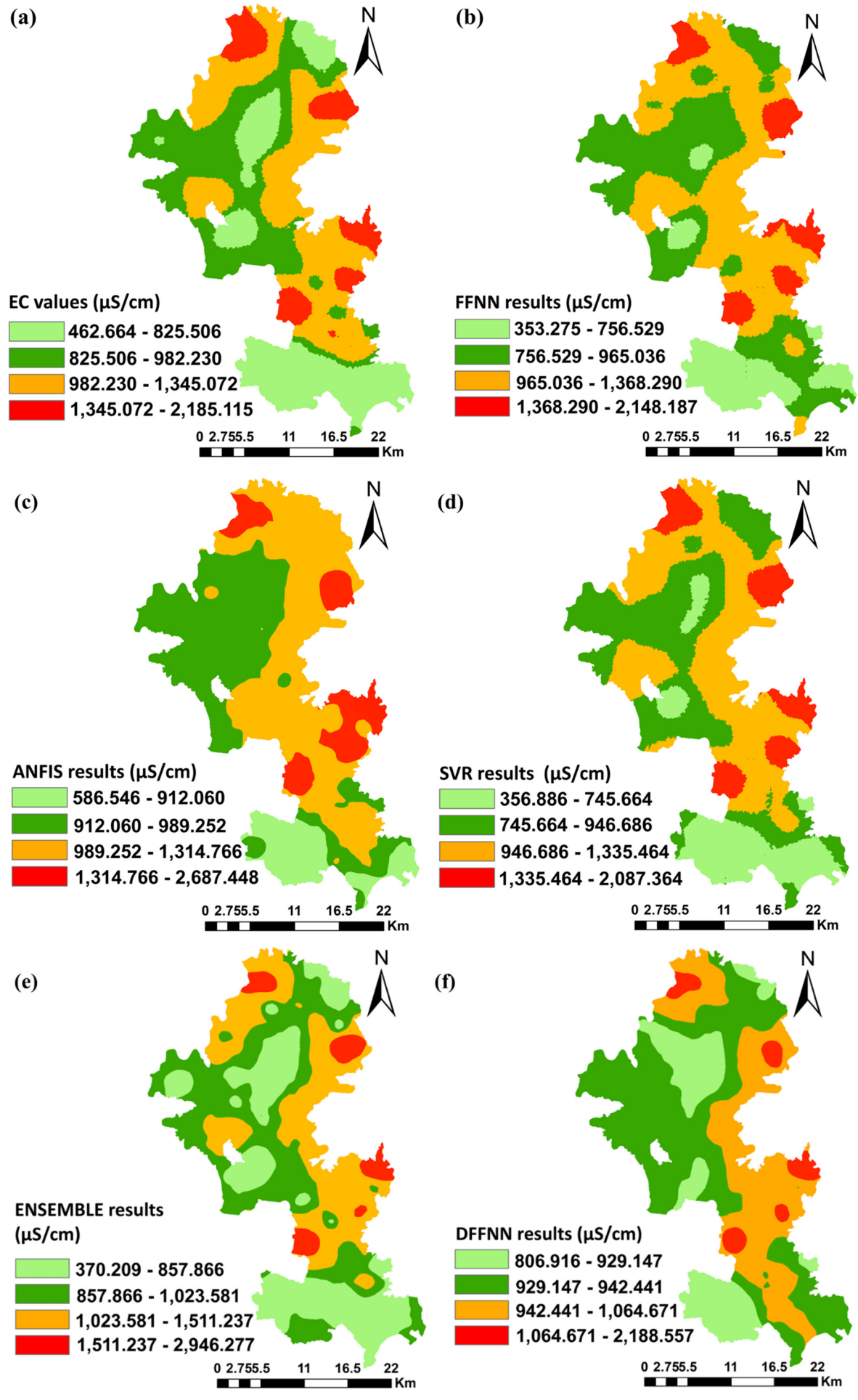


Figure 8. Spatial distribution of groundwater EC in the Salmas aquifer based on observational data and model predictions; (**a**) EC values from monitoring wells, (**b**) FFNN-predicted EC classes, (**c**) ANFIS-predicted EC classes, (**d**) SVR-predicted EC classes, (**e**) Ensemble model results, (**f**) DFFNN results, (**g**) BO-ELM model results.


(a)
N
EC values (µS/cm)
462.664 - 825.506
825.506 - 982.230
982.230 - 1,345.072
1,345.072 - 2,185.115
0 2.755.5 11 16.5 22
Km
(b)
N
FFNN results (µS/cm)
353.275 - 756.529
756.529 - 965.036
965.036 - 1,368.290
1,368.290 - 2,148.187
0 2.755.5 11 16.5 22
Km
(c)
N
ANFIS results (µS/cm)
586.546 - 912.060
912.060 - 989.252
989.252 - 1,314.766
1,314.766 - 2,687.448
0 2.755.5 11 16.5 22
Km
(d)
N
SVR results (µS/cm)
356.886 - 745.664
745.664 - 946.686
946.686 - 1,335.464
1,335.464 - 2,087.364
0 2.755.5 11 16.5 22
Km
(e)
N
ENSEMBLE results (µS/cm)
370.209 - 857.866
857.866 - 1,023.581
1,023.581 - 1,511.237
1,511.237 - 2,946.277
0 2.755.5 11 16.5 22
Km
(f)
N
DFFNN results (µS/cm)
806.916 - 929.147
929.147 - 942.441
942.441 - 1,064.671
1,064.671 - 2,188.557
0 2.755.5 11 16.5 22
Km

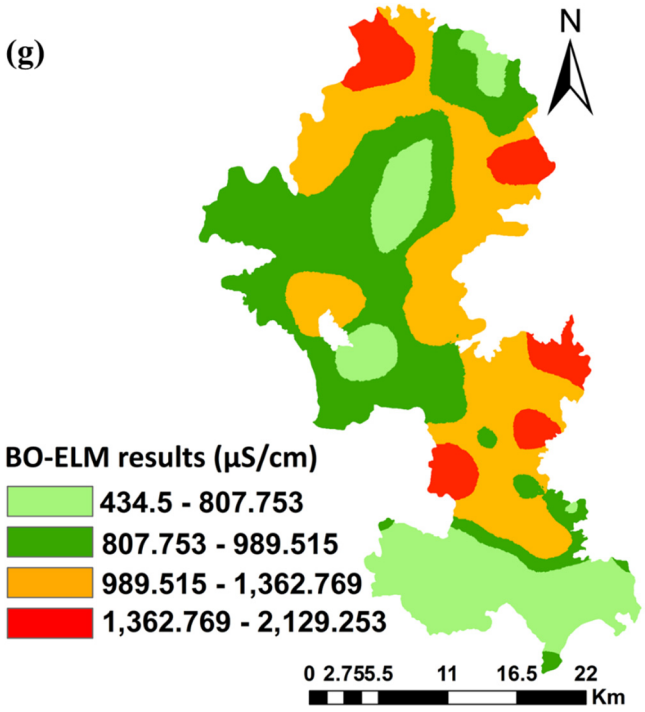


Figure 9. Spatial distribution of groundwater EC in the Urmia aquifer based on observational data and model predictions; (**a**) EC values from monitoring wells, (**b**) FFNN-predicted EC classes, (**c**) ANFIS-predicted EC classes, (**d**) SVR-predicted EC classes, (**e**) Ensemble model results, (**f**) DFFNN results, (**g**) BO-ELM model results.

3.2. Results of the AI-Based Approaches

As discussed in previous sections of this paper, assessing groundwater vulnerability is a crucial step in water resource management. Despite this, conventional and unsupervised methods (e.g., GALDIT and DRASTIC) may lead to errors and ambiguities in the estimation and evaluation of aquifer vulnerability due to the inclusion of expert judgment in the ranking and weighting of parameters. To increase the accuracy of the aquifer vulnerability assessment, some studies employed various techniques (e.g., see [21,63]). This study developed a methodology (Different AI-based approaches) to evaluate the surrounding aquifers susceptibility to salinization in Urmia Lake. The EC values from each well, along with other parameters, were added to GIS. After the interpolation process, the data were resampled, resulting in more than 1000 data points for each aquifer. This approach ensured a sufficiently large dataset, reducing the risk of overfitting and enhancing the reliability of the model, even when using deeper architectures with a higher number of trainable variables. It is worth mentioning that determining the optimal number of hidden layers and neurons is a standard preprocessing step in ML-based studies.

3.2.1. Results of FFNN & DFFNN Approaches

Following normalization of all the data, GALDIT parameters were applied as inputs, and EC values were utilized as outputs. The feature values were normalized to a standard range (0 to 1) to facilitate network convergence. By reducing the range of input values, network weights are updated more uniformly, and normalization prevents features with higher values from dominating the training process, improving model accuracy. The modeling was then conducted using FFNN and DFFNN, with 75% of the data for training and 25% for testing. After training the network, the test data were introduced to the model.

Using the evaluation metrics, the results for the Salmas aquifer show a significant performance improvement when using the DFFNN model compared to the FFNN model (See Table 3). For example, the RMSE (µS/cm) for the DFFNN model is 0.03 in both the training and testing phases, compared to 0.05 (training) and 0.06 (testing) for the FFNN model. This reduction in RMSE indicates that the DFFNN model provides more accurate predictions. Additionally, the DC values improved from 0.89 (training) and 0.82 (testing) in the FFNN model to 0.96 and 0.95, respectively, in the DFFNN model, suggesting a stronger correlation between observed and predicted values. Furthermore, the CC values also show improvement, with the DFFNN model achieving 0.98 (training) and 0.97 (testing), compared to 0.94 and 0.91 for the FFNN model. For the Urmia aquifer, the FFNN and DFFNN models were also evaluated. Similar to the Salmas aquifer, the DFFNN model outperformed the FFNN model, though the degree of improvement was less pronounced. The RMSE (µS/cm) values for the DFFNN model were 0.10

(training) and 0.11 (testing), which are slightly lower than the 0.12 and 0.13 values obtained by the FFNN model. This slight reduction in RMSE indicates a marginally better fit by the DFFNN model. The DC values increased from 0.70 (training) and 0.66 (testing) in the FFNN model to 0.80 and 0.75, respectively, in the DFFNN model, suggesting an enhanced correlation in the DFFNN model. The CC values also improved from 0.83 (training) and 0.81 (testing) in the FFNN model to 0.89 and 0.87 in the DFFNN model (See Table 3). As a result, the DFFNN model produced more accurate estimates of the specific vulnerability (EC values) of aquifers in the Salmas and Urmia regions. This method has multiple hidden layers that can influence the modeling process due to its unique characteristics. Despite this, the standard FFNN approach is powerful and efficient, and its results are reliable.

Table 3. Performance metrics of FFNN and DFFNN models for both aquifers.

Study Area	Model	* Best Structure	Epoch	RMSE (µS/cm)		DC		CC	
				Train	Test	Train	Test	Train	Test
Salmas aquifer	FFNN	5-4-1	50	0.05	0.06	0.89	0.82	0.94	0.91
	DFFNN	5-(8)-(8)-(8)-(1)-1	300	0.03	0.03	0.96	0.95	0.98	0.97
Urmia aquifer	FFNN	5-3-1	50	0.12	0.13	0.70	0.66	0.83	0.81
	DFFNN	5-(8)-(8)-(8)-(1)-1	300	0.10	0.11	0.80	0.75	0.89	0.87

* The structure denotes: inputs-hidden layers-output.

3.2.2. Results of ANFIS Approach

The ANFIS model for Urmia was tested using several types of membership functions. Among them, the Gaussian membership function (gaussmf) provided the best performance. Multiple trials were conducted up to 100 epochs, but the results showed only minor differences beyond 50 iterations, indicating that the model converged effectively within this range. Additionally, the system did not perform well when using four membership functions per input, likely due to increased in rule complexity and insufficient data support. Therefore, the final configuration for Urmia used Gaussian MF, a stable number of epochs (50), and a reduced number of MFs to ensure reliable performance. According to the results (See Table 4), the ANFIS model demonstrates satisfactory predictive performance for the Urmia plain. The RMSE value of 0.1033 indicates a very low deviation between the observed and predicted values. Similarly, the DC (0.808) and CC (0.822) confirm a strong agreement and correlation between the model outputs and the actual data.

For the Salmas aquifer, the ANFIS model was directly run using the Gaussian membership function, based on the superior performance observed in the Urmia trials. Since Gaussian MF provided strong stability and generalization, the same structure was applied successfully to Salmas without requiring additional MF comparisons. This ensured a consistent and effective fuzzy–neural representation of the nonlinear relationships between GALDIT parameters and EC in the Salmas dataset. For the Salmas plain, the ANFIS model exhibits even higher accuracy. The RMSE of 0.161 remains low, reflecting minimal prediction error. The efficiency of the model is further confirmed by the DC value of 0.9936 and the CC of 0.9827, both of which indicate excellent consistency and a near-perfect correlation with the observed data (See Table 4).

Table 4. Evaluation results for both aquifers using other models.

Model Type	Urmia Plain			Salmas Plain		
	RMSE	DC	CC	RMSE	DC	CC
ANFIS	0.1033	0.634	0.822	0.0161	0.9936	0.9827
SVR	0.108	0.76	0.895	0.075	0.930	0.958
Integrated Ensemble	0.015	0.995	0.895	0.0176	0.988	0.974
BO-ELM	0.082	0.864	0.830	0.009	0.995	0.997

3.2.3. Results of SVR Approach

In the SVR model, the Radial Basis Function (RBF) kernel was used due to its ability to model non-linear relationships in the data. A grid-search procedure was used to evaluate different combinations of the hyperparameters C (1, 10, 100) and gamma (1, 0.1, 0.01), resulting in nine candidate configurations. Each configuration was assessed through 5-fold cross-validation, and the model with the best overall performance was selected for further analysis. The corresponding validation results are presented below.

As shown in Table 4, for the Urmia plain, the SVR model produced an RMSE of 0.108, a CC of 0.895, and a DC of 0.76. These metrics indicate that the model captures the main trends of the observed data, although the level of agreement is moderate compared to other approaches. For the Salmas plain, the SVR results in Table 4

reveal an RMSE of 0.075, a CC of 0.958, and a DC of 0.930, indicating a strong correspondence between predicted and observed values and demonstrating that the model performs more reliably in this region than in Urmia.

3.2.4. Results of Integrated Ensemble Approach

The ensemble model for Urmia was implemented using a feedforward neural network (FFNN) as the meta-learner, with 50 neurons in the hidden layer (hn = 50) and a training duration of 20 epochs. This architecture was used to combine the outputs of the base models (ANFIS, SVR, and FFNN). The combination of a moderate hidden-layer size and a short training period yielded stable, accurate ensemble predictions while avoiding overfitting. This configuration consistently produced reliable integration of the individual model outputs for the Urmia aquifer.

For the Urmia plain, the Ensemble model provides reliable performance (See Table 4). The RMSE of 0.015 indicates a low level of prediction error. The DC value (0.995) shows that the model reproduces the observed variability with high consistency. Additionally, the CC of 0.895 reflects a strong correlation between the predicted and observed values. Overall, the results suggest that the Ensemble model performs effectively for this region.

For the Salmas aquifer, the same ensemble configuration was applied: 50 hidden neurons (hn = 50) and 20 training epochs. This structure again performed effectively, demonstrating strong generalization when merging the outputs of the base learners. The ensemble FFNN enhanced predictive performance by capturing nonlinear interactions among the model outputs while maintaining a simple, computationally efficient architecture. According to Table 4, for the Salmas plain, the Ensemble model also performs well. The RMSE of 0.0176 remains low, indicating acceptable accuracy in the predictions. The DC value (0.988) demonstrates that the model explains most of the variability in the observed data. Similarly, the CC of 0.974 indicates strong agreement between the predicted and measured values. These findings confirm that the Ensemble model provides dependable results for the Salmas plain.

3.2.5. Results of BO-ELM Approach

For the Urmia Plain, the BO-ELM model was optimized using Bayesian Optimization, which selected 720 hidden neurons and the tanh activation function. This configuration enabled the model to capture the complex nonlinear behavior of the dataset. The chosen optimization process also used a relatively large regularization parameter ($\lambda \approx 0.645$) to prevent overfitting, suggesting that the Urmia data contained substantial variability or noise. The best cross-validated RMSE achieved was 0.088, reflecting improved model stability and accuracy.

For the Salmas Plain, Bayesian Optimization selected 540 hidden neurons and the tanh activation function. The optimized regularization value ($\lambda \approx 0.0207$) was notably lower than that for Urmia, indicating that the Salmas dataset enabled more flexible learning due to its lower noise level. The best cross-validated RMSE was 0.01086, demonstrating excellent predictive accuracy and a strong match with observed values. Overall, the BO-ELM model provides robust and reliable predictions for both aquifers, with particularly high accuracy in the Salmas Plain.

In this study, the main objective is to enhance the conventional GALDIT model, which is expert-based and unsupervised, by leveraging various AI models. The conventional GALDIT approach relies on expert judgment for ranking and weighting its parameters. As an unsupervised method, it does not produce outputs that can be directly validated against observed data. In this work, EC values—representing groundwater salinity—were used to evaluate how well the AI-based models reproduce actual aquifer conditions. The second aim is to systematically compare the performance of the conventional GALDIT framework with the AI-based approaches. The results for each plain are presented below, demonstrating a clear improvement in predictive accuracy when AI models are applied.

In the Urmia Plain, the GALDIT index provides only a broad vulnerability pattern and does not accurately capture the observed salinity distribution, as expected given its unsupervised, parameter-driven nature. In contrast, the AI models show substantial improvement. For example, the DFFNN model achieves CC values of 0.87 and DC values of 0.75, indicating a much closer agreement with measured EC (Table 3). The BO-ELM model reaches an RMSE of 0.082 (Table 4), indicating its ability to learn the nonlinear patterns in the dataset. The Integrated Ensemble model yields the best overall performance, with DC values near 1 and significantly lower RMSE, confirming the effectiveness of combining multiple AI techniques.

A similar pattern is seen in the Salmas Plain. Although GALDIT identifies general vulnerability zones, it does not reflect the detailed EC variations observed in the aquifer. By contrast, the AI models provide highly accurate estimates. The DFFNN model achieves CC values close to 0.97 and DC above 0.9, while RMSE values are very low (Tables 3 and 4). The Integrated Ensemble model again performs exceptionally well, underscoring that data-driven approaches surpass expert-based methods in reproducing the spatial behavior of salinity.

Figures 8 and 9 provide a visual comparison of the GALDIT vulnerability patterns and the spatial distribution of EC, including both observed values and AI model predictions. These figures help illustrate how the expert-

based GALDIT index contrasts with, and in many areas diverges from, the detailed salinity patterns captured in the field and reproduced by the data-driven models. The present research recommends avoiding the disposal of contaminants in regions with medium to high vulnerability, particularly areas exposed to industrial effluents and agricultural waste. Maintaining consistent EC monitoring in these zones will further support reliable model calibration and sustainable groundwater management.

4. Conclusions

This research used the GALDIT method to evaluate the risk of SWI within the Salmas and Urmia aquifers. However, the method relies on expert judgment for rating and ranking parameters, which can introduce subjectivity and bias into the vulnerability assessment. Additionally, reliance on expert knowledge and variations across different geographic regions due to aquifer characteristics and morphology further limit its effectiveness. Also, incorporating a defined target variable for salinity modeling can help improve the expert-dependent nature of the conventional GALDIT method, allowing for more accurate modeling of coastal aquifer vulnerability. For those reasons, to improve modeling performance and achieve more reliable vulnerability assessments, AI-based models (FFNN and DFFNN) were utilized in this study. AI models can reduce subjectivity by learning from large datasets, minimizing the need for expert judgment in rating and ranking parameters. While AI models also require significant data, they can better leverage available data by identifying patterns and relationships that might not be evident with index-based methods.

The vulnerability index, derived from the GALDIT model, in the Salmas and Urmia aquifers, ranged from 4.3 to 7.8 and 3.8 to 6.8, respectively. It is worth mentioning that the difference in the GALDIT vulnerability index ranges for the Urmia and Salmas aquifers is due to hydrogeological variations between the two regions. The Salmas aquifer exhibits higher hydraulic conductivity (0.8–96 m/day) than the Urmia aquifer (0.03–46 m/day), leading to greater vulnerability to seawater intrusion. Additionally, Salt Flats and gypsum deposits in the Urmia region, along with its proximity to Lake Urmia, contribute to the observed differences in vulnerability classification. In contrast, the Salmas aquifer has relatively lower salinity pressure from the lake. However, it is still vulnerable due to intensive groundwater extraction, which alters the hydraulic gradient and facilitates inland saltwater migration. It is worth noting that proximity to Lake Urmia contributes to increased groundwater salinity in adjacent aquifers through multiple hydrogeological mechanisms. The highly permeable sediments at the river–lake junctions of Lake Urmia, combined with their high hydraulic conductivity and low hydraulic gradient, allow saline lake water to penetrate into the coastal aquifer more readily, thereby intensifying groundwater salinization [64].

Consequently, in this study, a comprehensive approach utilizing multiple artificial intelligence models—including FFNN, DFFNN, ANFIS, SVR, Ensemble, and Bo-ELM, was developed to assess groundwater vulnerability in the target area. Unlike the conventional GALDIT index, which is typically unsupervised and heavily reliant on expert judgment, this AI-based modeling framework was designed to provide both accurate predictions and an objective, data-driven assessment of aquifer susceptibility. The GALDIT parameters served as inputs across all models, while EC values were designated as outputs, given their effectiveness as indicators of water salinity and saltwater intrusion risk. Each model was selected for its unique strengths: FFNN and DFFNN for capturing nonlinear relationships; ANFIS for combining fuzzy logic with adaptive learning; SVR for robust regression under limited data; Ensemble methods for improving prediction stability and generalization; and Bo-ELM for fast training and high efficiency in modeling complex patterns. This integrated approach enables a more reliable and systematic evaluation of aquifer vulnerability, bridging the gap between expert-based assessments and predictive, data-driven modeling.

The conventional GALDIT index showed limited predictive performance in the Urmia plain, with high RMSE and very low DC and CC, reflecting its unsupervised, expert-based nature. In comparison, all AI-based models significantly improved prediction accuracy. For example, FFNN, DFFNN, and ANFIS reduced RMSE by roughly 63–71%, with DC values rising from near zero to above 0.6, highlighting their ability to capture nonlinear relationships. SVR also performed well with a notable improvement in correlation (CC = 0.895). Bo-ELM provided fast, efficient predictions, achieving about 77% lower RMSE than GALDIT. Finally, the Integrated Ensemble model outperformed all individual models, reducing RMSE by over 95% and achieving near-perfect DC, demonstrating the power of combining multiple models to enhance robustness and predictive reliability.

A similar trend was observed in the Salmas plain, where performance of GALDIT was again limited. AI-based models substantially improved predictions: FFNN, DFFNN, and ANFIS reduced RMSE by approximately 74–93%, with DC rising to above 0.82. SVR provided strong regression performance, while Bo-ELM and Integrated Ensemble achieved the highest accuracy, with DC exceeding 0.99 and CC approaching 1.0. These

results illustrate that ensemble and hybrid approaches offer the most reliable predictions by effectively integrating nonlinear, temporal, and spatial information.

Across both Urmia and Salmas plains, AI-based models clearly outperformed the conventional GALDIT index in predicting coastal aquifer vulnerability. The stem from their ability to model complex nonlinear relationships, capture temporal and spatial dynamics, and provide objective, data-driven predictions, thereby reducing reliance on subjective expert judgment. Among AI approaches, ensemble and hybrid models consistently deliver the highest predictive accuracy. In contrast, individual models such as SVR, and Bo-ELM also offer substantial improvements, making AI-based methods a robust and systematic framework for assessing aquifer susceptibility to saltwater intrusion.

Future research should focus on developing more robust and adaptive indices that integrate both conventional approaches, like GALDIT, and advanced AI-based modeling frameworks. Specifically, leveraging hybrid and recurrent neural networks, as well as ensemble learning techniques, can further enhance predictive accuracy and capture the complex nonlinear, spatial, and temporal interactions in coastal aquifer systems. Additionally, incorporating real-time monitoring data and uncertainty analysis into AI models could improve their generalizability and operational applicability. In summary, building on the findings of this study, future efforts should aim to develop integrated, data-driven frameworks that not only improve predictive performance but also support sustainable groundwater management and mitigate saltwater intrusion in vulnerable regions.

Author Contributions

V.N.: Conceptualization, supervision, methodology, writing—review and editing. E.B.K.: Conceptualization, data collection and analysis, visualization, methodology, writing—original draft. Sana Maleki: Formal analysis, methodology, writing—original draft. N.J.P.: Formal analysis, writing—original draft. E.S.: Validation, writing—review and editing. All authors were equally involved in the conceptualization and design of the study, and in writing the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The data used in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

Generative AI tools were not used. All scientific content, plots, analysis, interpretation, and conclusions were developed exclusively by the authors.

References

1. Carreira, P.M.; Marques, J.M.; Nunes, D. Source of groundwater salinity in coastline aquifers based on environmental isotopes (Portugal): Natural vs. human interference. A review and reinterpretation. *Appl. Geochem.* **2014**, *41*, 163–175.
2. Javadi, S.; Kardan Moghaddam, H.; Neshat, A. A new approach for vulnerability assessment of coastal aquifers using combined index. *Geocarto Int.* **2022**, *37*, 1681–1703.
3. Aslam, R.A.; Shrestha, S.; Pandey, V.P. Groundwater vulnerability to climate change: A review of the assessment methodology. *Sci. Total Environ.* **2018**, *612*, 853–875.
4. Machiwal, D.; Jha, M.K.; Singh, V.P.; et al. Assessment and mapping of groundwater vulnerability to pollution: Current status and challenges. *Earth-Sci. Rev.* **2018**, *185*, 901–927.

5. Shirazi, S.M.; Imran, H.M.; Akib, S.; et al. Groundwater vulnerability assessment in the Melaka State of Malaysia using DRASTIC and GIS techniques. *Environ. Earth Sci.* **2013**, *70*, 2293–2304.
6. Liggett, J.E.; Talwar, S. Groundwater Vulnerability Assessments and Integrated Water Resource Management. *Streamline Watershed Manag. Bull.* **2009**, *13*, 18–29.
7. Taghavi, N.; Niven, R.K.; Paull, D.J.; et al. Groundwater vulnerability assessment: A review including new statistical and hybrid methods. *Sci. Total Environ.* **2022**, *822*, 153486.
8. Mogaji, K.A.; Lim, H.S.; Abdullah, K. Modeling groundwater vulnerability prediction using geographic information system (GIS)-based ordered weighted average (OWA) method and DRASTIC model theory hybrid approach. *Arab. J. Geosci.* **2014**, *7*, 5409–5429.
9. Nourani, V.; Khajeh, E.B.; Paknezhad, N.J.; et al. Temporal evaluation of seawater intrusion vulnerability in Shabestar Plain using GALDIT and AI techniques. *Environ. Sci. Pollut. Res.* **2025**, *32*, 10855–10876.
10. Aller, L.; Bennett, T.; Lehr, J.H.; et al. DRASTIC: A standardized system for evaluating groundwater pollution potential using hydrogeologic setting. *J. Geol. Soc. India* **1987**, *29*, 23–37.
11. Chachadi, A.G.; Lobo Ferreira, J.P.C. Sea Water Intrusion Vulnerability Mapping of Aquifers Using the GALDIT Method. 2001. Available online: https://repositorio.lnec.pt/jspui/handle/123456789/5798 (accessed on 12 September 2025).
12. Civita, M. *Le carte Della Vulnerabilità Degli Acquiferi All'inquinamento*; Teoria e pratica; Pitagora Press: Bologna, Italy, 1994; p. 325.
13. Stempvoort, D.V.; Ewert, L.; Wassenaar, L. Aquifer vulnerability Index: A GIS-compatible method for groundwater vulnerability mapping. *Can. Water Resour. J.* **1993**, *18*, 25–37.
14. Emara, S.R.; Armanuos, A.M.; Shalby, A. Appraisal seawater intrusion vulnerability for the Moghra coastal aquifer, Egypt–application of the GALDIT index, sensitivity analysis, and hydro-chemical indicators. *Groundw. Sustain. Dev.* **2024**, *25*, 101166.
15. Chang, S.W.; Chung, I.M.; Kim, M.G.; et al. Application of GALDIT in assessing the seawater intrusion vulnerability of Jeju Island, South Korea. *Water* **2019**, *11*, 1824.
16. Ghorai, D.; Bhunia, G.S.; Shit, P.K. Coastal aquifer vulnerability for saltwater intrusion: A case study of Chennai Coast using GALDIT model and geoinformatics. In *Groundwater and Society*; Shit, P.K., Bhunia, G.S., Adhikary, P.P., et al., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 349–362. https://doi.org/10.1007/978-3-030-64136-8_16.
17. Pisciotta, A.; Suozzi, E.; Tiwari, A.K. A modified GALDIT-NUTS index to assess Favignana Island aquifer vulnerability. *Geocarto Int.* **2022**, *37*, 11706–11731.
18. Idowu, T.E.; Jepkosgei, C.; Nyadawa, M.; et al. Integrated seawater intrusion and groundwater quality assessment of a coastal aquifer: GALDIT, geospatial and analytical approaches. *Environ. Sci. Pollut. Res.* **2022**, *29*, 36699–36720.
19. Nguyen, A.H.; Pham, K.Q.; Le, Q.H. Assessment of Pleistocene Aquifer Vulnerability to Saline Intrusion in the Coastal Region of Ba Ria-Vung Tau Province Using GIS and Entropy-GALDIT. *Sustainability.* **2023**, *15*, 8107.
20. Goswami, S.; Rai, A.K. Identifying intrusion of seawater in coastal aquifers by modified GALDIT (M-GALDIT) index. *Groundw. Sustain. Dev.* **2024**, *25*, 101173.
21. Nourani, V.; Najafi, H.; Maleki, S.; et al. Z-number based assessment of groundwater vulnerability to seawater intrusion. *J. Hydrol.* **2024**, *632*, 130859.
22. Torkashvand, M.; Neshat, A.; Javadi, S.; et al. New hybrid evolutionary algorithm for optimizing index-based groundwater vulnerability assessment method. *J. Hydrol.* **2021**, *598*, 126446.
23. Moazamnia, M.; Hassanzadeh, Y.; Nadiri, A.A.; et al. Vulnerability indexing to saltwater intrusion from models at two levels using artificial intelligence multiple model (AIMM). *J. Environ. Manag.* **2020**, *255*, 109871.
24. Barzegar, R.; Razzagh, S.; Quilty, J.; et al. Improving GALDIT-based groundwater vulnerability predictive mapping using coupled resampling algorithms and machine learning models. *J. Hydrol.* 2021 Jul; 598:126370.
25. Bordbar, M.; Neshat, A.; Javadi, S.; et al. Improving the coastal aquifers' vulnerability assessment using SCMAI ensemble of three machine learning approaches. *Nat. Hazards* **2022**, *110*, 1799–1820.
26. Gharekhani, M.; Nadiri, A.A.; Khatibi, R.; et al. A study of uncertainties in groundwater vulnerability modelling using Bayesian model averaging (BMA). *J. Environ. Manag.* **2022**, *303*, 114168.
27. Aryafar, A.; Khosravi, V.; Zarepourfard, H.; et al. Evolving genetic programming and other AI-based models for estimating groundwater quality parameters of the Khezri plain, Eastern Iran. *Environ. Earth Sci.* **2019**, *78*, 69.
28. Sadikoglu, F.; Nourani, V.; Maleki, S.; et al. Application of Artificial Neural Network to Improve DRASTIC-Based Groundwater Vulnerability Assessment. In Proceedings of the 12th World Conference "Intelligent System for Industrial Automation" (WCIS-2022), Tashkent, Uzbekistan, 25–26 November 2022; Volume 718, pp. 273–281. https://doi.org/10.1007/978-3-031-51521-7_35.
29. Shahbazi, A.; Aydin, Y.; Semiz, G.D.; et al. The compounding effects of agricultural expansion and snow drought on Lake Urmia's drying crisis. *Sci. Rep.* **2025**, *15*, 38132.

30. Chachadi, A.G.; Lobo Ferreira, J.P. Assessing aquifer vulnerability to seawater intrusion using the GALDIT method: Part 2—GALDIT indicators description. In Proceedings of the Fourth Inter-Celtic Colloquium on Hydrology and Management of Water Resources, Guimarães, Portugal, 11–14 July 2005. Available: https://www.aprh.pt/celtico/PAPERS/26.PDF (accessed on 22 November 2025).
31. Fakhri, M.; Moghaddam, A.A.; Nadiri, A.A.; et al. Incorporating hydraulic gradient and pumping rate into GALDIT framework for salinity hazard assessment in coastal aquifers: A case study of Urmia Plain, Iran. *Research Square* **2024**, https://doi.org/10.21203/rs.3.rs-4186756/v.
32. Ghorbani, A.; Abbas Novinpour, E.; Ahangari, M. Identification of land subsidence areas in Salmas Plain using GARDLIF framework and learning machines. *Hydrogeology* **2025**, in press. https://doi.org/10.22034/hydro.2025.56975.1296.
33. Hemmati, F.; Khanri, S.; Alizadeh, A. Assessment of land subsidence variations in the Urmia Plain aquifer using differential interferometric synthetic aperture radar (DInSAR). *J. Hydrogeomorphol.* **2025**. https://doi.org/10.22034/hyd.2025.67902.1798.
34. Sheikholeslami, R.; Jahangiri, F. An uncertainty-informed water quality index: Incorporation of data uncertainty into water quality assessment. *Environ. Model. Softw.* **2026**, *196*, 106760.
35. Sundaram, V.L.K. Vulnerability assessment of seawater intrusion and effect of artificial recharge in Pondicherry coastal region using GIS. *Indian. J. Sci. Technol.* **2008**, *1*, 1–7.
36. Gupta, T.K.; Raza, K. Optimizing Deep Feedforward Neural Network Architecture: A Tabu Search Based Approach. *Neural Process Lett.* **2020**, *51*, 2855–2870.
37. Nourani, V. Investigating the Ability of Artificial Neural Network (ANN) Models to Estimate Missing Rain-gauge Data. *J. Environ. Inform.* **2012**, *19*, 38–50.
38. Nourani, V.; Sayyah Fard, M. Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes. *Adv. Eng. Softw.* **2012**, *47*, 127–146.
39. Sharghi, E.; Nourani, V.; Najafi, H.; et al. Emotional ANN (EANN) and Wavelet-ANN (WANN) approaches for Markovian and seasonal based modeling of rainfall-runoff process. *Water Resour. Manag.* **2018**, *32*, 3441–3456.
40. Fijani, E.; Nadiri, A.A.; Asghari Moghaddam, A.; et al. Optimization of DRASTIC method by supervised committee machine artificial intelligence to assess groundwater vulnerability for Maragheh–Bonab plain aquifer, Iran. *J. Hydrol.* **2013**, *503*, 89–100.
41. Torres, J.F.; Hadjout, D.; Sebaa, A.; et al. Deep learning for time series forecasting: A survey. *Big Data.* **2021**, *9*, 3–21.
42. Ketkar, N. *Deep Learning with Python*; Springer: Berkeley, CA, USA, 2017.
43. Lahmiri, S.; Bekiros, S. Deep Learning Forecasting in Cryptocurrency High-Frequency Trading. *Cogn. Comput.* **2021**, *13*, 485–487.
44. Menares, C.; Perez, P.; Parraguez, S.; et al. Forecasting PM2.5 levels in Santiago de Chile using deep learning neural networks. *Urban. Clim.* **2021**, *38*, 100906.
45. Amirshahi, B.; Lahmiri, S. Hybrid deep learning and GARCH-family models for forecasting volatility of cryptocurrencies. *Mach. Learn. Appl.* **2023**, *12*, 100465.
46. Behfar, N.; Booij, M.J.; Nourani, V. Assessing rainfall-runoff models for climate change: Simple and differential split-sample tests for conceptual and artificial intelligence models. *Hydrol. Sci. J.* **2024**, *69*, 861–877.
47. Jang, J.S.R. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man. Cybern.* **1993**, *23*, 665–685.
48. Jang, J.S.R.; Sun, C.T.; Mizutani, E. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*; Prentice Hall: Upper Saddle River, NJ, USA, 1997.
49. Vapnik, V.N. Statistical learning theory. New York: Wiley; 1998.
50. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222.
51. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297.
52. Sharghi, E.; Nourani, V.; Behfar, N. Earthfill dam seepage analysis using ensemble artificial intelligence based modeling. *J. Hydroinform.* **2018**, *20*, 1071–1084.
53. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501.
54. Huang, G.B.; Zhou, H.; Ding, X.; et al. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man. Cybern. B Cybern.* **2012**, *42*, 513–529.
55. Mockus, J. On Bayesian methods for seeking the extremum. In *IFIP Technical Conference on Optimization Techniques*; Springer: Berlin/Heidelberg, Germany, 1974; pp. 400–404.
56. Shahriari, B.; Swersky, K.; Wang, Z.; et al. Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE.* **2016**, *104*, 148–175.
57. Doveton, J.H. A Course on Log Analysis: Remote Sensing in the Subsurface. *J. Geol. Educ.* **1978**, *26*, 22–26.
58. Amoozegar, A.; Warrick, A.W. Hydraulic Conductivity of Saturated Soils: Field Methods. In *SSSA Book Series*; Klute, A., Ed.; Soil Science Society of America: Madison, WI, USA; American Society of Agronomy: Madison, WI, USA, 2018; pp. 735–770. https://doi.org/10.2136/sssabookser5.1.2ed.c29.

59. Boulton, N. The drawdown of the water-table under non-steady conditions near a pumped well in an unconfined formation. *J. Proc. Inst. Civ. Eng.* **1954**, *3*, 564–579.
60. Huan, H.; Wang, J.; Teng, Y. Assessment and validation of groundwater vulnerability to nitrate based on a modified DRASTIC model: A case study in Jilin City of northeast China. *Sci. Total Environ.* **2012**, *440*, 14–23.
61. Nakhaei, M.; Vadiati, M.; Mohammadi, K. Evaluation of vulnerability of Urmia Lake saline water intrusion to coastal aquifer using GALDIT model. *Geosciences* **2015**, *24*, 51–64.
62. Azizi Mobaser, J.; Masud Lak, M.; Rasoulzadeh, A. Evaluation of intrinsic vulnerability of Urmia plain groundwater pollution using original DRASTIC and DRASTIC modified models. *Iran-Water Resour. Res.* **2019**, *14*, 220–235.
63. Maleki, S.; Nourani, V.; Najafi, H.; et al. Z-numbers based novel method for assessing groundwater specific vulnerability. *Eng. Appl. Artif. Intell.* **2023**, *122*, 106104.
64. Amiri, V.; Nakhaei, M.; Lak, R.; et al. Geophysical, isotopic, and hydrogeochemical tools to identify potential impacts on coastal groundwater resources from Urmia hypersaline Lake, NW Iran. *Environ. Sci. Pollut. Res.* **2016**, *23*, 16738–16760.