

Received July 25, 2020, accepted August 4, 2020, date of publication August 7, 2020, date of current version August 19, 2020. Digital Object Identifier 10.1109/ACCESS.2020.3014908

Groundwater Salinity Susceptibility Mapping Using Classifier Ensemble and Bayesian Machine Learning Models

AMIRHOSEIN MOSAVI^{®1,2}, FARZANEH SAJEDI HOSSEINI^{®3}, BAHRAM CHOUBIN^{®4}, MASSOUD GOODARZI⁵, AND ADRIENN A. DINEVA^{6,7}

¹Environmental Quality, Atmospheric Science and Climate Change Research Group, Ton Duc Thang University, Ho Chi Minh City, Vietnam

²Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City, Vietnam

³Reclamation of Arid and Mountainous Regions Department, Faculty of Natural Resources, University of Tehran, Karaj 31585-77871, Iran
⁴West Azarbaijan Agricultural and Natural Resources Research and Education Center, Soil Conservation and Watershed Management Research Department,

⁵Soil Conservation and Watershed Management Research Institute (SCWMRI), AREEO, Tehran 13898-17611, Iran

⁶Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

⁷Kalman Kando Faculty of Electrical Engineering, Obuda University, 1034 Budapest, Hungary

Corresponding authors: Bahram Choubin (b.choubin@areeo.ac.ir) and Adrienn A. Dineva (adrienndineva@duytan.edu.vn)

ABSTRACT Risk and susceptibility mapping of groundwater salinity (GWS) are challenging tasks for groundwater quality monitoring and management. Advancement of accurate prediction systems is essential for the identification of vulnerable areas in order to raise awareness about the potential salinity susceptibility and protect the groundwater and top-soil in due time. In this study, three machine learning models of Stochastic Gradient Boosting (StoGB), Rotation Forest (RotFor), and Bayesian Generalized Linear Model (Bayesglm) are developed for building prediction models and their performance evaluated in the delineation of salinity susceptibility maps. Both natural and human effective factors (16 features) were used as predictors for groundwater salinity modeling and were randomly divided into the training (80%) and testing (20%) datasets. The models were evaluated using testing datasets after calibration using the selected features by recursive feature elimination (RFE) method. The RFE indicated that modeling with 8 features had better performance among 1 to 16 features (Accuracy = 0.87). Results of the groundwater salinity prediction highlighted that StoGB had a good performance, whereas the RotFor and Bayesglm had an excellent performance based on the Kappa values (> 0.85). Although spatial prediction of the models was different, all of the models indicated that central parts of the region have a very high susceptibility which matches with agricultural areas, lithology map, the locations with low depth to groundwater, low slope, and elevation. Additionally, areas near to the Maharlu lake and locations with a high decline in groundwater are also located in the very high susceptibility zone, which can confirm the effects of saltwater intrusion. The susceptibility maps produced in this study are of utmost importance for water security and sustainable agriculture.

INDEX TERMS Groundwater salinity, hazard, recursive feature elimination, stochastic gradient boosting, rotation forest, Bayesian generalized linear model.

I. INTRODUCTION

In arid areas, the essentials of life, economic growth, agriculture, and industries highly rely on groundwater as essential natural storage to make up against a shortage of surface water [1]–[4]. While the climate change, rapid urban development, unsewered disposal, constructions, mining, and other industrial pollutants are constantly threatening the quality

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Tucci¹⁰.

and quantity of groundwater; generally the oil and gas production, as well as agricultural practices of overusing pesticides and fertilizers more than any anthropogenic activities, are contributing to the groundwater contamination and scarcity [5]–[16]. Various salts (including nontoxic and toxic ions), minerals, heavy metals, microorganisms, pathogens, petroleum pollutants, and nutrients are among the major contaminators [17]–[22]. The salinity of groundwater may have various origins of natural and man-made causes [23]. As one of the main natural causes, the rock weathering

AREEO, Urmia 57169-63963, Iran

phenomenon realizes minerals, and over time, it is washed to the groundwater causing salinity [24], [25]. However, during the past decades, the excessive withdrawal of groundwater using numerous deep wells for irrigation purposes has been part of the major contributor to the rise of unnatural salinity level [26] and also saltwater intrusion (SWI) [27] resulting in a depletion of groundwater resources for sustainable and quality irrigation which has become a major environmental concern [24], [28]. Furthermore, irrigation using water with a high level of groundwater salinity (GWS) can permanently damage the top-soil through continuous salinization of the land surface [29]–[32]. Thus, risk mapping and salinity susceptibility assessment of the groundwater in the early stages is very important to protect soil, groundwater, and crop production [33]–[35].

Producing risk maps and susceptibility prediction models, particularly in arid areas, is essential for effective groundwater quality monitoring and management so as to regulate and enforce policies in due time. As the recovery from salinity and other contaminations may take a long time (due to the slow movement of groundwater), the environmental protection policies have been focused on prevention [36]. Delineation of high-risk zones, mapping vulnerable areas to salinity, and hazard prediction have become essential for sustainable land use planning, agricultural strategies and crop selection, water supply management, water security, and policy-making [37]–[42]. Literature includes various scenarios of risk maps and hazard prediction models used for informed policy-making in arid areas around the world [43]–[50].

Physically-based salinity modeling includes the geochemical, hydrogeochemical, hydro-chemical as well as solute transport modeling tools which are often used for identifying the source and dynamics of salinity and risk assessment of groundwater through considering isotopic or non-isotopic geochemical and physicochemical parameters of aquifer [51]-[55]. Nassery and Kayhomayoon [56] and Stephens et al. [57] present models based on geophysical log analysis and aqueous geochemical measurements for salinity mapping in the arid area of Isfahan, Iran, and California, USA, respectively. Geochemical and geostatistical models integrated with GIS-based mapping have recently been used for mapping, monitoring the salinity, and assessment of the groundwater governance [51], [55], [58], [59]. Computational hydrologic modeling platforms, e.g. DRAIN-MOD [60], have been widely used to predict salinity of drainage water used in irrigation [61]. A number of numerical models and finite element analyses have also been used for salinity prediction, e.g., SUTRA, RMA-10 [62]. Statistical based models such as multivariate statistical analysis [63], [64], cluster analysis and multiple linear regressions (MLR) [59], [65]-[67], linear discriminant analysis [68], transient modeling [69], Bayesian kriging [70] have been used to study groundwater salinity and its origin. Sofiyan Abuelaish and Camacho Olmedo [71] developed a statistical model based on linear and multiple regression analysis to predict the 5-year ahead salinity value in the arid and zoning map of salinity in the arid areas of Iran. Time series analysis for modeling of groundwater salinity can be found in the literature [74], [75]. For instance, Cook et al. [76] predicted the groundwater salinity using a distributed parameter recharge model in the SW Murray Basin. They used variables of depth to groundwater, recharge, water content of the unsaturated zone, and soil salinity as input data. Gundogdu and Akkaya Aslan [77] in a similar case predicted the salinity maps. However, time-series prediction for salinity is associated with a number of drawbacks including accuracy and generalization [78], [79]. Artificial intelligence (AI) methods and applications have already been reported as efficient and popular methods in groundwater modeling. However, among the AI methods, a limited number of soft computing models are available in literature for modeling the salinity hazard or irrigation risk mapping, e.g. Abdi et al. [80] proposed a fuzzy logic model and Attwa et al. [81] genetic algorithm (GA) in the arid areas of Iran and Egypt respectively. The use of machine learning methods also has been limited in a few modeling experiments even though they appeared more convenient and performed with higher accuracy, better generalization ability, more reliability, and efficiency compared to the statistical, numerical, and physical models. For instance, Alagha et al. [82] used artificial neural networks (ANNs) and support vector machine (SVM) to model salinization processes with higher performance and more simplicity compared to statistical methods. Yu et al. [53] also reported machine learning models of back propagation ANN (BP-ANN) and neuro-fuzzy (NF) superior to the conventional linear models in the prediction of groundwater salinity hazards. Although applications of machine learning methods in various realms of hydrological modeling have already been established, they have not received adequate attention for groundwater modeling, particularly for salinity mapping and hazard prediction. Thus, the literature misses the application of various machine learning methods for modeling spatial groundwater salinity and the evaluation studies on their performance and accuracy.

areas of Gaza, Palestine. Amiri-Bourkhani et al. [72] and

Ashrafzadeh et al. [73] used geostatistical data and krig-

ing methods to develop predictive models for probabilistic

This study in order to fill an existing gap in literature aimed to predict groundwater salinity probability for understanding susceptible areas and supporting manager communities to adopt the best decisions. Therefore, the main purposes of the study were: (i) to identify key variables related to the groundwater salinity, (ii) to predict the spatial probability of the groundwater salinity, (iii) to compare the performance of machine learning models in prediction of the groundwater salinity probability, and (v) to determine susceptible regions of the study area in view of the groundwater salinity.

II. MATERIAL AND METHODS

A. DESCRIPTION OF STUDY AREA

Sarvestan plain, which is a coastal plain in the south of Maharlu lake in the Fars province, Iran, is extended between latitudes of 29° 1' 11" to 29° 26' 52" N and longitudes of 52° 43' 45" to 53° 27' 49" E and was selected as the study area (Fig. 1). The area of the region is about 1688.2 square kilometers. Long-term yearly precipitation is about 280 mm in which usually summers are without precipitation, while yearly evaporation is about 2800 mm. Temperature averagely varies between 6.3 °C to 29.5 °C respectively for daily minimum and maximum temperatures. The climate of the study area is mostly semiarid-cold according to the extended-De Martonne method [83], nevertheless, some areas mostly around the Maharlu lake are semiarid-moderate. Land uses of the study area mostly are rangeland and agricultural areas (respectively about 65.7 % and 28.5 % of the study area), while other land uses comprise approximately 5.8 % of the study area (Fig. 2n). The main water source for irrigation purposes is groundwater. Due to the importance of the groundwater in this area, and because of the presence of salty formations [84] and existence of the Maharlu lake, a saline lake in Iran [85] located in the north of the study (Fig. 1), the groundwater can be degraded through dissolution of halite minerals and intrusion of saltwater owing to groundwater pumping [86].



FIGURE 1. Location of the study area.

B. GROUNDWATER SALINITY DATA

Salinity is the total amount of inorganic solid material dissolved in water and can be calculated directly by measuring the presence of various salts in water [87], [88]. But there is a feasible and indirect way to express salinity; by assessing the capability of the water to conduct an electrical current [56]. Therefore, the real and best way to express salinity is assessing the electrical conductivity (EC) of the water, which is a good measure of salinity susceptibility as it reflects the total dissolved solids [89], [90]. In this study, the researchers have collected EC data from 2003 to 2017 from the Iranian Water Resources Management Company (IWRMC). There were a number of 83 wells in the study area which have measured the groundwater quality data (Fig. 1). The data did not have an appropriate distribution for months or in seasons, so a yearly average value of EC for wells during the period of 2003 to 2017 was calculated. According to the literature, values of EC more than 2250 μ S/cm highlight an unsuitable water quality and it is even not suitable for irrigation [89]. Therefore, a threshold equal to 2250 μ S/cm for EC was considered to separate saline and non-saline wells and assignment of the values 0 and 1 to model groundwater salinity. The mean yearly EC data (from 2003 to 2017) for 83 monitoring wells varies from 736.4 to 70257.5 μ S/cm. Due to the high level of EC in the study area, understanding susceptible and safe zones in view of groundwater salinity could help managers to have better planning for groundwater utilization and assignments, as well as to control susceptible areas.

C. GROUNDWATER SALINITY CONDITIONING FACTORS (GSCF)

The origin of the groundwater salinity is because of the co-existence of both natural factors and human activities [54], [81]. According to a survey conducted by researchers within literature, e.g. [88], [91], both natural and human effective factors on groundwater salinity including topographic factors, topographic wetness index (TWI), distance from stream (DFS), distance from lake (DFL), distance from fault (DFF), depth of groundwater (DTGW), groundwater withdrawal (GWW), decline of groundwater level (DGWL), evaporation, precipitation, land use, lithology, and soil type of the study area were considered as GSCF in this study.

Topographic factors such as elevation, slope, aspect, and curvature (Fig. 2a,b,c,d) are effective in flushing and exporting the salty materials from the soil into fluvial plains, and transporting and accumulating it into lowlands. An ALOS PALSAR Digital Elevation Model (DEM) 12.5 \times 12.5 m (https://vertex.daac.asf.alaska.edu/) was used to extract topographic factors. TWI (Fig. 2e) and DFS (Fig. 2f) can affect the groundwater through waterlogging and infiltration of surface water, leaching and dissolving of solid materials within the unsaturated zone into groundwater [92]. Also, DFL (Fig. 2g) as an index indicating proximity to a lake for considering effects of the intrusion probability of saline water from the lake was considered. Existence of fault in the study area can affect the quality of the groundwater through conjunction with surface water, so, DFF (Fig. 2h) was considered. Layers of DFS, DFL, and DFF were calculated using the Euclidean tool in ArcGIS. Also, TWI was calculated using the SAGA GIS software.

From the groundwater, three indices DTGW (Fig. 2i), GWW (Fig. 2j), and DGWL (Fig. 2k) were calculated. DTGW (also called water-table depth) could be affected by the evaporation, irrigation water, and conjunction with rivers [88]. Also, groundwater salinity can be increased using GWW and DGWL [93], this increases the potential for saline water intrusion into the groundwater. The groundwater data for each well were received from the IWRMC, and a yearly average map from 2003 to 2017 was produced in the ArcGIS software.

Climate data such as evaporation and precipitation are the natural factors that have effects on the groundwater. Precipitation can cause groundwater salinity through



🧀 Entisols

FIGURE 2. Groundwater salinity conditioning factors: (a) elevation, (b) slope, (c) aspect, (d) curvature, (e) topographic wetness index (TWI), (f) distance from stream (DFS), (g) distance from lake (DFL), (h) distance from fault (DFF), (i) depth to groundwater (DTGW), (j) groundwater withdrawal (GWW), (k) decline of groundwater level (DGWL), (l) evaporation, (m) precipitation, (n) landuse, (o) lithology, and (p) soil type.

recharge, infiltration, and dissolving materials within the soil [91]. Also, evaporation can affect the groundwater because some parts of the study have low DTGW and were near to the earth's surface (about 3 m, Fig. 2i). The evaporation and precipitation data were received from IWRMC and a yearly average map from 2003 to 2017 for them was produced using the IDW method (Fig. 21 and m). Landuse such as domestic and industrial wastewaters, agricultural fertilizers, and irrigation water can be affected the groundwater salinity [33], [40], [52], [56]. Soil type and lithology, directly and indirectly, are effective in groundwater salinity. Direct effects include the presence of salts and evaporative rocks in the saturated and unsaturated zones, as well as water-rock interactions [91], and indirect effects are an influence on recharge and infiltration rates. Landuse, lithology, and soil type maps of the study area were received from IWRMC.

D. FEATURE SELECTION

Although features considered in this study (GSCF, Fig. 2) were based on a survey and influences on the groundwater, however, the existence of the redundant features could affect results. On the other hand, various combinations of features would have different results. So, a successful feature selection method was considered to identify key features. In this study, recursive feature elimination (RFE) [94] method was used to find and select the key feature combinations among the GSCF that was used in this study.

RFE is a machine learning technique that builds a model according to the predictors' subsets and computes them individually and then scores them through feature selection (FS). The model is reconstructed only based on the most important predictors and the less important predictors are eliminated. The number and the size of predictors' subsets can be adjusted as a mean for the model performance optimization. RFE has been widely used in model construction in various scientific fields. It has also been recently used for groundwater contamination prediction where it is used for identifying the best variable combination for each machine learning model [95]. However, the application of RFE has not been explored for salinity modeling and risk mapping. As the RFE is a wrapper, the random forest (RF) algorithm is used to train it [96]. Furthermore, selection of the best features was established upon the cross-validation with 5-fold and metric of accuracy using Caret R package [97]-[99].

E. MODELING APPROACHES

In this study, three machine learning models of probability prediction including stochastic gradient boosting, rotation forest, and Bayesian generalized linear model were employed to model groundwater salinity susceptibility:

Stochastic Gradient Boosting (StoGB): StoGB has been developed by Friedman [100] to improve the performance of his gradient boosting introduced earlier, through including randomness to the function estimation procedures. The randomness has been included using the adaptive bagging [101] for the least-squares fitting of additive expansions which

with the bagging learning. Further, at each boosting step, it substitutes random residuals for the ordinary ones. At each iteration, a subsample of the training data is drawn randomly from the training data. This randomly selected subsample is then used, instead of the full sample, to fit the base learner and compute the model update for the next iteration [100]. The model has four parameters including a number of trees (n.trees), maximum nodes per tree or number of leaves (interaction.d), learning rate (Shrinkage), and a minimum number of samples in tree terminal nodes (n.minobsinn), which are tuned within a 5-fold cross-validation procedure. The StoGB has been previously used for groundwater modeling for the purpose of prediction of the hydro-chemical composition of groundwater with promising results [102]. Further application of StoGB has been limited in hydrological modeling [103]. In this study, the R package of StoGB, which is known as gbm was used to implement this model. A more detailed description of the model is presented in Friedman [100].

replaces the base learner in gradient boosting procedures

Rotation Forest (RotFor): Rodríguez et al. [104] advanced the feature extraction machine learning method of RotFor for generating classifier ensembles. To handle the training data used for classifier ensembles, the feature set is randomly split into (K) number of subsets which are prepared by principal component analysis (PCA) [105] to retain the data variability. Then, to encourage simultaneous individual accuracy of the ensemble, equal to the number of subsets, axis rotations take place to form the new features for a base classifier. RotFor uses decision trees that perform well in showing a high level of sensitivity to the rotation of the feature axes. The RotFor has two parameters including the number of variable subsets (K) and Ensemble size (L). RotFor has been very recently used as the base classifier in spatial modeling of groundwater with promising results [106], and its popularity in other hydrological susceptibility mappings [107]-[109]. In this study, the model parameters are tuned within the 5-fold cross-validation procedure. The rotationForest R package [104] was applied to execute the RotFor. More description of the model can be found in Rodríguez et al. [104].

Bayesian Generalized Linear Model (Bayesglm): Generalized Linear Models (GLMs) [110] represents a group of modeling techniques e.g. Bayesian, logistic regression, through a general modeling approach to several useful distributions, including the gamma, Poisson, and binomial, suitable for a broad range of response modeling problems [111]. While modeling with GLM, the response can be specified with a higher degree of flexibility in the modeling, resulting in a more convenient structure. However, it needs parametrizations that are difficult to interpret and solutions that are not comprehensive. The observation model, linear predictor, and link function are the three building block components of GLM [110]. In a Bayesian GLM, or so-called Bayesglm, a normal prior is often chosen for the linear predictor's parameters. This conjugate prior to the normal linear model is a GLM with a Gaussian observation model and identity link

function. A normal prior resulting from expert opinion is therefore sought for the unknown coefficients in Bayesian GLMs. In this study, R's arm package [112] was applied to perform the Bayesglm. See Hosack *et al.* [111] and Hosseini *et al.* [113] for more details of the Bayesglm model.

F. MODEL CALIBRATION AND PERFORMANCE ESTIMATION

The database including predictors and predictand is randomly divided into the training (80 %) and testing (20 %) datasets. The models were evaluated using testing datasets after calibration using the selected features by RFE. The researchers used 5-fold cross-validation to optimize the models. The process was repeated 50 times and two key metrics of classification including Accuracy (Acc) and Kappa (K) were used to assess the models' performance by calculating a contingency table and hit and miss analysis. Accuracy indicates the percentage of correct classification. Kappa highlights the likelihood of agreement by chance using probabilities of the model classification [114]. This statistic is classified into five classes: poor (K < 0.4), moderate (0.4 < K < 0.55), good (0.55 < K < 0.85), excellent (0.85 < K < 0.99), and perfect (0.99 < K < 1.00); indicating the performance of the classification models [115]. These metrics are calculated as follows:

$$Acc = \frac{H + CN}{H + FA + M + CN} \tag{1}$$

where H, FA, M, and CN are calculated using the contingency table and respectively correspond with the number of hits, false alarms, misses, and correct negatives [116].

$$K = \frac{Acc - P_e}{1 - P_e} \tag{2}$$

where P_e is the expected probability of chance agreement [117], [118], which is calculated as follow:

$$P_e = \frac{(H + FA)(H + M) + (M + CN)(FA + CN)}{(H + FA + M + CN)^2}$$
(3)

III. RESULTS AND DISCUSSION

A. RESULTS OF FEATURE SELECTION

Results of the recursive feature elimination (RFE) method for determining the best number of features are reparented in Fig. 3. The accuracy metric was considered to evaluate the RFE performance. Each box indicates aggregated model performance for 5-fold cross validation during many times of model runs (about 700 runs). As can be seen, the RFE indicates that modeling with 8 features had better performance among 1 to 16 features (Accuracy is about 0.87) (Fig. 3). The average values of Accuracy and Kappa are represented in Table 1. The Accuracy and Kappa values for the number of 8 features are equal to 0.87 and 0.64, respectively, which are the highest in comparison with other numbers of features (Table 1).

Importance of the features was determined based on their occurrence in the model runs (Fig. 4). As can be seen from



Number of features

FIGURE 3. Performance of the recursive feature elimination method versus the number of features.

TABLE 1. Average values of the RFE performance.

Number of features	Accuracy	Kappa
1	0.74	0.39
2	0.80	0.54
3	0.81	0.56
4	0.76	0.42
5	0.77	0.45
6	0.81	0.57
7	0.80	0.57
8	0.87	0.64
9	0.80	0.54
10	0.80	0.49
11	0.78	0.47
12	0.81	0.56
13	0.78	0.45
14	0.78	0.43
15	0.80	0.49
16	0.80	0.58

Fig. 4, the depth of groundwater and elevation contributes as the input features for more than 96 % of the model runs. According to the RFE results, the number of 8 features must be selected for groundwater salinity modeling (Fig. 3). So, variables of depth to groundwater (DTGW), elevation, distance from fault (DFF), lithology, decline of groundwater level (DGWL), distance from lake (DFL), groundwater withdrawal (GWW), and distance from stream (DFS) were selected as key features respectively with occurrence frequency equal to 97.5%, 96.3%, 81.3%, 78.8%, 71.3%, 61.3%, 60%, and 50 % in all model runs (Fig. 4).

B. RESULTS OF MODEL EVALUATION

The StoGB model was optimized with 50 runs, and results indicated that the best value for learning rate (Shrinkage)



FIGURE 4. The frequency of features occurrence (%) in the RFE runs.

was equal to 0.1, also, 200 for the number of trees (n.trees), 3 for maximum nodes per tree (interaction.d), and 10 for the minimum number of samples in the tree terminal nodes (n.minobsinn). Fig. 5 shows the variation of Accuracy vs the number of trees (n.trees) and maximum nodes per tree (interaction.d). Predictive performance of the StoGB was calculated using the training data set, and results indicated good performance according to the Accuracy (80.01) and Kappa (0.79) (Table 2).



FIGURE 5. The variation of accuracy vs number of trees (n.trees) and maximum nodes per tree (interaction.d).

TABLE 2. Predictive performance of the models for the training dataset.

Statistic	StoGB	RotFor	Bayesglm
Accuracy (%)	80.1	86.7	86.7
Kappa	0.79	0.86	0.86

Results of the optimization for the RotFor model indicated that the best number of variable subsets (K) was equal to 2,

145570

and the best Ensemble size (L) was equal to 3 with the objective function of accuracy equal to 89.7 % (Fig. 6). Amounts of Accuracy and Kappa were respectively 86.7% and 0.86 indicating excellent performance (Table 2). After 5-fold crossvalidation, the final model result for Bayesglm indicated that accuracy was 86.7 and Kappa was 0.86 (excellent) (Table 2).



FIGURE 6. The variation of accuracy vs variable subsets (K) and Ensemble size (L).

Like the RotFor model, results of the Bayesglm indicated that the model had an excellent performance in the prediction of groundwater salinity (Accuracy and Kappa were respectively 86.7% and 0.86) (Table 2). The coefficients estimated by Bayesglm for each variable were represented in Table 3.

TABLE 3. The coefficients estimated by Bayesglm model for each variable.

Variable	Coefficient
Intercept	1.70
Depth to groundwater	-1.30
Elevation	-1.03
Distance from fault	-0.67
Lithology	-0.54
Decline of groundwater level	0.72
Distance from lake	0.16
Groundwater withdrawal	-0.08
Distance from stream	-0.19

According to the models' evaluation, the RotFor and Bayesglm indicate the superior performance of the StoGB. Although StoGB had a good performance, the RotFor and Bayesglm had an excellent performance based on the Kappa values (K > 0.85) [115].

C. SPATIAL PREDICTION OF GROUNDWATER SALINITY

Three individual machine learning algorithms (i.e., StoGB, RotFor, and Bayesglm) were applied to predict groundwater salinity in the area of the region. The models produced the



FIGURE 7. Probability (left) and susceptibility (right) maps of groundwater salinity: (a) StoGB, (b) RotFor, and (c) Bayesglm.

probability of groundwater salinity maps with a cell size of 12.5 m. The probability maps vary between zero to 1, which respectively indicates the lowest and highest probability of groundwater salinity (Fig. 7). The probability maps were classified using equal interval classification method with an interval 0.2 into five classes: very low (0 - 0.2), low (0.2 - 0.4), medium (0.4 - 0.6), high (0.6 - 0.8), and very high (0.8 - 1).

As can be seen, the very low class is highest for the StoGB model (1210.8 km^2) and includes the most area of the region, whereas the medium susceptibility class predicted by this model is lowest (70.8 km^2) (Table 4). Most of the study

area was placed in medium (mostly marginal parts with an area equal to 704.6 km²) and very high (mostly central parts with an area equal to 628.7 km^2) susceptibility classes by the RotFor model, whereas the class of high susceptibility was lowest (56.9 km²) (Table 4). The Bayesglm predicted that most of the study area has very low (702.2 km²) and very high (596.4 km²) susceptibility (Table 4), respectively located in the boundary and central of the study area (Fig. 7).

Spatial modeling results highlighted that two models of the RotFor and Bayesglm produced strong similarities in the development of the probability and susceptibility maps. It should be noted that these two models have performed

TABLE 4. Area (km²) of susceptibility class for groundwater salinity map of predictive models.

Hazard class	StoGB	RotFor	Bayesglm
Very low	1210.8	121.6	702.2
Low	144.8	176.4	102.3
Medium	70.8	704.6	114.6
High	113.9	56.9	172.7
Very high	147.9	628.7	596.4

better than the StoGB model (Table 2). Although prediction of the models was different, all of the models indicated that central parts of the region have a very high susceptibility in which this class matches through both probability and susceptibility maps predicted by models of the RotFor and Bayesglm (Fig. 7). The very high class of the susceptibility maps predicted by models, especially the RotFor and Bayesglm (Fig. 7), are matched with agricultural areas, lithology map, the locations with low depth to groundwater, low slope, and elevation (Fig. 2). Additionally, areas near to the Maharlu Lake and locations with a high decline in groundwater are located in the very high susceptibility zone.

However, it is obvious that a set of the abovementioned factors which may affect the salinity of the groundwater in the region are: (i) topographic effects, which can cause flushing and exporting the salty materials from the soil and accumulating it in lowlands; (ii) decline in groundwater and proximity to lake, which can cause intrusion of saltwater [86], [93]. In this regard, Jahanshahi and Zare [119] confirmed the contribution of the saltwater intrusion from Maharlu lake to deteriorate the Sarvestan plain groundwater; (iii) low depth to groundwater along with the agricultural area in the central parts, which can cause infiltration of irrigation water, leaching and dissolving solid material into groundwater [35], [92]. Also because of low depth to groundwater, Jahanshahi and Zare [119] demonstrated the effects of evaporation from shallow groundwater on the quality of groundwater in some parts of this plain; and (v) lithology effects, which includes salty formations of the Sachoun (lithology unit PeEsa includes the evaporate minerals of halite and gypsum), Razak (lithology unit Omr includes the marl interbedded with limestone), Aghajari (lithology unit MPlfgp which includes calcareous sandstone and red marl with interbedded gypsum), and Hormoz (lithology unit pC-Ch includes mainly the salts and other evaporate rocks) (Fig. 2o) [119]. Although mostly, these formations are located in the border of the very high susceptibility zones, however, drainage of the plain causes the salty materials to be flushed and deposited in lowlands (which have mostly Qft2 lithology: low-level piedmont fan, and valley terrace deposits) and penetrate into groundwater. This is confirmed by Raeisi et al. [84] who, using optical and chemical methods, indicated that the dissolution of gypsum and other evaporate materials, from surrounding formations of the plain, with water expelled from those are the main reasons for sodium-chloride water type and the poor quality of groundwater in this area.

IV. CONCLUSION

Producing the susceptibility map of salinity for groundwater is of utmost importance for the water security of arid areas and sustainable agriculture. In this study, the best features are selected by the RFE method, and the three machine learning models of StoGB, RotFor, and Bayesglm showed promising results in the prediction of susceptibility maps. All of the models indicated that central parts of the region have a very high susceptibility which matches with agricultural areas, lithology map, the locations with low depth to groundwater, low slope, and elevation.

In this study, due to irregular monitoring and lack of a continuous sampling for each month and season, modeling was conducted using mean yearly EC of wells during a period of 15 years (2003-2017). However, variations of EC during seasons and overtime may be effective on a susceptibility map, so it is recommended considering impacts of climate variability and land-use change on susceptibility map by dividing a longer time period into sub-periods for future studies in other regions. Although the sodium adsorption ratio (SAR), as well as the EC, affects groundwater salinity, investigation of the observed groundwater data in this study indicated low values of the SAR. However, integration of salinity susceptibility (from EC) and sodicity (alkalinity) susceptibility (from SAR) maps during the next studies can be worthwhile to indicate the status of the groundwater quality.

The susceptibility maps produced in this study can increase the awareness of farmers from susceptible areas to control their practices and to adapt cropping systems in order to deal with the groundwater salinity, and it also helps decision-makers to change both land use and land management.

REFERENCES

- [1] W. Xu and X. Su, "Challenges and impacts of climate change and human activities on groundwater-dependent ecosystems in arid areas—A case study of the Nalenggele Alluvial fan in NW China," *J. Hydrol.*, vol. 573, pp. 376–385, Jun. 2019, doi: 10.1016/j.jhydrol.2019.03.082.
- [2] J. Xue, Z. Huo, Q. Huang, F. Wang, J. Boll, G. Huang, and Z. Qu, "Assessing sustainability of agricultural water saving in an arid area with shallow groundwater," *Irrigation Drainage*, vol. 68, no. 2, pp. 205–217, Apr. 2019, doi: 10.1002/ird.2295.
- [3] N. Barreiras and L. Ribeiro, "Estimating groundwater recharge uncertainty for a carbonate aquifer in a semi-arid region using the Kessler's method," J. Arid Environ., vol. 165, pp. 64–72, Jun. 2019, doi: 10.1016/j.jaridenv.2019.04.005.
- [4] R. P. Sishodia, S. Shukla, W. D. Graham, S. P. Wani, J. W. Jones, and J. Heaney, "Current and future groundwater withdrawals: Effects, management and energy policy options for a semi-arid Indian Watershed," *Adv. Water Resour.*, vol. 110, pp. 459–475, Dec. 2017, doi: 10.1016/j.advwatres.2017.05.014.
- [5] A. E. Akay Demir, F. B. Dilek, and U. Yetis, "A new screening index for pesticides leachability to groundwater," *J. Environ. Manage.*, vol. 231, pp. 1193–1202, Feb. 2019, doi: 10.1016/j.jenvman.2018.11.007.
- [6] C. Q. Dai, "Analysis and prevention of the impacts of metro construction on groundwater environment," *Appl. Ecology Environ. Res.*, vol. 17, no. 2, pp. 4839–4854, 2019, doi: 10.15666/aeer/1702_48394854.

- [7] N. Jabbari, F. Aminzadeh, and F. P. J. de Barros, "Hydraulic fracturing and the environment: Risk assessment for groundwater contamination from well casing failure," *Stochastic Environ. Res. Risk Assessment*, vol. 31, no. 6, pp. 1527–1542, Aug. 2017, doi: 10.1007/s00477-016-1280-0.
- [8] D. J. Soeder, "Groundwater quality and hydraulic fracturing: Current understanding and science needs," *Groundwater*, vol. 56, no. 6, pp. 852–858, Nov. 2018, doi: 10.1111/gwat.12810.
- [9] L. Gaona, F. Bedmar, V. Gianelli, A. J. Faberi, and H. Angelini, "Estimating the risk of groundwater contamination and environmental impact of pesticides in an agricultural Basin in Argentina," *Int. J. Environ. Sci. Technol.*, vol. 16, no. 11, pp. 6657–6670, Nov. 2019, doi: 10.1007/s13762-019-02267-w.
- [10] R. Ghazavi and H. Ebrahimi, "Predicting the impacts of climate change on groundwater recharge in an arid environment using modeling approach," *Int. J. Climate Change Strategies Manage.*, vol. 11, no. 1, pp. 88–99, Jan. 2019, doi: 10.1108/IJCCSM-04-2017-0085.
- [11] A. L. Gimsing *et al.*, "Conducting groundwater monitoring studies in europe for pesticide active substances and their metabolites in the context of regulation (EC) 1107/2009," *J. Consum. Protection Food Saf.*, vol. 14, no. S1, pp. 1–93, Jul. 2019, doi: 10.1007/s00003-019-01211-x.
- [12] M. Larocque, J. Levison, A. Martin, and D. Chaumont, "A review of simulated climate change impacts on groundwater resources in Eastern Canada," *Can. Water Resour. J./Revue canadienne des Ressources Hydriques*, vol. 44, no. 1, pp. 22–41, Jan. 2019, doi: 10.1080/07011784.2018.1503066.
- [13] Z. Lyu, J. Chai, Z. Xu, and Y. Qin, "Environmental impact assessment of mining activities on groundwater: Case study of copper mine in Jiangxi Province, China," *J. Hydrol. Eng.*, vol. 24, no. 1, pp. 212–233, Jan. 2019, doi: 10.1061/(ASCE)HE.1943-5584.0001739.
- [14] W. Qiao, W. Li, S. Zhang, and Y. Niu, "Effects of coal mining on the evolution of groundwater hydrogeochemistry," *Hydrogeology J.*, vol. 27, no. 6, pp. 2245–2262, Sep. 2019, doi: 10.1007/s10040-019-01969-2.
- [15] T. Wen, A. Agarwal, L. Xue, A. Chen, A. Herman, Z. Lid, and S. L. Brantley, "Assessing changes in groundwater chemistry in landscapes with more than 100 years of oil and gas development," *Environ. Sci., Processes Impacts*, vol. 21, pp. 384–396, Jan. 2019, doi: 10.1039/c8em00385h.
- [16] E. Barth-Naftilan, J. Sohng, and J. E. Saiers, "Methane in groundwater before, during, and after hydraulic fracturing of the Marcellus Shale," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 27, pp. 6970–6975, Jul. 2018, doi: 10.1073/pnas.1720898115.
- [17] X. Jia, D. O'Connor, D. Hou, Y. Jin, G. Li, C. Zheng, Y. S. Ok, D. C. W. Tsang, and J. Luo, "Groundwater depletion and contamination: Spatial distribution of groundwater resources sustainability in China," *Sci. Total Environ.*, vol. 672, pp. 551–562, Jul. 2019, doi: 10.1016/j.scitotenv.2019.03.457.
- [18] M. G. Y. L. Mahagamage, P. S. Manage, and P. M. Manage, "Water quality and microbial contamination status of groundwater in Jaffna Peninsula, Sri Lanka," *J. Water Land Develop.*, vol. 40, no. 1, pp. 3–12, Mar. 2019, doi: 10.2478/jwld-2019-0001.
- [19] I. Mukherjee and U. K. Singh, "Groundwater fluoride contamination, probable release, and containment mechanisms: A review on Indian context," *Environ. Geochemistry Health*, vol. 40, no. 6, pp. 2259–2301, Dec. 2018, doi: 10.1007/s10653-018-0096-x.
- [20] R. Reshma and G. Sindhu, "Assessment of groundwater vulnerability to contamination: A case study," *Environ. Monitor. Assessment*, vol. 191, no. 6, Jun. 2019, Art. no. 356, doi: 10.1007/s10661-019-7493-1.
- [21] A. K. Tiwari, M. De Maio, and G. Amanzio, "Evaluation of metal contamination in the groundwater of the Aosta Valley Region, Italy," *Int. J. Environ. Res.*, vol. 11, no. 3, pp. 291–300, Aug. 2017, doi: 10.1007/s41742-017-0027-1.
- [22] X. Zhang, B. Hu, P. Wang, J. Chen, L. Yang, K. Xiao, and X. Zhang, "Hydrogeochemical evolution and heavy metal contamination in groundwater of a reclaimed land on Zhoushan Island," *Water*, vol. 10, no. 3, p. 316, Mar. 2018, doi: 10.3390/w10030316.
- [23] R. Afroza, Q. H. Mazumder, C. S. Jahan, M. A. I. Kazi, M. A. Ahsan, and M. A. Al-Mansur, "Hydrochemistry and origin of salinity in groundwater in parts of lower Tista Floodplain, Northwest Bangladesh," *J. Geological Soc. India*, vol. 74, no. 2, pp. 223–232, Aug. 2009, doi: 10.1007/s12594-009-0124-7.

- [24] S. M'nassri, L. Dridi, G. Schäfer, M. Hachicha, and R. Majdoub, "Groundwater salinity in a semi-arid region of central-eastern tunisia: Insights from multivariate statistical techniques and geostatistical modelling," *Environ. Earth Sci.*, vol. 78, no. 10, May 2019, Art. no. 288, doi: 10.1007/s12665-019-8270-8.
- [25] X. Wen, Y. Wu, J. Su, Y. Zhang, and F. Liu, "Hydrochemical characteristics and salinity of groundwater in the Ejina Basin, Northwestern China," *Environ. Geol.*, vol. 48, no. 6, pp. 665–675, Sep. 2005, doi: 10.1007/s00254-005-0001-7.
- [26] A. Pulido-Bosch, J. P. Rigol-Sanchez, A. Vallejos, J. M. Andreu, J. C. Ceron, L. Molina-Sanchez, and F. Sola, "Impacts of agricultural irrigation on groundwater salinity," *Environ. Earth Sci.*, vol. 77, no. 5, p. 197, Mar. 2018, doi: 10.1007/s12665-018-7386-6.
- [27] A. Lal and B. Datta, "Development and implementation of support vector machine regression surrogate models for predicting groundwater pumping-induced saltwater intrusion into coastal aquifers," *Water Resour. Manage.*, vol. 32, no. 7, pp. 2405–2419, May 2018, doi: 10.1007/s11269-018-1936-2.
- [28] S. Foster, A. Pulido-Bosch, Á. Vallejos, L. Molina, A. Llop, and A. M. MacDonald, "Impact of irrigated agriculture on groundwaterrecharge salinity: A major sustainability concern in semi-arid regions," *Hydrogeology J.*, vol. 26, no. 8, pp. 2781–2791, Dec. 2018, doi: 10.1007/s10040-018-1830-2.
- [29] A. E. K. Douaoui, H. Nicolas, and C. Walter, "Detecting salinity hazards within a semiarid context by means of combining soil and remotesensing data," *Geoderma*, vol. 134, nos. 1–2, pp. 217–230, Sep. 2006, doi: 10.1016/j.geoderma.2005.10.009.
- [30] A. Allbed, L. Kumar, and P. Sinha, "Mapping and modelling spatial variation in soil salinity in the al hassa oasis based on remote sensing indicators and regression techniques," *Remote Sens.*, vol. 6, no. 2, pp. 1137–1157, Jan. 2014, doi: 10.3390/rs6021137.
- [31] E. Asfaw, K. V. Suryabhagavan, and M. Argaw, "Soil salinity modeling and mapping using remote sensing and GIS: The case of Wonji Sugar Cane irrigation farm, Ethiopia," *J. Saudi Soc. Agricult. Sci.*, vol. 17, no. 3, pp. 250–258, Jul. 2018, doi: 10.1016/j.jssas.2016.05.003.
- [32] F. K. Khadim, H. Su, L. Xu, and J. Tian, "Soil salinity mapping in everglades national park using remote sensing techniques and vegetation salt tolerance," *Phys. Chem. Earth*, vol. 110, pp. 31–50, Apr. 2019, doi: 10.1016/j.pce.2019.01.004.
- [33] J. Huang, M. J. Prochazka, and J. Triantafilis, "Irrigation salinity hazard assessment and risk mapping in the lower Macintyre Valley, Australia," *Sci. Total Environ.*, vols. 551–552, pp. 460–473, May 2016, doi: 10.1016/j.scitotenv.2016.01.200.
- [34] J. Huang, R. Subasinghe, R. S. Malik, and J. Triantafilis, "Salinity hazard and risk mapping of point source salinisation using proximally sensed electromagnetic instruments," *Comput. Electron. Agricult.*, vol. 113, pp. 213–224, Apr. 2015, doi: 10.1016/j.compag. 2015.02.013.
- [35] J. Seydehmet, G.-H. Lv, A. Abliz, Q.-D. Shi, A. Abliz, and A. Turup, "Irrigation salinity risk assessment and mapping in arid oasis, Northwest China," *Water*, vol. 10, no. 7, p. 966, Jul. 2018, doi: 10.3390/ w10070966.
- [36] F. J. Alcalá and E. Custodio, "Atmospheric chloride deposition in continental Spain," *Hydrological Processes*, vol. 22, no. 18, pp. 3636–3650, Aug. 2008, doi: 10.1002/hyp.6965.
- [37] H. B. Gu and L. J. Wang, "Water and salt regulation mode of soil salinization based on change of land utilization," *Adv. Mater. Res.*, vols. 550–553, pp. 2551–2555, Jul. 2012, doi: 10.4028/www.scientific.net/AMR.550-553.2551.
- [38] D. Han, V. E. A. Post, and X. Song, "Groundwater salinization processes and reversibility of seawater intrusion in coastal carbonate aquifers," *J. Hydrol.*, vol. 531, pp. 1067–1080, Dec. 2015, doi: 10.1016/j.jhydrol.2015.11.013.
- [39] A. E. Khan, P. F. D. Scheelbeek, A. B. Shilpi, Q. Chan, S. K. Mojumder, A. Rahman, A. Haines, and P. Vineis, "Salinity in drinking water and the risk of (Pre)Eclampsia and gestational hypertension in coastal Bangladesh: A case-control study," *PLoS ONE*, vol. 9, no. 9, Sep. 2014, Art. no. e108715, doi: 10.1371/journal.pone.0108715.
- [40] D. Merchán, L. F. Auqué, P. Acero, M. J. Gimeno, and J. Causapé, "Geochemical processes controlling water salinization in an irrigated basin in Spain: Identification of natural and anthropogenic influence," *Sci. Total Environ.*, vol. 502, pp. 330–343, Jan. 2015, doi: 10.1016/j.scitotenv.2014.09.041.

- [41] A. Pérez-Gimeno, J. Navarro-Pedreño, M. B. Almendro-Candel, I. Gómez, and M. M. Jordán, "Environmental consequences of the use of sewage sludge compost and limestone outcrop residue for soil restoration: Salinity and trace elements pollution," *J. Soils Sediments*, vol. 16, no. 3, pp. 1012–1021, Mar. 2016, doi: 10.1007/s11368-015-1288-y.
- [42] A. Pérez-Gimeno, J. Navarro-Pedreño, M. B. Almendro-Candel, I. Gómez, and M. M. Jordán, "Nitrogen compound pollution associated with the use of sewage sludge compost and limestone outcrop residue under saline irrigation," *J. Soils Sediments*, vol. 16, no. 4, pp. 1345–1351, Apr. 2016, doi: 10.1007/s11368-014-0972-7.
- [43] G. Cui, Y. Lu, C. Zheng, Z. Liu, and J. Sai, "Relationship between soil salinization and groundwater hydration in Yaoba Oasis, Northwest China," *Water*, vol. 11, no. 1, p. 175, Jan. 2019, doi: 10.3390/w11010175.
- [44] G. Giannoccaro, A. Scardigno, and M. Prosperi, "Economic analysis of the long-term effects of groundwater salinity: Bringing the farmer's perspectives into policy," *J. Integrative Environ. Sci.*, vol. 14, no. 1, pp. 59–72, Jan. 2017, doi: 10.1080/1943815X.2017.1351993.
- [45] P. Li, H. Qian, and J. Wu, "Conjunctive use of groundwater and surface water to reduce soil salinization in the Yinchuan plain, North-West China," *Int. J. Water Resour. Develop.*, vol. 34, no. 3, pp. 337–353, May 2018, doi: 10.1080/07900627.2018.1443059.
- [46] C. Petheram, K. L. Bristow, and P. N. Nelson, "Understanding and managing groundwater and salinity in a tropical conjunctive water use irrigation district," *Agricult. Water Manage.*, vol. 95, no. 10, pp. 1167–1179, Oct. 2008, doi: 10.1016/j.agwat.2008.04.016.
- [47] E. Salameh, M. Alraggad, and A. Tarawneh, "Natural salinity sources in the groundwaters of jordan-importance of sustainable aquifer management," *Geochemistry*, vol. 74, no. 4, pp. 735–747, Dec. 2014, doi: 10.1016/j.chemer.2014.04.007.
- [48] E. Shaji, J. J. Gómez-Alday, S. Hussein, T. R. Deepu, and Y. Anilkumar, "Salinization and deterioration of groundwater quality by nitrate and fluoride in the Chittur Block, Palakkad, Kerala," *J. Geological Soc. India*, vol. 92, no. 3, pp. 337–345, Sep. 2018, doi: 10.1007/s12594-018-1017-4.
- [49] M. Shamsudduha, A. Zahid, W. G. Burgess, and K. M. Ahmed, "Security of deep groundwater against ingress of arsenic and salinity in Bangladesh: Policy aspects," in *Proc. 5th Int. Congr. Arsen. Environ. One Century Discovery Arsenicosis Latin Amer.*, 2014, pp. 891–893.
- [50] S. Khan, T. Rana, and M. A. Hanjra, "A cross disciplinary framework for linking farms with regional groundwater and salinity management targets," *Agricult. Water Manage.*, vol. 95, no. 1, pp. 35–47, Jan. 2008, doi: 10.1016/j.agwat.2007.09.005.
- [51] D. Han, X. Song, M. J. Currell, G. Cao, Y. Zhang, and Y. Kang, "A survey of groundwater levels and hydrogeochemistry in irrigated fields in the karamay agricultural development area, Northwest China: Implications for soil and groundwater salinity resulting from surface water transfer for irrigation," *J. Hydrol.*, vol. 405, nos. 3–4, pp. 217–234, Aug. 2011, doi: 10.1016/j.jhydrol.2011.03.052.
- [52] F. Hemmati and Z. Sajadi, "Salinity variation of ground water and hydro chemical characteristic of brines origin of salts and brine evolution of Borazjan, Iran," Adv. Environ. Biol., vol. 8, no. 11, pp. 294–300, 2014. Accessed: Jul. 07, 2020. [Online]. Available: https://go. gale.com/ps/i.do?id=GALE%7CA392176538&sid=googleScholar&v= 2.1&it=r&linkaccess=abs&issn=19950756&p=AONE&sw=w
- [53] S. Yu, J. Yang, G. Liu, R. Yao, and X. Wang, "Regional groundwater salinity dynamics forecasting based on neuro-fuzzy algorithm," *Nongye Gongcheng Xuebao/Transactions Chin. Soc. Agricult. Eng.*, vol. 30, no. 18, pp. 142–150, 2014, doi: 10.3969/j.issn.1002-6819. 2014.18.018.
- [54] A. Brenot, P. Négrel, E. Petelet-Giraud, R. Millot, and E. Malcuit, "Insights from the salinity origins and interconnections of aquifers in a regional scale sedimentary aquifer system (Adour-garonne district, SW France): Contributions of δ34S and δ18O from dissolved sulfates and the 87Sr/86Sr ratio," *Appl. Geochemistry*, vol. 53, pp. 27–41, Feb. 2015, doi: 10.1016/j.apgeochem.2014.12.002.
- [55] K. Rina, C. K. Singh, P. S. Datta, N. Singh, and S. Mukherjee, "Geochemical modelling, ionic ratio and GIS based mapping of groundwater salinity and assessment of governing processes in Northern Gujarat, India," *Environ. Earth Sci.*, vol. 69, no. 7, pp. 2377–2391, Aug. 2013, doi: 10.1007/s12665-012-2067-3.
- [56] H. R. Nassery and Z. Kayhomayoon, "Source of salinity in the groundwater of Lenjanat plain, Isfahan, Iran," *Environ. Earth Sci.*, vol. 68, no. 2, pp. 413–427, Jan. 2013, doi: 10.1007/s12665-012-1746-4.

- [57] M. J. Stephens, D. H. Shimabukuro, J. M. Gillespie, and W. Chang, "Groundwater salinity mapping using geophysical log analysis within the fruitvale and rosedale ranch oil fields, Kern County, California, USA," *Hydrogeology J.*, vol. 27, no. 2, pp. 731–746, Mar. 2019, doi: 10.1007/s10040-018-1872-5.
- [58] B. S. Karatas, G. Camoglu, and M. K. Olgen, "Spatio-temporal trend analysis of the depth and salinity of the groundwater, using geostatistics integrated with GIS, of the menemen irrigation system, western turkey," *Ekoloji*, vol. 22, no. 86, pp. 36–47, Mar. 2013, doi: 10.5053/ekoloji.2013.865.
- [59] I. Mullen and J. Kellett, "Groundwater salinity mapping using airborne electromagnetics and borehole data within the lower Balonne catchment, Queensland, Australia," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 9, no. 2, pp. 116–123, May 2007, doi: 10.1016/j.jag. 2006.07.004.
- [60] R. A. Brown, R. W. Skaggs, and W. F. Hunt, "Calibration and validation of DRAINMOD to model bioretention hydrology," *J. Hydrol.*, vol. 486, pp. 430–442, Apr. 2013, doi: 10.1016/j.jhydrol.2013.02.017.
- [61] K. Davoodi, A. Darzi-Naftchali, and G. Aghajani-Mazandarani, "Evaluating Drainmod's to predict drainage water salinity and groundwater table depth during winter cropping in heavy-textured paddy soils," *Irrigation Drainage*, vol. 68, no. 3, pp. 559–572, Jul. 2019, doi: 10.1002/ird.2339.
- [62] M. Fossati and I. Piedra-Cueva, "Numerical modelling of residual flow and salinity in the Río de la plata," *Appl. Math. Model.*, vol. 32, no. 6, pp. 1066–1086, Jun. 2008, doi: 10.1016/j.apm.2007.02.034.
- [63] L. Belkhiri and T. S. Narany, "Using multivariate statistical analysis, geostatistical techniques and structural equation modeling to identify spatial variability of groundwater quality," *Water Resour. Manage.*, vol. 29, no. 6, pp. 2073–2089, Apr. 2015, doi: 10.1007/s11269-015-0929-7.
- [64] S. M. Yidana, B. Banoeng-Yakubo, and T. M. Akabzaa, "Analysis of groundwater quality using multivariate and spatial analyses in the Keta Basin, Ghana," J. Afr. Earth Sci., vol. 58, no. 2, pp. 220–234, Sep. 2010, doi: 10.1016/j.jafrearsci.2010.03.003.
- [65] S. M. Praveena, M. H. Abdullah, K. Bidin, and A. Z. Aris, "Understanding of groundwater salinity using statistical modeling in a small tropical Island, East Malaysia," *Environmentalist*, vol. 31, no. 3, pp. 279–287, Sep. 2011, doi: 10.1007/s10669-011-9332-y.
- [66] F. Bouksila, M. Persson, A. Bahri, and R. Berndtsson, "Electromagnetic induction prediction of soil salinity and groundwater properties in a Tunisian Saharan Oasis," *Hydrological Sci. J.*, vol. 57, no. 7, pp. 1473–1486, Oct. 2012, doi: 10.1080/02626667.2012.717701.
- [67] V. Gholami, Z. Yousefi, and H. Zabardast Rostami, "Modeling of ground water salinity on the Caspian Southern Coasts," *Water Resour. Manage.*, vol. 24, no. 7, pp. 1415–1424, May 2010, doi: 10.1007/s11269-009-9506-2.
- [68] L. K. Lautz, G. D. Hoke, Z. Lu, D. I. Siegel, K. Christian, J. D. Kessler, and N. G. Teale, "Using discriminant analysis to determine sources of salinity in shallow groundwater prior to hydraulic fracturing," *Environ. Sci. Technol.*, vol. 48, no. 16, pp. 9061–9069, Aug. 2014, doi: 10.1021/es502244v.
- [69] R. Talebnejad and A. R. Sepaskhah, "Modification of transient state analytical model under different saline groundwater depths, irrigation water salinities and deficit irrigation for Quinoa," Gorgan Univ. Agricult. Natural Resour., Gorgan, Iran, Tech. Rep., 2016, doi: 10.22069/ijpp.2016.2903.
- [70] A. Khattak, N. Ahmed, I. Hussein, A. Qazi, S. Alikhan, A. Rehman, and N. Iqbal, "Spatial distribution of salinity in shallow groundwater used for crop irrigation," *Pakistan J. Bot.*, vol. 46, no. 2, pp. 531–537, 2014. Accessed: Jul. 07, 2020. [Online]. Available: http://www.academia.edu/download/33555616/18-1.pdf
- [71] B. S. Abuelaish and M. T. C. Olmedo, "Análisis y modelización de la salinidad del agua subterránea en la franja de gaza," *Cuadernos Geográficos*, vol. 57, no. 2, pp. 72–91, Jul. 2018, doi: 10.30827/cuadgeo.v57i2.5914.
- [72] M. Amiri-Bourkhani, M. R. Khaledian, A. Ashrafzadeh, and A. Shahnazari, "Vremenske i prostorne varijacije salaniteta podzemnih voda u ravnici Mazandaran, Iran, tijekom dugogodišnjeg razdoblja od 26 godina," *Geofizika*, vol. 34, no. 1, pp. 119–139, 2017, doi: 10.15233/gfz.2017.34.4.
- [73] A. Ashrafzadeh, F. Roshandel, M. Khaledian, M. Vazifedoust, and M. Rezaei, "Assessment of groundwater salinity risk using Kriging methods: A case study in Northern Iran," *Agricult. Water Manage.*, vol. 178, pp. 215–224, Dec. 2016, doi: 10.1016/j.agwat.2016.09.028.

- [74] D. Kaplan and R. Muñoz-Carpena, "Complementary effects of surface water and groundwater on soil moisture dynamics in a degraded coastal floodplain forest," *J. Hydrol.*, vol. 398, nos. 3–4, pp. 221–234, Feb. 2011, doi: 10.1016/j.jhydrol.2010.12.019.
- [75] O. Merlin, J. Walker, A. Chehbouni, and Y. Kerr, "Towards deterministic downscaling of SMOS soil moisture using MODIS derived soil evaporative efficiency," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3935–3946, Oct. 2008, doi: 10.1016/j.rse.2008.06.012.
- [76] P. Cook, G. Walker, and A. Telfer, "Potential for salinisation of the groundwater beneath mallee areas of the Murray Basin," Dept. Eng. Water Supply, Univ. South Australia, Adelaide, SA, USA, Tech. Rep. 707694, 1993.
- [77] K. S. Gundogdu and S. T. Akkaya Aslan, "Mapping multi-year ground water salinity patterns in irrigation areas using time-series analysis of ground water salinity maps," *Hydrological Processes*, vol. 22, no. 6, pp. 821–826, Mar. 2008, doi: 10.1002/hyp.6643.
- [78] W. M. Khadra and P. J. Stuyfzand, "Simulation of saltwater intrusion in a poorly karstified coastal aquifer in Lebanon (Eastern Mediterranean)," *Hydrogeology J.*, vol. 26, no. 6, pp. 1839–1856, Sep. 2018, doi: 10.1007/s10040-018-1752-z.
- [79] O. Dubovyk, G. Menz, C. Conrad, E. Kan, M. Machwitz, and A. Khamzina, "Spatio-temporal analyses of cropland degradation in the irrigated lowlands of uzbekistan using remote-sensing and logistic regression modeling," *Environ. Monitor. Assessment*, vol. 185, no. 6, pp. 4775–4790, Jun. 2013, doi: 10.1007/s10661-012-2904-6.
- [80] S. Abdi, A. Tajabadi Pour, H. Shirani, M. Hamidpour, and H. Shekofteh, "Assessing salinity and sodicity hazards of ground water for irrigation purposes using fuzzy logic," *Desalination Water Treatment*, vol. 57, no. 33, pp. 15547–15558, Jul. 2016, doi: 10.1080/19443994.2015.1072740.
- [81] M. Attwa, K. S. Gemail, and M. Eleraki, "Use of salinity and resistivity measurements to study the coastal aquifer salinization in a semi-arid region: A case study in Northeast Nile Delta, Egypt," *Environ. Earth Sci.*, vol. 75, no. 9, May 2016, Art. no. 784, doi: 10.1007/s12665-016-5585-6.
- [82] J. S. Alagha, M. Seyam, M. A. Md Said, and Y. Mogheir, "Integrating an artificial intelligence approach with k-means clustering to model groundwater salinity: The case of gaza coastal aquifer (Palestine)," *Hydrogeol*ogy J., vol. 25, no. 8, pp. 2347–2361, Dec. 2017, doi: 10.1007/s10040-017-1658-1.
- [83] A. Khalili, S. Hajam, and P. Irannezhad, "Integrated water plan of Iran," Univ. Tehran, Tehran, Iran, Tech. Rep., 1997.
- [84] E. Raeisi, O. Jehbez, and F. Moore, "Hydrochemical behavior of karstic and evaporitic formations surrounding Sarvestan Plain, Iran," *Theor. Appl. Karstology*, vol. 9, no. 9, pp. 165–174, 1996.
- [85] M. Zamanpoore, M. R. Daremipouran, M. R. Ghaed-Abdi, and N. K. Ahmadi, "Chemical and physical properties of Maharlu Salt Lake (Iran) prior to an extensive drought," *Ecopersia*, vol. 7, no. 1, pp. 59–67, 2019.
- [86] S. Selvakumar and N. Ch. (2012). Preliminary Investigation of Groundwater Quality along the Coastal Aquifers of Southern Tamil Nadu Using GIS Techniques. Accessed: Jul. 7, 2020. [Online]. Available: http://www.journal.bonfring.org/papers/iems/volume2/BIJ-002-1327.pdf
- [87] K. Pazand, A. Hezarkhani, Y. Ghanbari, and N. Aghavali, "Groundwater geochemistry in the Meshkinshahr Basin of Ardabil Province in Iran," *Environ. Earth Sci.*, vol. 65, no. 3, pp. 871–879, Feb. 2012, doi: 10.1007/s12665-011-1131-8.
- [88] I. Golchin and M. Azhdary Moghaddam, "Hydro-geochemical characteristics and groundwater quality assessment in Iranshahr Plain Aquifer, Iran," *Environ. Earth Sci.*, vol. 75, no. 4, Feb. 2016, Art. no. 317, doi: 10.1007/s12665-015-5077-0.
- [89] USDA Handbook No 60: Diagnosis and Improvement of Saline and Alkali Soils, USL Staff, Washington, DC, USA, 1954.
- [90] N. Aghazadeh and A. A. Mogaddam, "Assessment of groundwater quality and its suitability for drinking and agricultural uses in the Oshnavieh area, Northwest of Iran," *J. Environ. Protection*, vol. 1, no. 1, pp. 30–40, 2010, doi: 10.4236/jep.2010.11005.
- [91] V. Re and E. Sacchi, "Tackling the salinity-pollution nexus in coastal aquifers from arid regions using nitrate and boron isotopes," *Environ. Sci. Pollut. Res.*, vol. 24, no. 15, pp. 13247–13261, May 2017, doi: 10.1007/s11356-017-8384-z.
- [92] S. Foster, J. Chilton, M. Moencg, F. Cardy, and M. Schiffler, *Groundwater in Rural Development*. Washington, DC, USA: The World Bank, 2000.

- [93] S. Mann, M. C. Webb, and R. W. Bell, Sustainable and Profitable Crop and Livestock Systems in South-Central Coastal Vietnam. Canberra, NSW, Australia: ACIAR, 2015.
- [94] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Mach. Learn.*, vol. 46, nos. 1–3, pp. 389–422, 2002, doi: 10.1023/A:1012487302797.
- [95] B. Wang, C. Oldham, and M. R. Hipsey, "Comparison of machine learning techniques and variables for groundwater dissolved organic nitrogen prediction in an urban area," *Procedia Eng.*, vol. 154, no. 3, pp. 1176–1184, 2016, doi: 10.1016/j.proeng.2016.07.527.
- [96] C. Feng, M. Cui, B.-M. Hodge, and J. Zhang, "A data-driven multimodel methodology with deep feature selection for short-term wind forecasting," *Appl. Energy*, vol. 190, pp. 1245–1257, Mar. 2017, doi: 10.1016/j.apenergy.2017.01.043.
- [97] M. Kuhn, "Building predictive models in R using the caret package," J. Stat. Softw., vol. 28, no. 5, pp. 1–26, 2008.
- [98] K. Max, S. Weston, J. Wing, A. Williams, C. Keefer, and A. Engelhardt. (2011). Package 'Caret'. Classification and Regression Training, p. 138. Accessed: May 17, 2020. [Online]. Available: http://adsabs.harvard.edu/abs/2015ascl.soft05003K
- [99] M. Kuhn, Package 'Caret'—Classification and Regression Training. Boca Raton, FL, USA: CRC Press, 2019, doi: 10.1887/0750303123/b365c43.
- [100] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Ann. Stat., vol. 29, no. 5, pp. 1189–1232, 2001, doi: 10.1214/aos/1013203451.
- [101] L. Breiman, "Using iterated bagging to debias regressions," *Mach. Learn.*, vol. 45, no. 3, pp. 261–277, Dec. 2001, doi: 10.1023/A:1017934522171.
- [102] K. P. Singh, S. Gupta, and D. Mohan, "Evaluating influences of seasonal variations and anthropogenic activities on alluvial groundwater hydrochemistry using ensemble learning approaches," *J. Hydrol.*, vol. 511, pp. 254–266, Apr. 2014, doi: 10.1016/j.jhydrol.2014.01.004.
- [103] A. M. Filippi, I. Güneralp, and J. Randall, "Hyperspectral remote sensing of aboveground biomass on a river meander bend using multivariate adaptive regression splines and stochastic gradient boosting," *Remote Sens. Lett.*, vol. 5, no. 5, pp. 432–441, May 2014, doi: 10.1080/2150704X.2014.915070.
- [104] J. J. Rodriguez, L. I. Kuncheva, and C. J. Alonso, "Rotation forest: A new classifier ensemble method," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1619–1630, Oct. 2006, doi: 10.1109/TPAMI.2006.211.
- [105] H. Abdi and L. J. Williams, "Principal component analysis," Wiley Interdiscipl. Rev, Comput. Statist., vol. 2, no. 4, pp. 433–459, Jul. 2010, doi: 10.1002/wics.101.
- [106] S. A. Naghibi, M. Dolatkordestani, A. Rezaei, P. Amouzegari, M. T. Heravi, B. Kalantar, and B. Pradhan, "Application of rotation forest with decision trees as base classifier and a novel ensemble model in spatial modeling of groundwater potential," *Environ. Monitor. Assessment*, vol. 191, no. 4, Apr. 2019, Art. no. 248, doi: 10.1007/s10661-019-7362-y.
- [107] A. M. Al-Abadi, "Mapping flood susceptibility in an arid region of southern iraq using ensemble machine learning classifiers: A comparative study," *Arabian J. Geosci.*, vol. 11, no. 9, p. 218, May 2018, doi: 10.1007/s12517-018-3584-5.
- [108] G. D. Bian, J. K. Du, M. M. Song, Y. P. Xu, S. P. Xie, W. L. Zheng, and C.-Y. Xu, "A procedure for quantifying runoff response to spatial and temporal changes of impervious surface in Qinhuai River Basin of Southeastern China," *Catena*, vol. 157, pp. 268–278, Oct. 2017, doi: 10.1016/j.catena.2017.05.023.
- [109] R. Costache, "Flash-flood potential index mapping using weights of evidence, decision trees models and their novel hybrid integration," *Stochastic Environ. Res. Risk Assessment*, vol. 33, no. 7, pp. 1375–1402, Jul. 2019, doi: 10.1007/s00477-019-01689-9.
- [110] M. Hamada and J. A. Nelder, "Generalized linear models for qualityimprovement experiments," *J. Qual. Technol.*, vol. 29, no. 3, pp. 292–304, Jul. 1997, doi: 10.1080/00224065.1997.11979770.
- [111] G. R. Hosack, K. R. Hayes, and S. C. Barry, "Prior elicitation for Bayesian generalised linear models with application to risk control option assessment," *Rel. Eng. Syst. Saf.*, vol. 167, pp. 351–361, Nov. 2017, doi: 10.1016/j.ress.2017.06.011.
- [112] A. Gelman and I. Pardoe, "Bayesian measures of explained variance and pooling in multilevel (hierarchical) models," *Technometrics*, vol. 48, no. 2, pp. 241–251, May 2006, doi: 10.1198/004017005000000517.

- [113] F. S. Hosseini, B. Choubin, A. Mosavi, N. Nabipour, S. Shamshirband, H. Darabi, and A. T. Haghighi, "Flash-flood hazard assessment using ensembles and Bayesian-based machine learning models: Application of the simulated annealing feature selection method," *Sci. Total Environ.*, vol. 711, Apr. 2020, Art. no. 135161, doi: 10.1016/j.scitotenv.2019.135161.
- [114] H. Darabi, B. Choubin, O. Rahmati, A. Torabi Haghighi, B. Pradhan, and B. Kløve, "Urban flood risk mapping using the GARP and QUEST models: A comparative study of machine learning techniques," *J. Hydrol.*, vol. 569, pp. 142–154, Feb. 2019, doi: 10.1016/j.jhydrol.2018.12.002.
- [115] R. A. Monserud and R. Leemans, "Comparing global vegetation maps with the Kappa statistic," *Ecol. Modell.*, vol. 62, no. 4, pp. 275–293, 1992, doi: 10.1016/0304-3800(92)90003-W.
- [116] L. E. Johnson and B. G. Olsen, "Assessment of quantitative precipitation forecasts," *Weather Forecast.*, vol. 13, no. 1, pp. 75–83, 1998, doi: 10.1175/1520-0434(1998)013<0075:AOQPF>2.0.CO;2.
- [117] A. Beucher, A. B. Møller, and M. H. Greve, "Artificial neural networks and decision tree classification for predicting soil drainage classes in Denmark," *Geoderma*, vol. 352, pp. 351–359, Oct. 2019, doi: 10.1016/j.geoderma.2017.11.004.
- [118] A. Mosavi, F. Sajedi-Hosseini, B. Choubin, F. Taromideh, G. Rahi, and A. Dineva, "Susceptibility mapping of soil water erosion using machine learning models," *Water*, vol. 12, no. 7, p. 1995, Jul. 2020, doi: 10.3390/w12071995.
- [119] R. Jahanshahi and M. Zare, "Hydrochemical investigations for delineating salt-water intrusion into the coastal aquifer of Maharlou Lake, Iran," J. Afr. Earth Sci., vol. 121, pp. 16–29, Sep. 2016, doi: 10.1016/j.jafrearsci.2016.05.014.





FARZANEH SAJEDI HOSSEINI received the B.S. degree in rangeland and watershed management from Shiraz University, Shiraz, Iran, in 2011, and the M.S. degree in watershed management and engineering from Sari Agricultural Sciences and Natural Resources University, Sari, Iran, in 2015. She is currently pursuing the Ph.D. degree in watershed management with the University of Tehran, Tehran, Iran. Her research interests include hydrology and spatial and natural hazard modeline.

BAHRAM CHOUBIN received the B.S. degree in rangeland and watershed management from Shiraz University, Shiraz, Iran, in 2011, the M.S. degree in watershed management and engineering from the University of Tehran, Tehran, Iran, in 2013, and the Ph.D. degree in watershed management from Sari Agricultural Sciences and Natural Resources University, Sari, Iran, in 2017. From 2018 to 2019, he was a Postdoctoral Researcher at the University of Tehran. He has been an Assistant Professor at

the Soil Conservation and Watershed Management Research Department, AREEO, Urmia, Iran, since 2019. He has published more than 30 articles in JCR-indexed journals. His research interests include hydrometeorology, natural hazard, prediction in ungauged basin (PUB), remote sensing, and hazard modeling. He was a recipient of the Academic Award from the National Elites Foundation, Iran, in 2016. He is a scientific reviewer for more than 50 ISI journals.



MASSOUD GOODARZI received the Ph.D. degree in physical geography (climatology and planning) from the University of Tabriz, in 2011. He is currently an Assistant Professor at the Soil Conservation and Watershed Management Research Institute, AREEO, Iran. His research interests include climate change, drought, dust, flood, and other environmental issues.



AMIRHOSEIN MOSAVI received the Ph.D. degree in applied informatics from London Kingston University, U.K. He is an Alexander von Humboldt Research Fellow for big data, the IoT, and machine learning. He is currently a Senior Research Fellow at Oxford Brookes University. He is a Data Scientist for climate change, sustainability, and hazard prediction. He was a recipient of the Green-Talent Award, the UNESCO Young Scientist Award, the ERCIM Alain Bensoussan

Fellowship Award, the Campus France Fellowship Award, the Campus Hungary Fellowship Award, and the Endeavour-Australia Leadership.



ADRIENN A. DINEVA received the B.Sc. and M.Sc. degrees in mechatronics engineering from Obuda University, in 2013, and the Ph.D. degrees in applied informatics from the Universita degli Studi di Milano and Obuda University, in 2017. She is currently an Assistant Professor with Obuda University. Her research interests include the area of adaptive control of nonlinear systems and applied engineering, including machine learning and soft computing techniques, and also the area

of adaptive signal processing. She has authored more than 20 ISI articles in these fields.