ORIGINAL PAPER

Investigation of land-subsidence phenomenon and aquifer vulnerability using machine models and GIS technique

Adel Ghasemi[1] · Omid Bahmani[2] · Samira Akhavan[2] · Hamid Reza Pourghasemi[3]

Abstract

In this study, a land-subsidence vulnerability map has been prepared using Machine Learning (ML) models fusing, Random Forest (RF), Support Vector Machine (SVM) and a GIS technique in the Hamadan Province, Iran. The information layers used as input of the ML models in "R" software are comprised of elevation, slope, plan and profile curvatures, slope aspect, Topographic Wetness Index (TWI), Normalized Difference Vegetation Index (NDVI), soil texture, distance from rivers, distance from the fault, geology, land use, and groundwater level drawdown. The accuracy of the results obtained stood at 89% in the SVM, and up to 96% in the RF models. The RF model demonstrates a greater efficiency than the SVM model. To determine each parameter's effect on the land-subsidence of the study area, the Partial Least Squares (PLS) model has been used in the R software. The use of the PLS model shows a greater effect of elevation and groundwater level decline compared to the other parameters on the land-subsidence phenomenon. Finally, the raster vulnerability map in the GIS software was divided into four classes in terms of intensity as 'low,' 'medium,' 'high,' and 'very high' utilizing the natural break method. In the optimal RF model 45% of aquifers were assessed as being low, 23% as moderate, 20% as high, and 12% as very high. The study of the groundwater changing process, using GRACE satellite data in Google Earth Engine environment confirmed a decrease in groundwater level, which has led to land-subsidence in the aquifer.

Keywords Land-subsidence · Machine learning · Random forest · Support vector machine · PLS · GRACE satellite

✉ Omid Bahmani
  omid.bahmani@basu.ac.ir

  Adel Ghasemi
  adel.ghasemi@gmail.com

  Samira Akhavan
  s.akhavan@basu.ac.ir

  Hamid Reza Pourghasemi
  hr.pourghasemi@shirazu.ac.ir

[1] Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran

[2] Department of Water Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran

[3] Department of Natural Recourses and Environmental Engineering, Faculty of Agriculture, Shiraz University, Shiraz, Iran

1 Introduction

Supplying water in Iran has always been a concern for those involved in the fields of environment and natural resources. Water use and soil resources are dually important when considered from the perspective of (a) Operational management or the means by which these resources are exploited, and (b) the context and type of use of these resources. These two are both independently influential and to some extent interdependent. An increasing population, uncontrolled extraction of groundwater resources, and extensive mismanagement of water resources is currently imposing a great deal of damage to the environment; among which is land subsidence due to drought and groundwater decline in most plains (Sundell et al. 2017; Chai 2017). Land subsidence, also referred to in the literature as a silent earthquake, is a geological hazard that is caused by a variety of factors, such as uncontrolled extraction of groundwater aquifers, earthquakes, volcanic activities, floods, construction of large dams, and tectonic salt domes (Park et al. 2014). Due to the fact that subsidence spreads slowly and gradually, cracks caused by subsidence may not have the same impact as abrupt and catastrophic hazards such as floods and earthquakes; thus in an analogy to medicine subsidence is compared to the spread of a slow and silent cancer. It is of note that subsidence often occurs vertically, and is not noticeable in a short time. Generally, this phenomenon is local, and its mechanism depends on physical and natural processes. The natural process of land subsidence compounded by humanly induced stimulus factors has greatly intensified subsidence activity; in as such that in addition to surface morphological damage, it has brought about extensive financial and human losses, among which are compression and tension stress in buildings, flooding in downstream areas of coastal watersheds, damage to water well equipment due to the compression of sediments, the changing of the length and angle of pipes, and damage to water transport equipment (Martinez et al. 2013; Andreas et al. 2018). Due to its geographical location, the volume of precipitation in Iran is less than the amount of precipitation in many other parts of the world; moreover, the population surge in this country in recent years in tandem with an increase in social, economic, and industrial activities has led to a rise in the use of water resources, especially groundwater resources, beyond its existing capacity and potential. Over a 30 year period (1971–2001), the depth of groundwater aquifers has declined by at least 15 m (IDWRM, 2015) which means that the average aquifer level has dropped by an average of half a meter per year. According to existing reports, land subsidence due to the fall in aquifer levels, in some parts of Iran, has reached 50 cm per year (IWRM, 2016). Such a phenomena is unparalleled in the world, even though groundwater resources in large parts of central Iran, and the east, and south of Iran are the only sources of drinking water, irrigation, and for industrial use. In these areas, the existence of arid and semi-arid climates due to insufficient rainfall, the occurrence of long-term droughts, and the lack of permanent rivers have caused more than 90% of the water demand to be met through groundwater aquifers (Motagh et al. 2008). The high correlation between land subsidence, groundwater level reduction, and changes in the mechanical properties of subsurface layers has been widely identified and several attempts have been made to define this phenomenon, as a result studies on land subsidence in Iran are increasing, and are considered to be one of the research priorities among many companies and organizations involved in groundwater sustainability. Taking the current situation into account, it is believed that by recognizing the effective factors in the occurrence of this phenomenon, and creating a model for it, the formulating of a risk management program

to mitigate the damage caused by this phenomenon would be possible. It was assumed that by applying common numerical methods, which are mostly based on simplifying assumptions, it would be possible to model this phenomenon. In addition, remote sensing techniques and the use of satellite imagery which had been applied in a number of studies were also considered as a means of further developing the model. Lee et al. (2012) used an artificial neural network (ANN) to predict land subsidence and spatial modeling. The evaluation of the results of this model showed that the ANN model had a very high accuracy of 84.94%. Park et al. (2014) used ANN, frequency ratio (FR), logistic regression (LR), and a blend of these models to prepare the land-subsidence vulnerability map in South Korea. After preparing the maps, an ROC curve was used to determine the accuracy of the models. The results of the study showed that the accuracy of the combined models was higher than that of the models that were used independently. Pradhan et al. (2014) investigated land subsidence in the Kinta Perak area of Malaysia using GIS and RS. They also used the evidential belief function and a generalized additive model to assess the land subsidence process. The results showed that the evidential belief function had a greater accuracy as compared to other conventional methods. Castellazzi et al. (2016) studied land-subsidence vulnerability mapping using InSAR data along with hydrogeological data from five major Mexican cities. They concluded that the cities of Toluca and Aguascalientes have high subsidence rates of 10 cm per year, and the cities of Molarya and Celaya have a low subsidence rate of 2 to 5 cm per year. They also found that the rate of subsidence in the city of Querétaro decreased as a result of surface water management. Zhao et al. (2016) showed that the CART decision tree data-mining model with a correlation between variables, and the reduction in useless information could increase the validity of forecasting accuracy. A comparison of the CART model with the PSO-SVR model showed that the CART model has better accuracy and predictability in forecasting the groundwater level drop. Shrestha et al. (2017) assessed the risk of land subsidence in Kathmandu, Nepal. They showed that the northern and northeastern parts of the region are very sensitive to land subsidence, and an average of 6.1 mm subsidence occurred in these areas per year. Gonnuru and Kumar (2017) estimated the PsInSAR-based land subsidence in the Burgan oil field using TerraSAR-X. The main purpose of this study was to evaluate the ability of the PsInSAR technique to evaluate land subsidence in the Burgan oil field (Kuwait) between 2008 and 2011. The subsidence results of this study were compared with those of previous studies based on oil extraction in this region. Overall, it was found that the PsInSAR technique for monitoring land subsidence provides acceptable results after being corrected for atmospheric errors.

Taking into account the necessity of accuracy and speed in calculating and saving time, ML models are useful tools owing to their ability to learn effective factors and their relationships with dependent parameters. These models have a high capability in detecting the occurrence of subsidence phenomena in terms of using the estimation of distribution algorithms, data-based nature, and high repetition of the modeling process. In several GIS-based studies, these models proved their relative superiority over bivariate and multivariate statistical models. Oh et al. (2019) produced land subsidence vulnerability maps using ML models. To confirm the vulnerability map, the performance of the models was evaluated using an ROC curve. Among the models used, the logit boost model with a high accuracy of 91.44% provided the best performance among all others in preparing the subsidence risk map in South Korea. Zamanirad et al. (2019) investigated the effects of groundwater extraction on subsidence in the Kabudarahang Plain aquifer in the Hamadan Province, Iran using machine algorithms including random forest (RF), generalized additive model

(GAM), boosted regression trees (BRTs), and four anthropological and environmental fore-casters. The results showed that the GAM algorithm had a significantly higher accuracy than the BRT model. However, the performance of the RF forecast was lower than that of the GAM model.

. The aforementioned study area has faced many land-subsidence events due to a decrease in groundwater levels over the years (Rezaei et al. 2021). The authors of this paper have found no research to have been conducted in this area using machine models (artificial intelligence) for the subsidence phenomenon of the plain; thus, the authors considered the use of random forest (RF) and support vector machine (SVM) models based on identified subsidence locations, in addition to factors affecting their occurrence in the form of layers in the GIS environment in order to prepare a vulnerability map of the Kaboudrahang Plain to identify future solutions.

2 Materials and methods

2.1 Study area

The Kabudarahang study area, with a catchment area of 3470 km^2, is located in the north of the Hamadan province in Iran, and is considered to be part of the Salina catchment area. The study area is located between longitudes 48° 30´ and 48° 50´ E and between latitude 34° 50´ and 35° 40´ N (Fig. 0.1 (a, b, c)). Based on the numerical model map of the study area, the elevated area and the plain areas are 1217 km^2 and 2253 km^2 in extent, and the maximum and minimum elevation are 2834 asl and 1615 asl, respectively. The extent of the main water table of the Kabudarahang aquifer is 1471 km^2. The Kabudarahang Plain is located in a vast geological area in central Iran within the Sanandaj-Sirjan metamorphic zone; therefore, stones and the tectonic effects of both zones are clearly visible; whatsmore, a part of this area is in the Ghezel Ozan River tributary, and the Ali-Sadr Cave, a remarkable natural phenomenon, is to found in the lime formation in this region. This area has a semi-arid to arid and cold climate, and precipitation in this area is influenced by the Mediterranean winds, while the main sources of air humidity and rainfall are provided for by the western front. The average precipitation in the plain and the higher altitudes of Kabudarahang are respectively 304.2 mm and 340.2 mm. The average annual temperatures in the heights and plains of the Kabudarahang study area were calculated as 10.2 C° to 10.6 C°. Based on the evaporation curves in this region, the pan evaporations of plain and elevated areas are 2004.6 mm and 1837.8 mm per year (RWCH 2020). As a result of droughts and groundwater abstractions, this plain has experienced an extreme decrease in groundwater level in as such that over the past 30 years the aquifer unit hydrograph shows a 41.48 m loss in water level,in addition, during the past few years, large sinkholes have formed inside the plain. It was due to the current conditions that the area was selected by the authors of this paper for the land subsidence vulnerability mapping.

2.2 Land-subsidence inventory mapping and description of the modeling

A diagram of the process of land subsidence vulnerability modeling using machine models in the study area is presented in Fig. 2.

Field surveys (recorded by Global Positioning System (GPS) receivers) were used to determine the location of the actual subsidence in the study area. According to
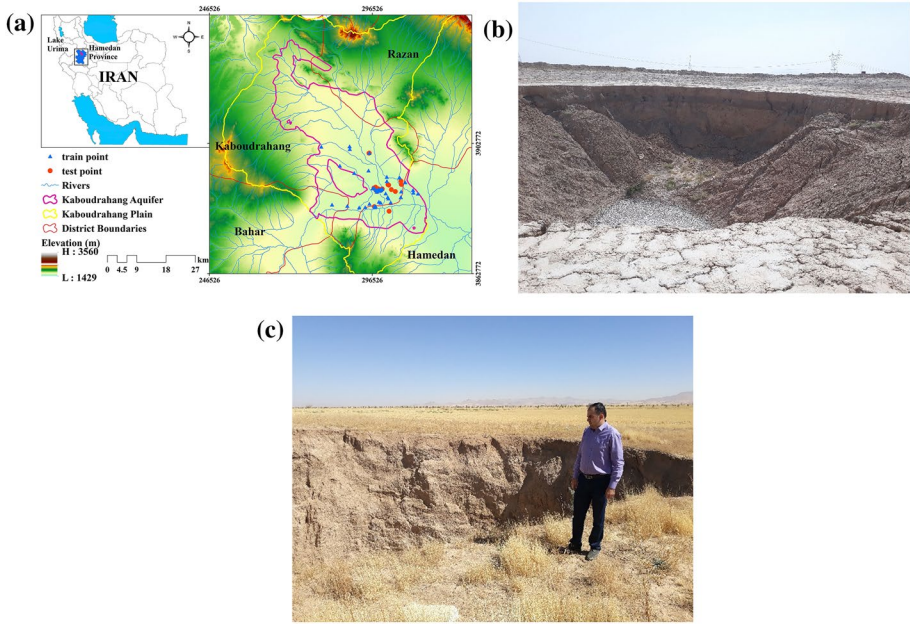
Fig. 1 Kabudarahang study area and aquifer and monitoring points (land subsidence features (**a**), (**b**) (**c**))
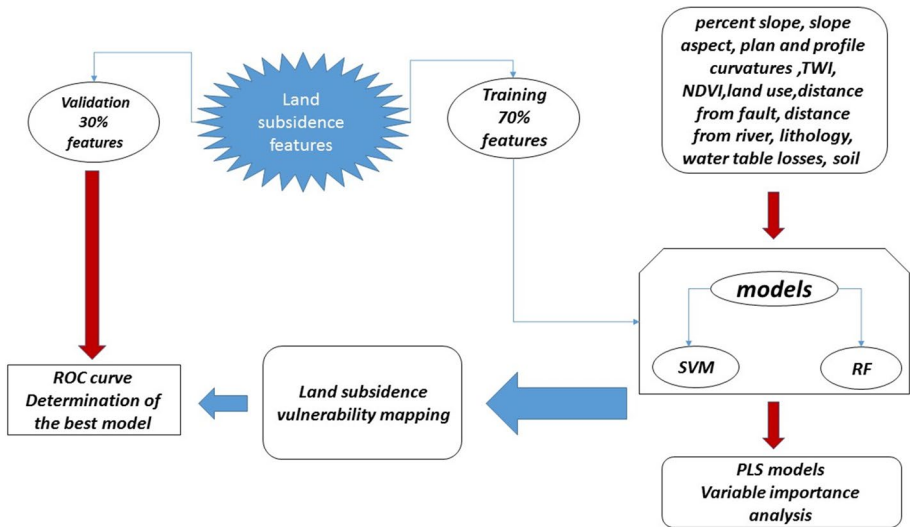

Fig. 2 Process of land-subsidence vulnerability modeling in the study are

investigation, 85 locations were identified of which 60 (70%) were selected for training and 25 (30%) were selected for testing (Zamanirad et al. 2019) (Fig. 1a). Randomly the locations are introduced to the model where the event occurred and not occurred is represented as the number (1) and (0), respectively (Mohammadi 2012). The most important action is

to determine the values of locations (0) and (1) from the map of independent variables (On base of Fig. 2, thirteen independent variables were classified and the value of species distribution in different classes was determined using GIS (Table 1)). At the first step, the map of each independent parameter is prepared and the values of each location are extracted one by one using the extract multi-value to point command in the GIS software and saved in Excel file format. Subsequently, based on the locations and independent variable layers, the model is generalized to the entire of study area and the classification map of the area is determined in terms of intensity and weakness in the field of vulnerability.

The values of each machine models were then independently computed according to their equations based on the proportion of pixels and species in each class. The derived values were then added to the various study layer classes, and a GIS map was prepared and using the GIS software's raster calculator function, the models were performed based on the Look Up maps. One base of 25 testing points and ROC curve, the comparison between models were done (Fig. 2). Finally with PLS model, influence of each of the independent parameters was determined.

The extended description of how to prepare the parameters and the process of modeling and evaluating them are as follows:

2.3 Determining the factors affecting land-subsidence events

The 10 m * 10 m DEM map was created on the basis of a topographic map with a scale of 1: 25,000. A river map was also prepared using DEM in ArcGIS 10.2. The layers, such as altitude, percent slope, slope aspect, plan, and profile curvatures were constructed with a spatial resolution of 10 m * 10 m based on the DEM map. All layers were classified using the natural break method in ArcGIS 10.2 software (Ghorbanzadeh et al. 2018; Rahmati et al. 2019). The altitude map of the study area was categorized into four classes:(1) 1620.95–1670.19, (2)1670.19– 1713.40, (3) 1713.40–1763.64, and (4)1763.64–1877.20 m (Fig. 3a). The slope percent map was categorized into five classes: (1) 0–0.34, (2) 0.34–1.03, (3)1.03–2.19, (4) 2.19–4.95 and (5) 4.95–29.39 (Fig. 3b).

The slope aspect was correlated with the solar energy in the region and has been categorized into nine classes (Dai and Lee 2002) (north, northeast, south, east, southeast, west, southwest, northwest, and flat; Fig. 4).

Profile curvature indicates the intensity of flow, amount of sediment, and amount of erosion (Yesilnacar 2005). The profile curvature map was categorized into three classes: (1)−4.73–0.12, (2) −0.12–0.06, and (3)0.06–4.74 (Fig. 5a). Plan curvature plays an important role in contributing to terrain instability and is created based on the intersection of a horizontal plane and the ground surface (Fernandez et al. 2004; Vijith and Madhu 2008). The plan curvature of the study area was also categorized into three classes: less than − 0.01 (concave), 0.01 –0.01 (flat) and larger than 0.01 (convex) (Conforti et al. 2014) (Fig. 5b).

To produce a distance from the river map, a network of streams and rivers was prepared from a topographic map and digitized in ArcGIS software, and the map of the network of streams was modified from the DEM of the region using SAGA-GIS software. The distance from the river map was distributed into four classes: (1) 0–152.64, (2)152.64–320.15, (3) 320.15–524.30, (4) 524.30–1220.90 (Fig. 6a).

In addition, a TWI map was built based on Eqs. (1) (Moore et al. 1991) using the SAGA-GIS software (Fig. 6b). The index indicates the level of participation of areas in the water outflow of the basin (Bevan and Kirkby 1979).

Table 1 previous studies with ML and input factors

Researchers	Country	Factors	Number of factors	Models*
Lee et al (2012)	Korea	Slope, Distance from drift, Groundwater level, Rock-mass rating, Depth of drift, Distance from lineament, Geology, Land use	8	ANN
Pradhan et al (2014)	Malaysia	Stream power index, Altitude, Slope, Aspect, Lithology, Distance from the fault, Distance from the river, NDVI, Soil, TWI, and Land use/cover	11	FR, EBF
Park et al (2014)	Korea	Geology, Slope, Depth of drift, Distance from drift, Groundwater level, Permeability,Land use	7	FR, LR, ANN
Ozdemir (2015)	Turkey	Geology, Altitude,Aspect,Fault density,Distance from fault, Groundwater level	6	FR
Ghorbanzadeh et al(2018)	Iran	Distance to well,Distance to fault,Altitude,Distance to river,Water depth, Land use,Slope	7	ANFIS
Rahmati and et al. (2019)	Iran	Distance from qanat, Land use, Distance from afforestation project,Distance from fault, Lithology, Groundwater level	6	MaxEnt,GRAP
Zamanirad et al (2019)	Iran	Altitude, Distance from river, Lithology, Groundwater level	4	BRT,GAM, RF
This study	Iran	Slope, Aspect, Altitude, TWI, Plan curvature, Profile curvature, Distance from fault, Distance from river, Lithology, Groundwater level, NDVI, Land use, Soil	13	RF, SVM

*Frequency ratio (FR), logistic regression (LR), artificial neural network (ANN), evidential belief function (EBF), support vector machine (SVM), boosted regression trees (BRTs), generalized additive model (GAM), maximum entropy (MaxEnt), genetic algorithm rule-set production (GARP), random forest (RF), adaptive neuro-fuzzy inference system (ANFIS)

Fig. 3 Altitude (**a**) and slope percent (**b**) layers
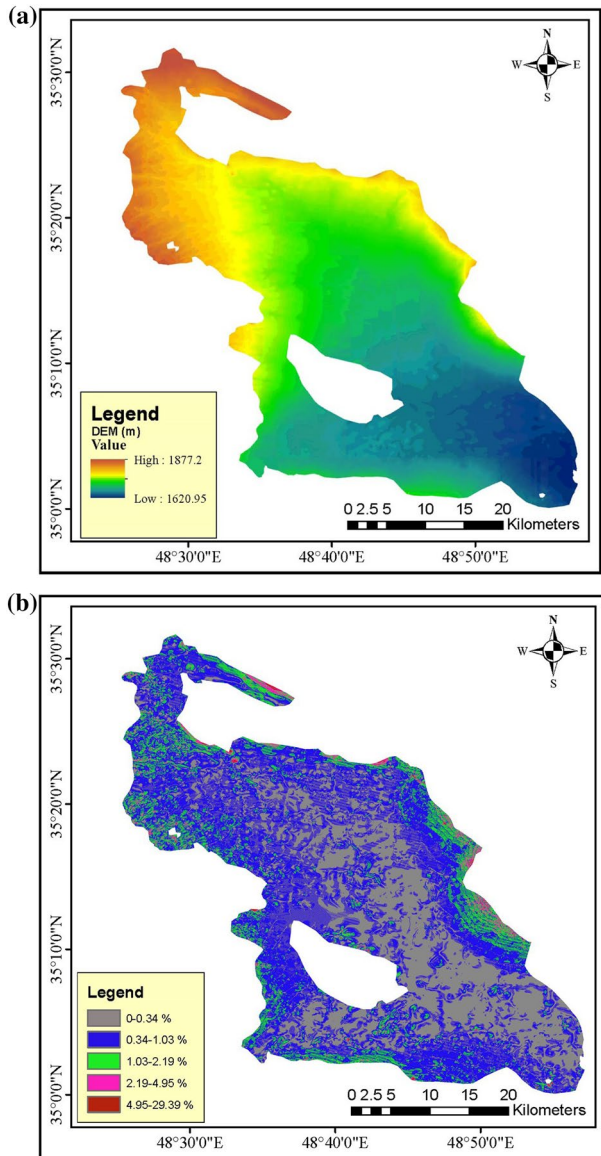
$$TWI = Ln\left(\frac{\alpha}{Tang\beta}\right) \tag{1}$$

TWI: Topographic Wetness Index. where α and β are the specific catchment area and slope angle of the area, respectively.

The geological map on a 100,000-scale of the region was acquired from the Geological Survey of Iran. Based on Table 2 and Fig. 7a, 16 different lithological classes can be considered for the Kabudarahang aquifer. Fault regions are also perceived as very important factors

Fig. 4 Slope aspect layer
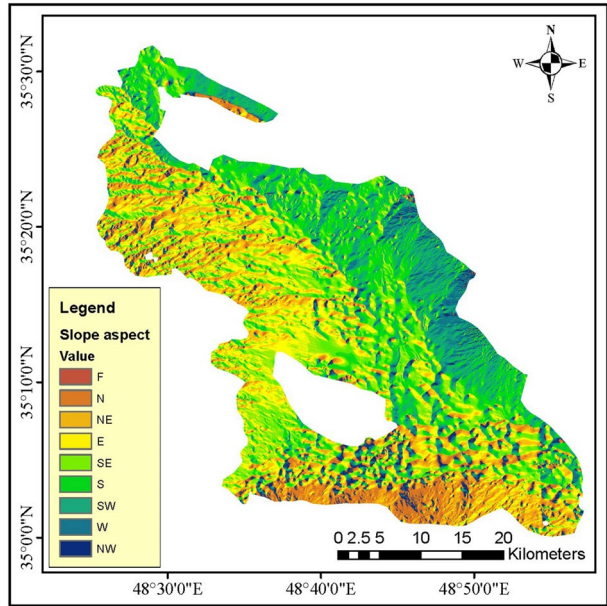
for land subsidence, sinkholes, and landslide vulnerability (Cevik and Topal 2003; Yilmaz 2009; Santo et al. 2011; Conforti et al. 2012; Ozdemir 2016). The fault map was exported from the geological map, and the distance from the fault map was prepared using ArcGIS 10.2 software. The distance from the fault was distributed into four classes: (1) 0–152.64, (2) 152.64–320.15, (3) 320.15–524.30, (4) 524.30–1220.90 (Fig. 7b).

The land use map and soil texture (1:250,000 scale) of the region were obtained from the Agricultural Research Center of the province and turned into raster layers. The land use classifications are bare rock, urban, barren land, agricultural, and range land. Based on the land use map, 3%, 7.5%, 88.5%, and 1% of the aquifers were in the urban, barren, agricultural, and range land zones, respectively (Fig. 8a). The soil texture was categorized into 10 classes, as shown in Fig. 8b and Table 3.

NDVI represents surface reflectance and can quantitatively compute vegetation growth and biomass (Hall et al. 1995; Yilmaz 2009). The NDVI values were calculated using Eqs. (2)) (Rouse et al. 1974; Tucker 1979):

$$NDVI = (IR - R)/(IR + R) \tag{2}$$

where IR and R are the infrared and red portions of the electromagnetic spectrum, respectively. The NDVI map was divided into six classes and was prepared using Google Earth Engine and 34 Landsat satellite images with radiometric and atmospheric correction, taking into account at least 10% of the cloud cover (Fig. 9a). The map of groundwater changes in the plains (groundwater level drawdown) was determined and calculated using both piezometric well data from 1989 to 2021 and the IDW interpolation method (Eq. 3 (Khan and et al. 2013; Park et al. 2014).

$$\lambda_i = D_i^{-\alpha} \Big/ \sum_{i=1}^{n} D_i^{-\alpha} \tag{3}$$

Fig. 5 Profile (**a**) and plan curvature (**b**) layers
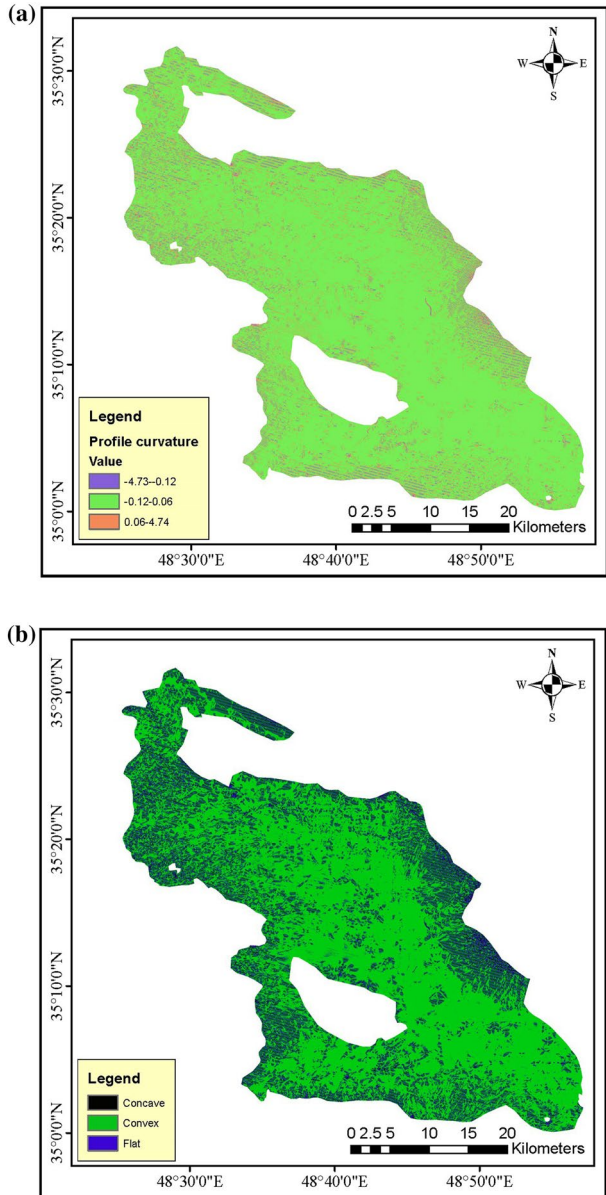
where λ_i is the point of I, D_i is the distance between point i and an unknown point, and α is equivalent to the weighing power.

The groundwater level drawdown was divided into four categories with intervals of (1) −86.44- −61.08, (2) −61.08—−44.39, (3) −44.39- −28.37, and (4) −28.37- −1.34 (Fig. 9b).

Fig. 6 Topographic wetness index (**a**) and distance from river (**b**) layers
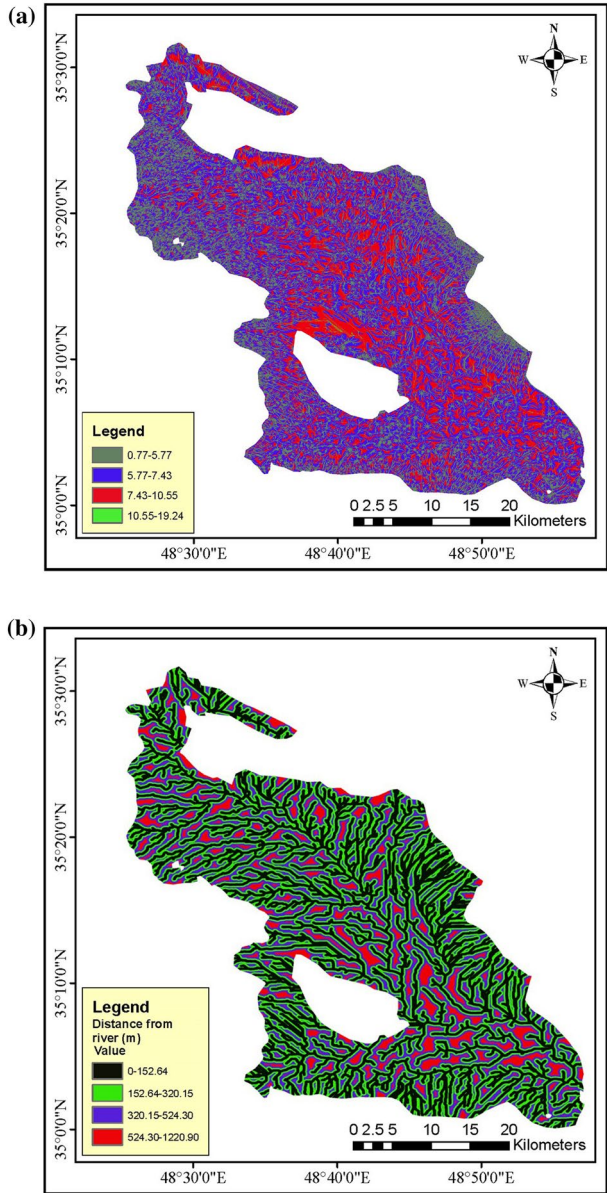
2.4 Determining the weight of classes of each factor using the FR model

Using a bivariate FR model, the weights of the classes for each effective factor were obtained (Eq. 4) (Bonham-Carter 1994; Mezughi et al. 2011).

$$FR = \left(\left(A/B \right) / \left(C/D \right) \right) \tag{4}$$

Table 2 Lithological characteristics

Class	Lithology	Class	Lithology
1	Basalt Pyroclastics	9	Young-medium& old terraces
2	Nummulitic limestone	10	Old terraces
3	Pyroclastics	11	Late-phyllite
4	Conglomerate	12	Marl
5	Pyroclastics-Lava	13	Conglomerate
6	Latite	14	Apelite
7	Gypsum	15	Andesite-pyroclastic
8	Conglomerate-sandstone	16	Volcanic rocks

where A is the amount of subsidence in each class, B is the total number of subsidence locations in the area, C is the area of each class, and D is the total area of the region.

2.5 Spatial modeling of the land-subsidence using ML models

For land-subsidence vulnerability mapping, SVM and RF ML models were used in R 3.6.0 statistical software.

2.6 Support vector machines (SVM)

One of the features of the SVM method is the joint classification and regression operations. Algorithms of the SVM model provide a general method for estimating functions. Their main purpose was to solve the second-order optimization problems. A set of separate linear training cells (Xi) was selected (i = 1, 2,…, n) Xi). The training cells consisted of two classes, defined as Yi = ± 1 (Cristianini 2000). The SVM model aims to determine the n-dimensional separation sheet that can establish the maximum distance and gap between the two classes and reduce the W variable. The mathematical expression for this is presented in Eqs. (5) and (6) (Xu et al. 2012):

$$1/2 \parallel w \parallel^2 \tag{5}$$

$$Y_i\big((W.X_i) + b\big) \geq 1 \tag{6}$$

where $\parallel W \parallel$ is the absolute value of the normal separation sheet, and b is the numerical base. In order to solve the above problem, a Lagrangian relation is used, which contains an incremental coefficient called λ_i. The goal of this relationship is to reduce the value of Lagrangian L by decreasing the coefficients of W and b and increasing λ_i. Thus, the general form of equation is given by Eq. (7) (Vapnik 2013):

$$L = \frac{1}{2 \parallel W \parallel^2} - \sum_{i=1}^{n} \lambda_i Y_i\big((W.X_i) + b\big) - 1) \tag{7}$$

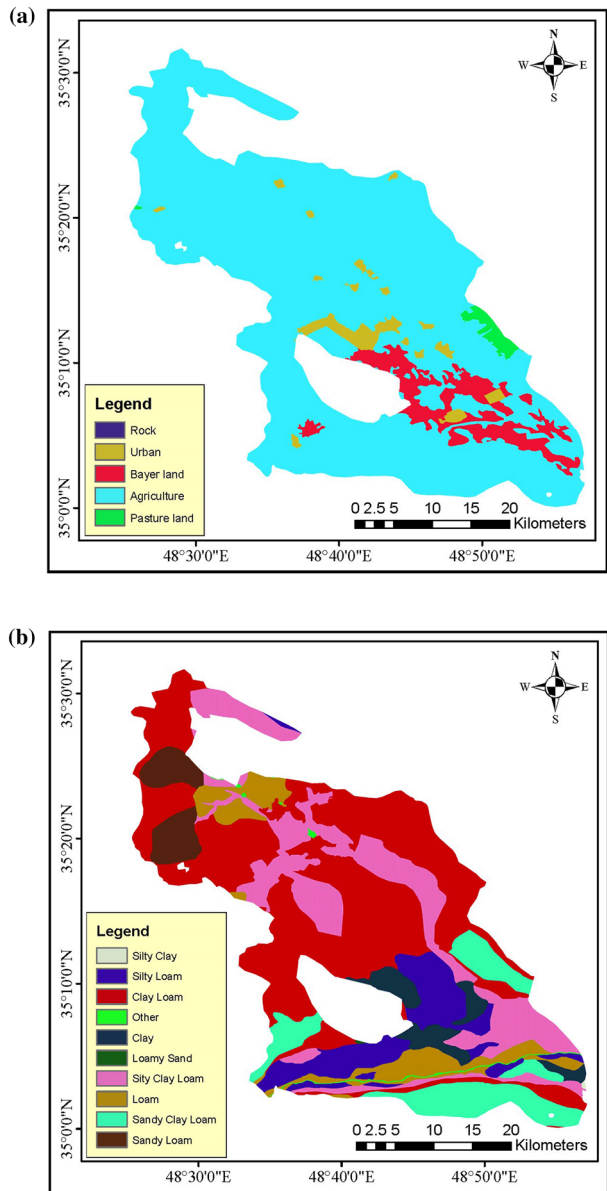**Fig. 7** Lithological (**a**) and distance from fault (**b**) layers
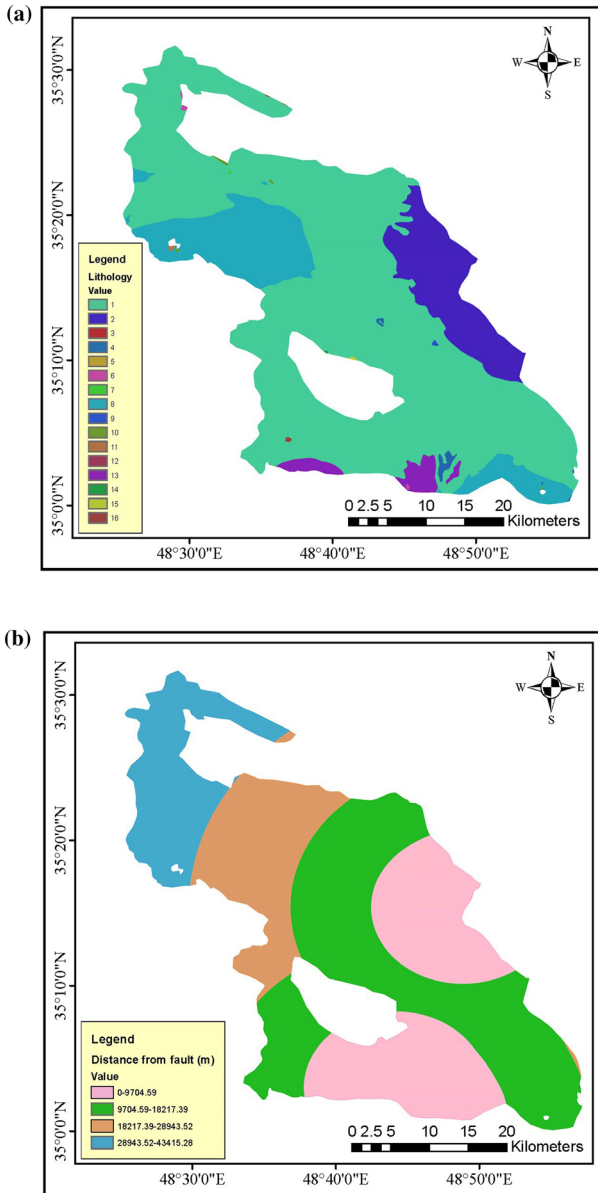
2.7 Random forest model (RF)

RF is a modern type of tree-based method that includes a multitude of classification and regression trees. The RF is created using a set of trees by considering n independent observation data:

$$\left(Y_i,\ X_i\right),\ i = 1, \ldots, n$$

Fig. 8 Land use (**a**) and soil texture (**b**) layers

This method is a compound of several decision trees, and by using a large number of bootstrap methods (for example, 2000 times) from the set of n samples of the initial observational data, performs sampling along with the placement. Then, a tree is spread on each sample of Bootstrap. After making the whole tree, the test data are introduced to the tree, and an output is obtained for the input vector of each tree. The final output of the model was calculated by averaging the outputs. Taking into account the experimental distribution of the outputs, the values of the percentiles and the range of uncertainty were calculated.

Table 3 Soil texture

Percentage	Soil type
0.1	Sandy clay
3.8	Sandy loam
17.8	Clay loam
1.8	Other
16.6	Clay
0.1	Loamy sand
8.1	Sity clay loam
30.2	Loam
4.0	Sandy clay loam
17.4	Sandy loam

The RF regression tree method is an efficient forecasting method, especially when the number of observations is relatively low compared to the number of forecasters (Svetnik et al. 2003).

2.8 Preparation of data layers

All the desired data layers were converted to 'asc' format using ArcGIS software to enter R software, and the land-subsidence vulnerability map was prepared using SVM and RF models in the R software. To export the map obtained from the SVM and RF models, after modeling and running in R software, the output of weights for each pixel (pixel by pixel) was transferred to the GIS software environment, based on which the final map was prepared. The output weight was in the range of zero to one. Pixels with zeros and ones were considered as completely stable and completely unstable regions, respectively. Finally, the vulnerability map obtained was divided into four classes: low, medium, high, and very high vulnerability. (Komac 2006; Sezer et al. 2011).

2.9 Evaluation of the final ML maps

The ROC curve characterizes the relative performance of each model. ROC is a curve in which the ratio of pixels that correctly predict the occurrence or non-occurrence of subsidence events is plotted on the horizontal axis (specificity), while the vertical axis shows the ratio of incorrect predictions (sensitivity) (Hanley 2014).

This curve was calculated and plotted using SPSS software. The area below this curve is called the AUC, and the model with the highest AUC has a higher relative performance. The AUC is equal to 0.5, indicating a neutral model; as this value approaches one, the efficiency of the model increases (Negnevitsky 2002).

2.10 Determining the importance of the parameters using the Partial least squares regression (PLS) model

The partial least squares regression model was used to eliminate the invalidity of general regression equations because of the existence of linearity in independent or explanatory variables. In this method, new orthogonal components are created, which are a

Fig. 9 NDVI index (**a**) and Drawdown of groundwater level (**b**) layers
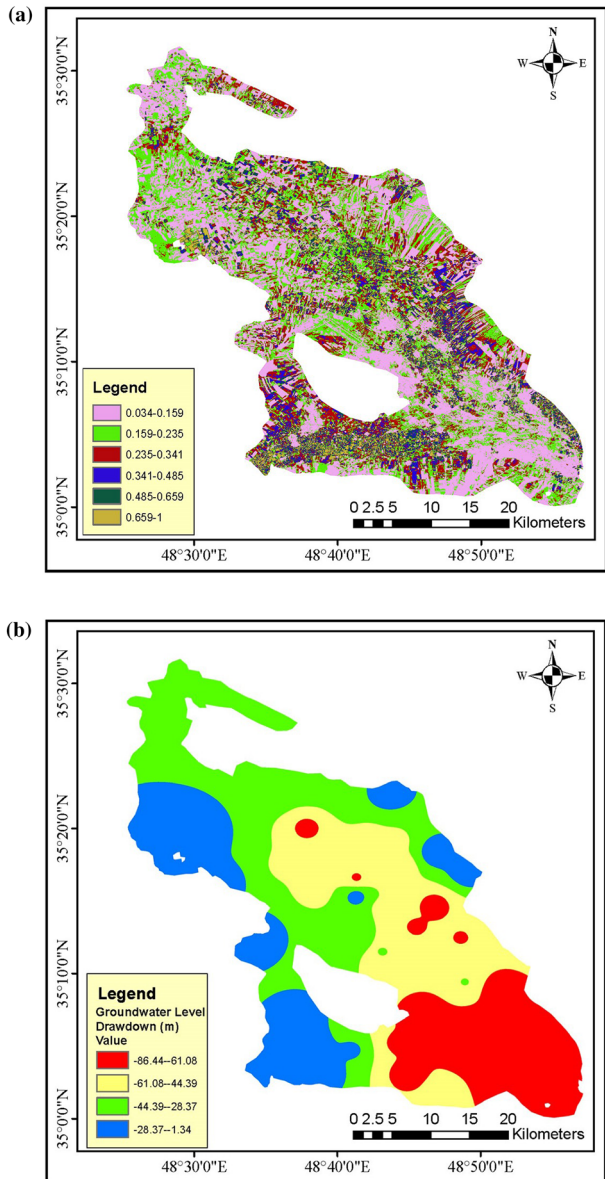
linear combination of the primary variables. Subsequently, these components were used to construct a regression equation (Vinzi and et al. 2010). In the PLS regression model, standard coefficients of variable importance in the projection (VIP) reflect the effect of individual Xs on Ys and can be easily seen in the PLS diagram. Therefore, the most effective variables and their degrees of importance were identified rapidly (Wold et al. 2010). In the current study, the PLS model in R software was used to determine the effect of each parameter on the subsidence of the Kabudarahang aquifer. Finally,

considering the importance of controlling the parameters affecting subsidence, altering behavior, supervisory methods, and planning strategies were considered.

2.11 GRACE data using

As changes in groundwater aquifers affect the gravity of the Earth, the level 2 data of the GRACE satellite, which measures the monthly gravity of the Earth, can be used as an indicator of groundwater level variations in the region (Voss et al. 2013; Joodaki and Swenson 2014; Saber et al. 2018). Groundwater fluctuations in the Kaboudrahang Plain were investigated using GRACE satellite data from the Google Earth Engine platform and its environmental coding. Data from three GFZ, CSR, and JPL centers were used. According to the nature of the platform's performance, the output of platform is the average water level change of the study area. The corresponding link of the groundwater change calculation code in the Google Erath Engine platform is also as follows: https://code.earthengine.google.com/f1d4e70c9f90a9a6d036eae5ca818437..

3 Results

3.1 Weight interpretation of classes for each effective factor

Owing to the identification of factors affecting the phenomenon of land subsidence, the frequency ratio was calculated based on Eq. 4. As shown in Table 4, in the lowest elevation class, the frequency of observation for subsidence locations was higher, indicating more subsidence in the plains and aquifers as compared to the elevated areas, which is in line with the studies of Dogan and Yelmaz (2011), Park et al. (2014) and Pradhan et al. (2014). On a lower slope, the number of recorded subsidence is higher, which is in line with the studies of Kim et al. (2009) and Pradhan et al. (2014). In addition, in the slope with class 2.19 to 4.95, the frequency ratio was 1.06. The slope aspect has a significant effect on soil moisture retention, so it has a direct effect on soil strength and vulnerability due to land subsidence (Pradhan et al. 2014). Accordingly, the eastern and southwestern slopes at the levels of their classification classes showed a higher frequency ratio than the other aspects. In addition, on the western slope at a rate of 0.54. In the present study, by increasing the distance from the faults, the frequency ratio decreased; therefore, this shows a direct effect of proximity to the faults in land subsidence (Hack 1965; Santo et al. 2011). According to the normalized vegetation index, the maximum frequency ratios in the 0.034–0.159 and 0.659–1 classes are 1.44 and 1.38, respectively, which indicates a higher rate of land subsidence in areas without vegetation and in areas that are irrigated by groundwater. TWI was considered as another factor. This index is a hydrological item that indicates the spatial changes in wetness in the drainage basin and in places where the rate of this index is higher, the amount of runoff will also be higher. In other words, only in times of drought, areas with high indices are involved in the production of runoff, and in areas where the index is low, runoff will occur only in saturated conditions (Yilmaz et al. 2013). The highest frequency ratio was 1.68 in class 7.43–10.55.

One of the main causes of land subsidence is the excessive use of groundwater (Ozdemi 2015), which is defined as groundwater level drop, while the highest frequency ratio was in the class with the highest groundwater level drop (−61.08–86.44). The plan curvature of the earth's surface in the convex surface class had a frequency ratio of 1.22,

Table 4 The frequency ratio of classes of factors affecting the land-subsidence

Factors	Classes	No. Pixels	Pixels (%)	No. LS	LS (%)	FR
Elevation (m)	1620.95–1670.19	3,892,218	31.07	54.00	90.00	2.90
	1670.19–1713.40	3,354,038	26.78	5.00	8.33	0.31
	1713.40–1763.64	3,384,890	27.02	1.00	1.67	0.06
	1763.64–1877.20	1,894,373	15.12	0.00	0.00	0.00
Slope percent	0–0.34	4,907,079	39.18	43.00	71.67	1.83
	0.34–1.03	6,040,683	48.23	13.00	21.67	0.45
	1.03–2.19	1,365,583	10.90	3.00	5.00	0.46
	2.19–4.95	197,454	1.58	1.00	1.67	1.06
	4.95–29.39	14,720	0.12	0.00	0.00	0.00
Distance from faults (m)	0–9704.59	3,534,443	28.22	40.00	66.67	2.36
	9704.59–18,217.39	4,891,844	39.06	19.00	31.67	0.81
	18,217.39–28,943.52	2,182,567	17.42	1.00	1.67	0.10
	28,943.52–43,415.28	1,916,665	15.30	0.00	0.00	0.00
NDVI	0.03–0.15	4,770,067	38.08	33.00	55.00	1.44
	0.15–0.23	3,807,746	30.40	10.00	16.67	0.55
	0.23–0.34	2,098,454	16.75	6.00	10.00	0.60
	0.34–0.48	909,297	7.26	6.00	10.00	1.38
	0.48–0.65	540,271	4.31	2.00	3.33	0.77
	0.65–1	399,330	3.19	3.00	5.00	1.57
TWI	0.77–5.77	3,511,675	27.70	10.00	16.67	0.60
	5.77–7.43	5,578,431	44.01	24.00	40.00	0.91
	7.43–10.55	3,269,825	25.80	26.00	43.33	1.68
	10.55–19.24	316,072	2.49	0.00	0.00	0.00
Piezometric wells data	−86.44- −61.08	2,639,834	21.08	49.00	81.67	3.87
	−61.08- −44.39	3,319,266	26.50	4.00	6.67	0.25
	−44.39- −28.37	3,920,595	31.30	5.00	8.33	0.27
	−28.37- −1.34	2,645,824	21.12	2.00	3.33	0.16
Plane curvature	Concave	1,256,506	10.03	2.00	3.33	0.33
	Convex	9,781,047	78.09	57.00	95.00	1.22
	Flat	1,487,966	11.88	1.00	1.67	0.14
Profile curvature	−4.73- −0.12	504,105	4.02	1.00	1.67	0.41
	−0.12–0.06	10,740,108	85.75	59.00	98.33	1.15
	0.06–4.74	1,281,306	10.23	0.00	0.00	0.00
Distance from river (m)	0–152.64	4,824,368	38.52	22.00	36.67	0.95
	152.64–320.15	3,985,720	31.82	16.00	26.67	0.84
	320.15–524.30	2,700,619	21.56	16.00	26.67	1.24
	524.30–1220.90	1,014,812	8.10	6.00	10.00	1.23
Aspect	Flat	54	0.00	0.00	0.00	0.00
	N	1,250,987	9.99	1.00	1.67	0.17
	NE	1,524,402	12.17	9.00	15.00	1.23
	E	2,324,231	18.56	16.00	26.67	1.44
	SE	2,155,802	17.21	11.00	18.33	1.07
	S	2,436,874	19.46	9.00	15.00	0.77
	SW	1,575,411	12.58	10.00	16.67	1.33

Table 4 (continued)

Factors	Classes	No. Pixels	Pixels (%)	No. LS	LS (%)	FR
	W	769,832	6.15	2.00	3.33	0.54
	NW	487,926	3.90	2.00	3.33	0.86
Land use	Rock	10	0.00	0.00	0.00	0.00
	Urban	372,461	2.97	1.00	1.67	0.56
	Bayer land	946,316	7.56	21.00	35.00	4.63
	Agriculture	11,072,910	88.41	38.00	63.33	0.72
	Pasture land	132,958	1.06	0.00	0.00	0.00
Soil texture	Silty Clay	3679	0.03	0.00	0.00	0.00
	Silty Loam	1,191,861	9.52	22.00	36.67	3.85
	Clay Loam	5,516,260	44.04	6.00	10.00	0.23
	Other	57,148	0.46	0.00	0.00	0.00
	Clay	514,904	4.11	14.00	23.33	5.68
	Loamy Sand	3126	0.02	0.00	0.00	0.00
	Silty Clay Loam	2,511,358	20.05	12.00	20.00	1.00
	Loam	936,342	7.48	6.00	10.00	1.34
	Sandy Clay Loam	1,250,033	9.98	0.00	0.00	0.00
	Sandy Loam	539,944	4.31	0.00	0.00	0.00
Lithological units	Basalt pyroclastics	9,229,563	73.69	59.00	98.33	1.33
	Nummulitic limestone	1,285,918	10.27	1.00	1.67	0.16
	Pyroclastics	987	0.01	0.00	0.00	0.00
	Conglomerate	41,431	0.33	0.00	0.00	0.00
	Pyroclastics-Lava	93	0.00	0.00	0.00	0.00
	Latite	7231	0.06	0.00	0.00	0.00
	Gypsum	1196	0.01	0.00	0.00	0.00
	Conglomerate-sandstone	1,690,204	13.50	0.00	0.00	0.00
	Young-medium& old terraces	3369	0.03	0.00	0.00	0.00
	Old terraces	8302	0.07	0.00	0.00	0.00
	Late-phyllite	3399	0.03	0.00	0.00	0.00
	Marl	60	0.00	0.00	0.00	0.00
	Conglomerate	246,316	1.97	0.00	0.00	0.00
	Apelite	1015	0.01	0.00	0.00	0.00
	Andesite-pyroclastic	2614	0.02	0.00	0.00	0.00
	Volcanic rocks	2957	0.02	0.00	0.00	0.00

which accounted for a larger share than the other classes. In a class of -0.12 to 0.06, the amount of frequency ratio is at its maximum point because of the profile curvature. The distance from the river was investigated because of the effect of water on land subsidence. In the class with 320.15–524.30 m distance from the river, the FR is 1.24. In terms of land use, the highest frequency ratio (4.63 belongs to barren lands. In terms of investigating the texture of the soils of the region for clay soils, the highest frequency ratio was 5.68. The subsidence with ratios of 1.33 and 0.16 were found in the Basalt-Pyroclastics and the nummulitic limestone class, respectively.

3.2 Land-subsidence vulnerability maps

Figure 10 shows the output of the machine learning models and four classes of low, medium, high, and very high subsidence. Based on the number of pixels in each class of the RF model, 45% of the study area was assessed as low (23%), moderate (20%), and very high (12%) in terms of the sensitivity of the aquifer to subsidence (Fig. 10a). The

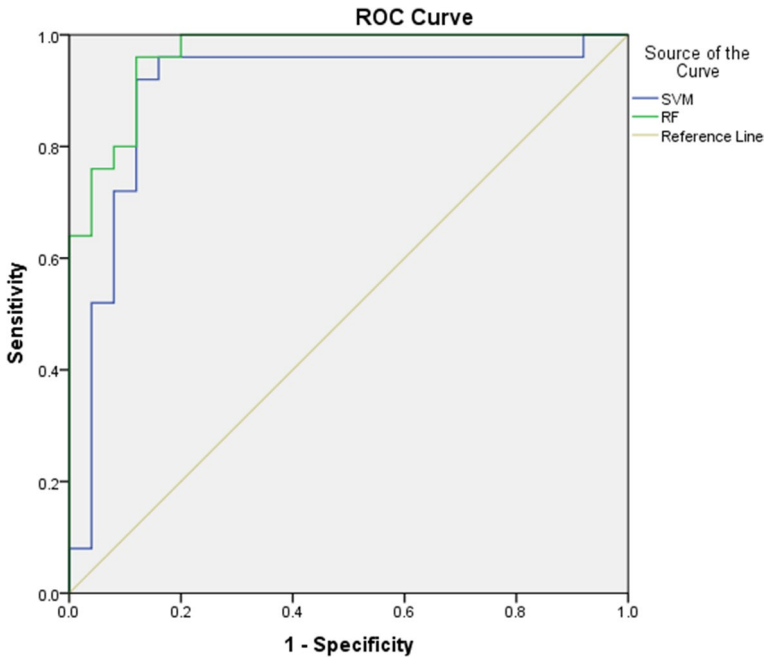**Fig. 10** Land-subsidence vulnerability (RF (**a**) and SVM (**b**)) layers
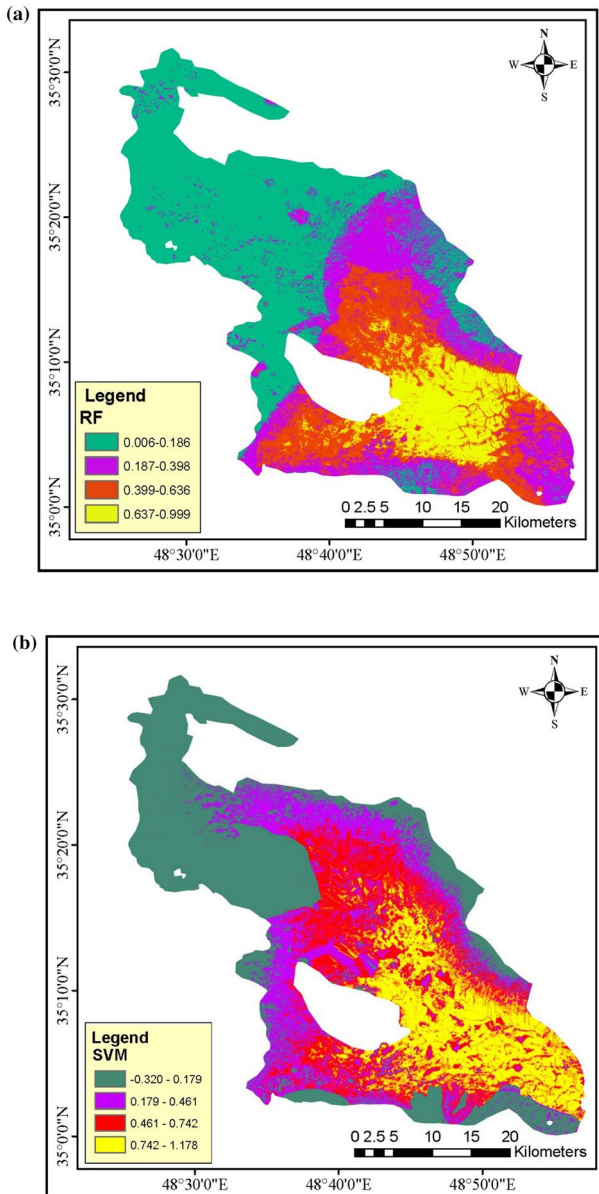
Fig. 11 Comparison of machine model accuracy in the ROC curve

Table 5 Comparison of ROC test output values between SVM and RF models

Models	AUC	Standard deviation error	Statistical significance	Asymptotic 95% confidence Interval	
				Lower bound	Upper bound
RF	0.965	0.022	0	0.922	1
SVM	0.899	0.052	0	0.797	1

results of the SVM model showed that 40% of the aquifer was low, 20% was moderate, 22% was high, and 18% was very high (Fig. 10b).

3.3 Assessment of the built ML models

The assessment of both ML models based on the ROC curve in the SPSS software is presented in Fig. 11 and Table 5. In the RF and SVM models, the area under the ROC curve (AUC) was determined to be 0.96 and 0.89, respectively. The performance of both models was highly suitable in this situation. However, the RF model performed better than the low-error SVM model. The higher ability of the RF model as compared to other machine models has been emphasized in the research conducted by Kotsiantis & Pintelas (2004), and Stumpf &Kerle (2011).

3.4 Importance of Variables using the PLS model

The results of the PLS model showed the major effect of topography on the prioritizing of abstraction, followed by groundwater changes as the second priority on the subsidence process as compared to other parameters (Fig. 12). The effect of the parameter of groundwater change on the rate of subsidence in this study has also been emphasized by Motagh et al. (2008), who investigated the main cause of subsidence in the plains of Iran. In addition, the results of Zamanirad et al. (2019) on subsidence modeling with machine models indicate the high impact of groundwater change as a major factor on the aforementioned phenomenon.

3.5 Groundwater changes in Kabudarahang plain

Groundwater changes in the plain were identified and presented based on GRACE data as per Fig. 13. Over 15 years, the trend of change has decreased. To make the changes tangible, JPL data were prepared from charts based on a polynomial trend in Excel. These changes indicate that the average water thickness has decreased by 30 cm over the 15 year period, which in itself has caused subsidence in the Kabudarahang area.
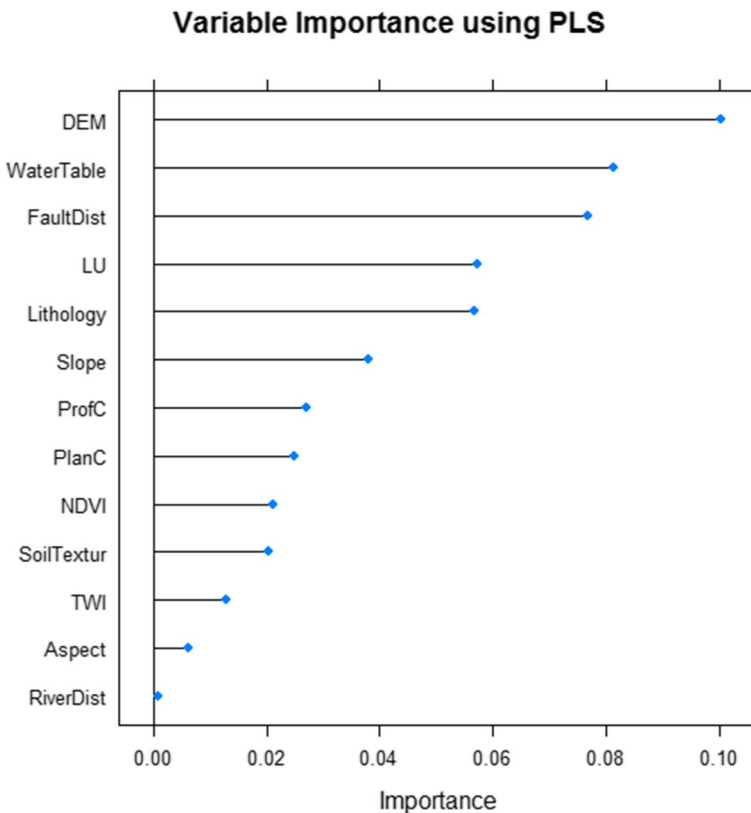

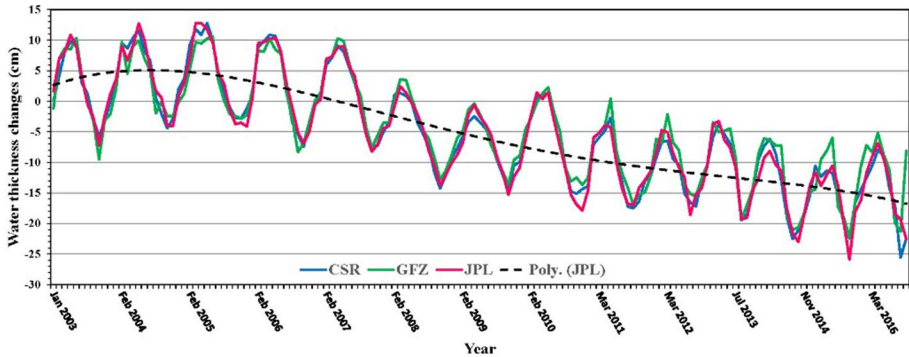**Fig. 12** The variables importance

Fig. 13 Groundwater change using GRACE satellite images

4 Discussion

Intricate relationships between the dependent factor (land subsidence) on the one hand and independent factors on the other, in addition to the complex application of the above-mentioned models, have produced differences in output when applying simple methods, and more sophisticated models. Machine models show different accuracies regarding the nature and relation of dependent and independent variables, as well as the number of independent variables (Teartisup and Kerdsueb 2013; Zhu et al. 2013). In this study, the number of input layers increased compared to other studies, and 13 layers were considered as the model input (Table 1); thus, the area under the ROC curve reached 96.5% and the accuracy was higher than that of the above-mentioned studies. It can be therefore concluded that adding independent parameters with a high correlation increases the model accuracy and decreases the errors. In addition, the accuracy of the RF model, as compared with RBDT, BRT, and CART, has been corroborated by researchers such as Rahmati et al. (2019). Therefore, considering the above-mentioned matters, in this study, both the RF and SVM models with an increase in input layers (independent variables) produced higher accuracies. Investigations have shown that the independent parameter of groundwater loss following the parameter of elevation has a greater effect on subsidence than other independent parameters (Mousavi et al. 2001; Karimzadeh 2015; Figueroa-Miranda et al. 2018; Ghorbanzadeh et al. 2018). Since elevation is ranked highest, it shows the hidden effect on precipitation when compared with other factors. In this research and others (Rahmati et al. 2019), groundwater level loss has been defined as an important and manageable factor that should be examined based on other decision-making items. In this case, drought in most parts of Iran and the limited period of rainfall in some central parts of Iran, compounded by the application of groundwater for agricultural and drinking purposes, aid water resource managers in the saving of additional water (artificial recharge projects), and the applying of such water in appropriate situations (Shi et al. 2018). In addition, determining the vulnerability to subsidence and the combination of vulnerability to subsidence and the concurrent real conditions (wells, subterranean canals, buildings, facilities, etc.) it is possible to define the existing risk of destruction. It is proposed that for future research applying other machine models as a modeling pattern, and investigating the relationship between groundwater level loss and subsidence should be considered.

5 Conclusions

Land subsidence in the Kabudarahang Plain in the Hamadan Province, Iran, was investigated using RF and SVM ML models. Although the accuracy of both models was deemed suitable, the RF model showed higher accuracy and efficiency than the SVM model in determining the vulnerability map of the Kabudarahang aquifer. This research confirmed the results obtained in previous research and reports regarding the fact that the decrease in groundwater level is the main reason for subsidence. The drop in groundwater levels has increased since 1988 based on observational well data, and since 2003, based on GRACE satellite data, due to excessive abstraction from groundwater aquifers as typified by the existence of unauthorized wells, in as such that extensive subsidence has occurred and many sinkholes have appeared in the region. In future studies, it is proposed that the application of other machine models be compared with the models used in this study. In addition, with due regards to the fact that the main factor identified in this study which can be administered to control land subsidence in the aforementioned area is the drop in groundwater level, it is proposed that governmental authorities should seek migratory measures by optimally planning for the management and operation of groundwater based changing crop patterns and other applications.

Declarations

References

Andreas HZ, Abidin H, Gumilar I, Teguh P, Sidiq TA, Sarsito D, Pradipta D (2018) Insight into the Correlation between Land Subsidence and the Floods in Regions of. Indonesia. https://doi.org/10.5772/intechopen.80263

Bevan KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology. Hydrol Sci Bull 24(1):43–69

Bonham-Carter GF (1994) Geographic information systems for geoscientists: modeling with GIS. In: Bonham-Carter F (ed) Computer methods in the geosciences. Pergamon, Oxford

Castellazzi P, Arroyo N, Martel R, I. Calderhead A, C. L. Normand J, Gárfias J, Rivera A (2016) Land subsidence in major cities of Central Mexico: Interpreting InSAR-derived land subsidence mapping with hydrogeological data. Int. J. Appl. Earth Obs. Geoinformation, 47, 102– 111

Cevik E, Topal T (2003) GIS-based landslide susceptibility mapping for a problematic segment of the natural gas pipeline, Hendek (Turkey). Environ Geol 44:949–962

Chai J, Suddeepong A, Liu MD, Yuan DJ (2017) Effect of daily fluctuation of groundwater level on land-subsidence. Int J Geosynth Ground Eng 3 (1), 1. https:// doi.org/https://doi.org/10.1007/s40891-016-0079-x

Conforti M, Robustelli G, Muto F, Critelli S (2012) Application and validation of bivariate GIS-based landslide susceptibility assessment for the Vitravo river catchmen(Calabria, south Italy). Nat Hazards 61:127–141

Conforti M, Pascale SR, obustelli G, Sdao F (2014). Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy) Catena, http://dx.doi.org/https://doi.org/10.1016/j.catena.2013.08.006

Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, 256 pages

Dai FC, Lee CF (2002) Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology 42:213–228

Dogan U, Yilmaz M (2011) Natural and induced sinkholes of the Obruk Plateau and Karapınar-Hotamı̧s Plain, Turkey. J Asian Earth Sci, 40, 496_508

Fernandez T, Irigaray C, Hamdouni RE, Chacon J (2003) Methodology for landslide susceptibility mapping by means of a GIS. Appl Contraviesa area (Granada, Spain). Natural Hazards 30, 297–308

Figueroa-Miranda S, Vargas JT, Ramos-Leal JA, Hernández-Madrigal VM, Villaseñor-Reyes CI (2018) Land subsidence by groundwater over-exploitation from aquifers intectonic valleys of Central Mexico: a review. Eng Geol 246:91–106

Ghorbanzadeh O, Rostamzadeh H, Blaschke T, Gholaminia K, Aryal J (2018) A new GIS-based data mining technique using an adaptive neuro-fuzzy inference system (ANFIS) and k-fold cross-validation approach for land subsidence susceptibility mapping. Nat Hazards 94(2):497–517

Gonnuru P, Kumar Sh (2017) PsInSAR-based land subsidence estimation of Burgan oil field using TerraSAR-X data. Remote Sens Appl: Soc Environ 9:17–25

Hack J.T (1965) Geomorphology of the Shenandoah Valley, Virginia and West Virginia, and origin of the residual ore deposits. U.S. Geology Survey Professional Paper 484. From http://pubs.usgs.gov/pp/0484/report.pdf Accessed 20 September 2012

Hall FG, Townshend JR, Engman ET (1995) Status of remote sensing algorithms for estimation of land surface state parameters. RemoteSens Environ 51:138–156

Hanley JA (2014) Receiver operating characteristic (ROC) curves. Wiley StatsRef: Statistics Reference Online https://doi.org/10.1002/9781118445112.stat05255.

Joodaki G, Wahr J, Swenson S (2014) Estimating the human contribution to groundwater depletion in the Middle East, from GRACE data, land surface models, and well observations. Water Resour Res 50(3):2679–2692

Karimzadeh S (2015) Characterization of land subsidence in Tabriz (NW Iran) using watershed and InSAR analyses. Acta Geodaetica Geophys, Springer 51:181–195

Khan MS, Khan SD, Kakar DM (2013) Land subsidence and declining water resources in Quetta Valley, Pakistan. Environ Earth Sci. DOI https://doi.org/10.1007/s12665-013-2328-9

Kim KD, Lee S, Oh HJ (2009) Prediction of ground subsidence in Samcheok City, Korea using artificial neural networks and GIS. Environ Geol 58(1):61–70

Komac M (2006) A landslide susceptibility model using the analytical hierarchy process method and multivariate statistics in perialpine Slovenia. Geomorphology 74(1):17–28

Kotsiantis S, Pintalas P (2004) Combining bagging and boosting. J Comput Intell 1(4):324–333

Lee S, Park I, Choi JK (2012) Spatial prediction of ground subsidence susceptibility using an artificial neural network. Environ Manag 49(2):347–358

Martínez j, Marín M , Burbey T, Cervantes N, Lozano J, De-Leon M , Pinto A (2013) Land subsidence and ground failure associated to groundwater exploitation in the Aguascalientes Valley, México. 164 (17): 172-186

Mezughi TH, Akhir JM, Rafek AG, Abdullah I (2011) Landslide susceptibility assessment using frequency ratio model applied to an area along the E-W highway (Gerik-Jeli) Am. J Environ Sci 7(1):43–50

Mohammadi M, Pourghasemi HR, Pradhan B (2012) Landslide susceptibility mapping at Golestan Province, Iran: a comparison between frequency ratio, Dempster-Shafer and weights of evidence models. J Asian Earth Sci 61:221–236

Moore ID, Grayson RB, Ladson A (1991) Digital terrain modeling: a review of hydrological, geomorphological, and biological applications. Hydrol Process 5: 3–30

Motagh M, Walter TR, Sharifi MA et al (2008) Land subsidence in Iran caused by widespread water reservoir overexploitation. Geophys Res Lett. https://doi.org/10.1029/2008GL033814

Mousavi SM, Shamsai A, Naggar MHE, Khamehchian M (2001) A GPS-based monitoring program of land subsidence due to groundwater withdrawal in Iran. Can J Civ Eng 28(3):452–464

Negnevitsky M (2002) Artificial intelligence—a guide to intelligent systems. Addison-Wesley Co., Great Britain

Oh HJ, Syifa M, Wook Lee C, Saro L (2019) Land subsidence susceptibility mapping using bayesian, functional, and meta-ensemble machine learning models. Appl Sci 9:1248

Ozdemir A (2016) Sinkhole susceptibility mapping using logistic regression in Karapınar (Konya, Turkey). Bull Eng Geol Environ 75(2):681–707

Ozdemir A (2015) Investigation of sinkholes spatial distribution using the weights of evidence method and GIS in the vicinity of Karapinar (Konya, Turkey). Geomorphology, 245, 40_50

Park I, Lee J, Lee S (2014) Ensemble of ground subsidence hazard maps using fuzzy logic. Center of Eur J Geosci 6(2):207–218

Pradhan B, Abokharima MH, Jebur NM et al (2014) Land subsidence susceptibility mapping at Kinta valley (Malaysia) using the evidential belief function model in GIS. Nat Hazards 73:1019–1042

Rahmati O, Golkarian A, Biggs T, Keesstra S, Mohammadi F et al (2019) Land subsidence hazard modeling: machine learning to identify predictors and the role of human activities. J Environ Manage 236:466–480. https://doi.org/10.1016/j.jenvman.2019.02.020

Regional Water Company of Hamedan (RWCH) (2020) Basic research reports of the Hamedan province water resources. 204pp (In Persian)

Iranian Department of Water Resources Management (IDWRM) (2015) The Report of Groundwater Drawdown in Plains of Iran.http://www.wrm.ir/index.php?l=EN accessed in May 2015

Iranian Department of Water Resources Management (IDWRM) (2016). Report of Groundwater Resource Monitoring and Land Subsidence Events in Iran. http://www.wrm.ir/index.php?l=EN

Rezaei Y, Dehghani M, Akhavan S, Sahebi MR (2021) Investigation of the effects of water table dropdown on land subsidence in the Kabudar Ahang plain of Hamedan by InSAR techniques. Appl Remote Sens. https://doi.org/10.1117/1.JRS.15.032005

Rouse J W, Haas RW, Schell JA et al (1974) Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation, NASA/GSFC Type III Final Rep., 371 pp., Greenbelt, Md.

Saber M, Abdel-Fattah M, Kantoush S, Sumi T (2018) Implications of land subsidence due to groundwater over-pumping: monitoring methodology using GRACE data. Int J Gen 41:52–59

Santo A, Ascione A, Del Prete S, Di Crescenzo G, Santangelo N (2011) Collapse sinkholes distribution in the carbonate massifs of central and southern Apennines. Acta Carsologica 40:95–112

Sezer EA, Pradhan B, Gokceoglu C (2011) Manifestation of an adaptive neuro-fuzzy model for landslide susceptibility mapping: Klang Valley. Malaysia Exp Syst Appl 38(7):8208–8219

Shi Y, Shi D, Cao X (2018) Impact factors and temporal and spatial differentiation of land subsidence in Shanghai. Sustain 10(9):3146

Shrestha PK, Shakya NM, Pandey VP, Birkinshaw SJ (2017) Model-based estimation of land subsidence in Kathmandu Valley. Nepal Geomatics, Natural Hazards, and Risks 8(2):974–996

Stumpf A, Kerle N (2011) Object-oriented mapping of landslides using random forests. Remote Sens Environ 115:25642577

Sundell J, Haaf E, Norberg T, Alén C, Karlsson M, Rosén L (2017) Risk mapping of groundwater-drawdown-induced land subsidence in heterogeneous soils on large areas. Risk Anal. https://doi.org/10.1111/risa.12890

Svetnik V, Liaw A, Tong C, Culberson J, Sheridan RP, Feuston BP (2003) Random forest: a classification and regression tool for compound classification and QSARmodeling. J Chem Inf Com Sci 43:1947–1958

Teartisup P, Kerdsueb P (2013) Land subsidence prediction in central plain of Thailand. Int J Environm Sci Develop 4(1):59–61

Tucker CJ (1979) Red and photographic infrared linear combinations for monitoring vegetation, Remote Sens. Environ 8:127–150

Vapnik V (2013). Nature of statistical learning theory Springer-Verlag New York, 314 pages

Vijith H, Madhu G (2008) Estimating potential landslide sites of an upland sub-watershed in Western Ghat's of Kerala (India) through frequency ratio and GIS. Environ Geol 55:1397–1405

Vinzi V E, ChinW W, Henseler J, Wang H (2010) Handbook of partial least squares: concepts, methods, and applications. Springer. Open access at http://www.springer.com/series/7286

Voss KA, Famiglietti JS, Lo MH, Linage CD, Rodell M, Swenson SC (2013) Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region. Water Resour Res 49(2):904–914

Wold S, Eriksson L, Kettaneh N (2010) PLS in data mining and data integration. Handbook of partial least squares, Springer 327–357

Xu C, Dai F, Xu X , Lee Y H (2012) GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China; Geomorphology, doi: https://doi.org/10.1016/j.geomorph.2011.12.040

Yesilnacar E, Topal T (2005) Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium-scale study, Hendek region (Turkey). Eng Geol 79(3–4):251–266

Yilmaz I (2009) A case study from Koyulhisar (Sivas-Turkey) for landslide susceptibility mapping by artificial neural networks. Bull EngGeol Environ 68(3):297–306

Yilmaz I, Marschalko M, Bednarik M (2013) An assessment on the use of bivariate, multivariate, and soft computing techniques for collapse susceptibility in GIS environment. J Earth Syst Sci 122:371–388

Zamanirad M, Amirpouya S, Sedghi H, Saremi A, Rezaee P (2019) Modeling the influence of groundwater exploitation on land subsidence susceptibility using machine learning algorithms. Natural Resour Res 1–15

Zhao Y, Li Y, Zhang L, Wang Q (2016) Groundwater level prediction of landslide based on classification and regression tree. Geodesy and Geodynam l7: 348–355

Zhu L, Gong H, Xiaojuan L, Yongyong L, Xiaosi S, Gaoxuan G (2013) Comprehensive analysis and artificial intelligent simulation of land subsidence in Beijing. China Chin Geogra Sci 23(2):237–248