

LAND SUBSIDENCE SUSCEPTIBILITY MAPPING USING MACHINE LEARNING ALGORITHMS

Z. Eghrari¹, M.R. Delavar^{2,*}, M. Zare³, A.Beitollahi⁴, B. Nazari⁵

¹MSc. Student, GIS Dept., School of Surveying and Geospatial Eng., College of Engineering, University of Tehran, Tehran, Iran - zahraeqrari@ut.ac.ir,

² Center of Excellence in Geomatic Eng. in Disaster Management, School of Surveying and Geospatial Eng., College of Engineering, University of Tehran, Tehran, Iran - mdelavar@ut.ac.ir

³International Institute of Earthquake Engineering and Seismology, Tehran, Iran - mzare@iiees.ac.ir

⁴ Road, Housing and Urban Planing Research Center, Ministry of Road and Urban Development - beitollahi@bhrc.ac.ir

⁵GIS Dept., School of Surveying and Geospatial Eng., College of Engineering, University of Tehran, Tehran, Iran - borzoo.nazari@ut.ac.ir

Commission VI, WG VI/3

KEY WORDS: Land Subsidence, GIS, InSAR, Machine Learning, RS, XGBoost Regression, Random Forest, Groundwater

ABSTRACT:

Land subsidence (LS) is one of the most challenging natural disasters that has potential consequences such as damage to infrastructures and buildings, creating sinkholes, and leading to soil destruction. To mitigate the damages caused by LS, it is necessary to determine the LS-prone areas. In this paper, LS susceptibility was assessed for Kashan Plain in Iran using Random Forest (RF) and XGBoost machine learning algorithms. For the susceptibility analysis, twelve influential factors including elevation, slope, aspect, curvature, topographic wetness index (TWI), groundwater drawdown (GWD), normalized difference vegetation index (NDVI), distance to stream (DtS), distance to road (DtR), distance to fault (DtF), lithology, and land use were taken into account. 291 LS points were used in this study which was divided into two parts of 70% and 30% for training and testing the models, respectively. The prediction power of the models and their produced LS susceptibility maps (LSSMs) were validated using the Root Mean Square Error (RMSE), R-Squared (R^2), and Mean Absolute Error (MAE) values. The results showed that the XGBoost had a higher R^2 equal to 0.9032 compared to that of the RF which was equal to 0.8355. XGBoost model had an RMSE equal to 0.3764 cm compared to that of the RF model which was equal to 0.4906 cm. MAE for the XGBoost model was 0.1217 cm and for the RF model was 0.3050 cm. Therefore, the achieved results proved that XGBoost had better performance in this research for predicting LS values based on the measured ones.

1. INTRODUCTION

Land subsidence (LS) is one of the most challenging natural disasters that has potential consequences of damage to infrastructures, creating sinkholes, leading to soil destruction and so on (Raspini et al., 2016; Shi et al., 2020). LS is an apparent and slow deformation or collapse of the earth surface which is caused by a number of natural and human factors (Ng et al., 2015; Zhou et al., 2017). LS can be a natural result of a number of natural and man made disasters such as earthquakes, dissolution of carbonate rocks, movement of faults, or an increase in the depth of groundwater (Arabameri et al., 2021a; Yang et al., 2012). LS has become a global threat, which has occurred in various countries such as China, Mexico, Italy, the United States of America, Spain, and Iran (Brown and Nicholls, 2015; Chaussard et al., 2014; Corbau et al., 2019; Galloway and Burbey, 2011; Tung and Hu, 2012). In recent decades, the rate of LS in Iran has increased widely (Motagh et al., 2008; Tarighat et al., 2021). One of the greatest causes of LS in Iran is the indiscriminate exploitation of groundwater for agricultural purposes (Foroughnia et al., 2019; Mohammady et al., 2019). Water is essential to sustain life on earth. However, its availability is not the same in space and time. Groundwater meets a major part of water demand, and the increasing dependence on this source has led to the depletion of groundwater in different

parts of the world. For decades, groundwater has been widely exploited for domestic, agricultural, and industrial purposes (Foroughnia et al., 2019; Mohammady et al., 2019). This requires artificial recharge to balance groundwater depletion and control LS. Water resources are a critical requirement for a sustainable food supply. The increase in temperature and the change in precipitation patterns in space and time due to climate change lead to frequent and severe droughts and floods, which reduce the ability to absorb and store water (Mirza, 2003). Population growth and the industrialization of societies have increased the demand for water resources and this trend will continue, as increasing amounts of water are needed to sustain societies (Dalín et al., 2017; Wada et al., 2010). Faced with the growing demand for access to fresh water, the use of groundwater is used as a vital resource to meet agricultural, industrial, and drinking water demands. The increase in water demand has led to a decrease in groundwater in many parts of the world including Iran (Foroughnia et al., 2019; Mohammady et al., 2019). In many areas, this issue has led to LS, which causes permanent loss of underground water storage (Smith et al., 2017), damage to infrastructure and arsenic contamination (Erban et al., 2013; Smith et al., 2017). Advances in remote sensing, geospatial information system (GIS) spatial analyses and artificial intelligence (AI), have helped in the modeling of several natural hazards such as LS to determine the LS prone areas. Machine

* Corresponding author

Learning (ML) methods are particularly important in natural hazard modeling due to their capacity to handle complex real world problems, as well as due to their high accuracy and efficiency. (Arabameri et al., 2021a; Chen et al., 2019; Ebrahimi et al., 2020; Feng et al., 2020; Lee et al., 2012; Yang et al., 2012). Previous research has verified that twelve factors that have the most important influence on LS are elevation, slope, aspect, distance to road (DtR), distance to fault (DtF), distance to stream (DtS), groundwater drawdown (GWD), normalized differentiate vegetation index (NDVI), topographic wetness index (TWI), curvature, lithology and land use (Arabameri et al., 2021b; Ebrahimi et al., 2020; Ranjgar et al., 2021). ML is a branch of AI (Zhou, 2021). ML has been used in various fields such as landslide (Arabameri et al., 2020), effects of climate change on the environment (Seyed Mousavi et al., 2022) and so on. In the previous study of LS, ML has been widely used to determine the relationship between the influencing factors and subsidence of the area (Ranjgar et al., 2021; Shi et al., 2020) and to estimate and predict LS (Mohammady et al., 2019; Ranjgar et al., 2021; Shi et al., 2020). In previous studies, the XGBoost model was once used to model the LS of the Beijing Plain, China (Shi et al., 2020), however, the influencing factors used in this study except groundwater level had not been taken into consideration. In this research, we have considered these influential factors as well in the modeling phase. Ensemble learning algorithms such as XGBoost and Random Forest (RF) improve the performance of the model by reducing the overall error rate (Zhou, 2012). Compared with traditional ML models such as SVM and ANN, XGBoost and RF have faster calculation speed and compared with deep learning algorithms, these models are good for tabular

data with fewer features (Shi et al., 2020). In this study, We have considered the twelve influencing factors in LS, and also a number of the sample points of subsidence were collected using radar interferometry. The data was divided into two parts, including training and test to train XGBoost and RF models and evaluate their results. The final evaluation of the model was undertaken using Root Mean Square Error (RMSE), R-Squared (R^2), and Mean Absolute Error (MAE) values. Finally, the LS susceptibility map of the study area was produced from the two models and a comparison was made between them.

The remaining parts of the paper is as follows. Section 2 has concentrated on the description of study area. Section 3, presents the research methodology. Section 4 elaborates the research results. Finally, section 5 concludes the paper and suggests some directions for future research.

2. DESCRIPTION OF THE STUDY AREA

Kashan plain is a part of Kashan city, which ends at Karkas mountain from the south and is located about 240 kilometers south of Tehran between the Longitudes of 51.05 and 51.54 degrees and the Latitudes of 33.45 and 34.23 degrees (Figure 1). Kashan plain with an area of 1570 Square Kilometers includes the city of Kashan, its central part, the cities of Aran and Bidgol and agricultural lands located in the plain. Kashan Plain is one of the least rainfall regions of Iran. The climate of the study area has two classes of arid and semi-arid. The temperature in this area ranges from 16°C to 22°C. Moreover, Elevation is between 799 m and 1336 m above mean sea level.

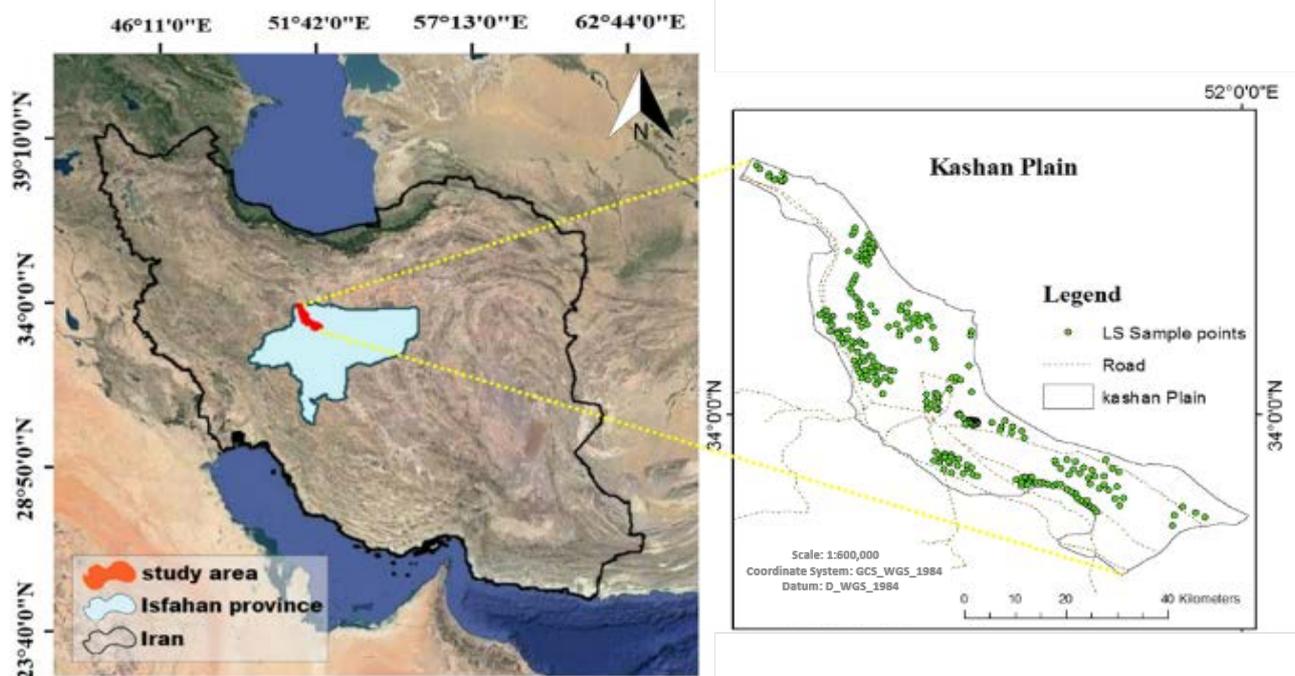


Figure 1. Study area and location of the employed LS sample points.

3. METHODOLOGY

The research methodology consists of four steps as follow which are illustrated in Figure (2):

- Selecting 291 LS points.

- Data preprocessing and production of maps of the twelve influencing factors in LS.
- Employing RF and XGBoost ML algorithms to map the LS susceptibility.
- Validation of the performance of each model using RMSE, R^2 and MAE.

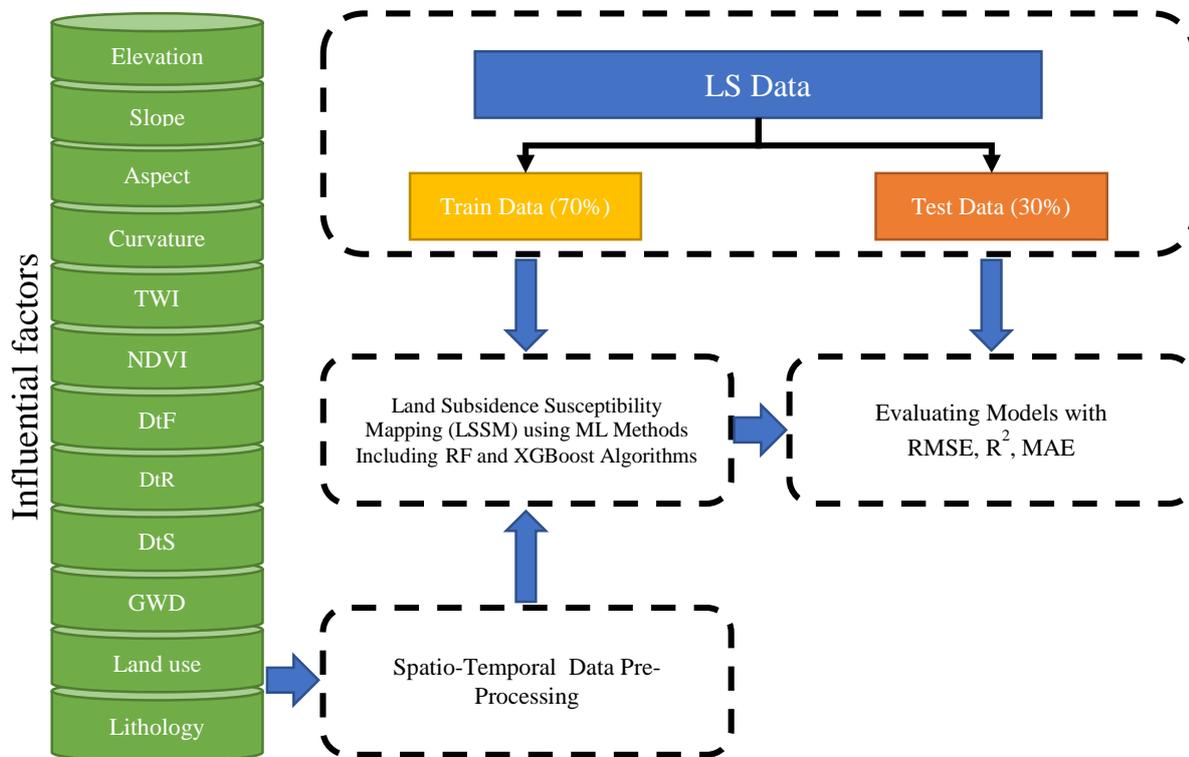


Figure 2. Flowchart of research methodology.

3.1 Selecting LS points

the spatial distribution of several LS regions was shown in Figure (1). The interferometric synthetic aperture radar (InSAR) map produced by the Geological Survey and Mineral Exploration of Iran (GSI) with centimeter accuracy in the first half of 2016 was used to prepare sample points of subsidence in the study area.

3.2 LS influential Factors

The twelve LS influential factors have been selected to produce the LSSM for this area (Table 1) which are illustrated in Figure (3). Various data sources were used to produce these influencing factors. For the digital elevation model (DEM), ALOS PALSAR RTC HR which has a 12.5 Meter spatial resolution was used to produce slope, aspect, elevation, curvature and TWI data layers.

To produce the land use and land cover map, Sentinel 2A images were used on Google Earth Engine platform and then classified in four classes including urban area, agricultural, barren land and poor range land. The land use map of the region was produced which is illustrated in Figure (3k). Sentinel 2A images were also used to produce NDVI map of the region (Figure 3d). The NDVI is calculated using Eq. (1) (Rouse et al., 1974):

$$NDVI = \frac{NIR-Red}{NIR+Red} \quad (1)$$

where Red = stands for the spectral reflectance measurements acquired in the red (visible) regions
 NIR = stands for the spectral reflectance measurements acquired in the near-infrared regions

Factors	Source	Scale	Resolution (m)
Slope	ALOS PALSAR RTC HR Dem		12.5 * 12.5
Aspect	ALOS PALSAR RTC HR Dem		12.5 * 12.5
Elevation	ALOS PALSAR RTC HR Dem		12.5 * 12.5
Curvature	ALOS PALSAR RTC HR Dem		12.5 * 12.5
TWI	ALOS PALSAR RTC HR Dem		12.5 * 12.5
Land Use	Sentinel - 2		10 * 10
NDVI	Sentinel - 2		10 * 10
DtR	GSI	1:100,000	
DtS	GSI	1:100,000	
DtF	GSI	1:100,000	
Lithology	GSI	1:100,000	
GWD	Regional Water Company of Kashan		12.5 * 12.5

Table 1. The details of the input influential factors.

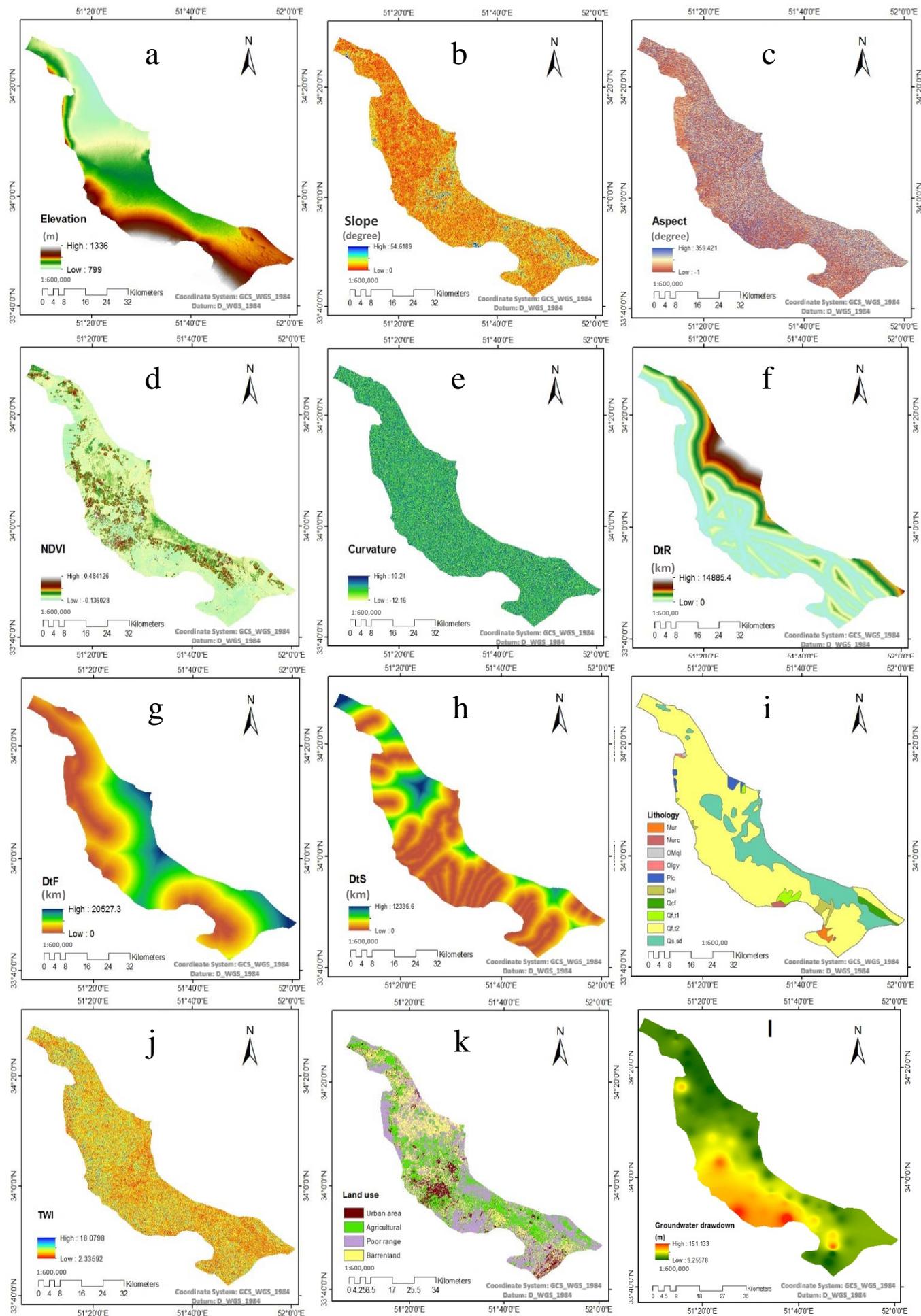


Figure 3. LS conditioning factors: (a) Elevation, (b) Slope, (c) Aspect, (d) NDVI, (e) Curvature, (f) DtR, (g) DtF, (h) DtS, (i) Lithology, (j) TWI, (k) Land use, and (l) GWD.

Maps of roads, faults, and stream at a scale of 1:100,000 were produced by GSI, and the maps of DtR, DtF, and DtS were produced (Figures 3f-h) in this research. the TWI is a secondary topographic variable (Figure 3j). The TWI is calculated using Eq. (2) (Beven and Kirkby, 1979):

$$TWI = \ln \frac{a}{\tan b} \quad (2)$$

where a = the local upslope area draining through a certain point per unit contour length
 b = the local slope in radians

The TWI map was produced using the DEM on the ArcGIS 10.3

platform whose value ranges from 2.33 to 18.07. The Lithology map of the study area was produced by GSI (Figure 3i). There are ten different classes of Geo Unit in this area (Table 2). The curvature map was shown in Figure (3e) whose values range from -12.16 to 10.24. Finally, the map layers had created in grids of 20 Meter * 20 Meter size in order to harmonize the data.

3.3 Employing RF and XGBoost

3.3.1 RF regression

RF is developed by Breiman (Breiman, 2001; Wang et al., 2020). RF Regression is a supervised ML decision tree-based algorithm, where the decision trees form with random samples from the train

Geo unit	Description	Age	Area
Mur	Pale-yellow to red sandstone and gypsiferous marl, siltstone and shale (Upper Red Fm.)	Miocene	1301
Murc	Conglomerate and sandstone	Miocene	810
Olgy	Gypsum	Oligocene	232
OMql	Massive to thick-bedded reefal limestone	Oligocene-Miocene	127
Plc	Poorly sorted, moderately consolidated, conglomerate, sandstone and mudstone	Pliocene	1886
Qal	Stream channel, braided channel and flood plain deposits	Quaternary	2330
Qcf	Clay flat	Quaternary	2371
Qf,t1	High level piedmont fan and valley terrace deposits	Quaternary	2599
Qf,t2	Low level piedment fan and valley terrace deposits	Quaternary	123814
Qs,sd	Unconsolidated wind blown sand deposit including sand dunes	Quaternary	36624

Table 2. Lithology of the study area.

-ing data (Breiman, 2001). Grid Search was used to adjust the parameters. n_estimator is the number of trees in the RF, and max_features is the number of features to consider when looking for the best split (Huang et al., 2016).

3.3.2 XGBoost regression

XGBoost is one of the quickest implementations of gradient boosted trees (Lu and Ma, 2020). XGBoost is an iterative decision tree algorithm, which uses residuals to improve the model. First, XGBoost supports parallel computing, second, it also supports regularization, which prevents model overfitting (Huang et al., 2022). Although the model is highly accurate, it

can be easily overfit. For this purpose, n_estimators must be Controlled. The eta is used to control the rate of iterations and prevent overfitting. Subsample controls the proportion of the extraction of example. xgb_model parameter is used for selecting a weak evaluator. The objective, max_depth, the alpha and lambda are used to select the loss function, specifies the maximum depth of each tree, control L1 and L2 regularization terms, respectively. We used Grid Search method to adjust the parameters (Table 3). Finally, to find the most important influential factors of the LS susceptibility, mean decrease in impurity (MDI) was calculated for RF (Figure 4) and XGboost models (Figure 5). The results showed that DtF, elevation and GWD hold the greatest impact on the LS occurrence.

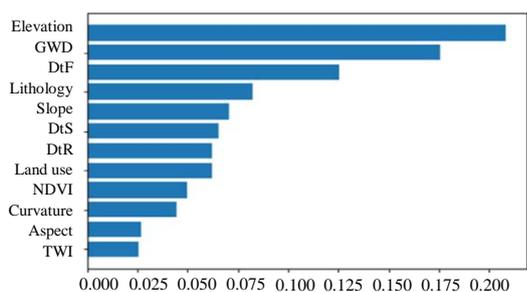


Figure 5. XGboost Feature importance based on MDI.

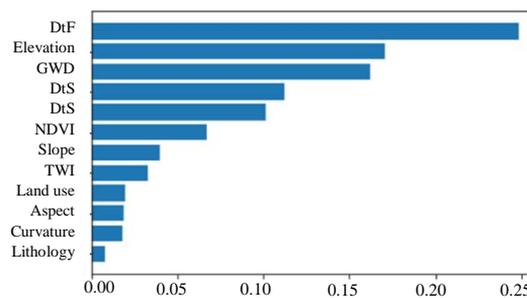


Figure 4. RF Feature importance based on MDI.

n_estimators	Eta	Subsample	xgb_Model	Objective	Alpha	Lambda	max_depth
1000	0.1	0.6	gbtree	reg:squarederror	0.1	1	10

Table 3. The XGBoost parameters.

3.4 Validation

To assess the efficiency of the models, R^2 , RMSE, and MAE were employed (Equations (3-5)). R -squared will give an estimate of the relationship between movements of a dependent variable based on an independent variable's movements. It represents the possible bias in the data and predictions. It does not mean whether the selected model is good or bad. The closer the R -squared to one, the better (Cameron and Windmeijer, 1997).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (3)$$

Root Mean Square Error (RMSE) is the standard deviation of the residual (prediction errors) which is one of the most commonly used measures for evaluating the quality of predictions. It tells

how concentrated the data is around the line of best fit. Naturally lower values indicate a better fit for the model (Barnston, 1992).

$$RMSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4)$$

MAE is the mean of the absolute errors which is the absolute value of the difference between the predicted and the measured values (Eq. 5) (Willmott and Matsuura, 2005).

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (5)$$

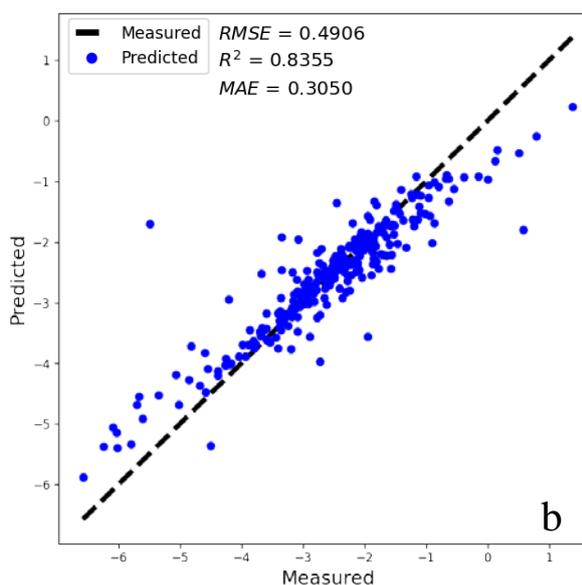
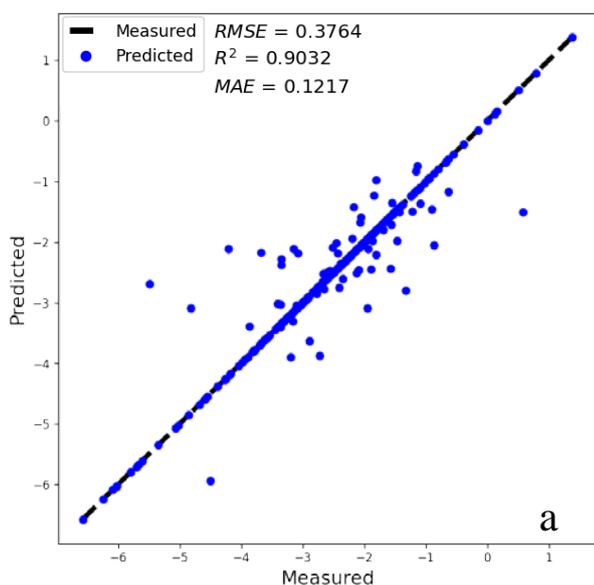


Figure 6. Compatibility between the measured data and the predicted data (a) XGBoost, (b) RF.

where \hat{y}_i = vector of predicted dependent variables with n data points
 y_i = vector of observed values of the variable being predicted
 \bar{y}_i = mean of the observed dependent variables

4. RESULTS AND DISCUSSION

By utilizing the twelve influential factors maps produced, the selected LS points and the methods mentioned in Section 3.3, the mapping and assessment of LS susceptibility for Kashan plain were undertaken. RMSE, R^2 , and MAE were calculated for RF and XGboost models. Table (4) demonstrates the comparison of RMSE, R^2 , and MAE for each models used. The results showed that the XGBoost had higher R^2 (0.9032) compared with that of the RF (0.8355). XGBoost model had less RMSE (0.3764 cm) than that of the RF model (0.4906 cm). MAE for the XGBoost model was equal to 0.1217 cm and for the RF model was equal

Model	Parameter	Validation
Random forest	RMSE (cm)	0.4906
	R^2	0.8355
	MAE (cm)	0.3050
XGBoost	RMSE (cm)	0.3764
	R^2	0.9032
	MAE (cm)	0.1217

Table 4. Validation.

to 0.3050 cm. Figure (6) demonstrates the compatibility between the measured data and the predicted data. As a result, the XGBoost model has a higher prediction accuracy than that of the RF model.

4.1 Results of the LS susceptibility Maps (LSSMs)

After applying the models and evaluating the accuracies, the maps of the twelve influencing factors, in the form of a stack on QGIS 3.16 platform was produced to be used as the model input. Then the LSSMs were produced. The values of the LSSMs prepared from the models were in the range of 1.441 cm to -7.497

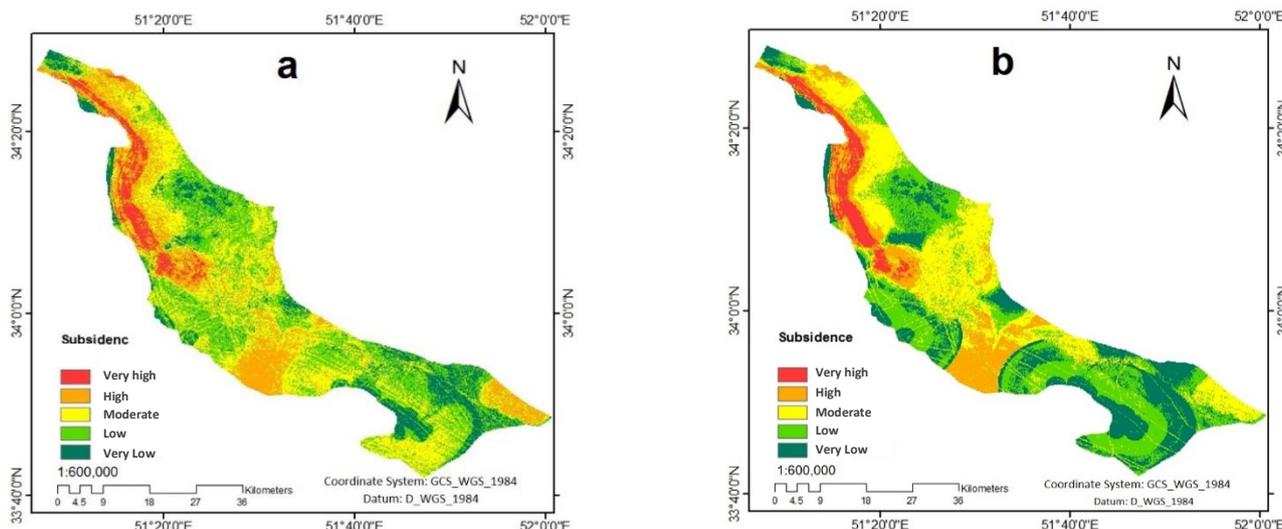


Figure 7. LSSM production using: (a). XGBoost, (b). RF.

cm. These LSSMs were then classified into five classes, using a natural break algorithm (Jenks and Caspall, 1971) illustrated in Figure (7).

5. CONCLUSION

Iran National Cartographic Center (NCC) reports show that the annual average subsidence rate has increased in Iran (<https://www.ncc.gov.ir/en/>). therefore, LS is an important issue in Iran. LS was affected by a number of factors. We selected the most important influencing factors to predict the LS rate in the Kashan Plain, and investigate the relationship of parameters in LS modeling. The conclusions are as follows:

- Phenomenon of LS is one of the most threatening natural hazards of the earth, which has great losses on the economy. For proper assessments of this issue, it is necessary to develop a suitable model of LS that can be used in any region. In previous studies, the XGBoost model was once used to model the LS of the Beijing Plain, China (Shi et al., 2020), however, the influencing factors used in this study except groundwater level had not been taken into consideration. In this research, we have considered these influential factors as well in modeling. The achieved results prove that the model is well established.
- The results have indicated that the XGBoost model had less RMSE (0.3764 cm) than that of the RF model (0.4906 cm). MAE for the XGBoost model was equal to 0.1217 cm and for the RF model was equal to 0.3050 cm. XGBoost had higher R^2 which was equal to 0.9032 compared to that of the RF which was equal to 0.8355 indicating better compatibility between the predicted and measured LS.
- As can be seen in both of the models used in this study, the highest rate of LS is in the northwest and west of Kashan Plain and the lowest rate of LS is observed in the south of Kashan Plain. In addition, in places where the DtR, the DtF and DtS are less, a higher LS rate has been observed. According to the GWD map, in the southwestern and the northwestern parts of the study area, the maximum GWD can be observed. Furthermore, in the southwest of Kashan plain, a high subsidence rate has been occurred.
- The major strength of this study was quality of the ensemble ML algorithms and their optimum prediction results in the LS mapping. There were some limitations in this research such as lack of implementing hydrological modelings which will be considered in our future research.

ACKNOWLEDGEMENTS

We are grateful to GSI for providing the subsidence data, lithology, fault, road and stream maps and Kashan Regional Water Company to provide water data to this research.

REFERENCES

- Arabameri, A., Lee, S., Rezaie, F., Chandra Pal, S., Asadi Nalivan, O., Saha, A., Chowdhuri, I., Moayedi, H., 2021a. Performance evaluation of GIS-based novel ensemble approaches for land subsidence susceptibility mapping. *Frontiers in Earth Science* 9, 307.
- Arabameri, A., Saha, S., Roy, J., Chen, W., Blaschke, T., Tien Bui, D., 2020. Landslide susceptibility evaluation and management using different machine learning methods in the Gallicash River Watershed, Iran. *Remote Sensing* 12, 475.
- Arabameri, A., Yariyan, P., Santosh, M., 2021b. Land subsidence spatial modeling and assessment of the contribution of geo-environmental factors to land subsidence: comparison of different novel ensemble modeling approaches.
- BEVEN, K.J., KIRKBY, M.J., 1979. A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Bulletin* 24, 43–69.
- Breiman, L., 2001. Random Forests. *Machine Learning* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Brown, S., Nicholls, R.J., 2015. Subsidence and human influences in mega deltas: the case of the Ganges–Brahmaputra–Meghna. *Science of the Total Environment* 527, 362–374.
- Cameron, A.C., Windmeijer, F.A., 1997. An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of econometrics* 77, 329–342.
- Chaussard, E., Wdowinski, S., Cabral-Cano, E., Amelung, F., 2014. Land subsidence in central Mexico detected by ALOS InSAR time-series. *Remote sensing of environment* 140, 94–106.
- Chen, W., Fan, L., Li, C., Pham, B.T., 2019. Spatial prediction of landslides using hybrid integration of artificial intelligence algorithms with frequency ratio and index of entropy in Nanzheng county, china. *Applied Sciences* 10, 29.

- Corbau, C., Simeoni, U., Zoccarato, C., Mantovani, G., Teatini, P., 2019. Coupling land use evolution and subsidence in the Po Delta, Italy: Revising the past occurrence and prospecting the future management challenges. *Science of the Total Environment* 654, 1196–1208.
- Dalin, C., Wada, Y., Kastner, T., Puma, M.J., 2017. Groundwater depletion embedded in international food trade. *Nature* 543, 700–704.
- Ebrahimi, H., Azadbakht, M., 2019. Downscaling MODIS land surface temperature over a heterogeneous area: An investigation of machine learning techniques, feature selection, and impacts of mixed pixels. *Computers and Geosciences* 124, 93–102. <https://doi.org/10.1016/j.cageo.2019.01.004>.
- Ebrahimi, H., Feizizadeh, B., Salmani, S., Azadi, H., 2020. A comparative study of land subsidence susceptibility mapping of Tasuj plane, Iran, using boosted regression tree, random forest and classification and regression tree methods. *Environmental Earth Sciences* 79, 1–12.
- Erban, L.E., Gorelick, S.M., Zebker, H.A., Fendorf, S., 2013. Release of arsenic to deep groundwater in the Mekong Delta, Vietnam, linked to pumping-induced land subsidence. *Proceedings of the National Academy of Sciences* 110, 13751–13756.
- Feng, W., Lu, H., Yao, T., Yu, Q., 2020. Drought characteristics and its elevation dependence in the Qinghai–Tibet plateau during the last half-century. *Scientific Reports* 10, 1–11.
- Foroughnia, F., Nemati, S., Maghsoudi, Y., Perissin, D., 2019. An iterative PS-InSAR method for the analysis of large spatio-temporal baseline data stacks for land subsidence estimation. *International Journal Of Applied Earth Observation and Geoinformation* 74, 248–258.
- Galloway, D.L., Burbey, T.J., 2011. Regional land subsidence accompanying groundwater extraction. *Hydrogeology Journal* 19, 1459–1486.
- Huang, J., Khan, S.D., Ghulam, A., Crupa, W., Abir, I.A., Khan, A.S., Kakar, D.M., Kasi, A., Kakar, N., 2016. Study of Subsidence and Earthquake Swarms in the Western Pakistan. *Remote Sensing* 8, 956. <https://doi.org/10.3390/rs8110956>.
- Huang, Y., Chen, C., Miao, Y., 2022. Prediction Model of Bone Marrow Infiltration in Patients with Malignant Lymphoma Based on Logistic Regression and XGBoost Algorithm. *Computational and Mathematical Methods in Medicine* 2022.
- Lee, S., Park, I., Choi, J.-K., 2012. Spatial prediction of ground subsidence susceptibility using an artificial neural network. *Environmental management* 49, 347–358.
- Lu, H., Ma, X., 2020. Hybrid decision tree-based machine learning models for short-term water quality prediction. *Chemosphere* 249, 126169.
- Mirza, M.M.Q., 2003. Climate change and extreme weather events: can developing countries adapt? *Climate policy* 3, 233–248.
- Mohammady, M., Pourghasemi, H.R., Amiri, M., 2019. Land subsidence susceptibility assessment using random forest machine learning algorithm. *Environmental Earth Sciences* 78, 1–12.
- Motagh, M., Walter, T.R., Sharifi, M.A., Fielding, E., Schenk, A., Anderssohn, J., Zschau, J., 2008. Land subsidence in Iran caused by widespread water reservoir overexploitation. *Geophysical Research Letters* 35.
- Ng, A.H.-M., Ge, L., Li, X., 2015. Assessments of land subsidence in the Gippsland Basin of Australia using ALOS PALSAR data. *Remote Sensing of Environment* 159, 86–101. <https://doi.org/10.1016/j.rse.2014.12.003>.
- Ranjgar, B., Razavi-Termeh, S.V., Foroughnia, F., Sadeghi-Niaraki, A., Perissin, D., 2021. Land Subsidence Susceptibility Mapping Using Persistent Scatterer SAR Interferometry Technique and Optimized Hybrid Machine Learning Algorithms. *Remote Sensing* 13, 1326. <https://doi.org/10.3390/rs13071326>.
- Raspini, F., Bianchini, S., Moretti, S., Loupasakis, C., Rozos, D., Duro, J., Garcia, M., 2016. Advanced interpretation of interferometric SAR data to detect, monitor and model ground subsidence: outcomes from the ESA-GMES Terrafirma project. *Natural Hazards* 83, 155–181.
- Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS.
- Seyed Mousavi, S.M., Akhoondzadeh Hanzaei, M., 2022. Monitoring and Prediction of the changes in water zone of wetlands using an intelligent neural-fuzzy system based on data from Google Earth Engine system (Case study of Anzali Wetland, 2000-2019). *Eng. J. Geospatial Inf. Technol.* 9, 19–42.
- Shi, L., Gong, H., Chen, B., Zhou, C., 2020. Land Subsidence Prediction Induced by Multiple Factors Using Machine Learning Method. *Remote Sensing* 12, 4044.
- Smith, R.G., Knight, R., Chen, J., Reeves, J.A., Zebker, H.A., Farr, T., Liu, Z., 2017. Estimating the permanent loss of groundwater storage in the southern San Joaquin Valley, California. *Water Resources Research* 53, 2133–2148.
- Tarighat, F., Foroughnia, F., Perissin, D., 2021. Monitoring of power towers' movement using persistent scatterer SAR interferometry in south west of Tehran. *Remote Sensing* 13, 407.
- Tung, H., Hu, J.-C., 2012. Assessments of serious anthropogenic land subsidence in Yunlin County of central Taiwan from 1996 to 1999 by Persistent Scatterers InSAR. *Tectonophysics* 578, 126–135.
- Wada, Y., Van Beek, L.P., Van Kempen, C.M., Reckman, J.W., Vasak, S., Bierkens, M.F., 2010. Global depletion of groundwater resources. *Geophysical research letters* 37.
- Wang, S., Liu, S., Zhang, J., Che, X., Yuan, Y., Wang, Z., Kong, D., 2020. A new method of diesel fuel brands identification: SMOTE oversampling combined with XGBoost ensemble learning. *Fuel* 282, 118848.
- Willmott, C.J., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research* 30, 79–82.
- Yang, Y., Jia, S., Luo, Y., Jiang, Y., 2012. Analysis and outlook of the land subsidence in Beijing. *Geol. Resour* 7, 20–26.
- Zhou, L., Guo, J., Hu, J., Li, J., Xu, Y., Pan, Y., Shi, M., 2017. Wuhan Surface Subsidence Analysis in 2015–2016 Based on Sentinel-1A Data by SBAS-InSAR. *Remote Sensing* 9, 982. <https://doi.org/10.3390/rs9100982>.
- Zhou, Z.-H., 2021. Machine learning. *Springer Nature*.
- Zhou, Z.-H., 2012. Ensemble methods: foundations and algorithms. *CRC press*.