Application Of Logistic Regression Method To Produce Landslide Susceptibility Map: A Case Study Of Tetouan Mazari, Morocco

L. Ait Brahim [1] And M. Elmoulat [1].

[1] University Mohamed V, Faculty of Science, Research Unit GeoRisk: Geological Risks, Remote Sensing and sustainable development, LGRN Ibn Battouta Avenue. Zip code 1014, Rabat – Agdal, Morocco.

laitbrahim@gmail.com

Abstract. The main purpose of this study is to use logistic regression (RL) model to map landslide susceptibility in and around the area of Tetouan Mazari in the Northern Morocco. Parameters, such as lithology, slope gradient, slope aspect, faults, drainage lines, and hillshade, were considered. Landslide susceptibility map was produced using RL method and then compared and validated. Before the modeling and validation, the observed landslides were separated into two groups. The first group was for training, and the other group was for validation steps. The accuracy of the model was measured by fitting them to a validation set of observed landslides. For validation process, the half landslides remaining was used. The final map was classified into five classes: Very High (32%), High (40%), Medium (7%), Low (7%) and Nil (15%). According to these values logistic regression was determined to be one of the most accurate method to generate landslide susceptibility map. Last but not least, logistic regression model can be used to manage and mitigate hazards related to landslides and to aid in land-use planning for the city of Tetouan..

1 Introduction

Landslides are significant natural hazards in many areas of the world. Every year they cause more than 100,000 deaths and injuries and more than one billion USD in damage [1] . Mapping the areas that are susceptible to be prone to landslides is mandatory for effective land-use management. Thus, there is an increasing need for methodologies to guide managers in choosing the best management strategies, including minimizing impacts from land-use activities in vulnerable slope areas [2]. A landslide susceptibility map indicates areas likely to undergo landslides in the future by correlating some of the factors that contribute to landslides with historic distributions of slope failures [3]. Many methods and techniques, often involving geographic information systems (GIS), have been proposed to produce landslide susceptibility maps [4]). Logistic regression has been applied for susceptibility mapping by various researchers including and [5].

The main objective of the present study is to create susceptibility map by using Logistic Regression (LR). For this purpose, initially, a detailed landslide inventory was prepared [6]. Then, the landslide locations mapped were separated into two groups for training and validation purposes. In addition to that six different controlling factors have been mapped to generate the susceptibility map.

2 Study area

Our site test is Mazari located in the province of Tetouan (Latitude: 35°45'N - 35°30'S; Longitude: 5°30'W - 5°9'E). It covers a total area of about 590 Km2. It is limited by Fnideq, the Mediterranean Sea, Chefchaouen, and Asilah on the North, East, South and West, respectively "Fig. 1". The study area is inserted within a hazardous and mountainous zone called the Western Rif (Northern Morocco). The Rif has been subject to a complex structural and Palaeogeographical evolution since the Triassic period until the Quaternary [7]. This tectonic zone is made up by three different paleogeographic domains.

2.1 The geological and structural sitting

The geological and structural sitting of the study area have been detailed in the geological map of Tetouan

sheet by (Kornprobst & Durand Delga 1985), which consist broadly of three main domains: (1) Internal domain: represented by the Ghomarides, which are metamorphic nappes of Paleozoic from Silurian to Carboniferous and Sebtides that are composed of polymetamorphic terrain and Paleozoic-Triassic cover called the unit of Federico. The Sebtides-Ghomarides are essentially thrust over calcareous units, namely, the Dorsale Calcaire that composed the third part of this domain. (2) Flyschs: formed by Cretaceous and Tertiary material. We distinguish three types of nappes: Numidian Flysch, Flysch of Tisirène and Flysch of Melloussa. (3) External zone: Represented by Tangier Unit.
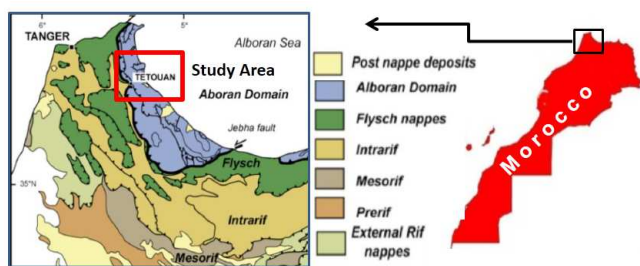

Fig 1. Location map showing the study area in the Province of Tetouan.

2.2 Seismic activity

The Tetouan region, although located in a moderate seismic zone, remains under the effect of the very active neighboring seismic zones (Alboran and Acores). However, no recent events of sufficient intensity (magnitude greater than or equal to 6, intensity VII or more MSK) occurred to slope instability.

2.3 Climate, Precipitation and Temperature

Climate: Our study area knows the succession of two different seasons in one year; to a warm and dry summer, which lasts for several months, succeeds a rainy and colder season corresponding to the end of autumn, winter and early spring. The rains are not continued but alternate during the whole year with periods of fine weather of long and short duration. The study area is part of an area which forms a transition between a Mediterranean and an Atlantic slope; in fact, the climate of the region benefits from two maritime influences, on the one hand a Mediterranean influence because the opening of the limestone chain at Tetouan allows to penetrate Mediterranean perturbations, on the other hand an Atlantic influence because it is exposed to the Atlantic winds of the West and South West. The rains are prolonged and very violent with a maximum in November, December and January, causing intense erosion on the slopes to flash flooding. For precipitation, 80% is recorded between November and April.

Precipitation: has a significant inter-annual variation, suggesting that the precipitation system is not stable from year to year, but also within a given year, which is an important hydrological system and erosion of the watershed.

Temperature: The temperatures are high, at the end of Spring until the end of the Autumn, especially in Summer when one attends a very strong Summer drought which causes desiccations into the soils. According to the doctoral work of [7] , the hottest months of the year are July and August with a high average (38°C for the two months), the coldest months are December and January (2.6°C and 2.2°C respectively). A thin blanket of snow falls in this area (4 to 5 days/year) and lasts few hours. The study area is part of sub-humid zone. There is a good relationship between mass movement and the height of precipitation level (Winter and Spring), itself dependent on altitude and the aspect; but the main feature of mass movements distribution remains the lithological factor, because the majority of these phenomena start in marly and schistose formations.

3 Methodology

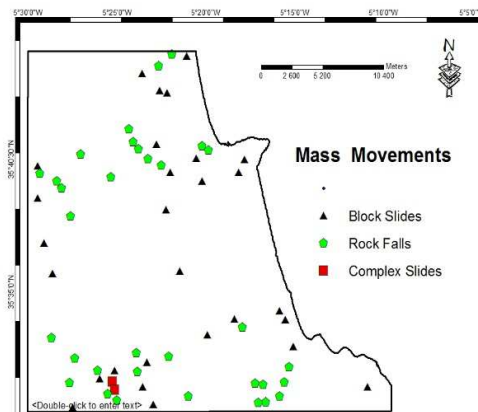### 3.1 Generation of database for LR model

Fig 2. Mass Movements of Tetouan Mazari Area

A total of 126 landslides were mapped. According to the classification made by Varnes [8], the type of landslides movement encounter in our case study were rock falls, and translational and rotational slides. A landslide inventory map was prepared by mapping the landslide locations in Mazari [9] using interpretation of aerial photographs, high resolution satellite imagery, previous reports and extensive field surveys [10]. A random partitioning of the training data was carried out in the area of Mazari, 50% of the whole landslides were selected randomly to calibrate the logistic regression analysis and the half remaining was used for validation purposes "Fig. 2". The pixel size of the landslide map and all the independent variable maps was 100m. Six parameters were created for the landslide susceptibility map of the study area. These parameters are lithology, fault density, hydrographic network density, slope gradient, slope aspect, hypsometry "Fig. 3". The previous description of the geological sitting shows that the study area consists predominantly of marsh sandstone Flyschs, detrital limestones, mudstones and conglomerates. All these formations are almost alternative, which have an effect on slope stability.
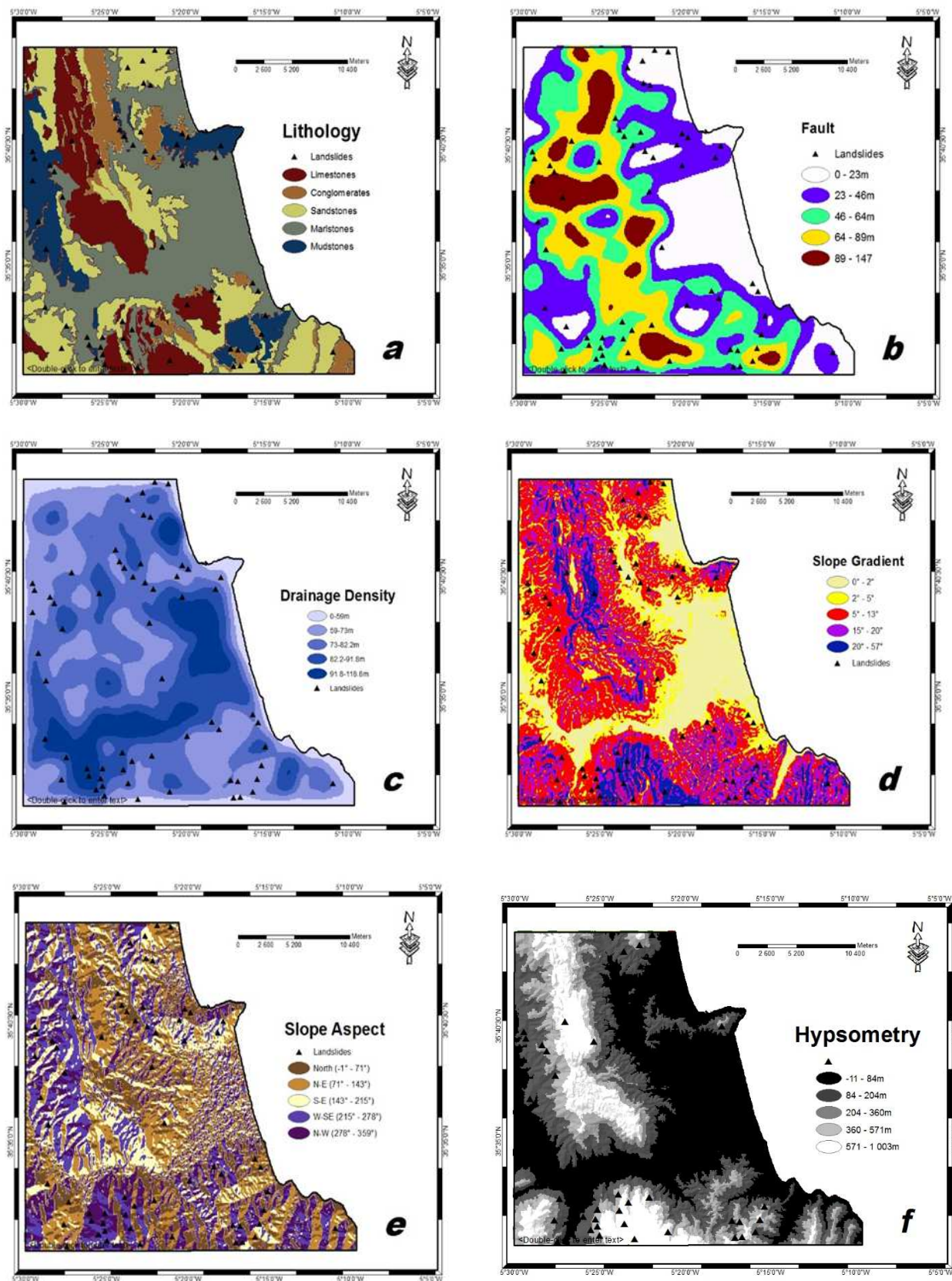
Fig 3. Various thematic data layers prepared in GIS for Tetouan-Ras-Mazari area

For example, mudstones become fluid in contact with water; hence, they are easily discharged when they are on surface. The alternation of limestones and sandstones with marlstones and clays is another unfavorable factor in terms of terrain stability [7].The lithological map used in this study has been derived form the geological map of the study area [11] and topographical map [12], combined to frequencies of landslide occurrence, and then were classified into five classes to create the lithology map "Fig. 3(a)".

According to previous work that has been done in the past by many researchers and experts in Western Rif [13], our case study is a faulted zone which presents favorable conditions for landslide occurrence. All faults are highlighted by the consistent movement caused within the features of different kinds of tectonic layering; they are posterior to thrust that accommodates horizontal shortening e geological and overthrusting. Hence, fault map of 341 types of tectonic movements were mapped and analyzed using geological map and Landsat 7 ETM+ 2015 satellite imagery observation and validated by ground truth studies [11] "Fig. 3(b)".

Water is considered as the primary factor to trigger landslide mechanisms. Rivers may induce failure of the banks due to slope undercutting. Many studies have shown that the proximity to the drainage lines is an important factor controlling the occurrence of landslides. This can be attributed to the fact that terrain modification caused by gully erosion may influence the initiation of a landslide. In order to map this parameter, we have used a DEM to extract the automatically The hydrographic network using ArcHydro tools. This extension was developed by ESRI along with a hands-on document that explained step-by-step how to perform drainage analysis on a terrain model. DEM Raster analysis is performed to generate data on flow direction, flow accumulation and watershed delineation. These data are then used to develop a vector representation of catchments and drainage lines. Using this information, a geometric network is constructed [14]. The final map of distance to drainage was validated using the topographic map of Tetouan scale 1:50,000 [12]"Fig. 3(c)".

Moreover, A digital elevation model (DEM) A DEM with 30×30m resolution representing our study area was a great source to generate various topographic factors, which influence the landslide activities in this zone. hence, the dem was modified to correct the artifacts associated with the aster sensor and to improve its visual quality. three thematic data layers were extracted which are slope (0° - 54°). aspect (0° - 360°), and hypsometry:

• Slope: The slope gradient is an important component of the slope stability analysis and is frequently used in landslide susceptibility studies. The more the slope gradient increases, the more it will correlate with increased likelihood of failure. However, soil thickness and strength are two factors that vary over a wide range between sites [15]. The slope gradient map of the study area was divided into 5 slope categories "Fig. 3(d)".

• Aspect: can be defined as the compass direction that a slope faces measured in degrees from north in a clockwise direction, ranking from 0° to 360°. The values of each cell in the output raster indicate the compass direction that the surface faces at that location. We especially used the aspect in relation to the effect of the solar radiation and the degree of humidity: The orientation of the slopes inthe region of Tetouan influences especially the degree of instability on the exposed slopes (Wet) North, NNW and NNE for the mudflows and landslides. The slopes (Dry) of the direction South, SSE, SSW show especially an erosion dominated by the streaming "Fig.3(e)".

• Hillshade: creates a shaded relief from a surface raster by considering the illumination source angle and shadows. It is obtained by determining illumination values for each cell in relation to neighboring cells "Fig.3(f)".

3.1.1 Overview of LR model

The binary logistic regression used in this study is one of the most popular statistical methods for determining landslide susceptibility and has often been applied for this purpose [16]. The application of logistic regression requires the inclusion of landslide triggering and/or conditioning parameters as independent variables. In general, the more independent variables that are included, the more complete the model will be, given that the consideration of variables plays a major role in determining the dependent variable. As it has been mentioned by Pham [17] in his paper that LR is another known popular method for landslide susceptibility assessment. Many authors [18], define LR is a mathematical model that used to predict the presence or absence of an event based on values of a set of controlling factors. It can be beneficial for either continuous or discrete, or any combination of both types, and they do not necessarily have normal distributions. The LR has been proven more efficient than other methods such as certainty factor, likelihood ration, artificial neural networks, and multi-criteria decision analysis for landslides susceptibility assessment [19].

$$p = 1/(1 + \square^{\wedge}(-z)) \qquad (1)$$

Where p is the probability of landslide occurrence, z is a parameter.

$$z = \alpha + \beta_1 x_2 + \beta_2 x_2 + \ldots + \beta_n x_n \qquad (2)$$

Where α is the intercept of the equation; x1, x2,….,xn are the predictor factors; β1, β2,…., βn are the coefficients needed to be estimated for the predictor factors [20].
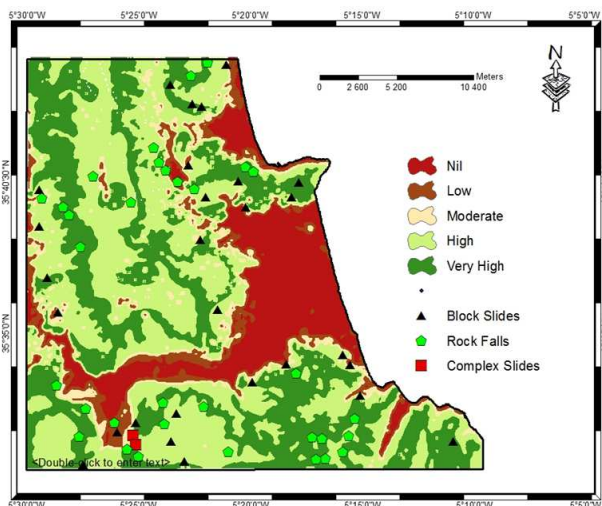
Fig 4. Final Susceptibility map for Tetouan Mazari area using RL model

4 Results and validation

Using the six independent variables as the input factors and the 63 landslide points as a dependent variable, forward stepwise logistic regression was used to analyze the data using the Logistic Regression tool implemented in The ArcGIS along with SDM utilities. An investigation of the susceptibility map highlights a high probability of landslides where mudstones and conglomerates outcrop. The susceptibility maps show also that the landslide is more prone in highly elevated areas greater than 57° with the Northeast, West and Southeast-facing slopes. This map reveals that landslides preferentially occur close to drainage lines and where illumination is present.

Validation of the results is considered to be one of the most important tasks, and without validation, the prediction model will have no scientific significance. The results of the landslide susceptibility map were validated using the test subset of other half remaining (63 landslides points) using the results from the Area Frequency Table and the receiver operating characteristic technique (ROC). The ROC curve is one of those statistical techniques that can be used to provide predictions of the performance and to compare the different models. The overall predicted accuracy of our model is 95.3% which is considered as an excellent according to many authors [21]. Finally, the susceptibility map has been classified into five classes: Very High (32%), High (40%), Medium (7%), Low (7%) and Nil (15%) by using natural breaks methods available in ArcGIS.

According to the results of the Table 1, we notice that the highest value of standard deviation (Sd) for lithology class is 4. This value is found in the formation of marlsones and mudstones or also called argillaceous limestone, which is an intimate mixture of clay and calcium carbonate. These formations regard as homogeneous from the geomorphological and engineering geological point of viewin terms of shallow

slope instability. The Table 2 emphasizes that the highest Standard Deviation of the Logistic Regression is 0.315; which can be explained by shallow landslides that has been developed in argillaceous formations of the area. Thus, the instability causative factors are represented by the poor geotechnical properties of the marl-argillaceous material such as high plasticity, high clay fraction and high content of expansive lattice-type minerals that affect the strength properties of the rocks. In addition to that, the swelling and/or shrinkage of marl-argillaceous formations appear to be a very important factor in the localization of landslide phenomena.

Table 1. Summary statistics of The Logistic Regression Model

Factor	Class	Area (Km²)	Landslides points	Std. C	Class Gen.
Lithology	Conglomerates	120.07	0	0	0
	Sandstones	234.35	21	0.569	99
	Mudstones	57.67	2	-1.308	99
	Marlstones	142.45	21	3.154	4
	Limestones	121.22	11	0.416	99
Fault	Very low	155.91	6	-2.168	1
	Low	95.77	10	0.876	99
	Midium	158.87	25	3.838	3
	High	148.45	12	-0.04	99
	Very high	114.84	2	-2.441	5
Slope	0° - 2°	202.7	15	-0.474	99
	2° - 5°	60.13	10	2.438	2
	5° - 13°	149.65	11	-0.411	99
	13° - 21°	217.62	17	-0.229	99
	21° - 50°	43.75	2	-0.882	99
Aspect	F-N	62.93	4	-0.548	99
	NE-E-SE	235.93	25	1.682	99
	S-E	84.11	6	-0.368	99
	W - SE	172.92	12	-0.68	99
	N - W	117.96	8	-0.602	99
Drainage	Very low	234.04	12	-2.064	1
	Low	223.7	17	-0.376	99
	Medium	162.74	18	1.539	99
	High	48.01	8	2.167	4
	Very high	5.35	0	0	99
Hypsometry	(-) 14 - 84 m	285.55	25	0.482	99
	84 - 204 m	133.62	15	1.434	99
	204 - 360 m	127.43	12	0.574	99
	360 - 572 m	111.8	3	-2.168	4
	572 - 1006 m	15.45	0	0	99

Table 2. The Logistic Regression coefficients obtained for the six independent factors

Factor	LR_Std_Dev
Lithology	0.559
Fault	0.559
Slope	0.39
Aspect	0
Drainage	0.362
Hypsometry	0.636

Regarding the fault parameter, the Table 2 shows that the highest value of the Standard Deviation of RL is 0.559, which is found in the low distance of 0-23m. Thus, we conclude that the areas to be prone to landslides are close to the faulted zone.

The drainage is considered as primary factor to trigger landslide mechanisms. According to the results of the Tables 1 & 2, we observe that the highest probability of the RL (0.362) are at low distances (0 – 59m). Consequently, the proximity to the drainage lines is an important factor controlling the occurrence of landslides.

For the slope, the landslide probability (0.390) is the highest in 0°-5° category. This result can be explained by the number of landslides occurred (10) in the slope's category.

Aspect is an important factor influencing the slope instability, since aspect related parameters, such as exposure to sunlight and drying winds, control the concentration of the soil moisture, which influence landslide occurrence. Based on the result of the final map, is it clear that the landslide percentage is relatively low and that it increases with the orientation angle reaching the maximum on the southwest aspect and then decreases.

Hillshade plays a significant role in landslide susceptibility studies. The results of the table 2 and the final susceptibility map display that the high value of the RL (0.646) is developed in moderate altitudes 204 – 360m and 360 – 572m.

5 Conclusion

The preparation of landslide susceptibility map is a major step forward in risk management. Nowadays, such maps can be prepared by GIS-based techniques and remote sensing data. There are many qualitative and quantitative techniques useful to analyze the relationship between landslides and their causative factors. This study used Logistic Regression (LR) to prepare a susceptibility map for the Tetouan Mazari area, Northern Morocco. Despite the diversity of suggested approaches to landslide susceptibility mapping, researchers still debate the performance and the accuracy of each method. The difficulty in obtaining data and the issue of selecting independent data to analyze, i.e., the parameters that are thought to be causally related, remain challenging in implementing the more data-based approaches. This problem has been discussed in detail by various researchers [21]. Otherwise, an alternative method should be identified because of the time, labor, and cost associated with the collection of inventory data. From the perspective of engineering, the pressure to solve a problem in the shortest period of time, with the most reliable data and in the most economical way requires that a given method should be selected carefully. Similar problems should be expected in the use of numerical methods based on expert opinions. The expert must propose the parameters that are causally related to landslides, and detailed land observations are needed to select the optimal parameters. Recently, the development of large-scale aerial photographs and high-resolution satellite images has considerably decreased the time and labor required for land studies. A total of 126 landslides were mapped using satellite imagery observation and field surveys. The causative factors considered are lithology, slope gradient, slope aspect, faults, drainage lines, and hillshade. After considering all of the potential parameter inputs, the problem solver should also consider the available opportunities and temporal constraints in selecting the optimal method. This study defined the problem, collected data possibly required for the solution, and identified the issues of how and from where these data could be obtained quickly and economically. This change might be attributed to the opportunity to perform inventory mapping more quickly and more easily due to the supply of high resolution images from satellite technologies, which overcome difficulties in time and labor in identifying and evaluating 126 of landslide locations in hundreds of square kilometers. LR-based susceptibility map was clearly more accurate. Consequently, the susceptibility map was successfully generated, validated and has a predicted accuracy of 95.3%. Then it has been classified to five classes: Very High (32%), High (40%), Medium (7%), Low (7%) and Nil (15%). Finally, the results of this study will be used in future for potential construction to mitigate landslide susceptibility.

References

1. R. Schuster, "Landslides: investigation and mitigation. Transportation Research Board National Research Council Socioeconomic significance of landslides," Turner A K, Schuster R L, Washington, 1996.
2. P. Gorsevski, P. Gessler, R. Folt and W. Elliot, "Spatial prediction of landslide hazard using logistic regression and ROC analysis," *Trans GIS,* vol. **10(3)**, p. 395–415, 2006

3.   C. Van Westen, "The modelling of landslide hazards using GIS," *Surveys in Geophysics,* vol. **21** (2–3), p. 241–255, 2000.

4.   T. Van Westen, T. Van Asch and R. Soeters, "andslide hazard and risk zonation - why is it still so difficult?," *Bull Eng Geol Environ,* vo**l. 65**, p. 167–184, 65.

5.   R. Bernknopf, R. Campbell, D. Brookshire and C. Shapiro, "A probabilitic approach to landslide hazard mapping in Cincinnati, Ohio, with applications for economic evaluation," *Bulletin American Association of Engineering Geologists,* vol. **25** (I), pp. 39-56, 1988.

6.   M. El Moulat, L. Ait Brahim, M. Mastere and A. I. Jemmah, "Mapping of Mass Movements Susceptibility in the Zoumi Region Using Satellite Image and GIS Technology (Moroccan Rif)," *International Journal of Scientific & Engineering Research,* vol. **6**(2), pp. 210-217, 2015.

7.   M. Kirat, "Essai de cartographie géomorphologique et étude des mouvements de terrain dans la vallée de l'Oued El Kbir (province de Tetouan, Rif occidental:Maroc Septentrional)," *University of Sciences and Technology of Lille I,* no. PhD diss, 1993.

8.   D. Varnes, "Slope movement types and processes," Schuster RL, Krizek RJ (eds) Landslides, analysis and control, 1978.

9.   L. Ait Brahim , P. Chotin, S. Hinaj , A. Abdelouafi , A. El Adraoui , C. Nakhcha, D. Dhont , M. Charroud , F. Sossey Alaoui , M. Amrhar, A. Bouaza, H. Tabyaoui and A. Chaouni , "Paleostress evolution in the Moroccan African margin from Triassic to Present," Tectonophysics, vol. **357**, no. 1-4, pp. 187-205, 2002.

10. L. Ait Brahim, F. Sossey Alaoui, H. Siteri, M. Tahri and B. Baghdada, "Prise en compte des glissements de terrain dans la quantification des pertes en sols dans le bassin versant Nakhla," *Bulletin du Réseau d'Erosion,* vol. **21**, pp. 50-64, 2002.

11. J. Kornprobst and M. Durand Delga, "Carte Géologique du Rif Tetouan-Ras-Mazari," Ministère de l'énergie et des mines, Direction de la Géologie, Vols. Editions de Service Géologique du Maroc, Notes et Mémoires (No. 292)., 1985.

12. D. Map, Topographical Map of Tetouan, Rabat: Ministry of Agriculture and Agrarian Reform, Department of land conservation and topographic works, 1970.

13. M. Mastere, B. Van Vliet Lanoë and L. Aït Brahim L, "Land use mapping and its relation to mass wasting and gullying in NorthWestern Rif (Morocco). " *Geomorpholog, relief, process, environment*, vol. **3**, pp. 335–352, 2013.

14. ESRI, "Arc Hydro Tools-Tutorial," 2011. [Online]. Available: http://downloads.esri.com/archydro/archydro/tutorial/doc/arc%20hydro%20tools%202.0%20-%20tutorial.pdf. [Accessed 26 March 2017].

15. F. Dai, C. Lee, J. Li and Z. Xu, "Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong," *Environmental Geology,* vol. **40**, pp. 381-391, 2001.

16. M. Ercanoglu and F. Temiz, "Application of logistic regression and fuzzy operators to landslide susceptibility assessment in Azdavay (Kastamonu, Turkey," *Environ Earth Sci,* vol. **64**, p. 949–964, 2011.

17. B. Pham, D. Bui, I. Prakash, L. Nguyen and M. Dholakia, "A comparative study of sequential minimal optimization-based support vector machines, vote feature intervals, and logistic regression in landslide susceptibility assessment using GIS," Environ Earth Sci , vol. **76**, p. 371, 2017.

18. S. Lee, "Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data," *Remote Sens ,* vol. **26**, pp. 1477-1491, 2005.

19. K. Devkota, "Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling-Narayanghat road section in Nepal Himalaya," *Nat Hazards,* vol. **65**, p. 135–165, 2013

20. C. Chang and A. Fabbri, "Validation of Spatial Prediction Models for Landslide Hazard Mapping," Natural Hazards, vol. **30**, p. 451–472, 2003.

21. J. Chacon , C. Irigaray , T. Fernandez and R. El Hamdouni R, " Engineering geology maps: landslides and geographical information systems," Bull Eng Geol Environ , vol. **65**, p. 341–411, 2006.