Proceedings of 12th IOE Graduate Conference
Peer Reviewed
ISSN: 2350-8914 (Online), 2350-8906 (Print)
Year: 2022 Month: October Volume: 12

Comparative Study of Landslide Susceptibility Mapping Along Mountainous Road Using Statistical and Machine Learning Approach

Sagar Karki [a], Bikash Sherchan [b], Krishna Prasad Bhandari [c]

[a, b] *Department of Geomatics Engineering, IOE, Pashchimanchal Campus, Tribhuvan University, Nepal*
[c] *Center for Space Science and Geomatics Studies, IOE, Pashchimanchal Campus, Tribhuvan University, Nepal*
[a] karkimjp77@gmail.com , [b] sherchan.bikash@gmail.com, [c] bhandari.krishna@wrc.edu.np

Abstract
Hilly and mountainous parts of Nepal are most vulnerable to earthquake and landslide-related risks. Every year during monsoon these area encounters highway blockage due to landslides but the concerned authority has not performed any kind of study along this highway to date. This study tends to find the best landslide susceptibility model in these mountainous terrains using statistical (Frequency Ratio) and machine learning (Random Forest) approaches. A total of 239 landslides were mapped using historical landslides, satellite images, and field surveys then these landslides were split into 80% training dataset and 20% testing dataset. Landslide susceptibility mapping was performed based on 12 landslide conditioning parameters under four groups mainly topographic factors (Slope, Aspect, Elevation, Profile curvature, Plan curvature), hydrological factors (Proximity to stream, Precipitation, Topographic wetness index), geological factors (Lithology, Fault line), and other factors (Proximity to road, Land use Land cover). The Landslide susceptibility map produced using both methods was classified into five classes very low, low, moderate, high, and very high. The validity and accuracy were tested by calculating the areas under the curve (AUC) value of the receiver operating characteristic (ROC) curve. The result illustrates the performance of both models where the RF model (AUC=0.902) performed better than the FR model (AUC=0.812). The final landslide susceptibility maps can be used for disaster risk reduction, land use planning, and early warning systems.

Keywords
Landslide, Machine Learning, Mountainous Road, Statistical, Susceptibility

1. Introduction

Landslides are among the most damaging natural hazards in the mountainous terrains of Nepal mainly due to earthquakes, fragile structures, and heavy rainfall. About 2/3 of the total area lies in the mountainous or hilly area with half of the country's population residing in this area. Every year country encounters many flood and landslide-related hazards in the high elevation area due to intense rainfall, especially in the monsoon season. As defined by [1] Landslide is movement of mass like debris flows and soil, debris, rock slides. The study of landslides are gaining popularity worldwide due to its impact on socio-economic factors and due to problem of village hollowing on the mountainous environment [2]. High mountain regions rely on roads for accessing the infrastructures such as hospitals, schools, and marketplaces for their long-term development and economic prosperity [3]. Therefore, understanding the distribution, frequency, and susceptibility of landslides along the transportation network is important for characterizing the impact on the population within this region.

Landslide susceptibility is the likelihood of a landslide occurrence in an area based on past landslides and local terrain conditions [4]. But it is possible to conduct effective landslide risk prediction and mapping based on the available technologies and existing geo-hazard research. But it is possible to conduct effective landslide risk prediction and mapping based on the existing geohazard research and available technologies [5]. Based on the literature review Landslide susceptibility mapping can be categorized as a heuristic, deterministic, statistical, machine learning, and other methods. A heuristic methods are the traditional way of finding the LSM which is based on expert knowledge [6, 7, 8]. Deterministic methods are based on the field survey data and laboratory test data which is only applicable in a small area because for the large area the resource and budget required are high [9, 10, 11]. Statistical methods are based on the mathematical calculation of the field data and expert knowledge. This method produces the result based on the past landslide data and triggering factors [12, 13]. Likewise, another method is machine learning which analyses the relationship between training data sets (past landslides) and triggering factors. Then, based on the relationship model predict the high potential area to landslide [14, 15, 16].

As mentioned above there are many approach for landslide susceptibility mapping among them statistical and machine learning approach are gaining popularity due to their robust calculation and accurate result. Statistical models are built based on the analysis between past landslides and triggering factors. There are various statistical analysis models that had been widely used such as frequency ratio (FR) [15], weights-of-evidence [17], evidential belief function [18], information model. Comparison between different statistical methods was done and the performance of FR model was generally better than others. For example, [11] compared landslide susceptibility mapping of the road section and found that the frequency ratio (FR) performed better than other statistical methods like statistical index (SI), and weights-of-evidence (WoE) approaches. Likewise, [19] compared five statistical methods particularly AHP, WoE, Logistic Regression (LR), FR and Weighting Factor (WF) and concluded that FR gives higher prediction rate 86.59 percent followed by WOE with 82.38 percent, AHP with 77.86 percent, WF with 77.58 percent and finally LR with 70.45 percent. With Advancement in technology and availability of huge amount of data many machine learning approaches have been developed like Support vector machine (SVM), Random forest (RF), LR, Decision tree (DT), Deep learning Neural Networks and so on. Based on various studies related to Landslide susceptibility mapping these models produce better result compared to heuristic and statistical models. But with the better performance they have complex modelling processes and require highly skilled manpower to use the model. Compared to other machine learning approaches, RF has out performance in both classification and prediction, unlike other methods it uses the combination of multiple decision trees and produces the result based on majority voting. Zhao, Liu, and Xu [20]compares different machine learning approach and concludes that RF model have high success rate compare to other models. Likewise, Chen[21] performs a comparative study of RF, Classification and Regression Tree (CART), and Logistic Model Tree (LMT) models and found RF provides an accurate prediction of 83.9 percent than LMT 82.6 percent and CART 77.3 percent.

Researchers around the world use different techniques for landslide susceptibility mapping mainly statistical and machine learning approach but few have done the comparative study between these methods. So, this study tends to perform the comparison between FR model which have good result among other statistical methods and RF model which performs better compared to other machine learning methods. Even though two method have their own algorithm, advantages and exhibit better results, comparative study between two methods can help researcher in model selection. Very less study have been done in the mountainous part of Nepal due to lack of high resolution data and difficult terrain So, this study will help researcher and policy maker to work in the area of landslide disaster prevention, land use plan and early warning system.

2. Materials and Methods

2.1 Study Area

Study was conducted in Pokhara-Beni highway of Gandaki province, Nepal. This highway runs east to north-west starting at province capital Pokhara in Kaski District and ending in Beni of the Myagdi District. Highway mostly runs along the hilly part from Pokhara to Nayapul and then it runs parallel to Modi river between Nayapul and Kushma, Beyond Kushma it follows Kali Gandaki river. It is only a proper roadway to connect hilly districts with the province capital Pokhara. Highway is 84 km in length and extends from latitude 28.209°N to 28.376°N and longitude from 83.564°E to 83.985°E. For this study, we defined a 3 km buffer from the centreline of the highway on both sides which covers an area of 280.23 km^2. Study area lies in subtropical region, with hot and rainy summers and relatively cold and rainless winter. Likewise, the study area is abundant with rainfall, as Pokhara and Lumle two of the highest rainfall area of the country lies along the highway. According to Department of hydrology and meteorology annual average rainfall along the highway ranges from 150mm to 450mm with most of the rainfall occurring during the monsoon season i.e. July-August ranging from 400mm to 1400mm. Most of the highway runs along the hilly area so the elevation ranges from 649m to 2376m above the mean sea level with slope angle up to 76 degrees.

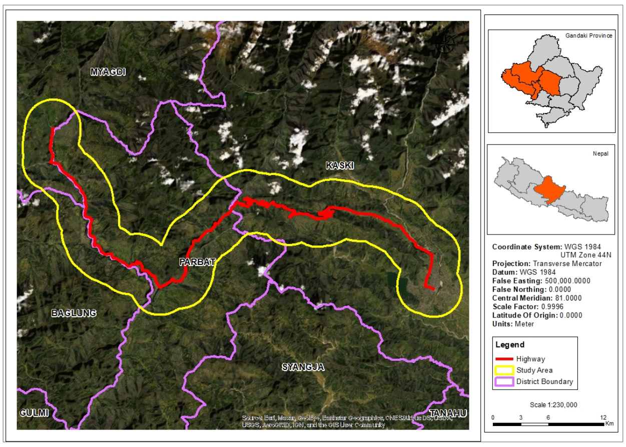

Figure 1: Study Area

2.2 Datasets

2.2.1 Landslide Inventory

The landslide susceptibility map uses existing landslide data (landslide inventory map) and triggering factors to find out the relationship between these two factors and predict the possibility areas where landslide can occur. Being one of the key factor for finding the LSM, landslide inventory mapping was done by evaluating historical landslide data, field surveys, and satellite imagery. Altogether 239 landslides were mapped inside the study area. These were the data of the dependent variables (landslide polygon) which was used to train the model. The dependent variable constitutes the occurrence (or not) of landslides. The size of landslides varies with the location and their topography with the largest landslide polygon with an area of 116476.28 km^2 and the smallest polygon with an area of 169.96 km^2 respectively. The average area of the landslide polygons was 5600.73 km^2. Along with the landslide data non-landslide data are also important because taking only landslide data as training data can result in overestimation of the model. Many study have suggested the ratio of landslide to non-landslide as 1:1 [22, 23]. Therefore, the same number of non-landslide data were randomly selected excluding the 500m buffer of the existing landslide and stream network in the study area.

2.2.2 Landslide Conditioning Factors

Another important factor uses to model the landslide susceptibility mapping is the landslide conditioning factor or triggering parameters. The relationship between past landslides and the environment i.e. conditioning factors in which the landslide has occurred can be used to predict the future event. To find out the landslide susceptible area we have analyzed the physical parameters that control landslide susceptibility i.e. landslide conditioning factors. We have categorized landslide triggering factors into four groups mainly topographic factors, hydrological factors, geological factors, and other factors. Altogether 12 parameters were used in this study under four category.

Topographic factors

Topographic factor consists of five parameters like elevation, slope, aspect, plan curvature and profile curvature. Elevation status was derived from the digital elevation model (DEM). The elevation is considered as an important triggering factor for landslides. Elevation of the study area ranges from 649m to 2376m (Fig. 2). Being the mid hill highway most of the area are located in higher altitude region. Slope angle is another important factor which have higher impact factor compare to other conditioning factors. Mostly area with higher slope exhibits more landslide as these regions are steep and can causes more stress on the environment. Slope of the study area varies from 0 degree to 83 degrees. The bulk of study area were very steep and rugged (Fig. 3). Aspect is another important conditioning factor used for LSM. It denotes the slope direction and have difference in the value of sun exposure. The aspect was categorized into nine classes based on the facing direction using angular values (Fig. 4). The curvature is defined as the rate of change of slope in a particular direction. Plan curvature shows the surface which is perpendicular to slope direction. It usually has three types laterally convex, laterally concave and linear denoted by positive, negative and zero value respectively (Fig. 5). Profile curvature shows the surface which is parallel to maximum slope direction. It usually has three types upwardly convex, upwardly concave and linear denoted by negative, positive, and zero value respectively (Fig. 6).

Hydrological factors

Hydrological factor consists of three parameters proximity to stream, precipitation and Topographic Wetness Index. Stream proximity is another important parameter for LSM as surface closer to stream exhibit soil erosion and landslides due to the scouring effect of water. Six buffer zone have been drawn around Rivers based on the Euclidean distance (Fig. 7). Precipitation is one of the most important triggering factors for a landslide. Most of the landslide occurred at place which encounters maximum rainfall. Our study area consists of 5 major rain gauge station in Pokhara, Lumle, Kushma, Baglung, and Beni. The data was obtained from hydrology and meteorology department and these data were interpolated in GIS using kriging technique. 32 (1990-2021) year average rainfall data from department of hydrology and meteorology was purchased and used for this study. It was categorizing into four classes ranging from 2000mm to 4000mm of yearly precipitation (Fig. 8). The topographic wetness index (TWI) is another important factor which contribute to LSM. It is prepared using the byproduct of DEM such as slope degree and flow accumulation. It is expressed as: TWI = $\ln(\alpha/Tan\beta)$ where α is the upslope value from flow accumulation i.e. upstream contributing area and Tanβ is slope angle. It usually shows the soil moisture content and tendency of an area to accumulate water. It was classified into five classes(Fig. 9).

Geological factors

Geological factor also consists of two parameters proximity to a fault line and Lithology. Both the data were obtained from the Department of mines and geology. Lithology mainly focuses on the types and formation of the rocks. As Nepal has young geology with the fragile structure it is vulnerable to landslide related risk. Study area comprises of four types of lithology mainly Gneiss migmatite, Quartzite, Fluvial non calcareous and Slate Phyllite (Fig. 10). Another geological factor that is important for LSM is fault line. These lines are usually the fracture in the earth's surface and earthquake occurs and affects these surface more than other normal surface so, the places close to fault line are more likely of experiencing mass movement. Six category was made based on the Euclidean distance from fault line (Fig. 11).

Other factors

All the remaining parameter comes under other factor which consist of land use land cover and proximity to a road. Proximity to Road is also one of the important factor which directly contribute to landslide. Like stream proximity the areas closer to road network have higher impact in landslide than area farther from road. Also, newly constructed road can cause more landslide as it will break the sloe stability and causes pressure on the upper slope of road. For this reason, Six category was made based on the Euclidean distance from centerline of roads (Fig. 12). Land Use Land Cover was another parameter use for landslide susceptibility mapping where it was categorize into five classes mainly agriculture, Riverine and Lake, forest, built-up and, barren where most of the area is covered by forest i.e. 52.67%, 8.3% of the land is covered by Built-up, 32.11 by agriculture, 1.69% by Riverine and Lake, and 5.22% by barren land (Fig. 13).

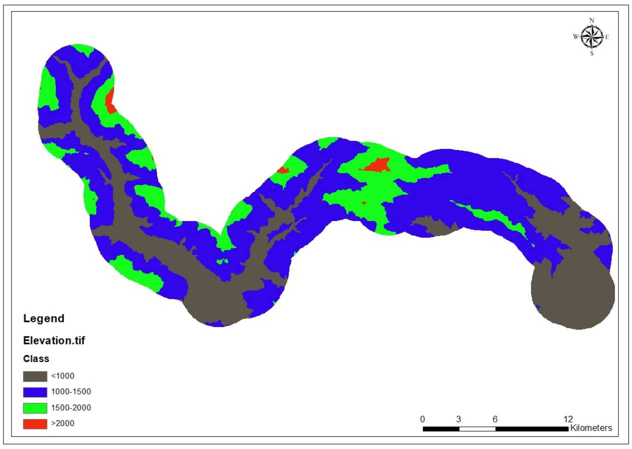

Figure 2: Elevation Map

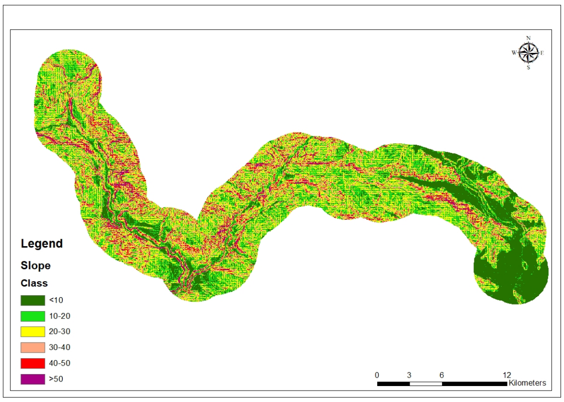

Figure 3: Slope Map

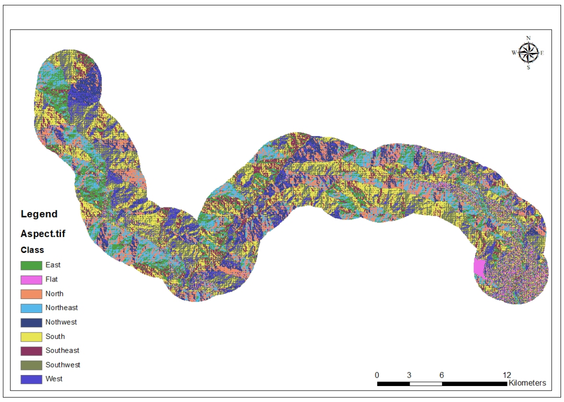

Figure 4: Aspect

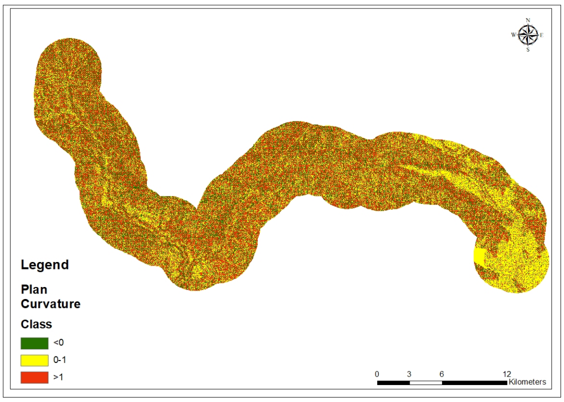

Figure 5: Plan Curvature

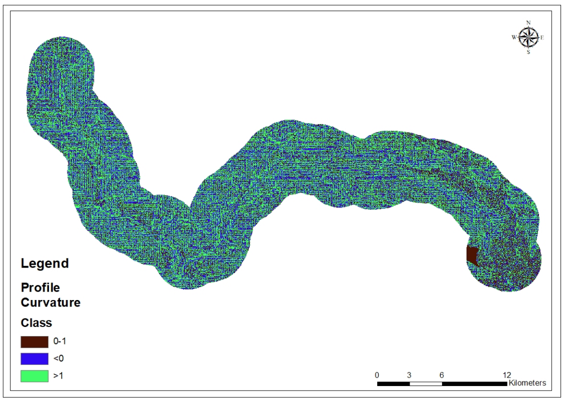

Figure 6: Profile Curvature

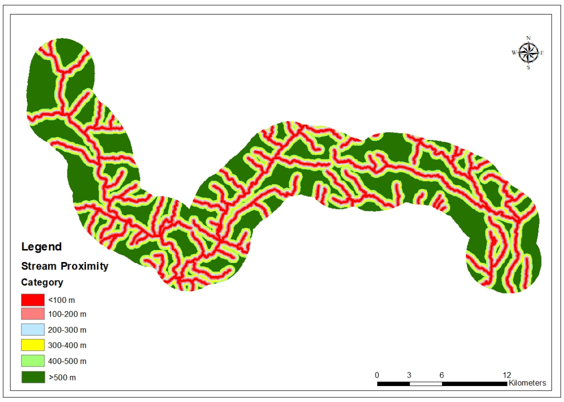

Figure 7: Proximity to Stream

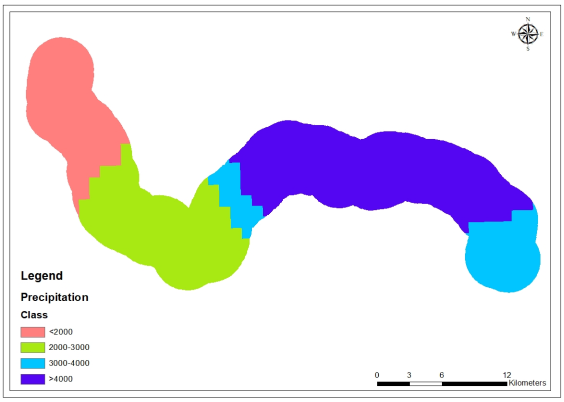

Figure 8: Precipitation

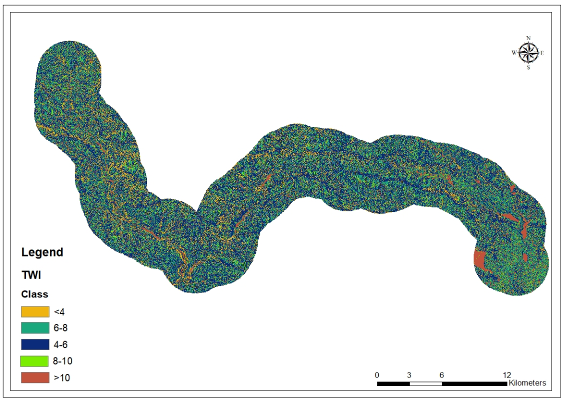

Figure 9: TWI

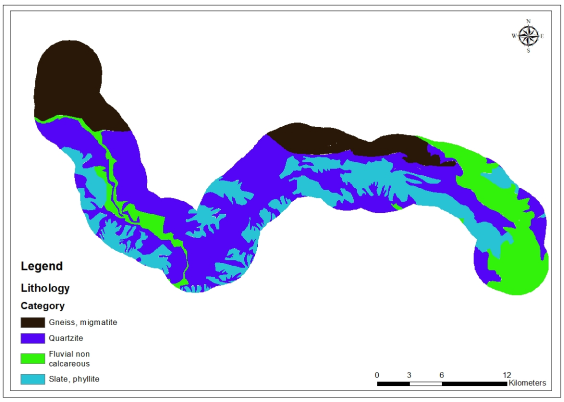

Figure 10: Lithology

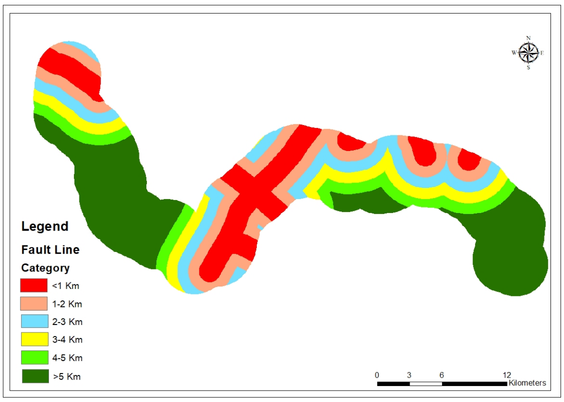

Figure 11: Fault Line

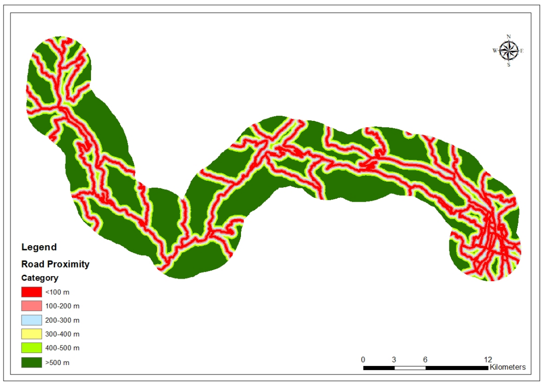

Figure 12: Proximity to Road

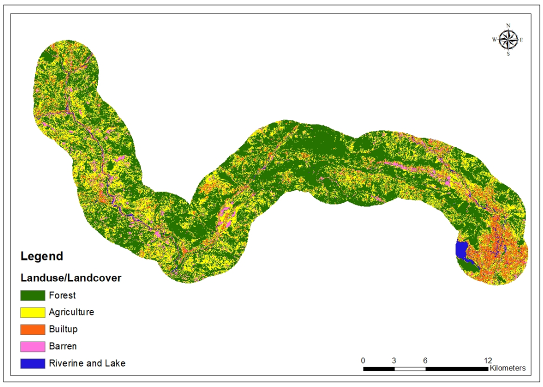

Figure 13: Landuse/Landcover

2.3 Methodology

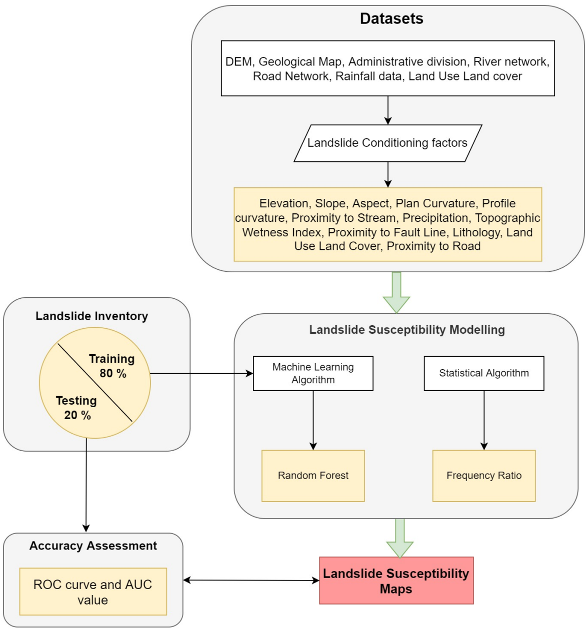

Figure 14: Flow chart of Methodology

Two methods statistical and machine learning were used in this study based on the landslide inventory. (Fig. 14) shows the detail process adopted for landslide susceptibility mapping. Usually both the method uses same datasets i.e. conditioning factors and landslide inventory.

2.3.1 Frequency Ratio (FR)

The FR is the ratio of the probability of the landslides occurrence to the probability of non-landslide in a given area. It is usually the quantitative relation between landslide inventory and conditioning factors. The frequency ratio has been calculated using Eqn 1

$$FR = \frac{(NLS/\sum NLSi) * 100}{(NC/\sum NCi) * 100} \quad (1)$$

where FR=frequency ratio value, NLS=number of landslide pixels in a class of a factor, NLSi=sum of all landslide pixels in the entire area, NC=number of pixels in a class of a factor and NCi=sum of all pixel class in the entire area. Frequency ratio value calculated using the Eqn. 1 was used to reclassify all the 12 factors and also use to find the prediction ratio. Finally, the frequency ratios of each factor's type or class were summed to calculate the landslide susceptibility index using Eqn. 2.

$$LSI = \sum FRi \quad (2)$$

where LSI is the landslide susceptibility index and is the FR of each factor range or class.

2.3.2 Random Forest (RF)

Random forest is one of the popular machine learning approach that builds multiple decision trees from different subsets of data which uses the bagging technique to randomly select samples from the training dataset for the classification and regression tree construction. It was first introduced by [24]. Comparing with other landslide division techniques, RF methods uses two random sampling i.e. features and samples. Unlike single decision tree RF method improve the accuracy of the model using randomly generated methods to select samples and features. Then the majority vote from multiple decision tree results the final output. R studio was used to develop the random forest model with the training dataset where the performance of the model was enhanced by using the 10fold cross-validation in the caret package in R. Other hyper parameters were also tuned in R to find the best value.

3. Result and Discussion

As we are performing the comparative analysis between statistical and machine learning approach. It is done based on three approaches: Landslide Susceptibility Mapping results, Conditioning Factor Importance and Model Performance evaluation and comparison.

3.1 Landslide Susceptibility Mapping results

For statistical approach we have used the frequency ratio method which generate the result based on quantitative relation between landslides and conditioning factors. LSM was calculated based on the equation 1 and equation 2 on ArcGIS platform. The Landslide susceptibility map was reclassified into five classes using natural breaks method [25]. Moderate class has largest area (26.61%), followed by high (23.65%), low (22.13%), very low (17.98%), and, very high (9.63%), (Fig. 15).

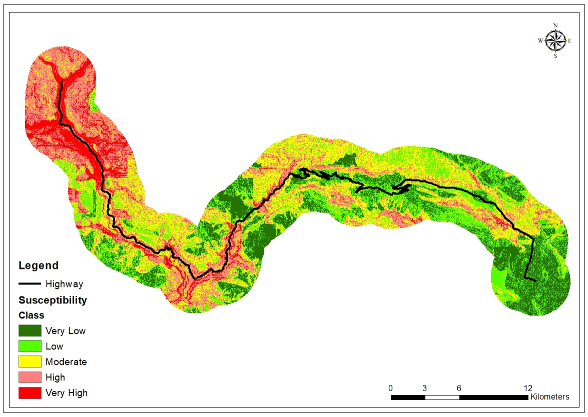

Figure 15: Suitability Map using FR method

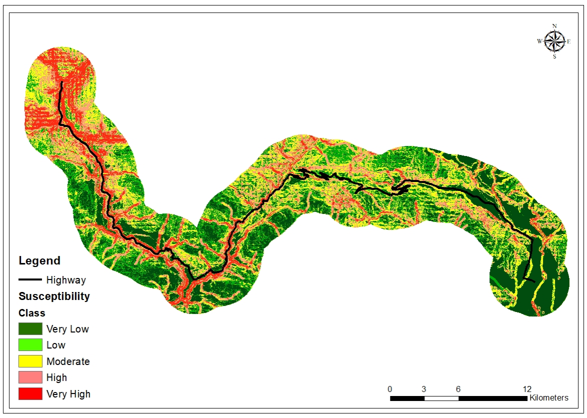

Figure 16: Suitability Map using RF method

Likewise, for the machine learning approach we have used Random Forest method in R studio and the LSM was generated which was also classified into five classes very low (25.71%0, low (22.96%), moderate (23.03%, high (17.75%) and, very high (10.55%), (Fig. 16).

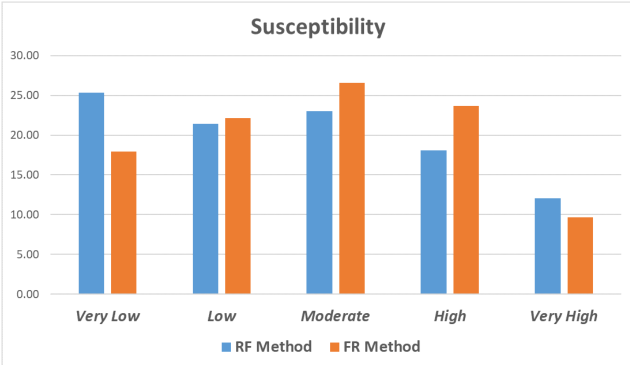

Figure 17: Suitability Map Comparison

Comparison between two methods is shown in the graph (Fig. 17). Result shows that PATICHAUR to DOBILLA and CHHAMARKE to GALESHWOR section of road are highly susceptible to landslides. After evaluating all the results from both frequency ratio and random forest method it was found that there was a certain variation in output. Firstly, landslide susceptibility maps prepared using both methods have some similarities in terms of their percent values for all classes but their distribution in the map is different. Especially, from Maldhunga to Galeshwor section the map prepared using the FR method shows a huge percent of high and very high zone whereas the map prepares using RF method shows comparatively less area at the high and very high zone with the majority portion under the moderate zone. After analysing the result and performing the field verification it was found that RF method shows the better result compared to FR method. Being one of the best statistical method frequency ratio still exist some heuristic approach for data analysis such as in case of defining parameter classes and FR method is based on the ratio between these classes and landslide inventory data. So, there is always a room for bias in defining the classes which result in variation of output based on expert opinion. Likewise, splitting of landslide inventory data into training and testing can also have bias because splitting is usually done randomly which can cause uneven distribution of data. But Random forest algorithm uses the machine learning approach for data splitting so there is no room for bias and also 10-fold cross validation is performed in RF method which uses the best combination of training data for model creation and hyper parameter optimization for better result. Apart from these fact RF method uses both landslide and non-landslide data for model training and testing which will reduce the problem of under fitting or over fitting.

3.2 Conditioning Factor Importance

Frequency ratio and random forest method uses 12 conditioning factors Elevation, Slope, Aspect, Plan curvature, Profile curvature, Fault line, Lithology, Topographic Wetness Index, Precipitation, Proximity to Stream, Proximity to Road, Land use/Land cover for LSM. Both the methods uses 80% training data to train the model. In FR method Elevation have greater importance followed by Precipitation, Slope and Lithology whereas Proximity to Stream, TWI, Proximity to Road have comparatively lower importance. Likewise, in RF method Slope, Elevation, Precipitation have high importance whereas TWI, Plan curvature, Profile curvature have lowest importance. Comparative analysis between the feature importance are shown in the (Fig. 18)

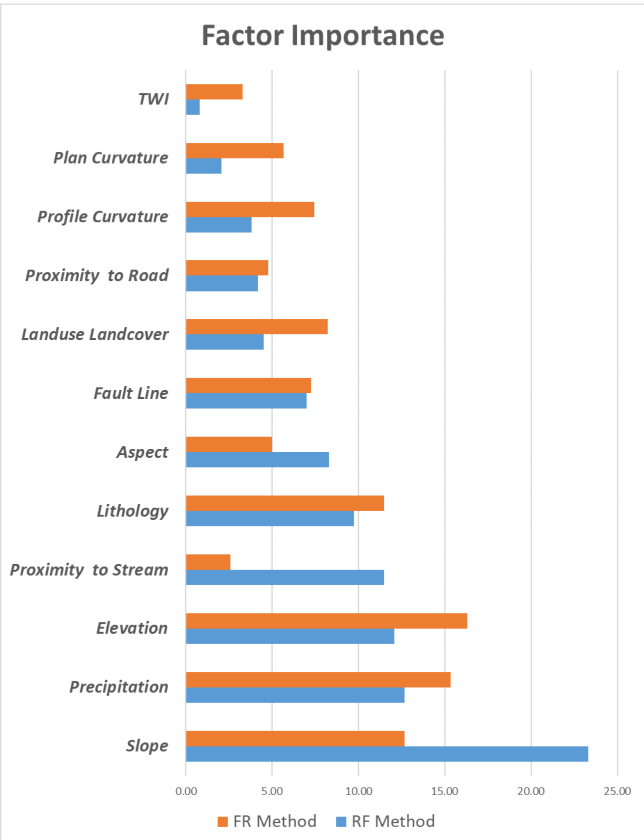

Figure 18: Factor Importance Comparison

3.3 Model Performance evaluation and comparison

Accuracy assessment was done for both the method using the 20% of the test dataset. Both the method uses area under the curve (AUC) value of Receiver Operating Characteristics (ROC) curve to test the model. In Frequency Ratio method AUC was calculated in ArcGIS platform whereas in Random Forest method model testing was done in R studio. Pourghasemi[13] describe that success rate curve is used to evaluate how well landslide susceptibility maps classify existing landslide areas. The AUC value ranges from 0 to 1, value closer to 1 means model have higher prediction accuracy whereas lower means lower accuracy. The AUC for both the model was obtained based on training and testing sample. (Fig. 19 and 20) illustrate that Random forest model have highest prediction accuracy with an AUC value of 0.902 whereas Frequency ratio model have AUC value of 0.812.

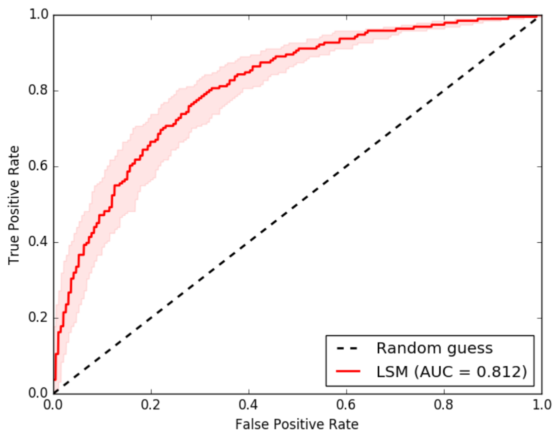

Figure 19: AUC of ROC curve

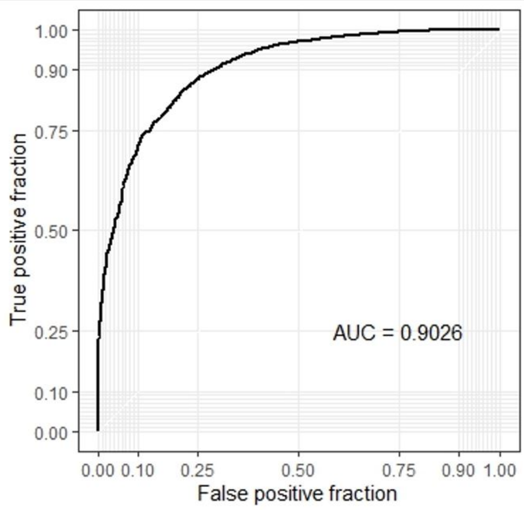

Figure 20: AUC of ROC curve

4. Conclusion and Recommendation

Landslide susceptibility mapping provides the possible causes and location of landslides which can be useful for evaluation and prevention of landslides. In this study we have used statistical and machine learning approach to find the LSM along the Pokhara-Beni highway and their comparative study was done. The main conclusion of the study are as follows: Total of 239 landslide samples were collected based on the historical landslides, satellite imagery, and field surveys. During the observation it was found that most of the landslides occurred during the monsoon season due to heavy rainfall and new road construction. The landslide inventory was split into 80-20% ratio for training and testing purpose. But for the RF method equal number of non-landslide were used and altogether 478 samples were split for training and testing. Susceptibility mapping was performed based on 12 landslide conditioning factors using training data. LSM using FR model shows that about 33% of the area falls under high and very high class and remaining 67% under other classes. Likewise, RF model shows about 28% of the area falls under high and very high class and remaining 72% under other classes. While evaluating the importance of conditioning factors both the model shows Slope, Elevation and Precipitation has high importance while other factors have relatively low importance. The result shows that Random forest model have highest prediction accuracy with an AUC value of 0.902 whereas Frequency ratio model have AUC value of 0.812. Result also shows that PATICHAUR to DOBILLA and CHHAMARKE to GALESHWOR section of road are highly susceptible to landslides. After evaluating all the factors, it was found that both the Frequency ratio and Random forest have yield better result in the mountainous terrain. So, we can use both the method for LSM but RF method is highly recommendable. Finally, the result from this study can be used for land use planning and disaster risk reduction. Since government is constructing Mid-Hill highway along the most section of the study area it is highly recommended to use the susceptible map and use early warning system in the area expose to population.

Acknowledgments

The authors would like to thank the Department of Geomatics Engineering, IOE, Pashchimanchal Campus, Tribhuvan University for supporting this research. The authors are grateful to Er. Rajan Kumar Adhikari and Er. Digvijaya Paudel for their valuable contribution during data collection and Reporting.

References

[1] David J Varnes. Slope movement types and processes. *Special report*, 176:11–33, 1978.

[2] Pietro Aleotti and Robin Chowdhury. Landslide hazard assessment: summary review and new perspectives. *Bulletin of Engineering Geology and the environment*, 58(1):21–44, 1999.

[3] Jeet Bahadur Sapkota. Access to infrastructure and human well-being: Evidence from rural nepal. *Development in Practice*, 28(2):182–194, 2018.

[4] Fausto Guzzetti, Paola Reichenbach, Mauro Cardinali, Mirco Galli, and Francesca Ardizzone. Probabilistic landslide hazard assessment at the basin scale. *Geomorphology*, 72(1-4):272–299, 2005.

[5] Yang Zhang, Weicheng Wu, Yaozu Qin, Ziyu Lin, Guiliang Zhang, Renxiang Chen, Yong Song, Tao Lang, Xiaoting Zhou, Wenchao Huangfu, et al. Mapping landslide hazard risk using random forest algorithm in guixi, jiangxi, china. *ISPRS International Journal of Geo-Information*, 9(11):695, 2020.

[6] Ali Yalcin. Gis-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in ardesen (turkey): comparisons of results and confirmations. *catena*, 72(1):1–12, 2008.

[7] Bakhtiar Feizizadeh and Thomas Blaschke. Gis-multicriteria decision analysis for landslide susceptibility mapping: comparing three methods for the urmia lake basin, iran. *Natural hazards*, 65(3):2105–2128, 2013.

[8] Thomas Stanley and Dalia B Kirschbaum. A heuristic approach to global landslide susceptibility mapping. *Natural hazards*, 87(1):145–164, 2017.

[9] CJ Van Westen. Geo-information tools for landslide risk assessment: an overview of recent developments. *Landslides: evaluation and stabilization*, 1:39–56, 2004.

[10] Renato Fontes Guimaraes, David R Montgomery, Harvey M Greenberg, Nelson Ferreira Fernandes, Roberto Arnaldo Trancoso Gomes, and Osmar Abılio de Carvalho Júnior. Parameterization of soil properties for a model of topographic controls on shallow landsliding: application to rio de janeiro. *Engineering Geology*, 69(1-2):99–108, 2003.

[11] Amar Deep Regmi, Krishna Chandra Devkota, Kohki Yoshida, Biswajeet Pradhan, Hamid Reza Pourghasemi, Takashi Kumamoto, and Aykut Akgun. Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in central nepal himalaya. *Arabian Journal of Geosciences*, 7(2):725–742, 2014.

[12] Biswajeet Pradhan and Saro Lee. Delineation of landslide hazard areas on penang island, malaysia,

by using frequency ratio, logistic regression, and artificial neural network models. *Environmental Earth Sciences*, 60(5):1037–1054, 2010.

[13] HR Pourghasemi, HR Moradi, and SM Fatemi Aghda. Landslide susceptibility mapping by binary logistic regression, analytical hierarchy process, and statistical index models and assessment of their performances. *Natural hazards*, 69(1):749–779, 2013.

[14] Taskin Kavzoglu, Ismail Colkesen, and Emrehan Kutlug Sahin. Machine learning techniques in landslide susceptibility mapping: a survey and a case study. *Landslides: Theory, practice and modelling*, pages 283–301, 2019.

[15] Halil Akinci, Cem Kilicoglu, and Sedat Dogan. Random forest-based landslide susceptibility mapping in coastal regions of artvin, turkey. *ISPRS International Journal of Geo-Information*, 9(9):553, 2020.

[16] Chenglong Yu and Jianping Chen. Landslide susceptibility mapping using the slope unit for southeastern helong city, jilin province, china: a comparison of ann and svm. *Symmetry*, 12(6):1047, 2020.

[17] Netra R Regmi, John R Giardino, and John D Vitek. Modeling susceptibility to landslides using the weight of evidence approach: Western colorado, usa. *Geomorphology*, 115(1-2):172–187, 2010.

[18] Biswajeet Pradhan, Mohammed Hasan Abokharima, Mustafa Neamah Jebur, and Mahyat Shafapour Tehrany. Land subsidence susceptibility mapping at kinta valley (malaysia) using the evidential belief function model in gis. *Natural hazards*, 73(2):1019–1042, 2014.

[19] Hamid Bourenane, Mohamed Said Guettouche, Youcef Bouhadad, and Massinissa Braham. Landslide hazard mapping in the constantine city, northeast algeria using frequency ratio, weighting factor, logistic regression, weights of evidence, and analytical hierarchy process methods. *Arabian Journal of Geosciences*, 9(2):1–24, 2016.

[20] Zhou Zhao, Zeng Yuan Liu, and Chong Xu. Slope unit-based landslide susceptibility mapping using certainty factor, support vector machine, random forest, cf-svm and cf-rf models. *Frontiers in Earth Science*, 9:589630, 2021.

[21] Wei Chen, Xiaoshen Xie, Jiale Wang, Biswajeet Pradhan, Haoyuan Hong, Dieu Tien Bui, Zhao Duan, and Jianquan Ma. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *Catena*, 151:147–160, 2017.

[22] Haoyuan Hong, Yamin Miao, Junzhi Liu, and A-Xing Zhu. Exploring the effects of the design and quantity of absence data on the performance of random forest-based landslide susceptibility mapping. *Catena*, 176:45–64, 2019.

[23] Yu Huang and Lu Zhao. Review on landslide susceptibility mapping using support vector machines. *Catena*, 165:520–529, 2018.

[24] L Breiman. Random forests mach learn 45 (1): 5–32, 2001.

[25] Faming Huang, Zhongshan Cao, Jianfei Guo, Shui-Hua Jiang, Shu Li, and Zizheng Guo. Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping. *Catena*, 191:104580, 2020.