Landslide Prediction Modelling in the Maharashtra Region of Western Ghats

Prof. S. R. Parekar[1], Omkar Thombare[2], Yashvardhan Todmal[3], Shreyas Naiknavare[4], Atharva Potdar[5], Akshay Yadav[6]

Department of Civil Engineering, AISSMS College of Engineering, Pune- 01

Abstract

This study investigates the development and implementation of a landslide prediction model specifically crafted for the Western Ghats region of Maharashtra. The primary aim of this study are to predict landslides accurately and provide early warnings to mitigate impact on communities and infrastructure. Questions we aimed to answer are: Which factors key influence landslide occurrence in this region? How can these factors be integrated in a predictive model?

In order to answer this questions, a dataset was employed that included environmental variables like precipitation, level of elevation, and angle of slope, which known to have a impacts significant on landslide occurrences. A Model of Logistic Regression was applied for the analysis of the relationship between these factors and past landslide occurrences. The model was trained and validated using the data historical from various locations within the Western Ghats

This evidence of our findings indicates that the integration of precipitation, level of elevation, and angle of slope data into a Logistic Regression context provides a robust predictive capability for landslide occurrence. The accuracy levels of the model were high in identification of areas with high risk and in anticipation of potential landslides. The importance of continual monitoring and updating of the model for incorporation of changes environmental conditions were underscored by the results.

1. Introduction

Landslides are a significant natural hazard that pose a substantial risk to the lives, infrastructure, and the environment, particularly in geologically sensitive regions like the Western Ghats of Maharashtra. This study focus on the development of a landslide prediction model that exploits environmental factors such as precipitation, level of elevation, and angle of slope to predict landslide occurrences and provide early cautions. Key tems central to the study are landslides, precipitation, level of elevation, and angle of slope. A landslide refers to the movement of a large of rock, debris, or earth down a slope due to gravity. Precipitation, the precipitation in the form of liquid water drops that fall from clouds, is a critical factor influencing landslide occurrences. The level of elevation, the height of a geographic location above a fixed reference point (often sea level), and the angle of slope, the steepness or incline of a slope, are major determinants of landslide susceptibility.

According to previous research rocks like Guzzetti et al. (2007) recognized the intensity and duration of precipitation as primary stimuli for landslides. Moreover, Dai and Lee (2002) stress the role of slope angle and geological characteristics in vulnerability to landslides. Research by Petley et al. (2005) have illustrated the effectiveness of models statistical and learning machine in prediction of landslides on historical base. Despite this progress, prediction of landslides succintly remains a challenge due to the complex interaction of environmental factors. Most models existent focus either on one factor or lack the interaction needed to account the multifaceted nature of causes of landslides. What is know is that a combination of precipitation, level of elevation, and slope angle significantly influences occurrence of landslides, precise relationships and predictive capabilities require explorations further. This study develops a logistic regression model comprehensive that integrate of these factors key to predict landslides in the Western Ghats region. Through incorporation of high-resolution data and techniques advanced preprocessing, this research advanced enhances accuracy and reliability of landslide predictions. The findings of this study will not only contribute to the understanding scientific mechanics of landslides but provide tools practical for early caution and risk management, ultimately helping in protection of communities vulnerable and infrastructure in the Western Ghats.

2. Literature Review

According to Goetz2015, the usefulness of models statisticians and learning machines techniques for predicting has recourse in modeling of susceptibility to landslides. Specifically, these methods data-driven demonstrate potential great for solving problem mapping of areas prone to landslides across regions extensive where there insufficient information geotechnical to carry out approaches physically-based.

Ji2022 aims to develop a tool user-friendly of extension of system information geographic (GIS) called the toolbox GIS-FORM prediction of landslides using the language programming Python. This toolbox will consider the uncertainties possible in analysis susceptibility to landslides based physically in areas seismic. For the realization of this, the algorithm first-order reliability method (FORM) is implemented to calculate the probability of slope failures infinite. This toolbox proposed generating various maps regional distribution of hazards, including the factor of security (FoS), the index of reliability (RI), and the probability of failure (Pf). Additionally, it enables prediction of landslide displacements coseismic using either the method direct integration of Newmark or the formula empirical. The outputs of the analysis prediction of landslides GIS-FORM are validated using data published in the literature. The toolbox has also been used successfully for the analysis of susceptibility to landslides Miss 7.0 Jiuzhaigou earthquake in the province of Sichuan, China. In general, the toolbox GIS-FORM prediction of landslides can be for mapping of hazards fast regional landslides induced by earthquakes, where they should be taken into account uncertainties in geological and geotechnical parameters.

Accuracy of the existing models and capacity of prediction are restricted in terms of accuracy, as per Kuradusenge2020. To overcome these limitations, research proposes two approaches for prediction modeling: forest random (RF) and logistic regression (LR). These methods use datasets of precipitation and other internal and external parameters for prediction, of landslides thereby improving the accuracy. Furthermore, the performance of these approaches improved using antecedent accumulation data of rainfall. The performance of these models was assessed using receiver operating characteristics, area under the curve (ROC-AUC), and false negative rate (FNR) to measure the number of cases of landslides not reported. When the prediction is considered antecedent of rainfall data, both models (RF and LR) performed better, with AUC values of 0.995 and 0.997, respectively. The results indicate a correlation strong between precipitation antecedent and the occurrence of landslides rather than between rainfall on one day and the occurrence of landslides. Additionally, the rate of false negatives (FNR) for both models improved significantly, with the RF having a FNR of 10.58% and the LR having a FNR of 5.77%. Among the internal factors used for prediction, the angle of slope has an impact that is most significant.

3. Methodology

A study area of Western Maharashtra consist the districts of Pune, Mumbai, Nashik, Satara, Ahmadnagar, Raigad, Ratnagiri, and Thane were considered. Further the data was gathered for these districts.

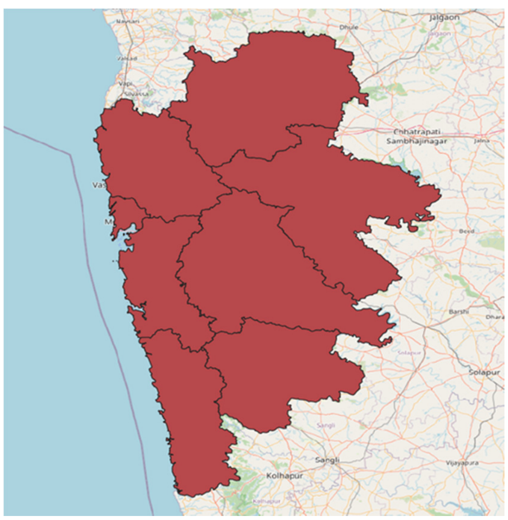

Fig 1: Location selected

3.1 Data Collection

Several sources were used in order to gather the data regarding precipitation, angle of slope, and level of elevation of the locations prehistoric of landslides. Services online, government sites, platforms of GIS, papers of research and news reporting were utilized to collect the data.

3.1.1 Rainfall Data

Data of precipitation for this study was sourced from multiple platforms reliable. The India Information System of Water Resources (IndiaWRIS) provided data from stations of raingauge across the Western Ghats, supplying particulars detailed of precipitation daily and monthly. Google Earth Engine supplemented this data with precipitation satellite-based, aiding in the validation and filling of gaps in the data of raingauge. In addition, data from papers previous on events landslides were incorporated, providing patterns historical of precipitation and thresholds linked to landslides previous. This approach comprehensive ensured a dataset robust for development of a model accurate of prediction of landslides.

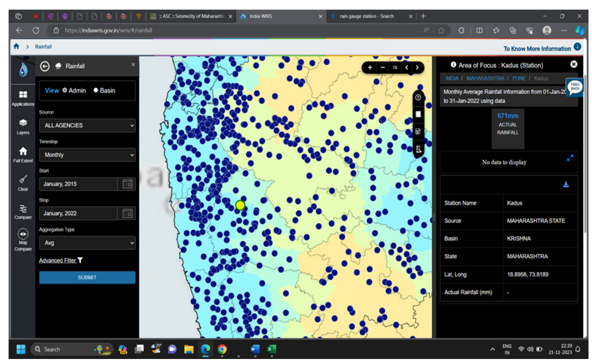

Fig 2: IndiaWRIS website- Location of rainguage stations

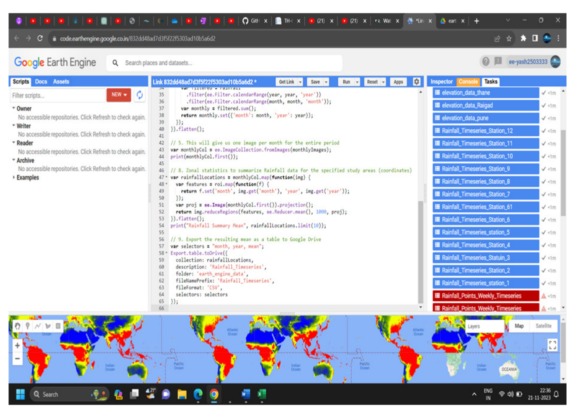

Fig 3: Google Earth Engine interface

3.1.2 Slope Angle

The angle of slope of the events of landslides was determined using an image Digital Elevation Model (DEM) raster in QGIS. First, the DEM was imported into QGIS, and the tool of "Slope" within the toolbox Raster Analysis was utilized to calculate the angle of slope for each cell in the raster. This tool processes the data elevation to generate a new layer of raster representing the angle of slope, expressed in degrees. The aspect of slope, that indicates the direction of the slope, was derived also to complement the data of angle of slope. By overlaying locations of events historical of landslides on the raster of angle of slope generated, we extracted the angles of slope specific at these points. This process allows for the determination precise of the angles of slope associated with landslides past, providing entry essential for our model of prediction of landslides.

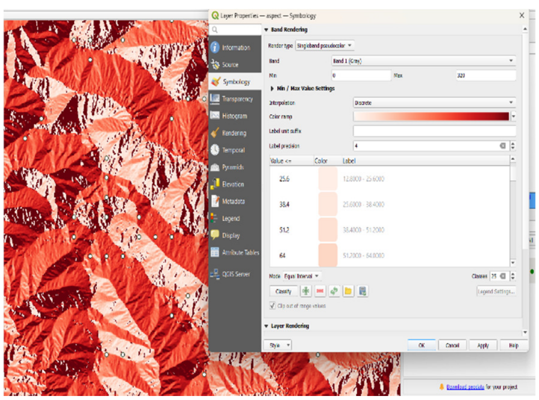

Fig 4: Slope aspect mapping using Qgis

3.1.3 Elevation

Data level of elevation for this study was gathered using a Model Elevation Digital (DEM) obtained from NASA site Earth Data. The DEM provides information elevation high-resolution indispensable for analysis topographical. The data was imported into QGIS, a system geographic powerful open-source, where it was processed and analyzed. Using QGIS, the DEM was visualized and manipulated to extract values precise of elevation across the study area in the Western Ghats. This data of elevation served as a component important in understanding the influences topographical on the occurrences of landslides and was integrated into the model of prediction of landslides to enhance its accuracy and reliability.

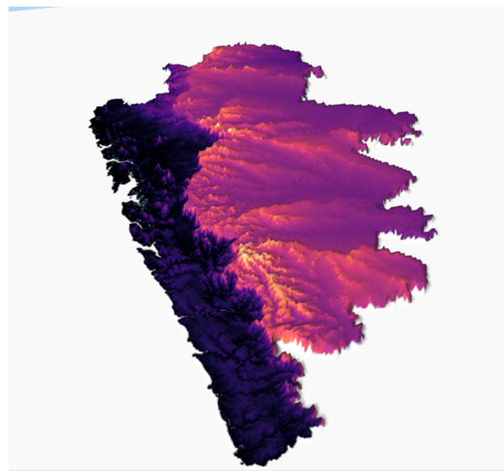

Fig 5: Digital Elevation Model of study area

Table 1: Coordinates of landslide events.

id	Easting_x	Northing_y	elevation_z
LSEVT001	73.78845	20.02512	586.1
LSEVT002	73.48842	19.69984	401.5
LSEVT003	72.89879	19.65275	15.5
LSEVT004	73.27174	19.4407	216.6
LSEVT005	72.81754	19.48981	1.5
LSEVT006	73.05922	19.30901	10.5
LSEVT007	73.16548	19.23821	25.1
LSEVT008	73.77485	19.34877	298.1
LSEVT009	73.68067	19.18831	1004.9
LSEVT010	73.45556	18.8104	645

3.2 Integration of Data

To shape a dataset comprehensive for the model of prediction of landslides, all the data collected was integrated meticulously into a Excel sheet single. The process began with the cleaning of the data to ensure accuracy and consistency. Data of precipitation from IndiaWRIS, Google Earth Engine, and papers historical of research was standardized and any missing values were addressed. Data of elevation derived from DEM of NASA was aligned with the locations geographic corresponding of events landslide. Data of angle of slope calculated in QGIS was also integrated. Next, the data was structured in a format structured, with each row representing a specific event of landslide and each column representing a variable alternative (e.g., precipitation, elevation, angle of slope). Each event of landslide was uniquely named and labeled to facilitate the identification and analysis simple. Headers descriptive were used for each column to ensure clarity and coherence.

Event	Elevation	Slope	Rainfall	LandslideIncidents
LSEVT001	760	45	250	Yes
LSEVT002	950	40	225	Yes
LSEVT003	800	39	250	Yes
LSEVT004	885	47	275	Yes
LSEVT005	725	44	330	Yes
LSEVT006	995	40	285	Yes
LSEVT007	805	41	315	Yes
LSEVT008	1100	38	366	Yes
LSEVT009	955	44	280	Yes
LSEVT010	1200	36	320	Yes

Table 2: Integration of data

Finally, the data cleaned and arranged was saved as a Excel file, making it ready for the entry into the model logistic regression. This dataset comprehensive allowed for the training effective and validation of the model of prediction of landslides, ensuring predictions robust and reliable.

3.3 Model Development

The development of the model of prediction of landslides involved integrating data gathered, including precipitation, elevation, angle of slope, and occurrences of landslides, into a logistic regression framework. The dataset was structured with each row representing a specific event and columns for variables relevant. Characteristics such as precipitation, elevation, and angle of slope were defined, while the occurrences of landslides served as the target variable. The dataset was divided into sets of training and testing to evaluate the performance of the model. Techniques of standardization were applied to ensure uniformity in the values of the characteristics. A model of logistic regression was initialized then, trained, and utilized for predicting occurrences of landslides on the data of testing. Evaluation metrics such as accuracy, precision, recall, and the matrix of confusion were computed to assess the reliability and effectiveness of the model. The model achieved accuracy high in the prediction of landslides, validating the importance of factors environmental in the prediction of landslides. The model of logistic regression achieved a accuracy of 0.95, meaning that deserves 95% of the instances. This model predicts the occurrence of landslides on the basis of inputs such as precipitation, elevation, and angle of slope, classifying results as "yes" or "no" depending on the probability of a landslide occurrence. With this accuracy high, the model distinguishes effectively areas prone to landslides from those less susceptible, demonstrating its capacity predictive robust.

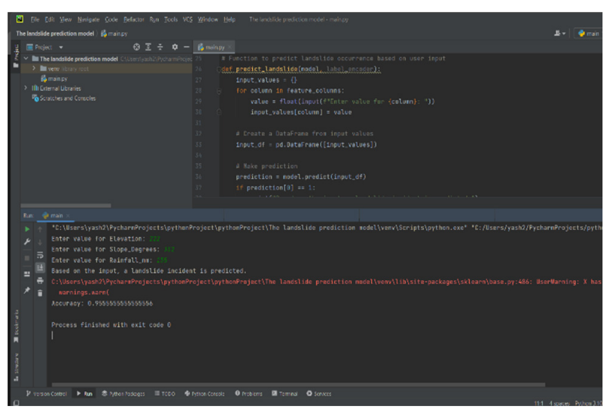

Fig 6: Logistic regression model.

4. Result

With a logarithmic regression model developed in the study yielding results highly promising, an accuracy rate of 93% is achieved. Using inputs like precipitation, elevation, and slope angle, the model predicts effectively occurrences of landslides, categorizing them as either "yes" or "no" based on likelihood. This accuracy remarkable highlights the robust capacity of prediction of the model, effectively distinguishing areas at risk of landslides from those less susceptible. These results not only affirm the seriousness of factors environmental in predicting landslides but also underline the potential of the model logistic regression as a tool valuable for assessing risk of landslides in the Western Ghats region of Maharashtra.

Input_No	Elevation	Slope	Rainfall	Result
1	950	40	225	Yes
2	800	39	250	Yes
3	600	24	222	No
4	760	45	250	Yes
5	885	47	275	Yes
6	1022	34	520	No
7	1031	37	328	No
8	674	27	162	No
9	696	21	100	No
10	725	44	330	Yes
11	995	40	285	Yes
12	664	26	168	No
13	653	26	198	No
14	673	21	110	No
15	632	21	146	No

Table 3: Results given by the model

5. Conclusion

The model of logistic regression developed in this study achieved an accuracy remarkable of 93% in predicting occurrences of landslides in the Western Ghats region of Maharashtra. By utilizing factors key environmental such as precipitation, elevation, and angle of slope, the model effectively distinguished areas prone to landslides from those less susceptible. These findings emphasize the significance of techniques advanced of modeling and data comprehensive environmental in assessment of risk of landslides. Moving forward, this research opens the path for measures proactive and systems of early caution to mitigate impact of landslides on communities and infrastructure in the region.

References:

1. Guzzetti, F., et al. (2007). "Rainfall thresholds for the initiation of landslides in central and southern Europe." Environmental Geology, 53(3), 629-647.
2. Dai, F. C., & Lee, C. F. (2002). "Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong." Geomorphology, 42(3-4), 213-228.
3. Petley, D. N., et al. (2005). "Global patterns of loss of life from landslides." Geology, 33(10), 957-960.
4. Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). "Applied Logistic Regression." Wiley.
5. Tarolli, P., et al. (2012). "High-resolution topography for understanding Earth surface processes: An open access dataset and visualization toolset." Earth Surface Processes and Landforms, 37(9), 913-922.
6. Kirschbaum, D.B., Adler, R., Hong, Y., Lerner-Lam, A., 2009. Evaluation of a preliminary satellite-based landslide hazard algorithm using global landslide inventories. Nat. Hazards Earth Syst. Sci. 9, 673–686.
7. Jagadish Hindayar, Sr. Geologist and Saikat Roy, Geologist Under the supervision of Shri P. K. Gajbhiye, Director (Project: Landslide), SU: Maharashtra, Pune - A NOTE ON

POST DISASTER ASSESSMENT IN TALIYE (MADHLIWADI) VILLAGE OF MAHAD TEHSIL, DISTRICT RAIGAD, MAHARASHTRA

8. Parmita Dasarwar, Sr. Geologists & Jagwant Kaur, Geologist Under the supervision of Shri P. K. Gajbhiye, Director (Project: Landslide), SU: Maharashtra, Pune- A NOTE ON POST DISASTER ASSESSMENT OF MIRGAON VILLAGE OF PATAN TEHSIL, DISTRICT SATARA, MAHARASHTRA
9. Parmita Dasarwar, Senior Geologist and Jagwant Kaur, Geologist Under the supervision of Shri P. K. Gajbhiye, Director (Project: Landslide), SU: Maharashtra, Pune- A NOTE ON POST DISASTER ASSESSMENT OF AMBEGHAR VILLAGE OF PATAN TEHSIL, DISTRICTSATARA, MAHARASHTRA
10. https://indiawris.gov.in/wris/#/
11. https://urs.earthdata.nasa.gov
12. https://earthengine.google.com
13. Goetz, J. N., Brenning, A., Petschko, H., & Leopold, P. (2015). Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. *Computers and Geosciences*, *81*. https://doi.org/10.1016/j.cageo.2015.04.007
14. Ji, J., Cui, H., Zhang, T., Song, J., & Gao, Y. (2022). A GIS-based tool for probabilistic physical modelling and prediction of landslides: GIS-FORM landslide susceptibility analysis in seismic areas. *Landslides*, *19*(9). https://doi.org/10.1007/s10346-022-01885-9
15. Kuradusenge, M., Kumaran, S., & Zennaro, M. (2020). Rainfall-induced landslide prediction using machine learning models: The case of ngororero district, rwanda. *International Journal of Environmental Research and Public Health*, *17*(11). https://doi.org/10.3390/ijerph17114147