International Journal of Geography and Geography Education

Research Article/Araştırma Makalesi

To Cite This Article: Abujayyab, S. K. M. (2024). Enhanced landslide susceptibility prediction with 3D ALOS PALSAR imagery and neural networks: A data-efficient framework. *International Journal of Geography and Geography Education (IGGE),* 51, 115-126.
http://dx.doi.org/10.32003/igge.1380504

ENHANCED LANDSLIDE SUSCEPTIBILITY PREDICTION WITH 3D ALOS PALSAR IMAGERY AND NEURAL NETWORKS: A DATA-EFFICIENT FRAMEWORK

Sohaib K. M. ABUJAYYAB*

Abstract

Landslide susceptibility mapping (LSM) founded on DEM is a growing research field with profound implications for human safety and infrastructure preservation. Many existing methods rely on extensive input data to enhance predictive accuracy. This paper aims to introduce a remote sensing-data-requirement framework for LSM. Our approach exclusively leverages a single ALOS PALSAR image, comprising three key steps: (1) Pre-processing, (2) derivation of explanatory variables, and (3) neural network modeling. To begin, we extracted 22 input variables from the ALOS PALSAR image. These variables played a pivotal role in developing the Neural Network (NN) predictor. The predictor structure consists of 22 variables in the input layer, 150 neurons in the hidden layer, and a single output layer. Our model was trained using 5,829 sample points, and subsequently, it was employed to generate landslide susceptibility (LS) map with 745,810 points. Based on the Overall accuracy metric, the model exhibited impressive performance accuracy, achieving 89.3% training and 82.3% testing accuracies. Additionally, it demonstrated a strong performance of 95.22% during training and 84.7% during testing according to the ROC curve. In conclusion, the implementation of our proposed method underscores its ability to develop remarkable accuracy model with remote sensing-data-requirement. This framework offers valuable insights for future progress in regions with challenging conditions and extensive data coverage. Moreover, it effectively handles data quality inconsistencies and data updating issues.

Keywords: Landslide Susceptibility, Data-Efficient Framework, ALOS PALSAR, Neural Networks, Topographic Attributes

* **Sorumlu Yazar:** International College for Engineering and Management, Muscat, 112, Oman, sohaib@icem.edu.om

Geliş Tarihi/Recieved: 24.08.2023 **Kabul Tarihi /Accepted:** 03.12.2023 **Ç. Yayınlanma Tarihi/Online Published:** 25.01.2024

INTRODUCTION

Landslides, recognized as one of the most significant geohazards worldwide, pose substantial threats to infrastructure, property, natural ecosystems, and human lives, particularly in mountainous regions. Statistics from various sources reveal that landslides account for a substantial portion of global natural disaster-related fatalities, amounting to approximately 17% (Chae et al., 2017). Moreover, the research indicates that the overall economic impact in Europe amounted to around 4.7 billion Euros annually (Haque et al., 2016). Turkey, in particular, faces an ongoing challenge due to the growth of urban areas and continuous deforestation, which worsens landslide risks. Additionally, the regions susceptibility to landslides is further amplified by the increased regional precipitation associated with climate fluctuations (Yilmaz, 2009a). Landslides, mostly caused by prolonged or heavy precipitation and are influenced by various geomorphic parameters, including topography, land cover, forest cover, soil composition, and geology. Consequently, the landslide susceptibility (LS) prediction has come to be a pressing concern for international engineering, geology, and geomorphology communities (Nefeslioglu et al., 2012).

The development of LS maps involves the selection of relevant variables that have contributed to historical landslide occurrences, followed by their utilization to construct predictive models for identifying potential landslide-prone areas in the future (Song et al., 2012). To achieve this, a quantitative framework is necessary. Such frameworks are designed to extract valuable knowledge from past landslide events by leveraging pertinent parameters associated with landslides. In recent years, a range of methodologies has been proposed for developing these predictors, including data mining, statistical analysis, probability assessment, and more advanced techniques such as Neural Networks (NN) in conjunction with remote sensing (Chaudhary et al., 2015; Lee et al., 2020; Z. Liang et al., 2023; Pradhan & Lee, 2010; Sameen et al., 2020). The NN approach, in particular, has gained prominence in recent times and is increasingly applied across various research domains (Lee et al., 2020; Z. Liang et al., 2023; Yilmaz, 2009b).

The development of LS predictors is a task profoundly influenced by the chosen methods and the quality and quantity of input data. When it comes to performing LS predictions on a large scale, the significance of these factors becomes even more noticeable. Gathering comprehensive and geographical, geological information, and historical landslide data is a time-consuming endeavor, particularly when considering extensive areas (Arca et al., 2019).

Moreover, in regions characterized by mountainous terrain, gathering pertinent information on landslide and explanatory parameters from ground poses additional challenges due to the challenging and frequently hazardous circumstances (Nefeslioglu et al., 2012). To overcome these challenges and streamline the process of LS prediction, researchers have increasingly turned to the utilization of 3D satellite imagery. These images, obtained from various satellite missions, have become valuable resources for landslide susceptibility analysis. Prominent examples of such satellite imagery include ALOS PALSAR 12.5m, SRTM 30m, and ASTER 30m. These three-dimensional satellite images provide the opportunity for the automated extraction of a wide array of relevant parameters for LS predictor development, making the process more efficient and accessible.

The aim of this work is to present a remote sensing method for the development of landslide susceptibility models. This method primarily relies on the utilization of 3D ALOS PALSAR images in combination with the prediction process executed through neural networks. The introduction of this remote sensing-requirement framework is highly significant in the field of LS prediction. It addresses the time-consuming and resource-intensive nature of collecting vast amounts of data by capitalizing on the wealth of information offered by 3D satellite imagery. This approach not only mitigates the challenges associated with data acquisition in mountainous regions but also streamlines the entire process of LS prediction. By utilizing 3D ALOS PALSAR images and neural networks, this approach offers a more efficient and accessible way to develop landslide susceptibility predictors, thus contributing to enhanced landslide susceptibility assessment and risk mitigation strategies.

METHODS

Case Study

The research was conducted within the northwestern region of Turkey, specifically in the state of Karabuk. Geographically, this area is situated between approximately 40°58'30"N to 41°07'40"N in latitude and 32°43'10.09"E to 32°53'15.53"E in longitude, as visually represented in Figure 1. The total land area of this region encompasses roughly 153 square kilometers, constituting approximately 3.7% of the entirety of Karabuk state. As of the year 2018, the population of Karabuk state was recorded at 250,269 individuals.

Climatically, Karabuk falls under the classification of a warm and temperate zone. It experiences a notable amount of precipitation throughout the year, with a regular temperature of 9.9 °C. Precipitation levels are substantial, reaching approximately 733 mm annually. The area is situated within a local watershed, predominantly characterized by mountainous terrain. The average elevation in this region stands at 1024 m above the mean sea level (MSL), with the highest peak reaching an elevation of 1590 meters.

This distinctive topography has been witness to several historical landslides, reflecting the inherent susceptibility of the area to such events. The unique characteristics of the case make it highly fit the application of LS model and analyses, given the elevated risk associated with its mountainous landscape.

Data Collection

To generate LS and facilitate NN modeling, it is imperative to identify the pertinent factors that will influence the outcome. In the following section, all the necessary factors derived from a digital elevation model (DEM). DEM data gathered from the Advanced Land Observing Satellite (ALOS). The ALOS platform equipped with PALSAR sensor. PALSAR, as an active microwave instrument utilizing L-band frequency offered the advantage of enabling day-and-night and cloud-free land monitoring.

These PALSAR DEM images cover a temporal extent from May 16, 2006, to April 21, 2011, with a repeat coverage cycle of 46 days. PALSAR images provide worldwide data coverage. Notably, these PALSAR data feature a spatial resolution of 12.5 meters, which is highly conducive for LS mapping. This high spatial resolution plays a pivotal role in enhancing the accuracy of predictors, particularly in scenarios involving small-scale landslide occurrences. Furthermore, the high resolution has the capacity to compensate for data insufficiencies or the exclusion of moderately accurate input data, such as soil maps with limited accuracy. The images consist of 16-bit pixel depth, a significant improvement over past DEM datasets. This richer pixel depth enables a more precise representation of real elevations, especially in complex terrains such as mountainous regions.

It is important to note that all prediction parameters in this research are exclusively extracted from the DEM data. Lastly, The active landslide inventory map acquired from the General Directorate of Mineral Research and Exploration (Maden Tetkik Arama Genel Müdürlüğü MTA) (Çan et al., 2013; Duman et al., 2005), as illustrated in Figure 1.

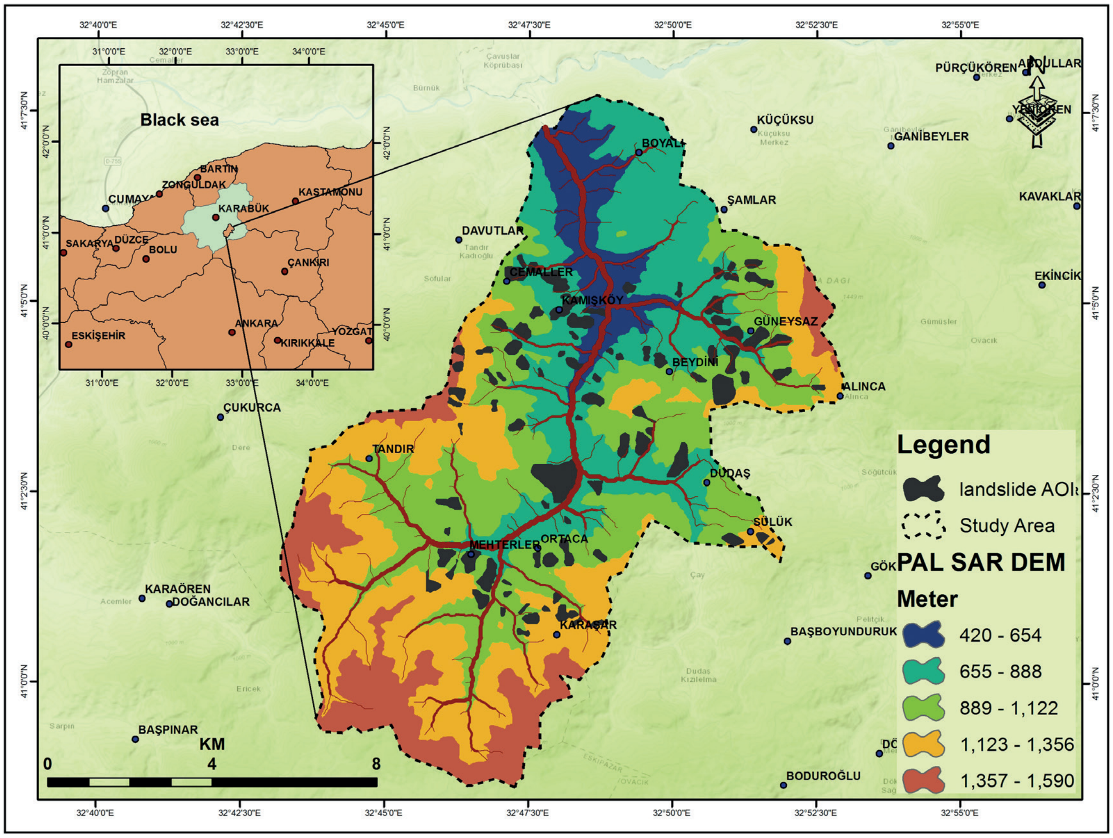


Figure 1. Geographical Region of Interest, PALSAR Digital Elevation Model, and Locations of Past Landslides.

Propose Framework

This section introduces a remote sensing-data constraint approach for LS mapping, as illustrated in Figure 2. The method is primarily based on DEM data and consists of three key stages: (A) Pre-processing. DEM data gathered and processed within the ESRI software. Subsequently, the data is confined to the study border. To address potential holes in the data, the Fill tool is employed, ensuring a more complete input image. (B) Deriving explanatory parameters from the DEM data. Exactly, 22 variables are derived from the DEM data, encompassing surface topography, geomorphological texture, temperature, geomorphometry, and moisture factors (as outlined in Table 1). (C) The final stage revolves around the utilization of NN for LS model training. The trained network is then applied to generate LS maps. This stage requires the construction of an NN training dataset, which is achieved through sampling. Samples are extracted from the derived 22 factors and the target image, which represents the locations of past landslides that obtained from the General Directorate of Mineral Research and Exploration (Maden Tetkik Arama Genel Müdürlüğü MTA).

Table 1. provides a catalog of explanatory variables.

N.	Group of variables	Variables
1	Geomorphometry	Flow Direction, Flow Accumulation, Stream Order
2	Topographic surfaces and Geomorphology texture	Hillshade, Profile Curvature, Roughness, Slope, General Curvature, Landform, Plan Curvature, Dissection, Slope Position, Surface Area Ratio, Aspect, Surface Relief Ratio, Elevation
3	Temperature and moisture	Compound Topographic Index, Site Exposure Index, 2nd Derivative Slope Integrated Moisture Index, Slope/Aspect Transformation, Heat Load Index,

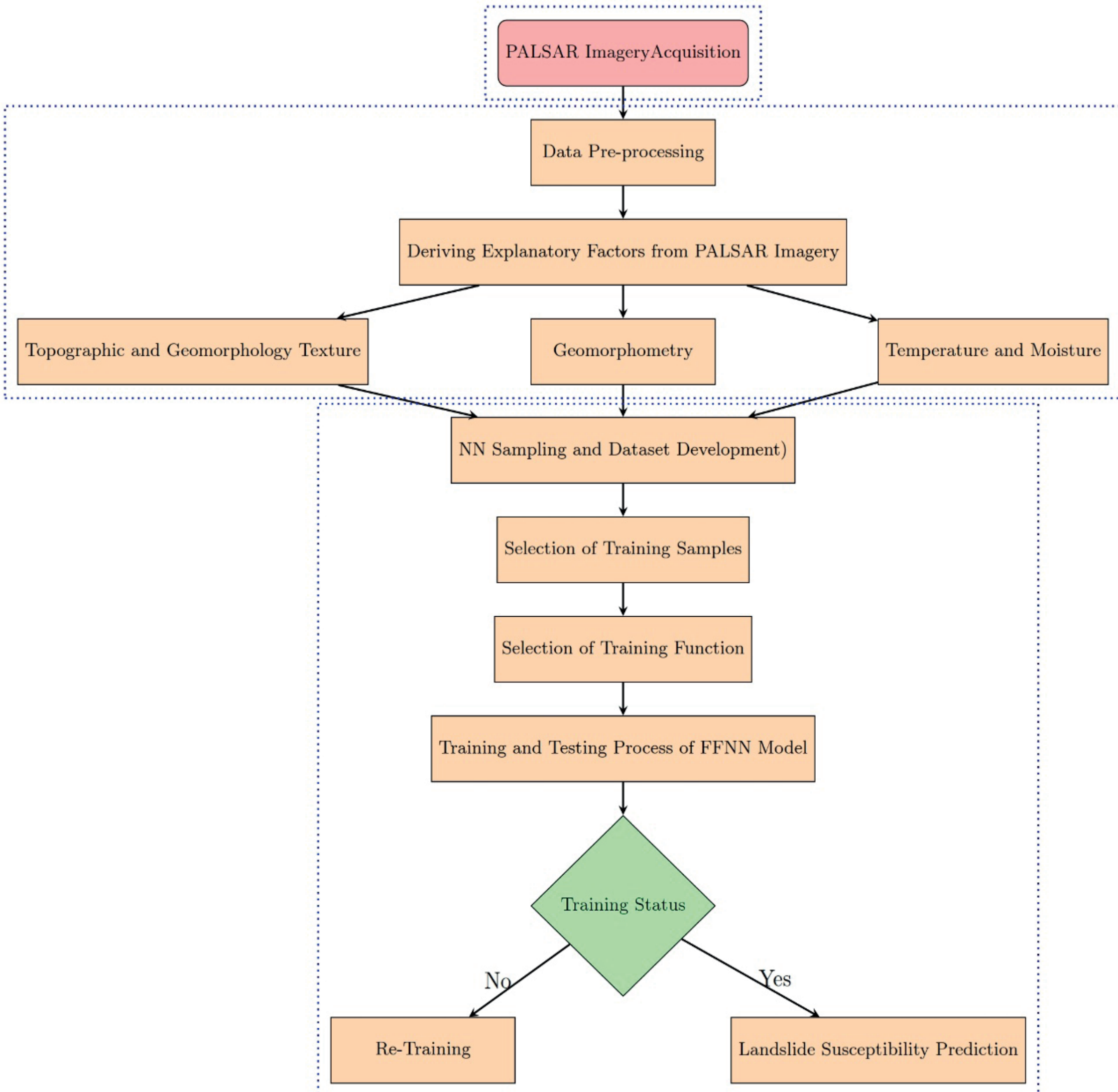
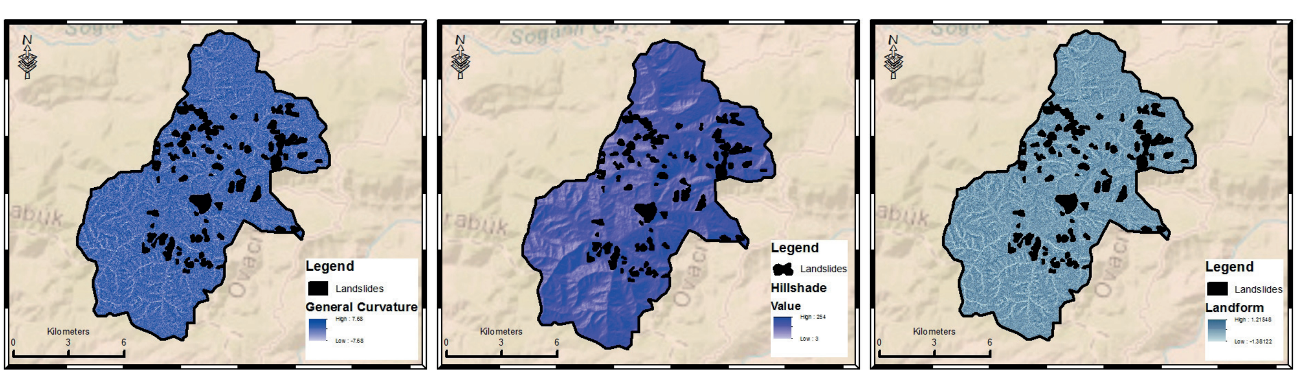


Figure 2. Remote sensing-based method for LS prediction.



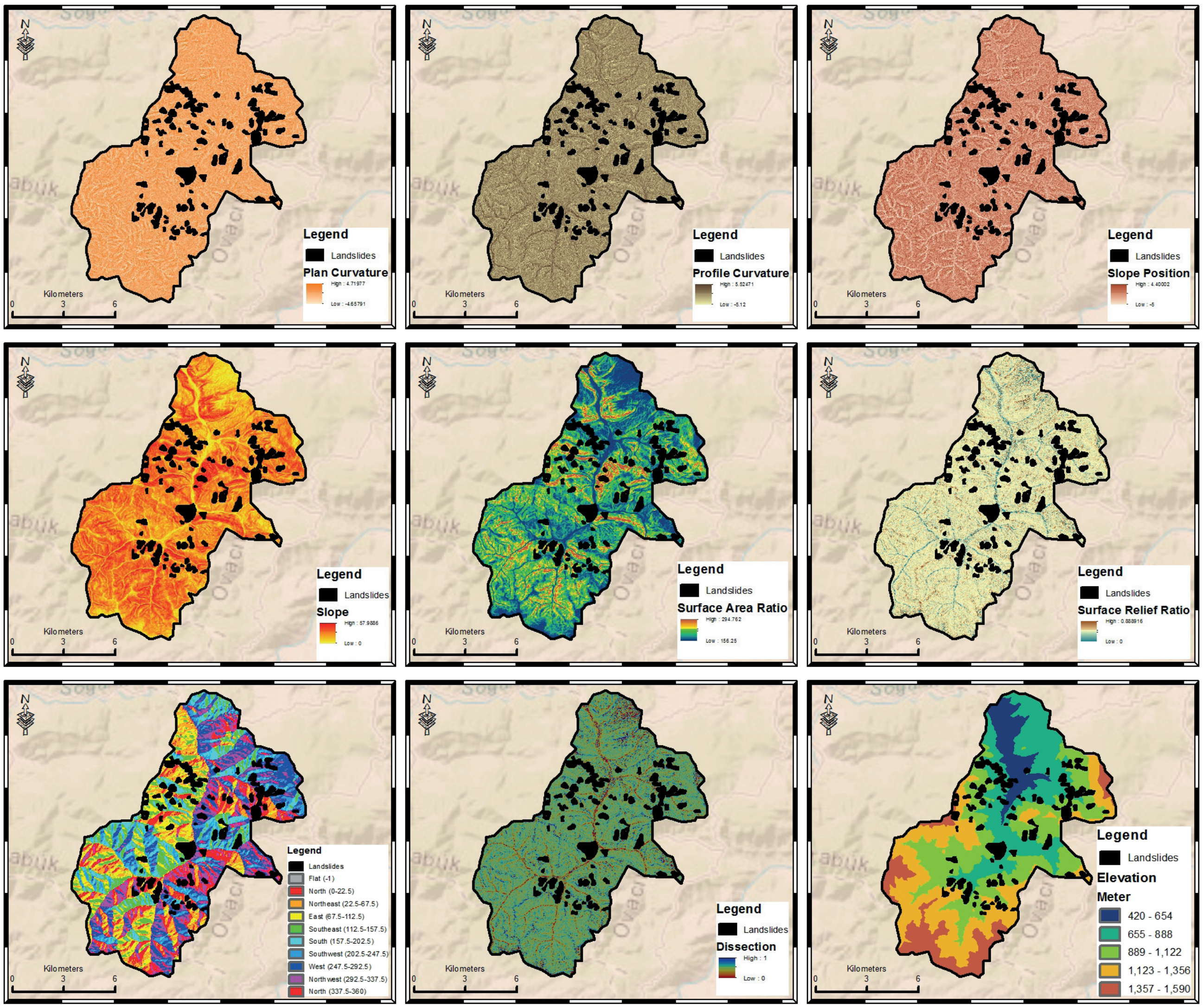


Figure 3. Landslide conditioning factor maps: Landform, Aspect, Surface Area Ratio, Plan Curvature, Slope, General Curvature, Slope Position, Roughness, Surface Relief Ratio, Hillshade, Dissection, Profile Curvature, and Elevation,

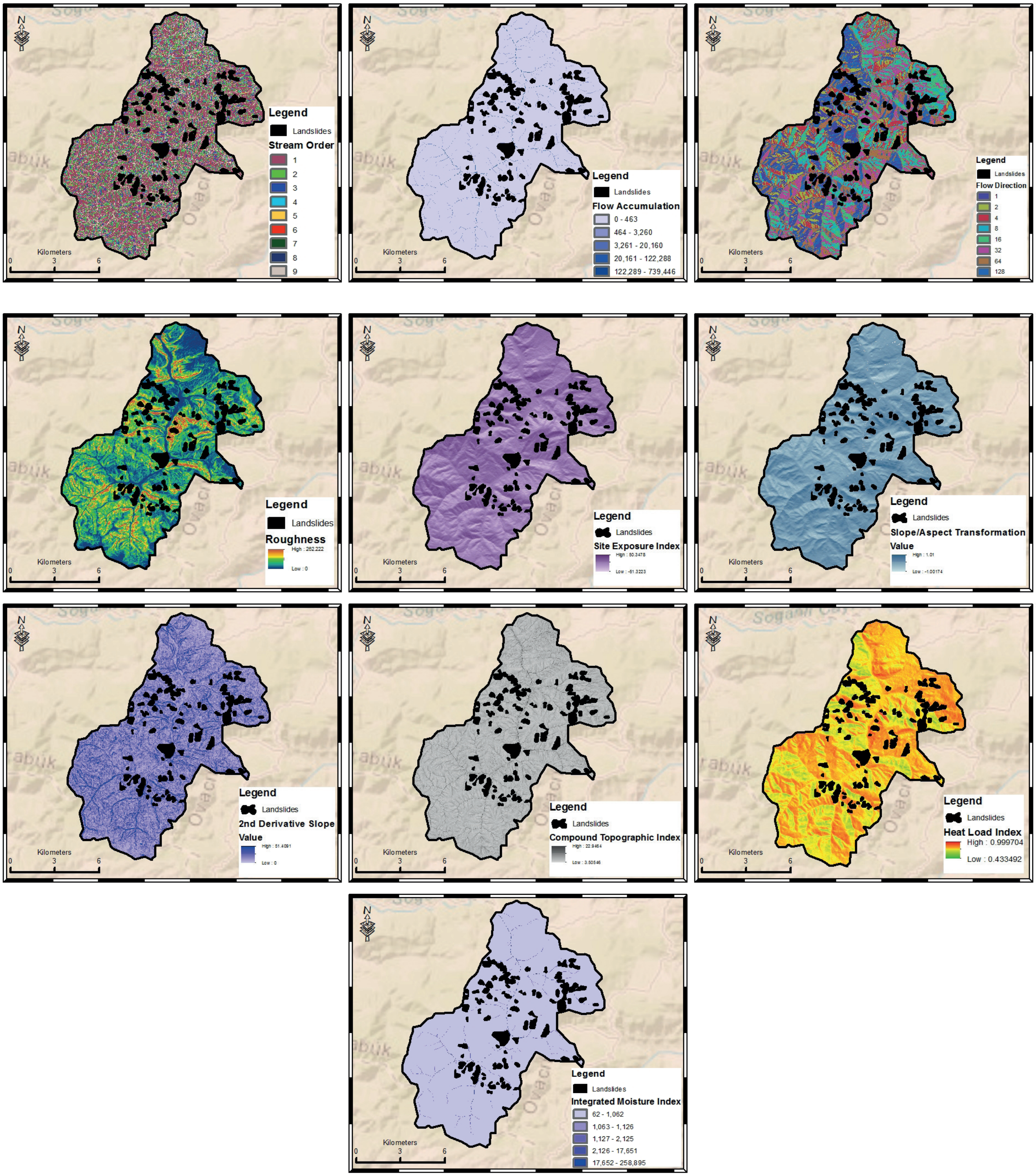


Figure 4. Landslide conditioning factor maps: Stream Order, Site Exposure Index, Flow Direction, Integrated Moisture Index, 2nd Derivative Slope, Slope/Aspect Transformation, Flow Accumulation, Compound Topographic Index, and Heat Load Index,

Accuracy Metrics

The primary focus of this study's NN predictor is the assessment of model accuracy, which determined by evaluating the differences among the predicted and actual target. The accuracy explained through the Overall accuracy (OA), as described by the following equations (1) (D. Liang et al., 2015).

$$OA = 1 - \left(\frac{\begin{matrix} Number\ of\ samples \\ correctly\ classified \\ TP + TN \end{matrix}}{\begin{matrix} Total\ number\ of\ samples \\ TP + FP + FN + TN \end{matrix}} \right) * 100 \tag{1}$$

Where;

OA = Overall Accuracy,

FP = false positive,

TP = true positive,

FN = false negative

TN = true negative,

LS Model Development

In this section, the development of the Landslide predictor has described based on a Neural Network (NN). The NN used 3-layered NN model. The 5829 training points randomly generated to represent the historical landslides and non-landslide. Subsequently, the sample dataset was migrated to Matlab software for the construction and training of the NN model.

The NN model structure developed using 150 neurons and was determined by conducting some experiments, varying the number of neurons from ten to two hundred. Various learning methods tested through an optimizer, and the most suitable algorithm employed for the model. As a result, the model structure defined as 22 factors for the input layer, 150 neurons for the hidden layer as demonstrated in Figure 5.

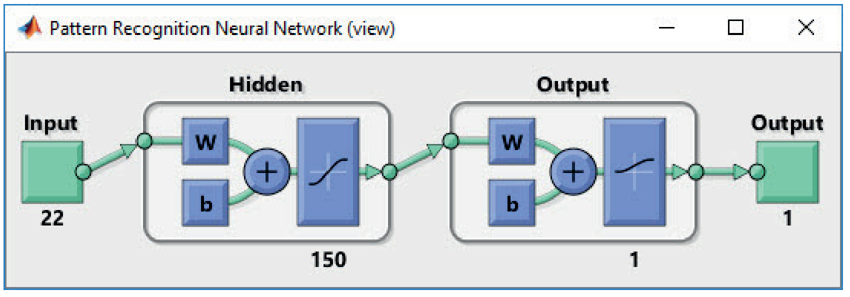


Figure 5. illustrates the architecture of the NN predictor for landslide susceptibility (LS)

RESULTS

The study involved in the deployment of a Neural Network (NN) predictor to generate landslide susceptibility (LS) within the Karabuk region. The process commenced with the ingestion of 5829 index characteristic values extracted from the training dataset into the NN predictor. Following network training, pertinent output factors were derived. These output factors played a pivotal role in the accurate prediction of LS within. It is important to note that the NN predictor was accurately developed based on tabular data originating from the DEM data only.

To streamline the LS prediction process across the entire study area, tabular data consisting of 745,810 pixels was prepared and subsequently transposed into the Matlab environment. This dataset was then fed into the NN model, yielding susceptibility map as depicted in Figure 6. Concurrently, susceptibility predictions were tested and contrasted against historical landslide data. Figure 6 illustrates the anticipated levels of LS within the study area, thoughtfully categorized into five distinct groups: The classification of susceptibility levels reveals distinct proportions across the landscape. The category labeled as “very high susceptibility” encompasses an area of 3.53 km², constituting 3.03% of the total expanse. Meanwhile, the “high susceptibility” classification spans 31.72 km², accounting for 27.22% of the overall area. In contrast, the “moderate susceptibility” category covers 27.70 km², representing 23.77% of the total landscape. Additionally, the “low susceptibility” designation occupies 22.39 km², equivalent to 19.22% of the entire area. Lastly, the “very low susceptibility” category spreads over 31.16 km², comprising 26.74% of the total landscape. These diverse susceptibility levels provide a nuanced understanding of the distribution of susceptibility across the studied region. The levels of landslide susceptibility were classified using equal interval methods within ArcGIS, dividing the values equally based on the specified number of categories. Furthermore, Figure 6 clarifies that the very high landslide susceptibility category predominantly prevails within the altitudinal range of 423 to 1590 meters. This particular category exhibits a markedly higher vulnerability to landslides in comparison to the other categories. Our findings reveal a robust correlation between the “very high susceptibility” category and eleven variables. These relevant variables encompass heat load index, surface relief ratio, roughness, landform, slope/aspect transformation, dissection, general curvature, aspect, slope position, plan curvature, and profile curvature. Our analysis underscores the varying degrees of influence wielded by these factors, with the “very high” LS category being particularly responsive to the three curvature factors.

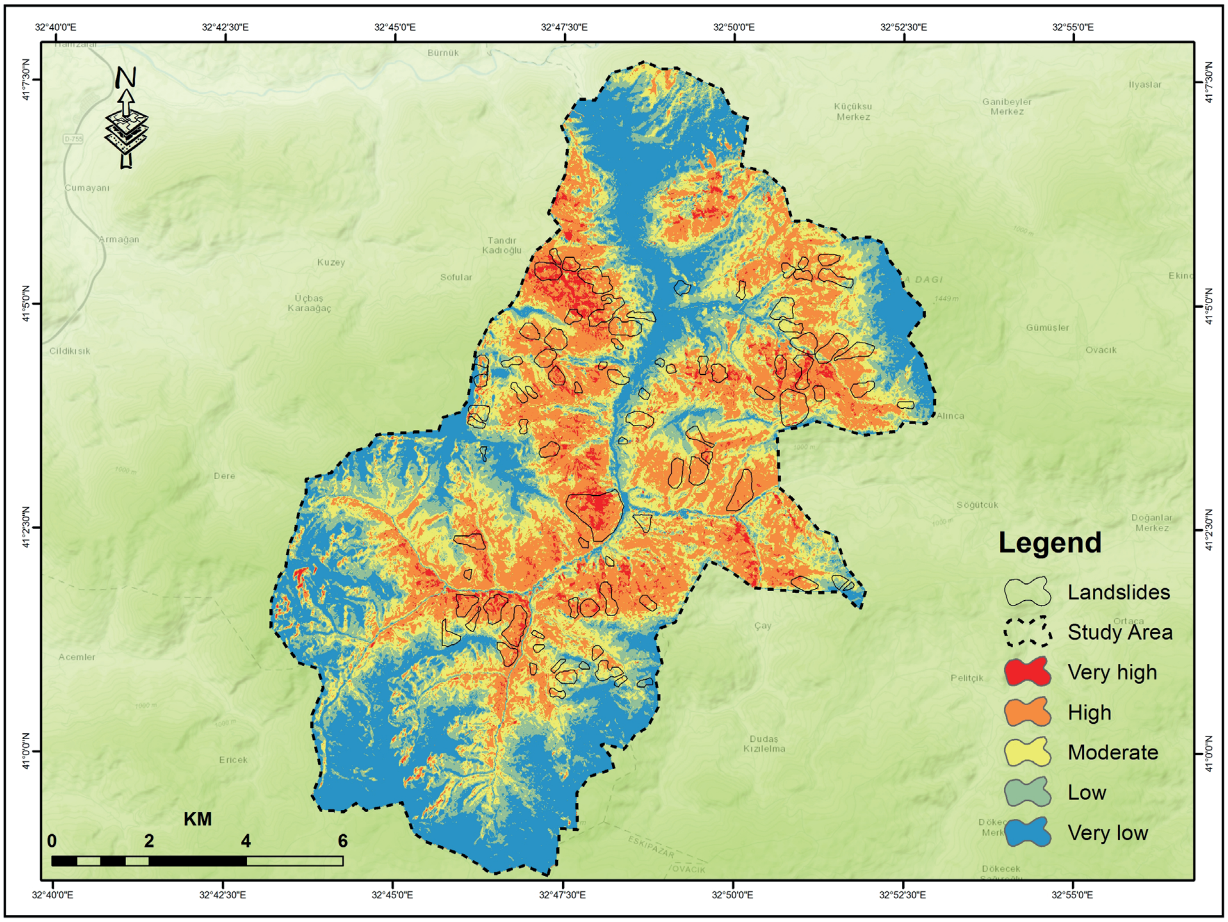


Figure 6. illustrates the landslide susceptibility assessment utilizing the NN predictor.

Spatial analysis conducted during this study revealed the heightened relevance of areas situated in close proximity to rivers and streams to landslide. Additionally, the heat load index emerged as a noteworthy variable influencing the landslides incidence. To measure the validity of the NN model in LS prediction, confusion matrices (CM) presented in Table 2. Notably, the training dataset boasted an accuracy rate of 89.3%, while the testing dataset secured an accuracy rate of 82.3%. Meanwhile, Figure 7 depicts the performance accuracy of the NN model via ROC curves, accordingly calculating the area under the curve (AUC). This metric exhibited an AUC of 95.22% for the training dataset and 84.7% for the testing dataset. Significantly, these consequences serve to expose any concerns related to overfitting, as the training dataset's performance accuracy exceeded that of the testing dataset. The rate of false positives in the training dataset was found to be minimal, and in the testing dataset, it remained within acceptable bounds.

Our model demonstrated a high accuracy of 89.3%, surpassing various models in the literature. Comparison with existing studies underscores our model's excellent performance across all evaluation aspects, particularly in terms of accuracy. Notably, Liang et al. (2023) achieved an 85.28% accuracy with a CNN model in a study on landslide susceptibility in Yadong Country, Tibet (Z. Liang et al., 2023)especially in the Himalayan areas. Landslide susceptibility mapping (LSM. Adel et al. (2023) reported a 76.6% reliability rate using a fuzzy gamma operator model in Northern Tunisia (Adel et al., 2023)a case study of Mogods and Hedil (Northern Tunisia. Zhu, Liu, and Yu (2023) achieved an impressive total accuracy of 92.38% with the SGCN-LSTM model in landslide susceptibility prediction (Zhu et al., 2023). Fu et al. (2023) utilized an integrative sampling approach, achieving an AUC of 0.92 in landslide susceptibility mapping in Songyang County, China (Fu et al., 2023). The accuracy percentages in these studies range from 76.6% to 92.38%, and our model's accuracy of 89.3% falls within this range. While some studies achieved slightly higher accuracy, it is important to note that their methods often necessitated extensive data collection, making them inefficient for large-scale areas.

Table 2. Illustrates the accuracy of the NN predictor's using the information derived from CM

	Training		Testing	
= true positive	5629	31.1%	819	25.6%
= true negative	10554	58.2%	1812	56.7%
= false positive	477	2.6%	162	5.1%
= false negative	1462	8.1%	405	12.75
Overall Accuracy		89.3%		82.3%

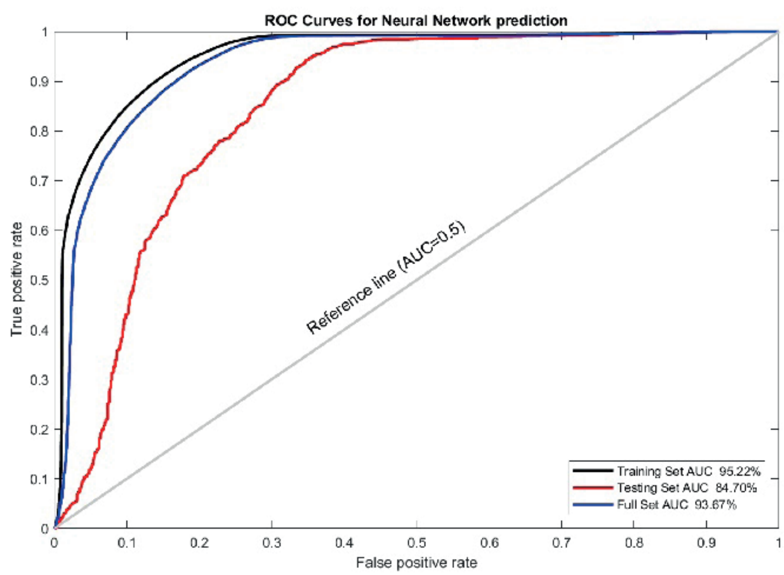


Figure 7. illustrates the NN predictor's performance accuracy using ROC curves. ROC curves provide a depiction the model accuracy through AUC.

In summation, the findings of our study underscore the successful application of neural networks in predicting landslide susceptibility, validating the robustness of the NN model. The remote sensing-data-requirement framework presented herein yields promising results in the area of LS prediction. The LS map holds substantial value for land administration and planning departments, aiding future development initiatives. Such maps, can effectively guide construction and infrastructure projects towards areas of very low susceptibility while highlighting the need for rigorous geotechnical and geological engineering considerations in regions of very high susceptibility.

CONCLUSION

In conclusion, our research successfully deployed a Neural Network (NN) predictor to assess landslide susceptibility (LS) in the Karabuk region. This process began with the input of 5829 training dataset values into the NN predictor, which was then trained to derive key output factors for precise LS prediction. Efforts were made to streamline the LS prediction process for the entire study area, resulting in the preparation and transformation of tabular data comprising 745,810 pixels into the Matlab environment. This dataset was subsequently fed into the NN predictor, producing susceptibility values categorized into five distinct levels: very high susceptibility, high susceptibility, moderate susceptibility, low susceptibility, and very low susceptibility.

Our findings indicated that regions at altitudes between 423 and 1590 meters are significantly more vulnerable to landslides, with the very high susceptibility category showcasing a robust correlation with 11 key input factors, including dissection, surface relief ratio, roughness, heat load index, landform, and others. Furthermore, our analysis revealed that areas in proximity to rivers and streams are exceedingly pertinent to landslide occurrences, with the heat load index being a notable influencing parameter.

However, our approach has some limitations. It requires a considerable number of neurons for high accuracy, leading to data processing costs. The need for numerous explanatory factors further contributes to data processing expenses, especially in large-scale analyses. Yet, the periodic nature of LS prediction somewhat mitigates these challenges.

For future research, we recommend integrating LIDAR (Light Detection and Ranging) imagery for the second stage of prediction, with a specific focus on very high susceptibility areas. This will refine our understanding of landslide susceptibility and contribute to more informed decision-making in geotechnical and geological engineering. This research showcases the potential of remote sensing and neural networks in landslide susceptibility prediction, offering valuable insights for future studies and practical applications in the field.

REFERENCES

Adel, K., Katlane, R., Haddad, R., & Rabia, M. C. (2023). Landslide susceptibility mapping by Frequency Ratio and Fuzzy logic approach: A case study of Mogods and Hedil (Northern Tunisia).

Arca, D., Keskin Citiroglu, H., & Tasoglu, I. K. (2019). A comparison of GIS-based landslide susceptibility assessment of the Satuk village (Yenice, NW Turkey) by frequency ratio and multi-criteria decision methods. *Environmental Earth Sciences*, 78(81), 4–13. https://doi.org/10.1007/s12665.019.8094-6

Çan, T., Duman, T. Y., Olgun, Ş., Çörekçioğlu, Ş., Karakaya-Gülmez, F., Elmacı, H., Hamzaçebi, S., & Emre, Ö. (2013). Türkiye heyelan veri tabanı. *TMMOB Coğrafi Bilgi Sistemleri Kongresi*.

Chae, B.-G., Park, H.-J., Catani, F., Simoni, A., & Berti, M. (2017). Landslide prediction, monitoring and early warning: A concise review of state-of-the-art. *Geosciences Journal*, 21(6), 1033–1070. https://doi.org/10.1007/s12303.017.0034-4

Chaudhary, P., Chhetri, S. K., Joshi, K. M., Shrestha, B. M., & Kayastha, P. (2015). Application of an Analytic Hierarchy Process (AHP) in the GIS interface for suitable fire site selection: A case study from Kathmandu Metropolitan City, Nepal. *Socio-Economic Planning Sciences*. https://doi.org/10.1016/j.seps.2015.10.001

Duman, T. Y., Emre, Ö., Çan, T., Nefeslioğlu, H. A., Keçer, M., Doğan, A., Durmaz, S., & Ateş, Ş. (2005). Türkiye heyelan envanteri haritası-1: 500.000 ölçekli Zonguldak Paftası. *MTA* Özel *Yayınlar Serisi*-4, Ankara.

Fu, Z., Wang, F., Dou, J., Nam, K., & Ma, H. (2023). Enhanced absence sampling technique for data-driven landslide susceptibility mapping: A case study in Songyang County, China. *Remote Sensing*, 15(13). https://doi.org/10.3390/rs15133345

Haque, U., Blum, P., da Silva, P. F., Andersen, P., Pilz, J., Chalov, S. R., Malet, J.-P., Auflič, M. J., Andres, N., Poyiadji, E., Lamas, P. C., Zhang, W., Peshevski, I., Pétursson, H. G., Kurt, T., Dobrev, N., García-Davalillo, J. C., Halkia, M., Ferri, S., … Keellings, D. (2016). Fatal landslides in Europe. *Landslides*, 13(6), 1545–1554. https://doi.org/10.1007/s10346.016.0689-3

Lee, D.-H., Kim, Y.-T., & Lee, S.-R. (2020). Shallow landslide susceptibility models based on artificial neural networks considering the factor selection method and various non-linear activation functions. *Remote Sensing*, 12(7). https://doi.org/10.3390/rs12071194

Liang, D., Tsai, C.-F., & Wu, H.-T. (2015). The effect of feature selection on financial distress prediction. *Knowledge-Based Systems*, 73, 289–297. https://doi.org/10.1016/j.knosys.2014.10.010

Liang, Z., Peng, W., Liu, W., Huang, H., Huang, J., Lou, K., Liu, G., & Jiang, K. (2023). Exploration and comparison of the effect of conventional and advanced modeling algorithms on landslide susceptibility prediction: A case study from Yadong Country, Tibet. *Applied Sciences*, 13(12). https://doi.org/10.3390/app13127276

Nefeslioglu, H. A., San, B. T., Gokceoglu, C., & Duman, T. Y. (2012). An assessment on the use of Terra ASTER L3A data in landslide susceptibility mapping. *International Journal of Applied Earth Observation and Geoinformation*, 14(1), 40–60. https://doi.org/10.1016/j.jag.2011.08.005

Pradhan, B., & Lee, S. (2010). Landslide susceptibility assessment and factor effect analysis: Backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environmental Modelling & Software*, 25(6), 747–759. http://dx.doi.org/10.1016/j.envsoft.2009.10.016

Sameen, M. I., Pradhan, B., & Lee, S. (2020). Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. Catena, 186 (September), 104249. https://doi.org/10.1016/j.catena.2019.104249

Song, K.-Y., Oh, H.-J., Choi, J., Park, I., Lee, C., & Lee, S. (2012). Prediction of landslides using ASTER imagery and data mining models. *Advances in Space Research*, 49(5), 978–993. https://doi.org/10.1016/j.asr.2011.11.035

Yilmaz, I. (2009a). A case study from Koyulhisar (Sivas-Turkey) for landslide susceptibility mapping by artificial neural networks. *Bulletin of Engineering Geology and the Environment*, 68(3), 297–306. https://doi.org/10.1007/s10064.009.0185-2

Yilmaz, I. (2009b). Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat-Turkey). *Computers and Geosciences*, 35(6), 1125–1138. https://doi.org/10.1016/j.cageo.2008.08.007

Zhu, L., Liu, L., & Yu, C. (2023). Landslide Susceptibility Prediction Modeling Based on Self-Screening Deep Learning Model. arXiv Preprint arXiv:2304.06054.