Manuscript number: nhess-2023-73

My co-authors and I would like to express our gratitude to the reviewers for their constructive feedback and suggestions for strengthening our research. The changes we have made to the attached file in response to such feedback and suggestions have been highlighted in blue to facilitate their identification. I would also like to offer my apologies for the length of time it took us to prepare this response. We also record our deep appreciation for the efficient handling of the manuscript.

Response to Reviewer#1

General remarks: The paper is interesting for two main reasons. Firstly, the work represents a step towards the prediction of the intensity of landslide events, and second because analyze the pattern among the landslide induced debris and causal influencing factors.

I would suggest to improve the definition of the landslide types, and to clarify the differences among the geomorphological processes of table 1.

We would like to thank the reviewer for taking the time to examine the preprint and provide insightful comments and suggestions. We appreciate your positive assessment of the paper's significance and contribution to the fields of landslide prediction and analysis of causal influencing factors. Your suggestions will undoubtedly aid in enhancing the overall quality of the manuscript.

Comment 1: Line no. 32- Who is referring? Landslides, debris produced by the landslides?

Response: Thank you for your insightful review. The phrase 'To mitigate their effects from different landslides' refers to taking steps to reduce the adverse consequences caused by rainfall-induced landslides and the resulting debris. These impacts encompass severe property damage, financial losses, and loss of human lives resulting from such landslides. Consequently, the sentence has been revised in the manuscript as follows:

‘The resultant debris from these landslides leads to severe property damage, substantial financial repercussions, and loss of human lives. To address these consequences, various deterministic and probabilistic models of landslide susceptibility mapping are frequently used.’

Comment 2: Susceptibility map are useful for planning or adaptation actions not properly for mitigation actions.

Response: Thank you for your comment. You're absolutely right that susceptibility maps are generally used for planning and adaptation rather than direct mitigation actions. These maps help identify regions more likely to have certain natural disasters, which is vital for informed decision-making and preparedness. As susceptibility maps don't directly mitigate the impact of events, they play a critical role in guiding mitigation efforts by highlighting areas requiring heightened attention and specific mitigation strategies. Accordingly, the sentence has been revised in the manuscript as follows:

'Rainfall-induced landslides frequently occur in the mountainous region of the Korean peninsula due to heavy summer rainfall. The resultant debris from these landslides leads to severe property damage, substantial financial repercussions, and loss of human lives. To address these consequences, various deterministic and probabilistic models of landslide susceptibility mapping are frequently used. However, the existing landslide susceptibility models identify regions with potential landslides but do not quantify their size.'

Comment 3: Line no. 33- which method?

Response: Thank you for your insightful review. The phrase 'these methods' in line no 33 (preprint version) refers to the existing landslide susceptibility methods. Accordingly, the sentence has been revised in the manuscript as follows:

'However, the existing landslide susceptibility models identify regions with potential landslides but do not quantify their size.'

Additionally, we thoroughly reviewed the manuscript, and such mistakes have been corrected in the revised manuscript.

Comment 4: Line no 37- What is it?

Response: Thank you for your insightful comment. We carefully read the entire manuscript, and such typos and linguistic errors have been corrected in the revised manuscript. The revised version of the sentence is as follows:

'In addition, the exploratory data analysis of the RFIL debris' dataset has been conducted to examine patterns and relationships between RFIL debris volume, triggering (rainfall) and influencing factors.'

Comment 5: Line no 88: 'Further, most of studies were performed on a small scale and only predicted the occurrence, not the size.' Please see Lombardo et al., 2018; Lombardo et al., 2023 to extend the topic.

Response: Thank you for your comment. As suggested, the extensive literature review has been performed, accordingly the topic has been extended in the introduction as follows:

'Globally, numerous researchers have attempted to predict the landslide magnitude through different statistical approaches (Lombardo et al., 2020). For example, Dai and Lee (2001) analyzed the relationship between landslide volume, cumulative frequency, and the connection between rainfall and landslide occurrence. Malamud et al. (2004) proposed frequency and size distribution for landslides to quantify the magnitude of landslide events. Shirzadi et al. (2017) compared popular statistical and machine-learning methods for simulating the volume of landslides. Lombardo et al. (2018) introduced the concept of estimating landslide intensity to complement susceptibility measures. They used the Poisson distribution for spatial estimates of the landslide intensity within terrain units. Further, Lombardo et al. (2021) explored advanced techniques, leveraging Bayesian versions of a Generalized Additive Model and Log-Gaussian model to estimate landslide susceptibility and intensity. The existing literature lacks a widely applied machine-learning model capable of capturing and predicting landslide sizes (volume). To address this gap and assess the potential of machine learning methods for predicting landslide volume in South Korea, we used the proportional odds logistic regression (POLR), random forest (RF), support vector machine (SVM), and extreme gradient boosting (EGB) methods to evaluate the relationship between various influencing & triggering factors and RFIL debris volume. Consequently, the present study aims to predict the RFIL debris volume based on the triggering and influencing factors. The quantification of debris volume may be useful in land management by highlighting regions prone to higher RFIL debris to know whether economic activities may be carried out in the given region, so that those activities may not be vulnerable to extreme landslide hazards.'

Comment 6: Rainfall is the trigger not an influencing factor (Figure 2)?

Response: Thank you for your insightful observation. Figure 2 has been updated in the revised manuscript as follows:

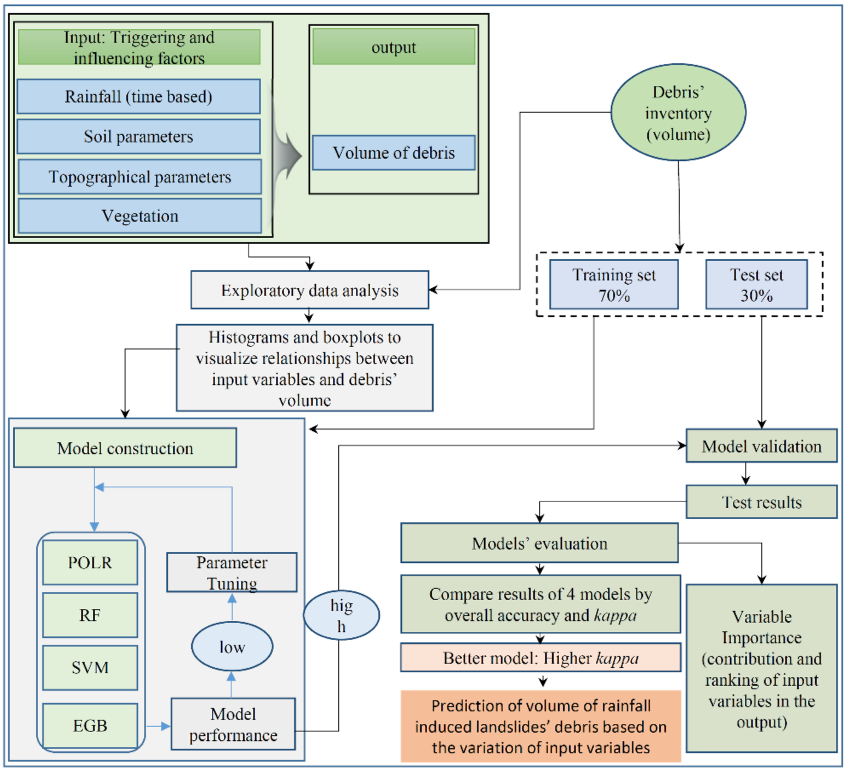

Figure 2. Modeling workflow process for the prediction of RFIL debris volume.

Comment 7: "There are different types of landslides; which are classified based on the cause or shape after occurrence (Causes 2001). Landslides may result from liquefaction, earthquakes, intense surface water flow due to precipitation, underground water, ice melting, human activities, tectonic movements etc (Alexander 1992; Causes, 2001; van der Beek, 2021; McColl 2022)"

Is relevant to the purpose of your work?

Response: Thank you for your comment. We greatly appreciate your input. We have carefully reviewed and revised the manuscript based on your suggestion to streamline the content. Unnecessary text has been removed, allowing us to better focus on the main aspects of our research.

Comment 8: The only landslide type is falling rock; the others are erosion or morphological features (Table 1).

Response: Thank you for your insightful comment. We appreciate your observation that the different landslides mentioned in Table 1 are not all technically distinct but rather variations in erosion processes or morphological features. The landslide classes mentioned in Table 1 (preprint version) describe different geomorphological processes that lead to various landslide forms. For instance, valley erosion occurs when material moves due to a valley's erosion, and falling rocks describe the tumbling of loose rocks down slopes (Causes, 2001). While mixed or complex landslides involve multiple processes, slope failures result from a slope's collapse (Wang et al., 2016). Landslides resulting from scour are a consequence of erosion, while curved wedge-shaped slides display a distinctly curved structure (Ritchie, 1958). Circular arc-shaped landslides take on a semi-circular shape, and plane-shaped ones occur on inclined planes (Causes, 2001). Each type emerges from specific geological actions, defining their distinct characteristics and appearances.

However, in the present study, we considered 455 landslide inventory data based on the magnitude of the landslides: below 500m^3, between 500-2000m^3, 2000-5000m^3, and above 5000m^3 (Fig. 1b). Consequently, we analyzed the relationship between independent variables and debris size. As a result, Table 1 has been updated in the revised manuscript to provide a detailed summary of data features rather than focusing solely on different geomorphological processes. Additionally, a summary of continuous variables is provided in Table 2, including minimum, mean, median, maximum, standard deviation, and associated units for each considered feature.

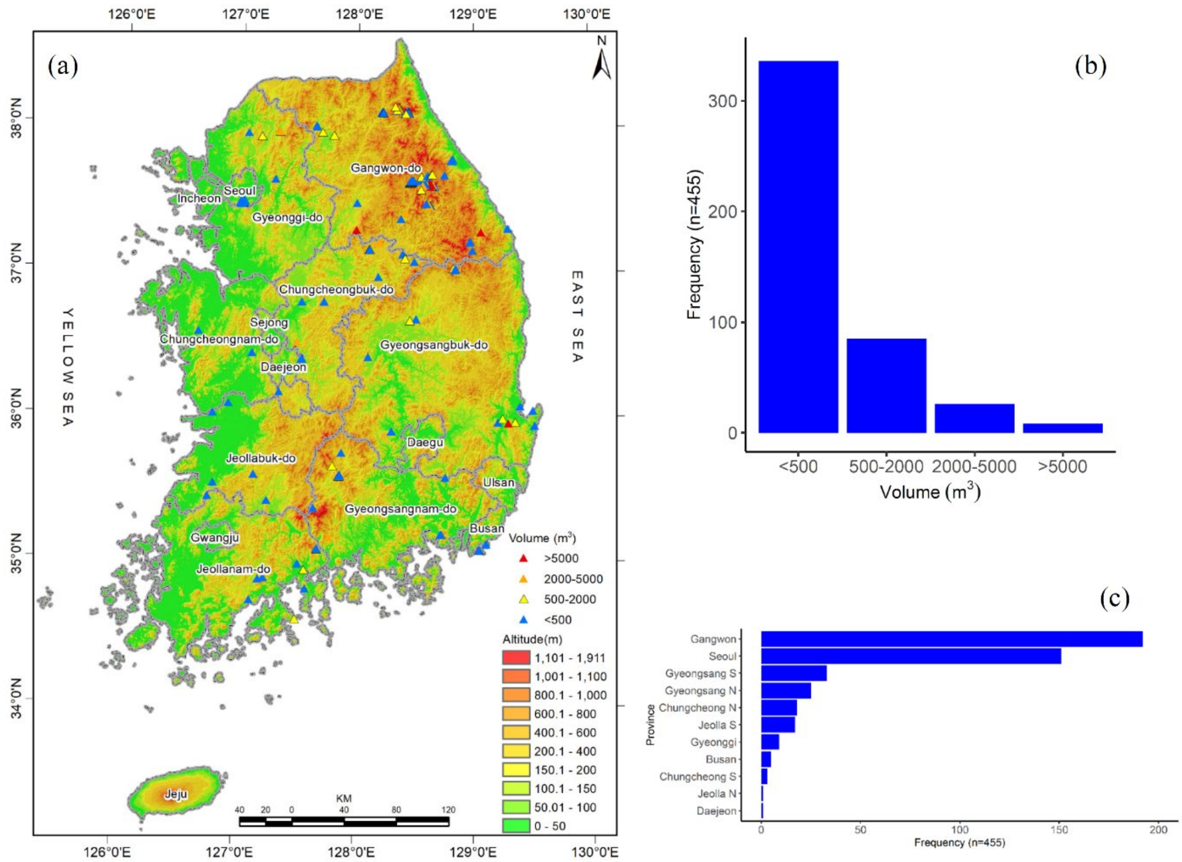

Figure 1. (a) The distribution of RFIL in South Korea, (b) Histogram of RFIL debris volume, and (c) Province-wise RFIL debris frequency distribution (Data source: elevation data acquired from NGII, 2018).

Table 1: The detailed description of continuous and categorical variables.

Causative factors	Variables	Type	Importance
Rainfall	Continuous hourly rainfall	Continuous	Triggering factor; effect on soil moisture (Ngo et al., 2021)
	Three hours rainfall		
	Six hours rainfall		
	Nine hours rain		
	Twelve hours rain		
	One day rain		
	Three days rain		
	Seven days rain		
	Two weeks rain		
	Three weeks rain		
	Four weeks rain		
Soil	Soil depth (cm)	Categorical 1) <20, 2) 20-50, 3) 50-100	Permeability, infiltration, surface runoff, and soil strength affect slope stability (Meena et al., 2022)
	Soil type	Categorical 1) Sandy loam, 2) Silt loam, 3) Lithosol, 4) Clay	
Forest features	Leafage	Categorical 1) Broad-leaved, 2) Mixed, 3) Pines	Effect on slope stability i.e., vegetation roots cause more stability (Ngo et al., 2021)
	Size of wood	Categorical 1) No tree, 2) Small, 3) Average, 4) Large	
	Age of tree	Categorical 1) No tree, 2) 1-5, 3) 5-15, 4) 15-25, 5) 25-35, 6) 35-45, 7) 45-60, 8) >60	
	Forest density	Categorical 1) No tree, 2) Low, 3) Medium, 4) High	It reflects the inhibitory effect of landslide occurrence (Huang et al., 2020)
	Forest Fire history	Categorical 1) No, 2) Yes	Effect on soil erosion
Topographical and geomorphological parameters	Slope (degree)	Continuous	Effect on infiltration process, shear stress, and gravity. Landforms with a steep slope and high slope length are usually more susceptible to collapse (Pham et al., 2018)
	Slope length (m)	Continuous	
	Slope aspect	Categorical (8 directions) North, Northeast, south, East,	Effect on rainfall, soil moisture, and vegetation cover (Dahal et al., 2008)

Causative factors	Variables	Type	Importance
		Southeast, South, Southwest, West, Northwest	
	Altitude (m)	Continuous	Effect on rainfall, vegetation cover, and soil depth (Raja et al., 2017)
	Drainage	Categorical 1) Bad, 2) Good, 3) Very good	Effect on water flow, saturation, soil moisture, and valley landslides (Shahabi and Hashim, 2015)

Table 2. Summary statistics of continuous variables.

Variables	**Observation**	**Min**	**Mean**	**Median**	**Max**	**SD**	**unit**
Maximum hourly rainfall	455	0	48.2	48	78.5	20.262	mm
Continuous hourly rainfall	455	0	285.341	327	549.5	106.279	
Three hours rainfall	455	0	87.716	79.5	171	60.166	
Six hours rainfall	455	0	114.381	89	240.5	79.493	
Nine hours rainfall	455	0	136.317	95	284.5	85.988	
Twelve hours rainfall	455	0	150.161	99	447	95.431	
One-day rainfall	455	0	201.598	162	538.5	111.62	
Three-days rainfall	455	0	279.6	283.5	549.5	85.875	
Seven-days rainfall	455	0.5	323.16	330	633.5	87.895	
Two-weeks rainfall	455	0.5	385.033	399.5	663	89.754	
Three-weeks rainfall	455	85.5	503.989	533	914.4	114.888	
Four-weeks rainfall	455	108	586.585	561	1135	159.945	
Slope	455	10	34.004	34.004	65	7.938	Degree
Slope length	455	1.8	21313	13	180	22.623	m
Altitude	455	9	390.789	272	1324	273.069	m

References

Causes, L., 2001. Landslide types and processes. US Geological Survey: Reston, VA, USA.

Dahal, R. K., Hasegawa, S., Nonomura, A., Yamanaka, M., Masuda, T., Nishino, K., 2008. GIS-based weights-of-evidence modeling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. Environmental Geology, 54(2), 311-324.

Dai, F. C., Lee, C. F., 2001. Frequency–volume relation and prediction of rainfall-induced landslides. Engineering geology, 59(3-4), 253-266.

Huang, F., Chen, J., Du, Z., Yao, C., Huang, J., Jiang, Q., ..., Li, S., 2020. Landslide susceptibility prediction considering regional soil erosion based on machine-learning models. ISPRS International Journal of Geo-Information, 9(6), 377.

Lombardo, L., Opitz, T., Huser, R., 2018. Point process-based modeling of multiple debris flow landslides using INLA: an application to the 2009 Messina disaster. Stochastic environmental research and risk assessment, 32, 2179-2198.

Lombardo, L., Opitz, T., Ardizzone, F., Guzzetti, F., Huser, R., 2020. Space-time landslide predictive modelling. Earth-Science Reviews, 209, 103318.

Lombardo, L., Tanyas, H., Huser, R., Guzzetti, F., Castro-Camilo, D., 2021. Landslide size matters: A new data-driven, spatial prototype. Engineering Geology, 293, 106288.

Malamud, B. D., Turcotte, D. L., Guzzetti, F., Reichenbach, P., 2004. Landslide inventories and their statistical properties. Earth Surface Processes and Landforms, *29*(6), 687-711.

Meena, S. R., Puliero, S., Bhuyan, K., Floris, M., Catani, F., 2022. Assessing the importance of conditioning factor selection in landslide susceptibility for the province of Belluno (region of Veneto, northeastern Italy). Natural hazards and earth system sciences, 22(4), 1395-1417.

Shahabi, H., Hashim, M., 2015. Landslide susceptibility mapping using GIS-based statistical models and Remote sensing data in tropical environment. Scientific reports, *5*(1), 9899.

Shirzadi, A., Shahabi, H., Chapi, K., Bui, D. T., Pham, B. T., Shahedi, K., Ahmad, B. B., 2017. A comparative study between popular statistical and machine learning methods for simulating volume of landslides. Catena, 157, 213-226.

NGII, (2018). Digital elevation model, NGII (National Geographical Information Institute), the Ministry of Land, Infrastructure and Transport, Korea.

Ngo, P. T. T., Panahi, M., Khosravi, K., Ghorbanzadeh, O., Kariminejad, N., Cerda, A., Lee, S., 2021. Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. Geoscience Frontiers, 12(2), 505-519.

Pham, B. T., Tien Bui, D., Prakash, I., 2018. Bagging-based support vector machines for spatial prediction of landslides. Environmental Earth Sciences, 77(4), 1-17.

Raja, N. B., Çiçek, I., Türkoğlu, N., Aydin, O., Kawasaki, A., 2017. Landslide susceptibility mapping of the Sera River Basin using the logistic regression model. Natural Hazards, 85(3), 1323-1346.

Ritchie, A. M., 1958. Recognition and identification of landslides. Landslides and Engineering Practice, Highway Research Board Special Report, 29, 48-68.

Wang, B., Vardon, P. J., Hicks, M. A., 2016. Investigation of retrogressive and progressive slope failure mechanisms using the material point method. *Computers and Geotechnics, 78*, 88-98.