Mapping of landslide-prone areas in the Lisu river basin
Barru Regency based on binary logistic regression

I G Sudarman and A Ahmad
Department of Soil Science, Faculty of Agriculture, Hasanuddin University, Makassar, Indonesia

E-mail: asmita.ahmad@agri.unhas.ac.id

Abstract. Barru Regency Government once issued an emergency response status for floods and
landslides on December 28, 2018. According to data from the Barru Regency Regional Disaster
Management Agency, from 2013 to 2019, there were always landslides in the Districts of Tanete
Riaja, Pujananting, Tanete Rilau, and Barru. This area is included in the Lisu Watershed. This
will produce a map of the distribution of the landslide-prone regions in the Lisu watershed. This
study uses Binary Logistic Regression (BLR) and NDVI (Normalized difference vegetation
index) analysis. The data used are landslide points, soil types, lithology, slopes, land use, rainfall,
soil texture, and river distance. One hundred thirty-seven landslide points were recorded that
were successfully obtained through field surveys and NDVI image analysis. The research area
has a type B climate according to the Schmidt-Ferguson classification system. The variables of
rainfall, land-use class, lithology type, and slope class significantly affect regression modeling
with significance values of 0.042, 0.000, 0.003, and 0.000, respectively. Variables of rainfall,
distance from the river, slope class, lithology type, and land use class significantly affect
regression modeling. Landslides occur on slopes from rather steep to very steep slopes. A total
of 60 landslides occurred in the Camba Formation. A total of 93 landslides occurred in the scrub
land-use class. Landslide-prone areas with a slightly hazardous class covering an area of 4386
hectares, 4031 hectares of prone, and 4275 hectares of very prone areas are generally scattered
in the Lisu River Basin's southern region.

1. Introduction
Landslide disasters had a major impact on society's economic system, so the government and research
institutions make various approaches to assess the level of vulnerability, danger, and risk of landslides,
then produce a landslide hazard map for disaster mitigation management purposes [1,2].
Badan Nasional Penanggulangan Bencana (BNPB) released data in early 2019, which reported 366
incidents, 76 of which were landslides. The number of landslides in January 2018 was 49 incidents, up
to 76 incidents in January 2019. It is stated that 189 people were victims of this incident, and 76 houses
were damaged. According to [3] Barru Regency has areas that are prone to landslides around 61,786
hectares.

The Lisu River Basin is one of the watersheds that plays an important role and has a strategic position
that is important in Barru Regency's development [4]. The Lisu River Basin is potential agricultural
areas. Data from BPDAS Jeneberang Walanae in 2014 informed that the critical land in Lisu watershed
in 2013 was 18,854 ha (48.67% of Lisu watershed area). The data of the landslides event can be seen in
table 1. So it is necessary to map the level of landslide hazard in Lisu area.
Table 1. Landslide event data in Barru Regency.

| Year | Landslide Event | District |
|------|-----------------|----------|
| 2013 | 19              | Tanete Riaja, Pujananting, Barru |
| 2014 | 7               | Tanete Riaja, Pujananting |
| 2015 | 17              | Tanete Riaja, Mallusetasi, Barru |
| 2016 | 14              | Tanete Riaja, Pujananting, Tanete Rilau, Barru, Sopeng Riaja, Balusu, Mallusetasi |
| 2017 | 12              | Tanete Riaja, Pujananting, Tanete, Rilau, Barru, Balusu, Dan Mallusetasi |
| 2018 | 16              | Tanete Riaja, Pujananting, Barru, Mallusetasi |
| 2019 | 7               | Pujananting, Tanete Riaja |

[5] stated that the binary logistic regression method has an accuracy value of 85.20%, a statistical index of 80.37%, and a process hierarchy analysis of 75.70%. [6] stated that the Binary Logistic Regression (BLR) method had been widely used throughout the world, is easy to use, and data processing is available to make it easier to work at the analysis stage. This research aims to produce a landslide hazard distribution map in the Lisu River Basin based on the Binary Logistic Regression method.

2. Methods
This research took place in Barru Regency and focused on the Lisu River Basin, which covers the Districts of Tanete Riaja, Pujananting, and Barru (figure 1). Soil analysis was carried out at the Soil Physics Laboratory, Faculty of Agriculture, Hasanuddin University.

Figure 1. Location of the Lisu Watershed.
The tools used in this study include ArcGIS 10.3, SPSS, Binary Logistic Regression Tools, Normalized Difference Vegetation Index (NDVI), GPS, Google Earth Pro, laboratory tools, and camera. The materials used in this study are 2009-2019 rainfall data, National Digital Elevation Model (Demnas), river distribution maps, 2009-2019 Sentinel2 imagery, geological maps, and soil type maps.

2. 1. Landslide Point Distribution

Recording and tracking landslide points are carried out in two ways: normalized difference vegetation index (NDVI) analysis and direct field surveys. Landslide point distribution data can be built using OBIA in the form of NDVI. Landslide points can be detected in image data due to changes between the pixel value before and after a landslide. Landslides will cause loss of vegetation in a field of land. This loss of vegetation will then change the NDVI value of the image recording in the area. By comparing the NDVI values before and after landslides, the points that are suspected of being landslides will be found. This point is validated during field surveys and image observations.

\[
:\text{NDVI} = \frac{(\text{IR} - R)}{(\text{IR} + R)}\]

The hypothesis H0 is accepted when the regression analysis test states no independent variables affect the dependent variable. H1 is accepted when there are independent variables that affect the dependent variable. For this study, the independent variables included the variable distance from the river, rainfall, land use, lithology type, slope, soil type, and texture class. The dependent variable in this study is the occurrence of landslides. Variables of lithology type, soil type, texture class, and land use are classified as categorical data.

2. 2. Binary Logistic Regression

This binary logistic regression (BLR) application requires a basic concept that states what factors can cause landslides, expressed in measured parameters, independent variables. BLR used to produce landslide hazard maps requires data in activities, points, and the distribution of all landslides that occur in the study area. BLR will estimate the maximum likelihood after changing the dependent variable into a logical variable representing the natural logarithm of probability and states that a phenomenon occurs or does not occur.

\[
\text{p} = \frac{1}{1 + e^{-Z}}
\]

where \( P \) is the estimation result that will state the probability of landslides; this value ranges from 0 to 1 and will form S curve. The Z value is a linear combination that has the following equation

\[
Z = \text{intercept} + b1x1 + b2x2 + b3x3 + \cdots bn xn
\]

Where \( b1, b2, b3, \) and \( b_n \) are the coefficients of each variable in the logistic regression model. The values of \( x1, x2, x3, \) and \( xn \) are independent variables. The intercept in this equation is a constant from the regression [5].

The hypothesis H0 is accepted when the regression analysis test states no independent variables affect the dependent variable. H1 is accepted when there are independent variables that affect the dependent variable. For this study, the independent variables included the variable distance from the river, rainfall, land use, lithology type, slope, soil type, and texture class. The dependent variable in this study is the occurrence of landslides. Variables of lithology type, soil type, texture class, and land use are classified as categorical data.

2. 3. Validation

Validation is an activity to match the accuracy of the information from the landslide-prone area distribution map produced with the actual conditions in the field. This activity is a field survey by visiting certain points on the map to do a ground check and recording the coordinates of landslides at the research location. The resulting regression model is validated using ROC (Receiver Operating Characteristic) curve analysis. Validation to assess regression models can be done using the Omnibus Test, Nagelkerke R Square, and the Hosmer and Lemeshow Test. This study uses a level of confidence (\( \alpha \)) = 0.05 or 5%.

The flowchart of the study and the landslides parameters factor used in this study can be seen in figure 2 and figure 3.
3. Results

NDVI analysis and field surveys recorded 137 landslide points in the southern part of the Lisu watershed. Moderately steep has the widest area, around 10692 ha (27.1%), steep with an area of 10178 ha (25.8%), nearly level has an area of 9432 ha (23.9%), moderately sloping have an area of 7477 ha (18.9%), and very steep has an area 1710 ha (4.3%) (figure 4). From slope condition data, it explains that this research area has a predominantly mountainous topography. This sloping area is spread in the eastern part, while the flat slope class is spread out in the northwest.

The Camba Formation is the most extensive in the study area covering 13191 ha (33.4%), the Tonasa Formation covering 9028 ha (22.9%), the Balangbaru Formation covering 4641 ha (11.8%), Alluvium covering 3980 ha (10.1%), the Bantilmala Complex covering 3021 ha (7.7%), Basalt covers 2261 ha (5.6%), Diorite covers 1550 ha (3.9%), Trakit covers 1218 ha (3.1%), and the Malawa Formation covers 613 ha (1.6%) (figure 4).

Dystropepts dominates the Lisu Watershed with 24799 ha, Eutropepts covering 5715 ha, Fluvaquents covering 2497 ha, Tropudults covering 2280 ha, Tropaquents covering 1848 ha, Tropudalfs covering 520 ha, Tropopsamments covering 471 ha, and Eutrandepts covering 91 ha (figure 4). This research used five land use classes in forest, garden, settlement, agriculture, and shrubs. The widest use class is the forest with an area of 20010 hectares, agriculture with 8948 hectares, shrubs with 7793 hectares, settlements with 1391 hectares, and gardens with 1024 hectares (figure 4).

Rainfall in the study area ranges from 2300 to 2500 mm/year (figure 5 and figure 6). The rainfall class division is divided into two classes, class 2300 - 2400 mm/year covering an area of 16066 ha in the north and class 2400 - 2500 mm/year covering 23400 ha in the south.
Figure 4. Map of landslide variable.

Figure 5. Chart of annual rainfall 2009-2019 in the Lisu Watershed.

Figure 6. Chart of monthly rainfall 2009 – 2019 in the Lisu Watershed.
2016 is the year with the highest rainfall, which is 3503 mm. The monthly rainfall analysis shows that December has the highest rainfall, which is 398 mm. Based on the Schmidt-Ferguson climatic classification, the study area has eight months of wet months (January, February, March, April, May, October, November, December) and two months of the dry season (July and August) with the Q value is 25%, so it is included in the type B climate which is described as a wet climate.

This study uses data sourced from the Lisu River and all tributaries in the Lisu River Basin. The distance class <100 has an area of 20366 hectares, 100-200m has an area of 11765 hectares, 200-300m has an area of 4743 hectares, and <300m has an area of 3769 hectares. The Lisu River Basin has a dominant dendritic flow pattern and partially parallel flow pattern.

The area with the clay texture class is the largest area with 15947 ha, clay with an area of 7466 ha, dusty loam with an area of 4445 ha, dusty clay with an area of 2794 ha, clay sand with an area of 1681 ha, sandy clay loam with an area of 877 ha, and the smallest is sandy loam with an area of 306 ha. The dominant fraction found in the southern region is clay to dust.

The Omnibus Test analysis states the appropriateness of using independent variables to predict the dependent variable. The resulting analysis states that this model is appropriate because it has a significant value (0.00) smaller than the confidence level (0.05). The data analysis results have resulted in Cox & Snell R2 values of 0.547 and Nagelkerke R2 of 0.751 (table 2). The Nagelkerke R2 value stated that the independent variables could explain the dependent variable (landslide) by 75%.

Table 2. Model Summary.

| Step | -2 Log likelihood | Cox & Snell R Square | Nagelkerke R Square |
|------|------------------|----------------------|--------------------|
| 1    | 196.311          | 0.547                | 0.751              |

The Hosmer and Lemeshow Test's value has a significant value of 0.232 (table 3), which is greater than the confidence level of 0.05; this indicates that the data model is worth analyzing by BLR. The classification table explains that 116 landslide points from a total of 137 landslide events can be predicted by the BLR and have an accuracy of 88.5% (table 4). The regression equation can explain the relationship between the independent and dependent variables. Analysis of the Omnibus Test, Nagelkerke R Square, and the Hosmer and Lemeshow Test as a goodness of fit to determine regression model is acceptable or not [5]. The Omnibus Test, Nagelkerke R Square, and the Hosmer and Lemeshow Test stated that this study's regression model is acceptable.

Table 3. Hosmer and Lemeshow test.

| Step | Chi-square | Df | Sig. |
|------|------------|----|------|
| 1    | 10.491     | 8  | 0.232|

Table 4. Classification tablea.

| Observed | Predicted Longsor 0.00 1.00 Percentage Correct |
|----------|-----------------------------------------------|
| Step 1   | Longsor                                      |
| .00      | 224 23 90.7                                  |
| 1.00     | 21 116 84.7                                  |

| Overall Percentage | .88.5 |
|--------------------|-------|
| a. The cut value is .500. |

BLR analysis states that the slope variable has a significant value smaller than the confidence level, meaning that the slope variable affects the regression model (table 5). The value of the slope variable for the coefficient (B) is 2.33 so that it has a positive correlation with landslides. This value indicates that the higher the slope class, the greater the potential for landslides. According to [7], the slope condition is the most important factor determining how stable the soil is in an area. According to [8],
soil and slopes are two important factors that can affect an area's condition and stated that 90% of landslides in his research occurred in slope classes of more than 15%. [5] stated that in areas with a mountainous topography, the soil mass movement would be influenced by slope and slope length factors. The number of landslide points found in a steep area with a greater number than very steep slopes is very much influenced by the area, where the area with a steep class is higher than the very steep class.

### Table 5. Result of binary logistic regression analysis.

|          | B    | S.E.  | Wald     | df | Sig.  | Exp(B) | 95% C.I for EXP(B) |
|----------|------|-------|----------|----|-------|--------|-------------------|
| Step 1a  |      |       |          |    |       |        |                   |
| CH       | .840 | .413  | 4.133    | 1  | .042  | 2.317  | 1.031 - 5.211     |
| Sungai   | -.224| .222  | 1.019    | 1  | .313  | .799   | .517 - 1.235      |
| LU       | -.809| .110  | 53.946   | 1  | .000  | .445   | .359 - .553       |
| Litologi | .252 | .086  | 8.712    | 1  | .003  | 1.287  | 1.088 - 1.522     |
| Lereng   | 2.332| .281  | 69.105   | 1  | .000  | 10.298 | 5.943 - 17.846    |
| Tanah    | -.712| 1.065 | .447     | 1  | .504  | .490   | .061 - 3.955      |
| Tekstur  | -.060| .096  | .391     | 1  | .532  | .942   | .781 - 1.136      |
| Constant | -6.402| 2.741| 5.454    | 1  | .020  | .799   | .445 - .809       |

a. Variable(s) entered on step 1: CH, Sungai, LU, Litologi, Lereng, Tanah, Tekstur.

The rainfall variable has a significant value smaller than the confidence level, so this variable significantly affects landslides. The rainfall variable has a coefficient value (B) in the regression model of 0.84, which indicates that rainfall has a positive correlation with landslides. [9] and [10] describe that rainfall is an important hydrological parameter for the landslide scheme's modeling analysis as one of the triggers for landslides. [11] explained that the heavy level of rainfall at sloping locations would have the potential to cause landslides. [12] stated that the washing process on the source rock would produce clay minerals that have low permeability, inhibiting the percolation process of groundwater in sloping areas, and increasing the clay content of the soil, then the soil stability will decrease and tends to trigger landslides.

The lithology type in the regression analysis results has a significantly smaller value than the confidence level, so this variable affects landslide events. The lithology variable has a coefficient (B) of 0.252 which means that the lithology type positively correlates to the regression model. [13] explained that The Tonasa Formation contains sedimentary rocks in sandstone, siltstone, and clay rock formed during the geological age from Eocene to Miocene. This formation is the largest on the western side of South Sulawesi and is bounded by the Walanae fault.

The land-use class variable in the regression analysis results has a significantly smaller value than the confidence level, so this variable affects landslide events. The land-use class variable has a coefficient value (B) of 0.809, so that this variable has a negative correlation with landslides. [14] describe that land use will determine the type, variety, and density of vegetation in the area. This growing vegetation will affect the environment, where the vegetation's roots will affect soil formation and slope stability. Furthermore, [15] explained that land-use change would affect slope stability and impact hydrological processes in the form of infiltration, surface runoff, and soil strength.

The distance variable from the river has a significant value greater than the confidence level, so this variable has no significant effect on the regression model. The distance variable from this river has a coefficient value of (B) -0.22 which means that this variable has a negative correlation with landslides, so the farther the point is from the river, the smaller the potential for landslides. [16] explained that areas close to rivers with steep slopes would experience heavy erosion, which in the event of high-intensity rain and sloping topography will cause landslides. [17] explained a relationship between river density and the strength of the soil bearing capacity against landslides in an area; the relationship connects to the permeability of rocks in passing water.

The soil type in the regression analysis has a significant value greater than the confidence level, so that it does not have a significant effect on the regression model. This soil type variable has a coefficient value of (B) -0.71 which means that this variable negatively correlates with landslides. The Dystropepts...
soil type dominates the southern part of the Lisu watershed. [8] state a relationship between soil types and landslides that landslides can occur on Inceptisols. [7] stated that inceptisols are lands that have not undergone much development. [18] The landslide will remove soil solum in the sloping area and show the soil type’s material and source rock.

The texture class has a significant value greater than the confidence level so that it does not have a partially significant effect on the regression model. The texture class variable in the regression analysis has a coefficient (B) of -0.06, which means that this variable negatively correlates with landslides. Several studies have been conducted on the relationship between particle size, macro-micro pore scale, and soil texture to landslides. According to [19] and [20], landslides occur as a result of the release of a volume of soil on a relatively waterproof layer that is saturated with water, the layer contains a high level of clay, and after saturation, the water will serve as a slippery field. The more clay or smooth a soil texture, the more micro pores space is formed, filled with water and air. [21] explained that soils with a loamy texture that slips could trigger large landslides that cover a very large area. According to [22], soils with sand and dust textures are very susceptible to landslides compared to clay textures with better water holding power. The dominance of texture in the soil will significantly affect the soil's mechanical processes in sloping areas and can cause landslides.

The very high susceptible level has 4275 hectares with 90 events landslide points, the high susceptible level has 4031 hectares with 24 landslide points, the moderate susceptible level has 4386 hectares with 12 landslide points, and the low susceptible level has 26314 hectares with 11 landslide points (figure 7). The very high susceptible, highly susceptible, and moderate susceptible levels are scattered in the southern part of the Lisu River Basin, while the areas that are low susceptible levels are mostly in the northern part (table 6). The southern part of the Lisu watershed has been dominated by very steep has lithology types of the Camba Formation, Balangbaru Formation, Malawa Formation, Basal Formation, Bancuh Bantimala Complex, and Tonasa Formation, has forest and scrub classes, and is dominated by soil types Dystropepts. This map of the distribution of landslide-prone areas can be used by regional designers, policy-makers and field workers to carry out land management in areas prone to landslides to minimize the potential for landslides with good management.

![Figure 7. Landslide susceptible maps in the Lisu Watershed.](image)

![Figure 8. ROC curve.](image)
Table 6. Landslide Frequency in Lisu Watershed.

| Susceptible Level   | Total Area (Ha) | Landslide |
|---------------------|-----------------|-----------|
| Low Susceptible     | 26314           | 11        |
| Moderate Susceptible| 4386            | 12        |
| High Susceptible    | 4031            | 24        |
| Very High Susceptible| 4275          | 90        |

The binary logistic regression equation model used to produce a landslide hazard distribution map in the Lisu watershed has a performance value of 0.96 based on ROC analysis (figure 8); this was right and valid as according to [22] stated that the regression model would be categorized as good when it has a performance value greater than 0.7.

4. Conclusions
The moderate susceptible level covering 4386 hectares, the high susceptible level covering 4031 hectares, and the very high susceptible level covering 4275 hectares are generally scattered in the Lisu River Watershed region. The occurrence and distribution of landslides are influenced by rainfall variables, land use, lithology type, and slope class.

References
[1] Felícísimo Á M, Cuartero A, Remondo J and Quirós E 2013 Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: A comparative study Landslides 10 175–89
[2] Ahmad A, Lopulisa C, Imran A, Baja S and Solle M S 2020 Spatial analysis of landslide vulnerability in Enrekang District , South Sulawesi Spatial analysis of landslide vulnerability in Enrekang District , South Sulawesi IOP Conf. Ser. Earth Environ. Sci. 486 1–8
[3] Leprina S, Usman A and Budirman B 2017 Identification of types and Landslide Characteristics in Lisu Watershed, Barru regency 1–10
[4] Semu Y D, Arsyad U and Umar A 2018 Indikator Kinerja dan Peran Stakeholder dalam Pengelolaan Daerah Aliaran Sungai (DAS) Lisu J. Hutan dan Masy. 10 257
[5] Pourghasemi H R, Moradi H R and Fatemi Aghda S M 2013 Landslide susceptibility mapping by binary logistic regression, analytical hierarchy process, and statistical index models and assessment of their performances Nat. Hazards 69 749–79
[6] Lombardo L and Mai P M 2018 Presenting logistic regression-based landslide susceptibility results Eng. Geol. 244 14–24
[7] Conforti M, Pascale S, Robustelli G and Sdao F 2014 Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy) Catena 113 236–50
[8] Eslami M, Shadfar S, Mohammadi-Torkashvand A and Pazira E 2019 Assessment of density area and LNRF models in landslide hazard zonation (Case study: Alamout watershed, Qazvin Province, Iran) Acta Ecol. Sin. 39 173–80
[9] Chang K, Huang J, Kao S and Chiang S 2009 Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications Data Assim. Atmos. Ocean. Hydrol. Appl. 127–45
[10] Ahmad A, Poch R M, Lopulisa C, Imran A M and Baja S 2018 Identification of Soil Characteristic on North Toraja Landslide, Indonesia ARPN J. Eng. Appl. Sci. 13 8381–5
[11] Busthan., S afruddim. and Pachi. 2020 Identification of landslide disaster potential based on weathering grade of rock in Parepare City South Sulawesi, Indonesia IOP Conf. Ser.: Mater. Sci. Eng. 875 1-7
[12] Maulana A, Jaya A, Irfan U R and Farida M 2019 Field geological characteristic of Mesozoic paremba sandstone, Bantimala Complex, South Sulawesi IOP Conf. Ser. Earth Environ. Sci. 279 1-5
[13] Bartelletti C, Giannecechini R, D'Amato Avanzi G, Galanti Y and Mazzali A 2017 The influence
of geological-morphological and land use settings on shallow landslides in the Pogliaschina T. Basin (northern apennines, Italy) J. Maps 13 142–52

[14] Chen L, Guo Z, Yin K, Pikha Shrestha D and Jin S 2019 The influence of land use and land cover change on landslide susceptibility: A case study in Zhushan Town, Xuan'en County (Hubei, China) Nat. Hazards Earth Syst. Sci. 19 2207–28

[15] Wang F, Xu P, Wang C, Wang N and Jiang N 2017 Application of a gis-based slope unit method for landslide susceptibility mapping along the longzi river, southeastern tibetan plateau, China ISPRS Int. J. Geo-Information 6

[16] Sobirin. Sitanala, F.T.H.R. Ramadhan M 2004 Analisis Potensi Dan Bahaya Bencana Longsor Menggunakan Modifikasi Metode Indeks Storie Di Kabupaten Kebumen Jawa Tengah Ind. Res. Work. Natl. Semin. Politek. Negeri Bandung 8 59–64

[17] Trustrum N A and De Rose R C 1988 Soil depth-age relationship of landslides on deforested hillslopes, taranaki, New Zealand Geomorphology 1 143–60

[18] Priyono 2015 Hubungan klasifikasi longsor, klasifikasi tanah rawan longsor dan klasifikasi tanah pertanian rawan longsor Gema 27 1602–17

[19] Sartohadi J, Harlin Jennie Pulungan N A, Nurudin M and Wahyudi W 2018 The ecological perspective of landslides at soils with high clay content in the middle Bogowonto watershed, Central Java, Indonesia Appl. Environ. Soil Sci. 2018

[20] Ahmad A, Lopulisa C, Baja S and Imran A 2019 The correlation of soil liquid limit and plasticity index for predicting soil susceptibility: A case study on landslides area in South Sulawesi The correlation of soil liquid limit and plasticity index for predicting soil susceptibility: A case study on IOP Conf. Ser. Earth Environ. Sci. 235 1–10

[21] Isra N, Lias S A and Ahmad A 2019 Karakteristik Ukuran Butir Dan Mineral Liat Tanah Pada Kejadian Longsor (Studi Kasus: Sub Das Jeneberang) J. Ecosolum 8 62

[22] Margarint M C, Grzavu A and Patriche C V 2013 Assessing the spatial variability of coefficients of landslide predictors in different regions of Romania using logistic regression Nat. Hazards Earth Syst. Sci. 13 3339-3355