LANDSLIDE SUSCEPTIBILITY ASSESSMENT OF KOKAP AREA USING MULTIPLE LOGISTIC REGRESSION

Somyot Makealoun[*1], Doni Prakasa Eka Putra[2], and Wahyu Wilopo[2]

[1]*Water Resources Department, Faculty of Engineering, National University of Laos.*
[2]*Geological Engineering Department, Gadjah Mada University, Indonesia.*

Abstract

A number of landslides have occured in Kokap Sub-District, Kulon Progo Regency, Yogyakarta Special Province, Indonesia, which have influenced the communities. The natural disaster is commonly associated with a few days of heavy rainfall events. To mitigate the impact of landslides in this area, a landslide susceptibility assessment needs to be carried out. The main objective of this research is to develop a landslide susceptibility zonation in the research area by applying a logistic regression (LR) method. Field observation was conducted at 68 locations in the research area, in which 46 landslides occured. Data of slope angle, lithology, geologic structure and groundwater conditions were collected. The relationship between landslide occurrence and the slope angle, lithology, geologic structure and groundwater conditions was analysed using the LR method. The analysis results showed a 0.984 standard error, implying a good-fit model. The study area was classified into very low, low, moderate, high and very high landslide susceptibility zones with 0–20%, 20–40%, 40–60%, 60–80%, and 80–100%, respectively, probabilities of occurrence. A 60% area of the total study area was classified as a moderate to very high susceptibility to landslide. From 47 landslides, 80% landslides occured in high and very high landslide susceptibility zones, 17% landslides occured in the moderate susceptibility zone and 2% landslides occured in the low susceptible zone. None of landslides occured in the very low landslide susceptibility zone. The analysis results show that LR method is a very useful method for landslide prediction.
Keywords: *landslide susceptibility, multiple logistic regression, Kokap Kulon Progo-Indonesia*

1 Introduction

Indonesia is one of the countries in the world that has suffered frequently from natural disasters, such as flood, drought, volcanic eruption, earthquake, tsunami and landslide. The disasters occur due to natural, climate, or environmental conditions. As landslides are defined as the movement of slope forming materials under the influence of gravity (Varnes, 1978), they usually occur in hilly or mountainous areas with high slope angle. Kokap Sub-district of Kulon Progo is located on Kulon Progo Mountains, which are about 30 km from Yogyakarta City of Indonesia (Figure 1). Rockfall and debris flow types of landslides have commonly occurred in this area and are commonly associated with a few days of heavy rainfall events and steep slope angle. The occurrence of landslides cannot be averted, but understanding of the hazard can lead to proper mitigation strategies and therefore significant reduction of the impacts. One of important means in landslide hazard mitigation is a development of a landslide hazard zonation or landslide susceptibility map. In this paper, the term of susceptibility

*Corresponding author: S. MAKEALOUN, Water Resources Department, Faculty of Engineering, National University of Laos. E-mail: makealoun.somyot@gmail.com
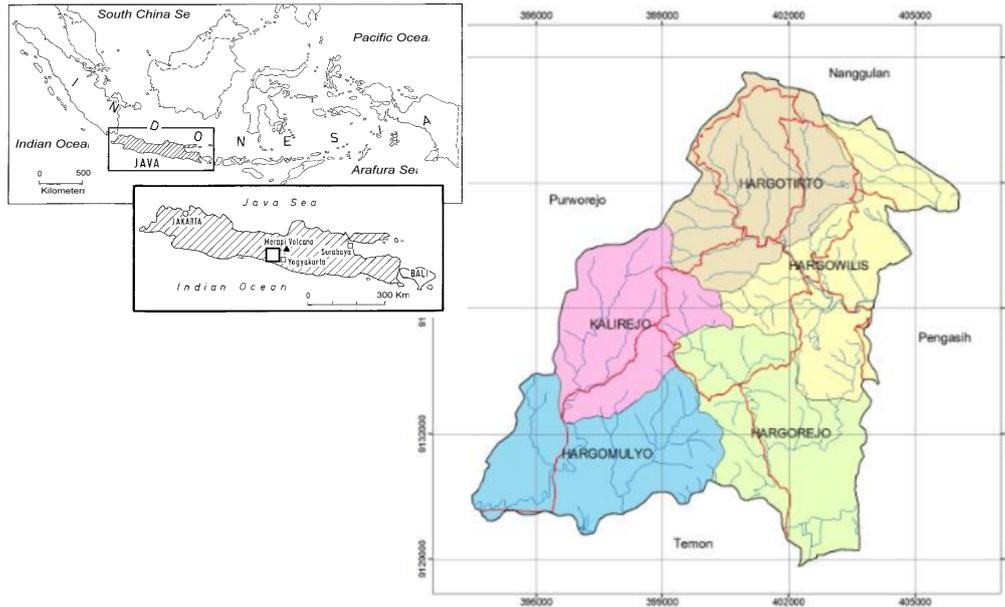
Figure 1: Location of the study area.

is used to represent the propensity of an area to undergo a landslide.

According to Van Westen *et al.* (1999), there are two methods of landslide susceptibility zonation, which are direct and indirect methods. The direct method is determined by experience and knowledge of the terrain and the indirect method is determined based on statistical method, such as LR analysis or deterministic method or other indirect approaches. A number of applications of LR method in landslide susceptibility assessment has been reported by Rupert et.al. (2002), Chau and Chan (2005), Das *et al.* (2010) and Bai *et al.* (2010). Chau and Chan (2005) explicitly mentioned that according to Rowlotham and Dudycha (1998), the LR method was the most useful method for landslide susceptibility assessment. The LR method is therefore applied in this study to assess the susceptibility of the research area to landslide.

2 Geological condition

Kulon Progo mountains are composed of volcanic rocks, which are dominated by volcanic clastic materials and andesite intrusion (Syafri *et al.*, 2013). Rahardjo *et al.* (1995) named this volcanic rocks as the Old Andesite Formation. Below the volcanic rocks of the Old Andesite

Formation, the oldest stratigraphic unit in the area is known as Nanggulan Formation, consisting of sandstones with lignite, sandy marl, mudstone with limonite, marl, sandstone, tuff (Rahardjo *et al.*, 1995). The whole succession was overlain by sandy tuff and limestone and the youngest rock found in the area is quaternary alluvial deposits. The morphology of Kulon Progo mountains forms an elliptical-shape dome (Syafri *et al.*, 2013). The slope with a number of radially pattern faults can be recognized in the Geological Map of Yogyakarta (Rahardjo *et al.*, 1995). The study area has an undulating to steep hilly morphology with minimum and maximum elevations of about 6 m and 850 m above the sea level, respectively (Figure 2a–c). The slopes in the study area range from about 3 to 50°. The exposed lithologies in the study area (i.e., andesite breccias, sandstone, limestone) are mostly weathered (Figure 2d–f).

3 Logistic regression

Logistic regression has been used extensively in the health and social sciences since the late 1960s to predict a binary response from explanatory variables (diagnostic – prediction model). Logistic regression is a statistical
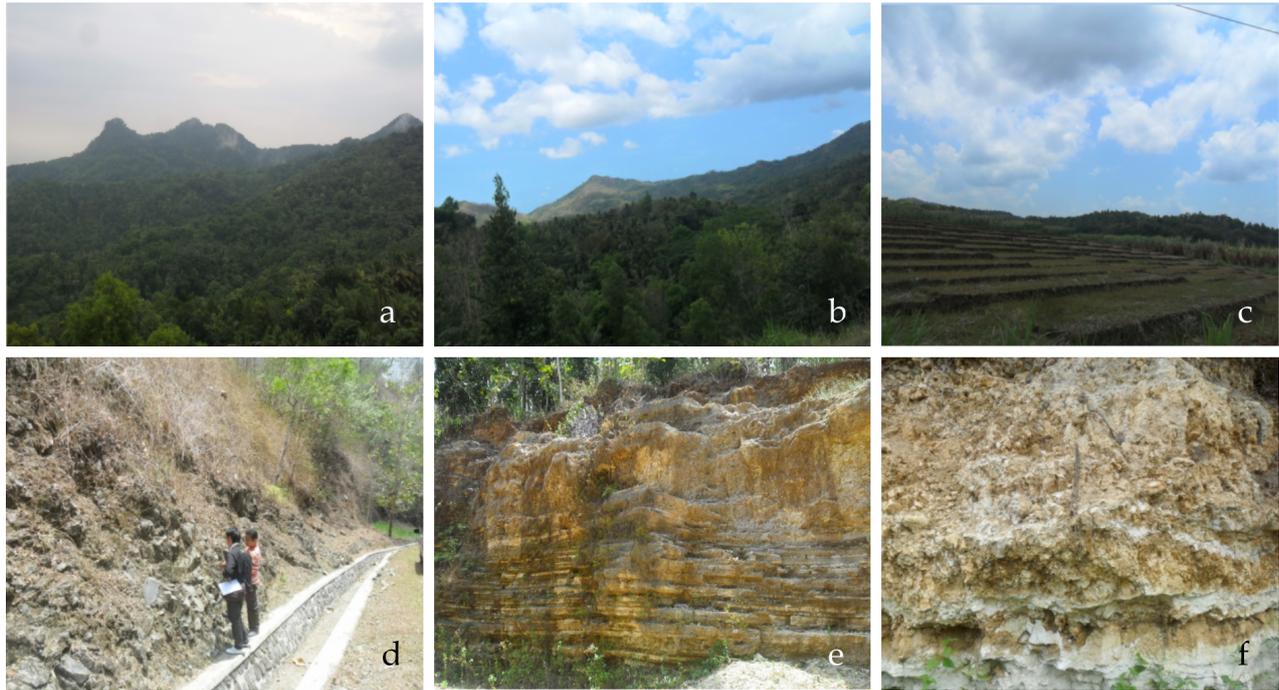
Figure 2: (a–c) Steep hilly to undulating morphology in the study area; (d) Andesite breccia; (e) Tufaceous sandstone, and (f) Limestone, which are exposed in the study area.

method that predicts a probability of occurrence. It is similar to a linear regression model but it is suited to models where the dependent variable is dichotomous or binary (Hosmer *et al.*, 2013). In term of landslide, binary variable will be "0" for "no landslide" and "1" for "yes there is a landslide occurrence" on an observed locations. The equation of multiple logistic regression is shown in Eq. 1 below:

$$P_y = \frac{e^{\alpha + \beta 1 X1 + \beta 2 X2 + \cdots + \beta i Xi}}{1 + e^{\alpha + \beta 1 X1 + \beta 2 X2 + \cdots + \beta i Xi}} \qquad (1)$$

where: P_y = probability of occurrence (0 to 1); X = independent variable; α = statistically constant coefficient; β = statistically derived coefficient.

The equation of LR analysis is called a LR model. The performance of LR model can be evaluated using several statistical measures. Individual regression coefficients are considered to be statistically significant if the p-value of the Wald statistic is less than or equal to 0.05 (95-percent confidence or greater). The Wald statistic follows a chi-squared distribution and is used to indicate where the model coefficients are significantly different from zero. A simple but sometimes subjective technique to evaluate if the model is statistically stable is to inspect the magnitude of the standard error (SE) of each variable. According to Chan (2004), the acceptable magnitude of the SE is around 0.001–5.0. However, there are no fixed criteria on how small the SE should be but a matter of judgment.

4 Landslide causative factors

According to Karnawati (2005), causative factors of landslide can be differentiated into two factors, which are controlling and triggering factors. The controlling factors of landslide include the slope morphology, slope stratigraphy, slope material, structural geology (joints, faults), groundwater condition and also some intrinsic factors related to the slope condition. The triggering factors are natural triggering factors, such as rainfall and earthquake. Human influence, such as forest devastation or change of land use can also become the triggering factors. Chau and Chan (2005) mentioned that the causative factors of landslide may vary from one region to another. Therefore, it is necessary

and reasonable that local landslide inventory in conjunction with local slope condition must be well examined to obtain reliable landslide predictions.

5 Methodology

In this study, the causative factors of landslide used to predict a landslide in the study area are slope angle, lithology, distance from faults or lineaments and groundwater condition. Slope angle and distance from lineaments were assessed from evaluation of Digital Elevation Model (DEM) in GIS and were also checked in the field. Data of lithologies and faults were obtained from field observation and literature. Data of groundwater condition were obtained from field measurements of groundwater depth.

The landslide susceptibility assessment was performed by quantifying each of the causative factors of landslide at 68 random observation locations. The quantification of the causative factors of landslide is shown in Table 1. All of the 68 data were then analysed by multiple LR to obtain the LR model of landslide prediction in the study area by applying the landslide occurrence as dependent variable and the causative factors of landslide (i.e., slope angle, slope lithology, distance from faults or lineaments and groundwater depth) as independent variables. The data of each causative factors are shown in Figure 3 to Figure 6. The LR equations obtained from this analysis were used to develop a landslide susceptibility map of the study area using the Geographical Information System (GIS).

6 Results and discussion

The LR analysis results are presented in Table 2. From this table, it can be seen that the SE of the multiple logistic regression is between 0.001 and 5.0, meaning that the LR model is acceptable. Based on the value of constant and derived coefficient from Table 3, the final equation of logistic regression of landslide occurrences versus landslide causative factors (i.e., lithology type, slope angle, distance from faults and lin-

eaments and groundwater condition) can be obtained.

By applying these coefficients of LR equation to the GIS, a final map of landslide susceptibility of the study area can be produced, as shown in Figure 7. The LR analyses show that probabilities of landslide occurrence in the study area are from 0% (no landslide will be occurred in this area) to 100% (absolute occurrence of landslide in this area). Based on the probability, the study area can be classified into very low, low, moderate, high and very high landslide susceptibility zones with 0–20%, 20–40%, 40–60%, 60–80%, and 80–100%, respectively, probabilities of occurrence. Based on the landslide susceptibility map, the moderate to very high landslide susceptibility zones occupied 60% area of the whole study area.

Comparison between the number of landslide occurrence and landslide susceptibility can be used to investigate the reliability of the developed susceptibility map. Table 3 shows the comparison between the numbers of landslide and the class of landslide susceptibility. It can be concluded that the occurrence of landslides matches with the class of susceptibility. From 47 landslides occurred in the study area, a 38% of the landslides occurred in the very high susceptibility class, a 43% of the landslides occurred in the high susceptibility class, a 17% of the landslides occurred in the moderate susceptibility class, one landslide occurred in the low susceptibility class and no landslide occurred in the very low susceptibility class.

Although the ratio of landslide occurrence appears to match with the susceptibility class, predictions of landslide occurrence in the high and very high landslide susceptibility classes were not too precise. Both classes of susceptibility show almost similar percentages of landslide occurrence. This may occur due to different local conditions, such as amount of rainfall, sunshine, and vegetation cover, that have not been examined in this study.

7 Conclusion

Despite the unsatisfactorily prediction of landslide occurrence in the high and very high sus-

Table 1: Quantification of data for landslide prediction by using LR analysis in this study.

| Parameters	Description	Susceptibility class to cause landslide	Relative Score for quantification
Landslide Occurrence	Yes		1
	No		0
Slope angle	0 ° - 26 °	Low	1
	26 ° - 40 °	Moderate	2
	> 40 °	High	3
Slope Lithology	Augite andesite	Very low	1
	Andesite breccias and lava andesite	Low	2
	Limestone	Moderate	3
	Sandstone	High	4
	Alluvial deposits	Very high	5
Distance from faults	0 – 500 m	Low	1
	500 - 1000 m	Moderate	2
	> 1000 m	High	3
Distance from lineaments	0 – 100 m	High	3
	100 – 200 m	Moderate	2
	> 200 m	Low	1
Ground water depth	0-2 m	High	3
	2-5 m	Moderate	2
	> 5 m	Low	1
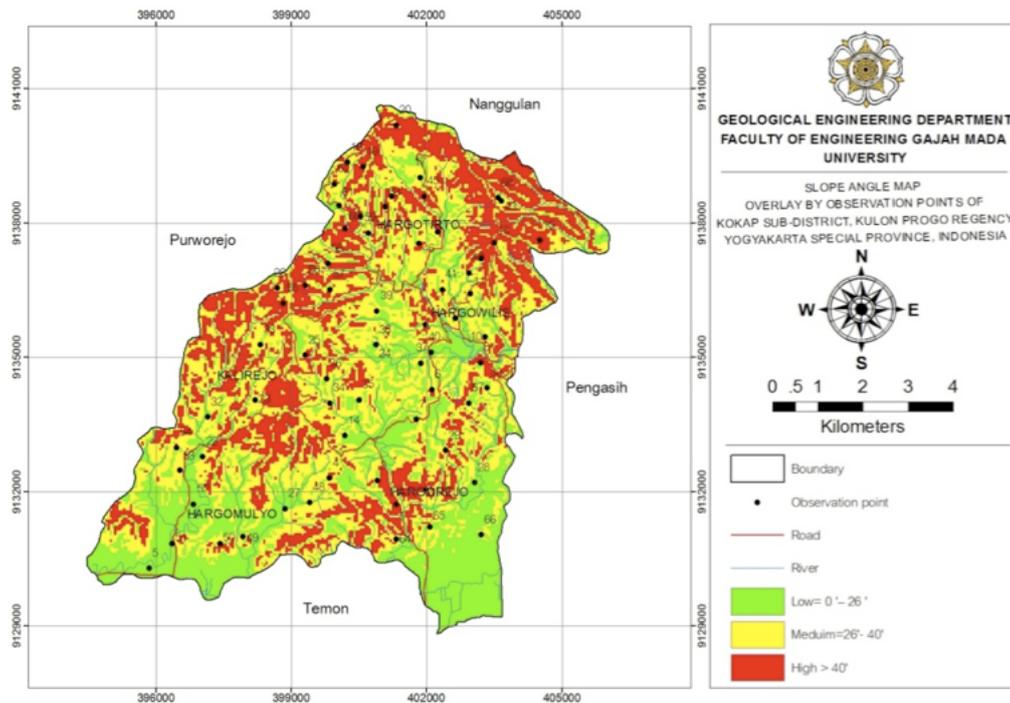
Figure 3: Slope angle distribution in the study area.
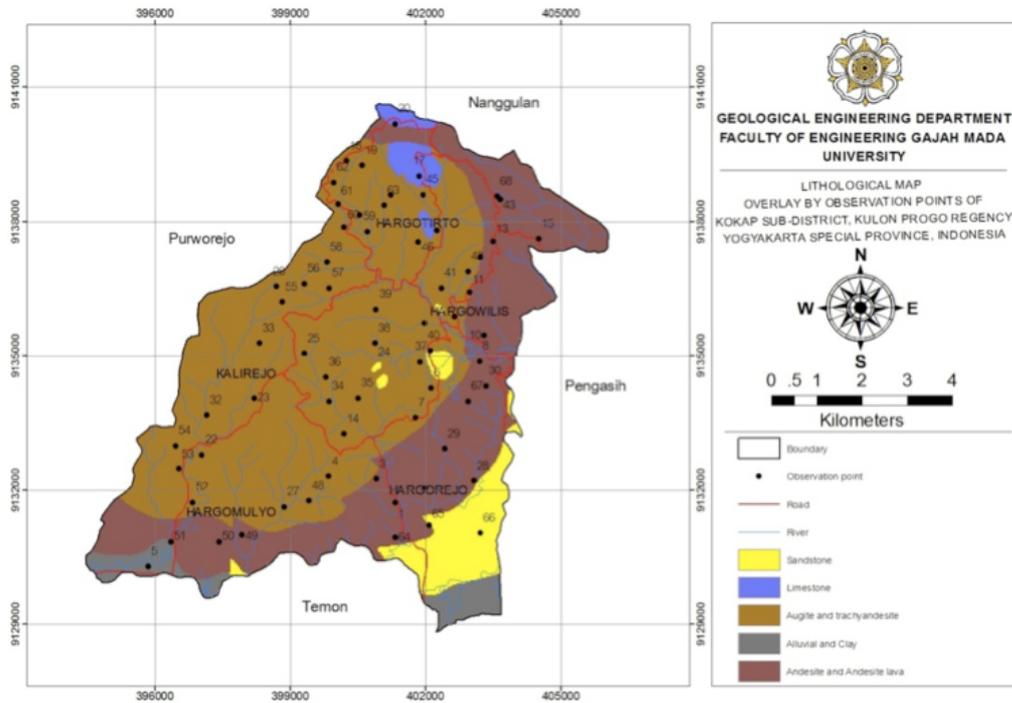
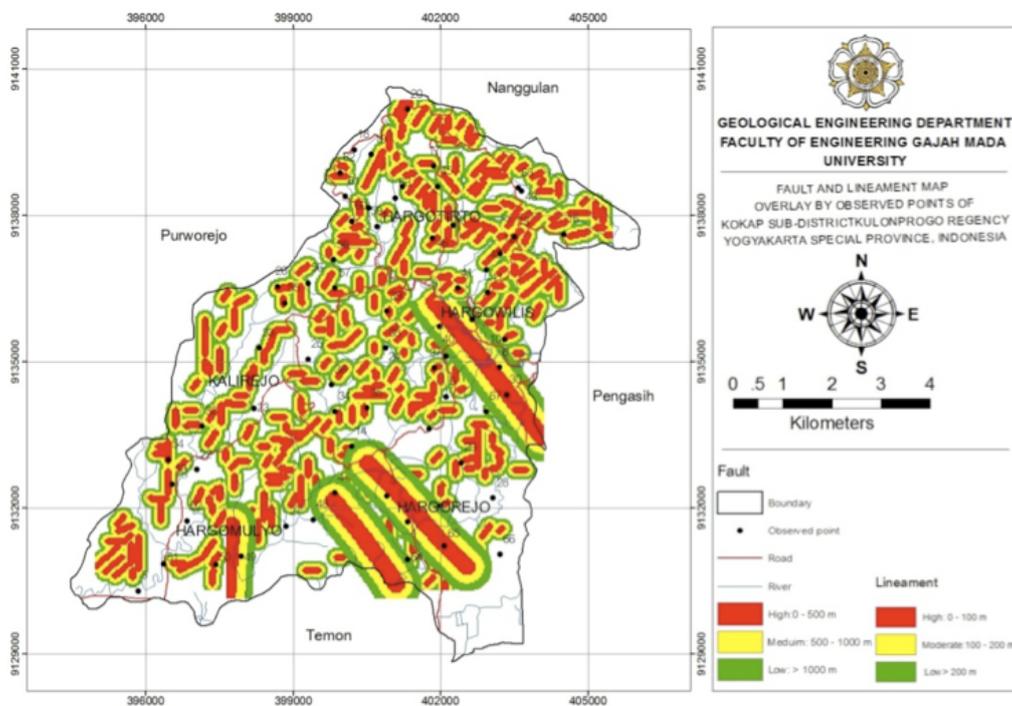Figure 4: Lithology distribution in the study area.

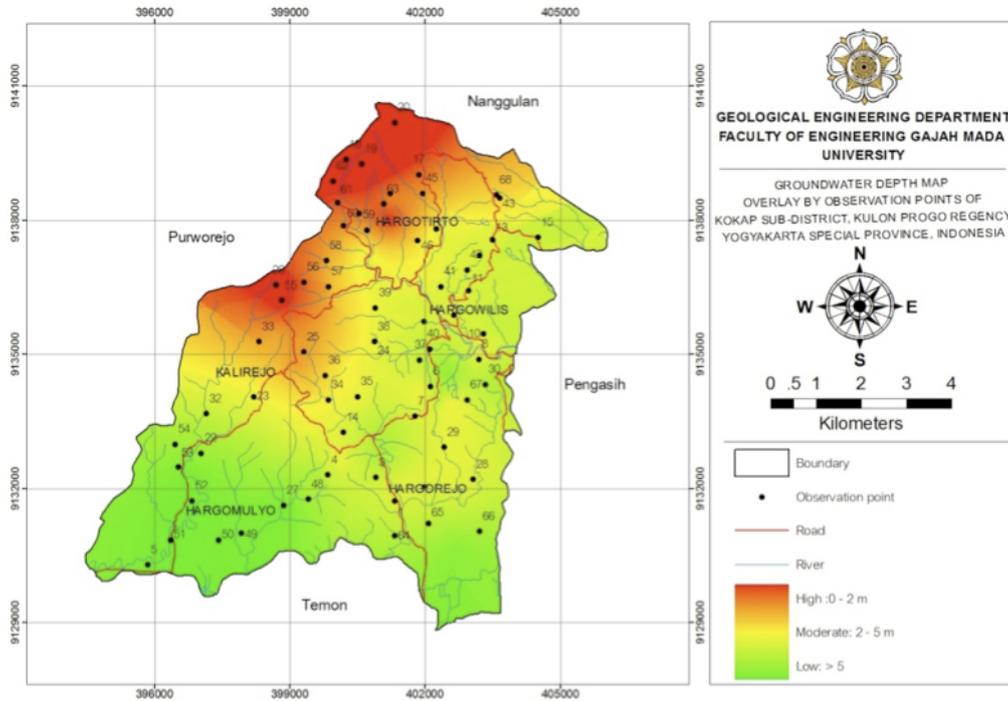Figure 5: Map showing distance from faults and lineaments in the study area.

Figure 6: Groundwater depth distribution in the study area.

Table 2: Logistic regression model coefficient and SE value for prediction of the probability of landslide occurrence in the study area.

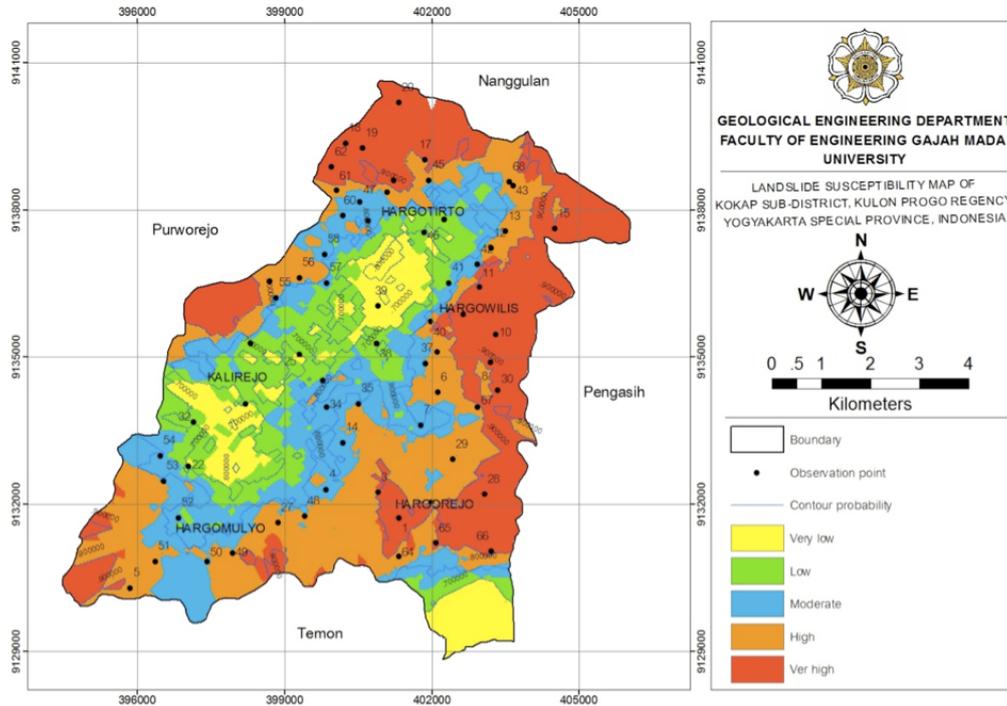| Model Variable		β	SE
Slope	1	-0.805	0.984
	2	-0.871	0.989
	3	-0.094	0.815
Lithology	1	0.672	0.781
	2	0.685	0.761
	3	1.379	0.566
	4	1.387	1.766
	5	0.748	1.698
Distance from faults/lineaments	1	0.861	1.019
	2	-0.619	1.011
	3	-0.212	1.019
Groundwater depth	1	0.232	1.177
	2	1.113	1.012
	3	0.876	2.032
Constant	α	0.876	

Figure 7: The LR landslide susceptibility map of the study area.

Table 3: Comparison between zones of susceptibility and number of landslide occurrence.

Class of Landslide susceptibility	Number of Landslide	Percentage landslide occurrence
Very Low	None	0
Low	1	2%
Moderate	8	17%
High	20	43%
Very High	18	38%

ceptibility classes, the LR method has been proven to be a reliable method for the development of landslide susceptibility map in the study area. Satisfactorily prediction may be obtained if more landslide causative factors are taken into account in the analyses, particularly factors related to the local conditions.

Aknowledgement

References

Bai, S., Wang, J., Lu, G., Zhou, P., Hou, S., Xu, S. (2010) GIS-based Logistic Regression for Landslide Susceptibility Mapping of the Zhongxian Segment in Three Gorges Area, China., Geomorphology, vol. 115, p. 23–31.

Chan, Y.H. (2004) Biostatistic 202: Logistic Regression Analyses, Singapore Medical Journal, vol. 45(4), p. 149–153.

Chau, K.T. and Chan, J.E. (2005) Regional Bias of Landslide in Generating Suceptibility Maps using Logistic Regression: Case of Hong Kong Island, Landslides, vol. 2, p. 280–290.

Das, I., Sahoo, S., Van Westen, C., Stein, A., and Hack, R. (2010), Landslide susceptibility assessment using logistic regression and its comparison with a rock mass classification system, along a road section in the northern Himalayas (India), Geomorphology, vol. 114, p. 627–637.

Hosmer, Jr., D.W., Lemeshow, S., and Sturdivant, R.X. (2013) Applied Logistic Regression, 3rd, John Wiley & Sons, 500p.

Karnawati, D. (2005) Bencana Alam Gerakan Massa Tanah di Indonesia dan Upaya Penanggulangannya. Yogyakarta, Jurusan Teknik Geologi, UGM.

Rahardjo, W, Sukandarrumidi and Rosidi, H.M.D. (1977) Geological map of Yogyakarta Quadrangle, Geological Survey of Indonesia, Bandung.

Rupert, M.G., Cannon, S.H., and Gartner, J.E. (2002) Using Logistic Regression to Predict the Probability of Debris Flows Occuring in Areas Recently Burned By Wildland Fires, Open-File Report of 03-500, USGS.

Syafri, I., Budiadi, E., and Sudradjat, A. (2013) Geotectonic Configuration of Kulon Progo Area, Yogyakarta, Indonesian Journal of Geology, vol. 8(4), p. 185–190.

Van Westen, C.J., Seijmonsbergen, A.C., and Mantovani, F. (1999) Comparing Landslide Hazard Maps, Natural Hazards, vol. 20, p. 137–158.

Varnes, D. J. (1978). Slope movement types and processes. In R.L. Schuster and R.J. Krizek (eds), Landslide: Analysis and control Transportation Research Board, National Research Council, Washington DC, 11–33.