MODELLING OF POPULATION EXPOSURE TO LANDSLIDE RISK IN SUKABUMI, INDONESIA USING GIS

S. Wiguna[1], J. Gao[2, *]

[1]School of Environment, University of Auckland, Auckland, 1142, New Zealand - swig831@aucklanduni.ac.nz
[2]School of Environment, University of Auckland, Auckland, 1142, New Zealand - jg.gao@auckland.ac.nz

KEY WORDS: Landslide exposure, Landslide susceptibility zoning, Landslide determinants; Population distribution

ABSTRACT:

Modelling population exposure to landslide risk is essential for mitigating the damage of landslides. This research aims to assess population exposure to the modelled landslide risk in the Sukabumi region, Indonesia. Also assessed in this study is the importance of 10 environmental variables and their spatial association with past landslide occurrence using the Weight of Evidence (WOE) method. The accuracy of the modelled landslide susceptibility is assessed using the AUC ROC method. Village level population was spatially redistributed via dasymetric modelling, and overlaid with the modelled landslide susceptibility map differentiated by the source zone and the runout zone. It is found that slope, curvature, and soil are the three most influential variables of landslides. The WOE method is able to achieve a similar success rate (0.877) and prediction rate (0.876) in modelling landslide susceptibility. In 2017, medium (114,588 ha) and high (106,337 ha) susceptibility levels were the two largest classes while low (94,778 ha), very high (52,560), and very low (51,910 ha) susceptibility classes are much less extensive. An absolute majority of the population faces a high (1,081,875 people or 38.98% of the total population), and a medium (1,036,080 people or 37.33%) level of landslide risk. Those facing a low (409,658 people or 14.76%), very high (168,193 people or 6.06%), and very low susceptibility (79,656 people or 2.87%) account for slightly more than one fifth of the total population. These findings demonstrate the critical role of GIS in assessing the exposure of population to landslide risk from a diverse range of variables.

1. INTRODUCTION

Landslides are one of the major disasters in Indonesia. Over the last decade, they have caused 1,591 fatalities and displaced over 200,000 people (BNPB, 2018). A landslide is controlled by its determinants, such as geological factors (e.g., lithology and structure), topography, soil, and land use. Topography influences shallow landslide initiation through both concentration of subsurface flow and destabilisation of slopes at a high gradient (Montgomery, Dietrich, 1994) In general, the steeper the slopes, the more likely they are to fall (van Westen et al., 2005). In terms of land cover, the role of vegetation in slope stability has been widely recognised (Reichenbach et al., 2014; van Westen et al., 2005). Trees are able to increase stability in two ways. First, they can retain a considerable amount of water that may cause landsliding. Second, their root systems increase shear resistance of the earth material mass, and through the creation of a negative pore pressure that increases soil cohesion. Therefore, an area with dense vegetation is considered to be stable (Reichenbach et al., 2014; van Westen et al., 2005).

Some of the aforementioned determinants, such as climate and land use, are dynamic and may change in a short period of time (Reichenbach et al., 2014). Such changes will affect land susceptibility to landslides, and the population exposed to landslide risk. So far population exposure to landslide hazard has been modelled from various at-risk components. Althuwaynee and Pradhan (2016) quantified the landslide risks in Kuala Lumpur, Malaysia from estimated population and infrastructure density. Landslide exposure has also been analysed via dasymetric cartography (Garcia et al., 2016). Pereira et al. (2016) and Zezere et al. (2017) quantitatively evaluated landslide risks at the municipal scale based on a rainfall triggering scenario by considering buildings and infrastructure in Portugal. However, nobody has modelled population exposure to landslide risk by considering its runout zone.

This study aims to assess regional population exposure to landslide risks in Sukabumi, Indonesia by considering a whole range of landslide-affecting variables and landslide runout zone. The specific objectives are: (a) to identify the important environmental factors that are critical to past landslide occurrence using Weight Of Evidence (WOE) analysis; (b) to evaluate the accuracy of modelled landslide susceptibility from its determinants; and (c) to assess population exposure to different levels of landslide susceptibility.

2. STUDY AREA

The study area is located in West Java Province at 106°21'35" - 107°19'30" E and 7°27'46" - 6°42'20" S (Figure 1), covering 4,210 km². Its climate is characterised by a moderate humidity, moderate wind speed, and high rainfall (Marpaung, 2006) with an annual temperature of 18°-30° C. Annual rainfall ranges from 2,000 to 4,000 mm, with more in the north than in the south. Precipitation is controlled mainly by monsoonal winds with distinct wet and dry seasons. Major land covers include forest (15.62%), mixed farmland (66.85%), built-up areas, paddy fields, and water. Natural forest is located mainly in the national parks in the north. Built-up areas (8.52%) and paddy fields (8.17%) are mainly located in the capitals, Sukabumi city and Pelabuhanratu. They also include villages in rural areas. Paddy fields are generally located in low lands.

* Corresponding author

Elevation ranges between 0 and 3,017 m a.s.l. The low-lying region is mainly located in the southwest. The north is mountainous with a high elevation. Slope gradient ranges from 0°-63.4°. Generally, steeper slopes are coincidently located in the mountainous region in the north, the middle, and in the southeastern river valleys. The 53 rock types fall into four major classes of Jampang, Tapos breccia, Bentang and Beser formation. Among them, less permeable breccia rocks facilitate landslides in some locations. These impermeable layers often become slip surfaces for a weathered rock found atop of them (Kasbani, 2019; PVMBG, 2016). Five out of the 18 soil types are dominant, accounting for 82% of the total area. In particular, Typic Hapludalfs in the middle and toward the south is well-developed, and considered as one of the most susceptible to landsliding in Lampung, Indonesia (Arifin et al., 2006).

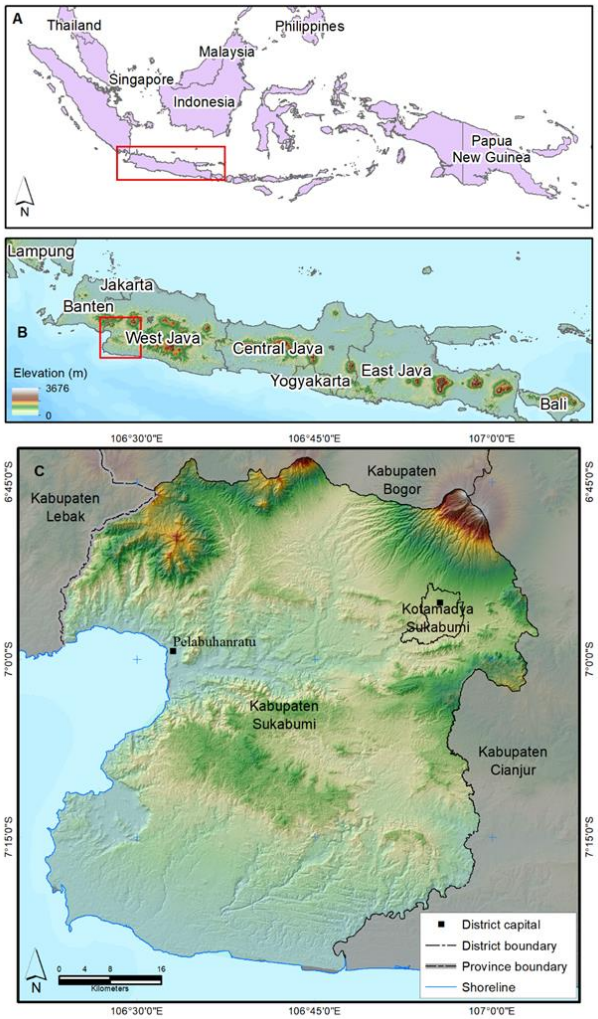

Figure 1. Location of the study area and its topography.

The study area has a population of 2.7 million in 2017 (BPS, 2018). People live mostly in urban areas. In rural areas, they are settled commonly along transportation routes (Izhom, 2012). Some settlements are located in slopes as steep as 45°, making them highly exposed to landsliding risk. This area is selected for study because it has suffered from severe landslides. The most common landslides take the form of slides on hillslopes in the mountainous areas, and debris flows along roads without a proper drainage system (Sugianti *et al.*, 2016). Rainfall and earthquakes are two common landslide triggers, with rainfall being the dominant one (Kasbani, 2019).

3. METHODOLOGY

3.1 Datasets used

A wide range of data are used for landslide susceptibility modelling, including landslide inventory, digital elevation model (DEM), slope loading measured by topographic wetness index (TWI), stream power index, soil, lithology, land cover, and village-level population. The DEM has 1 arc second (about 30 m) resolution. A SPOT 7 image of 2017 was sourced from LAPAN (Lembaga Penerbangan dan Antariksa Nasional/National Institute of Aeronautics and Space). It has a spatial resolution of 1.5 m for the panchromatic band and 6 m for the multispectral bands. Additional public domain images, including Landsat and high-resolution commercial satellite images of GeoEye and SPOT were accessed via Google Earth (GE) (Fisher, 2012). Ancillary data such as a 1:25,000 topographic map of 2013 were obtained from the portal http://tanahair.indonesia.go.id/portal-web in the shapefile format. It comprises various layers, including land cover, roads, rivers, administrative boundary up to the village level. The land cover map shows fourteen classes that were regrouped into six: forest, built-up (settlement) area, mixed farmland, paddy fields, water body, and other classes. A soil map published in 2016 by the Ministry of Agriculutre at a scale of 1:50,000 shows 18 soil types. Population data in 2017 was collected from the BPS (*Badan Pusat Statistik*/Central Bureau of Statistics) at the village level.

Landslide data of 2010-2018 were acquired from the inventory data maintained by the Centre of Volcanology and Geological Hazard Mitigation and District Disaster Management Authority. These government agencies retain mainly those landslides that have affected humans. Thus, the data often do not encompass landslides in remote areas. The landslide information presented in the tabular format shows the location, time, and impacts of landslides on humans and their livelihood, but not landslide magnitude. In order to maintain a high location accuracy, only those marked with a village name were excluded from this study. The inventory data of 257 retained landslides were supplemented with 155 more landslides published in Andriono (2012); Badrudjaman (2016); Hermansyah (2015); Izhom (2012); Ristya (2018); and Sugianti et al. (2016). They all represented only the landslide location by a pair of coordinates. There is no information on the date, magnitude, extent, and impacts of the recorded landslides. This lack of details makes it difficult to estimate the probability of frequency and magnitude that are required for a complete landslide risk assessment.

In order to augment the landslide data, more landslides were visited in the field during 22-24 September 2018. Field inspection also aimed at ensuring the quality of the landslide inventory data. The field route of inspection was determined based on accessibility and landslide vulnerability. Most of the landslides visited were found in slope-cut areas for road construction. In the field, the observed landslides were then logged using GPS. To minimise double counting, all the landslides were plotted into a GPS base map, with all duplications removed, resulting in 11 unique landslides.

3.2 Variables studied

Variables studied include topography (slope, aspect, curvature); stream power index (SPI), TWI, and distance to rivers which indicates hydrological influence; lithology, soil types, land cover, and distance to roads (Table 1). These variables are studied as they are the major known drivers and constraining factors of landslides.

Variables	Total C*	Rank	Attributes having high C value
Curvature	0.52	1	Convex and Concave
Soil type	0.14	2	Lithic Udorthents (1.84), Sulfic Endoaquepts (1.64), Typic Hapludands (1.41)
Slope	0.13	3	Above 18.7°
Lithology	0.03	4	Badui formation (2.82), Cikotok formation (2.69), Lava flow (2.78), Citorek tuff (1.95), Older volcanic deposits (1.79), and Dacite (1.02).
Land cover	-0.13	5	Natural forest (0.99) and built-up areas (0.49)
SPI	-0.15	6	6.1 – 7.3, > 14
TWI	-0.28	7	Small value of TWI: -0.5 – 1.7
Distance to rivers	-0.35	8	High value: > 120 m
Distance to roads	-0.55	9	Low value: < 30 m
Aspect	-1.57	10	West-facing slopes

*: C = Contrast; SPI= Stream power index, TWI = Topographic wetness index

Table 1 Landslide variables studied and their spatial association with landslide events as judged by a high total C value.

Slope gradient, aspect, and terrain curvature were derived from the DEM. The calculated slope gradient (°) was classified into nine classes using natural breaks. Slope aspect was expressed in nine classes of flat (-1), and other eight directions at an interval of 45°. Overall curvature (both profile and cross-section) values were classified into 9 classes using natural breaks. Stream Power Index (SPI) was derived from the DEM using Eq. (1) (Danielson, 2013):

$$SPI = \ln[(facc + 0.001)* (slope / 100 + 0.001)] \quad (1)$$

where ln = natural logarithm
facc = flow accumulation
slope = slope gradient (%)

The SPI values were categorised into nine classes using natural breaks.

TWI was produced from the DEM, and classified into nine classes using natural breaks. Buffer analysis was used to approximate the influence of roads and rivers on landslide occurrences. Five buffer zones were generated along all roads and rivers at an interval of 30 m, namely, 30 m, 60 m, 90 m, 120 m, and above 120 m. The 30 m interval is identical to the DEM grid size. All the collected remotely sensed datasets were analysed in ArcGIS and remote sensing (RS) systems. ArcGIS was used for data pre-processing, data analyses, as well as results presentation.

3.3 Mapping of land covers and landslides

The 2017 land cover was mapped by updating the 2013 land cover map, over which the 2017 SPOT satellite image was overlaid via on-screen visual interpretation. The image has been pan-sharpened by merging the 6 m multi-spectral bands with the 1.5 m panchromatic band. The changes in land cover between 2013 and 2017 were then digitised onscreen in accordance with the same land cover classes as in 2013.

More landslides of 2012-2018 were visually interpreted from GE high-resolution satellite images as polygons. Because of the minimum mapping unit issue, this interpretation may cause the miss of small landslides retained in the landslide inventory data and published studies, even though it enables identification of landslides in remote areas. All 98 observable and unique landslides were digitised on screen. The output from the GE system is in the Keyhole Markup Language format with the geographic coordinate system. It was subsequently converted to the ESRI shapefile format, and rasterised at a grid size of 30 m × 30 m. In many cases, the rasterisation caused spatially extensive slope failures to be represented by multiple grid cells. In total, there are 884 landslide cells (521 landslide paths) from the inventory data, published studies, field inspection, and visual interpretation of GE images after duplications are removed.

Also rasterised are the vector datasets such as geology, soil, and land cover, as well as intermediate vector data from buffer analysis along roads and rivers at 30 m resolution. Of all the data types, only landslide data were preserved in the point format because the WOE calculation requires the input to have this format. Finally, all the raster data were standardised to the same spatial extent, cell size (30 m), and coordinate system (UTM Zone 48S, Datum WGS 1984).

3.4 Land susceptibility modelling

Landslide susceptibility was differentiated into source zone and runout zone susceptibility. The former was evaluated using the WOE model, and the latter was estimated using the multiple flow direction with the source zone as the input. The process of landslide susceptibility modelling is illustrated in Figure 2.

3.4.1 Weighting landslide variables: The weight assigned to different landslide variables is determined using one of the famous bivariate statistics analysis methods called WOE. It uses Bayesian statistics to calculate the strength of the spatial association between a training set (past landslide events) and predictors (determinants), and assign weights to them to produces a susceptibility map (Schmitt, 2010). WOE has been used to assign weights for each predictor's attribute based on landslide densities (Corominas et al., 2013). All the 10 variables shown in Figure 2 are considered as the potential predictors of landslides. These variables are presumably independent of each other (Bonham-Carter, 1994; Neuhäuser et al., 2011).

WOE is based on a Bayesian rule in a log-linear form using the prior and posterior probability. The prior (unconditional) probability is the probability of a past event in a given period of time (e.g., the spatial probability of a landslide). The posterior (conditional) probability is the change in probability owing to the additional information given to re-evaluate the prior

probability (Samodra et al., 2017). For example, additional use of lithology as a predictor may change the probability of landslide occurrence based on slope gradient.

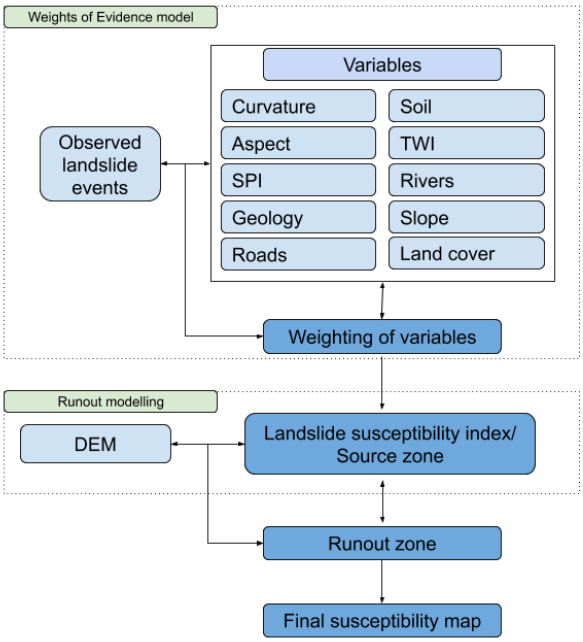

Figure 2. Flowchart of modelling landslide susceptibility.

By overlaying the landslide layer with each of the landslide predictor layers, a spatial association between the landslides and a unique attribute of the predictor (evidence) can be established according to their significance for the landslide occurrence (Neuhäuser et al., 2011). WOE is expressed in a pair of weights, positive weight W^+ (Eq. 2) and negative weight W^- (Eq. 3). (W^+ indicates the presence of spatial association, and W^- shows the absence of spatial relationship. The magnitude of the weight W suggests the level of spatial association. The evidence map is converted to a binary predictor map of presence and absence. As indicated in Bonham-Carter (1994), the weight is then calculated for each landslide predictive factor (F) based on the presence or absence of a landslide (L) within a grid cell.

$$W+ = ln \frac{P\{F|L\}}{P\{F|\bar{L}\}} \quad (2)$$

$$W- = ln \frac{P\{\bar{F}|L\}}{P\{\bar{F}|\bar{L}\}} \quad (3)$$

where P = probability
F = presence of landslide predictive factor
$\bar{F}$ = absence of a landslide predictive factor
L = presence of a landslide
$\bar{L}$ = absence of a landslide

Three special combinations of weights need particular consideration: (a) if W^+ is positive and W^- is negative, the predictor is favourable for landslide occurrence; (b) if W^+ is negative but W^- is positive, the predictor is not favourable for landslide occurrence; and (c) if $W^+ = 0$ and $W^- = 0$, the predictor is not correlated with landslides.

The relationship can also be measured by weight contrast C, which is a difference between W^+ and W^-:

$$C = W^+ - W^- \quad (4)$$

A positive C indicates the existence of a spatial association between the landslide and a predictive factor, and *vice versa*. The positive spatial association indicates that more landslides occur in a unique attribute of a predictor than would be expected due to chance match. In contrast, a negative spatial association shows that there are fewer landslide events in the attribute than expected due to chance match (Bonham-Carter et al., 1988). The total of contrast values forms the WOE-based landslide susceptibility map. It is expressed as:

$$LSI = \sum Fc \quad (5)$$

where Fc = attribute contrast of each variable

Treated as the weighting factor, Fc was used as multiple factors for each variable. All the weighted attributes of the same variable are then summed to form the susceptibility for the landslide source zone. The output landslide susceptibility map was classified into five levels using natural breaks: very high, high, moderate, low, and very low.

3.4.2 Runout zone modelling: Runout modelling utilised the landslide source zone obtained from the previous step and the DEM. The source zones refer to the sources of flow of mobilised earth material, and their delineation was facilitated by the DEM. In the runout modelling, the area affected by landsliding is based on the propagation of the mobilised material, e.g., a combination of flow direction and runout distance. While the former calculates the flow path of the displaced material, the latter determines the runout distance (Horton et al., 2008). Flow direction was generated through the multiple flow direction algorithm D-infinity (Tarboton, 1997), and the runout distance was generated using the D-infinity Avalanche Runout in TauDEM, an ArcGIS extension downloadable at http://hydrology.usu.edu/taudem/. Both algorithms are underpinned by the assumption that the debris flows stop when they reach a terrain of 5°. This runout modelling was implemented for each of the five susceptibility levels, resulting in five classes of runout susceptibility mapping.

Landslide susceptibility is mapped by integrating the slope failure susceptibility as the source of slides and the modelled landslide runout path (Corominas et al., 2003; Guinau et al., 2007; Kritikos, Davies, 2014). In the mapping, if a site is characterised by more than one susceptibility level, then it receives the highest level. For example, if a cell has a very low susceptibility but falls into the runout path in the moderate and low susceptible zones, then it receives an overall value of moderate susceptibility. The final landslide susceptibility map comprises of five classes of very low, low, medium, high and very high susceptibility.

The produced landslide susceptibility map was validated for its accuracy and predictive capacity (Chung, Fabbri, 2003) by comparing with the real-world datasets (observed landslide events). The landslide cells (884) were divided into two groups randomly. The first group (70% or 619 cells) was used for landslide susceptibility mapping, so-called training points. The second group of 30% (265 cells), so-called testing points, was used for validation.

3.5 Landslide exposure assessment

Exposure assessment aims to estimate the number of people living in zones having various levels of landslide susceptibility. Population exposure to landslide risk is determined through spatial intersection of the modelled landslide susceptibility map with population distribution (Corominas et al., 2013; Pellicani et al., 2013). The newly created binary map of susceptibility is overlaid with the dasymetrically redistributed population map (multiplication) to extract the total population only in the targeted susceptibility level. This process is repeated four more times, each time for one of the four remaining susceptibility classes.

4. RESULTS

4.1 Influence of landslide variables

Landslides in the study area tend to occur on steep slopes. Approximately, 60% of them occur on slopes with a gradient > 18.7° (Figure 3). Such terrain only occupies <20% of the study area. Landslide occurrence is more common at steeper terrain. The association between the two changes from positive to negative at 18.7°, indicating that slopes gentler than 18.7° are unfavourable for landslides while slopes steeper than this threshold are more prone to slope failures.

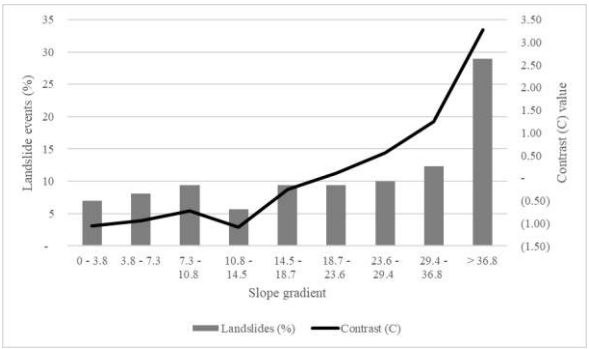

Figure 3 Distribution of landslide events (left vertical axis) and contrast C value (right vertical axis) with slope gradient (°).

The overall curvature has a value ranging from -10 (concave) to 2.5 (convex) (Figure 4). A larger curvature is more favourable for landslide occurrence regardless of its sign. However, a negative curvature bears a slightly stronger relationship with landslide occurrence than a positive curvature. Besides, the lower the curvature (both positive and negative), the more unfavourable it is for slope failure occurrence. The relationship between curvature and landslides becomes negative over the curvature range of -0.8 to 0.7.

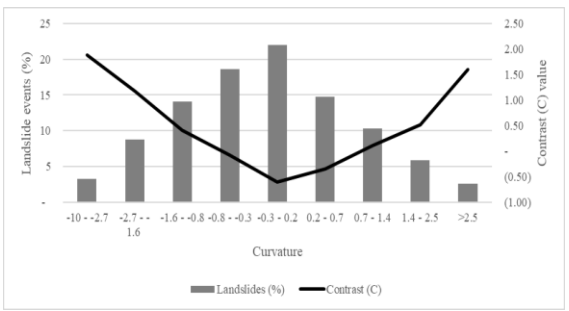

Figure 4 Landslide distribution (%) (left vertical axis) and contrast value (right vertical axis) in different curvature ranges.

Landslides have occurred in all aspects, including the flat surface (Figure 5). However, some aspects have a negative spatial relationship with the landslide events. In general, slopes facing the east tend to be unfavourable for landslide occurrence, including northeast-, east-, and up to the southeast-facing slopes, judging from their negative contrast value. The exception is the southeast aspect that has a positive value (0.13). In comparison, west-facing slopes, including northwest, west, and southwest, all have a positive contrast value, indicating that these aspects are favourable for landslide occurrence.

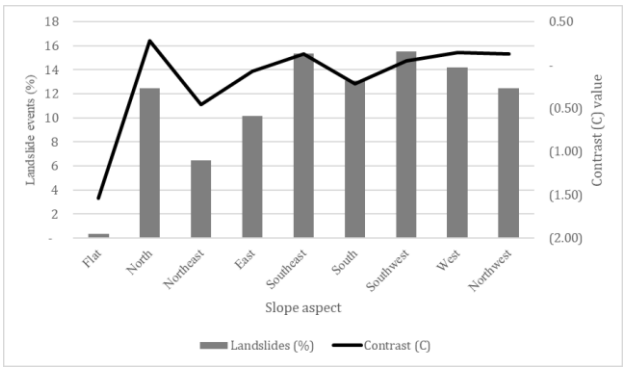

Figure 5 Landslide distribution (%) (left vertical axis) and contrast value (right vertical axis) in different slope aspects.

Of the TWI range of -0.5 – 19, only two classes (-0.5 – 1.7 and over 12.8) are favourable for landslide occurrence. While the first class has a C value of 0.97, the second class has a C close to 0. The majority of landslides occurred in the first three classes (Figure 6). However, since these classes cover a vast ground area, the spatial relationship between landslides and TWI value is indistinct. It can be inferred that topographic wetness does not play a critical role in landsliding in the study area.

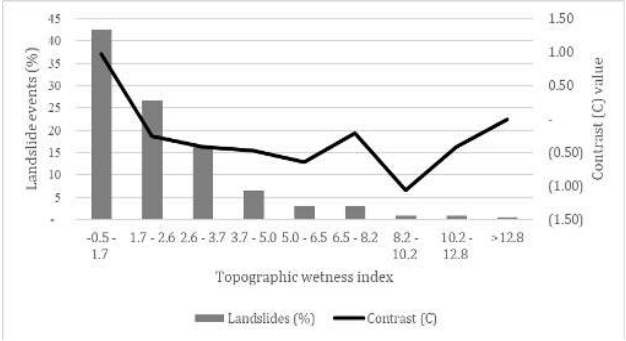

Figure 6 Landslide distribution (%) left vertical axis) and contrast value (right vertical axis) in relation to topographic wetness.

Generally, a low SPI is more favourable to landsliding than a higher magnitude (Figure 7). The threshold value for this favourable-unfavourable influence is SPI=8.7. Although there are some exceptions, the ≤8.7 values tend to have a positive association with the landslides. In contrary, the association becomes negative at SPI>8.7. The exceptions appear in the classes of -0.23 to 4.5 and 11.9 to 14.0, which have a negative value in a positive trend and a positive association in a negative trend, respectively. The majority of landslides is found in the first four SPI classes which comprise a major proportion of the study area, even though the positive SPI value for each class is generally lower than 1.

Landslide occurrence is negatively associated with proximity to rivers within 0-30 m and up to the distance of 90-120 m (Figure 8). Beyond 120 m, the C value is positive. In other words, more landslides have occurred in the area beyond 120 m from drainage systems. This indicates that the lateral erosion triggered by streams is not significant in the present study.

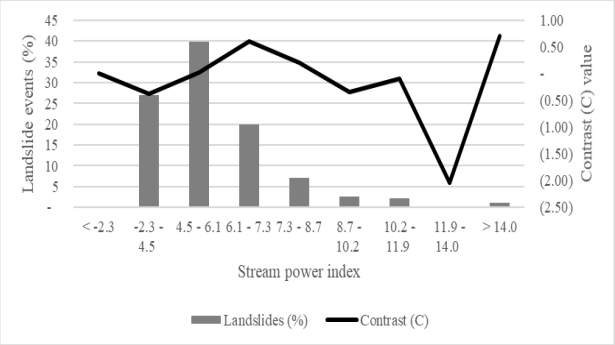

Figure 7 Landslide distribution (%) (left vertical axis) and contrast value (right axis) in relation to stream power index.

Some soil types have a strong spatial association with landslide occurrence than others. Most soil types are unfavourable for landslide occurrence as their C value is negative (magnitude range: -0.05 to -1.5). Among the approximately 53 rock types, some rock types have a strong spatial relationship with the landslide events, such as Badui formation, Cikotok formation, Dacite, Lava flow, Citorek tuff, and Older volcanic deposits. Other rock types have a weaker positive association with the landslides, and some others even have a negative relationship with the landslide occurrence. The highest negative correlation occurs to the alluvium and coastal deposits with a magnitude of -1.89.

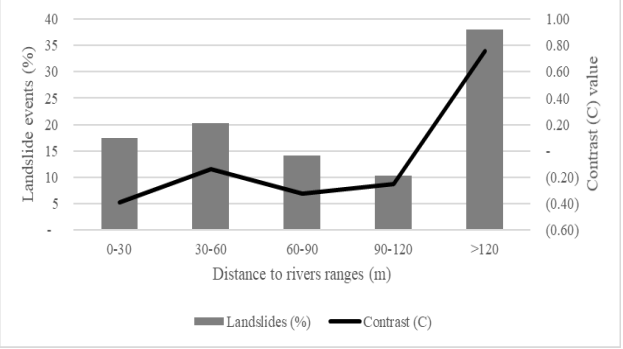

Figure 8 Landslide distribution (%) (left vertical axis) and contrast value (right axis) in distance to rivers ranges.

Among the six types of land cover, water, mixed farmland, and paddy fields have a negative C value. In contrast, forest and built-up have a positive relationship with landslides. Both negative and positive values are < 1. For example, forest cover has the highest positive value of just only 0.99. Thus, the overall relationship between land cover and landslide occurrence in this study is not strong. Mixed farmland is associated with the largest number of landslide cells (300), probably because it occupies a large proportion (66%) of the study area. In other words, fewer landslide cells coincide with forest cover than would be expected due to chance match.

Unlike proximity to rivers, areas closer to roads are considered to be more favourable for landslide occurrence. According to the C value, areas between 0-30 m from roads have a positive relationship with the landslide events. Landslides are more frequent along roads due to inappropriate cutting of slopes during construction (Dahal et al., 2007; Samodra et al., 2017), or due to the vibrations generated by vehicles that may trigger landslides (Samodra et al., 2017). Furthermore, areas beyond 30 m from roads are unfavourable for landslide occurrence, suggesting that construction and vehicle-triggered vibration of the ground cease to have any detectable impact on terrain stability or its vulnerability to landslides.

Overall, curvature, soil type and slope are the most influential to landsliding (Table 1), all having a positive C value with past landslide occurrence. On the other hand, aspect, distance to rivers and roads all have a relatively large negative C value, suggesting that they are conducive to landsliding.

4.2 Accuracy of the modelled landslide susceptibility

The accuracy of the landslide susceptibility model is assessed using the area under the curve (AUC) Receiver Operating Characteristics (ROC) method. The ROC is basically a comparison of true positive rate (TPR) and false positive rate (FPR). True false refers to the number of landslide cells falling in safe zones, while the false true rate indicates the number of landslide cells falling in unsafe zones. Both TPR and FPR are then used to calculate the AUC rate in RStudio. The AUC ROC can be applied to assess the success rate and the predictive rate. The success rate measures the agreement between the training points and the landslide susceptibility map, and quantifies the goodness of fit between the two sets of data assuming that the model is correct. It can yield the success rate of the susceptibility model through comparing the training sample data with the landslide susceptibility map. The success rate indicates how well the model separates the landslides among different susceptibility classes. The prediction rate is obtained from a comparison between the landslide susceptibility map and validation data (Vakhshoori, Zare, 2018). It is indicative of the predictive power of a model. The prediction rate measures the agreement between the training data and the susceptibility map, and is indicative of the predictive capability of a model (Chung, Fabbri, 2003; Vakhshoori, Zare, 2018). The AUC value ranges from 0 to 1. The accuracy or the predictive rate increases as the rate approaches the maximum value.

The assessment shows a similar success rate (0.877) and a prediction rate of 0.876. The results are consistent with those of previous findings by Althuwaynee and Pradhan (2016) and Pradhan et al. (2010). They achieved a slightly higher success rate than the prediction rate. Overall, the accuracy assessment indicates that the model is of excellent accuracy (Hosmer Jr et al., 2013).

4.3 Population exposure to landslide risk

The dasymetric population distribution map indicates that areas of a large population are situated in the northeast, up to the north, and in the central west (Figure 9). The total population is dasymetrically modelled at 2,775,463, the same as the statistic figure. However, dasymetric modelling represents more realistic population distribution than the conventional choropleth maps (Garcia et al., 2016) while still retaining the original statistics figure.

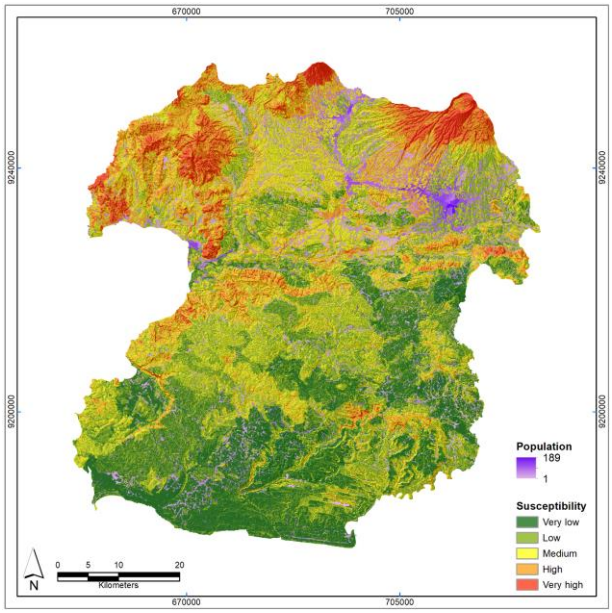

Figure 9. Level of population exposure to landslide risk in relation to population density in 2017.

The terrain highly susceptible to landslides is located at the top of hills (Figure 9). In comparison, the runout zone of the very high susceptibility shows that the possibly devastated area by the landslides is not confined only to the top of hills. Besides, it may also affect the downslope region that has a lower susceptibility. Therefore, when these two types are combined, the final susceptibility class may have a larger area than that of the slope failure susceptibility class. The northern part of the study area is more susceptible to landslides than the southern part. The high and very high susceptibility areas are located in the northern study area, especially the northernmost, the northwest, and the northeast. Central Sukabumi, the west, and a small portion in the southeast all have a high and very high susceptibility level.

Quantitatively, one third of the study area is of the medium landslide susceptibility level. The proportion of the very low class is nearly similar to that of the very high susceptibility classes at approximately 12% of the study area. Both the low and high susceptibility classes have nearly the same area of approximately 22.56 and 25.31%. The medium susceptibility level is the most widespread (114,588 ha). The second highest area has a high susceptibility (106.337 ha). The other three classes are less common in the descending order of low (94,778 ha), very high (52,560 ha), and very low (51,910 ha).

Quantitatively, in 2017, approximately 45% of Sukabumi population lived in areas having a (very) high landslide susceptibility. The majority of people live in areas of high landslide risk (38.98%). The least populated area has a very low landslide risk (2.87%). The second, third, and fourth most inhabited areas have a medium (37.33%), low (14.76%), and very high (6.06%) susceptibility level, respectively.

5. CONCLUSIONS

Curvature, soil type and slope gradient play a critical role in landslide susceptibility occurrence in the study area. While curvature and slope gradient indicate a high contribution of topography to slope instability, soil type highlights the role of surface material in landslide occurrence. These three most influential variables of landslides all have a positive contrast value over 0.5 (Table 1). In particular, the landside occurrence in the study area is closely connected with concave and convex slopes. The WOE method achieved a similar success rate (0.877) and prediction rate (0.876) in modelling landslide occurrence from a diverse range of environmental and land use variables. In 2017, the medium susceptibility level holds the majority area (114,588 ha). The second highest area has a high susceptibility (106.337 ha). The other classes have a descending order of low (94,778 ha), very high (52,560 ha), and very low (51,910 ha). The majority of people live in areas of high landslide risk (38.98% or 1,081,875 people). The least populated area has a very low landslide risk (2.87% or 79,656 people). The second, third, and fourth highest inhabited areas have a medium (37.33% or 1,036,080 people), low (14.76% or 409,658 people), and very high (6.06% or 168,193 people), and very low susceptibility (2.87% or 79,656 people). Such a high level of exposure means that future settlement development must be restricted to low-lying gentle terrain to minimise the potential damage by and loss of properties and lives to landslides.

REFERENCES

Althuwaynee, O. F., & Pradhan, B. 2016. Semi-quantitative landslide risk assessment using GIS-based exposure analysis in Kuala Lumpur City. *Geomatics, Natural Hazards and Risk, 8*(2), 706-732. 10.1080/19475705.2016.1255670

Andriono, B. 2012. *Wilayah Rentan Tanah Longsor di Sepanjang Alur Ci Tarik DA Ci Tarik Kabupaten Sukabumi.* (Bachelor), Universitas Indonesia, Depok.

Arifin, S., Carolita, I., & Winarso, G. 2006. Implementasi Penginderaan Jauh dan SIG untuk Inventarisasi Daerah Rawan Bencana Longsor (Propinsi Lampung). *Jurnal Penginderaan Jauh, 3*(1), 77-86.

Badrudjaman, S. H. A. 2016. *Bahaya Longsor di Daerah Vulkanik Kabupaten Sukabumi Bagian Utara.* (Bachelor), Institut Pertanian Bogor, Bogor.

BNPB. (2018). Data dan Informasi Bencana Indonesia. Retrieved 23 December 2018, 2018, from http://dibi.bnpb.go.id/

Bonham-Carter, G. 1994. Geographic Information Systems for Geoscientists, Volume 13: Modelling with GIS (Computer Methods in the Geosciences).

Bonham-Carter, G., Agterberg, F., & Wright, D. 1988. Integration of geological datasets for gold exploration in Nova Scotia. *Photogrammetric engineering and remote sensing, 54*(11), 1585-1592.

BPS. (2018). *Sukabumi Municipality in Figures 2017*

Chung, C.-J. F., & Fabbri, A. G. 2003. Validation of spatial prediction models for landslide hazard mapping. *Natural Hazards, 30*(3), 451-472.

Corominas, J., Copons, R., Vilaplana, J. M., Altimir, J., & Amigó, J. 2003. Integrated landslide susceptibility analysis and hazard assessment in the principality of Andorra. *Natural Hazards, 30*(3), 421-435.

Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J. P., Fotopoulou, S., . . . Smith, J. T. 2013. Recommendations for the quantitative analysis of landslide risk. *Bulletin of Engineering Geology and the Environment* 10.1007/s10064-013-0538-8

Dahal, R. K., Hasegawa, S., Nonomura, A., Yamanaka, M., Masuda, T., & Nishino, K. 2007. GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. *Environmental Geology, 54*(2), 311-324. 10.1007/s00254-007-0818-3

Danielson, T. 2013. Utilizing a high resolution Digital Elevation Model (DEM) to develop a Stream Power Index (SPI) for the Gilmore creek watershed in Winona County, Minnesota.

Garcia, R. A. C., Oliveira, S. C., & Zêzere, J. L. 2016. Assessing population exposure for landslide risk analysis using dasymetric cartography. *Natural Hazards and Earth System Sciences, 16*(12), 2769-2782. 10.5194/nhess-16-2769-2016

Guinau, M., Vilajosana, I., & Vilaplana, J. M. 2007. GIS-based debris flow source and runout susceptibility assessment from DEM data? A case study in NW Nicaragua. *Natural Hazards and Earth System Science, 7*(6), 703-716.

Hermansyah. 2015. *Wilayah Bahaya Longsor Menggunakan Metode SINMAP, Studi Kasus: Kecamatan Simpenan, Kabupaten Sukabumi, Provinisi Jawa Barat.* (Bachelor degree), Universitas Indonesia, Depok.

Horton, P., Jaboyedoff, M., & Bardou, E. (2008). *Debris flow susceptibility mapping at a regional scale.* Paper presented at the Proceedings of the 4th Canadian Conference on Geohazards: from causes to management.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398): John Wiley & Sons.

Izhom, M. B. 2012. *Kerentanan Wilayah Tanah Longsor Di Daerah Aliran Ci Catih, Kabupaten Sukabumi, Jawa Barat.* (Bachelor), Universitas Indonesia, Depok.

Kasbani. 2019. *Laporan singkat pemeriksaan gerakan tanah di Kecamatan Cisolok, Kab. Sukabumi, Provinsi Jawa Barat*. Bandung: Pusat Vulkanologi dan Mitigasi Bencana Geologi.

Kritikos, T., & Davies, T. 2014. Assessment of rainfall-generated shallow landslide/debris-flow susceptibility and runout using a GIS-based approach: application to western Southern Alps of New Zealand. *Landslides, 12*(6), 1051-1075. 10.1007/s10346-014-0533-6

Marpaung, F. 2006. *Penyusunan Model Spasial untuk Memprediksi Penyebaran Malaria (Studi Kasus Kabupaten Sukabumi, Jawa Barat).* (Bachelor), Institut Pertanian Bogor, Bogor.

Montgomery, D. R., & Dietrich, W. E. 1994. A physically based model for the topographic control on shallow landsliding. *Water resources research, 30*(4), 1153-1171.

Neuhäuser, B., Damm, B., & Terhorst, B. 2011. GIS-based assessment of landslide susceptibility on the base of the Weights-of-Evidence model. *Landslides, 9*(4), 511-528. 10.1007/s10346-011-0305-5

Pellicani, R., Van Westen, C. J., & Spilotro, G. 2013. Assessing landslide exposure in areas with limited landslide information. *Landslides, 11*(3), 463-480. 10.1007/s10346-013-0386-4

Pereira, S., Garcia, R. A. C., Zêzere, J. L., Oliveira, S. C., & Silva, M. 2016. Landslide quantitative risk analysis of buildings at the municipal scale based on a rainfall triggering scenario. *Geomatics, Natural Hazards and Risk, 8*(2), 624-648. 10.1080/19475705.2016.1250116

Pradhan, B., Oh, H.-J., & Buchroithner, M. 2010. Weights-of-evidence model applied to landslide susceptibility mapping in a tropical hilly area. *Geomatics, Natural Hazards and Risk, 1*(3), 199-223. 10.1080/19475705.2010.498151

PVMBG. 2016. *Laporan Singkat Bencana Gerakan Tanah Kec. Cisolok, Kab. Sukabumi Provinsi Jawa Barat*. Bandung: Pusat Vulkanologi dan Mitigasi Bencana Geolog.

Reichenbach, P., Busca, C., Mondini, A. C., & Rossi, M. 2014. The influence of land use change on landslide susceptibility zonation: the Briga catchment test site (Messina, Italy). *Environ Manage, 54*(6), 1372-1384. 10.1007/s00267-014-0357-0

Ristya, Y. 2018. *Comparison of Landslide Potential Areas by Using SINMAP, SMORPH, and Index Storie Methods in Kecamatan Pelabuhanratu and Surroundings Areas.* (Undegraduate Thesis), Univesity of Indonesia, Depok.

Samodra, G., Chen, G., Sartohadi, J., & Kasama, K. 2017. Comparing data-driven landslide susceptibility models based on participatory landslide inventory mapping in Purwosari area, Yogyakarta, Java. *Environmental Earth Sciences, 76*(4) 10.1007/s12665-017-6475-2

Schmitt, E. 2010. Weights of evidence mineral prospectivity modelling with ArcGIS. *Vancouver: Department of Earth, Ocean & Atmospheric Studies, University of British Columbia*

Sugianti, K., Sukristiyanti, S., & Tohari, A. 2016. Model Kerentanan Gerakan Tanah Wilayah Kabupaten Sukabumi Secara Spasial Dan Temporal. *Jurnal RISET Geologi dan Pertambangan, 26*(2), 117. 10.14203/risetgeotam2016.v26.270

Tarboton, D. G. 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water resources research, 33*(2), 309-319. 10.1029/96wr03137

Vakhshoori, V., & Zare, M. 2018. Is the ROC curve a reliable tool to compare the validity of landslide susceptibility maps? *Geomatics, Natural Hazards and Risk, 9*(1), 249-266. 10.1080/19475705.2018.1424043

van Westen, C. J., van Asch, T. W. J., & Soeters, R. 2005. Landslide hazard and risk zonation—why is it still so difficult? *Bulletin of Engineering Geology and the Environment, 65*(2), 167-184. 10.1007/s10064-005-0023-0

Zezere, J. L., Pereira, S., Melo, R., Oliveira, S. C., & Garcia, R. A. C. 2017. Mapping landslide susceptibility using data-driven methods. *Sci Total Environ, 589*, 250-267. 10.1016/j.scitotenv.2017.02.188