Implementing Landslide Susceptibility Map at Watershed Scale of Lompobattang Mountain South Sulawesi, Indonesia

Abdul Rachman Rasyid, Netra P. Bhandary and Ryuichi Yatabe

Abstract Landslides represent a serious hazard, particularly in the mountainous area currently. This recent study attempts to predict future landslide occurrence at watershed scale and to calculate the potency of mass movements for each sub-watershed at Lompobatang Mountain. In order to produce landslide susceptibility map (LSM) using the statistical model, we identified landslide inventories that occurred in the past, and predicted the prospective future landslide occurrence by correlating it with causal factors. In this study, six parameters were used namely, distance from fault, slope, aspect, curvature, distance from river and land use. This research proposed the Weight of Evidence (WoE) model to produce a landslide susceptibility map. Success and predictive rate were also used to evaluate the accuracy by using Area Under Curve (AUC) of Receiver Operating Characteristic (ROC). The result would be useful for land use planner and decision makers, in order to devise a strategy for disaster mitigation.

Keywords: landslide susceptibility; WoE model; watershed scale; disaster mitigation; Lompobattang Mountain

1. Introduction
As a developing country, Indonesia could not afford to recover from any natural disaster in short given time like landslide, flood, earthquake etc. The lack of spatial modeling in land use planning and information about the disaster make this country highly vulnerable to hazards. Hence, an effort to minimize these problems is to incorporate all disaster related information by preparing susceptibility map.

In Indonesia, floods and landslide occurrences are increasing and thus the combination of natural factors like the influence of global warming and human factors like land use change are becoming suspects. The land use change as part of development is a dominant factor for these disasters. Nowadays, landslides are serious problems, particularly in hilly terrains. Landslides trigger debris flows/flash, flood disaster that occurs every year often during or after a heavy rainfall in Indonesia. The Government and research institutes have attempted to minimize the loss through land-use planning and by incorporating all the information that contribute to future landslide susceptibility.

Indonesian National Board for Disaster Management /BNPB (2016) reported that disturbed watershed environment also caused floods and landslides. Currently, watershed damage in Indonesia is an extraordinary condition. In 2016, there are 118 watersheds in critical condition from total 450 while in 1984 there were only 22 critical and in 2007 it increased to 80 damaged watersheds (references?).

Landslide susceptibility map is the tool of land use planner and decision makers. In land use spatial planning management, the main subject is the area that includes the administrative authority and natural boundary like watershed. Watershed is sources of a set of valuable ecosystem because it provides high quality freshwater, carbon sequestration, nutrient retention, and bio diversity. Despite, watershed areas also have potential hazards like flash floods, landslides, and forest fires [Band, et al., 2012]. Based on the watershed boundary, the scope of losses during disasters like landslide is becoming obvious.
Beside landslide inventory, as a first stage of landslide hazard mitigation, landslide susceptibility mapping must provide important information to support decisions for urban development that considerably reduces the potential landslide damages. In other words, landslide susceptibility maps are produced to help humans to recognize and adapt to landslide hazard mitigation procedures [Pourghasemi, et al., 2012].

To deal with the impact of landslide disaster, this study also considers the landslide density and number of settlements at sub watershed scale in order to set the land use planning and management. A spatial planning based on disaster facilitates the governments and decision makers to implement a better policy. The landslide susceptibility map provides locations which have a high probability of landslide which helps to avoid the fatalities and to formulate preventive measures.

This paper presents GIS aided procedures for creating the landslide susceptibility map using a Weight of Evidence one of the statistical approaches. GIS represent and visualize the world through geographic layers containing object pixels, points, lines, and polygons along with associated attributes [Malczewski, 2004]. In GIS technique, a model can be used to combine a set of input maps or factors employing a function to produce an output map. The function can take many forms including linear regression, multiple regression, conditional analysis and discriminate analysis, etc. [Fell, et al., 2008].

The WoE initially uses Bayesian probability model and landslide causal factors are prepared as input maps. The WoE method calculates the weight for a certain category of a predictive factor based on the presence or absence of the landslide. The result of this method is map that tends to show the probability of occurrence and the associated uncertainty of the probability estimates of landslide occurrences [Neuhausser & Terhorst, 2007].

A variety of purposes using the WoE model have been applied not only in study related to landslide mapping. Agterberg et al (1993) employ the WoE model for study which concerned in potential mineral mapping and applying in a ground subsidence spatial hazard near abandoned underground coal mines by [Oh & Lee, 2010] or to analyze the relationship among groundwater productivity data to produce the regional groundwater productivity potential map [Lee, et al., 2012].

A comprehensive of the mathematical relationships was described in [Dahal, et al., 2008b] and [Regmi, et al., 2010]. In order to get positive and negative weights, the landslide locations overlayed with each of the causal factors in pairs and the spatial relationship was calculated subsequently (Table 1).

The WoE modeling is compatible with most GIS software packages [Dahal, et al., 2008a] and relatively easy, and less time consuming.

A set of landslide inventory map was prepared. It will show the location of visible landslides. The map was used in landslide susceptibility mapping by weights-of-evidence modeling because the analysis requires an inventory map. The literature shows the way to compile the landslide data beside field survey like collected from aerial photographs [Ayalew & Yamagishi, 2005] and remotely sensed images interpretation based on the spectral characteristics, shape, contrast and the morphological expression [Kanungo, et al., 2006].

The main objective of this study is to create a landslide susceptibility map at Lompobattang Mountain in watershed scale.

2. The Methods

A general rule of thumb is that water flow direction is perpendicular to contour lines. In the case of the isolated hill, water flows down on all sides of the hill (Fig. 1). The highest point upstream is the head of the watershed beyond which the land slopes away into another watershed. At each point on the stream, the land slopes up on each side to some high point then down into another watershed. Finally, the watershed boundary was created by connecting all of the high points around the stream in lines. High points are generally hilltops, ridgelines, or saddles.

The study area set at Lompobattang Mountain in Sulawesi Selatan province. The land use was classified into agriculture area and forest area . The study area was located at 119°50’ – 120°04’ E and 5°12’ -5°28’ S. It has maximum altitude of 2,876 meter above sea level covering a total area of 351.742 km² (Fig. 2).

There are 93 settlements in this area and the hydrologic system contains six watersheds, i.e.; Jeneberang, Lantebong, Kelara, Apparan, Bijawang and Tangka.

This study was developed with the general assumption that the future landslides will triggered by the same condition that causes landslides in the past. In general, there are two main data including for analysis i.e. landslide data inventories (dependent variable) and landslide causal factors (independent variable). In order to achieve the objective, the major procedures were conducted: (1) mapping of past landslides and landslide causal factors, (2) development of
landslide data inventories, including non-landslide areas, (3) preparation of probabilistic model of landslide susceptibility using weight of evidence model (4) Validation of model using Receiver Operating Characteristic (ROC) curve approach, (5) landslide density analysis at sub watershed scale in order to determine the level of susceptibility of settlements in this area. This study obtains landslide inventory data which was divided into landslide data for training and for validation. ArcGIS is used in spatial analysis, Microsoft excel to calculate the relationship between causal factors with landslide and SPSS statistical software was used to count the AUC of the ROC curve for validation of the model.

Table 1. Computed weights for causal factor classes based on landslide occurrences

<table>
<thead>
<tr>
<th>Topography</th>
<th>ΣA</th>
<th>ΣB</th>
<th>W+</th>
<th>W-</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspect Class</td>
<td>15</td>
<td>192998</td>
<td>0.05</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Flat</td>
<td>5</td>
<td>1933</td>
<td>-1.13</td>
<td>-1.13</td>
<td></td>
</tr>
<tr>
<td>North</td>
<td>159</td>
<td>22786</td>
<td>-0.13</td>
<td>-0.14</td>
<td></td>
</tr>
<tr>
<td>North East</td>
<td>353</td>
<td>53942</td>
<td>-0.20</td>
<td>-0.23</td>
<td></td>
</tr>
<tr>
<td>East</td>
<td>207</td>
<td>57312</td>
<td>-0.79</td>
<td>-0.88</td>
<td></td>
</tr>
<tr>
<td>South East</td>
<td>242</td>
<td>65237</td>
<td>-0.77</td>
<td>-0.87</td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>540</td>
<td>56441</td>
<td>0.18</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>South West</td>
<td>834</td>
<td>46535</td>
<td>0.81</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>372</td>
<td>30493</td>
<td>0.43</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>North West</td>
<td>231</td>
<td>35952</td>
<td>-0.22</td>
<td>-0.24</td>
<td></td>
</tr>
<tr>
<td>North</td>
<td>174</td>
<td>20199</td>
<td>0.08</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Distance From fault (m)</td>
<td>1500</td>
<td>296</td>
<td>35483</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>2000</td>
<td>177</td>
<td>27949</td>
<td>-0.23</td>
<td>0.02</td>
<td>-0.25</td>
</tr>
<tr>
<td>3000</td>
<td>411</td>
<td>50855</td>
<td>0.01</td>
<td>0.00</td>
<td>-0.02</td>
</tr>
<tr>
<td>4000</td>
<td>263</td>
<td>47642</td>
<td>-0.37</td>
<td>0.04</td>
<td>-0.41</td>
</tr>
<tr>
<td>6000</td>
<td>63</td>
<td>81203</td>
<td>-2.33</td>
<td>0.21</td>
<td>-2.54</td>
</tr>
<tr>
<td>8000</td>
<td>69</td>
<td>53027</td>
<td>-1.81</td>
<td>0.12</td>
<td>-1.94</td>
</tr>
<tr>
<td>Landuse Class</td>
<td>994</td>
<td>99453</td>
<td>0.23</td>
<td>-0.09</td>
<td>0.32</td>
</tr>
<tr>
<td>Primary Dry Forest</td>
<td>757</td>
<td>46591</td>
<td>0.71</td>
<td>-0.15</td>
<td>0.86</td>
</tr>
<tr>
<td>Secondary Dry Forest</td>
<td>977</td>
<td>83215</td>
<td>0.39</td>
<td>-0.14</td>
<td>0.52</td>
</tr>
<tr>
<td>Mix Dryland Agriculture</td>
<td>382</td>
<td>133251</td>
<td>-1.02</td>
<td>0.29</td>
<td>-1.31</td>
</tr>
<tr>
<td>Grass Land</td>
<td>7</td>
<td>1112</td>
<td>-0.24</td>
<td>0.00</td>
<td>-0.24</td>
</tr>
</tbody>
</table>

ΣA = Number of landslide pixels, ΣB = Number of class factor pixels, C = Contrast

Landslide data inventories were collected from Google Earth image interpretations in area where step scarp of noticeable landscape modification exist that lies at the altitude of 500 m above sea level. From this, a total of 158 landslides were identified which covers an area of 3.44 km². Most of the landslide are shallow landslides with maximum and minimum area of 708 m2 and 512.765 m2 (0.51 km²) respectively.

Six landslide causal factors (distance from fault, slope, curvature, aspect, distance from river, and land use) were selected as independent variables for landslide susceptibility mapping. Figure 3 describes the stages of this research.

Remote sensed images were used to identify rock fractures that is indicated by lineaments. The fractures affects surface material structures and make a large contribution to terrain permeability leading to slope instability [Kanungo, et al., 2006]. Faults and lineaments are structural features representing zone of weakness along fractures planes where landslide susceptibility is higher. Thus, distance from fault and
landslide occurrences were analyzed to identify their relationship. The identification of faults was based on Geology Map at a scale of 1:250,000 [Sukamto and Supriatna, 1982]. The proximity distance from fault was identified using multilevel buffer from lineament or fault map (300 m to 8000 m in 8 categories).

Topography is one of the important factors in the development of landslide susceptibility map. The topographic data were analyzed to determine slope, aspect and curvature. These data were derived from ASTER DEM with spatial resolution of 30 m. ASTER DEM was extracted to be elevation thematic map. Then, Triangular Irregular Networks (TIN) were produced in ArcGIS and converted into DEM raster with a pixel size 30 m. This allows to calculate slope angle, slope aspect and curvature. Slope is the most substantial cause for landslides. On a slope of uniform isotropic material, increasing slope angle correlates with increasing likelihood of failure. In this study, six slope angle categories were made i.e., 0-5°, 5-10°, 10-20°, 20-30°, 40-50°, and above 50°. In many landslide susceptibility studies, slope angle is an important factor in inducing slope instability study. Aspect is defined as the direction of maximum slope of the terrain surface which have an indirect influence on slope instability. The aspect map also plays a significant role in slope stability assessment [Chauhan, et al., 2010]. In this study, aspect is divided into 9 classes for the study area, namely, flat, N, NE, E, SE, S, SW, W, and NW. To describe the variances among classes, aspect map display the distribution of each direction in the topography by using different color to each cell of the study area [Quan & Lee, 2012]. Profile curvature was reclassified into 3 classes, namely, concave, flat and convex. The curvature values represent the morphology of the topography. Profile curvature is generally related to the puddle condition after heavy rainfall. Profile curvature slope contains more water and retain water from heavy rainfall for a longer period.

The distance from river was identified by buffering and analyses of river lines that were derived from topographic map with scale 1:50,000. The class starts from 0-50 m and ends with > 300 m. Drainage lines and landslide occurrence in hilly area have strong association between them due to erosional activity. The last causal factor is land use. Land use (cover) is also a key factor responsible for landslide occurrences. The occurrence of landslide is inversely related to the vegetation density. The land use map was derived from Landsat 7 with 30 m x 30 m, and it was established by one board from Ministry of Environment and Forestry 2014 (Balai Pengelolaan Daerah Aliran Sungai Jeneberang Walanae. in. Bahasa). The land use map is classified into several classes: forest (including primary and secondary), bushes, crop land (agriculture), and grass land were identified. Dependent and independent variables were converted into a spatial form using Geographic Information System (GIS). Arc GIS 10.0 was used for preparing a spatial database, particularly using the spatial analyst and 3-D analyst tools. All independent and dependent variable were digitized and then processed by converting all the datasets into a raster format with 30 m x 30 m pixel size. The total numbers of cells are 390,837 pixels and the numbers of landslide pixels are 3,827. Microsoft Excel was used to compute, manage tabular data and calculate the weight of independent variable values related to dependent variable based on weight of evidence model (WoE).

Statistical approaches can be categorized into bivariate and multivariate approaches. The typical procedure for bivariate statistical analysis that each map which depicts the spatial extent of causal factor of landslide will be combined with the landslide data inventory. Then, weight values based on landslide density are calculated for each mapped category of the causal factor. In general, there are two models in bivariate
approaches in landslide susceptibility study i.e. frequency ratio (FR) and WoE. This current research uses the WoE to calculate and produce landslide susceptibility.

Originally, the purpose of the WoE model is to provide flexible tools to modify the value in continuous and categorical predictor variables into discrete categories automatically and to assign to each category as “weight” value. The weight of evidence tells the predictive power of an independent variable in relation to the dependent variable. It is calculated by taking the natural logarithm (log to base e) of division of % of non-events and % of events. F represents the presence and represents absence of a potential landslide causal factor. L and represent consecutively the presence and the absence of landslide, then the WoE method calculates the positive and negative weights of the respective factor classes based on the probability ratios as follows:

\[ W^+ = \log \left( \frac{p[F|L]}{p[F|L]} \right) \]

\[ W^- = \log \left( \frac{p[F|L]}{p[F|L]} \right) \]

In order to calculate the weights of each landslide factor classes for landslide susceptibility mapping, Eq. (1) and Eq. (2) were expressed in terms of the numbers of cell (pixel) as follows:

\[ W^+ = \log \left( \frac{npix1}{npix4} \right) \]

\[ W^- = \log \left( \frac{npix2}{npix4} \right) \]

and

\[ W = \log \left( \frac{npix3}{npix4} \right) \]

where npix1 is the number of landslide pixels present on a given causal factor class, npix2 is the number of landslides pixels not present in a given causal factor class, npix3 is the number of pixels in a given causal factor class in which no landslide pixels are present and npix4 is the number of pixels in which neither landslide nor the given factor is present.

A positive weight (W+) the presence of a causal factor in the landslide and its magnitude is an indication of the positive correlation between presence of the causal factor and landslides. A negative weight (W−) indicates an absence of the causal factor and the magnitude indicates negative correlation. The C was calculated in order to quantify the spatial association between a map class and the occurrence of landslides which is known as the weight contrast, where \( C = W^+ - W^- \). The magnitude of contrast reflects the overall spatial association between the causal factor and landslides.

3. Results and Discussion

The six landslide causal factor maps were combined with the landslide data for training. The procedure of combining the maps based on GIS spatial analysis tools and Microsoft Excel was used to calculate the weight of their relationship. The pixels of each class of causal factors and landslide were arranged in spatial database and they were subsequently exported into Microsoft Excel format.

The result of total weight (i.e. contrast) as shown in figure 4, directly indicates the importance of each causal factor. In case of slope angle factor, the contrast increases with the slope angle class from 30° to > 50°, indicating a higher positive correlation with the landslide locations. Related to aspect factor, the graph is positively correlated in south west. Concerning distance from fault and distance from river class, same condition is seen within the distance from 500 m to 1000 m and >300 m respectively, which means that these two factors have high probability of occurrence. From five classes of land use factors, class of secondary dry forest is highest contrast means has a significant positive relationship with the landslide.

Landslide susceptibility index map (Fig. 5) can be obtained by combining the contrast of each causal factor. This is expressed in the following equation ;

\[ LSI = \sum_{i=1}^{n} C_{ij} \]

where, LSI is a Landslide Susceptibility Index, Cij is the contrast for class i of causative factor j; the summation is over all causative factors from 1 to n.

The weights are assigned to the classes of causal factor maps to produce weighted causal factor maps. Then, all the maps were overlaid and numerically added according to Eq. (5) to create a LSI map. The result of the LSI values varies from -8.22 to 6.15.

The landslide susceptibility map in Fig. 5 was created from the five classification of landslide susceptibility index map by using natural breaks classification method or Jenks optimization method. In this study area, the five classes include very low, low,
moderate, high and very high landslide susceptibility.

The percentages of landslides susceptibility zonation map is shown in Fig. 6. The ratio of each class of susceptibility was found by dividing the area of each class with the total area that area which falls in high to very high, moderate, and very low to low comprises of 28.96%, 24.29 %, and 46.75% respectively. The landslide data for validation subsequently overlay on the LSM to calculate number of landslide pixels in each class of susceptibility. Fig. 6 showed that the ratio is 82.96% in high to very high classes, 24.29% in moderate class, and 8.17% in very low to low classes. It indicates, that most of the landslide data for validation pixel fall into high to very high class of susceptibility.

This study establishes two stage of validation. First, the Area Under Curves (AUC) of Receiver Operational Characteristic (ROC) curve was used to measure the level of accuracy of the model and the probability result. The second stage is the ratio of landslide data for validation which laid on the LSM that indicate good results (82.96%).
The AUC represents the success rate and predictive rate. The AUC of the success rate was found by combining the LSI map values with landslide data for training. Likewise, in case of the AUC for the predictive rate was set up by combining the LSI with landslide data for validation. As general rule, Hosmer and Lemeshow (2002), classified the AUC value for validation. The model is failed if the AUC less than 0.50, acceptable: 0.70-0.80, excellent: 0.80-0.90 and outstanding: 0.90-1.00. The result shows that the model exhibit similar performance for both success rate and predictive rate, i.e. 0.849 and 0.847 (indicate 84.9 % and 84.7 % accuracy) respectively.

The density of landslide susceptibility in sub watershed area was found by summarizing each class of landslide susceptibility in sub watershed area divided by total area. The number of settlements which were located in each sub watershed was counted as well means found that the probability of property loss will be high for instance houses and road. Fig. 7 shows Bialo, Bijawang, and Kelara have high percentage of landslides, which fall in high to very high (VH+H) around 6% to 7% comparing with another sub watershed. In the case of a number of settlements, Bialo and Tangka is > 15. The ratio indicates that these sub watershed need to get more concern on land use planning to minimize the probability of loss.

This recent study tried to create a map which is describing the vulnerability of landslide using 6 (six) causal factors with the high level of accuracy, meanwhile, some researchers use more. It can conclude that the combination of factors also has an important role to achieve the best result [Meten et.al. 2015].

4. Conclusions
This study has demonstrated the simple of procedures for landslide susceptibility assessment by WoE methods with (six) thematic maps. This landslide susceptibility map will be used to provide an important information particularly to determine the suitability of land. This research also describe about the good level of accuracy of prospective landslide prediction using the AUC curve and by overlying landslide inventories that were not used in model building over the landslide susceptibility map. It is recommended to consider the sensitive natural factors such as slope angle and land use/land cover for spatial planning purposes. Nevertheless, the result of the landslide susceptibility map will be more accurate on a larger scale, particularly at topographic map and geological map. In order to prevent any fatalities in future causes by landslide disaster, the availability of data regarding population at this area will be helpful and the result of the planning decisions will be more comprehensive.

Acknowledgements
This research was supported by DIKTI Scholarship Batch 1,2013 and collaboration between Ehime University of Japan and Hasanuddin University of Indonesia.

References


