RESEARCH ARTICLE

Landslides Susceptibility Mapping in R Program (Case Study in Lima Puluh Kota Regency)

Ahyuni[a],*, Bigharta Bekti Susetyo[a], Isra Haryati Diva[a], Zakiyah Mar'ah[b], Hamdi Nur[c], Adenan Yandra Nofrizal[d], Azwirda Aziz[e]

[a] Department of Geography, Universitas Negeri Padang, Indonesia; [b] Department of Statistika, Universitas Negeri Makassar, Indonesia; [c] Department of Urban and Regional Planning, Universitas Bung Hatta, Indonesia; [d] Graduate School of Integrated Science and Technology, Shizuoka University, Japan; [e] Accounting Departement, STIE Swadaya, Indonesia

Abstract Landslide susceptibility zonation is necessary to be considered in regional land use planning. Various approaches and analytical methods using GIS software have been applied to zone the area. However, there are constraints in its use, such as the cost to get license software and a given source code that cannot be accessed and manipulated by users. The recent development of free and open-source software such as R programming language for data analysis and graphics has opened up opportunities for users to reevaluate, modify and develop further the programming algorithm. Considering these issues, the study aims to develop a coding and syntax model of dominant factors controlling landslides and generate a landslides susceptibility map in R program which can be accessed by users. This study applies the WoE (Weight of Evidence) method as a part of the statistical approach which is appropriate for assessing landslides in a medium-scale area to inform the regional spatial planning. The case study of modeling is Limapuluh Kota Regency in West Sumatra Province of Indonesia, a hilly and mountainous region where most of its districts are prone to landslides. Eight of eleven factors (geology, landform, land cover, elevation, the density of vegetation greenness, slope, rainfall intensity, and proximity to stream) were considered to control landslides which set up four classes of landslide susceptibility zone (very low, low, moderate, and high).

Keywords: Landslide susceptibility zone, Weight of Evidence, R program.

Introduction

Landslide is soil or rock mass movement, or a mixture of both, down or out of the slope. The natural properties of slope stability influence its susceptibility. The study of landslide susceptibility evaluation and prediction are broadly grouped into geomorphological mapping, geotechnical of physically-based, and analysis of landslides inventories. It uses heuristic, deterministic and statistical approaches, respectively [1]. The heuristic method applies geomorphological mapping to large-scale areas based on experts' judgment of variables such as slopes, faults, and geology. The deterministic method applies to the small-scale area by analyzing the geotechnical stability condition of parameters. A statistical approach is a new approach to mapping landslide hazards by combining the possibility of landslides from statistical data and the physical parameters of landslides. This approach is appropriate for assessing landslides in a medium-scale area which helps inform the regional spatial planning [2, 3]. Research on landslides has been widely applied using a method or comparing them [4, 5, 6, 7, 8].

Nowadays, free and open-source Geographic Information System (GIS) software can process statistical models [9]. One of them is R program, which has cutting-edge spatial packages to behave as a fully-featured GIS [10]. Several advantages of the utilization of R language for spatial analysis such as its command-line interfaces allow a rapid description of workflow and reproducibility, has sophisticated and customizable graphics, and have an extensive range of functions through an additional package, integrated processing, analysis, and modeling framework. R statistics has a wide range of functions and libraries that allow using all statistical tools with advanced visualization capabilities [11]. The recent updates of the libraries attached to R environment made the output and result very handy and without the need to change the working environment or data format, which will reduce the uncertainty of switching back and forth between different geospatial and statistical analysis platforms [8]. Some studies have analyzed land susceptibility using R Program [12, 13, 14].

This study uses R program to develop a coding and syntax model of dominant factors controlling landslides and generate a landslides susceptibility map. Users can access its source code, and hence the program functions can be applied, modified, or developed further in analyzing regional landslides susceptibility in another context. In order to determine the factors of landslides, this study applies the WoE (Weight of Evidence) method based on bivariate statistical analysis. The advantage of using this method is its evidence-based and easy to measure. The case study of this study is a region of Limapuluh Kota Regency, West Sumatra Province of Indonesia.

Study Area

The study is located in Lima Puluh Kota Regency, West Sumatra Province of Indonesia. This regency is located between $0^0 25'28.71''$N - $0^0 22'14.52''$ South Latitude and $100^0 15'44.10''$E - $100^0 50'47.80''$ East Longitude. It consists of 13 sub-districts with approximately 3,354.30 Km^2 and, according to the 2020 national census, has 383.525 people. The geological of this region are varied, influenced by the dynamic Sumatra Fault System [15], which is evidenced by the discovery of igneous rocks such as basalt, tuff, andesite, pumice, sedimentary rocks such as sandstone, limestone, and metamorphic rocks such as slate, quartz, and phyllite [5]. The topography ranges from flat, undulating, hilly, and mountainous with a height between 110 meters and 2,261 meters from sea level. There are three inactive volcanoes in this area, namely Mount Sago (2,261 m), Mount Bungsu (1,253 m), and Mount Sanggul (1,495 m), and also 13 large and small rivers. The research area is shown in figure 1.

Floods, landslides, and tornadoes frequently hit this regency. Especially for landslides, almost all of its districts are prone to this hazard. Officially, the Local Disaster Management Agency (BPBD) had recorded 127 landslide points based on data from 2016 to 2019, which had deteriorated facilities, houses, roads, and agricultural land.
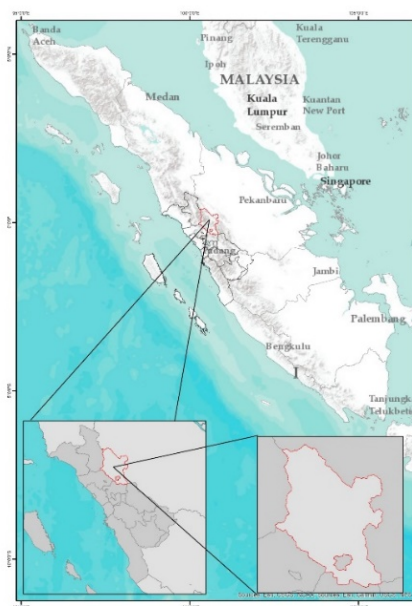

Figure 1. The research area

Methods

This research applies the Weight of Evidence (WoE) method, a type of the bivariate statistics model, to determine significant factors controlling landslides. This method assumes that landslide susceptibility factors were not inter-related and be based on the Bayes' rule [9,10]. Measurement of Weight of Evidence (WoE) is as follows:

$$\mathrm{Wij}^+ = \mathrm{Log\ e} \frac{P(B/D)}{P(B/\bar{D})}$$

$$\mathrm{Wij}^- = \mathrm{Log\ e} \frac{P(\bar{B}/D)}{P(\bar{B}/\bar{D})}$$

Where:
W_{ij}^+ = The positive weights of evidence of the jth parameter class of ith landslide
W_{ij}^- = The negative weights of evidence of the jth parameter class of ith landslide
P   = Probability
B   = presence of the landslide evidential feature
D   = The number of landslides belonging to the evidential feature
$\bar{B}$ = The total area on the map where the evidential feature is absent
$\bar{D}$ = The number of landslides not belonging to the evidential feature

The positive value of W_{ij} indicates the importance of the presence of the factor for the occurrence of the landslides. Otherwise, the negative value indicates the factor is not favorable or absent; the greater the value, the stronger the predictive ability for susceptibility analysis. The class between 0.1-0.5 is middle predictive, 0.5-1 is moderately predictive, 1-2 is strongly predictive, and greater than 2 is extremely predictive [16].

Cij is the contrast value of the jth parameter class of ith evidential map with a positive or negative value. It measures the spatial association between the analyzed parameter and landslide occurrence. A class of parameters has greater predictive power when C > 1 and is very significant when the value is approximately 2 [17].

$$\mathrm{Cij} = \mathrm{Wij}^+ - \mathrm{Wij}^-$$

The next step is calculating the value of the Area Under Curve (AUC), which measures the reliability of the association of the parameters to landslides occurrence. The AUC in each parameter class is measured by dividing the percentage of the total area of the parameter class by the percentage of the total number of landslides. This step validates by comparing the susceptibility measurement of the training sets that were used for building the models and the test set that were not used during the model building process. The AUC value of the predictive factors of landslides is 0.6 to 1 and is considered as most predictive when its value is closer to 1. Value of AUC classification are: < 0.6 (a poor model), 0.7-0.8 (a medium or reasonable model), 0.8-0.9 (a good model), and > 0.9 (very good model) [18, 19].

There were 142 landslides points collected from field surveys and historical records data from The Regional Agency for Disaster Countermeasure (BNPBD). The data was divided into 83 train and 59 test data sets in a proportion of 60:40 (as shown in figure 2). Data of parameters such as elevation, slope, aspect of slope, curvature, and the distance from the stream were processed from the image of the National Digital Elevation Model (DEM Nas) available in the Inageo portal of Indonesia Geospatial Information Agency (BIG), with a resolution of 8.3 meters. Land cover data from Image Spot 7 2020. NDVI data is processed from Landsat 8 imagery and then pan-sharpened to get a resolution of 15 meters, in pixels of 25 x 25 meters, and presented at a scale of 1:50,000. Rainfall intensity data is obtained from the West Sumatra Regional Agency for Water Resources Management (PSDA) and processed using the IDW method. Lithology data was collected from the geological map and landforms and soil types from a semi-detailed soil map of the Indonesian Center for Agricultural Land Resources Research and Development (BBSDLP).

Model of Code and Syntax in The R Program

The flowchart of research to develop a coding and syntax model of dominant factors controlling landslides and generating a landslide susceptibility map using R program can be seen in figure 3.
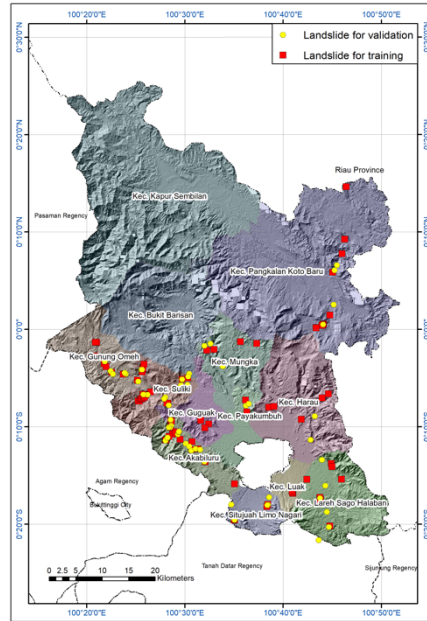
Figure 2. Landslide points occurrence comprises train and test data set
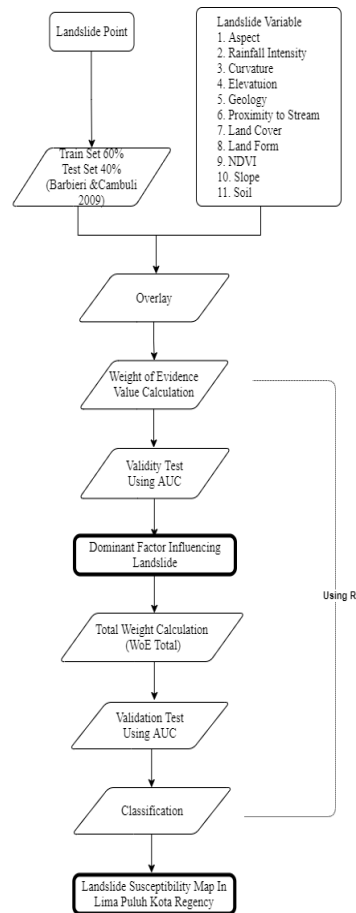

Figure 3. Flow chart of the research

The steps of creating code and syntax of dominant factors controlling landslides and generating landslide susceptibility maps are described as follows.

A. Set up data in Microsoft excel software with the following steps:
   1. Prepare maps of variables that affect landslides in the form of raster data. Each map contains various categories based on predefined classes (in the GIS arc) with a pixel resolution of 25 meters.
   2. Prepare a map of landslide event points in raster data divided into data trains and tests (60:40) (GIS arc).
   3. Overlaying map of landslide event points (data trains) with variable maps that affect landslides using the "map algebra" tool. The attribute table of overlay results was then moved to processed in R program.

B. Identify factors that affect landslides by using packages "readxl and packages "writexl") and then the data is stored with the following syntaxes:

   calling data
   > excel_sheets("Landslide variable.xlsx")
   > data<-read_xlsx("Landslide variable.xlsx", sheet="Aspect")
   > data

   1. Arrange the syntax for the WoE value calculation function.
   WoE is a function of the values x and y, where x is the value of the number of pixels and y is the number of landslide events. The calculation is as follows:

```
#function WoE
WOE<-function(x,y){
  tc=sum(x)
  factor<-0
  for (i in 1:length(x)){
    factor[i]<-x[i]/tc
  }
  tot_Ls=factor+y
  class_dens=tot_Ls/x
  tls=sum(tot_Ls)
  map_dens=tls/tc
  fr=class_dens/map_dens
  ivm=log(fr)
  ww=log((tot_Ls/tls)/((x-tot_Ls)/(tc-tls)))
  w=log(((tls-tot_Ls)/tls)/((tc-x-tls+tot_Ls)/(tc-tls)))
  c=ww-w
  tw=sum(w)
  wwoe=ww+tw-w
  woe=round(1000*wwoe)
  results<
as.data.frame(cbind(x,y,factor,tot_Ls,class_dens,map_dens,fr,ivm,ww,w,c,wwoe,woe))
  colnames(results)<-
c("Area","Landslide","Factor","Total_Ls","Class_Dens","Map_Dens","FR","IVM","W+","W-
","C","WeightWOE","WOE")
  return(results)
}
```

   2. Calculate the WoE value
   Before performing the WoE calculation, run the WoE function that has been defined. For the next syntax writing, variable x is defined in a column that contains pixel values and variable y in the landslide event number column. The calculation of WoE is as follows:

```
#count WOE
```
   > results<-WOE(x=data$Area,y=data$Landslide)
   > results

   To perform repeated analysis or different data, simply rename x and y. if the variable x in excel data is named "a" then the Code "pixel value" is replaced with]"a". if the variable y in excel data is named "b", then the code of "landslide event" is changed to "b"

   3. Save data in Microsoft excel form
   > write_xlsx(hasil,"results Aspect.xlsx")

   4. Compose syntax for the AUC calculation function
   AUC is defined as a function consisting of the values of w, x, and y. WoE value is sorted from largest to smallest and expressed in w. Several pixels are expressed in x and the number of landslide events in y.

```
#function AUC
```

```
>        AUC<-function(w,x,y){
a<-data.frame(w,x,y)
a<-a[order(-a$w),]
tot_area=cumsum(a$x)
prop_area=tot_area/sum(a$x)
total_ls=cumsum(a$y)
prop_ls=total_ls/sum(a$y)
auc<-0
for (i in 1:length(a$x)){
  if (i==1)
    auc[i]=prop_ls[i]*prop_area[i]*0.5
  else
    auc[i]=(prop_ls[i]+prop_ls[i-1])*(prop_area[i]-prop_area[i-1])*0.5
}
results<-as.data.frame(cbind(a$w,a$x,a$y,tot_area,prop_area,total_ls,prop_ls,auc))
colnames(results)<-
c("WOE","Area","Landslide","Tot_Area","Tot_Area_%","Total_Ls","Tot_Ls_%","AUC")
  return(results)
}
```

5. Calculate AUC and Total AUC values

After define the AUC function syntax, the next step is to calculate of AUC. The syntax used is as follow:

```
#count AUC (w=value woe)
>   results2<-AUC(w=results$WOE,x=data$Area,y=data$Landslide)
>   results2
#count total AUC
>   tot_auc<-round(sum(results2$AUC),4)
>   tot_auc
```

6. Create AUC graph

The AUC is the proportion of pixel values in each variable category and the number of landslide events. The syntax is as follows:

```
# Plot AUC
>   PLOT_AUC<-function(x,y,auc){
    plot(x,y,type="s",lwd=2, col="red",
        xlab="Tot_Area_%",
        ylab="Tot_Ls_%",
        main="AUC Curve")
      legend(0.6,0.6,auc,title="AUC",cex=1)
    }
>   PLOT_AUC(x=results2$`Tot_Area_%`,y=results2$`Tot_Ls_%`,auc=tot_auc)
```

7. Save AUC calculation results

The results of AUC calculations can be stored in Microsoft excel form to make it easier to see the data.

```
#save the results in excel format
>   write_xlsx(results2,"NDVI 10.xlsx")
```

C. Analysis of Landslide Susceptibility Map

The steps to create a map are as follows:

1. Calculate the sum of WoE in Arc GIS

Selected variable with the AUC value > 0.6, and then calculate the sum of WoE by overlay using algebra map tools with addition function. "WoE Total = WoE Geology + WoE land forms + WoE land cover + WoE slope slope + WoE vegetation density + WoE distance from river+WoE Slope Direction + WoE Rainfall"

2. Overlay the sum of WoE value with data test.

The value of total WoE that has been calculated is then overlayed in Arc GIS with the data test set by using the multiplication function of algebra map tools.

3. AUC calculation for the map-making process is the same as step in B 6, which uses the total WoE result of the calculation of C1 and C2.

4. Classify Landslides Susceptibility

```
#classification
    CLASS<-function(data,prop,class_1,range_1a,range_1z,
        class_2,range_2a,range_2z,
        class_3,range_3a,range_3z,
        other) {
```

```
  classification<-0
  for (i in 1:length(prop)){
    if (prop[i] >= range_1a &prop[i] <= range_1z) {classification[i]<- class_1}
else if (prop[i] > range_2a &prop[i] <= range_2z) {classification[i]<-class_2}
else if (prop[i] > range_3a &prop[i] <= range_3z) {classification[i]<-class_3}
else {classification[i]<-other}
  }

  code<-0
  for (i in 1:length(prop)){
    class_1=1
    class_2=2
    class_3=3
    other=4
    if (prop[i] >= range_1a &prop[i] <= range_1z) {code[i]<- class_1}
    else if (prop[i] > range_2a &prop[i] <= range_2z) {code[i]<-class_2}
    else if (prop[i] > range_3a &prop[i] <= range_3z) {code[i]<-class_3}
    else {code[i]<-other}
  }

  results1<-cbind.data.frame(data,classification,code)
  return(results1)
}
```

#only valid for 2 to 4 classes

```
>   description<-CLASS(data=results,prop=results$`Tot_Ls_%`,
        class_1="High", range_1a=0, range_1z=0.7,
        class_2="Moderate", range_2a=0.7, range_2z=0.85,
        class_3="Low", range_3a=0.8, range_3z=0.95,
        other="Very Low")
>   description
```

#save classification

```
>   library(writexl)
>   write_xlsx(description,"classification AUC Total .xlsx")
```

5. Mapping of Landslides Susceptibility
    Mapping landslide susceptibility using the classification table data on number C. 4. The
    syntax is as follows:
    #read data

```
>   library(readxl) #rstudio 4.0.5
>   data<-read_excel("classification AUC Total .xlsx")
>   data
```

#Mapping
#function only valid for 2 to 4 classes

```
>   library(raster)
>   library(rgdal)
>   r<-raster("WoE_rev1.tif")
>   MAP<-function(raster,area,code,
     map_name,
     legend,legend_name){
     graphics.off()
     rat<-ratify(raster,count=T)
     levs<-levels(rat)
     vals<-unique(values(raster))
     vals<-na.omit(vals)
     vals<-sort(vals)
     gab<-data.frame(levs,vals)
     kod<-cbind.data.frame(area,code)
     kod<-kod[order(kod$area),]
     recl<-cbind.data.frame(gab$ID, kod$code) #yang dirubah. titik lokasi adalah ID-nya
     mapping<-reclassify(raster, rcl=recl)
     par(xpd=F)
plot(mapping,col=c("red","yellow","#00FF33","#003300"),legend=F,axes=F,main=map_name)
```

```
                        par(xpd=T)
                        legend("topright",legend=legend,
                            fill=c("red","yellow","#00FF33","#003300"), title=legend_name)
                        results<-cbind.data.frame(gab,kod)
                        return(results)
                    }
            #code = class code
                >    map<-MAP(raster=r,
                            area=data$WOE,
                            code=data$code,
                            map_name="Landslide Susceptibility Map In Lima Puluh Kota Regency",
                            legend=c("High","Moderate","Low","Very Low"),
                            legend_name="Susceptibility Level")
```

Results and Discussion

Eight of eleven parameters become the determining factors in controlling landslides (figure 4). These factors are: geology (0.81), landform (0.73), land cover (0.65), aspect of slope (0.64), slope (0.64), NDVI (0.64), proximity to stream (0.62), and rainfall intensity (0.61). The other three parameters: soil type (0.60), elevation (0.59), and curvature of the slope (0.54), are not taken into account as significant factors for controlling landslides. Soil type value was at the minimum threshold, but this study considered it was not significant enough to control landslides.
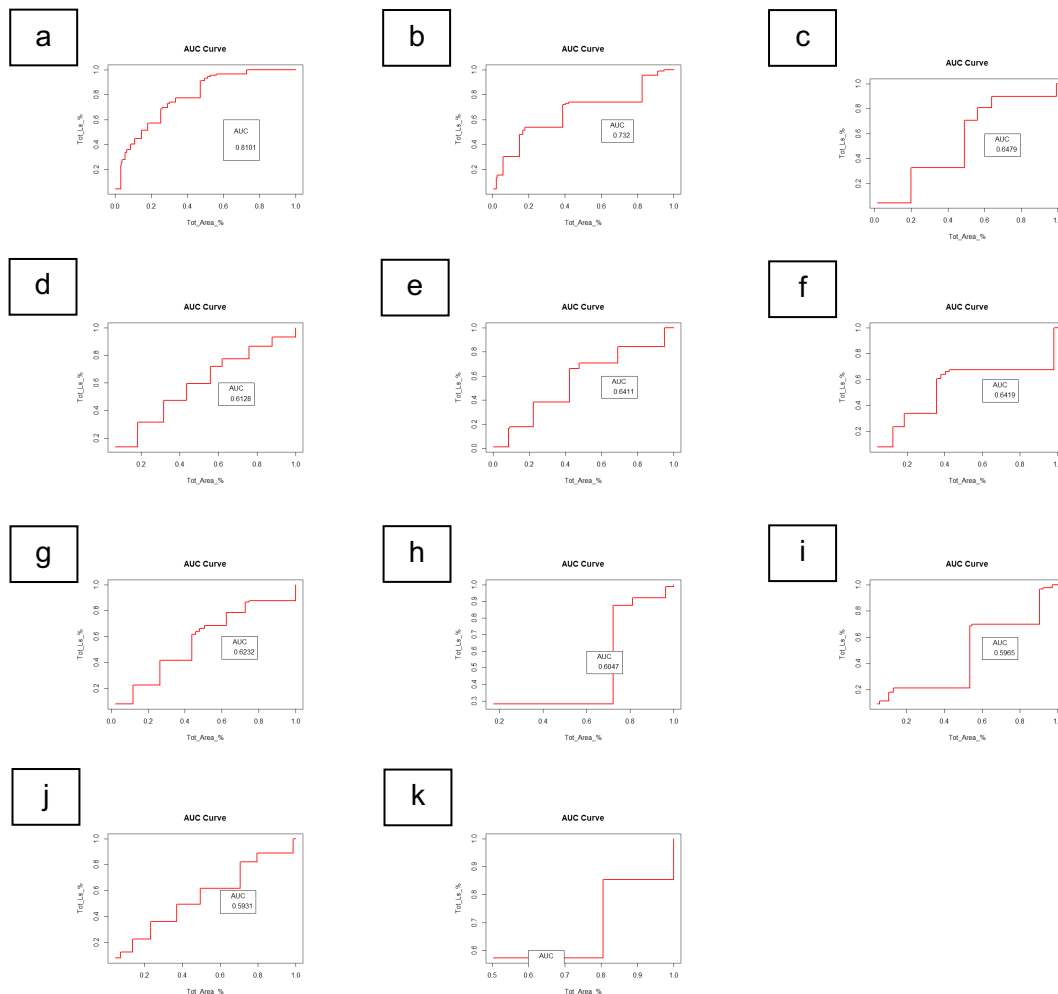
Figure 4. AUC value of parameters (a) Geology, (b) Land form, (c) Land cover, (d) Aspect, (e) Slope, (f) NDVI, (g) Proximity to stream, (h) Rainfall intensity, (i) Soil type, (j) Elevation, (k) Curvature

The analysis of AUC by using 83 (60%) train data set get a value of 0.61. Then, a test was carried out to measure the validity of the prediction of the parameter model by using 59 (40%) test data sets, which are recorded landslide point data that were not used previously in the analysis process. The value of AUC by a test of validity is 0.75, which means that the parameter and inventory database model for prediction is at a reasonable model.
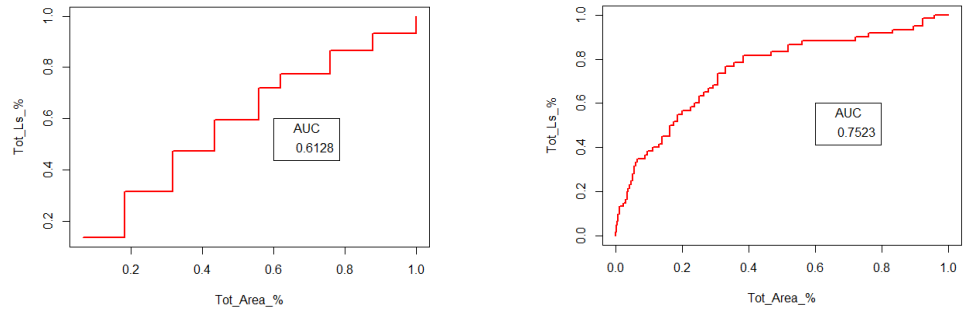
AUC curve and value by using 83 (60%) train data set

AUC curve and value by using 59 (40%) test data set

Figure 5. AUC curve and value

Significance factors controlling landslides in the region describe as follows: (1) lithology types in geological factor were found as the most significant factor to predict landslides, (2) volcanic intrusion landform in landform factor is strongly associated with the landslides, and respectively followed by old volcanic hills, lower volcanic slope, and volcanic plain, (3) land cover factor that was found to be associated with landslides are paddy fields, plantations, and settlements. Notably, the terraced form of the paddy field, which is found vastly in this regency, is one of the explanations for the occurrence of its landslide, (4) aspect of slope $0-22.5^0$ is most associated with the landslide, (5) the dense vegetation class in NDVI factor is associated with landslides, (6) in slope factor, the class of slope 164-180% is a very significant factor for controlling landslides and the slope 42-69% is moderately significant but the other classes in not significant, (7) in rainfall intensity factor, areas with rainfall intensity class between 1000-2000 correlated with the occurrence of landslides, and (8) the landslide is associated with proximity to the river (100-200 meters) and in a considerable distance from the river (2000-3000 meters). Factors associated with landslides based on the calculation of Weight of Evidence are shown in Table 1.

Table 1. Weights (W), contrast values (C), and Area Under Curve (AUC) of parameters classes

Classes	The class area from the total area (%)	Landslide occurrence	Weight+	Weight -	C	AUC
Geology						0.81
Sandstone	23.5	9	-0.747	0.150	-0.897	
Quartzite, sandstone	8.6	7	-0.013	0.001	-0.014	
Quartz sandstone	7.1	5	-0.146	0.010	-0.156	
Rude Conglomerate	2.5	3	0.389	-0.012	0.401	
Andesite, Basalt	12.6	29	1.021	-0.297	1.318	
Phylite, Quartzite	3.8	11	1.265	-0.104	1.370	
Pumice	3.3	9	1.190	-0.082	1.271	
Conglomerate	0.1	1	2.090	-0.011	2.100	
Granite	1.4	4	1.229	-0.035	1.264	
Phillyte, flakes	1.8	3	0.692	-0.019	0.710	
Silt, sand, gravel	0.4	1	0.997	-0.008	1.005	
Landform						0.73
Colluvial plain	0.4	1	0.997	-0.008	1.005	
Tectonic plain	8.5	2	-1.215	0.063	-1.278	
Volcanic plain	3.4	9	1.156	-0.080	1.236	
Tectonic plate escarpment	1.3	1	-0.062	0.001	-0.062	
Volcanic intrusion	0.3	4	2.709	-0.046	2.756	
Lower Volcanic slope	1.9	6	1.356	-0.056	1.412	
Middle Volkanic slope	1.5	1	-0.193	0.003	-0.195	

Classes	The class area from the total area (%)	Landslide occurrence	Weight+	Weight -	C	AUC
Tectonic mountains	40.8	18	-0.609	0.273	-0.881	
Old volcanic mountains	8.9	14	0.647	-0.093	0.740	
Tectonic hills	20.7	17	0.002	0.000	0.002	
Old volcanic hills	1.7	8	1.743	-0.085	1.827	
Tectonic plate ridge	1.8	1	-0.369	0.006	-0.375	
Land cover						0.65
Forest	35.4	6	-1.532	0.358	-1.890	
Plantation	29.2	34	0.345	-0.187	0.532	
Settlement	1.5	2	0.496	-0.010	0.506	
Paddyfield	7.9	16	0.903	-0.134	1.036	
Shrubs	7.2	3	-0.661	0.037	-0.698	
Dry agriculture	18.8	21	0.305	-0.086	0.392	
The aspect of slope (^{0})						0.65
0-22.5	6.6	12	0.714	-0.076	0.790	
22.5-67.5	14.0	8	-0.436	0.056	-0.492	
67.5-112.5	13.4	14	0.162	-0.027	0.189	
112.5-157.5	12.4	11	-0.001	0.000	-0.001	
157.5-202.5	11.7	11	0.051	-0.007	0.058	
202.5-247.5	12.3	6	-0.588	0.060	-0.649	
247.5-292.5	11.8	6	-0.552	0.055	-0.607	
292.5-337.5	11.7	16	0.423	-0.073	0.496	
-337.5-360	6.2	5	-0.095	0.006	-0.101	
Slope (%)						0.64
0 – 12	21.4	9	-0.791	0.139	-0.930	
12 – 28	25.7	24	-0.007	0.002	-0.009	
28 – 42	19.9	18	-0.037	0.009	-0.046	
42 – 56	13.1	18	0.378	-0.072	0.449	
56 – 69	8.5	10	0.222	-0.023	0.245	
69 – 81	5.5	1	-1.604	0.046	-1.650	
94 – 106	1.4	1	-0.265	0.003	-0.269	
164 – 180	0.1	13	5.152	-0.146	5.298	
NDVI						0.64
Nonvegetation	4.4	4	0.108	-0.005	0.113	
Lower dense	4.1	4	0.181	-0.008	0.189	
Dense	5.9	7	0.362	-0.028	0.390	
Higher dense	13.8	28	0.898	-0.265	1.163	
Highest dense	71.8	39	-0.406	0.616	-1.022	
Proximity to stream (m)						0.62
100-200	2.6	4	0.706	-0.028	0.735	
200-300	2.6	1	-0.654	0.013	-0.666	
300-400	2.5	2	0.085	-0.002	0.087	
400-500	2.5	1	-0.594	0.011	-0.606	
500-1000	11.7	9	0.036	-0.005	0.041	
1000-2000	21.0	16	0.031	-0.008	0.039	
2000-3000	17.5	22	0.526	-0.159	0.685	
3000-4000	14.5	9	-0.172	0.026	-0.199	
4000-5000	12.4	4	-0.811	0.076	-0.887	
5000-6000	9.6	6	-0.169	0.016	-0.185	
Rainfall intensity (mm)						0.61
<1000	3.7	1	-1.081	0.025	-1.106	
1000-1500	17.1	24	0.532	-0.157	0.688	
1500-2000	55.1	50	0.100	-0.138	0.238	
2000-2500	15.2	2	-1.769	0.139	-1.908	
>2500	8.9	5	-0.371	0.030	-0.401	

Note:
- The table does not present the parameter with an AUC value $\leq$ 0.6
- The table does not present the classes of parameters with no existing landslide occurrence

The landslide susceptibility zone map is derived based on the AUC and WoE calculation, which comprises four classes: very low, low, moderate, and high. For comparing results, data is processed in ArcGIS software using a table containing class_ID of classification information. The Total WoE calculation in Arc GIS in step C.1 combined with a table contains information of class_ID. The result of the landslides susceptibility map in R and Arc GIS is seen in Figure 6. The image shows no difference between modeling algorithm mapping in R program and Arc GIS. It means that the arranged algorithm is correct.
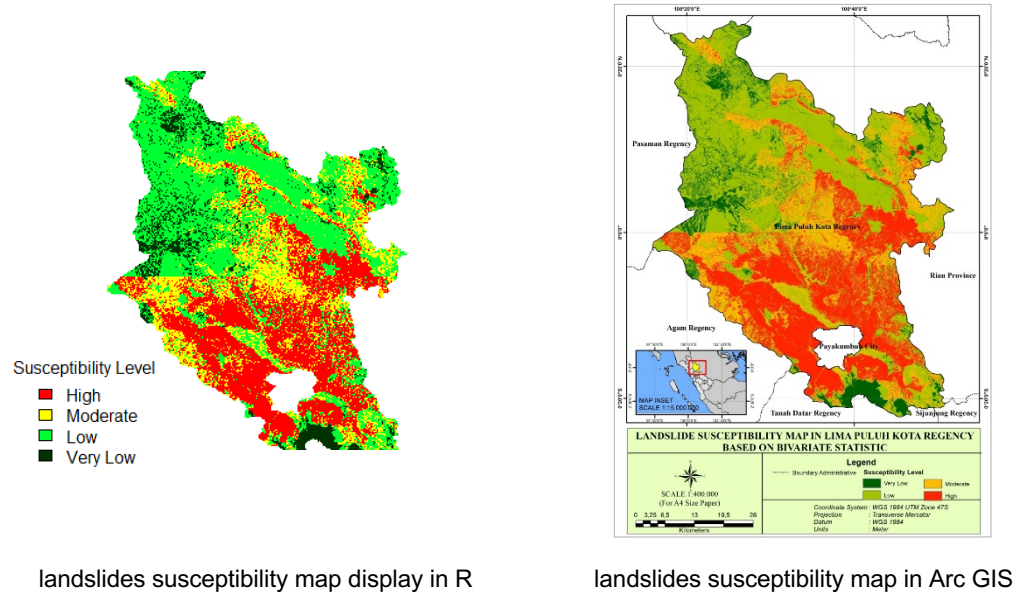

landslides susceptibility map display in R   landslides susceptibility map in Arc GIS

Figure 6. Display of landslides susceptibility map

Conclusions

R is an open-source programming language widely used because it can integrate data, analysis, and graphs in a single narrative. We use this program to model landslide susceptibility algorithm using the WoE method and apply it to a region. The result of this model is no different from using ArcGIS software. However, creating a landslides susceptibility algorithm in R model has an advantage in that other researchers can reinterpret and reevaluate the program by modifying its syntax and codes to get a more comprehensive and appropriate model applying in a specific region.

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

Acknowledgment

References

[1]   Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. Earth-Science Reviews, 180, 60–91.
[2]   Rossi, M., Reichenbach, P., "LAND-SE: A software for statistically based landslide, "Geoscientific Model Development, pp. 9533-9543, 2016.
[3]   Fausto, G., Alberto C., Mauro C., Paola R., "Landslide hazard evaluation: A review of current techniques and their application in a multi-scale study, Central Italy," Geomorphology, pp. 181-216, 1999.
[4]   Anis, Z., Wissem, G.,Vali, V., Smida, H., Essghaie, G.M. (2019). GIS-based landslide susceptibility

mapping using bivariate statistical methods in North-western Tunisia.Open Geosciences, 11:708–726 October 2019.

[5]   Pamela, Sadisun, I. A., Arifianti, Y. (2018). Weights of Evidence Method for Landslide Susceptibility Mapping in Takengon, Central Aceh, Indonesia. IOP Conference Series: Earth and Environmental Science, 118, 012037.

[6]   Silalahi, F.E.S., Pamela, Arifianti, Y., Hidayat, F. (2019). Landslide susceptibility assessment using frequency ratio model in Bogor, West Java, Indonesia. Geoscience Letters, 6(1).

[7]   Singh Pradhan, A. M., Dawadi, A., Kim, Y. T. (2012). Use of different bivariate statistical landslide susceptibility methods: A case study of Khulekhani watershed, Nepal. Journal of Nepal Geological Society, 44, 1–12.

[8]   Sumaryono, Muslim D., Sulaksana N., Triana Y.D. (2015)Weights of Evidence Method for Landslide Susceptibility Mapping in Tandikek and Damar Bancah, West Sumatra, Indonesia, International Journal of Science and Research (IJSR), Vol. 4, Issue 10, October 2015, 1283-1290.

[9]   Christos P., Christos C., "Comparison and evaluation of landslide susceptibility maps obtained from the weight of evidence, logistic regression, and artificial neural network models," Natural Hazards, vol. 93, no. August 2018, pp. 249-274, 2018.

[10]   Roger, S. B., Edzer, P. and Virgilio G.R., Applied Spatial Data Analysis with R, New York: Springer, 2013.

[11]   Hadley, W., "Tidy Data," Journal of Statistical Software, Vol. 59, no. 10, pp. 1-23, 2014.

[12]   Althuwaynee, O. F., Musakwa, W., Gumbo, T., Reis, S. (2017). Applicability of R statistics in analyzing landslides spatial patterns in Northern Turkey. 2017 2nd International Conference on Knowledge Engineering and Applications (ICKEA).

[13]   Thinnukool O., Kongchouy, N., Choonpradub C. (2014). Detection of Land Use Change Using R Program (A Case Study of Phuket Island, Thailand). Research Journal of Applied Sciences, 9: 228-237.

[14]   Tonini, Marj Abellan, Antonio. (2013). Rockfall detection from terrestrial LiDAR point clouds: A clustering approach using R. Journal of Spatial Information Science. 8. 10.5311/JOSIS.2014.8.123.

[15]   Darman, Herman (Ed) et al. 2011. Sedimentology Newsletter Sumatra. Indonesian Sedimentologist Forum (IAGI): Indonesia.

[16]   Sumaryono, Muslim D., Sulaksana N., Triana Y.D. (2015). Weights of Evidence Method for Landslide Susceptibility Mapping in Tandikek and Damar Bancah, West Sumatra, Indonesia, International Journal of Science and Research (IJSR), Volume 4 Issue 10, October 2015, 1283-1290.

[17]   Pamela, Sadisun, I. A., Arifianti, Y. (2018). Weights of Evidence Method for Landslide Susceptibility Mapping in Takengon, Central Aceh, Indonesia. IOP Conference Series: Earth and Environmental Science, 118, 012037.

[18]   Pimiento E (2010) Shallow landslide susceptibility: modeling and validation.Dept of Physical Geography and Ecosystem Science—Lund University.Thesis, pp 25-29

[19]   Silalahi, F.E.S., Pamela, Arifianti, Y., Hidayat, F. (2019). Landslide susceptibility assessment using frequency ratio model in Bogor, West Java, Indonesia. Geoscience Letters, 6(1).