Research article

Evidence-based Modeling of Landslide Hazard in Wahig-Inabanga Watershed, Bohol, Philippines using GIS and Statistical Models

TOMAS D. REYES, JR.*
Bohol Island State University, Philippines
Email: tomseyer@yahoo.com, tomseyer@gmail.com

Received 3 March 2015   Accepted 30 November 2015   (*Corresponding Author)

Abstract. The study combined geographic information system (GIS) and statistical models in predicting landslide hazard in Wahig-Inabanga Watershed, Bohol, Philippines. The bivariate statistical analysis (BSA) and logit regression (LR) were employed for class and factor weighting, respectively, to determine landslide prone areas using eleven significant landslide-related instability factors such as elevation, slope, aspect, lithology, soil order, soil type, fault line proximity, river proximity, road proximity, rainfall, and land cover. The very satisfactory results of model evaluation warranted the application of the LR model in evidence-based landslide hazard assessment. Out of the eleven instability factors, only soil order and soil type were determined not significant. The first three most important instability factors based on the values of regression coefficients were elevation, slope, and lithology. Landslide hazard assessment revealed around 7,063 ha or 11.33% of the total area of the watershed had high to very high landslide hazard ratings. The study showed that GIS, in tandem with useful models, provided pertinent results which could be used as scientific basis for watershed management and land use planning in relation to landslide disaster risk reduction and management.

Keywords bivariate statistical analysis, GIS, landslide hazard, logit regression

INTRODUCTION

As a natural phenomenon, landslide usually occurs in high elevation, and sloping areas. Its severity is definitely influenced by the type of vegetation cover and the geomorphological structure of the soil. This geologic hazard is often triggered by excessive and continuous rainfall, and seismic events. One of the most unforgettable landslide events which struck the Philippines was that which happened in Guinsaugon, Leyte where in a matter of minutes, the homes and families in barangay of Guinsaugon were wiped out and buried under a mountain of soil and rock (Catane et al., 2006).

The province of Bohol, in particular, had experienced frequent landslides and massive erosion in the upland, and floods at lowland areas. Its susceptibility to landslides and floods had been recorded by the Mines and Geosciences Bureau (MGB) in 2007. There were more than 100 barangays identified prone to these hazards out of the 46 municipalities covered in the survey. Specifically for landslide events, 64 barangays within the Wahig-Inabanga watershed were rated with moderate to high vulnerability to landslides based on 13 out of 16 municipalities evaluated. The 52-hectare Mayana landslide which happened on July 13, 2005 was triggered by a surface-wave magnitude 4.9 earthquake with its epicenter recorded somewhere in Sierra Bullones on March 31, 2005. The landslide's average slope was only about 13%, described as elongated, oriented east-west and had a total length of 1.4 km (Catane et al., 2005).

Geospatial technology is increasingly being used in spatial decision support systems. This has been the trend since its functions and applications have been made known to the public, especially to experts, practitioners, and policy makers around the globe. In the Philippines, GIS, in tandem with remote sensing and modeling technologies, are used in many applications since the 1990s from resource assessment, land use classification, and mapping to environmental hazards assessment and management.

In the Philippines, the utilization of logit regression model and GIS for landslide susceptibility mapping in the country was initially, and perhaps solely, conducted by Lee and Evangelista (2005) in Baguio City. The method which has been used in this present study was a modification from their work with the intention of increasing the predictive power of the LR model by initially applying the numerical values of landslide frequencies as class weights derived using bivariate statistical analysis (BSA) prior to the actual logit regression analysis.

This current study assessed the applicability of GIS and logit regression combined with BSA in determining landslide prone areas in Wahig-Inabanga Watershed, Bohol, Philippines.

METHODOLOGY

Study Site

The Wahig-Inabanga Watershed is the biggest watershed in the province of Bohol, Philippines (Fig. 1). From the northwest bay of Inabanga, the river dissects the central part of the island embracing a total land area of more than 610 km^2. According to Genson (2006) and Ludevese (2006), this watershed is more or less 15.20% of the total land area of the province.

It is geographically located between 124^{o}3’36” and 124^{o}23’24” East longitude, and between 9^{o}43’48” to 10^{o}4’48” North latitude. It has two headwaters namely the Wahig and Pamacsalan Rivers, and its outlet is geographically located at 10^{o}4’12” North latitude and 124^{o}4’12” East longitude.

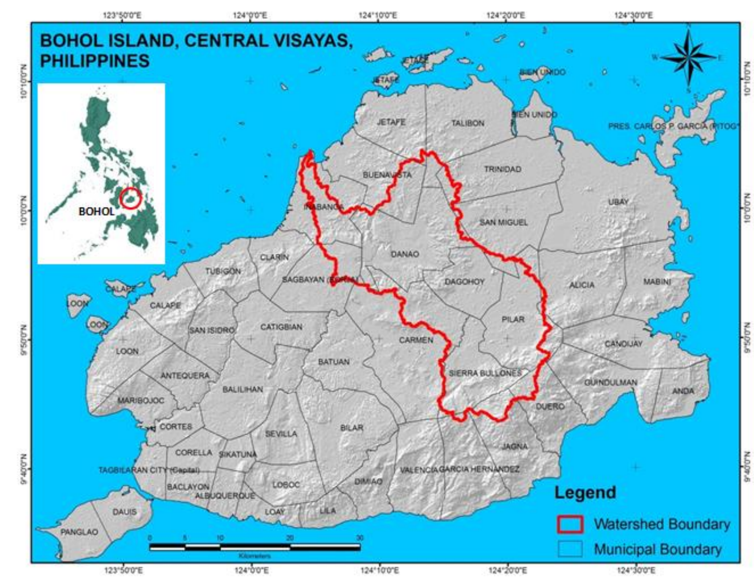

Fig. 1 Location map of Wahig-Inabanga Watershed, Bohol, Philippines

Acquisition of GIS Input Files

Thematic layers and other GIS input data which were used in this study were obtained from several Philippine government authorities and well-known sources. The digital elevation model which was used to derive elevation, slope, and aspect layers were downloaded from the ASTERGDEM website.

The SPOT5 image, river and stream network, land use and land cover, contour, and spot height layers were taken from the National Mapping and Resources Information Authority (NAMRIA). The soil order and lithology layers were obtained from the Mines and Geosciences Bureau (MGB) while the soil series and soil type layers were solicited from the Bureau of Soils and Water Management (BSWM). The political boundaries and road network layers were acquired from the Bohol Provincial Planning and Development Office (PPDO), and the updated ground rupture map was secured from the Philippine Volcanology and Seismology (PHIVOLCS).

Inventory of Landslide

The identification and mapping of existing landslides are prerequisite to perform statistical analysis on the relation between the distribution of landslides and influencing parameters (Saha et al., 2005). Any landslide susceptibility or hazard assessment must begin with the collection of information on where landslides are located. Collection of information is done in many forms. Most of the researchers utilize aerial photographs and satellite image interpretation to locate landslides. Others employ the same interpretation method in tandem with ground validation. In this study, both image interpretation and extensive field surveys were performed.

The ortho-rectified Spot 5 Image (scene: 313330) acquired from NAMRIA was used to delineate visible landslides. Ground validation was carried out during the field surveys. The complexity of image interpretation for a huge watershed like Wahig-Inabanga led the conduct of extensive field surveys. Field surveys were done mostly in, but not limited to, areas within the watershed. The buffer zone of 100 m along the watershed boundary was also considered. Major [= deep-seated] and minor [= shallow] landslides were identified and their GPS coordinates were taken either at the center or along the landslide boundaries. A total of 215 landsides were recorded throughout the study area.

GIS Processing

Mapping of landslides: The geographic coordinates taken in the field were retrieved from GPS receivers using Garmin BaseCamp software (version 4.2.1). Coordinates were identified as to those taken around the corners of huge and medium-sized landslides and those taken at the center of small landslides and classified as to size of the landslides.

To prepare the landslide inventory map, coordinates were all geo-processed in ArcMap. Polygon shapefiles were created for landslides with corner coordinates, while buffer function in ArcMap interface was used for landslides with coordinates taken at the center. These, including the digitized landslides from Spot 5 Image, were all clustered into one shapefile using merge function and then projected to WGS84_UTM_Zone_51N. To complete the landslide inventory map, the projected landslide polygons were combined with the Wahig-Inabanga Watershed mask and was finally converted to raster format with a pixel size of 10-m. The landslide pixels were given a value of 1 and non-landslide pixels as 0. There were a total of 1,891 landslide pixels and 6,749,911 non-landslide pixels.

Sample size computation and selection: The power calculator (http://calculator.stat.ucla.edu /powercalc, accessed on December 2013) was used in the computation of the sample size. Setting the confidence level at 95% and confidence interval of 3, sample sizes of 1,058 and 1,067 were calculated to represent all landslide and non-landslide pixels, respectively. The rasterized landslide inventory map facilitated the preparation of combined landslides and non-landslides vector grid map with an aid of Hawth's Analysis Tool, a freeware ArcGIS extension. Landslide and non-landslide sample selection was also performed from the vector grid map using the same tool. Two sets of samples were selected, one for model generation and another for accuracy evaluation.

Statistical analysis: The landslide hazard assessment using logit regression (LR) adopted in this study followed the method of Ayalew et al., (2005) and Ayalew and Yamagishi (2005) which required class and factor weighing.

Bivariate statistical analysis (BSA): BSA was applied to the first set of samples to determine class weights through the calculation of landslide frequencies on input parameter classes. Overlay analysis and extract by samples in ArcGIS were used in this respect and the output dbase file was imported in Microsoft Excel for landslide frequency computation.

Logit regression analysis: Logit regression was employed to define factor weights. The result of BSA in excel was imported in SPSS and the method of stepwise backward selection was used. The outputs of the regression analysis included the measures of model fit, model prediction probability, and regression coefficients. The regression coefficient of each factor [also known as predictor or parameter] served as the factor weight.

The model equation adopted in the study is shown below:

$$Y = b_o + b_1 (P_1) + b_2 (P_2) + b_3 (P_3) + \ldots + b_{11} (P_{11})$$

where: Y = landslide occurrence (presence or absence)

b_o = constant value (Y intercept)

$b_1, b_2, b_3 \ldots b_{11}$ = regression coefficients

$P_1, P_2, P_3 \ldots P_{11}$ = input parameters or instability factors

Below was the equation used to compute for the probability of landslide event:

$$P = \Pr [Y=1] = 1/1+e^{(-X*B)}$$

where: P = probability of landslide occurrence

= Pr [Y=1] = 0 to 1

e = exponential function

X*B = Y value of the logit regression function

RESULTS AND DISCUSSION

Model Assessment

Out of 11 parameters, only soil order and soil type were found not significant and were eliminated in the final model. Model rerun identified elevation as the most significant parameter based on the value of regression coefficient [*B*=0.007162]. This was followed by slope and lithology (Table 1).

Table 1 Logit regression output for the regression coefficient estimation using the model of significant parameters

Predictor (Factor)	Regression Coefficient (*B*)	Std Error	P-Value
Elevation	0.007162	0.001	0.000
Slope	0.004933	0.001	0.000
Aspect	0.002949	0.001	0.000
Lithology	0.003533	0.000	0.000
Fault line proximity	0.000825	0.000	0.000
River proximity	-0.001247	0.000	0.000
Road proximity	-0.001002	0.000	0.000
Rainfall	0.000717	0.000	0.000
Land cover	0.001842	0.000	0.000
Constant	-6.069435	0.535	0.000

Model assessment in SPSS was very satisfactory. The model of goodness of fit was significant [*p*-value < 5%], model chi-square was not significant [*p*-value > 5%], and the pseudo R^2 was relatively high [pseudo R^2 = 0.607] (Table 2). The P-value of less than 5% for the goodness of fit test shows that the LR model fits with the selected data. This is supported by the model chi-square P-value of more than 5% which implies the model fit is good. In addition, the pseudo R^2 value 0.607 means that the 60.7% of the total variation in the data is explained by the model. According to Clark and Hosking (1986) as cited by Ayalew et al., (2005) and Ayalew and Yamagishi (2005), a pseudo R^2 greater than 0.2 indicates a relatively good fit.

Moreover, the prediction probability was determined very high [PredProb = 83.2%] (Table 3). This value denotes a very acceptable prediction accuracy of the LR model for both cases (presence [84%] and absence [82.4%] of landslides).

Table 2 Summary comparison of model fit statistics of the logit regression function

Statistics	Value
Goodness of Fit	
Omnibus Test Chi-square	1290.114
P-value	0.000
Model Chi-square	
Hosmer and Lemeshow Test	17.330
P-value	0.051
Pseudo R^2	
Nagelkerke R^2	0.607

Table 3 Classification table showing the results of the landslide model prediction probability

Model	Observed	Predicted		Percentage Correct
		Absence	Presence	
LR with significant factors only	Absence	879	188	82.4
	Presence	169	889	84.0
Overall Percentage				83.2

Landslide Hazard Assessment

Result of the landslide hazard assessment indicated that more than 60% of the total area of the watershed (about 38,180 ha) was identified to have very low probability of landslide occurrence. About 16.63% or 10,360 ha had low landslide hazard, while roughly 6,692 ha or 10.74% was estimated to fall under the moderate landslide class. Conversely, high and very high landslide ratings were predicted for areas mostly situated on the upper elevations of the watershed, as shown on Figure 2, having 4,101 ha and 2,962 ha, respectively. These amounted to a total of 7,063 ha or 11.33% of the total area of the watershed.

CONCLUSION

Logit regression, with bivariate statistical analysis, is best applied if landslide inventory is available. The landslide inventory aids in the quantitative assessment of landslide prediction probability and serves as a means of determining the accuracy of the model.

The application of bivariate statistical analysis in determining landslide frequencies and the utilization of these computed frequencies as quantitative substitutes to nominal variables of landslide predictors simplify the logit regression equation and the interpretation of its results.

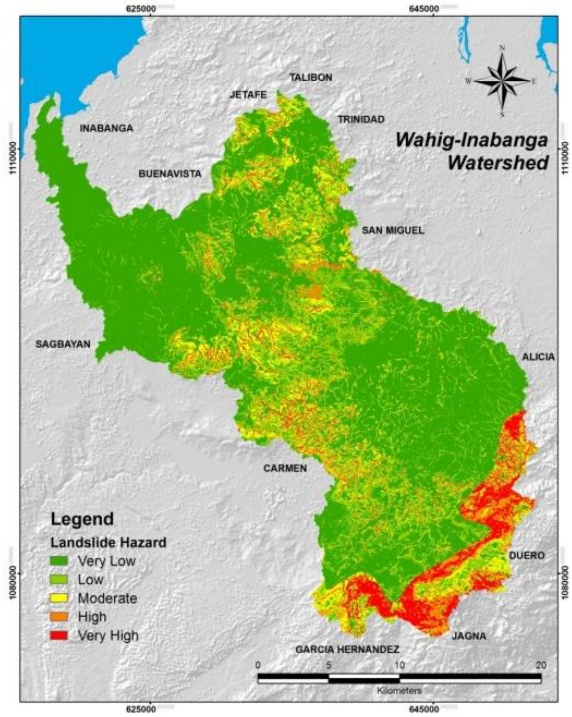

Fig. 2 Landslide hazard map generated from the LR model

Modeling landslide is very important to measure the relationship between each causative factor with every single landslide location. The relationship between landslide and its causative factors vary according to area, time, and climate. By modeling landslide, the inherent characteristics of landslide activities can be quantified. This is very important in order to identify which causative factor plays a major or a minor role. Such information can then help the authority to plan the activities and land utilization in areas prone to landslides.

REFERENCES

Ayalew, L., Yamagishi, H., Marui, H. and Kanno, T. 2005. Landslides in Sado Island of Japan, Part II. GIS-based susceptibility mapping with comparisons of results from two methods and verifications. Engineering and Geology, 81, 432-445.

Ayalew, L. and Yamagishi, H. 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology, 65, 15-31.

Catane, S.G., Foronda, F. and Tamorong, C. JR. 2005. Assessment of hazards resulting from July 11, 2005 landslide, Barangay Mayana, Jagna, Bohol, A report to the local government unit of Jagna, Bohol. Unpublished Technical Report. National Institute of Geological Sciences. University of the Philippines, Diliman, Quezon City.

Catane, S.G., Foronda, F. and Tamorong, C. JR. 2006. Extremely rapid rockslide-debris avalanche at St. Bernard, Southern Leyte, Philippines. Earth Surface Processes and Landforms, 28, 925-950

Genson, I.C. 2006. Erosion and water resources assessment in the upper Inbanga Watershed, Philippines, Application of WEPP and GIS tools. Masteral Thesis. University of Western Sydney, Hawkesbury Campus, Richmond, New South Wales, Australia.

Lee, S. and Evangelista, D.G. 2005. Landslide susceptibility mapping using probability and statistics models in Baguio City, Philippines.

Ludevese, E.G. 2006. Land suitability evaluation of the Wahig-Inabanga river watershed forest (Bohol, Philippines) Using Geographic Information System. Masteral Thesis. Visayas State University, VISCA, Baybay City, Leyte.

Mines and Geosciences Bureau (MGB). 2007. Report on landslide and flood prone areas in Bohol. Unpublished Technical Report from Mines and Geosciences Bureau, Region 7.

Saha, A.K., Gupta, R.P., Arora, M.K., Sarkar, I. and Csaplovics, E. 2005. An approach for GIS-based.