Bulletin of the Department of Geology

Bulletin of the Department of Geology, Tribhuvan University, Kathmandu, Nepal, Vol. 10, 2007, pp. 63–70

GIS-based quantitative landslide hazard prediction modelling in natural hillslope, Agra Khola watershed, central Nepal

*Prem Bahadur Thapa[1] and Tetsuro Esaki[2]

[1]*Department of Geology, Tri-Chandra Campus, Tribhuvan University, Kathmandu, Nepal*

[2]*Institute of Environmental System, Faculty of Engineering, Kyushu University, Fukuoka, Japan*

ABSTRACT

The influence of geological and geomorphological variables were spatially integrated to develop landslide hazard prediction model in the Agra Khola watershed of central Nepal where a large number of landslides triggered off due to extreme weather event of July 19–21, 1993. A quantitative technique of multivariate analysis was performed to predict elements or observations of landslides successfully into different hazard levels in the area. The predicted landslide hazard was validated and spatially relevancy of the prediction is established. The GIS-based prediction model possessed objectivity and reproducibility, and also improved the landslide hazard mapping in the natural hillslope.

INTRODUCTION

A quantitative approach of statistical analysis was selected to develop a predictive model for landslide hazard in the Agra Khola watershed of central Nepal due to existence of limitation in the deterministic models as it needs geo-material data (mechanical properties, water saturation, etc.), which are difficult to obtain over large areas (Terlien et al. 1995). Multivariate model of the relation between the potential of landslide (dependent variable) and a set of intrinsic properties (independent variables) were constructed by using Geographic Information System (GIS). The GIS-based landslide hazard analysis illustrated the spatial relevancy of the prediction with objectivity of the model in the similar terrains.

STUDY AREA

The study area lies in the central Nepal and is bounded by the latitudes 27°36’ and 27°45’N, and the longitudes 84°58’ and 85°7’E. The Agra Khola is the main draining river in the study area and the altitude of the study area ranges from 600 to 2480 m. The total area of the watershed is 111.8 sq. km.

The high intensity rainfall of 19–21 July 1993 triggered off numerous rockslides and soil slides in the Agra Khola watershed. During that rainfall event, the volume of maximum precipitation recorded within 24 hours in the nearest rain gauge station at Tistung was 540 mm. The entire catchment received intense rainfall of 300–500 mm between 19 and 21 July of 1993 near the Tistung–Palung region (DHM 1993). Isohyetal map of the Agra Khola area and hyetograph at the Tistung Station is shown in Fig. 1.

The geomorphology of the area exhibits highly dissected and rugged topography in the south whereas it is gentle in the north. Surficial deposits in hillslope are mainly of residual soil and colluvium including debris flow and other slope debris deposits. The residual soils have variable depths and areal extents. The colluvium is scattered in distribution along the foothills and occurs as ribbon like deposits filling drainage courses. Small alluvial deposits occur along the river valleys but alluvium is mostly confined to fans developed downslope of the colluvial deposits.

---

*Corresponding author:

E-mail address: geoscithapa@yahoo.com

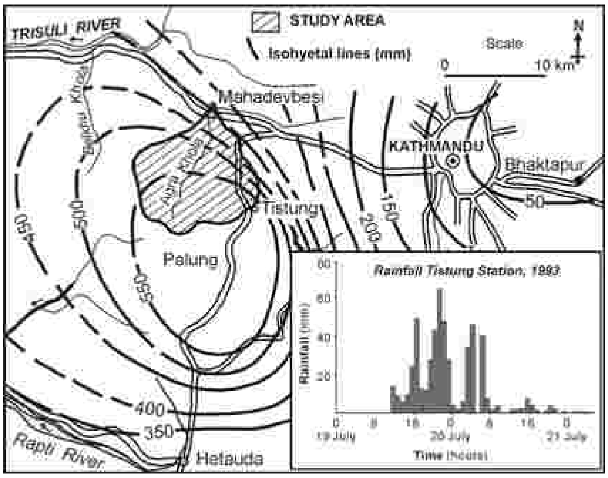


Fig. 1 Isohyetal map (24 hours) of the Agra Khola watershed area with hyetograph at the Tistung Station during July 19–21, 1993

A regolith or mantle of weathered rock occurs over most of the area.

The bedrock geology of the study area consists predominantly of low- to medium-grade metamorphic rocks. The rocks are metasandstone, slate, phyllite, marble, quartzite and schist, and their age ranges from Precambrian to Palaeozoic (Stöcklin and Bhattarai 1977; Stöcklin 1980). Intrusive rock, granite is found in southwestern part of the area whereas sedimentary strata comprising limestone crop out in eastern region.

LANDSLIDE CHARACTERISTICS

The major slope failures were occurred at the uppermost catchment of the study area due to extreme weather event of July 19–21, 1993. The large deep-seated rockslides were found in the north-facing dip-slopes, whereas shallow slides were observed on the counter dip-slopes and the area occupied by granite. During the incidence, upper catchment area of the watershed was severely damaged causing the 42 casualties (Thapa and Dhital 2000). However, number of landslides were frequently seen in the middle reaches of the watershed with huge debris deposition in the middle section of the rivers. A four span bridge located at Mavadevbesi of the Prithvi Highway connecting the capital city, Kathmandu was washed away early in the morning of July 20, 1993. Segments of the Tribhuvan Highway, which lies in the Agra Khola watershed, were also damaged from many rock and soil slides.

Landslide characteristics in the study area were derived by using GIS analytical procedure. The procedure utilized the detailed spatial database of the area to integrate the evidence of the past slope failures and their causuative variables for landslide occurrences.

Spatial database and variable analysis

The variables for the derivation of landslide characteristics consisted of various thematic maps acquired from different sources (Table 1 and Fig. 2). The basic data layers produced from GIS included landslide inventory and variables responsible for causing slope failures. Landslide inventory map was prepared and attribute data were assigned to individual landslide to analyze quantitatively the relationship between landslides and their causative factors. Landslide conditioning variable maps were obtained as derivative layers by spatial analysis or digitized layers of field-surveyed maps. A Triangulated Irregular Network (TIN) model was created from digital topographic map of 1:25,000 scale and was used to generate Digital Elevation Model (DEM) at a resolution of 10 x 10 m grid size. The DEM was utilized to produce very significant derivative layers, including slope angle and slope aspect, which are main predisposing factors for the landslide activity. Drainage lines were extracted from DEM by hydrologic modelling and Strahler classification function was used to rank the stream order into five different classes.

Table 1: Data layers for GIS analysis in the study area

Classification	Coverage	Spatial data	Attribute data
Geological hazard	Landslide	Polygon	Nominal
Damageable object	Building, Road	Point, polyline	Nominal
Basic map	Topography	Point, polyline	Nominal, interval
	Geology	Polygon	Nominal
	Eng. geology	Polygon	Nominal
	Land use	Polygon	Nominal
Hydrologic data	Precipitation	Point, polyline	Nominal, interval

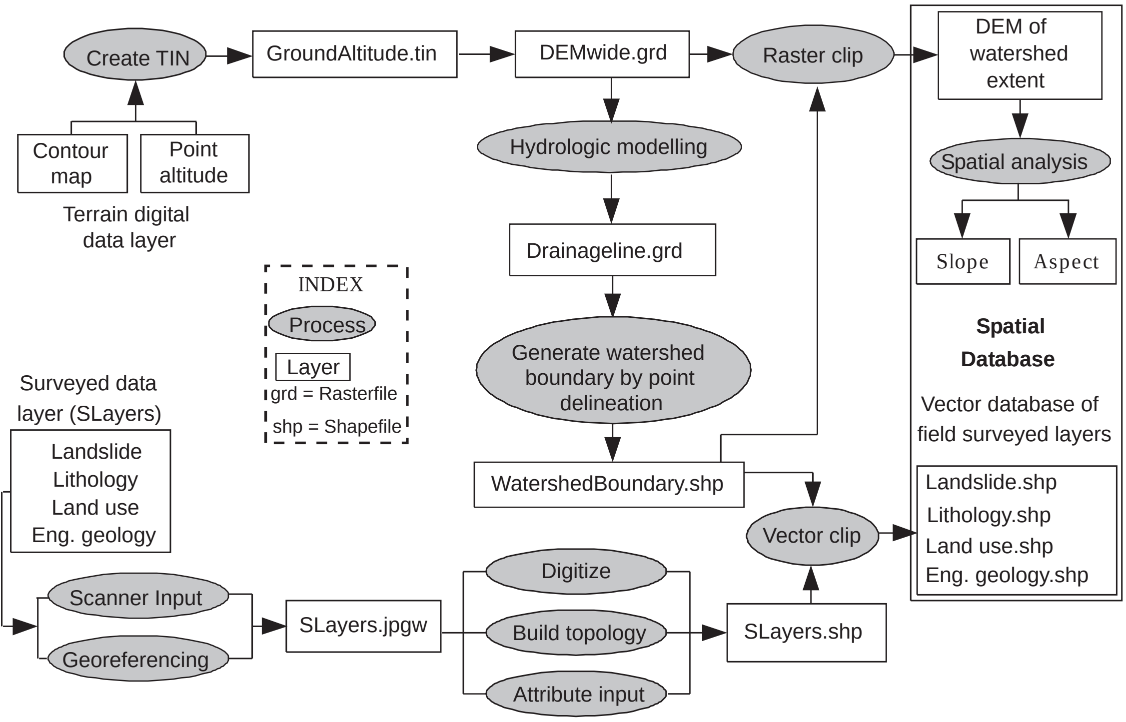


Fig. 2 GIS process for spatial database creation

Vector layers of field surveyed physical variables like lithology, land use, and engineering geological maps were prepared. Lithology was obtained from the geological map which consisted of six different lithological units. Engineering geological map was represented in a discrete way. Soil depths were estimated in the field and two different subclasses were considered: thin soil (1–3 m) and thick soil (>3 m). The soil thickness <1 m was considered as rocky terrain and was categorized into low, medium, and high rock mass strengths, which were based on measurement of intact rock and rock mass properties. The land use layer was produced by editing the existing digital polygon shapefile with field verification and was classified into five categories.

All thematic vector layers were rasterized for analytical purpose and grid resolution of 10 m was utilized to obtain the detailed morphological information. The grid format was considered optimum for this kind of process, as the sizes of the smallest landslide occurrence represented in the analysis format was 10 m. Rasterization of linear element was done by creating appropriate buffer around the linear features. Continuous variables were transformed into discrete classes with reclassification criteria that were made between the required limited number of classes, which sufficiently represented the wide range of original categories in each class. Natural slope angle was classified into five categories. The slope aspect variable was classified in eight major orientations with the addition of flat areas. Elevation map was reclassified with interval of 200 m.

A landslide inventory map was converted to 10 m grid file. Each cell was assigned ‘‘0’’ if no landslide is present or ‘‘1’’ if a landslide is present and ‘‘no data’’ code was assigned if the cell is outside the study area. The landslide grid file and causative variable grid files were logically compared to ensure that covered a common area then combined and converted to centre of each cell. All locations of pixel of landslides studied were used to extract from the existing data layers, the physical parameters that characterize landslide locations. A base map of point feature with 10 m distance interval was prepared and

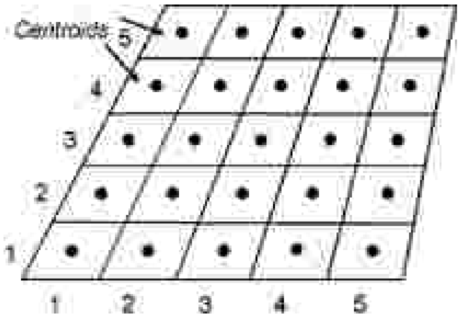


Fig. 3 Centroid of spatial extent

these points were centroid of the grid cells in raster layers (Fig. 3). An algorithm was executed to extract pixel value to point layer from all raster layers of causative variables. Thus, all attributes extracted were stored in point base map file indicating the presence and absence of the landslide occurrence. The categorical variable maps (e.g. lithology) were extracted as nominal value and their corresponding pixel counts. Variables like elevation, slope angle, and slope aspect were continuous values and were further reclassified to discrete classes by using Spatial Analysis Module in GIS.

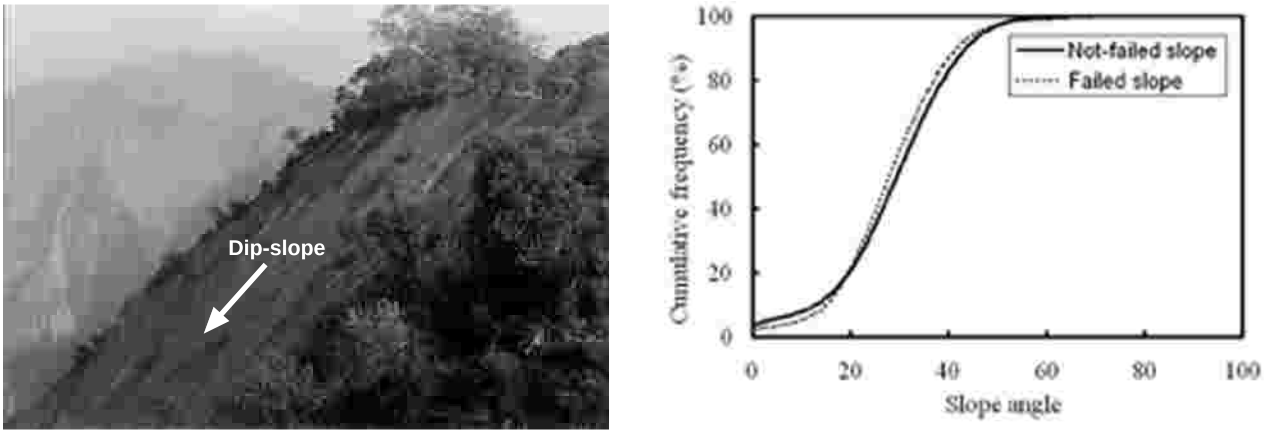
Landslide characteristics were computed using attribute column values in the landslide base point map to quantify the influence of various causative variables. New thematic map layers were generated from the calculated values and were visualized. Zonal statistics was also used to extract landslide attribute in each variable class of the thematic data layers and frequency of failures occurrence were evaluated.

Landslide and causative variables relationship

To examine the physical variables contributing to the initiation of landslides, the landslides that occurred in the study area were correlated with those parameters considered to have influence on their occurrence. The univariate probability analysis was adopted to quantify the relationship between landslide frequency and the each variable class. The predominant slope failures are found in quartzite and schist unit, which is due to favourable orientation of foliation in the direction of slope face (Fig. 4a).The occurrence of slope failures is reached maximum at slope angle of 27^o within slope range of 25^o–35^o (Fig. 4b). When slopes exceed 45^o, a sharp decrease in the landslide frequency values is observed and slope failures are more apt to fail as rockslides. Ii is evident that slope gradient and geological variables are the most distinct pre-disposing factors for landslide events in the study area.

QUANTITATIVE LANDSLIDE HAZARD MODELLING

Spatial prediction modelling of landslide hazard was performed by integration of several spatial data layers (i.e., variables). Quantitative technique of spatial prediction model was developed to aim at a higher degree of objectivity and better reproducibility of the hazard prediction. The model assumes the



(a)

(b)

Fig. 4 (a) Spatial localization of slope failures due to dip-slopes and (b) Plot of not-failed and failed slopes likelihood of landslides occurrence by statistical relationships between past landslides of a given type and spatial data layers of causative variables.

The acquisition of the variables for the prediction model was carried out by extracting causative variables for all the landslides by considering the relevancy of their influence on the occurrence of landslides. The overlay function was used and the result from the overlays was then transferred to a matrix of causative variables. Spatial design in the model was represented by mapping units (slope units). The slope units were automatically identified in the entire watershed from DEM by hydrologic analysis using "Arc Hydro Tool" ESRI and "Slope Unit Tool" (Esaki et al. 2004).

The main steps in the procedure are shown in Fig. 5 that discusses the flow chart for hazard prediction. In the prediction model, the optimum mapping units were defined and preliminarily selection of variables were carried by univariate analysis.

Due to the binary character of the response and predictor variables, and the dubious normality of some of the variables, a logistic regression procedure of multivariate analysis was selected. All input variables were grouped into a few meaningful classes. For categorical data, such as lithology, engineering geology, and land use did not involve any subjective judgment. For continuous variables, such as slope angle or elevation, the selection of the number of classes and class limits requires a significant amount of guess work guided by previous knowledge of the causal relationships between slope failures and instability factors (Guzzetti et al. 1999).

Using the overlay capabilities of GIS, the attribute data were georeferenced to the mapping units (slope units). An important aspect was the conversion of various parameters from nominal to numeric, e.g., lithology, land use. This can be done through the creation of dummy variable matrix automatically (Fig. 6).

The regional GIS database was then exported to statistical software for analysis (SPSS 1997). The technique of logistic regression yields coefficients for each variable based on data derived taken across a study area. These coefficients serve as weights in an algorithm, which were used in the GIS database to produce a map depicting the probability of landslide occurrence. Quantitatively, the relationship between the occurrence and its dependency on several variables is expressed as:

$$\Pr(\text{event}) = 1/(1+e^{-Z}) \qquad (1)$$

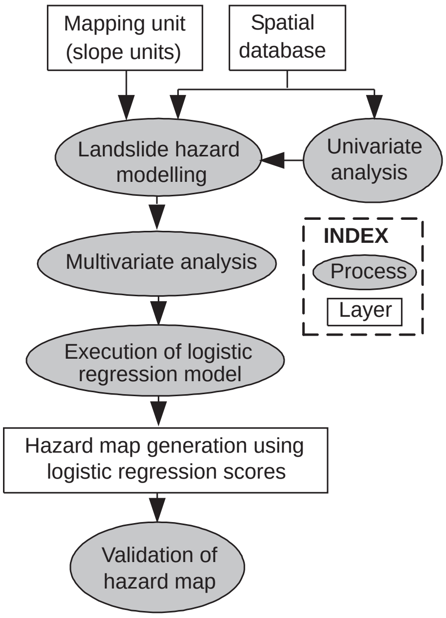


Fig. 5 Landslide hazard prediction modelling procedure

$$\begin{bmatrix} 21 & 0 & 0 & 11 & 23 & 0 \\ 0 & 76 & 56 & 0 & 15 & 0 \\ 0 & 0 & 87 & 0 & 98 & 7 \\ 98 & 0 & 0 & 90 & 0 & 0 \\ 34 & 41 & 33 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Fig. 6 An illustration of matrix conversion from integer to dummy variable coding (Esaki et al. 2005)

Where Pr(event) is the probability of an event occurring. In the present situation, the Pr(event) is the estimated probability of landslide occurrence. As Z varies from $-\propto$ to $+\propto$, the probability varied from 0 to 1 on an S-shaped curve. Z is the linear combination:

$$Z = B_0+B_1X_1+B_2X_2+...+B_nX_n \qquad (2)$$

where, B_i (i=0, 1,..., n) is the coefficient estimated from the sample data, n is the number of independent variables (i.e. landslide-related physical parameters), and X_i (i=1, 2,..., n) is the independent variable.

The coefficients for the final logistic regression are shown in Table 2 and Fig. 7. It should be noted that all the variables in the model were binary variables representing presences or absences of the corresponding variables. For each variable, the last category was used as the default reference category, and the coefficient of that category was thus overridden.

Table 2: Regression coefficient for the hazard prediction model

Variables	Coef.
Slope angle	
$<15^{o}$	-0.217
$15^{o}–25^{o}$	0.074
$25^{o}–35^{o}$	0.778
$35^{o}–45^{o}$	0.417
$>45^{o}$	-0.145
Slope aspect	
Flat	-0.385
North (N)	0.117
North East (NE)	0.253
East (E)	0.590
South East (SE)	-0.177
South (S)	0.195
South West (SW)	-0.348
West (W)	-0.027
North West (NW)	0.333
Engineering geology	
Thin soil [1–3 m] (TnSl)	-0.203
Thick soil [>3 m] (TkSl)	-0.272
Colluvium (Clv)	-0.876
Alluvium (Alv)	0.240
High Rock Mass Strength (HRMS)	-0.868
Medium Rock Mass Strength (MRMS)	0.222
Low Rock Mass Strength (LRMS)	1.420
Slope complexity	
Granite slope (GS)	-1.662
Oblique slope (OS)	-0.349
Dip-slope<slope (DS-EL)	1.357
Dip-slope>slope (DS-G)	-0.023
Counter dip-slope (CDS)	-0.163
Fractured zone (FZ)	- 1.030
Land use	
Forest (Fo)	1.536
Shrub land (SrL)	-0.124
Grassland (GrL)	0.880
Cultivated land (CuL)	0.657
Barren land (BaL)	2.845
Constant	-3.640

Modelling Procedure, Result and Validation

In logistic regression analysis, variables were evaluated for removal one by one if they do not contribute sufficiently to the regression equation. The likelihood-ratio test was used for determining whether variables should be added to the model. If the observed significance level was greater than the probability for remaining in the model (0.1 in this study), the variable was removed from the model and the model statistics were recalculated to see if any other variables are eligible for removal. The elevation and distance to drainage variables were removed from the model due to their lower significance values. Slope complexity was used in analysis as the slope failures are more influenced by bed rock layering than the different rock units. The final coefficients of logistic regression values from SPSS were imported back into the GIS and analyzed to prepare landslide hazard map.

The landslide hazard map generated from the analysis was classified into different levels of hazard for general use. Practically, there is no straightforward statistical rule to categorize continuous data automatically and always unclear in landslide hazard mapping. Most of the researchers use their expert opinion basis along with available classification methods to develop class boundaries. The hazard levels were categorized into very low, low, medium, high, and very high hazards using natural junk break method and overlaying of past landslide map for the adjustment of class boundaries.

In prediction modelling, the most important and the essential components are to carry out a validation of the prediction results. Some works have addressed the issue of map validation (Carrara 1983; Brabb 1984; Yin and Yan 1988; Carrara et al. 1991; van Westen 1993; Carrara et al. 1995; Chung et al. 1995; Luzi and Pergalani 1996; Chung and Fabbri 2003; Remondo et al. 2003). The methodology presented here is rather simple and useful procedure for the

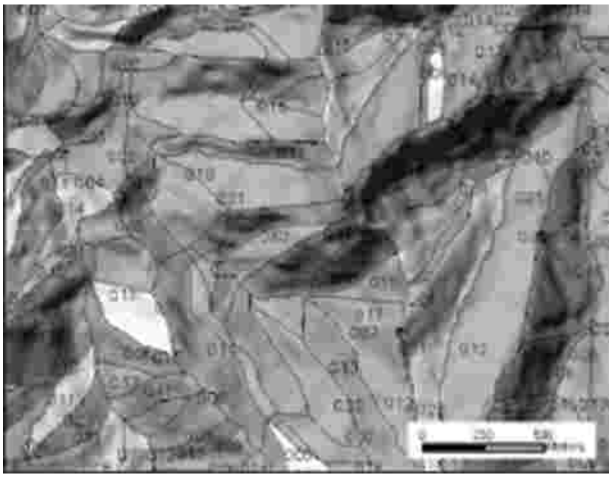


Fig. 7 An illustration of numerical probability value based on logistic regression coefficient in slope units

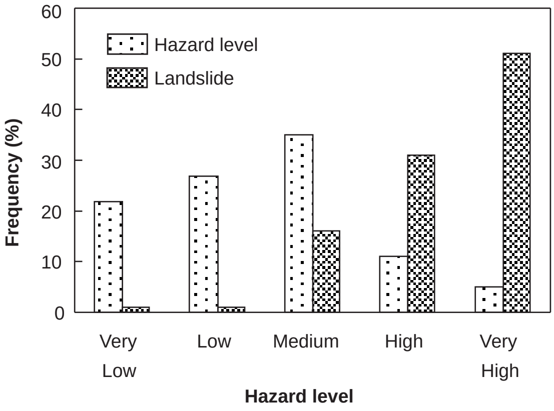


Fig. 8 Validation of predicted landslide hazard (Thapa et al. 2004)

validation, so that the prediction results can be interpreted meaningfully with respect to the future landslide hazard. Map crossing between predicted hazard map and slope failure of past events was carried out on raster maps. The calculated and classified hazard levels were found in good agreement with occurrence of pre-existing landslides (Fig. 8). Validation has shown useful not only to determine the predictive value of maps but also improved the different steps in map-making process or to obtain hazard map.

Spatial relevance of hazard prediction

Application to the study area aims to establish whether and to what extent a prediction can be extended, in space, to neighbourhood areas with similar geomorphology or geology. The accuracy is obtained by multivariate technique due to its functionality of forward parameter selection and backward removal. Validated results with respect to past incidence of landside events clarified the spatial relevancy of predicted landslide hazard in the natural hillslope terrain.

CONCLUSIONS

Likelihood occurrence of slope failures with respect to causative variables verified that slope and bedrock layering variables are found to be the most influential factors. Natural slope angle is the distinct pre-disposing factor for slope failures and maximum failures were found at slope angle of 27^o within range of 25^o–35^o. The predominant failures were seen in quartzite and schist where thin soil resting on the bedrock.

Statistical approaches of the quantitative method have established the successful modelling of landslide hazard in terms of spatial relevancy. The analysis indicated that the use of geomorphological and geological variables in the analysis improved the overall accuracy of the final hazard map considerably. Also, demonstrated that prediction modelling minimizes the effect of error in selecting input variables. The methodology was typically data driven and therefore highly objective.

The modelling result showed that very high hazardous zones were mainly confined in upper and middle reaches of the study area and such a spatial localization of very high hazards was attributed to the presence of favourable steep topography and orientation of foliation in the direction of slope face. A few localized very high hazardous zones were embedded in the medium to high hazardous zones in some parts of the Agra Khola watershed, central Nepal.

ACKNOWLEDGEMENT

Authors are thankful to Ministry of Education, Culture, Sports, Science and Technology, Government of Japan for providing Monbukagakusho scholarship to this study. The study was carried out while persuing the Doctoral Research in Kyushu University.

REFERENCES

Brabb, E.E., 1984. Innovative approaches to landslide hazard and risk mapping. In the Proceedings of 4th International Symposium on Landslides, 1, Toronto, Canada, pp. 307–324.

Carrara, A., 1983. A multivariate model for landslide hazard evaluation. Mathematical Geology, 15, pp. 403–426.

Carrara, A., Cardinali, M., Detti, R., Guzetti, F., Pasqui, V. and Reichenbach, P., 1991. GIS techniques and statistical models in evaluating landslide hazard. Earth Surface Processes and Landforms, 16, pp. 427–445.

Carrara, A., Cardinali, M., Guzzetti, F. and Reichenbach, P., 1995. GIS-based techniques for mapping landslide hazard. In: Geographical Information Systems in Assessing Natural Hazards, Carrara, A. and Guzzetti, F. (eds.), Kluwer Publications, Dordrecht, The Netherlands, pp. 135–176.

Chung, C.F., Fabbri, A.G. and Van Westen C.J., 1995. Multivariate regression analysis for landslide hazard zonation. In: Geographic Information Systems in Assessing Natural Hazards. Carrara, A. and Guzzetti, F. (eds.), Kluwer, Dordrecht, pp. 107–133.

Chung, C.F. and Fabbri, A.G.., 2003. Validation of Spatial Prediction Models for Landslide Hazard Mapping. Natural Hazards, Kluwer Academic Publishers, 30, Netherlands, pp. 451–472.

DHM, 1993. Climatological Records of Nepal, 1963-1993. Department of Hydrology and Meteorology (DHM), Ministry of Water Resources, Kathmandu, Nepal.

Esaki, T., Xie, M., Mitani, Y. and Zhou, G., 2004. Development of GIS-based spatial three-dimensional slope stability analysis system: 3DSlopeGIS. Jour. of Applied Computing in Civil Engineering, 13, pp 21–32.

Esaki, T., Thapa, P.B., Mitani, Y., and Ikemi, H., 2005. Landslide and Debris flow Characteristics and Hazard Mapping in Mountain Hill-slope Terrain by Using GIS, Central Nepal. In: Landslide Risk management, Hungr, O., Fell, R., Couture, R., Eberhardt, E. (eds.), Proceedings of the International Conference on Landslide Risk Management, Vancouver, Canada, A. A. Balkema, pp. 657–668.

Guzzetti, F., Carrarra, A., Cardinali, M., and Reichenbach P., 1999. Landslide hazard evaluation: A review of current techniques and their application in a multi-scale study, central Italy. Geomorphology, 31, pp. 181–216.

Luzi, L. and Pergalani, F., 1996. A methodology for slope instability zonation using a probabilistic method. Actas VI Congreso Nacional y Conferencia Internacional de Geología Ambiental y Ordenación del Territorio, 1, pp. 537–556.

Remondo, J., González, A., De Terán, J.R.D., Cendrero, A., Fabbri, A. and Chung, C.F., 2003. Validation of Landslide Susceptibility Maps; Examples and Applications from a Case Study in Northern Spain. Natural Hazards, 30: Kluwer Academic Publishers, Netherlands, pp. 437–449.

SPSS, 1997. SPSS advanced statistics 7.5. SPSS, Chicago.

Stöcklin, J. and Bhattarai, K D., 1977. Geology of Kathmandu Area and Central Mahabharat Range, Nepal Himalaya. Kathmandu: HMG/UNDP Mineral Exploration Project, Technical Report, New York, 64 p. (Unpublished Rept.).

Stöcklin, J., 1980. Geology of Nepal and its Regional Frame. Jour. Geol. Soc., London, 137, pp. 1–34.

Terlien, M.T.J., van Westen, C.J. and van Asch, T.W.J., 1995. Deterministic modeling in GIS based landslide hazard assessment. In: Geographic Information Systems in Assessing Natural Hazards, Carrara, A. and Guzzetti, F. (eds.), Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 57–78.

Thapa, P.B. and Dhital, M.R., 2000. Landslides and debris flows of 19-21 July 1993 in the Agra Khola watershed of central Nepal. Jour. Nepal Geol. Soc., v. 21, pp. 5 –20.

Thapa, P.B., Esaki, T., and Mitani, Y., 2004. Landslides and debrisflows characteristics and hazard mapping in the natural hill-slope terrain by using GIS. 2nd International Workshop on Earth Science and Technology, Kyushu University, Fukuoka, Japan, pp. 63–71.

van Western, C.J., 1993. GISSIZ: Training package for geographic information systems in slope instability zonation. Volume 1-Theory, International Institute for Aerospace Survey and Earth Science (ITC), 15, 245p.

Yin, K.L. and Yan, T.Z., 1988. Statistical prediction models for slope instability of metamorphosed rocks. In Proceedings of the 5th international symposium of landslides, Bonnard, C. (ed.), Lausanne, Switzerland, 2, Rotterdam, The Netherlands, Balkema, pp. 1269–1272.