# Comparing an Artificial Neural Network and a physically-based model for landslide hazard zonation



Mihai Pavel FPInnovations, Feric Division, Vancouver, BC, Canada R. Jonathan Fannin Dept. Civil Engineering and Dept. Forest Resources Management – University of British Columbia, Vancouver, BC, Canada John D. Nelson Dept. Forest Resources Management – University of British Columbia, Vancouver, BC, Canada

## ABSTRACT

This study compares two methods for landslide hazard zonation. The first method uses an Artificial Neural Network (ANN) and consists of replicating an existing subjective geomorphic mapping produced by terrain specialists. The second method employs a physically-based method and is implemented as a software model. The mappings produced by the two methods are compared and the advantages, disadvantages, and practical applicability of each method are discussed.

## RÉSUMÉ

Cette étude compare deux méthodes pour la zonation de risque des glissements de terrain. La première méthode emploie un réseau neurologique artificiel et reproduit une carte géomorphique existant produite par des spécialistes en terrain. La deuxième méthode utilise une méthode physique-basée et est mise en application comme un logiciel. Les résultats produits par les deux méthodes sont comparés et les avantages, les inconvénients, et l'applicabilité pratique de chaque méthode sont discutés.

#### 1 INTRODUCTION

Landslide hazard assessment is a complex activity for which numerous methods of analysis have been developed, among which subjective geomorphic mapping is probably the most commonly used (Keaton and DeGraff 1996; Soeters and van Westen 1996; Carrara et al. 2003). Subjective geomorphic mapping consists of delineating terrain polygons that are relatively uniform with respect to surficial materials, landforms, and geomorphic processes. The method relies heavily on the skills and experience of the mapper, and therefore, its major drawbacks are the high costs and a potential for lack of consistency between products generated by different terrain mappers. To address these shortcomings Pavel et al. (2008) developed an approach for costeffective and consistent replication of subjective geomorphic mappings by using a type of Artificial Neural Network (ANN) named Learning Vector Quantization (LVQ).

An alternate method for landslide hazard assessment utilizes physically-based models. These models use principles of soil mechanics and groundwater hydrology combined with digital elevation models (DEM) of the terrain to assess landslide hazard. Various assumptions about groundwater flow and distribution of soil parameters over the area analyzed are usually employed. One of the more sophisticated physically-based models for landslide hazard zonation is SINMAP (Pack et al. 2005) which was used in this study.

The main objective of this study is to conduct a comparison between the LVQ-based method and the SINMAP method in the specific context of forest development planning in British Columbia (BC). As with other models developed for natural resources management, these models are not meant to replace the analyst but to assist in the terrain stability mapping. These models are assessed with respect to their ability to delineate areas of potential instability, and identify zones that require detailed ground checks, thereby limiting expensive ground checks to only the most landslide-prone areas.

#### 2 STUDY SITE

This study was conducted using data from the Lower Seymour watershed within the Metro Vancouver (formerly, Greater Vancouver Regional District), situated in southwestern BC. The Seymour watershed lies within the Pacific Ranges of the Coast Mountains, and has a total area of 56.7 km<sup>2</sup>. The watershed was included in a comprehensive ecological and geomorphologic study Anon. (1999) which constitutes the main data source for this analysis. The terrain is characterized by rugged topography, with elevation ranging from 40 m to about 1 500 m. Fig. 1 displays a Digital Elevation Model (DEM) of the study area, based on a 20-m cell raster using the ArcView<sup>TM</sup> Geographic Information System (GIS). The locations of 212 existing landslides were identified from aerial photographs and ground-truthing.



Figure 1. Seymour - DEM, roads, streams, and location of existing landslides.

Most common surficial deposits in Seymour consist of silt and coarse gravel. More recent deposits, mostly coarse-textured, have been deposited on gentle slopes by streams and debris flows, and on steeper slopes by rock fall. The valley floors represent a particular challenge with respect to stability, because the surficial materials overlie glaciolacustrine deposits. Although these areas are characterized by low slopes, a high clay content makes them prone to instability, as confirmed by the relatively large number of landslides in this lower part of the watershed.

The terrain stability map of Seymour developed by Anon. (1999) is also presented in Fig. 2. Within Seymour, 397 polygons were delineated according to the BC terrain classification system (Howes and Kenk 1997) and stability classes from Class I (stable) to Class V (most unstable) were assigned to each polygon based on principles reported in Province of BC (1999). Terrain stability classes provide a relative ranking of the likelihood of a landslide occurring after timber harvesting or road construction.



Figure 2. Seymour - terrain stability mapping and location of existing landslides.

#### 3 LANDSLIDE HAZARD ZONATION USING THE LEARNING VECTOR QUANTIZATION (LVQ) ALGORITHM

In general, Artificial Neural Networks (ANN) are information-processing systems that have certain performance characteristics in common with biological neural networks, and have been developed as generalizations of mathematical models of human cognition and neural biology. In an ANN-based analysis, each terrain unit is assigned a series of topographic and geomorphic attributes relevant to stability, e.g., elevation, slope, aspect, type of surficial material and texture, existing geomorphic processes, etc. Thus, in our study, raster cells are represented by n-dimensional vectors and the terrain classification problem consists of analyzing these high-dimensional data sets.

The Learning Vector Quantization (LVQ) neural network used in this study is a relatively simple and yet efficient classification algorithm which was purposely developed for statistical pattern recognition, especially when dealing with very noisy, high-dimensional stochastic data (Kohonen 2001). The main benefit of LVQ is good recognition accuracy while at the same time significantly reducing the number of computing operations when compared with more traditional statistical methods. In this study LVQ was used for supervised classification, based on the assumption that a number of examples (spatial entities) already classified are available for purposes of learning. The task is to assign new entities to various classes based on how similar they are to the examples included in these classes. Essentially, the LVQ algorithm is able to 'learn' the patterns of instability from the first dataset and then use these to classify new data.

For the purpose of this study, and also considering the practical applicability of the method, the LVQ algorithm was used to identify two groupings of terrain stability class. Taking into account the mapping presented in Fig. 2, terrain identified as Class I - III was considered stable, and terrain Class IV - V was considered unstable. The terrain stability analysis in Seymour was conducted in two steps. In the first step the LVQ algorithm was trained using a portion of the data. Next, the trained network was used to classify the remaining data. The quality of results was assessed by comparing the mapping produced by the LVQ algorithm with the existing subjective geomorphic mapping. The amount of data to separate into training and testing sets is problem dependent. Most practitioners recommend randomly splitting the data into two thirds for training and one third for testing. For our study site, the total number of raster cells in the DEM was split in accordance with this same rule. A complete description of the LVQ-based analysis for Seymour is presented in Pavel et al. (2008).

#### 4 LANDSLIDE HAZARD ZONATION USING SINMAP

SINMAP is designed to work as an extension for the ArcView<sup>TM</sup> GIS and it is described by the following features:

- It is based on an infinite slope stability model.
- Topography is represented as a DEM.
- Spatial distribution of groundwater is based on shallow subsurface flow convergence and topographic slope.
- Uncertainty of parameters is incorporated through ranges of soil and hydrologic parameters.
- It is interactively calibrated.

The measure of stability established by SINMAP (its main output) is the so-called Stability Index (SI). Essentially, a Stability Index is computed for each cell in the DEM. The formula for the SI is derived starting from the infinite slope formulation (Hammond et al. 1992), through a series of successive transformations and by using a probabilistic representation of some parameters. SINMAP uses terrain slope derived from the DEM and a series of user inputs consisting of a unique value for soil density (kg/m<sup>3</sup>), and three soil parameters (given as ranges): (i) dimensionless cohesion defined in SINMAP as a function of root cohesion, soil cohesion, soil thickness and soil density; (ii) angle of internal friction for the soil (deg.); and (iii) the ratio transmissivity / recharge. Transmissivity is a function of hydraulic conductivity and soil thickness, and recharge is a measure of water input at a certain point (this ratio has units of m). If available, the location of existing landslides is also used to improve

the accuracy of the prediction. A complete description of the method is provided in Pack et al. (2005).

Essentially, SINMAP delineates potential landslide initiation zones. The model applies only to shallow translational landslides controlled by shallow groundwater flow convergence. Apart from SI, SINMAP also outputs a calibration graph which reflects the distribution of the mapped area by stability classes. If available, existing slides are represented in the graph as well, and based on their distribution by (predicted) stability classes the input parameters are adjusted. The analysis is repeated until a satisfactory state is reached. At the end of this process, the map obtained is considered the prediction obtained with SINMAP.

The parameters used in SINMAP for this study were selected based on the field tests conducted in the Seymour area by Wilkinson (1996), Jaakkola (1998), and based on the experience of the authors. The value for soil density ( $\rho$ ) was set to 1800 kg/m<sup>3</sup>. The other parameters are input as ranges. During the calibration process these ranges were adjusted, and the list of initial and final parameters is presented in Table 1.

Table 1. The list of initial and final parameters used in SINMAP.

Parameters	Values	
	Initial	Final
Dimensionless cohesion - C	0 - 2	0 - 1.2
Angle of internal friction - φ (deg.)	28 - 47	35 - 47
Transmissivity /	200 -	1000 -
Recharge -T/R (m)	3000	3000

#### 5 RESULTS

#### 5.1 Results produced by the LVQ algorithm

The LVQ-based landslide hazard mapping is presented in Fig. 3, which depicts stable and unstable terrain as identified by this method. In addition to the LVQ-based mapping, terrain polygons Class IV and V (unstable terrain) and the location of existing landslides are also illustrated. Fig. 3 shows that terrain identified as unstable by the LVQ algorithm includes all the existing landslides.

The similarity to the actual mapping produced by the terrain specialists achieved by the LVQ-based mapping is 90.9%. The misclassifications consist of a relatively small area classified as unstable in the 'stable' zone (as identified by the geomorphic mapping), and also of some polygons or portions of polygons Class IV and V classified as stable by the LVQ algorithm. Pavel et al. (2008) discuss in detail this result and analyze the intrinsic differences between the subjective geomorphic method and the LVQ algorithm, and propose techniques for accuracy improvement in future studies.



Figure 3. Terrain stability prediction using Learning Vector Quantization.

# 5.2 Results produced by SINMAP

The mapping produced by SINMAP is presented in Fig. 4 In this figure, six distinct zones are automatically identified corresponding to various stability indices as follows (in decreasing order of stability): stable, moderately stable, quasi-stable, lower threshold, upper

threshold, and defended. In addition to the SINMAP mapping, terrain polygons Class IV and V (unstable terrain) and the location of existing landslides are also illustrated.



Figure 4. Terrain stability prediction using SINMAP. The entire area and detail.

In SINMAP, although reasonable effort was put into tuning the attributes, the system still failed to identify 21 landslides (18 near the valley floor and 3 in the rest of the area). Very often, areas classified as unstable include isolated pixels or clusters of pixels of stable areas. However, in three cases, the stable pixels include existing landslides, as identified in the detail presented in Fig. 4.

Although there is no direct correspondence between the stability zones identified by SINMAP and the subjective geomorphic mapping, it seems reasonable to compare the area of the most unstable three classes (lower threshold, upper threshold, and defended) with terrain Classes IV and V. The similarity to the actual mapping produced by the terrain specialists achieved by the SINMAP mapping is 82%. Similarly to the LVQ-based prediction, the area classified as unstable in the 'stable' zone (as identified by the geomorphic mapping) is relatively small. The majority of the misclassified areas are represented by portions of polygons Class IV and V classified as stable by SINMAP.

# 6 DISCUSSION

The results obtained in this study indicate that the LVQbased mapping yields better results than the physicallybased model. However, it is very likely that the quality of the LVQ-based mapping was influenced by the random selection of DEM cells included in the training and testing data sets. Random selection is a commonly used procedure which ensures that both data sets were representative and the analysis is unbiased. Through this selection/splitting of data, adjacent pixels similar in their attributes were assigned to the training and testing data sets, respectively. Hence, it seems that random selection of cells influenced the good quality of classification by the virtues of spatial autocorrelation (Legendre 1993): i.e., if the pattern represented by one cell was correctly learned in the training phase, then the adjacent cell was correctly classified in the testing phase. This was obviously the case for terrain (and existing landslides) correctly classified on the valley floor, on the southern end of the study site. On the same issue, despite the efforts for calibrating SINMAP, the model was not successful at identifying the respective area as unstable. These results identify a major advantage of the 'pattern recognition' (LVQ) approach versus the physically-based model: the LVQ approach was able to 'learn' the patterns of instability existing on the entire watershed, whereas the physically-based approach is restricted by its principles, and was not successful in predicting instability in complex conditions.

Another difference between the mappings produced by the two methods presented in this study relates to the contiguity of areas identified as stable or unstable. As presented in Fig. 3, the LVQ-based mapping produced relatively large contiguous areas of similar terrain (stable or unstable). This increases the practical applicability of the method, and confirms the principle of selforganization of surficial geomorphic deposits that was also identified in other studies (Werner 1999; Mann 2003; Pavel 2003). In contrast, the SINMAP mapping creates a mosaic of zones of various sizes with different stability indices. This may be problematic for practical applications as field assessments cannot be directed based only on the size of the unstable areas (i.e., cannot neglect small unstable areas).

The last and probably most important difference between the two approaches compared in this study refers to the assumptions used. The LVQ mapping implicitly assumes that a mapping for a similar terrain is available. However, similarity between various areas is a subjective call and this is a weakness of the approach. The assumptions used by SINMAP are presented in Section 4 of this paper and refer to the groundwater flow and distribution of soil parameters over large areas. Experience to-date suggests the assumptions of the LVQbased mapping are more acceptable than those of the physically-based model. Also, given its main principle, the SINMAP analyses have an element of subjectivity (i.e., parameters have to be input by the user), and therefore predictions may vary with the experience of the analyst. Future investigations of these approaches will clarify their applicability for landslide hazard zonation.

#### ACKNOWLEDGEMENTS

Funding for this study was provided by the Science Council of British Columbia, University of British Columbia, Natural Sciences and Engineering Research Council of Canada, and Western Forest Products, Ltd. The authors are very grateful to the Greater Vancouver Regional District for providing the data used in this study.

#### REFERENCES

- Anon. 1999. Greater Vancouver Regional District (GVRD) watershed ecological inventory program. Study coordinated by Acres International Limited. 3 volumes, 500 + pp.
- Carrara, A., Crosta, G., and Frattini, P., 2003. Geomorphological and historical data in assessing landslide hazard. Earth Surf. Process. Landforms 28: 1125-1142.
- Hammond, C., Hall, D., Miller, S., and Swetik, P. 1992. Level I Stability Analysis (LISA). Documentation for Version 2.0. USDA, Intermountain Research Station. General Technical Report INT-285. 128 pp.
- Howes, D.E., and Kenk, E. (contributing eds.) 1997. Terrain classification system for British Columbia (version 2). Ministry of Environment and Surveys, and Ministry of Crown Lands. Victoria, BC. 101 pp. <u>http://ilmbwww.gov.bc.ca/risc/pubs/teecolo/terclass/co</u> <u>ntent.htm</u> (accessed April, 2008)
- Jaakkola, J. 1998. Forest Groundwater Hydrology: Implications for Terrain Stability in Coastal British Columbia. MASc Thesis. UBC.
- Keaton, J.R., and DeGraff, J.V., 1996. Surface observation and geologic mapping. In: Turner, A.K., Schuster, R.L. (Eds.), Landslides - Investigation and Mitigation. National Acad. Press, Transp. Res. Board Special Rep. 247, Washington, DC, pp. 178-230.
- Kohonen, T., 2001. Self-organizing Maps. Third edition. Springer-Verlag, Berlin.
- Legendre, P. 1993. Spatial autocorrelation: Trouble or New Paradigm ? Ecology, 74 (6): 1659 - 1673.
- Mann, D. 2003. On patterned ground. Science 299: 354 355.
- Pack, R.T., Tarboton, D.G., Goodwin, C.N., and Prasad, A. 2005. SINMAP 2: A Stability Index Approach to Terrain Stability Hazard Mapping. User's Manual. 73 pp.

### http://hydrology.neng.usu.edu/sinmap2/

(accessed April, 2008)

- Pavel, M. 2003. Application of Artificial Neural Networks for terrain stability mapping. Ph.D. thesis, University of British Columbia, Vancouver.
- Pavel, M., Fannin, R.J., and Nelson, J.D. 2008. Replication of a terrain stability mapping using an artificial neural network. Geomorphology 97 (3 – 4): 356 – 373.

- Province of BC (British Columbia), 1999. Mapping and assessing terrain stability guidebook. Second Edition. Forest Practices Code of British Columbia. Ministry of Forests and Range, Victoria, BC, 36 pp. <u>http://www.for.gov.bc.ca/TASB/LEGSREGS/FPC/FPC</u> <u>GUIDE/terrain/index.htm</u> (accessed April, 2008)
- Soeters, R., and van Westen, C.J., 1996. Slope instability recognition, analysis, and zonation. In: Turner, A.K, Schuster, R.L. (Eds.), Landslides - Investigation and Mitigation. National Academy Press, Transp. Res. Board Special Rep. 247, Washington DC, pp. 129-177.
- Werner, B.T. 1999. Complexity in natural landform patterns. Science 284: 102 104.
- Wilkinson, J.M.T. 1996. Landslide initiation: a unified geostatistical and probabilistic modeling technique for terrain stability assessment. M.A.Sc. Thesis, UBC.