Analyzing Landslide Susceptibility in St. Vincent and the Grenadines Using Co-Kriging and Logistic Regression

Jacob Gaskill[1], Brian Zuber[1], Erik Nordman Ph. D.[2]

[1]Grand Valley State University, Masters Candidate, Biology
[2]Grand Valley State University, Associate Professor, Department of Natural Resources and Management

Abstract:

Landslides around the world cause human fatalities and significant economic losses for governments and private citizens. Therefore, there is great value in being able to predict future landslides. We use logistic regression to model factors causing landslides on the island of Saint Vincent in the eastern Caribbean. We found that slope explained 18.7 percent of the variation in our model, while aspect was insignificant. We also use co-kriging to spatially show the areas on Saint Vincent where hazard is greatest for future landslides. Our analyses should serve as a preliminary landslide susceptibility assessment for managers on Saint Vincent.

1. Introduction

Landslides cause fatalities, significant economic losses (both to governments and individuals), and damage to infrastructure in many regions around the world. Many studies have attempted to measure "risk" in landslide-prone areas. Risk refers to "the expected losses including persons injured, damage to property and disruption of economic activity due to a particular damaging phenomenon for a given area and reference period" and ultimately is a function of hazard, exposure, and vulnerability (Varnes 1984). Hazard takes into account the spatial factors that cause landslides while exposure is a measure of the temporal probability of a landslide event occurring within a given time period (Varnes 1984). Vulnerability references the social factors that are affected by landslides. Vulnerability is usually measured as loss of human life, economic losses, or a combination of both (Glade and Crozier 2005; Jaiswal et al. 2010). Vulnerability depends on the probability that important social factors such as population density, infrastructure, and buildings (van Westen et al. 2006) will be affected by a landslide. Often data are limited, incomplete, or nonexistent which explains why many of the case studies concerning landslides do not calculate true total risk (van Westen et al. 2006). Instead, many model landslide susceptibility or where landslides are most likely to occur in the future.

1.1 Study Area

Our study was the island of Saint Vincent, the largest island in the country of Saint Vincent and the Grenadines (Figure 1). It is part of the Windward Islands which comprises the southern Antilles south of Guadeloupe and north of Trinidad and Tobago.

Figure 1: Saint Vincent is the site of the landslide susceptibility assessment.

Saint Vincent topography includes drastic slopes and elevation changes culminating in the island's active volcano, La Soufriere. The recent eruptions from this volcano, along with the island's relatively short geologic history, explain the fertile but shallow soil layer (DeGraff et al. 1989). This shallow soil contributes to the frequency and magnitude of landslides on the island.

Table 1: Saint Vincent information (Boruff and Cutter 2007).

Area	317.06 km^2
Highest Elevation	1,234 m
Population	76,762
Location	13.25°N, 61.20°W

The second largest economic industry (behind tourism) on Saint Vincent is agriculture, with most of the agricultural exports consisting of bananas (Boruff and Cutter 2007; DeGraff et al. 1989). This means that forested land is increasing replaced with banana plantations, a process that increases landslide susceptibility. Banana roots are much shallower than the native vegetation and do not stabilize the soil nearly as well (DeGraff et al. 1989). DeGraff et al. (1989) reports an eyewitness account of a Saint Lucian (neighboring island) farmer watch his entire five hectare field get destroyed by a landslide. The farmer believed that a combination of the steep slope and absence of trees with "firm roots" that led to the landslide.

1.2 Significance

Reliance of Saint Vincent and the Grenadines economy on both tourism and a single cash crop makes the island vulnerable to landslides and other natural disasters (Boruff and Cutter 2007). Most of the agriculture on Saint Vincent is small scale meaning the destruction of one field or plantation can place a great financial hardship on the farmer (DeGraff et al. 1989). DeGraff et al. (1989) estimate that one landslide event can destroy 25-50% of a farmer's annual income. This figure does not take into account the indirect costs associated with landslide damages, a figure that Sterlacchini et al. (2007) believe to be as costly as direct damages. Globally, landslides cause more damage to personal property than any other natural disaster (Garcia-Rodriguez et al. 2008), with estimates totaling over one billion (US) dollars (Schuster et al. 1996).

To the best of our knowledge, the number of fatalities due to landslides on Saint Vincent is unknown. However, Boruff and Cutter (2007) estimate that 78% of the population is living in at risk areas for future landslide events. Consequently, preparing for and preventing future damages from landslides is a priority for the government of Saint Vincent and the Grenadines. Mitigation initiatives must come in two forms: legislation that reduces development in risk-prone areas and sets building standards to promote disaster resistant structures, *and* implementing educational programs informing communities about the steps they can take to minimize risk and the effects of landslides (Anderson et al. 2010; Holcombe and Anderson 2010). Managers must know where landslide risk is greatest on the island. This information will help managers chose the best mitigation options, government officials to budget funds to cover damages from landslides, and decision makers during the planning and site selection of building projects (Sterlacchini et al. 2007; Durman et al. 2006).

1.3 Purpose

Here we present a preliminary evaluation of landslide risk on the island of Saint Vincent. While we were not able to create a "true" risk map due to incomplete and non-existent data sets, our analyses provide valuable information for both managers and other analysts to target at-risk areas on the island. We created a logistic regression model identifying significant explanatory variables. We then created a co-kriging output map using the significant variables as explanatory variables to spatially evaluate the amount of risk across the island. Unfortunately, the landslide datasets that currently exist on Saint Vincent are sparse, limiting the amount of explanatory variables we could include in our model. However, it is important to begin the process of evaluating risk on Saint Vincent.

2. Methods

We obtained our historical landslide data from Jerome DeGraff (pers. comm.). DeGraff et al. (1988) produced detailed maps of areas impacted by landslides (as well as others from Caribbean Islands) in a report for the Organization of American States. In this report, landslides were identified as "definite" and "possible." The maps were georeferenced and the landslide areas were digitized as points. When we digitized the points, we included all (both definite and possible) as landslide events. In this way, we were able to compile a historical landslide database of 463 points for Saint Vincent. Absence points were created by erasing a 200 meter buffer around known landslide points. A total of 400 random points were created with a minimum distance of 100 meters between the random points.

We obtained our DEM data from the ASTER sensor through the USGS Earth Explorer database. This sensor is on board NASA's Terra satellite and captured the images we used of Saint Vincent in 2011. Slope and aspect were derived from the DEM data using their respective spatial analyst tools (Moore et al. 1991). Slope and aspect values were then extracted from the rasters and appended to the point data. Of the 863 landslide and non-landslide points, 663 random points were selected as training points leaving 200 points for validation.

Finally, we isolated developed areas for Saint Vincent using the land cover dataset offered by the GeoSUR database (http://www.geosur.info/geosur/index.php/en/). The image was georeferenced and the light and heavy urban areas were separated from the other land use data and combined.

We used the binary logistic regression tool in SPSS (IBM, Version 20.0.0) to model landslide susceptibility. Logistic regression is commonly used for this purpose because it handles categorical variables (like landslide occurrence) better than other modeling techniques. Logistic regression evaluates which independent variables are significant, and the output coefficient can be used to determine the importance of each variable in explaining variation with relation to other variables (Kincal et al. 2009). In our model, only slope was significant ($p<0.05$). Significant variables (slope) were then used for our co-kriging analysis. First, slope had to be tested for autocorrelation (a requirement of co-kriging). Spatial autocorrelation was validated using Moran's I (Moran's Index: 0.51, p=0.00).

The 663 training points were used to create the simple co-kriging model. Landslide presence / absence, the first variable, was not transformed. The second variable, slope, was declustered prior to being transformed using normal score.

The co-kriging model values were extracted and appended to the validation points. Two validation points found outside of the model area were removed leaving 198 validation points. Predicted values were then calculated into risk / no risk with predicted values <0.5 being no risk and values >= 0.5 being risk. These values were then compared to the actual presence (risk) / absence (no risk) landslide data.

3. Results

Analysis of the validation points revealed that the model had a total accuracy of 82% (Table 2). From the actual data, 81 points were no risk with 117 points being at risk (actual landslides). The model predicted that 93 points would have no risk and 105 points being at risk.

		Actual		
		No Risk	Risk	**% Correct Validation Points**
Predicted	No Risk	69	24	**74%**
	Risk	12	93	**89%**

Table 2: Accuracy of co-kriging model of landslides on Saint Vincent using 198 validation points.

Our model predicts that most of the urban areas have relatively low risk of landslides (Figure 1). The areas all along the Eastern and Southern coasts have very low risk where the northern and western coasts are more at risk. The interior portion of the island has scattered areas of high to low risk.

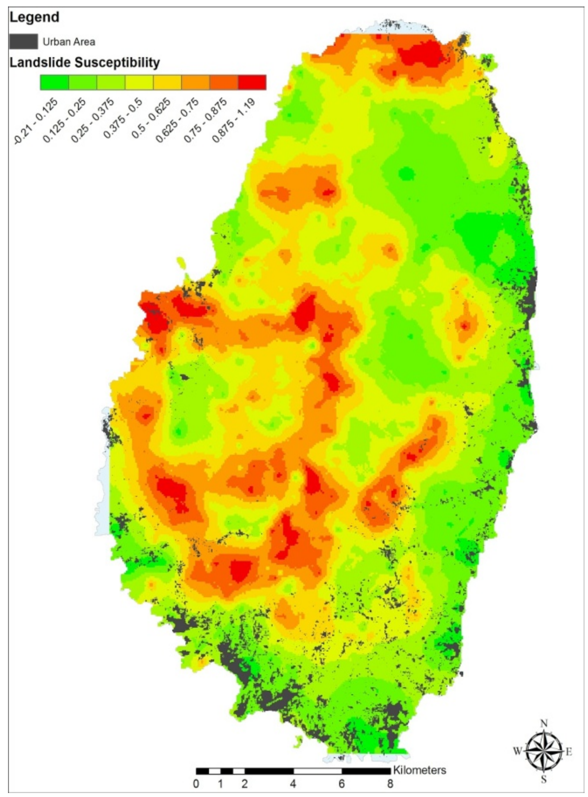

Figure 1: Saint Vincent co-kriging analysis using landslide data and slope. Red represents the highest likelihood of a landslide, where green represents the lowest likelihood. The dark grey areas symbolize urban development.

With all of the actual landslide points overlaid on the co-kriging output we can visually see the model's accuracy (Figure 2). Very few actual landslide events fall on areas with less than 0.5 risk.

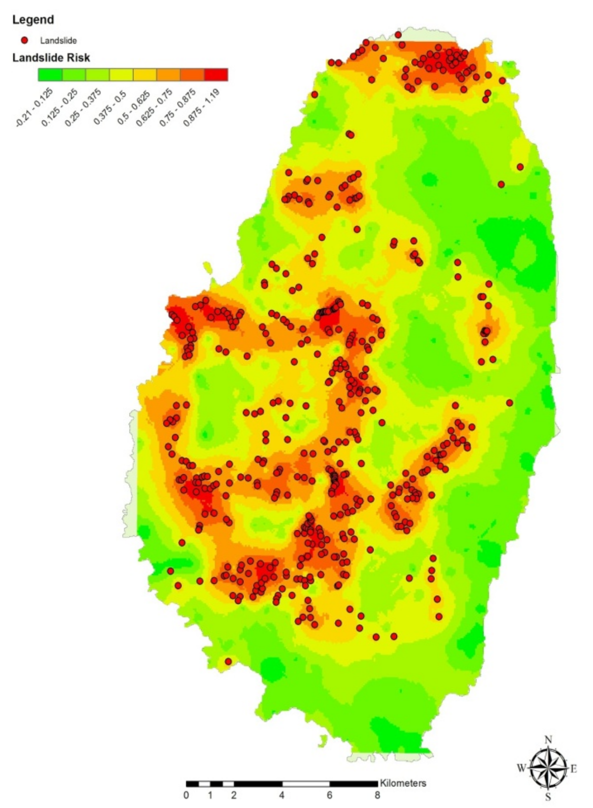

Figure 2: Saint Vincent co-kriging analysis using landslide data and slope. Red represents the highest likelihood of a landslide, where green represents the lowest likelihood. The red points symbolize 463 known landslide points.

The co-variance graph shows the degree of similarity between surface points and is used to generate the co-kriging map (Figure 3). The model suggests that similarity in landslide risk drops off sharply within 800 meters of a landslide point and stabilizes at a distance of about 5,000 meters (Table 3, Figure 3).

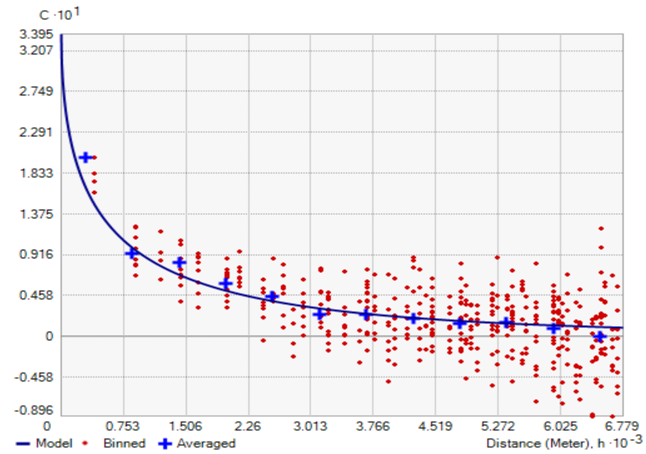

Figure 3: Landslide x Slope covariance plot calculated by the co-kriging analysis on Saint Vincent.

Table 3: Covariance models generated during the co-kriging analysis. LS = Landslide; Sl = Slope.

	Constant	**Nugget**	**Range**	**Sill**
LS x LS	0.074458	0.17421	4870.4	0.51465
LS x Sl	0	0.33945	4870.4	0.51464
Sl x Sl	0	1.11240	4870.4	0.51465

4. Discussion

The spatial explanatory variables we used for our logistic regression analysis were slope and aspect. Of these, only slope was significant ($p<0.05$) but was able to explain 18.7 percent of the variation in the model. While this is not a high *total* amount of variation explained for most models, it is relatively high considering our model only has one independent variable. A greater number of variables would explain a greater amount of variation, but unless each explains a large amount of the variation, the model is not as effective (Ayalew and Yamagishi 2004). Our findings reiterate those of similar studies conducted in other areas of the world: slope is very important in being able to measure landslide susceptibility.

When logistic regression was used for this purpose in Turkey, slope gradient explained the greatest amount of variation (Kincal et al. 2009; Duman et al. 2006; Akgun et al. 2012; Akgun 2012). The same was true for studies in India (Jaiswal et al. 2010; Jaiswal et al. 2011) and Japan (Ayalew and Yamagishi 2005).

Unlike slope, our findings concerning aspect are not in concordance with the work of others. Initially, aspect was included as an independent variable in our logistic regression analysis but was removed when it was found to be insignificant ($p>0.05$). This contradicts the findings of Duman et al. (2006) and Ayalew and Yamagishi (2004) who both found aspect to explain large amounts of variation in each of their respective logistic regression models. We believe that the discrepancy between our study and others highlights the localized nature of landslides. In fact, there is a lack of agreement in the literature about the importance of aspect in explaining landslides, and many conflicting viewpoints exist (Carrara et al. 1991).

4.1 Limitations

GIS based statistical models operate under the assumption that landslides will most likely occur in the same location and because of the same reasons that they have in the past. However, this might not always be the case as each landslide event changes the surrounding terrain, sometimes drastically, and thus changes conditions from how they were prior to the landside happening (van Westen et al. 2006). While it is common for landslides to occur close to each other (resulting in "hot spots"), it is not very likely for two landslides to occur in the exact same place, a characteristic that is unlike some other natural disasters and contributes to the difficulty in modeling landslides (van Westen et al. 2006). Furthermore, there are many types of landslides (DeGraff et al. 1989), and the factors that lead to each often differ (van Westen et al. 2006). Despite these limitations, logistic regression remains one of the best and most common modeling techniques for evaluating landslide susceptibility. Akgun (2012) demonstrated that landslide susceptibility maps generated from significant logistic regression independent variables were more accurate than other prediction methods evaluated. Thus we are confident in the figures generated from our analyses.

4.2 Value as a Preliminary Investigation

Despite the limitations of our datasets and inherent limitations associated with GIS based statistical modeling, we believe our analyses provide great value. To the best of our knowledge, no other studies attempt to model landslide susceptibility on Saint Vincent. Thus our results provide a preliminary investigation for future studies to build upon. Furthermore, our co-kriging map (Figure 1) can be compared with dense population centers to evaluate which areas are most at risk for landslide exposure, as well as identify areas where land should not be developed. Our study is unique in that the total number of points comprising our historical landslide dataset were very high (n=463). Many of the other studies we examined used fewer than 87 historical landslide points in their analyses (Ayalew and Yamagishi 2010; Akgun et al. 2012). Our large sample size strengthens our model by reducing the effects of outliers.

4.3 Applications for Saint Vincent

A model is only as good as the reliability, accuracy, and thoroughness of its data inputs (Kincal et al. 2009). Currently, the available data relative to landslides and their causes on Saint Vincent is limited. Spatially, only slope and aspect data are available and the historical landslide dataset used in our analysis only include landslides that happened prior to 1988 (Jerome DeGraff, pers. comm.). While Jaiswal et al. (2010) used historical landslide data dating back 23 years for their analysis, such long time delays can be problematic. Landscapes change over time. Vegetation, land use, and to a limited extent topography, might have changed between the time our independent variable datasets were collected (2011) and the time the landslide occurred. This again highlights the need for governments to keep long-term landslide databases.

For this reason, we advocate the creation of a long term landslide database for Saint Vincent and the Grenadines. The type of data to include in this database should be based on the findings of other landslide case studies from around the world. Specifically, Kincal et al. (2009) and Duman et al. (2006) found that weak underlying bedrock explained a large amount of variation in their logistic regression models, so lithology data should be collected for Saint Vincent. However, the bedrock in these case studies is sedimentary. The bedrock of Saint Vincent is volcanic and thus it might not be as influential in causing landslides. But considering lithology only needs to be sampled once (bedrock remains stable over short timescales) and could potentially provide important information, its collection should become a priority. Another metric to collect should be the proximately of landslide events to roads. Ayalew and Yamagishi (2005) found that proximity to roads was almost as important a predictor of landslide occurrence as slope gradient in their Kakuda-Yahiko mountains (Japan) study.

Most importantly, future landslide databases on Saint Vincent should include information on a multitude of potential factors as the explanatory variables in one region of the globe might not be as influential as in another (Ercanoglu and Gokceoglu 2004).

Finally, the importance of triggering events on explaining landslide occurrences has been well documented (van Westen et al. 2006). Unfortunately, GIS based statistical models are constructed on spatial variation (i.e. elevation) and do not account for temporal variation like triggering events (van Westen et al. 2006). For a study to truly be a risk assessment analysis and most accurately predict landslides, it must include temporal variation (to meet the "exposure" part of the risk equation) or, in other words, the triggering event (van Westen et al. 2006). This contradiction can be circumvented by justifying the exclusion of triggering events (Ayalew and Yamagishi 2005), or by using other statistical procedures to search for correlations between potential triggering events and landslides (Dai et al. 2002). For either of these techniques to work, data about potential triggering events must exist, something that is not the case for Saint Vincent. Landslide triggering factors are site-specific and need to be collected before the landslide occurs. Even when historical records for landslides exist (this is not always the case), a lack of information about the triggering factor(s) behind each landslide event makes predicting future landslides difficult (van Westen et al. 2006).

When the triggering factor is measured, precipitation seems to be the triggering event most frequently behind landslide occurrences (Jaiswal et al. 2011; Ayalew and Yamagishi 2005; Jaiswal et al. 2010), though seismic activity has also been demonstrated as a causal factor as well (Akgun et al. 2012, Dai et al. 2002). Based on past newspaper articles and the large amount of rain that falls on Caribbean countries (DeGraff et al. 1989), precipitation is most likely the case for Saint Vincent as well, though the amount of precipitation that leads to landslides here is unknown (DeGraff et al. 1989). In the future, Saint Vincent authorities should measure precipitation spatially across the island to aid in the predictive value of future models.

The high frequency of landslide events and lack of a reliable landslide database on the island of Saint Vincent make it difficult to model landslide risk on the island. However, the ability to do this has a great potential to decrease losses, both from an economic and human life perspective. We found that slope explains a large part in landslide variation while aspect is insignificant. However, our model could have been strengthened with more data. For this reason, we call for increased consistent and reliable data collection on Saint Vincent and hope that our preliminary analysis provides managers with further information with which to make more informed decisions.

5. Acknowledgements

We would like to thank the DigitalGlobe Foundation for their generosity in sharing with us high resolution IKONOS images for preliminary analyses. We would also like to thank Jerome DeGraff for sharing with us his previous work on Saint Vincent and answering our questions. Our historical landslide dataset came directly from Mr. DeGraff's 1988 survey. His help and past work were an integral part of this project, and for that we extend our sincere gratitude.

6. References

Anderson, M. G., Holcombe, E., Esquivel, M., Toro, J., & Ghesquiere, F. (2010). The efficacy of a programme of landslide risk reduction in areas of unplanned housing in the eastern Caribbean. *Environmental Management*, 45, 807-821.

Akgun, A. (2012). A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at Izmir, Turkey. *Landslides*, 9, 93-106.

Akgun, A., Kincal, C., & Pradhan, B. (2012). Application of remote sensing data and GIS for landslide risk assessment as an environmental threat to Izmir city (west Turkey). *Environmental Monitoring and Assessment*, 184, 5453-5470.

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

Boruff, B. J., & Cutter, S. L. (2007). The environmental vulnerability of Caribbean island nations. *Geographical Review*, 97, 24-45.

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., & Reichenbach, P. (1991). GIS techniques and statistical models in evaluating landslide hazard. *Earth Surfaces Processes and Landforms*, 16, 427-445.

Carrara, A., Cardinali, M., Guzzetti, F., & Reichenbach, P. (1995). GIS technology in mapping landslide hazard. In: Carara, A. & Guzzetti, F. (eds), Geographical Information Systems in Assessing Natural Hazards. Kluwer Academic Publishing, Dordrecht, The Netherlands, 136-176.

Dai, F. C., Lee, C. F., & Ngai, Y. Y. (2002). Landslide risk assessment and management: an overview. *Engineering Geology*, 64, 65-87.

DeGraff, J. V., Bryce, R., Jibson, R. W., Mora, S., & Rogers, C.T. (1989). Landslides: Their extent and significance in the Caribbean. West Indies-Final Report. Washington DC, Organization of American States.

Duman, T. Y., Can, T., Gokceoglu, C., Nefeslioglu, H. A., & Sonmez, H. (2006). Applicatin of logistic regression for landslide susceptibility zoning of Cekmece Area, Istanbul, Turkey. *Environmental Geology*, 51, 241-256.

Ercanoglu, M. & Gokceoglu, C. (2004). Use of fuzzy relations to produce landslide susceptibility map of a landslide prone area (West Black Sea Region, Turkey). *Engineering Geology*, 75, 229-250.

Garcia-Rodriguez, M. J., Malpica, J. A., Diaz, M. (2008). Susceptibility assessment of earthquake triggered landslides in El Salvador using logistic regression. *Geomorphology*, 95, 172-191.

Glade, T., & Crozier, M. J. (2005). A review of scale dependency in landslide hazard and riskVar analysis. In T. Glade, M. G. Anderson, & M. J. Crozier (Eds.), Landslide risk assessment (pp. 75-138). Chichester: Wiley.

Holcombe, E. A., & Anderson, M. G. (2010). Implementation of community-based landslide hazard mitigation measures: the role of stakeholder engagement in 'sustainable' project scale-up. *Sustainable Development*, 18, 331-349.

Jaiswal, P., van Westen, C. J., & Jetten, V. (2010). Quantitative landslide hazard assessment along a transportation corridor in southern India. *Engineering Geology*, 116, 236-250.

Jaiswal, P., van Westen, C. J., & Jetten, V. (2011). Quantitative estimation of landslide risk from rapid debris slides on natural slopes in the Nilgiri hills, India. *Natural Hazards and Earth System Sciences*, 11, 1723-1743.

Kincal, C., Akgun, A., Koca, M. Y. (2009). Landslide susceptibility assessment in the Izmir (West Anatolia, Turkey) city center and its near vicinity by the logistic regression method. *Environmental Earth Sciences*, 59, 745-756.

Moore, I.D., Grayson, R.B., & Ladson, A.R. (1991). Digital terrain modeling: a review of hydrological and biological applications. *Hydrological Processes*, 5, 3-30.

Schuster, R. L. (1996). Socioeconomic significance of landslides. In: Turner, A. K., Schuster, R. L. (eds), Landslides: investigation and mitigation. Transportation Research Board National Research Council, Special Report no. 247, Washington, DC.

Sterlacchini, S., Frigerio, S., Giacomelli, P., & Brambilla, M. (2007). Landslide risk analysis: a multi-disciplinary methodological approach. *Natural Hazards and Earth System Sciences*, 7, 657-675.

van Westen, C. J., van Asch, T. W. J., & Soeters, R. (2006). Landslide hazard and risk zonation – why is it still so difficult? *Bulletin of Engineering Geology and the Environment*, 65, 167-184.

Varnes, D. J. (1984). Landslide hazard zonation: a review of principles and practice (p. 63). Paris: UNESCO.