Elsevier Editorial System(tm) for Geomorphology
Manuscript Draft

Manuscript Number: GEOMOR-925

Title: Combining computational models for landslide hazard assessment of Guantánamo province, Cuba

Article Type: Research Paper

Section/Category:

Keywords: landslide; hazard assessment; Guantánamo; Cuba; ANN; Weight of Evidence

Corresponding Author: Sr. Enrique Armando Castellanos Abella, M.Sc.

Corresponding Author's Institution: Institute of Geology and Paleontology

First Author: Enrique Armando Castellanos Abella, M.Sc.

Order of Authors: Enrique Armando Castellanos Abella, M.Sc.; Cees J Westen, PhD; Caterina Melchiorre, PhD

Manuscript Region of Origin:

Abstract: As part of the Cuban system for landslide disaster management, a methodology was developed for regional scale landslide hazard assessment, which is a combination of different models. The method was applied in Guantánamo province at 1:100 000 scale. The analysis started with an extensive aerial photointerpretation to produce a landslide inventory map. Five main types of landslide movements were identified: slides (186), rockfalls (22), debrisflows (26), topples (18), and large rockslides (29). The causal factors for each landslide type were analysed using twelve indicators: lithology, geomorphology, landuse, soil, slope angle, internal relief, slope orientation, drainage density, distance to roads and faults, rainfall intensity, and peak ground acceleration. A specific set of causal factors was identified for each of the five landslide types. Artificial Neural Networks (ANNs) were applied for the assessment of susceptibility to slides and the weights of evidence for the assessment of susceptibility to the other landslide types. Success rates were used for evaluating the performance of the model. In all cases more than 80% of landslides were classified within the 10% of most susceptible areas. The five susceptibility maps were reclassified and

spatial and temporal probabilities were calculated for each class, based on the landslide densities and on geomorphological evidences. Around 13 % of the province was classified as high susceptible to at least one landslide type, 17 % as moderate and 24 % as low. The results show that it is possible to derive at semi-quantitative landslide hazard maps, for different landslide types, making use of a combination of heuristic reasoning and probabilistic derived weights.

Enschede, The Netherlands
Tuesday, January 29, 2008

Dear Dr. Takashi Ogushi,

Maybe you still remember part of the presentation of our work during you visit at ITC. We are close to finish the research and we are submitting part of the output for publication. Please, find attached the manuscript of the paper we would like to publish in Geomorphology journal titled: "Combining computational models for landslide hazard assessment of Guantánamo province, Cuba".

We will appreciate any comment from you and the reviewers.

Sincerely

Enrique A. Castellanos Abella

E. A. Castellanos Abella
PhD Candidate
International Institute for Geo-Information Science and Earth Observation (ITC)
P.O. Box 6, 7500 AA, Enschede
The Netherlands

Contact:
0031 53 48 74 504 (ITC office)
0031 53 48 74 444 (ITC reception)
0618823693 (mobile)
e-mail: castellanos@itc.nl
http://eacastellanos.googlepages.com/

Combining computational models for landslide hazard assessment of Guantánamo province, Cuba

E. A. Castellanos Abella [a], C. J. Van Westen [b], C. Melchiorre[c]

[a] Corresponding Author: Instituto de Geología y Paleontología (IGP), Vía blanca y Carretera Central, San Miguel del Padrón, CP 11000, Ciudad de La Habana, Cuba. Phone: (537) 696 7232, Fax: (53 7) 696 7004, E-mail: castellanos@itc.nl
[b] International Institute for Geoinformation Science and Earth Observation (ITC) ITC, PO, Box 6, 7500 AA Enschede, The Netherlands E-mail: westen@itc.nl
[c] Geological Survey of Norway (NGU), 7491 Trondheim, Norway. E-mail: caterina.melchiorre@ngu.no

Abstract
As part of the Cuban system for landslide disaster management, a methodology was developed for regional scale landslide hazard assessment, which is a combination of different models. The method was applied in Guantánamo province at 1:100 000 scale. The analysis started with an extensive aerial photointerpretation to produce a landslide inventory map. Five main types of landslide movements were identified: slides (186), rockfalls (22), debrisflows (26), topples (18), and large rockslides (29). The causal factors for each landslide type were analysed using twelve indicators: lithology, geomorphology, landuse, soil, slope angle, internal relief, slope orientation, drainage density, distance to roads and faults, rainfall intensity, and peak ground acceleration. A specific set of causal factors was identified for each of the five landslide types. Artificial Neural Networks (ANNs) were applied for the assessment of susceptibility to slides and the weights of evidence for the assessment of susceptibility to the other landslide types. Success rates were used for evaluating the performance of the model. In all cases more than 80% of landslides were classified within the 10% of most susceptible areas. The five susceptibility maps were reclassified and spatial and temporal probabilities were calculated for each class, based on the landslide densities and on geomorphological evidences. Around 13 % of the province was classified as high susceptible to at least one landslide type, 17 % as moderate and 24 % as low. The results show that it is possible to derive at semi-quantitative landslide hazard maps, for different landslide types, making use of a combination of heuristic reasoning and probabilistic derived weights.

Keywords: landslide; hazard assessment; Guantánamo; Cuba; ANN; Weight of Evidence

1. Introduction

Landslide susceptibility mapping have been carry out for many years worldwide with achievements either in methods applied as well as in knowledge of landslide occurrence (Carrara, A. et al., 1991). The susceptibility maps, however, should be made for different landslide types occurring and each failure mechanism should have its own combination of factor maps and classes to be included. Moreover, depending on the characteristics of each landslide type different methods could be applied and combined. The proposed method follows a spatial data integration approach, based on the analysis of a landslide inventory and a set of causal factors. This approach is generally considered as the most useful at medium to small scales (Aleotti and Chowdury, 1999). A central component in the proposed method is the evaluation of different landslide types, and failure mechanisms, which may be controlled by different combinations of causal factors. Therefore a combination of models is evaluated, and for each failure mechanism the optimal combination of causal factors and data integration technique is assessed. There are

still relatively few publications that report results of a separate analysis of different failure mechanisms at regional scale landslide susceptibility and hazard assessment (Castellanos, E. et al., 1998). Even fewer find the optimal model for each failure mechanism (Guzzetti, F., et al., 1999, Lutfi, M. and Doyuran, V., 2004, Yesilnacar, E. and Topal, T., 2005). For carry out this research the province of Guantánamo in Cuba was used.

Landslides in Cuba are mostly associated with hurricanes, tropical storms, prolonged periods of rainfall, and occasionally due to earthquakes (Pérez, 1983; Iturralde-Vinent, 1991; Pacheco and Concepción, 1998; Castellanos, 2000). The damage due to landslide is mostly not separately recorded from the main triggering event, so that overall landslide damage information is lacking in Cuba. Unfortunately, there is no comprehensive landslide inventory for Cuba available. In a report presented by the National Civil Defence Headquarters it was estimated that 45,000 inhabitants are vulnerable to landslides (EMNDC, 2002). Cuba is considered a model in hurricane risk management by the United Nations (Sims and Vogelmann, 2002; ISDR, 2004) because hurricanes in Cuba cause considerable less casualties as compared to neighboring countries with a different economical, social and political context. However, landslides were not appropriately considered in the Cuban national system for disaster prevention. Against this background the Cuban National Defence Organization initiated a programme for the development of a national system for multi-hazard risk assessment. This requires a component of landslide hazard and risk assessment, which includes the design of a national landslide inventory system, and proposed methodologies for landslide hazard and risk assessment at different scales (national, provincial, municipal and local) corresponding to the administrative subdivisions of the Civil Defence Organization. The method proposed at the national scale makes use of Spatial Multi-Criteria Evaluation, in combining a series of hazard indicators with vulnerability indicators for the entire country, and defining a risk index per municipality or province (Castellanos and Van Westen, 2007). At the municipal scale the proposed method uses a combination of geomorphological mapping, expert weighting, and physical modeling (Castellanos and Van Westen, 2008). In this paper the hazard assessment method for the provincial scale is presented.

A province is the middle level in the political and administrative framework of Cuba, and its disaster risk management responsibilities are a combination of coordinating and executing tasks. The main objectives for landslide hazard assessment at the provincial level are to: i) identify areas where landslide events have happened and characterize them as detailed as possible; ii) locate, in the provincial territory, the areas with high susceptibility and identify the main causes in terms of the indicators; iii) use the susceptibility for a hazard assessment and as an input for landslide risk management. Considering these objectives the purpose of the study at provincial scale was to design a method for landslide hazard assessment that takes into account the requirements and available data at this level, and to test it out for Guantánamo province.

2. Case study area: Guantánamo province

Guantánamo (which in the pre-colonial language means "the land of the rivers") is the most eastern province of Cuba with a surface area of 6186 km^2, comprising 5.5% of the national territory. The population is 511 224 (ONE, 2007), which is 4.6% of the national population (Figure 1). About 75% of the area is mountainous with the highest point of 1 181 m., located in Maisí municipality in the East. Most of the northeast part is mountainous, while the southwest is covered by a large valley, which also forms a separate hydrographic basin draining into Guantánamo bay. The northeast basin is drained by the Toa River, which has the highest discharge in Cuba. In terms of climate Guantánamo contains both the most humid (in the

North) and dry (in the South) zones of the Country. Other natural characteristics are discussed in section 3.1 and 3.2.

Fig 1. Location map

The province has 10 municipalities (Figure 1). Agriculture is the most important economic income for the province and is based on sugar cane, coffee, cacao, wood, and coconut. The last four are cultivated in mountainous regions. The industries include an iron foundry, and factories for coffee, agricultural tools, furniture, food, sugar cane, and salt. Guantánamo has a record of 49 devastating hurricanes measured over the period 1789-2003, which are more frequent in September and October (ONE, 2004). Natural and man-made forest fires are also a major concern. Since 1997 to 2002 there were 93 fires reported, affecting an area of 3043 hectares. Other disasters, like landslides are rarely recorded in the official statistics. For that reason an extensive landslide investigation was required for this study as explained below.

3. Materials and methods

3.1. Landslides in Guantánamo province

The provincial landslide hazard assessment started with a detailed landslide inventory. Except for the disastrous landslides in Jagüeyes de Caujerí, which happened during hurricane Flora in 1963, there almost no information available on historical landslide location and damage before the start of the study. In 1991 a book was published containing one chapter on an inventory of landslides in the sea-facing slopes of the coastal hills in the South by the Institute of Tropical Geography of Cuba (Magaz et al., 1991), in which these landslides in uplifted marine coral reef terraces were considered to be related to Pliocene-Quaternary earthquakes.

The current research is the follow-up of an earlier work carried out in 1998 at 1:100,000 scale in the eastern part of Cuba for the mapping of the principal hazardous geological processes (Castellanos et al., 1998). In the framework of the project the main types of landslides were identified. An inventory and a hazard map were produced. The hazard map was made using a multivariate statistical analysis also indicated as a "data driven method" (Guzzetti, F et al. 1999) . Lack of experience in landslide photo-interpretation and problems associated to having a complete causal factors dataset, made a review of the results necessary.

For the present study landslides were photo-interpreted in 300 aerial photos (format 23 x 23cm) from the year 2000 at 1:25 000 scale covering the entire Guantánamo province. The photo-interpretation was transferred from the photos to base maps and then digitized. The landslide boundaries were crosschecked using two Landsat ETM+ satellite images from 2001-03-08 (path 010, row 046) and 2001-02-27 (path 011, row 046). The information from band 8 (with 15 m spatial resolution) was used because it provides the highest spatial resolution and contrast of this sensor. The crosscheck allowed differentiating those landslides that were showing signs of activity, mainly presented by unvegetated surfaces, from the ones that were completely vegetated. Unfortunately no multi-temporal image interpretation could be carried out, which made it difficult to establish the age of the landslides. A spatial database was produced including information on size and type of the landslides. The database also contains 12 landslides that were reported by the local civil defence authorities during the fieldwork campaign. Other characteristics of the landslides (e.g., activity) were not described, as they were not possible to be identified at this level of analysis during the photo-interpretation. A map was produced including the landslides, the two orthorectified Landsat TM images, and several cartographic elements such

as roads, topographic names, etc. This map was presented to the provincial authorities as the first step in reducing landslide disaster risk.

Due to the large size of the province, the small scale of the study, and the relatively small size of most of the landslide events, the actual landslide area could not always be digitized to scale. Especially in case of rockfalls, topples and debrisflows events, the source and depositional areas could not always be separately displayed. Some landslide events were very small and difficult to map while others were very large and could even be subdivided into smaller units. The representation of landslides, which have been active over various time periods, also posed a problem at this scale. This had some implications during the analysis, which will be discussed later.

Fig. 2 landslide photos

Figure 2 shows some examples of landslides in Guantánamo province. In total 281 landslides were identified covering 7971 pixels (about 19.92 km^2). From this inventory, four main types of landslides were determined including 22 rockfalls (517 pixels), 26 debrisflows (501 pixels), 18 topples (510 pixels) and 215 slides (6443 pixels). Further analysis showed that slide-type movements were basically of two genetically and morphometrically different types: a group of 29 larger rockslides (3356 pixels) located in a tectonically affected area in the Sierra de Caujerí, and a group of 186 smaller slides (3087 pixels) dispersed all over the province. Although topples are defined as a landslide type (IGME, 1987; UNESCO-WP/WLI, 1993) and some research have been made separately on this type of movement (Nichol et al., 2002; Parise, 2002; Tamrakar et al., 2002; Andriani et al., 2005; Duman et al., 2005), it may be also considered as an initial step toward a rock fall. In this study toppling events were considered separately. They are recognized by a series of detached blocks on a slope, separated by tension cracks or small scarps without necessarily fell down. A toppling movement may culminate in an abrupt falling or sliding but the type of the movement is tilting without collapse (Varnes, 1978; Dikau et al., 1996). Statistical analysis revealed that the size of all landslides varies from 7982 m^2 up to 941,142 m^2 with an average of 70989 m^2. The frequency distribution of landslide area by landslide type (see Figure 3) shows a distribution similar to other landslide inventories (Malamud et al., 2004). Due to the small number of events per type (except for slides) the dataset was not sub-divided in events with different magnitude.

Fig. 3 size frequency distribution

3.2. Collection of potential causal factor maps

A total of 12 potential causal factors were selected, after evaluating the literature (Carrara et al., 1991; Soeters and van Westen, 1996; Guzzetti et al., 1999) and the data available in Cuba. They were divided into fours main groups: ground condition factors, distance related factors, geomorphometric factors and triggering factors. A number of these factors are presented in Figure 4 and their location in the methodology is shown in Figure 5.

Ground condition factors

Ground condition factors are those that represent characteristics of the ground condition in which landslides might take place, related to the geomorphology, geology, soils and landuse.

- The geomorphological map at 1:100 000 scale was created from the national geomorphological map at 1:1,000,000 of the National Atlas of Cuba (Portela et al., 1989). The digital boundaries were superimposed on stereo-images generated from the digital elevation model and different band combinations from satellites images. Interactively, these

boundaries were adjusted and in some cases more detailed units were made, based on the same legend. The geomorphology of Guantánamo province is very diverse. The map contains 97 areas within 34 legend units of mountains, hills and plains of different types and genesis. The mountains are the most extensive covering 64% of the surface. According the classification of Cuban relief the mountains have been divided in low mountains (1,000-1,500 m), small mountains (500-1,000 m), pre-mountains, and sub-mountains (300-500 m) (Portela et al., 1989; Díaz et al., 1990).

- The geological factor map was obtained from the Institute of Geology and Palaeontology (IGP). The map was supplied in CARIS GIS format (CARIS GIS, 2007) with legend and styles according to the Geological Information System of Cuba called SIGEOL (Castellanos et al., 2003). In order to mosaic the 10 maps covering the province, editing and polygonization were carried out in ILWIS. Then the whole map was exported to ArcGIS. Guantánamo province has 44 geological units distributed in 621 polygons. A total of 45.5 % of the province is covered by only three units: San Luis Formation with 1051.6 km^2 distributed mainly in the western part, Sierra del Purial Formation with 907 km^2 mostly located in the east, and Maquey Formation with 860.1 km^2 mainly outcropping in the central part of province. Each of the other units covers less than 6% percent of the territory.
- The landuse map was obtained from the Institute of Physical Planning (IPF). It was digitized, edited and polygonized in ArcGIS. Guantánamo province has 16 landuse types distributed in 633 polygons. A total of 55 % of the territory is covered by natural forest and 14 % by natural pasture. The rest of the territory is covered by sugar cane, coffee and cultivated pasture and grassland. The landuse does not show very dynamic changes, but there has been a tendency to gradually increase the forested area. Besides, the sugar cane lands have been largely decreased after many sugar factories were closed in the middle of the 90s.
- The soil map used was obtained from the Soil Institute (Mesa et al., 1992). It was digitized, polygonized, and linked to a database. Each unit is classified by a combination of a Group, a Sub-Group, and parent material ("genus"). Guantánamo province has 45 unique combinations of soil map units distributed in 101 polygons. The predominant soil map unit (21 % of the area) is the Tropical Grayish Calcareous formed from limestone, marls, and carbonated detrital material. No other soil map units cover more than 10% of the area; those with most area are the "Typical Grayish Tropical Saline", the "Grayish-Red Calcareous" and the "Typical Saline". Most of the soils are less than 30cm deep, which is relevant for landslide occurrence.

Distance related factors

Another group of potential causal factors are those that relate to the distance of a feature that might have influence on landslide occurrence. Two of these were taken into account: roads and faults.

- As road construction in hilly or mountainous terrain is often a causal factor for landslide occurrence, a buffer map was made with zones adjacent to roads. The road map was obtained from the digital topographic map at 1:100,000 scale produced by the national cartographic agency Geocuba. The edited vectors include: highways, first and second order roads, streets in populated zones, unpaved and enhanced-unpaved roads, trails and tracks, wide and narrow railways. (Figure 4H).
- Tectonic factors play relevant role in producing landslide in some cases. In order to explore the spatial relation with landslide the fault map of the province was used. The faults were obtained from the Geological Information System of Cuba (SIGEOL) recently designed and implemented (Castellanos et al., 2003). Different faults types were used including: inferred faults, certain faults, supposed faults, thrust faults and reversed faults. A distance map was created (Figure 4G).

Fig. 4. hazard indicators

Geomorphometrical factors

The digital elevation model (DEM) of the province was created in ArcGIS using the tool "topo to raster", which considers different data sources in the interpolation process such as: elevation points (usually geodetic points), contours lines, drainage network, sinks, lakes or water surfaces and the limit of the area. During the process sinks were identified and those erroneously created were removed interactively until the process was successfully completed. From the DEM, with a spatial resolution of 50 meters, 4 geomorphometric parameter maps were extracted: slope steepness, slope orientation (aspect), internal relief (vertical dissection), and drainage density.

- The slope angle map was divided in 10 classes by quantiles. Due to the large pixel size and the slope calculation algorithm used, the maximum slope angle was 48 degrees (Figure 4D), which occurs in the side slopes of the marine terraces in the coastal zone of Maisí, the main river valleys and the ridges of some mountain system such as the Sierra de Caujerí.
- The slope orientation (Aspect) map contains 8 classes representing the main compass directions (north, northeast, east, etc), and one class for flat areas. The slope direction classes have more or less the same frequency with a slight tendency for south facing slopes (Figure 4A).
- Internal relief or vertical dissection considers the height difference per unit area. This map (Figure 4B) was created in ArcGIS obtaining first the minimum and maximum elevation per hectare using a moving filter. Then the height difference was calculated for each pixel. The maximum internal relief in the province was 636.6 m/ha and the average was 152.2 m/ha (105 SD). The areas with higher values are located at the edge of the terraces and in the mountain ridges. The Sierra del Purial was the physiographic region with highest internal relief values.
- The drainage density map was obtained from the detailed drainage vector map that was made by EMPIFAR Company, including all natural and artificial drainage networks. The length per unit area of drainage lines that fall into a specific search radius of 1 square kilometre for each pixel was calculated. The drainage density map (Figure 4C) shows concentrations of high values in specific zones, especially in the eastern part of Sierra del Purial and other mountain zones. The average value was 2.94 km/km^2 (0.97 SD) and the maximum was 6.7 km/km^2.

Triggering factors

Apart from the above described environmental factors, the spatial variation of two triggering factors was used in the landslide hazard assessment: rainfall and earthquakes.

- A raster map of maximum expected rainfall in 24 hours for a 100 year return period was obtained from Planos et al. (2004). This map was made for the whole country, but for the provincial analysis it was cut and resampled. The values range from 148.6 to 853.96 mm/24h with an average of 301.6 mm/24h (127.77 SD). Figure 4E shows a high contrast between the northern and southern parts of the province. Rainfall in the northern part, close to Baracoa, usually comes from the northeast (Atlantic region) and is controlled by the relief. The mountains in the central eastern part serve as a barrier for rain clouds and the area south of that is a semi-arid zone.
- The second triggering factor was a map of the peak ground acceleration (PGA) with 100 years return period obtained from García (2003). In this region the PGA values (Figure 4F) are highly influenced by the Caribbean-North American plate boundary located south of the province and by the high seismic activity zone south of Santiago de Cuba city. Intra-plate seismic activity has been also recently detected in the mountainous part of Guantánamo province.

3.3. Methods for susceptibility and hazard analysis

A schematic overview of the complete methodology for landslide risk assessment at provincial level is given in Figure 5. The method started with a comprehensive landslide inventory, and the collection of input data on landslide causal factors and elements at risk, represented in the upper part of the Figure. The next step was to generate a number of landslide susceptibility maps for the five different landslide types, using a combination of heuristic approaches, probabilistic models (the Weights of Evidence) and Artificial Neural Networks. The susceptibility maps were converted into hazard maps, based on the landslide densities of the susceptibility classes and the temporal probability of landslide occurrence. This resulted in five hazard maps (SHazard to LSHazard, indicated in the middle part of Figure 5). The vulnerability and risk assessment methodology, represented in hatched area in Figure 5, are discussed in Castellanos and van Westen (2009).

Fig. 5 Flowchart

As part of the hazard analysis two models were applied for estimating the spatial probabilities of landslides occurrence: Weights of Evidence (WofE) modelling and Artificial Neural Networks. Before selecting the appropriate model, a detailed heuristic analysis was carried out in order to recognize the spatial relationship between the factors and the landslides. The authors do not intend to give a comprehensive explanation on these models since there are many publications on this (Bonham-Carter, 1996; Arora et al., 2004; Lee et al., 2004; Kanungo et al., 2006; Neuhauser and Terhorst, 2007). Here the reasoning behind the selection of the methods and the data preparation are briefly explained.

Basic estimation of weights and heuristic analysis
The first step was to estimate the prior probabilities of each specific type of landslide in the province. Guantánamo province has an area of 6188.85 km^2, corresponding to 2 475 543 pixels of 50 by 50 m. The calculated prior probability values were very low, due to the very large area of the province as compared to the relatively small area of the landslides. For example, the prior probability for rockfall is only 0.0002088. Other landslide types have similar low prior probability values: 0.0002060 for topples, 0.0002024 for debrisflows, 0.001247 for slides and 0.0013557 for large rockslides.

The second step was to prepare the 12 factors maps for the analysis, which were described above. The factor maps with classes (geology, geomorphology, landuse and soil) were rasterized at 50 m per pixel. The continuous value maps (slope angle, aspect, internal relief, drainage density, fault distance, road distance, maximum rainfall, and maximum acceleration) were converted in 10 classes maps using statistical quantiles. After this, the map with the five landslide types was overlaid (crossed in ILWIS terminology) with the 12 factor maps. As a result of each overlay operation a joint frequency table (or crosstable) was obtained with the areas occupied by each landslide type for all factor classes. The statistical processing of these tables allowed the extraction of the density of each of the five landslide types for each of the classes of the 12 factor maps, which were used as the basic information for the susceptibility analysis, and which were compared with the prior probabilities discussed above. For example, the density of rockfall events in geological unit 13 (Fm. Rio Maya) was calculated by dividing the number of pixels of rockfall events in this unit by the total number of pixels in this unit (in this example 256/73 145=0.00349, which is almost 17 times higher than the prior probability). This calculation allowed evaluating the importance of each of the classes of the 12 factor maps, for the spatial prediction of the 5 landslide types.

After generating the landslide density, relative weights were calculated, expressed as the ratio between landslide density in the class and the landslide density in the entire map. All classes with density ratios substantially higher than 1 were considered relevant for the analysis, because they have higher frequency than the regional average. The higher the ratio, the higher is the weight of the class in the susceptibility model. For the example of rockfall and the Fm. Río Maya the resulting density ratio is 0.00349/0.00021 = 16.76. After calculating this weight value for all classes, two estimators were applied to quantify the importance of the factors, namely the "accountability" and the "reliability" (Greenbaum et al., 1995a; Greenbaum et al., 1995b). The former calculates the percentage of the total landslide population accounted by each factor as:

$$Accountability = \frac{\sum Npixsld1}{\sum Npixsld} \qquad \text{[eq.1]}$$

Where *Npixsld1* are landslide pixels in those classes of the factor map having weight values larger than 1 and *Npixsld* are landslide pixels over the entire study area. The reliability is defined as the percentage area of the variable corresponding to landslides and it is compute for each variable as:

$$\mathrm{Re}liability = \frac{\sum Npixsld1}{\sum Npixcls} \qquad \text{[eq.2]}$$

Where *Npixcls* are landslide and non-landslide pixels in the classes with weights larger than 1. Both indicators provide different but relevant results to predict landslides, although reliability is more important (Greenbaum et al., 1995a; Greenbaum et al., 1995b). Based on this type of analysis it is possible to recognize the relevance of the individual factor classes as possible contributors to landslide occurrence.

The next step is to make a selection of important factors for each landslide type, by evaluating the weights, accountability and reliability values for each factor and their classes. After analyzing the importance for the occurrence of the 5 types of landslides, the actual landslide susceptibility assessment was carried out. Two specific methods were used: weights of evidence modelling and artificial neural networks.

Weights of evidence (WofE) modeling

The weights of evidence (WofE) (Bonham-Carter, 1996) is a well proven model for landslide susceptibility assessment, and therefore only a short explanation is given here. Many landslide susceptibility studies have been carried out using this method (van Westen, 1993; Fernández, 2003; van Westen et al., 2003; Lee and Choi, 2004; Lutfi and Doyuran, 2004; Suzen and Doyuran, 2004; Neuhauser and Terhorst, 2007). Essentially, the WofE method is a bivariate statistical technique that calculates the spatial probability and odds of landslides given a certain variable. The method was very well explained with GIS examples by van Westen (1993) and Bonham-Carter (1996). The weight of evidence values estimated for each class are calculated by two weights: the positive weight (W^+) gives the importance of the presence of the class for prediction landslides, and the negative weight (W^-) gives the importance of the absence of the factor for landslide prediction. The positive and negative weights (Wi^+ and Wi^-) are defined as:

$$W_i^+ = \log_e \frac{P\{B_i \mid S\}}{P\{B_i \mid \overline{S}\}} \qquad \text{[eq. 3]}$$

and

$$W_i^- = \log_e \frac{P\{\overline{B_i} \mid S\}}{P\{\overline{B_i} \mid \overline{S}\}} \qquad \text{[eq. 4]}$$

Where,

B_i = presence of a potential landslide conditioning factor,
$\overline{B_i}$ = absence of a potential landslide conditioning factor,
S = presence of a landslide, and
$\overline{S_i}$ = absence of a landslide

The contrast factor, C_n, was also calculated. It is expressed as the difference between the weights W^+ and W^- and quantifies the predictability of each class. Finally, the positive weights calculated for all the classes of the factor maps that occur simultaneously in a certain location (raster cell), and all negative weights of all the other classes, that do not occur at that location, are summed up, to produce the final weight, $W_{map.}$. The main steps in the GIS-based WofE analysis, which was carried out in a script file, are presented in Figure 6.

Fig. 6. Weight of evidence flowchart

The landslide inventory map was overlain with each factor map (geology, soil, etc.) to produce joint frequency tables with the number of pixels in each combination of the two maps. Some table calculations need to be accomplished to estimate the weights of evidence for each class and produce the weights maps corresponding to each factor. Finally, the weight maps were added to obtain the final weights of evidence map. This map is then again crossed with the initial landslide map to calculate the success rate (Chung and Fabbri, 1999).

The WofE has shown good success rates in many applications including mineral exploration (Carranza, 2002). The main drawback of the WofE, however, is that the model assumes conditional independence between variables, which makes that it may lead to relatively higher weights in certain areas that do not fit with the landslide data perfectly. However, as mentioned by Bonham-Carter (1996), WofE provides a simplification that, when used carefully, is useful for evaluating the relative contributions of the separate factor classes, and allows a good selection of the most relevant factors to be used in the analysis.

Artificial Neural Network (ANN)

One of the relatively new computational models applied for landslide hazard and susceptibility assessment is artificial neural networks (ANNs). Lee at al. (2003a; 2003b) present examples of case studies in Korea. They also show a combination of ANN for determination of weights used with spatial probabilities to create a landslide susceptibility index map (Lee et al., 2004). Rainfall and earthquake scenarios as triggering factors for landslides have been used in hazard assessment with ANNs (Lee and Evangelista, 2006; Wang and Sassa, 2006). Several studies recognize ANN as a promising tool for these applications and most of them use a Multi-layer Perceptron (MLP) network and a back-propagation algorithm for training the network (Arora et al., 2004; Ercanoglu, 2005; Ermini et al., 2005; Gomez and Kavzoglu, 2005; Wang et al., 2005).

More critical analyses compare ANNs with statistical models and fuzzy weighing (Lu, 2003; Miska and Jan, 2005; Yesilnacar and Topal, 2005; Kanungo et al., 2006). Melchiorre et al. (2006) did further research on the behaviour of a network with respect to errors in the conditioning factors by performing a robustness analysis. While above studies use ANN in landslide hazard or susceptibility zonation, Neaupane and Achet (2004) applied ANNs for monitoring mass movement.

ANNs are defined as non-linear function approximators extensively used for pattern recognition and classification (Bishop, 1995; Haykin, 1999). Neurons are the basic units of a neural network.

They are organized to compute a non-linear function of their input(s). A neuron receives input(s) with an assigned weight (s), which influence the overall output of the neuron. It is possible to allocate more than one layer of neurons and pass the information and weights from one layer to the next one. The structure of layers, the weights and the connections, known as network topology, determine the behaviour of the network (see Figure 7).

Fig. 7 Spatial ANN

There are several possible network topologies. Figure 7 shows a multi-layer feed-forward (i.e., no loops are present) with a 3-layer structure. This topology can approximate any non-linear continuous function (Bishop, 1995) and so it is considered to be suitable to evaluate landslide susceptibility. In order to make a spatial assessment of landslide susceptibility, the analysis was carried out using a pixel-based approach. A vector of factor values (x) for each pixel is the input of the network. Initially, the network needs to be trained. During the training phase the weights are optimised to minimize an error function by applying any back-propagation algorithm (Rumelhart et al., 1986), standard optimization techniques (Press et al., 1992), or a randomized algorithm (Montana and Davis, 1989).

The topology is crucial for the training capabilities: simple topologies make it more difficult to approximate complex functions, whereas an increase of complexity leads to a reduction of generalization capability. There are some techniques used to estimate the appropriated topology and improve generalization. We used the early stopping technique (Caruana et al., 2000).

Preparing the factors maps is also a relevant task. ANNs can work with different types of variables such as class maps (like geology) and continuous data (like rainfall) in the same dataset. Since ANNs can approximate any non-linear function, thematic maps such as geology, geomorphology or soil, do not need to be ranked, which is a clear advantage over the WofE method. It is recommendable, however, to use some ranking relationship with respect to landslide occurrence (Melchiorre et al., 2007). Another advantage with respect to the weights of evidence method is the possibility to work directly with continuous maps without any need of reclassification. The network is forced to find de relationship between the given classes, or continuous variables and the landslide occurrences. The whole process could be divided in two phases: training and simulation (Figure 8).

Fig. 8 ANN flowchart

In this study, the analysis was performed by using the Levenberg-Marquardt algorithm (Marquardt, 1963; Hagan and Menhaj, 1994) to train the Multi-Layer Perceptron (MLP) network and the early stopping technique to improve its generalization capability. The landslide inventory database was randomly subdivided in three subsets: a training set (75% of the landslides) used to optimize the weights, a validation set (12.5%) used to stop the network algorithm before the network starts learning noise in the data, and a test set (12.5%) to evaluate the prediction capability of the network. An equal number of samples were also randomly selected from non-landslide areas. In order to obtain results representative of the whole data and not conditioned by a specific subdivision of the dataset, a "10 cross-folder validation procedure" was introduced in the analysis. This procedure consists of subdividing the landslide inventory database 10 times in training, validation, and test set. The initial topology with only one neuron in the hidden layer and the training set goes into the training process. The training is stopped when the error on the validation set starts to increase. This process was repeated 20 times increasing one neuron in the hidden layer each time. The ANN topology with minimum error was used in the simulation phase to produce a predicted landslide map.

The ANN performance measurements, used to estimate the prediction capability on the test set, are: sensitivity, specificity, and overall accuracy. The sensitivity is the percentage of correctly classified landslides (i.e., true positive); the specificity is the correct samples classified as no landslide (i.e., true negative); the accuracy assesses the goodness of the classification, since it evaluates the correct samples classified as landslide or as not landslide.

4. Results

The results of accountability and reliability for each indicator are shown in Table 1. Geomorphology and geology are predominant factors in most of landslide types. Regarding triggering factors, large rockslides, topples and slides are more associated to PGA values, whereas rockfalls and debrisflows are better predicted by rainfall. The weights of the classes of the 12 factor maps were used to analyze the role of each factor map in predicting the occurrence of the five different landslide types. For each type, the factors and classes to be included in the analysis were selected and the best method (ANN and WofE) were choose for the susceptibility assessment. The following sections describe the conceptual models for each landslide type (see also Figure 2 for some illustrations of the types).

Table 1 Accountability and reliability

Rockfalls
Rockfall events are mostly located in the east, southeast and along the south coast of Guantánamo province. They are associated to different types of marine terraces, and biodetritic limestones of Jaimanitas, Río Maya and Cabo Cruz formations, but also to small outcrops of marble and crystallized dolomites of the Chafarina formation. The relationship between rockfalls and slope angle is three times higher in slopes with more than 23 degrees. The internal relief did not show the expected relation, due to the fact that it was created in GIS using a neighbourhood operation with a window of 1 km and therefore the local terrace scarps, where landslides are most frequent, are not adequately represented. Rockfalls are also associated with a low drainage density (< 2.58 km/km^2) as they occur in a semi-arid environment mostly due to the physical weathering. Initially it was considered that rockfalls may be associated with roads cuts, but the analysis showed no statistical relationship. It was also not possible to find a clear relationship with fault distance. As far as triggering factors are concerned, rockfalls are associated with relatively low values of maximum daily rainfall intensity ranging from 225 up to 300 mm in 24 hours, indicating that sporadic rainfall (chemical weathering) and long dry periods with high solar radiation (physical weathering) are playing a major role in the limestones for rockfall occurrence. The most important relationship was found with high values of peak ground acceleration, as the area where rockfalls occur is close to the Caribbean/North American plate boundary, which implies a periodical shaking and a slow but continuous uplifting.

Topples
Toppling features exhibit similar relationships as rockfalls, but they predominately occur in the south, along the coast on uplifted fluvio-marine terraces, tectonic-erosive small mountains, with folded and monoclinal structures, with karst phenomena. They are associated with the following geological formations: Cabo Cruz, Río Maya, Yateras and Maquey, composed of detritic, biodetritic and biogenic limestones, marls and to a lesser associated with sandstones, clays and limonites. The area is mostly covered by red limestone soils, with natural forest and pastures. For toppling features the relationship with slope angles is not straightforward. Although the densities increase with higher slope values almost all classes of slope angle have toppling events. The

internal relief shows a clear increase in weights with values higher than 211 m/km^2. They are also associated with very low rainfall (0 to 225 mm/24h) and high PGA values (0.2755 to 0.2900 g) which is explained by their geographic location along the south coast. The drainage density, and roads distance were not considered relevant factors, based on the statistically derived weights, and expert knowledge on their occurrence. The derived weights didn't support the expert assumption that toppling phenomena are very closely related to faulting, probably due to the generalization of the fault types in the buffer zone creation.

Debris flows

Debris flow features are distributed over several different geomorphological units in Guantánamo province but happen most frequently in structural-tectonic hills and mountains with horst and graben structures. Obviously, the accumulational zones of debris flows appear often on alluvial fans. Debrisflows are strongly related to colluvial deposits and also to mafic, ultramafic and granodiorite lithological complexes. The soil relationship was rather diverse with predominant occurrence in tropical black soils and the grayish tropical latosols. The prominent landuse types are natural pasture and natural forest. A positive relationship was found with slope angles (> 18 degrees), with slopes facing east and southeast, and with high values of internal relief. Debrisflows are also associated with moderate values of drainage density and road distance. There was no relationship with distance to faults, maximum daily rainfall and PGA. Debrisflows events were encountered in very high dry areas as well as in areas with high rainfall amounts.

Slides

Slide type movements, further on referred to as slides, are the most abundant type of mass movements and are distributed throughout the study area. Therefore, their spatial pattern related to geology, geomorphology, soil and landuse is not well defined which indicates that the same landslide type occurs under different conditions and one single model could not be defined for all cases. Therefore only the main controlling factors are mentioned here. Many slides are located in small massive mountains or pre-mountains and less common in structural-tectonic hills. The main lithologies related to slides were the ophiolite complex, but they are also found in marine-palustral deposits from Middle-Late Pleistocene. The soils in these lithologies are typical latosols and less evolved latosolic soils on basic and ultrabasic igneous rocks as well as grayish tropical saline and grayish limestone rich tropical soils. The slides happen most predominantly on landuse types such as natural forest and partially cacao plantations. The slides occur in slopes steeper than 14 degrees with an increasing relationship in higher slopes classes. The slopes facing north to southeast have an increasing amount of slides. Also there is a positive correlation between slides and internal relief. Drainage density, fault distance and road distance, however, show not clear relationship with slides. Slides were found most in areas with either very high or very low values of maximum daily rainfall, which could be due to two different types of slides. One based mostly on dry conditions with sporadic precipitation and other type characteristic of high rainfall expectance. There is no relation with PGA values for this type of event.

Large rock slides

Large rockslides were considered as a separate group because they occur under different environmental conditions as small landslides. The rock slides in the Guantánamo province are predominantly related to landforms in the Caujerí fault scarp and in the northern part of Baitiquirí (See Figure 1). They contain alluvial and colluvial deposits originated from limestones of Yateras formation located in an area with natural pasture and forest cover. Although large volumes of colluvial materials are present, related to older landslides, the soil formation in this area is shallow and very limited. These slides are strongly related to the highest slope classes facing east and southeast and also to the highest internal relief classes. The large slides are

located in areas with low drainage density probably due to the presence of old landslides and karstic phenomena. There are not many roads in this area, so these are not considered as contributing factor. Although tectonic movements play an important role in the occurrence of these slides, there are many minor faults in the area. Since fault types were not differentiated when creating the distance to fault factor map, no clear relation was found between tectonic lineaments and landslide occurrence. Landslides are predominantly related to high PGA values and low values of maximum daily rainfall, but they are known to be caused by extreme rainfall during a hurricanes, as was the case with the Jagüeyes landslide.

4.1. Generation of susceptibility maps

Once the contributing factors were analyzed for each landslide type, WofE or ANNs was selected for susceptibility assessment. The weights of evidence model was selected for the generation of the susceptibility maps for debris flows, rockfalls, topples and large rock slides, because for each of them it was possible to outline the main controlling factor classes, and define the conditions under which they occur. However, for slide movements, no clear sets of conditional factors classes could be differentiated. We decided to select ANNs to assess slide susceptibility, since ANNs are able to model non-linear relationships and interactions among variables. After the generation of the resulting weight maps for the five landslide types, success rate curves were generated. These curves were also used to classify them in 4 classes: no susceptibility, low, moderate and high susceptibility. In Figure 9 these susceptibility maps are presented and in Figure 10 the success rates for each landslide type are shown. The numbers of pixels with and without landslides per class are shown in Table 3. The success rate for rockfalls, topples and rock slides show very good results, as less than 10 percent of the susceptibility maps with the highest scores contain over 80 percent of the existing landslides. This is due to the fact that these events occur under specific conditions in the area, which can be identified clearly based on the combination of specific classes of the causal factor maps.

Fig. 9 susceptibility maps

Fig.10 success rates

Regarding the slide type movements, the susceptibility was carried out by means of weights of evidence and ANNs. ANNs were used in order to obtain a model with higher prediction capability. According to accountability and reliability values, several combinations of factors were tested in the weight of evidence analysis. We chose to perform the analysis by means of ANNs using 2 sets of data: the first one contained class and continuous variables, the second one only class factors. In general, the models obtained by ANNs showed better performance, as the success rate curve is much steeper (See Figure 10). Between the 2 models carried out by using ANNs, the one obtained with both class and continuous variables has a higher success rate. When the continuous variables are classified, ANNs are forced to find relationship between those classes and landslide occurrence, but those classes can not be optimal for the modelling. This can explain the better results when using continuous variables. Table 2 shows the performance measurements and the selected number of neurons in the hidden layer for each of the 10 subdivisions of the landslide inventory database. Taking into consideration the accuracy, the network with 14 neurons was chosen (Melchiorre et al., 2007). Using this network, the whole study area was classified and the map in Figure 9E was obtained. The regional influence of geology, rainfall and more locally the slope angle can clearly be noted.

Table 2 ANN performance

As provincial authorities need a single map for management and planning, the 5 individual susceptibility maps were combined in an overall susceptibility map (See Figure 9F), by selecting the highest susceptibility class for each pixel among the five maps. As a result 12.88 % of the province was classified having high susceptibility of at least one landslide type, 17.45 % were classified as moderate and 23.98 % of the area as low.

4.2. From susceptibility to hazard

The analysis presented so far only resulted in landslide susceptibility maps, which indicate the relative likelihood of an area for landslide occurrence. However, in order to be able to use these maps for landslide risk assessment, information on the spatial and temporal probability should also be included. Unfortunately no temporal information was available on the landslide occurrences, either through historical records or multi-temporal image interpretation. Therefore a simplified procedure was followed, in which the hazard is calculated by multiplying three probabilities.

$$H = P_E * P_S * P_T \qquad \text{[eq. 5]}$$

In which:

- H is hazard;
- P_E is the event probability, defined as the probability that if a landslide occurs of a given type, it happens in the particular susceptibility class.
- P_S is the spatial probability, defined as the probability that if a landslide occurs within a given susceptibility class, a pixel in this class might be hit.
- P_T is the temporal probability, P(T), defined as the annual probability of occurrence of a particular landslide type.

The calculation was made based on the combination of the landslide inventory map (with 5 landslide types) and the 5 susceptibility maps, classified in four classes (low, moderate, high, very high). This method is selected because the output of a heuristic approach or a WofE approach is a susceptibility map with scores that are not considered to be spatial probabilities. The values resulting from ANN are considered to be spatial probability values. However, for uniformity, the five maps were analyzed similarly. The hazard only considers the landslide initiation. Runout is not considered given that this study is done on a regional scale.
The event probability is calculated as follows:

$$P_E = \frac{Npix(sld, cls)}{\sum Npix(sld, cls)} \qquad \text{[eq.6]}$$

In which Npix(sld,cls) refers to the number of pixels of landslides of a certain type within a particular susceptibility class. The event probability assumes that the landslides happening in the various susceptibility classes have more or less the same dimensions. The event probability is directly related to the success rate and normally the high susceptibility class has an event probability of 0.8 or higher.
The spatial probability is calculated using the following equation:

$$P_S = \frac{Npix(sld, cls)}{Npix(cls)} \qquad \text{[eq.7]}$$

In which Npix(cls) refers to the number of pixels of a susceptibility class. Susceptibility maps having a very steep success rate, like the ones shown in Figure 9 will have high values for both the event probability as well as the spatial probability.

The temporal probability could not be obtained statistically, due to the lack of sufficient temporal landslide data. Therefore, an estimation was made based on detailed photointerpretation and on field observation using geomorphological evidence. Debris flow events happen frequently, but their depositional features are more easily masked out by vegetation growth, and human intervention. Small debris flows and slides happening in the heavily forested and wet part in the north of the province are more difficult to recognize by photointerpretation and after about 20 years vegetation is recovered again. In the southern part with dry condition and scarce vegetation, erosional processes affect the landslide features and make them less visible over time. The scarps remain visible for a longer time, especially in bedrock and can still be identified after 20 years. Rockfalls that happen along the marine terraces in the south remain clearly visible for a large number of years, whereas the ones occurring along road cuts are more rapidly modified. Large rock slides occur with very high magnitude and their features remain clearly visible over very long periods. These events are associated to weak tectonic areas and to extreme rainfall events or earthquakes. Topples are also associate with tectonic areas but they have a higher frequency than the rockslides. Based on geomorphological reasoning and by using the triggering factors of rainfall and PGA for different return periods, this study assumed a temporal probability of 1/100 years for large rock slides, 1/50 years for rockfall and topples and 1/20 for debris flows and slides. For the calculation of the hazard using equation [5] this means that the landslide area of the small slides used to calculate the event probability and spatial probability has originated in a 20 year period, and the one for rockslides in a 100 year period. Table 3 shows the calculation of hazard for the five landslides types. The highest hazard values were obtained for large rockslides mainly due to the large size of these events, and the concentration of these events in clearly defined areas, making the spatial and event probability very high.

Table 3 hazard calculation

The results obtained for the hazard analysis allowed to characterize the landslides problems in Guantánamo province for each landslide type. However, a complete quantitative hazard assessment can only be carried out if good historical records are available of landslide occurrences, coupled with analyses of triggering events. As both types of data are not sufficiently available for Cuba, the hazard, and therefore the risk can only be expressed semi-quantitatively.

5. Discussion and conclusions

A landslide hazard assessment study was carried out in Guantánamo province at a regional scale (most input maps were at scale 1:100 000, and the raster maps had a pixel size of 50 by 50 meters). The study started by analyzing the landslide inventory in the territory and the potential causal factors related to the occurrences of different landslide types. Based on the photointerpretation, fieldwork observations and the spatial correlation with the hazard indicators, each landslide type was described as detailed as possible to create a descriptive model of causal factors. The study proved that the five landslide types that were differentiated were controlled by substantially different combinations of causal factors, and that in the susceptibility analysis, each of these should be analysed separately. The results also showed that some of the landslide types were related to clearly definable classes of the causal factor maps, resulting in very steep success rate curves, in particular for the rockfall, rockslides, and toppling features.

The study demonstrated the usefulness of selecting different types of techniques, by combining a heuristic approach with WofE and ANN. Despite its limitations the weights of evidence (WofE) is considered a useful tool in landslide susceptibility assessment basically as an exploratory tool in analysing the importance of individual causal factor classes. This model doesn't look at the combined effect of combinations of factor classes, unless these are specifically constructed by the expert. Only in this way the WofE can be a useful supporting tool in an interactive, mainly heuristic, approach. The relatively good success rate curves for several of the landslide types analyzed, clearly illustrate this. However, WofE produces much worse results in situations where there is no clear relationship between the causal factor classes and landslide occurrences, or when there are several failure mechanisms for the same landslide type. In this study, ANNs were also applied to analyse the susceptibility to slides. The weight of evidence values did not show clear relations between slides and classes of factors. This can be an indication of the presence of more than one landslide generation model or a complex non-linear generator model with variables interaction. Because ANNs can cope with this problem, better results were obtained as compared with those of WofE. Applying ANNs with multiple training and included sensitivity analysis showed good results that could be replicated in other areas.

The user should make a selection of the optimal models to use, considering the type of landslides, the complexity of the spatial relationship with causal factors, and the degree of interactivity. Weights of evidence can be selected for rather simple cases, and is computationally much simpler, but doesn't provide optimal solutions in complex situations. In such conditions, ANNs or other multivariate models are more appropriate.

The conversion from susceptibility to hazard has proven to be very difficult in this area, due to lack of temporal landslide information. Temporal probability was assumed based on expert opinion on the landslide processes and the triggering factors, but not on actual frequency analysis of the landslide inventory due to inexistent of this information. The study did not allow taking into account different return periods for the same landslide type, as there was no magnitude-frequency information available. The effects of landslide runout were also not considered, given the regional scale of the study and the large pixel size that was used for the large study area.

6. Acknowledgements

The research is carried out within the United Nations University- ITC School on Disaster Geo-Information Management and a national research project for landslide risk assessment in Cuba. Authors would like to thanks to national and provincial civil defence authorities in Cuba, which support this investigation and data collection.

References

Aleotti, P., Chowdhury, R., 1999. Landslide hazard assessment: summary review and new perspectives. Bulletin of Engineering Geology and Environment 58, 21-44.

Andriani, G.F., Walsh, N. and Pagliarulo, R., 2005. The influence of the geological setting on the morphogenetic evolution of the Tremiti Archipelago (Apulia, Southeastern Italy). Natural Hazards and Earth System Sciences, 5: 29-41.

Arora, Das, G. and Gupta, 2004. An artificial neural network approach for landslide hazard zonation in the Bhagirathi (Ganga) Valley, Himalayas. International Journal of Remote Sensing, 25(3): 559-572.

Bishop, C.M., 1995. Neural networks for pattern recognition. Oxford University Press, Inc.

Bonham-Carter, G.F., 1996. Geographic Information Systems for Geoscientists: Modeling with GIS. Pergamon, Elsevier Science Ltd., 398 pp.

CARIS GIS, 2007. CARIS GIS User Manual. CARIS, Fredericton, New Brunswick, Canada.

Carranza, E.J., 2002. Geologically-Constrained Mineral Potential Mapping. Examples from the Philippines. ITC Publication No. 80, ITC, Enschede, 496 pp.

Carrara, A. et al., 1991. GIS techniques and statistical models in evaluating landslide hazard. Earth surface processes and landforms, 16(5): 427-445.

Caruana, R., Lawrence, S. and Giles, C.L., 2000. Overfitting in Neural Nets: Backpropagation, Conjugate gradient, and Early Stopping, Proceedings of Neural Information Processing Systems 2000, pp. 402-408.

Castellanos, E., 2000. Design of a GIS-Based System for Landslide Hazard Management, San Antonio del Sur, Cuba, case study. M.Sc. Thesis, International Institute for Aeroespace Survey and Earth Sciences (ITC), Enschede, 108 pp.

Castellanos, E. et al., 2003. SIGEOL: Diseno del sistema de informacion geológica de Cuba : escala 1:100.000, Instituto de Geología y Paleontología (IGP), La Habana, Cuba.

Castellanos, E. et al., 1998. Cartografía de los Principales Procesos Geológicos Amenazantes del Este de la Región Oriental. 14655, Instituto de Geología y Paleontología, MINBAS, La Habana, Cuba.

Castellanos, E. and Van Westen, C.J.., 2007. Generation of a landslide risk index map for Cuba using spatial multi-criteria evaluation. Landslides, 4 (4): 311-325

Castellanos, E. and Van Westen, C.J., 2008. Qualitative landslide susceptibility assessment by multicriteria analysis; a case study from San Antonio del Sur, Guant´anamo, Cuba. Geomorphology, 94 (3-4): 453-466

Castellanos Abella, E.A. and van Westen, C.J. 2009. Landslide vulnerability and risk assessment of Guantánamo province, Cuba. (in preparation)

Chung, C.-J.F. and Fabbri, A.G., 1999. Probabilistic prediction models for landslide hazard mapping. Photogrammetric Engineering & Remote Sensing, 65(12): 1389-1399.

Díaz, J.L., Magaz, A.R., Portela, A., Bouza, O. and Hernández, J., R., 1990. El relieve de Cuba. Ciencias de la Tierra y el Espacio, 18: 33-44.

Dikau, R., Brunsden, D., Schrott, L. and Ibsen, M.-L., 1996. Landslide Recognition. Identification, Movement and Courses. Wiley & Sons, Chichester, etc., 251 pp.

Duman, T.Y. et al., 2005. Landslide inventory of northwestern Anatolia, Turkey. Engineering Geology, 77: 99-114.

Ercanoglu, M., 2005. Landslide susceptibility assessment of SE Bartin (West Black Sea region, Turkey) by artificial neural networks. Natural Hazards and Earth System Sciences, 5(6): 979 - 992.

Ermini, L., Catani, F. and Casagli, N., 2005. Artificial Neural Networks applied to landslide susceptibility assessment. Geomorphology, 66(1-4): 327-343.

EMNDC, 2002. Informe Nacional presentado por la delegación cubana en la consulta regional "Gestión local y reducción de riesgo en los asentamientos humanos de la Cuenca del Caribe", Estado Mayor Nacional de la Defensa Civil (ENMDC) La Habana, Cuba.

Fernández, T., 2003. Methodology for Landslide Susceptibility Mapping by Means of a GIS. Application to the Contraviesa Area (Granada, Spain). Natural Hazards, 30(3): 297-308.

García, J., Slejko, D., Alvarez, L., Peruzza, L. and Rebez, A., 2003. Seismic Hazard Maps for Cuba and Surrounding Areas. Bulletin of the Seismological Society of America, 93(6): 2563-2590.

Gomez, H. and Kavzoglu, T., 2005. Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. Engineering Geology, 78(1-2): 11-27.

Greenbaum, D., Bowker, M.R., Dau, I., Bropsy, H., Greally, K.B., McDonald, A.J.W., Marsh, S.H., Northmore, K.J., O'Connor, E.A., Prasad, S. and Tragheim, D.G., 1995a. Rapid methods of landslide hazard mapping: Fiji case study. Technical Report WC/95/28, British Geological Survey (BGS), Natural Enviromental Research Council, Keyworth, Nottingham.

Greenbaum, D., Tutton, M., Bowker, M.R., Browne, T.J., Buleka, J., Greally, K.B., Kuna, G., McDonald, A.J.W., Marsh, S.H., O'Connor, E.A. and Tragheim, D.G., 1995b. Rapid methods of landslide hazard mapping: Papua New Guinea case study. Technical Report WC/95/27, British Geological Survey (BGS), Natural Enviromental Research Council, Keyworth, Nottingham.

Guzzetti, F., Carrara, 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(1-4): 181-216.

Hagan, M.T. and Menhaj, M., 1994. Training feedforward networks with the Marquardt algorithm. IEEE Transactions on Neural Networks, 5(6): 989-993.

Haykin, S., 1999. Neural Networks. A Comprehensive Foundation (2nd Edition). Prentice Hall, New Jersey, USA.

IGME, 1987. Manual de Taludes. Línea Punto Tres, Madrid, Instituto Geológico y Minero de España 456 pp.

ISDR, 2004. Cuba: a model in hurricane risk management. In: B. Leoni (Editor), ISDR. Media room. ISDR, Geneva, pp. Press release by ISDR

Iturralde-Vinent, M., 1991. Deslizamientos y descensos del terreno en el flanco meridional de la Sierra Maestra, Cuba sudoriental. In: Colectivo de Autores (Editor), Morfotectónica de Cuba Oriental. Editorial Academia, La Habana, pp. 24-27.

Kanungo, D.P., Arora, M.K., Sarkar, S. and Gupta, R.P., 2006. A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. Engineering Geology, 85(3-4): 347-366.

Lee, S. and Choi, J., 2004. Landslide susceptibility mapping using GIS and weight of evidence model. International Journal Geographical Information Science, 18(8): 789-148.

Lee, S. and Evangelista, D.G., 2006. Earthquake-induced landslide-susceptibility mapping using an artificial neural network. Natural Hazards and Earth System Sciences, 6(5): 687 - 695.

Lee, S., Ryu, J.-H., Lee, M.-J. and Won, J.-S., 2003a. Use of an artificial neural network for analysis of the susceptibility to landslides at Boun, Korea. Environmental Geology, 44: 820-833.

Lee, S., Ryu, J.-H., Min, K. and Won, J.-S., 2003b. Landslide susceptibility analysis using GIS and artificial neural network. Earth surface processes and landforms, 28: 1361-1376.

Lee, S., Ryu, J.-H., Won, J.-S. and Park, H.-J., 2004. Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. Engineering Geology, 71(3-4): 289-302.

Lu, P., 2003. Artificial Neural Networks and Grey Systems for the Prediction of Slope Stability. Natural Hazards, 30(3): 383-398.

Lutfi, M. and Doyuran, V., 2004. A comparison of the GIS based landslide susceptibility assessment methods: multivariate versus bivariate. Environmental Geology, 45: 665-679.

Magaz, A. et al., 1991. El Complejo de Formas del Relieve Gravitacional en la Franja Costera Baitiquirí-Punta Maisí, Provincia de Guantánamo, Cuba. In: Colectivo de Autores (Editor), Morfotectónica de Cuba Oriental. Editorial Academia, La Habana, pp. 28-43.

Malamud, B.D., Turcotte, D.L., Guzzetti, F. and Reichenbach, P., 2004. Landslide inventories and their statistical properties. Earth Surface Processes and Landforms, 29(6): 687-711.

Marquardt, D., 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. SIAM 11, 431-441. Journal of Applied Mathematics, 11: 431-441.

Melchiorre, C., Castellanos, E. and Matteucci, M., 2007. Analysis of sensitivity in Artificial Neural Network models: application in landslide susceptibility zonation, Guantánamo Province, Cuba. 4th EGU General Assembly. April 15-20, 2007 Vienna, Austria. Geophysical Research Abstracts, 9: 10615.

Melchiorre, C., Matteucci, M. and J, R., 2006. Artificial neural networks and robustness analysis in landslide susceptibility zonation, International Joint Conference on Neural Networks, July 16-21, 2006, Vancouver, BC, Canada, pp. 8808-8814.

Mesa, A., Colom, C., Trémols, A.J., Pena, J. and Suaréz, O., 1992. Características edafológicas de Cuba. Según el mapa en escala 1:50,000. Editorial Científico - Técnica, 189 pp.

Miska, L. and Jan, H., 2005. Evaluation of current statistical approaches for predictive geomorphological mapping. Geomorphology, 67(3-4): 299-315.

Montana, D.J. and Davis, I., 1989. Training feed forward neural networks using genetic algorithms, Proceedings of the Third International Conference on Genetic Algorithms 3, pp. 762-767.
Neaupane, K.M. and Achet, S.H., 2004. Use of back propagation neural network for landslide monitoring: a case study in the higher Himalaya. Engineering Geology, 74(3-4): 213-226.
Neuhauser, B. and Terhorst, B., 2007. Landslide susceptibility assessment using "weights-of-evidence" applied to a study area at the Jurassic escarpment (SW-Germany). Geomorphology, 86(1-2): 12-24.
Nichol, S.L., Hungr, O. and Evans, S.G., 2002. Large-scale brittle and ductile toppling of rock slopes. Canadian Geotechnical Journal, 39(4): 773-788.
ONE, 2004. Anuarios estadísticos terroriales, año 2003, CD-ROM. Oficina Nacional de Estadísticas (ONE), Ciudad de La Habana.
ONE, 2007. Anuario Estadístico de Cuba 2006. Anuario Estadístico de Cuba. Oficina Nacional de Estadísticas (ONE), La Habana.
Pacheco, E. and Concepción, L., 1998. Factores que Originan los Deslizamientos de Tierra. Afectaciones en la Municipio Mariel. In: C.N.d.I.G. (CNIG) (Editor), III Congreso de Geología y Minería. Sociedad Cubana de Geología (SCG), La Habana, Cuba, pp. 528-530
Parise, M., 2002. Landslide hazard zonation of slopes susceptible to rock falls and topples. Natural Hazards and Earth System Sciences, 2: 37-49.
Pérez, N., 1983. Aspectos ingeniero geológicos del levantamiento geológico de Cuba Oriental. In: Editorial Científico Técnica (Editor), Contribución a la geología de Cuba Oriental. Instituto de Geología y Paleontología. ACC., pp. 173-185.
Planos, E., Limia, M. and Vega, R., 2004. Intensidad de las precipitaciones en Cuba, Instituto de Meteorología, Ciudad de La Habana.
Portela, A., Díaz, J., Hernández, J.R., Magaz, A. and Blanco, P., 1989. Sección Relieve IV.3.2-3 Geomorfología. In: Instituto de Geografía Tropical (Editor), Nuevo Atlas Nacional de Cuba. Instituto Geográfico Nacional de España, La Habana, Cuba.
Press, W.H., Teukolsky, S.A., Flannery, B.P. and Vetterling, W.T., 1992. Numerical Recipes in C: The Art of Scientific Computing. University Press, Cambridge, UK.
Rumelhart, D.E., Hinton, G.E. and Williams, R.J., 1986. Learning representations by back-propagating errors. Nature, 323: 533-536.
Sims, H. and Vogelmann, K., 2002. Popular mobilization and disaster management in Cuba. Public Administration and Development, 22: 389-400.
Soeters, R. and van Westen, C.J., 1996. Slope Instability. Recognition, analysis and zonation. In: A.K. Turner and R.L. Schuster (Editors), Landslide: Investigations and Mitigation. Special Report 247. Transportation Research Board. National Research Council. National Academy Press., Washington, D.C, pp. 129-177.
Suzen, M.L. and Doyuran, V., 2004. Data driven bivariate landslide susceptibility assessment using geographical information systems: a method and application to Asarsuyu catchment, Turkey. Engineering Geology, 71(3-4): 303-321.
Tamrakar, N.K., Yokota, S. and Osaka, O., 2002. A toppled structure with sliding in the Siwalik Hills, midwestern Nepal. Engineering Geology, 64(4): 339-350.
UNESCO-WP/WLI, 1993. Multilingual Landslide Glossary. Bitech Publishers Ltd, Richmond, 34 pp.
van Westen, C.J., 1993. GISSIZ : training package for Geographic Information Systems in slope instability zonation : Part 1. theory. ITC Publication Number 15. ITC, Enschede, 245 pp.
van Westen, C.J., Rengers, N. and Soeters, R., 2003. Use of geomorphological information in indirect landslide susceptibility assessment. Natural Hazards, 30(3): 399-419.
Varnes, D.J., 1978. Slope movements types and processes. In: R.L. Schuster and R.L. Krizek (Editors), Landslides: Analysis and Control. Special Report 176. Transportation Research Board, National Academy of Sciences, Washington, D.C., pp. 11-33.
Wang, H.B. and Sassa, K., 2006. Rainfall-induced landslide hazard assessment using artificial neural networks. Earth surface processes and landforms, 31: 235-247.
Wang, H.B., Xu, W.Y. and Xu, R.C., 2005. Slope stability evaluation using Back Propagation Neural Networks. Engineering Geology, 80(3-4): 302-315.
Yesilnacar, E. and Topal, T., 2005. Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). Engineering Geology, 79(3-4): 251-266.

Figure legends

Figure 1. Location map of Guantánamo province. Black line represents municipal boundaries and light gray lines represent main roads. The landslides mapped in this study are indicated as black dots.

Figure 2. Examples of landslides in Guantánamo province. A: Marine terraces with rock fall in southern coast. B: Translational rockslide in Sierra del Purial. C: Coastal rockfall in San Antonio del Sur. D: debris flow deposit in Caujerí valley. E: Rockslide in Baracoa, F: Jagüeyes landslide in Sierra de Caujerí.

Figure 3. Graph for area and frequency of landslides in Guantánamo province. Area in m^2.

Figure 4. Some of the potential causal factor maps for the landslide susceptibility assessment. A: Slope direction (aspect), B: Internal relief, C: Drainage density, D: Slope angle, E: Rainfall intensity, F: Peak ground acceleration, G: Fault distance and H: Road distance. The maps of Geomorphology, Soils, landuse and geology have too many legend units to be able to show here.

Figure 5. Flowchart for landslide risk assessment carried out in Guantánamo province. Non-shaded area corresponds to hazard assessment published in this paper. The abbreviations in the figure refer to the following aspects: D: Debrisflows, R: Rockfalls, S: Slides, LS: Large rock slide and T: Topples.

Figure 6. Schematic representation of weight of evidence method implemented in GIS.

Figure 7. Spatial implementation of generalized multi-layer feed-forward ANN scheme. X_1-X_n are input maps for the analysis, I_1-$I_n$are neurons in the input layer, W are weight values, h_1-h_n are neurons in the hidden layer and g is goal neurons or pixel predicted.

Figure 8. General flow-chart for spatial landslide hazard assessment with ANN.

Figure 9. Landslide susceptibility maps. A: debris flow, B: large rockslides, C: rockfalls, D: topples, E: slides and F: all susceptibility maps.

Figure 10. Success rates for all landslide types (left) and for slide type movements with different methods (right). ANN: artificial neural network.

Tables

Table 1. Accountability (A) and reliability (R) for each indicator and landslide type in Guantánamo.

	Rockfall		Topples		Debrisflows		Slide		Large rockslides	
Variables	A	R	A	R	A	R	A	R	A	R
Geomorphology	100.00	0.23	100.00	9.02	93.01	3.71	82.44	10.31	97.20	67.80
Slope	92.26	0.02	79.61	1.64	69.66	1.41	80.63	10.05	82.60	11.20
Aspect	76.21	0.02	77.84	1.61	52.69	1.07	57.73	7.23	72.35	9.84
Internal Relief	83.95	0.02	99.41	4.15	83.83	2.10	91.32	12.50	81.76	13.74
Drainage density	97.49	0.07	79.22	2.02	68.46	1.54	60.45	7.54	75.00	11.23
Road distance	89.56	0.02	76.08	1.92	39.32	0.80	69.97	8.73	63.92	8.66
Geology	100.00	0.23	97.06	4.66	85.63	2.19	66.70	8.82	86.86	27.78
Faul distance	70.99	0.01	80.20	1.99	57.29	1.26	93.26	12.68	96.99	15.81
Landuse	99.61	0.03	93.33	1.70	98.00	2.41	92.78	13.18	77.68	12.54
Soil	95.16	0.07	100.00	5.62	94.21	3.44	67.18	10.61	96.90	37.32
Rainfall	100.00	0.05	91.96	4.59	73.45	1.87	64.17	8.00	100.00	43.74
PGA	79.50	0.03	100.00	10.48	66.67	1.35	71.75	8.95	100.00	34.20

Table 2. Performance measurements and number of neurons obtained.

Database Subdivisions	Number of Neurons	Performance		
		Sensitivity	Specificity	Accuracy
1	11	86.22	82.88	84.26
2	12	89.47	70.47	78.93
3	16	91.72	76.42	84.67
4	11	86.25	72.48	79.07
5	8	72.16	73.94	72.91
6	8	86.05	73.7	79.32
7	13	87.25	82.88	84.92
8	10	82.15	83.87	83.1
9	14	89.58	82.83	85.52
10	10	85.04	82.38	83.67

Table 3. Landslide hazard estimations per landslide susceptibility class. * Total with susceptibility

Susceptibility Rockfall	area in pixels	Perc area	Slides in pixels	Event probability	Spatial probability	Temporal probability	Hazard $X10^{-4}$
None	2414586	97.54	0	0.0000	0.0000	0.02	0.0000
Low	37200	1.50	22	0.0448	0.0006	0.02	0.0005
Moderate	15587	0.63	81	0.1650	0.0052	0.02	0.0171
High	8158	0.33	388	0.7902	0.0476	0.02	0.7517
Total*	2475531	100.00	491	1.0000	0.0002	0.02	0.7693
Large rockslides							
None	2358294	95.26	12	0.0036	0.0000	0.01	0.0000
Low	79118	3.20	267	0.0797	0.0034	0.01	0.0027
Moderate	28156	1.14	711	0.2122	0.0253	0.01	0.0536
High	9963	0.40	2361	0.7046	0.2370	0.01	1.6697
Total	2475531	100.00	3351	1.0000	0.0014	0.01	1.7259
Topples							
None	2346673	94.79	3	0.0059	0.0000	0.02	0.0000
Low	77146	3.12	48	0.0941	0.0006	0.02	0.0012
Moderate	41658	1.68	120	0.2353	0.0029	0.02	0.0136
High	10054	0.41	339	0.6647	0.0337	0.02	0.4483
Total	2475531	100.00	510	1.0000	0.0002	0.02	0.4630
Debris flows							
None	1642390	66.34	1	0.0020	0.0000	0.0500	0.0000

Low	491839	19.87	25	0.0499	0.0001	0.0500	0.0001
Moderate	211754	8.55	125	0.2495	0.0006	0.0500	0.0074
High	129548	5.23	350	0.6986	0.0027	0.0500	0.0944
Total	2475531	100.00	501	1.0000	0.0002	0.0500	0.1019
Slides							
None	1377092	55.63	61	0.0198	0.0000	0.0500	0.0000
Low	569395	23.00	94	0.0305	0.0002	0.0500	0.0003
Moderate	327319	13.22	773	0.2504	0.0024	0.0500	0.0296
High	201725	8.15	2159	0.6994	0.0107	0.0500	0.3743
Total	2475531	100.00	3087	1.0000	0.0012	0.0500	0.4041

Figure 1

Click here to download high resolution image

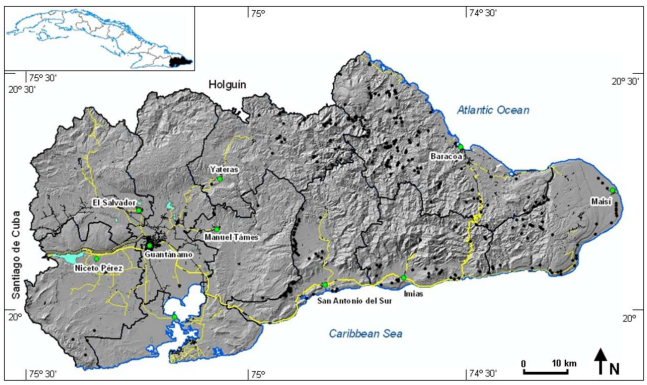

Figure 2

Click here to download high resolution image

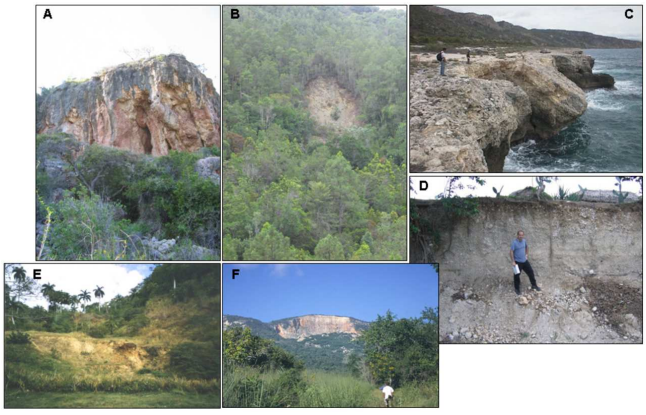

Figure 3

Click here to download high resolution image

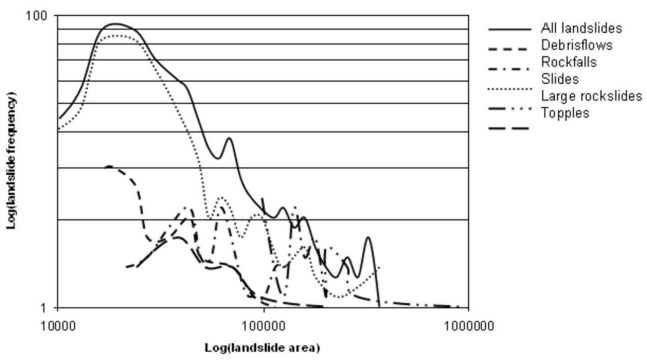

Figure 4

Click here to download high resolution image

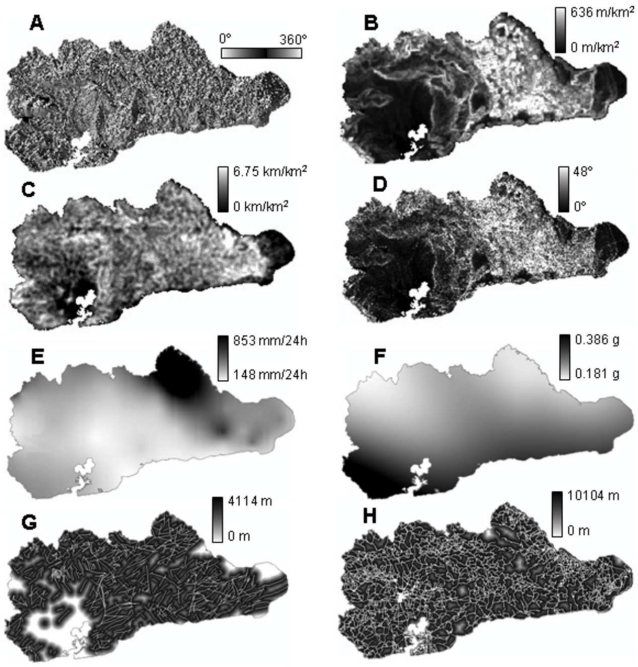

Figure 5

Click here to download high resolution image

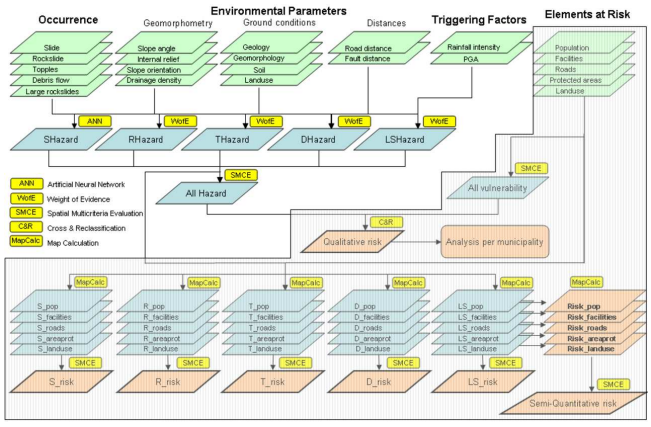

Figure 6

Click here to download high resolution image

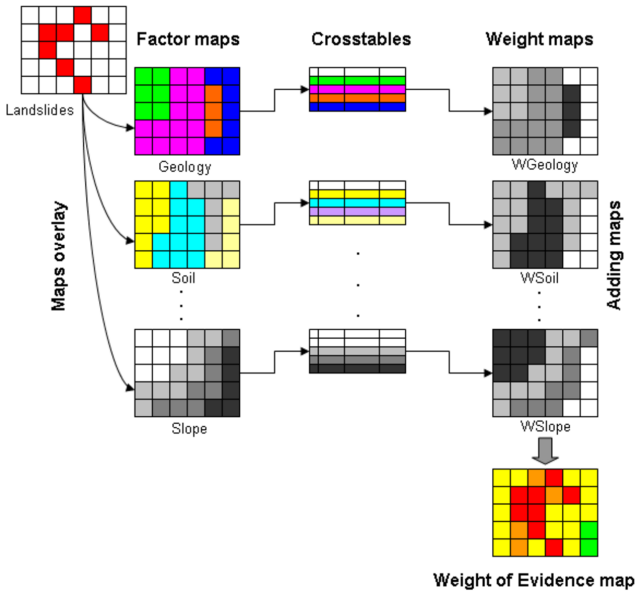

Figure 7

Click here to download high resolution image

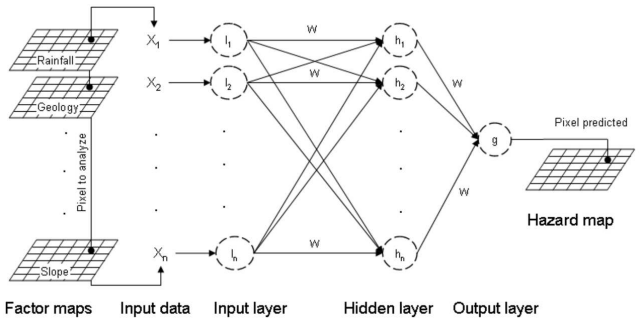

Figure 8

Click here to download high resolution image

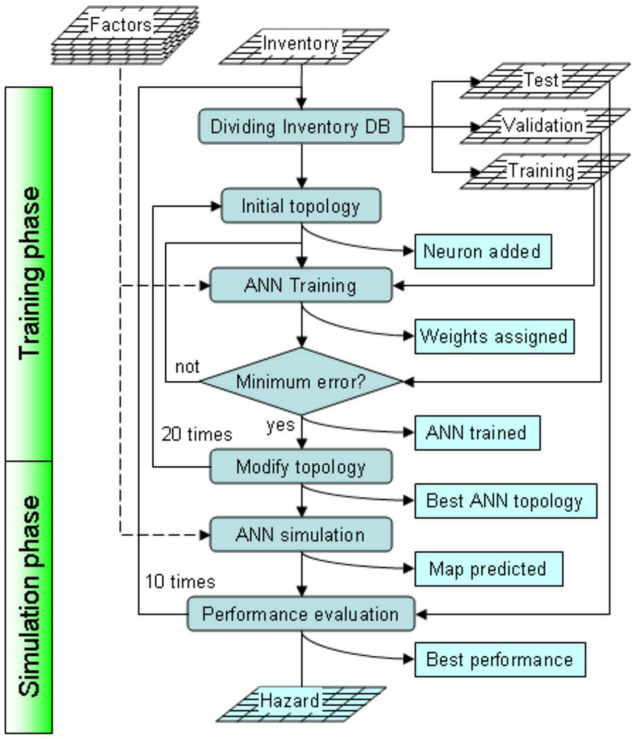

Figure 9

Click here to download high resolution image

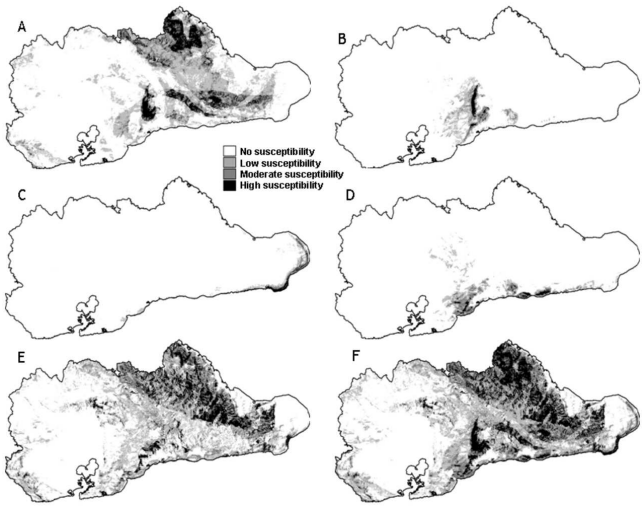

Figure 10

Click here to download high resolution image