



Automated Geo-Spatial System for Generalized Assessment of Socio-Economic Vulnerability due to Landslide in a Region

Devanjan Bhattacharya^{1*}, Jayanta Kumar Ghosh², Piero Boccardo³ and Jitka Komarkova¹

¹Inst. of Systems Engg & Informatics, FES, University of Pardubice-53210, Pardubice Czech Republic

²Dept. of Civil Engg, Indian Institute of Technology Roorkee-247667, Roorkee Uttarakhand India

³ITHACA, Politecnico di Torino, 10138, Torino, Italy

*Corresponding author, e-mail address: devanjan@gmail.com

Abstract

The paper explains a system to assess automatically vulnerability due to landslide on socio-economy of a region by categorizing landslide hazard using spatial as well as temporal causative factors. The expert system has input, understanding, expert & output modules & uses digital spatial data of causative factors of landslide. Input accepts thematic images of contributing factors for landslides, Understanding module interprets to extract relevant information as required by expert module consisting of a Knowledge Base & Inference strategy categorizing region into different susceptibilities of landslide. Overlaid on socio-economic parameters in output module for vulnerability maps of landslide on population, forestry, urban, rural, agriculture separately to ascertain the impact of landslide on socio-economy of the Tehri-Garhwal region lower Himalayas, India.

Keywords: Geo-Spatial, socio-economic vulnerability, landslide, knowledge-base, expert system.

Introduction

Experiences during critical development of landslides have shown a lack of methodologies and above all a non-systematic approach of interpreted risks [Cruden and Varnes, 1996; Boccardo and Rinaudo, 2011]. Risk management is in practice accomplished by local and regional authorities only during the critical event in a necessarily improvised way [Boccardo and Tonolo, 2008; Boccardo, 2010]. This “reactive” approach induces negative consequences on the identification procedure. For instance, the thrust is given in the direction of very expensive studies which hide the immediate nature of the problem as also the practical questions concerning the management of the risk [Uddin and Engi 2002; Michaels and Headley 2003]. Therefore, a system that can address the need of knowing immediate susceptibility due to landslide on socio-economic parameters such as forestry,

urban expanse, rural expanse, agriculture, and population under threat, will be very useful for disaster management and decision making.

Furthermore, risk can extend well beyond local damage for instance, risk of river damming which may induce major hydrological hazards: floods [Ajmar et al., 2008], inundation of sewage plants, loss of drinking water resources, so that it must be considered in a wide perspective [Australian Standard, 1999; Lindell and Prater, 2003; Ajmar et al., 2011]. So thoughts must be given towards development of a system which can cover a wide perspective as in multiple socio-economic parameters for landslide susceptibility. Such a system does not exist in the literature as has been put forward by [Einstein, 1988; Cherubini et al., 1993; Canuti and Casagli, 1994]. Besides only a few works have been done [Leone et al., 1996; Gottardi and Tonni, 2003] on vulnerability. Also there does not exist any well defined procedure to include the results of the risk analyses in land planning. Then, it is foreseen to develop the vulnerability and risk analysis for several landslides considering short and long term perspectives, direct and indirect consequences, as well as technical and social impacts, using tree event techniques. The combination of the results obtained in relation with hazard analysis and through the risk approach will allow the development of a new practical and quantified risk assessment programs which will be applied to several sites.

Many methods and techniques have been proposed to prepare landslide vulnerability maps [Varnes, 1984; Calligaris et al., 2013; Castagnetti et al., 2013]. The accuracy of these maps depends on the number and types of causative factors those have been taken under consideration for assessment. It is often necessary to combine large amount of information contained in different maps to generate a hazard map. These operations require participation of experts and are very laborious as well as time consuming. Since computers can process large amounts of data and perform repetitive and tedious operations with great precision, it is desirable to develop an automated system to produce vulnerability maps with the expert knowledge framed inside a knowledge-base. A system having the capability to correlate different landslide contributory factors is expected to improve efficiency. Further, the versatility of an automated system lies in its getting updated with new sources of data and information, along with the conventional and readily available data and information and in its flexibility to make it widely prevalent.

One of the major drawbacks of these existing systems is that they are site-specific. So the objective of this study is to develop a knowledge-based system for preparation of a landslide vulnerability map for a region of interest. It should have the capability to categorize the given region into levels of threats of landslide, to provide a map of landslide susceptibility zone and overlay the perceived threat on thematic maps pixel by pixel so that extent of threat on developmental work can be ascertained.

Background

The implementation of geo-location establishment of spatial data is a very important prerequisite. The data or information that identifies the geographic location of features and objects classes on the surface of the Earth, such as natural or constructed features, and more, is geo-spatial data [Martin et al., 2008; Sterzai et al., 2010]. Spatial data is usually stored as coordinates and topology that can be mapped. Mapping and geo-coding spatial data includes collection, information extraction, storage, dissemination, and exploitation of geodetic, imagery, topographic, and other data accurately referenced to a precise location on

the earth's surface [Backhaus et al., 2002]. It may be presented in the form of printed maps, charts, and publications [Cockings and Leung, 2008] in digital simulation and modeling databases; in photographic form; or in the form of digitized maps and charts or attributed centerline data [Oosterom et al., 2005; Martin, 2008]. Geospatial services include tools that enable users to access and manipulate spatial data. The geospatial algorithms also include programmed instruction sets, and can be modeled on artificial intelligence training modules, for the use with geospatial data [Lee and Zlatanova, 2008] to enhance efficiency and accuracy.

To geo-reference something means to define its existence in physical space of the earth. That is, establishing its location in terms of map projections or coordinate systems. The term is used both when establishing the relation between raster / vector images and coordinates and also when determining the spatial location of other geographical features [Martin et al., 2005]. Examples would include establishing the correct position of remote sensing data within a map or finding the geographical coordinates of a place-name or street-address. This procedure is thus imperative to geo-spatial data modeling and other cartographic methods. When data from different sources need to be combined and then used in a geo-spatial application, it becomes essential to have a common referencing system [Kasturi and Alemany, 1988; Martin et al., 2008; Erener and Düzgün, 2009]. This is brought about by using various geo-referencing techniques. Most geo-referencing tasks are undertaken either because the user wants to produce a new map or because they want to link two or more different datasets together by virtue of the fact that they relate to the same geographic locations.

The algorithm to geo-reference an image has the following steps: it is first needed to establish control points, input the known geographic coordinates of these control points, choose the coordinate system and other projection parameters and then minimize residuals. Residuals calculations are important because these depict the difference between the actual coordinates of the control points and the coordinates predicted by the geographic model created using the control points. Residuals provide a method of determining the level of accuracy of the geo-referencing process in any system [Sechidis et al., 2001]. There are various GIS tools available that can transform image data to some geographic control framework [Dubey et al., 2005; Chen et al., 2008], like Arc Map [ESRI, 2011], PCI Geomatica [PCIG, 2011], or ERDAS Imagine [ERDAS, 2011]. It is possible to geo-reference a set of points, lines, polygons, images, or 3D structures. For instance, a GPS device will record latitude and longitude coordinates for a given point of interest, effectively geo-referencing this point. A geo-reference must be a unique identifier. In other words, there must be only one location for which a geo-reference acts as the reference [Lee and Zlatanova, 2008; Villa et al., 2007, 2008]. The geo-referencing of a raster image is shown through Figure 1a and Figure 1b where the geo-location is divided latitude-longitudinally from the four corner points of the image area.

A map of a region which has not been geo-referenced yet has been shown in Figure 1a [Darjeeling admin., 2007]. The principle of geo-referencing requires at least four well-oriented known coordinates on the map to be geo-referenced. The coordinates should be prominently identified on the map. The spacing of the known coordinates should be such that they form some sort of quadrilateral or expansive geometry, not lie in a line. The map in Figure 1a is shown bounded in Figure 1b in the four corners having known coordinates. The images are the exact overlaps of each other which has been ensured through zero mismatches in the x-direction and y-direction of the image layout.

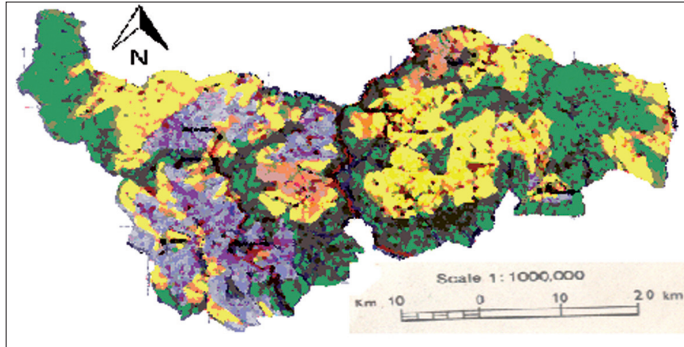


Figure 1a - Test factor-map (Land Use Land Cover) to be geo-referenced [Darjeeling admin., 2007].

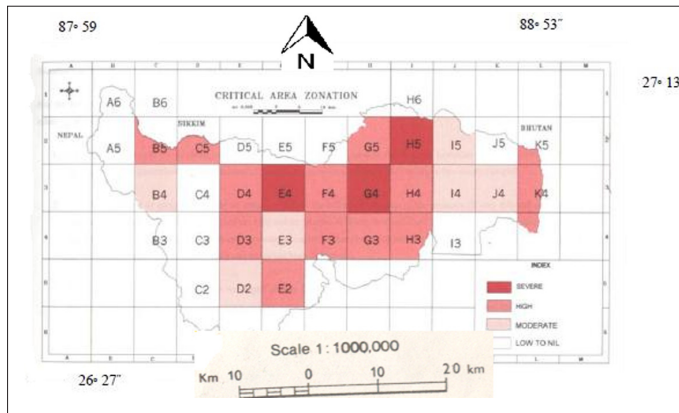


Figure 1b - Base for geo-referencing of the previous map by grids [Darjeeling admin., 2007].

A simple algorithm is sufficient to check the one-to-one pixel correspondence between two such images the first of which is non-referenced. The resultant image would have a specified location available for each pixel. The location details are to be stored in a separate array from the array storing the pixel values. The two arrays would have mutual correspondence between the base addresses.

Figure 1b demonstrates the correspondence between grids and geo-referencing. The whole area shown in the figure is bound between designated latitude and longitude. And subsequently, grids are assigned to the entire expanse. In the figure, 13 x 5 grids completely cover the area and dividing the bounding latitude-longitude uniformly, rest of the regions can be geo-referenced in respective grids accurately. The other task is to establish correspondence between the geo-referenced grids and the pixels of the image read by the system. The solution is to keep the number of grids same as the number of pixels in the image. This enables one-to-one relationship with the grids and the pixels. The four known coordinates have been shown in Figure 1b and these are at the four corners.

The top-left corner corresponds to (0,0)th pixel, and the rightmost bottom corner is the last pixel of an $M \times N$ sized image i.e. the $(M-1, N-1)$ th pixel. The other two corners are similarly identified in Figure 1b. The rest of the pixels can be geo-referenced by dividing the length and breadth as captured by the latitude-longitude with the pixel-wise length and breadth. This fundamental approach [Ghosh and Bhattacharya, 2010] towards developing an algorithm for geo-referencing would suffice when the purpose is hazard zonation and broadcast communication [Martin et al., 2005].

Architecture of the Developed System

The system is being developed on thematic maps based information. The architecture of the developed system has been outlined as in Figure 1. It consists of four modules based on their functionalities. These are Input Module, Understanding Module, Expert Module and Output Module [Ghosh and Bhattacharya, 2008]. The expert module consists of the Knowledge Base (KB) and Inference Engine of proposed system.

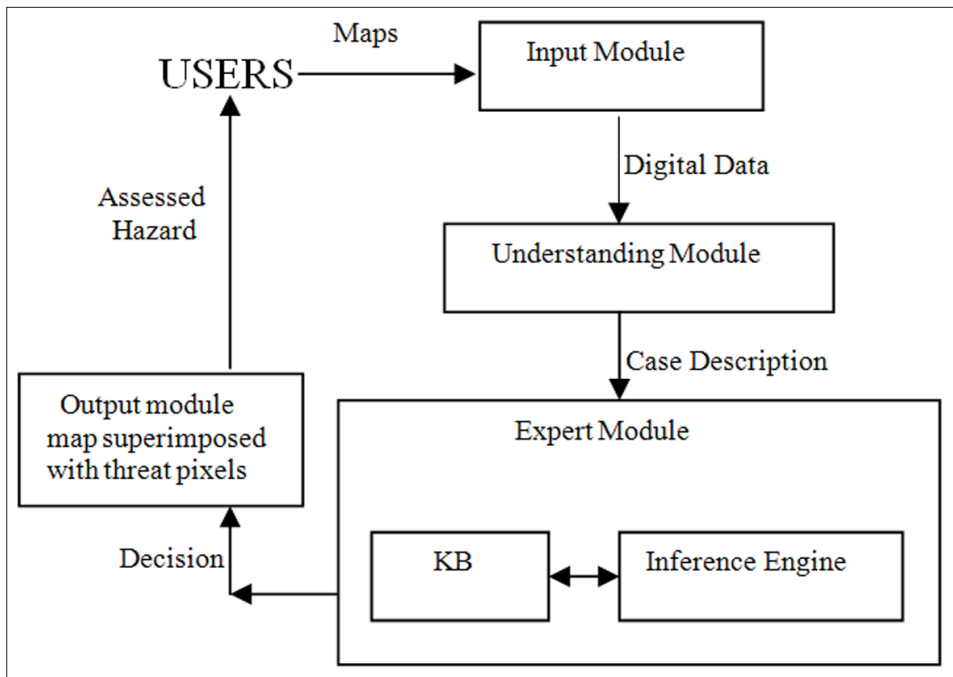


Figure 2 - Architecture of the system.

Input Module

The input module accepts thematic–maps in digital form.

Understanding Module

The Understanding Module consists of a matching algorithm that emulates the map interpretation capability of a human interpreter. The algorithm is based on correlating the code

of the legends with the different regions present in the map. This is achieved by matching the digital values of the legends with that of the digital values of the pixels present in the scanned maps (made available through input module). This leads to an understanding of the digital maps to correlate the information with the next functional module (i.e. the expert module).

Expert Module

The expert module is developed to get decision using the knowledge base within the inference strategy.

Knowledge Base

The knowledge base consists of expert information necessary to decide the intensity of landslide hazard. The expertise available in Indian Standard code IS:14496:1998 has been implemented [IS 14496, 1998] in the developed system. The KB made available in the system through a knowledge representation scheme (KRS). An object-oriented scheme [Bhattacharya and Ghosh, 2008] has been used to build KRS in the developed system. In this KRS, the rules are represented using the concepts of object-oriented paradigm. String functions are used to manipulate a string or change or edit the contents of a string. They also are used to query information about a string. They are usually used within the context of a computer programming language. The most basic example of a string function is the *length(string)* function, which returns the length of a string (not counting any terminator characters or any of the string's internal structural information) and does not modify the string. For example, *length("hello world")* returns 11. There are many string functions which exist in other languages with similar or exactly the same syntax or parameters [Kruse et al., 1997; Langsam et al., 2000; Cormen et al., 2001]. For example in many languages the length function is usually represented as *len(string)*. Even though string functions are very useful to a computer programmer, a computer programmer using these functions should be mindful that a string function in one language could in another language behave differently or have a similar or completely different function name, parameters, syntax, and results [Langsam et al., 2000; Cormen et al., 2001].

String searching algorithms, sometimes called string matching algorithms, are an important class of string algorithms that try to find a place where one or several strings (also called patterns) are found within a larger string or text [Kruse et al., 1997; Langsam et al., 2000; Cormen et al., 2001]. The simplest and least efficient way to see where one string occurs inside another is to check each place it could be, one by one, to see if it's there. So first it is searched if there's a copy of the string in the first character of the phrase; if not, it is searched to see if there's a copy of the needle starting at the second character of the haystack; if not, the third character is looked into, and so forth. One of the searching techniques maintains a sub-string that recognizes inputs with the string to search for as a suffix, another algorithm starts searching from the end of the string, so it can usually jump ahead a whole string-length at each step [Kruse et al., 1997; Langsam et al., 2000; Cormen et al., 2001].

While another type of algorithm keeps track of whether the previous *j* characters were a prefix of the search string, and is therefore adaptable to approximate string searching. In computing, approximate string matching (often colloquially referred to as fuzzy string searching) is the technique of finding approximate matches to a pattern in a string. The closeness of a match is measured in terms of the number of primitive operations necessary

to convert the string into an exact match. This number is called the edit distance between the string and the pattern [Kruse, 1997; Heileman, 2002]. The most common application of approximate matchers until recently has been spell checking. Faster search algorithms are based on preprocessing of the text. After building a substring index, for example a suffix tree or suffix array, the occurrences of a pattern can be found quickly. Some common string functions are:

Substring

Definition substring(string, startpos, endpos) returns string

Description Returns a substring of string between starting at startpos and endpos, or starting at startpos of length numChars. The resulting string is truncated if there are fewer than numChars characters beyond the starting point. endpos represents the index after the last character in the substring.

Contains

Definition contains(string, substring) returns boolean

Description Returns whether string contains substring as a substring. This is equivalent to using #Find and then detecting that it does not result in the failure condition listed in the third column of the #Find section. However, some languages have a simpler way of expressing this test.

Find

Definition find(string, substring) returns integer

Description Returns the position of the start of the first occurrence of substring in string. If the substring is not found most of these routines return an invalid index value – -1 where indexes are 0-based, 0 where they are 1-based – or some value to be interpreted as Boolean FALSE.

The inference engine

The inference scheme will pick up the facts from the input images and apply searching and matching logic to fire a rule. The inference strategy will be forward chaining definite clause inferencing. The searching and matching in this case will be for the string derived from the legend with the strings in the KB to come up with a match using the same complete matching technique [Kasturi and Alemany, 1988; Ghosh and Suri 2005] with exact string match, as used in understanding module. As soon as match will be found, the hazard rating of that factor will be put in a variable, from the KB. This will be repeated for all the causative factors, and their ratings to be stored in variables. Total estimated hazard (TEH) rating is to be calculated based on hazard ratings variables A, B, C ... [IS 14496, 1998] as: $TEH = A + B + C + \dots$, which can have maximum value 10 and the value so calculated determines the landslide susceptibility in that pixel region into one of three broad categories low, medium or high.

Inference is the act of drawing a conclusion by deductive reasoning from given facts. The algorithmic conclusion drawn is also called an inference [Ng and Abramson, 1990]. The laws of valid inference are studied in the field of logic. Human inference (i.e. how humans draw conclusions) is traditionally studied within the field of cognition, and also artificial intelligence researchers develop automated inference systems to emulate human inference [Goodenough et al., 1987; Lucks and Gladwell, 1993]. Statistical inference allows for inference from quantitative data. The process by which a conclusion is inferred from

multiple observations is called inductive inferencing [Adelman, 1989; Bhattacharya and Ghosh, 2008]. The conclusion may be correct or incorrect, or correct to within a certain degree of accuracy, or correct in certain situations.

Conclusions inferred from multiple observations may be tested by additional observations. The validity of an inference depends on the form of the inference. That is, the word “valid” does not refer to the truth of the premises or the conclusion, but rather to the form of the inference. An inference can be valid even if the parts are false, and can be invalid even if the parts are true. But a valid form with true premises will always have a true conclusion [Beerel, 1987]. For example, consider the form of the following:

1. All A are B. Premise 1
2. C is a B. Premise 2
3. Therefore, C is an A. Conclusion

Artificial Intelligence (AI) systems first provided automated logical inference and elaborate areas of research topics, leading to industrial applications under the form of expert systems and business rule engines. An inference system’s job is to extend a knowledge base automatically [Goodenough et al., 1987; Adelman, 1989]. The knowledge base (KB) is a set of propositions that represent what the system knows about the world. Several techniques can be used by systems to extend KB by means of valid inferences. An additional requirement is that the conclusions the system arrives at are relevant to its task. In the domain of programming languages, they are based on a subset of predicate calculus. Their main job is to check whether a certain proposition can be inferred from a KB (knowledge base) using an algorithm called backward chaining.

The Output Module

The output module provides a meaningful description of the intensity of landslide hazard prevalent in the region. It gets decision as input from the expert module and conveys the same to the user through display and storage. The output module displays pixels into one of the three hazard categories by three different colour codes green, yellow or red respectively for low, moderate and high hazard. Since the system works in a per pixel mode, the hazard category (once determined for a pixel) is stored pixel by pixel in an array until the whole image eventually get categorized into different hazard zones and stored by the output module in the array. The display of the array together appears as a map with the legend of the map showing the type of hazard delineated. The technique used for vulnerability assessment is the method of superimposition wherein the pixel locations stored as hazardous in the array are placed in red colour on any map with socio-economic theme, in this case urban and rural habitation. This is explained next.

Method Of Working Of System

The working of the system consists of the following broad steps:

The first step consists of accepting the input thematic maps in digital form and to make it compatible to the understanding module. It is being achieved by providing the scanned images of input thematic maps and getting it converted into identical code for regions having similar thematic information. Preprocessing has been done by passing the input of scanned thematic maps through a filter so as to clean off the noise elements crept in by scanning, and retain only the well defined regions of color. Areas of similar color get identified and

stored. The system is able to preprocess .bmp and .tif file format in which after the header of any of these two file formats are read, the actual data is then taken up for processing using a 3x3 majority filter which replaces a pixel with the most commonly occurring value in the neighboring 8 pixels. The effect is to replace the noise pixels with the most likely data value. Now the file is stored as a two-dimensional array of digital numbers for further processing. Contextual classification is done where each color in the legend is represented by a fixed tight range of digital numbers and this range is distinct from the representations of other colors. To ascertain all of the ranges of color present in a map is to do a contextual classification since in a thematic map there are uniform regions signifying some contexts. This becomes a tool for interpretation as we progress by utilizing the legend. Since the six thematic maps that have been used have standard legends, this has made the task of identifying color coded regions and storing their digitized values along with the string defining the segments a single time task. Whenever a set of input files is received, the system segregates the possible regions based on the prior information.

Once this classification is over, each pixel in a distinct region of uniform color is assigned a single code which can be referenced much faster when matching with the legend. For example, suppose a soil map for a region has five types of soil shown by five different colors. After having identified the digital value ranges for each color on the map the five ranges are replaced by numbers from 1 to 5. So each pixel in any segregation would now have a unique single representation. Similar is the case for other regions. Here, it is of utmost importance to mention that there is the presence of a segment outside the working area of a scanned map that usually has the highest value scanned. The highest value pixels in a scan are assigned the code 0 by the system. The overlapping digital values of low contrast segments are resolved by fuzzy min or max assignment. Finally the pixels of all segments are denoted by one of 1 to 5 and these five numbers are associated with the strings defining the legend.

At this stage legend understanding is achieved where the arrays are stored along with corresponding legend files. This is the stage where the system processes one pixel at a time, and only when a pixel location on each of the six maps is fully processed by interpretative algorithms and assigned a class, is a new pixel location accessed. Two loops are run on two-dimensional arrays storing pixel values of the scanned maps. The (i,j) th location for each array is accessed and the number stored there is compared with entries stored in the legend file and the explanatory string retrieved that corresponds to the number. This is the task of legend understanding as the system is relating each pixel to the description in the legend file through a common code that has replaced the color in the map. String variables store the retrieved string for each pixel for further searching and matching in the knowledge base.

The next step is the interpretation of input layers of information by means of matching algorithm. This is the understanding or interpretation module where the task of deciphering the meaning of this digital data is done, where the membership of each pixel scanned from the map is based on its proportionate intensity [Kasturi and Alemany, 1988; Kuan et al., 1989]. The subsequent step consists of decision making by the system in order to categorize region to different zones of landslide hazard. The decision making process involves searching and matching. The above interpretation is only possible when the data read in and understood finds corresponding rule for it in the rule base. Otherwise the program asks to update the rule base or cancel the query. The system currently understands soil, lithology, relief, land use and land cover characteristics from thematic maps.

Inference scheme picks up the facts from the KB (Tab. 1) and applies the operations defined for it and checks whether the rule can be fired. The searching and matching in this case is of the string derived from the legend with the strings in the KB to come up with a match. As soon as a match is found, the hazard rating of that factor is put in a variable, from the KB (Tab. 1). This is repeated for all the six causative factors and their ratings are stored in variables as given in Table 2.

Table 1 - Window to the KB [IS 14496, 1998; Gupta and Anbalagan, 1995].

| S.No. | Contributory Factor | Category | Rating |
|-----------|-----------------------------|-----------------------------------|-------------------|
| | String / alphanumeric value | String / alphanumeric value | Numeric |
| 1. | Lithology Rock | Limestone Granite Quartzite | 0.2 0.3 0.4 |
| | Lithology Soil | Clayey Sandy Barren | 1.0 1.2 1.4 |
| 2. | Slope | 0 | |
| | Very gentle | 15 | 0.5 |
| | Gentle | 25 | 0.8 |
| | Moderately steep | 35 | 1.2 |
| | Steep | 45 | 1.7 |
| | Escarpment | 90 | 2.0 |
| and so on | | | |

Total estimated hazard (TEH) rating is calculated based on these variables as: $TEH = A + B + C + D + E + F$, which can have maximum value 10. And the value so calculated determines the landslide hazard in that pixel region according to the distribution as shown in Table 3.

Case Study

The landslide hazard zones have been determined for a region of which thematic maps of the causative factors for landslide are available.

Area of study

The district of Tehri-Garhwal in Uttarakhand lying between $30^{\circ} 4' N$ and $30^{\circ} 52' 5'' N$ latitudes and $77^{\circ} 50' E$ and $79^{\circ} 3' E$ longitudes has been chosen as the area for case study on landslide hazard zonation (LHZ). For this area, the features that are taken up for a landslide hazard study are the soil type, rock type, relief, land use land cover, rainfall, slope, and groundwater condition. The KB works with these parameters where the region specific matching of attributes is done according to the prevailing conditions such as the soil being one of "clayey", "sandy", "mixed" or "fluvial"; rock formations being "phyllite", "quartzite", "limestone", "granite", "shale" or "slate".

Table 2 - Proposed Maximum LHZ Rating [Gupta and Anbalagan, 1995].

| Variable | Contributory Factor | Maximum LHZ Rating |
|----------|---------------------------------|--------------------|
| A | Lithology | 2 |
| B | Structure | 2 |
| C | Slope | 2 |
| D | Relative Relief | 1 |
| E | Land Use Land Cover | 2 |
| F | Ground Water Condition/Rainfall | 1 |
| TEH | Total (maximum) | 10 |

Input Data

Published by National Atlas and Thematic Mapping Organization (NATMO) [NATMO, 2001] under the India district planning series-index, five thematic maps of Tehri-Garhwal district at a scale of 1: 1000,000 have been provided as input. The maps are of soil, rock, relief and slope, Land Use & Land Cover (LULC), rainfall. The scanned maps and their respective legends in colour actually stores the images in a 24 – bit palette in which each of R, G, B has its assigned 8-bit representation in digital numbers. This digital number in each of RGB is what is manipulated by the different modules. The scanning has kept each pixel as $0.25^\circ \times 0.25^\circ$ latitude x longitude at the ground. The system adopted for correlating scanned pixels with actual ground expanse in lat-long has been independent of resolution in the sense that if scanned at 1000 x 1000 pixel resolution the geographical divisions per pixel will automatically be distributed by the system as it would for a 640 x 480 pixel scan or any such.

Table 3 - Classification of LHZ [adapter after Gupta and Anbalagan, 1995].

| Zone | Value of TEH | Description of LHZ |
|------|--------------|--------------------|
| I | < 5.0 | Low Hazard |
| II | 5.1 – 7.5 | Moderate Hazard |
| III | 7.5 – 10.0 | High Hazard |

Automated output

The scanned-in legend associated with the map along with explanation string describing the type of causative factor has been used in the understanding module. The input string for each factor map is searched and matched in the instance of KB and hazard rating has been found in the inference module. The procedure adopted here is based on per pixel processing so the hazard class to each pixel is assigned first and an output array is filled pixel by pixel until the final pixel has been classified. There are three hazard classes – *high*, *moderate* and *low* hazard for each pixel. There are some ambiguous or outlier pixels which could not be placed in the above mentioned three hazard classes so a fourth category of *unclassified* pixels is kept. The combination of causative factors by way of interpretation of thematic maps by the system has produced a landslide hazard zonation map as shown in Figure 3. For each pixel showing high hazard, the x,y co-ordinates are saved and superimposed on

a map of the area having a socio-economic theme. The procedure is, assuming the outside portion of the image i.e. the boundary of the study area and beyond be represented by gray scale value of 0, the first pixel count starts at the first non-zero pixel-value encountered. The pixel-value of the color-code for high threat is known before-hand so starting from the first encountered image-pixel, a count is kept till high threat pixel is reached and its location becomes the count till this pixel. Similarly, all threat-prone pixels are located and their positions saved till the end of image is reached. Now these saved positions are located on an exact same map of the area having a socio-economic relevant theme like forestry, habitation etc. and the color-code for the threat generated in those pixel locations. The associated legend gives the level of threat.

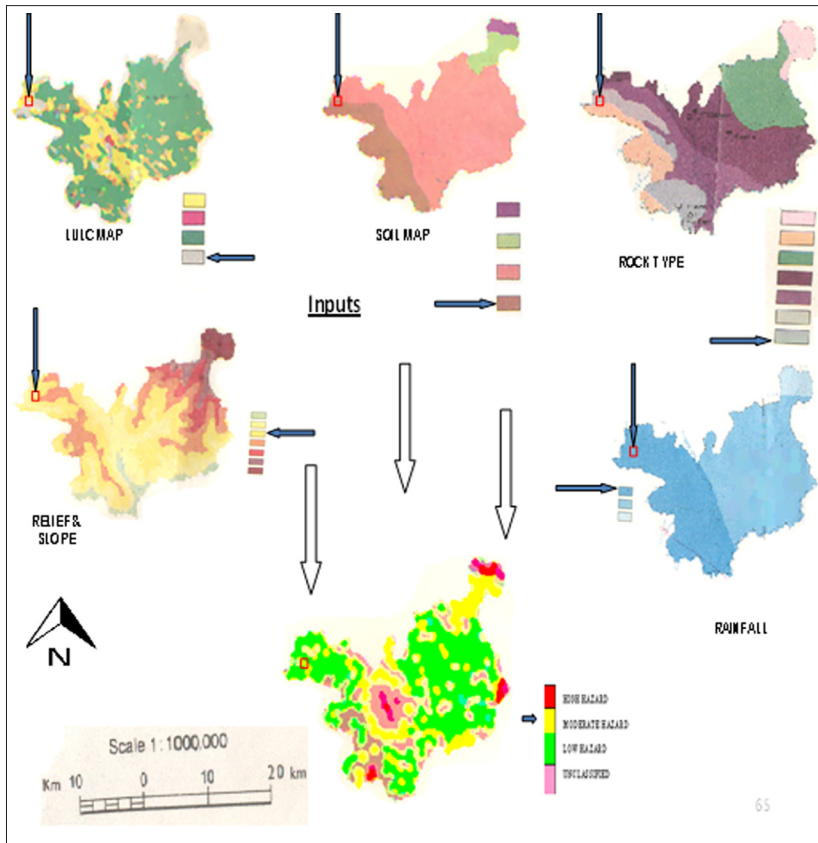


Figure 3 - The processing and output of evaluated landslide hazard from the developed system.

Results and Assessment

Conceptually, a map understanding system must have the logic programmed into it to follow map reading as a human interpreter does. And that is the capability to correlate the colour code of the legend with the colours in the map to know what the map says. This has to be

done in the digital environment and the solution is to scan the maps so that each colour is captured as digital values on a pixel by pixel basis. Same is done with the legend. Now the digital values of the coloured regions on map can be matched with the digital values in the legend and the string defining it stored in a string variable.

The logic and algorithms embedded into it are all for interpreting thematic maps. And to read a thematic map the legend associated with the map is to be taken help from. Each region on the map has explanation in the legend according to some color code. In the digital procedure, scanning both map and its legend yields digital numbers which can be compared for match inside a range. For this match there is a string defining the region. This is stored in the knowledge base. The combination of all the extracted strings gives the required output. As far as the distribution of threat on habitation is concerned, the zones of high hazard in the district of Tehri – Garhwal seems to be concentrated around the Bhagirathi valley passing through Tehri township as shown in Figure 4, 5 and 6a, b. Around Tehri hazard extends from $30^{\circ} 22' N$ to $30^{\circ} 28' N$ latitudes and $78^{\circ} 24' E$ to $78^{\circ} 28' E$ longitudes. Needless to say, moderate hazard has been found around this township. Inside the latitudes and longitudes of $30^{\circ} 10' N$ to $30^{\circ} 13' N$ and $78^{\circ} 16' E$ to $78^{\circ} 18' E$ respectively are high hazard areas which are between Muni-ki-Reti and Narendranagar. Moderate hazard areas have also been found around these townships too. Moderate hazard areas have been found around the Devaprayag and Kirtinagar habitations within latitudes and longitudes $30^{\circ} 40'$ to $30^{\circ} 46' N$ and $78^{\circ} 40'$ to $78^{\circ} 50' E$, respectively.

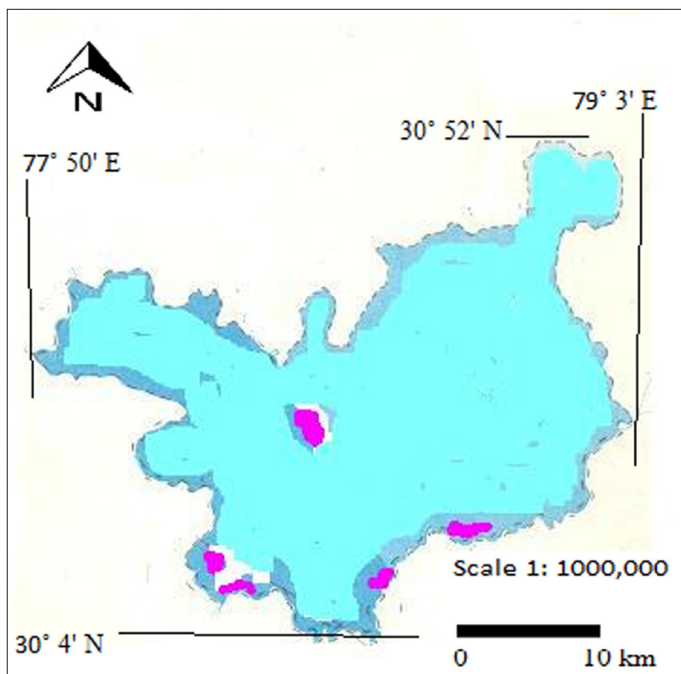


Figure 4 - Location extents of study area in latitudes and longitudes with major settlements shown in pink.

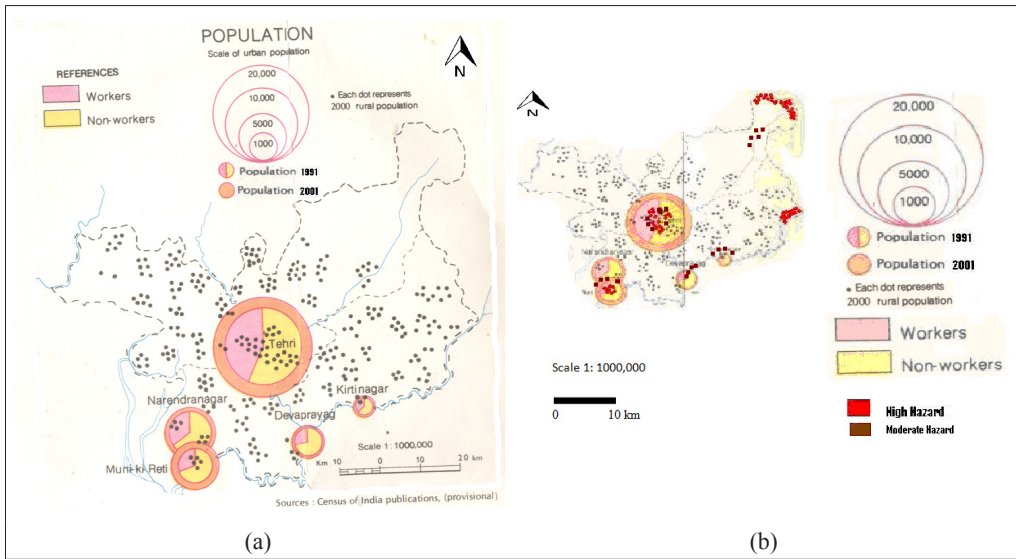


Figure 5 - Landslide hazard vulnerability on population of the area. (a) Original population map with sections of the district [NATMO, 2001]; (b) High and moderate hazard on populations.

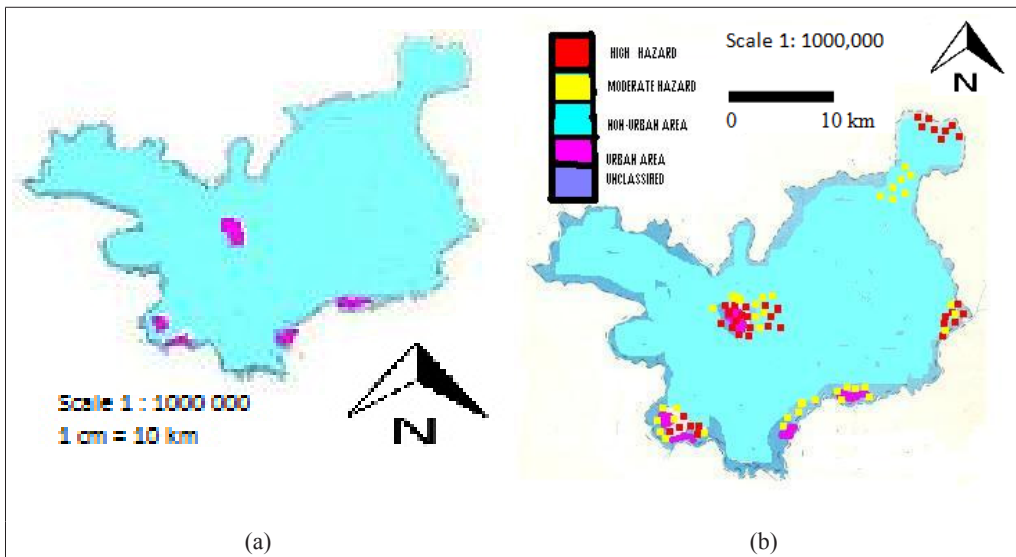


Figure 6 (a) - Urban-area of the study area; (b) Landslide hazard vulnerability on Urban area in the study area.

Also the road connecting Rishikesh - Badrinath passing through Devprayag is in high hazard zone as shown in Figure 6. This lies between $30^{\circ} 26'$ to $30^{\circ} 31'$ N latitude, and $78^{\circ} 50'$ to 79° longitude. Towards the high altitude in North-East and South-West where Alkananda river passes, are high hazard zones of the district, but it is uninhabited. The latitudinal and longitudinal extents are $30^{\circ} 47'$ to $30^{\circ} 52'$ N and 79° to $79^{\circ} 3'$ E, respectively.

Under the socio-economic parameter of urban dwellings, the regions under threat are the Tehri township between $30^{\circ} 22' N$ to $30^{\circ} 28' N$ latitudes and $78^{\circ} 24' E$ to $78^{\circ} 28' E$ longitudes. Urban dwellings of Narendranagar, Muni-ki-Reti, Devaprayag and Kirtinagar have moderate threat perception as shown in Figure 6b.

Rural dwellings are scattered across the terrain as shown in Figure 7 and have moderate to high threat perception extending between $30^{\circ} 22' N$ to $30^{\circ} 28' N$ latitudes and $78^{\circ} 24' E$ to $78^{\circ} 28' E$ longitudes. Towards the southern parts of the region where rural dwellings abound, there is moderate threat extending from $30^{\circ} 40'$ to $30^{\circ} 46' N$ and $78^{\circ} 40'$ to $78^{\circ} 50' E$.

The region is predominantly agriculture based so the study of threat to areas under cultivation is important. The arable land as shown in Figure 8a and b comes under moderate threat in the southern to south-eastern part of the Tehri-Garhwal region between the extents of $30^{\circ} 40'$ to $30^{\circ} 46' N$ and $78^{\circ} 40'$ to $78^{\circ} 50' E$. Also, there are scattered regions in the central locations of the district extending between $30^{\circ} 22' N$ to $30^{\circ} 28' N$ latitudes and $78^{\circ} 24' E$ to $78^{\circ} 28' E$ longitudes where agricultural fields are there and threat varies from moderate to high.

Forests abound in the region so it is of economic interest to know the extent to which forestry is affected by landslide threat. Towards the western to south western areas, the threat to forests by landslides is moderate to high as shown in Figure 9a and b. The southern areas have moderate landslide threat and the south-western areas have moderate to high threat perception, together lying in $30^{\circ} 10'$ to $30^{\circ} 15' N$ to $78^{\circ} 40'$ to $78^{\circ} 50' E$. Towards the higher altitudes, between $30^{\circ} 40'$ to $30^{\circ} 46' N$ and $78^{\circ} 40'$ to $78^{\circ} 50' E$ we have moderate threat to forestry.

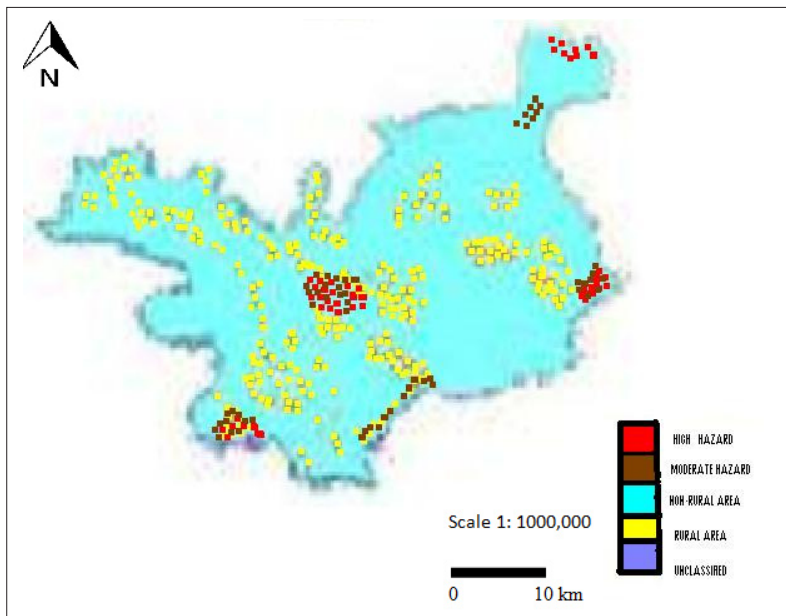


Figure 7 - Landslide hazard vulnerability on Rural area.

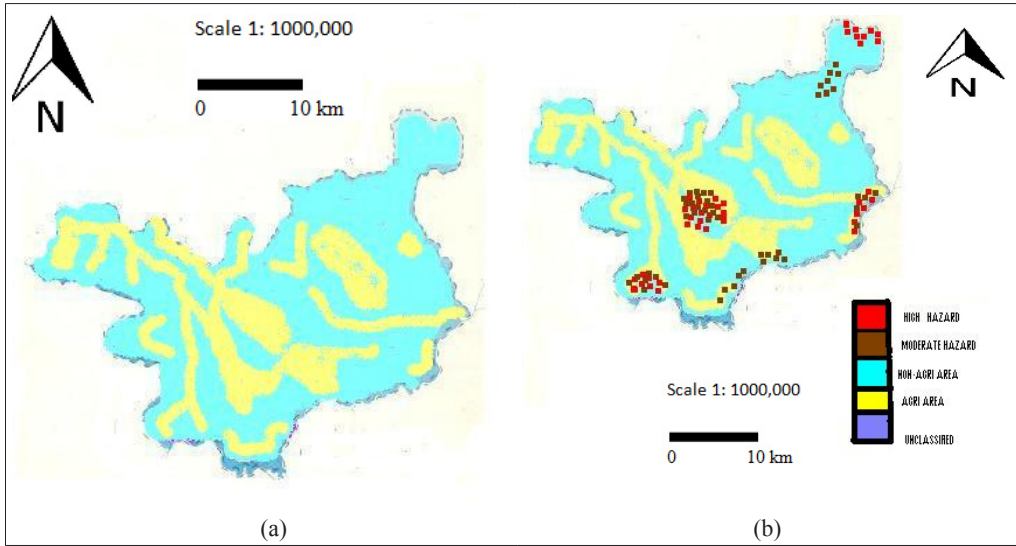


Figure 8 - (a) Agricultural map of the study area; (b) Landslide hazard vulnerability of Agricultural area in the study area.

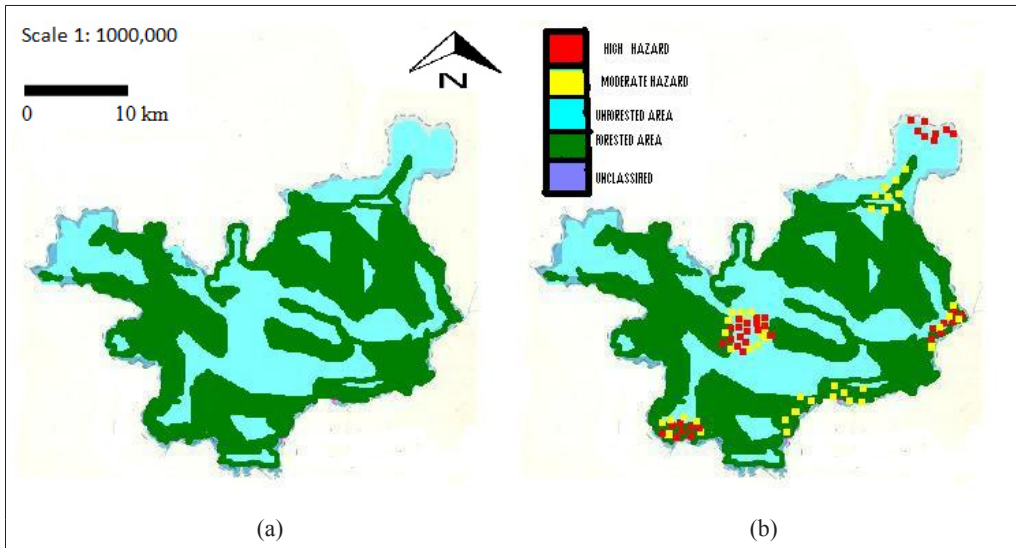


Figure 9 - (a) Forestry map of the study area; (b) Landslide hazard vulnerability on forestry of the study area.

Finally, a superimposition on the entire land use land cover map of the region with threat perception due to landslides is shown in Figure 10a and b. It gives a comprehensive idea of level of threat due to landslide on the various resources. The people having the socio-economic responsibilities like the local and governmental authorities can be notified of danger in real time and this has been a topic of consideration in some researches [Ghosh et al., 2010; Bhattacharya et al., 2011].

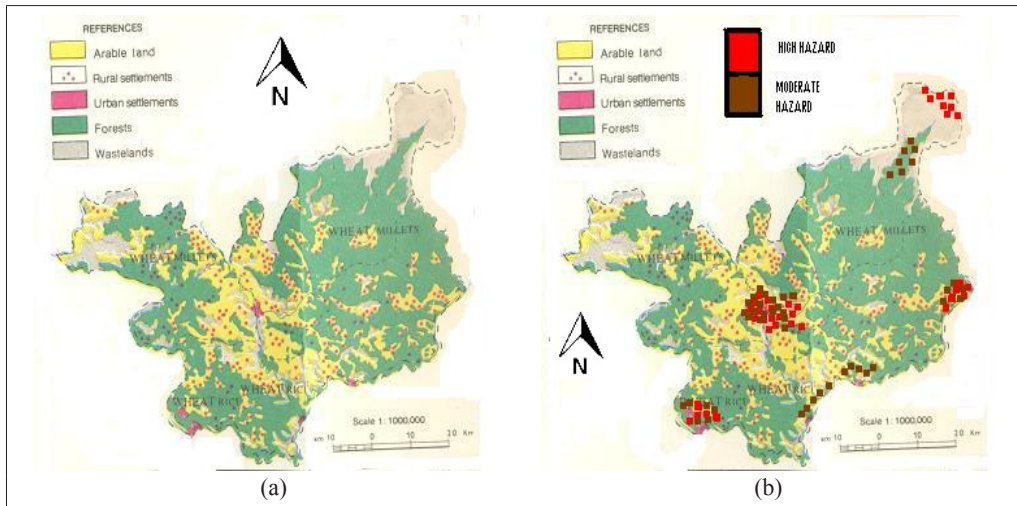


Figure 10 - (a) Land Use Land Cover map of the study area [NATMO, 2001]; (b) Landslide hazard vulnerability on the Land Use Land Cover map.

Conclusion

The generation of a landslide hazard zonation map through the system has been a fast but restrictive alternative to the more prevalent and detailed zonation by advanced models. The trade-off was expected from the start of the development of the system and through this sample study it now stands confirmed that an automated system such as this is worthwhile to be pursued for further advancements. The impact of landslide on communities can be greatly reduced if the risk is defined beforehand and appropriate management strategies are put in place. In order to assist with the protection of life and property from landslide, it is necessary to have a coordinated approach to the use of resources, an understanding of the hazard and consequences associated with landslide, and it is also possible to model landslide scenarios and their impact upon communities. In this way appropriate planning will mitigate the economic and social costs which landslides impose. Planning will reduce the impact to a community by creating an agreed set of arrangements and actions that will improve the community's ability to manage the threat. This will include devising solutions for the many problems which can be foreseen and which would, without the plan, have to be managed in an ad hoc fashion. The system issued maps could be used by land use planners, emergency managers, government at all levels, government authorities, business, the community, the insurance industry, and others as appropriate.

References

- Adelman L. (1989) - *Measurement issues in knowledge engineering*. IEEE Transactions on Systems, Man and Cybernetics, 19 (39): 483-488. doi: <http://dx.doi.org/10.1109/21.31055>.
- Ajmar A., Boccardo P., Tonolo F.G. (2011) - *Earthquake damage assessment based on remote sensing data. The Haiti case study*. Italian Journal of Remote Sensing, 43 (2): 123-128. doi: <http://dx.doi.org/10.5721/ItJRS20114329>.
- Ajmar A., Boccardo P., Disabato F., Tonolo F. G., Perez F., Sartori G. (2008) - *Early impact procedures for flood events February 2007 Mozambique flood*. Italian Journal of Remote

- Sensing, 40: 65 - 77. doi: <http://dx.doi.org/10.5721/ItJRS20084036>.
- Australian Standard (1999) - *Risk Management*. Standards Association of Australia/New Zealand.
- Backhaus R., Bock M., Weiers S. (2002) - *The Spatial Dimension of Landscape Sustainability*. Environment, Development and Sustainability, Springer-Netherlands, 4 (3): 237-251.
- Beerel A.C. (1987) - *Expert Systems. Strategic Implications and Applications*. John Wiley, New York.
- Bhattacharya D., Ghosh J.K. (2008) - *Evaluation of Knowledge Representation Schemes as a pre-requisite Towards the Development of a Knowledge Based System*. Journal of Computing in Civil Engineering, 22 (6): 348-359. doi: [http://dx.doi.org/10.1061/\(ASCE\)0887-3801\(2008\)22:6\(348\)](http://dx.doi.org/10.1061/(ASCE)0887-3801(2008)22:6(348)).
- Bhattacharya D., Ghosh J.K., Boccardo P., Samadhiya N.K. (2011) - *Wireless Hazard Communication System*. Journal of Systems and Information Technology, Emerald Publishing (United Kingdom), 13 (4): 408-424. doi: <http://dx.doi.org/10.1108/13287261111183997>.
- Boccardo P. (2010) - *Introduction to the special issue: Geo Information for Disaster Management (editorial)*. Italian Journal of Remote Sensing, 42: 33-36. doi: <http://dx.doi.org/10.5721/ItJRS2010423a1>.
- Boccardo P., Rinaudo F. (2011) - *Gi4DM-geo-information for disaster management-Torino*. Applied Geomatics, 3 (2): 59-60. doi: <http://dx.doi.org/10.1007/s12518-011-0051-3>.
- Boccardo P., Tonolo F.G. (2008) - *Natural disaster management: Activities in support of the UN system*. Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences: ISPRS Congress Book, pp. 385-396.
- Calligaris C., Poretti G., Tariq S., Melis M.T. (2013) - *First steps towards a landslide inventory map of the Central Karakoram National Park*. European Journal of Remote Sensing, 46: 272-287. doi: <http://dx.doi.org/10.5721/EuJRS20134615>.
- Canuti P., Casagli N. (1994) - *Some considerations about the evaluation of landslides risk*. Proc. of Convegno Fenomeni Franosi e Centri Abitati, Bologna, pp. 1-58.
- Castagnetti C., Bertacchini E., Corsini A., Capra A. (2013) - *Multi-sensors integrated system for landslide monitoring: critical issues in system setup and data management*. European Journal of Remote Sensing, 46: 104-124, doi: <http://dx.doi.org/10.5721/EuJRS20134607>.
- Chen J., Jiang J., Nayak S. (2008) - *Proceedings of the XXI ISPRS Congress, Part B4-TYC IV*, Beijing, China.
- Cherubini C., Giasi C.I., Guadagno F.M. (1993) - *Probabilistic approaches of slope stability in a typical geomorphological setting of Southern Italy*. Risk and reliability in ground engineering, Thames Telford, London, pp. 144-150.
- Cockings S., Leung S. (2008) - *Geo-Refer: Geographical Referencing Resources for Social Scientists*. NCRM Research Methods Festival 2008 Conference, St Catherine's College Oxford, 30 June - 3 July. URL: http://www.restore.ac.uk/geo-refer/files/RMF2008_GeoRefer.pdf (last accessed 15/01/2012).
- Cormen T.H., Leiserson C.E., Rivest R.L., Stein C. (2001) - *Introduction to Algorithms*. 2nd ed., MIT Press and McGraw-Hill, ISBN 0-262-53196-8.
- Cruden D.M., Varnes D.J. (1996) - *Landslide Types and Processes*. Landslides, Investigation and Mitigation, Special Report 247, Transportation Research Board, National Research

- Council, National Academy Press Washington DC.
- Darjeeling admin. (2007) - Darjeeling District Administration, URL: <http://www.darjeeling.gov.in/geography.html> (last accessed 26/01/2012).
- Dubey C.S., Chaudhry M., Sharma B.K., Pandey A.C., Singh B. (2005) - *Visualization of 3-D digital elevation model for landslide assessment and prediction in mountainous terrain: A case study of Chandmari landslide, Sikkim, eastern Himalayas*. Geosciences Journal, 9 (4): 363-373. doi: <http://dx.doi.org/10.1007/BF02910325>.
- Einstein H.H. (1988) - *Special lecture: Landslide risk assessment procedure*. Proc. 5th Int. Symp. on Landslides, Lausanne, Switzerland, 2: 1075-1090.
- ERDAS (2011) - *Earth Resources Data Analysis System* (ERDAS Imagine package) URL: erdas.com/products/ERDASIMAGINE/ERDASIMAGINE/Details.aspx (last accessed 17/01/2012).
- Erener A., Düzgün H.S.B. (2009) - *A methodology for land use change detection of high resolution pan images based on texture analysis*. Italian Journal of Remote Sensing, 41(2): 47-59. doi: <http://dx.doi.org/10.5721/ItJRS20094124>.
- ESRI (2011) - *Environmental Systems Research Institute*. URL: www.esri.com/software/arcgis/eval-help/quick-start-guide-93.pdf (last accessed 29/01/2012).
- Ghosh J.K., Bhattacharya D. (2008) - *Landslide Hazard Automated Zonation (LHAZ) system*. The 12th International Conference of International Association for Computer Methods and Advances in Geomechanics (IACMAG), Goa, India, IACMAG. available online at: www.civil.iitb.ac.in/~dns/IACMAG08/pdfs/G07.pdf (last accessed 17/01/2012).
- Ghosh J.K., Bhattacharya D. (2010) - *A Knowledge Based Landslide Susceptibility Zonation System*. Journal of Computing in Civil Engineering, 24 (4): 325-334. doi: [http://dx.doi.org/10.1061/\(ASCE\)CP.1943-5487.0000034](http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000034).
- Ghosh J.K., Bhattacharya D., Boccardo P., Samadhiya N.K. (2010) - *A Landslide Hazard Warning System*. ISPRS Tech. Comm. Symp.- VIII, Kyoto, Japan, pp. 1-10. Available online at: http://www.isprs.org/proceedings/xxxviii/part8/pdf/W01OE3_20100304002241.pdf (last accessed 14/01/2012).
- Ghosh J.K., Suri S. (2005) - *A knowledge based system for assessment of landslide hazard*. Proc., Indian Geot. Conf., Vol. 1, IGS, Ahmedabad, India, pp. 393-396.
- Goodenough D.G., Goldberg M., Plunkett G., Zelek J. (1987) - *Role of AI in the integration of remotely sensed data with GIS*. IEEE Transactions on Geoscience and Remote Sensing, GE-25 (3): 330-347.
- Gottardi G., Tonni L. (2003) - *Interpretation of Slope Movements Induced by Adjacent Large Landslide in Northern Italy*. Natural Hazards Review, 4 (2): 71-77. doi: [http://dx.doi.org/10.1061/\(ASCE\)1527-6988\(2003\)4:2\(71\)](http://dx.doi.org/10.1061/(ASCE)1527-6988(2003)4:2(71)).
- Gupta P., Anbalagan R. (1995) - *Landslide hazard Zonation, Mapping of Tehri-Pratapnagar area, Garhwal Himalayas*. Journal of Rock Mechanics and Tunnel Technology, 1 (1): 41-58.
- Heileman G.L. (2002) - *Data Structures, Algorithms and Object Oriented Programming*. First Edition, Tata McGraw Hill.
- IS 14496 (1998) - *Indian Standard Preparation of Landslide Hazard Zonation Maps in Mountainous Terrains - Guidelines*. Part 2 Macro-Zonation. Bureau of Indian Standard, New Delhi, pp. 1-19.
- Kasturi R., Alemany J. (1988) - *Information Extraction From Images of Paper-Based Maps*. IEEE Transactions on Software Engineering, 14 (5): 671-675. doi: <http://dx.doi.org/10.1109/32.1000000>.

- org/10.1109/32.6145.
- Kruse R., Tondo C.L., Leung B. (1997) - *Data Structures and Program Design in C*, Second Edition, Pearson Education.
- Kuan D., Shariat H., Dutta K. (1989) - *Constraint-based Image Understanding Sys. for Aerial Imagery Interpretation*. IEEE Proc., Annl. AI Systems in Govt. Conf., Vol. 1, IEEE, Washington, D.C., pp. 141-147.
- Langsam V., Augenstein M.J., Tanenbaum A.M. (2000) - *Data Structures Using C and C++*. Second Edition, Prentice Hall of India.
- Lee J., Zlatanova S. (2008) - *A 3D data model and topological analyses for emergency response in urban areas*. Geospatial information technology for emergency response (ISPRS book series), Taylor & Francis Group, London, UK, pp.143-168. URL: http://www.gdmc.nl/publications/2008/3D_data_model_topological_analyses.pdf (last accessed 13/01/2012).
- Leone F., Asté J.P., Leroi E. (1996) - *Vulnerability assessment of elements exposed to mass-movement: working toward a better risk reception*. Proc. of 7th Int. Symp. on Landslides, pp. 263-270.
- Lindell M.K., Prater C.S. (2003) - *Assessing Community Impacts of Natural Disasters*. Natural Hazards Review, 4 (4): 176-185. doi: [http://dx.doi.org/10.1061/\(ASCE\)1527-6988\(2003\)4:4\(176\)](http://dx.doi.org/10.1061/(ASCE)1527-6988(2003)4:4(176)).
- Lucks M., Gladwell I. (1993) - *A knowledge representation for multiple criteria assessment problems*. Procs. IEEE Ninth Conference on Artificial Intelligence for Applications, pp. 100-105. doi: <http://dx.doi.org/10.1109/CAIA.1993.366655>.
- Martin D., Cockings S., Leung S. (2005) - *Geo-Refer: Geographical referencing resources for social scientists*. ESRC Researcher Development Initiative Launch, Royal College of Physicians, London, 9 November. URL: <http://www.restore.ac.uk/geo-refer/files/GeoReferpre-start.pdf> (last accessed 23/01/2012).
- Martin D., Lewis A., Cockings S., Leung S. (2008) - *Geographical referencing & SASPAC*. A practical workshop, GeoRefer and SASPAC training event, City University London, 9 December. URL: <http://www.restore.ac.uk/geo-refer/files/London%20GeoRefer-SASPAC%20Dec08FINAL.pdf> (last accessed 19/01/2012).
- Martin D. (2008) - *Linking and mapping geographically referenced data*. NCRM's 2008 Autumn School for Early Career Researchers - Exploring new data sources in the social sciences, Jury's Inn Southampton, 12-14 November. <http://www.restore.ac.uk/geo-refer/files/NCRM%20Autumn%20School%20-%20Nov%202008.pdf> (last accessed 19/01/2012).
- Michaels S., Headley W. (2003) - *Actively Searching for Hazards Information*. Natural Hazards Review, 4 (4): 186-196. doi: [http://dx.doi.org/10.1061/\(ASCE\)1527-6988\(2004\)5:1\(10\)](http://dx.doi.org/10.1061/(ASCE)1527-6988(2004)5:1(10)).
- NATMO (2001) - *National Atlas and Thematic Mapping Organization*. District Planning Map Series, Govt. of India. <http://natmo.gov.in> (last accessed 05/02/2007).
- Ng K.C., Abramson B. (1990) - *Uncertainty management in expert systems*. IEEE Expert Systems, 5 (2): 29-48. doi: <http://dx.doi.org/10.1109/64.53180>.
- Oosterom V.P., Zlatanova S., Fendel E. (2005) - *Geo-information for Disaster Management*. Springer, Berlin-Heidelberg, New York, pp. 1434, ISBN 3-540-24988-5.
- PCIG (2011) - *PCI Geomatica*. URL: http://www.pcigeomatics.com/pdf/tutorials/Exporting_map.pdf (last accessed 20/01/2012).

- Sechidis L., Tsioukas V., Patias P. (2001) - *Geo-referenced 3D Video as visualization and measurement tool for Cultural Heritage*. International Archives of CIPA, Vol. XVIII-2001, pp. 293-299, ISSN 0256-1840.
- Sterzai P., Vellico M., Berti M., Coren F., Corsini A., Rosi A., Mora P., Zambonelli F., Ronchetti F. (2010) - *LiDAR and hyperspectral data integration for landslide monitoring: The Test Case Of Valoria Landslide*. Italian Journal of Remote Sensing, 42 (3): 89-99, doi: <http://dx.doi.org/10.5721/ItJRS20104237>.
- Uddin N., Engi D. (2002) - *Disaster Management System for Southwestern Indiana*. Natural Hazards Review, 3 (1): 19-30. doi: [http://dx.doi.org/10.1061/\(ASCE\)1527-6988\(2002\)3:1\(19\)](http://dx.doi.org/10.1061/(ASCE)1527-6988(2002)3:1(19)).
- Varnes D.J. (1984) - *Landslide hazard zonation; a review of principles and practice*. Natural Hazards 3, UNESCO, 63 pp.
- Villa P., Reitz T., Gomasasca M. (2007) - *The HUMBOLDT project: implementing a framework for geo-spatial data harmonization and moving towards an ESDI*. pp. 29-36. URL: http://www.earsel.org/symposia/2007-symposium-Bolzano/pdf/ch1_1_4.pdf (last accessed 21/01/2012).
- Villa P., Reitz T., Gomasasca M. (2008) - *Humboldt Project For Data Harmonization In The Framework Of GMES And ESDI: Introduction And Early Achievements*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Beijing 2008, Vol. XXXVII, Part B4, pp. 1741-1745, URL: http://www.isprs.org/proceedings/XXXVII/congress/4_pdf/303.pdf (last accessed 20/01/2012).

Received 14/02/2012, accepted 22/11/2012