Proceedings of the *Open source GIS - GRASS users conference 2002* - Trento, Italy, 11-13 September 2002

A GRASS GIS based Shell script for Landslide Susceptibility zonation by the Conditional Analysis method

Aldo Clerici

Dipartimento di Scienze della Terra, Università degli studi di Parma, Parco Area delle Scienze 157/A, 43100 Parma, Italy, tel. ++39 0521905362 fax ++39 0521905305 e-mail aldo.clerici@unipr.it

1 Introduction

Landslides are an environmental phenomenon of great socio-economic significance in many world regions. For this reason they attract the interest of a wide range of professionals, including engineers, geologists, geomorphologists, planners and policy makers. Besides the studies to improve investigation techniques and mitigative measures, there has been a constant development of quantitative methods to assess the probability of future landslide occurrence. These methods can be divided into two main distinct categories: the deterministic and the statistical methods. The former base their predictions on empirical geotechnical laws that require the collection of geotechnical parameters. The extreme variability in three dimensional space of geotechnical characteristics means that their application to wide areas tends to be expensive and so this kind of analysis is usually limited to small areas and, more often than not, to single slopes. The statistical methods, on the other hand, and especially the multivariate ones, appear to be better suited for assessing landsliding probability on a regional scale. In recent decades a remarkable leap forward has been made with these methods because of the diffusion of progressively more efficient and cheaper GISs. Since the early 1970s, dozens of papers proposing many different GIS based procedures for the assessment of future landslide probability have been published. A thorough examination of the methodological aspects, together with examples and exhaustive reference lists, can be found in Carrara *et al.* [1], Soeters and van Westen [2] and Guzzetti *et al.* [3] to quote just the most recent works. All the proposed methods are based on the assumption that 'slope-failure in the future will be more likely to occur under those conditions which led to past and present instability' [1]. In other words, landslide prediction for areas currently free of landslides is effected by evaluating the similarity between the conditions in such areas and those conditions that have led to landslides in the past. Therefore the conceptual model consists in:

1) The mapping of the landslides.

2) The mapping of a set of environmental factors which are supposed to be directly or indirectly correlated with slope instability.

3) The classification of the land surface according to the degree of landslide susceptibility on the basis of the detected statistical relationships between instability factors and instability phenomena.

Among the different multivariate techniques, i.e. those techniques based on the relationship between landslides and a number of instability factors considered simultaneously, the so called Conditional Analysis method [1] [4] is conceptually straight forward and highly compatible with GIS operating features. This method is profitably applied in relation to a particular land surface subdivision in the so called Unique Condition Units [1] or Unique Condition Subareas [4].

The method considers a number of environmental characteristics, or factors, which are thought to be directly or indirectly connected with landslide occurrence. The data layers, in which each factor is subdivided into a convenient number of classes, are crossed in order to obtain all the possible combinations of the various classes of the different factors. Each specific combination represents a Unique Condition Unit (UCU). Subsequently, the landslide presence, usually represented by the landslide area and only rarely by the landslide number, is determined within each UCU and the landslide density is computed. Assuming the already mentioned principle that slope-failure in the future is more likely to occur under those conditions which led to slope-failure in the past, the computed landslide density is equivalent to the future landslide probability. Formally, the conditional probability is given by:

P(L|UCU) = landslide area / UCU area

i.e. the probability of landslide occurrence (L) given a unique combination of factors (UCU) is given by the landslide density in that specific UCU.

Despite its conceptual simplicity, the method has some problematic aspects, mainly regarding the notion of landslide in terms of the assessment of landslide presence and the evaluation of instability factors. To overcome these problems and improve the method, in this paper these aspects are discussed and a procedure is proposed in which the upper edge of the landslide main scarp is assumed as the geomorphological feature representing the genetical part of a landslide.

Furthermore, although the method is particularly suited to GIS processing, it requires a long and well defined sequence of operations that makes the procedure complex and error prone even for well trained GIS users. So a shell script, or shell program, mainly based on GIS commands, has been created that, starting from a Landslide Inventory map and a number of factor maps, all in raster format, automatically carries out the entire procedure up to the construction of the final map portraying the different probabilities of future landslide occurrence. This makes the method accessible to a wide range of users consenting a swift analysis of large areas with a high spatial resolution.

As regards terminology, according to Brabb [5] and Soeters and van Westen [2] the term 'Landslide hazard map' often used as a generic term to define different sorts of maps portraying probabilities of future landsliding, should be restricted to maps whose prevision of landsliding is both spatial and temporal. As the map produced by the method proposed here only takes into account the spatial probability of landslide occurrence, the term 'Landslide Susceptibility map', suggested by the above mentioned authors, is preferable.

In the following pages, after an analysis of the methodological characteristics introduced in the method, the shell script is described and an application to a relatively small basin illustrated.

2 Problematical aspects of the method and characteristics of the proposed procedure

One of the limits of the Conditional Analysis method is the necessity to introduce a limited number of factors subdivided into a limited number of classes into the analysis. Otherwise, a high number of UCUs of small dimensions, and therefore of little statistical significance, could result from the crossing of the data layers. Nevertheless, this does not seem to be a big limitation as, generally, only a few factors subdivided into few classes turn out to be relevant in evaluating the stability in a given region. Furthermore, it must be considered that small UCUs, even if they are numerous, occupy small and therefore usually irrelevant portions of the investigated area.

Other aspects of the method are more open to criticism. In particular, the definition and assessment of instability factor combinations at the landslides. It is well known that, in a landslide, it is possible to identify two genetically and morphologically distinct zones: the depletion zone, i.e. the upper part of the landslide where the failure is effectively generated, and the accumulation zone, i.e. the lower part which is simply affected by the arrival of the depleted material (Fig.1).

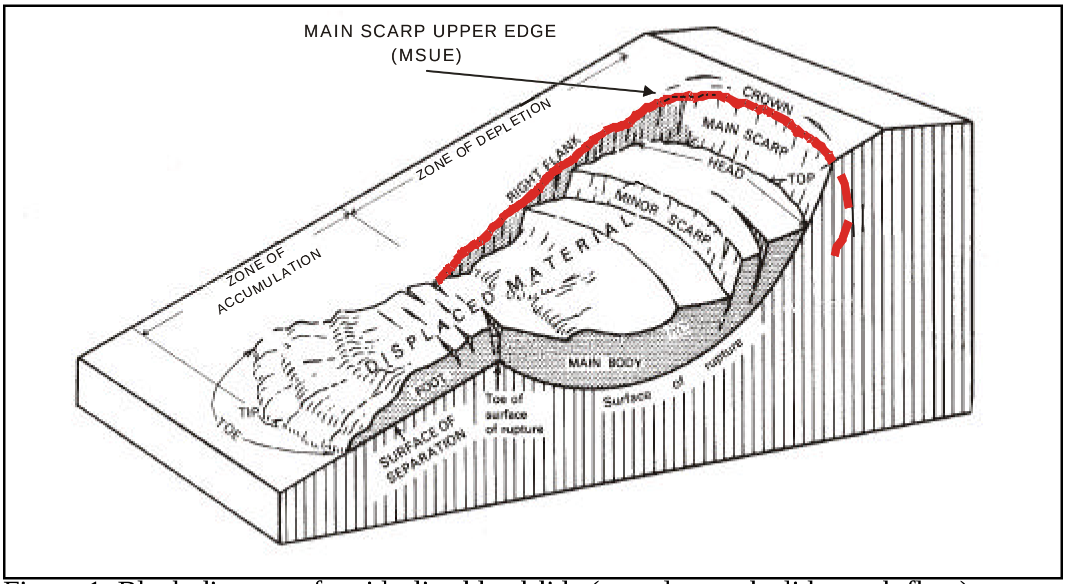

Figure 1: Block diagram of an idealized landslide (complex earth slide-earth flow). Modified from Cruden D.M and Varnes D.J., 1996.

If the whole landslide body is considered in assessing the factor combination, as it would appear was the case in the previous applications of the method, the combinations present in the accumulation zone are also considered to favour landsliding. Since the method claims to identify the conditions under which landslides are generated, the factor combinations should only be considered in the area in which the landslide originated, i.e. the depletion zone. Yet this morphological feature is generally difficult to outline in its entirety in as much as it is occupied by the displaced material (often to a great extent). Only the higher portion of this zone, the so called main scarp, and, in particular, its upper edge, is clearly evident on account of the slope angle difference with the surrounding unfailed zones, both at the contact with the landslide crown and along the flanks of the depletion zone. So the Main Scarp Upper Edge (MSUE) is the most evident morphological feature of a landslide and can be easily identified and mapped both by aerial photographs and on the field.
It must also be considered that all the statistical methods based on the conceptual model expressed above, require knowledge of the factor conditions before the landslide occurrence. With the exception of very recent landslides, the available maps and photographs obviously represent the situation after landsliding. Some factors, such as the bedrock lithology, can be supposed not to have been modified by the landslide occurrence, so the currently observable situation at the depletion zone, or along the MSUE, can be considered to be the same as at time of failure. On the contrary, other features, and in particular morphometric factors like slope angle or aspect, are substantially modified by landsliding and cannot be defined in the current depletion zone or along the MSUE. If it is virtually impossible to reliably reconstruct the old topography, it nevertheless seems plausible to assume that the morphometric characteristics which most resemble those present at the failure time are those currently observable in the immediate surroundings of the depletion zone, i.e. along the unmodified external belt of the MSUE.
Furthermore, this linear feature has the further advantage of a linear development that is practically independent of how much the material has shifted in the depletion zone, so its length is strictly proportional to the depletion zone area, at least for the same type of landslide. Therefore the MSUE length provides a reliable measurement of the landslide size.
For all these characteristics the MSUE can be considered to be an easily measurable morphological element which is particularly representative of the genetical part of a landslide and can be profitably used for determining both the instability factor

combination and the landslide dimension. So in the procedure here proposed the Landslide Factor Combination (LFC) is defined along, or very close to, the MSUE, and the total length of the MSUEs pertaining to each specific LFC is computed. The landslide density for each UCU is then obtained by dividing this length (in meters) by the area (in km^2) of the UCU with the same factor combination as the LFC. The landslide density is therefore expressed in m/km^2 and the above reported landslide probability is modified as follows:

P(L|UCU) = MSUE length / UCU area

By simply dividing the density values by the MSUEs' mean length, in the procedure the final susceptibility classes are also expressed as landslide frequency, i.e. in number of landslides / km^2.

3 The shell script

The high number of GIS commands and the operational complexity make the creation of a shell script to all intents and purposes indispensible, as has already been said. In a previous work [6] a relatively simple shell script was proposed for the construction of Landslide Susceptibility maps considering the whole landslide surface, as proposed in the original formulation of the conditional analysis method. The adoption of the MSUE as the geomorphological feature representative of a landslide, both for landslide presence quantification and for factor definition, makes the procedure much more complex requiring a new specific shell script.

In the script, commands of the Bash shell [7] [8], of the GRASS GIS [9], of the text processing language 'gawk' [10] [11] and, albeit in a limited manner, of the non-interactive line editor 'sed' [12] [13] are used. Considering that the script was implemented under Linux O.S., only free software has been used. The script, available through anonymous ftp at ftp.geo.unipr.it in pub/condan3/, contains many comments describing in detail the various operational phases. Therefore, in the following pages only the input and output characteristics and the main steps of the procedure are described.

3.1 Input and output

The shell script, named condan3.sh, works in a non interactive way, so all the parameters defining the input and output names and values are specified in the command line, as follows (the arguments in square brackets are optional):

condan3.sh **title** =*"map name"* **landsl**=*rastername* **warea**=*rastername* [**msc**=*number*] **factor1**=*rastername* [**recfact1**=*filename*] [**dst1**= *0-3*]
....
[**factor10** = *rastername* **recfact10** = *filename* **dst10** = *0-3*]
report = *filename* [**ucutable** = *filename*]
[**psmapname** = *filename* **psmapsize** = *sizename* **psmapscale** = *value*]

The parameters stand for the following:

- title: the final map header, also used as the job title.

- landsl: the name of the Landslide Inventory map. As previously described, in the procedure the MSUE is taken as the landslide's representative feature for defining both the landslide extension and the LFC. In the script, this morphological element is automatically extracted from the mapped main scarps that must therefore be distinguished from the accumulation areas by a different category value. The value assigned to the former is specified by the 'msc' parameter described below, whereas the latter can have any value other than zero. The areas outside the landslides must have no value.

- msc: the category value assigned to main scarps in the Landslide Inventory map. If not otherwise specified, the default value 1 is assumed.

- warea: the polygon delimiting the study area. All the cells inside the polygon must have any value other than 0, the external cells a 0 or null value.

- factor1 ...factor10: the names of the raster maps containing the instability factors subdivided into classes. At least one, but no more than 10, factor maps must be specified. If only one map is introduced, a univariate analysis is performed and the result is a simple frequency distribution of the landslides in the factor classes. This can help in some circumstance in deciding a preliminary class subdivision of a factor.
If a zero or a null category value occurs in any factor layer, the combinations, or UCUs, containing that category are not considered in the subsequent phases of the analysis and the corresponding areas are assigned to an 'undefined' class in the final map, even if other factor maps contain non-zero values in the same areas.
- recfat1 ....refact10: the factor reclassification files, i.e. ASCII files containing the rules for the creation of a new raster map by changing the category values, or classes, of an existing factor map. Each record of the file contains the value of the category (or categories) to be reclassified, followed by the equal sign and the value of the new category. Optionally an alphanumeric label for the new category can also be added. These files must obviously be introduced only for those factors for which a change of category is required. The original categories eventually omitted in the reclassification file, are excluded from the subsequent analysis and the corresponding areas are considered 'undefined' in the final map. The numerical suffix of the parameter, from 1 to 10, links the reclassification file to the factor to be reclassified.
- dst1....dst10: represent the distance from the MSUE, expressed as a cell number from 0 to three, at which a factor value is defined. In fact, as previously discussed, some instability factors can be evaluated in coincidence with the MSUE, others at a certain distance from it. If dstx = 0, the factor values are defined as being in coincidence with the MSUE cells. If dstx = 1, in the cells immediately adjacent to the MSUE. If dstx = 2 or dstx = 3, the factor cells are chosen at a distance of two and three cells respectively. For the omitted dstx values, a 0 value is assumed as default. Also in this case, the argument is related to a factor by the suffix number.
- report: the name of the file containing a job report (see figure 5 as an example). As well as the arguments specified in the command line, this report contains the following information:
- The region limits and cell size.
- The extension of study area as delimited by the 'warea' polygon.
- The theoretical number of UCUs, i.e. the number of all possible combinations of all the factor classes in the investigated area; the actual number of UCUs, i.e. the number of the combinations effectively occurring in the area; the total UCU area, that can differ from the extension of the study area if undefined UCUs are present.
-The MSUE total length, i.e. the extension of all MSUEs present in the area; the MSUE mean density, obtained by dividing the total length by the overall UCU area. The number of MSUEs, corresponding to the number of landslides in the area; the MSUE mean length, obtained by dividing the total MSUE length by the number of MSUEs.
- As some MSUE cells could be excluded from the analysis, as explained later, the above characteristics are also computed and reported for the 'actual' MSUEs, i.e. for the MSUE cells actually considered in the analysis.
- A table with the susceptibility class number and name, the class value expressed both as density and frequency and the class extension, in km^2 and in %. In the program, the computed landslide density values are grouped into five susceptibility classes that are defined so that the mean density value of the entire area is located in the middle position of the middle (third) class. The extension of the undefined areas (if any) and of landslides are also reported.
- ucutable: the name assigned to the optional output file containing UCU statistics. More precisely, for each UCU the table lists (see table 2 as an example):
- composition, i.e. the combination of the factor classes; factor classes are listed in the order defined by the suffix number assigned to each factor in the command line.
- extension, in number of cells and in km^2.
- MSUE length, number, frequency and density.

This table is useful for assessing the effectiveness of each single factor or the appropriateness of the factor subdivision, through multivariate statistical techniques like discriminant analysis, as performed by Carrara *et al.* [1], or weighted regression analysis, as proposed by Chung *et al.* [4].

The last three parameters are optional and must be specified only if a PostScript map is required.
- psmapname: the name of the PostScript file containing the map.
- psmapsize: the size of the map; can assume the values a0, a1, a2, a3 or a4.
- psmapscale: the scale of the final map in the form 1:xxxxxx; the scale is reduced automatically if too big to generate the map in the specified size.

3.2 Landslide susceptibility calculation

The steps of the entire procedure are schematically shown for a very simple case in figure 2.
After reading the command line arguments, the first step in the procedure is the reclassification (if requested) of the factor maps.
The reclassified maps are then crossed in order to construct an UCU map, i.e. a map containing all the factor class combinations in the study area. Each combination is distinguished in the map by a different identification number, while the factor combination is stored as a category label. If a zero or no value occurs in some portion of any factor map, as shown for the Factor B map in figure 2, or if a category is not defined in any reclassification file, the corresponding areas are excluded from the definition of the UCUs and therefore from the subsequent processing. In the final Landslide Susceptibility map, these areas are assigned to an 'undefined' category. From this map a file reporting the combination of factor classes and the extension of each UCU in the investigated area is created.
Subsequently the MSUEs are extracted from the main scarps in the Landslide Inventory map. Each extracted MSUE is made up of a linear feature with a single cell width including all the cells along the border of the main scarps that are common to the main scarp area and to the enlarged external landslide area, as illustrated in figure 3.
The length, in meters, of each MSUE cell is then computed considering half of the side length for each lateral or vertical contact with the adjacent cell and half of the diagonal length for each diagonal contact. It is worth noting that the MSUE length can increase with an increase of the spatial resolution, as the smaller the adopted cell size is, the better the MSUE cells fit the mapped MSUE. The sum of the length values of all the MSUE cells defines the MSUE total length which, divided by the overall area occupied by the UCUs, defines the MSUE total density.
Then the MSUE number is calculated, which, except in specific cases, is equivalent to the number of landslides included in the Landslide Inventory map. These specific cases occur in situations in which two or more MSUEs, mapped as pertaining to different landslides, come into contact. The adopted procedure does not register the different origin and a single MSUE is calculated.
By dividing the MSUE total length by the MSUE number, the MSUE mean length is computed.
Then, for each MSUE cell, the class value of each factor is defined. This is performed by a paired analysis of the MSUE map and each single factor map. As specified previously, to each MSUE is assigned the factor class value present in a cell of the factor map located at a distance from the MSUE cell ranging from a minimum of 0 to a maximum of three cells, as selected by the user on the basis of the type of factor being considered. The process is illustrated in figure 4 for a single landslide. In the case of a 0 distance (figure 4b), the operation is extremely simple as the values are determined in the cells of the factor map coinciding with the cells of the MSUE.
When distance = 1 (figure 4c), the factor cell search is performed by a simple neighbourhood operation using a 3x3 cell moving matrix centred on each MSUE cell. The conditions for the selection of the factor cell among the eight cells of the matrix

perimeter is that it be outside any landslide area but, obviously, inside the study area. The search proceeds clockwise starting from north and considers first the four vertical and horizontal cells in order to favour the selection of the nearest cell. If none of these cells has the necessary requisites, the more distant diagonal cells are inspected.

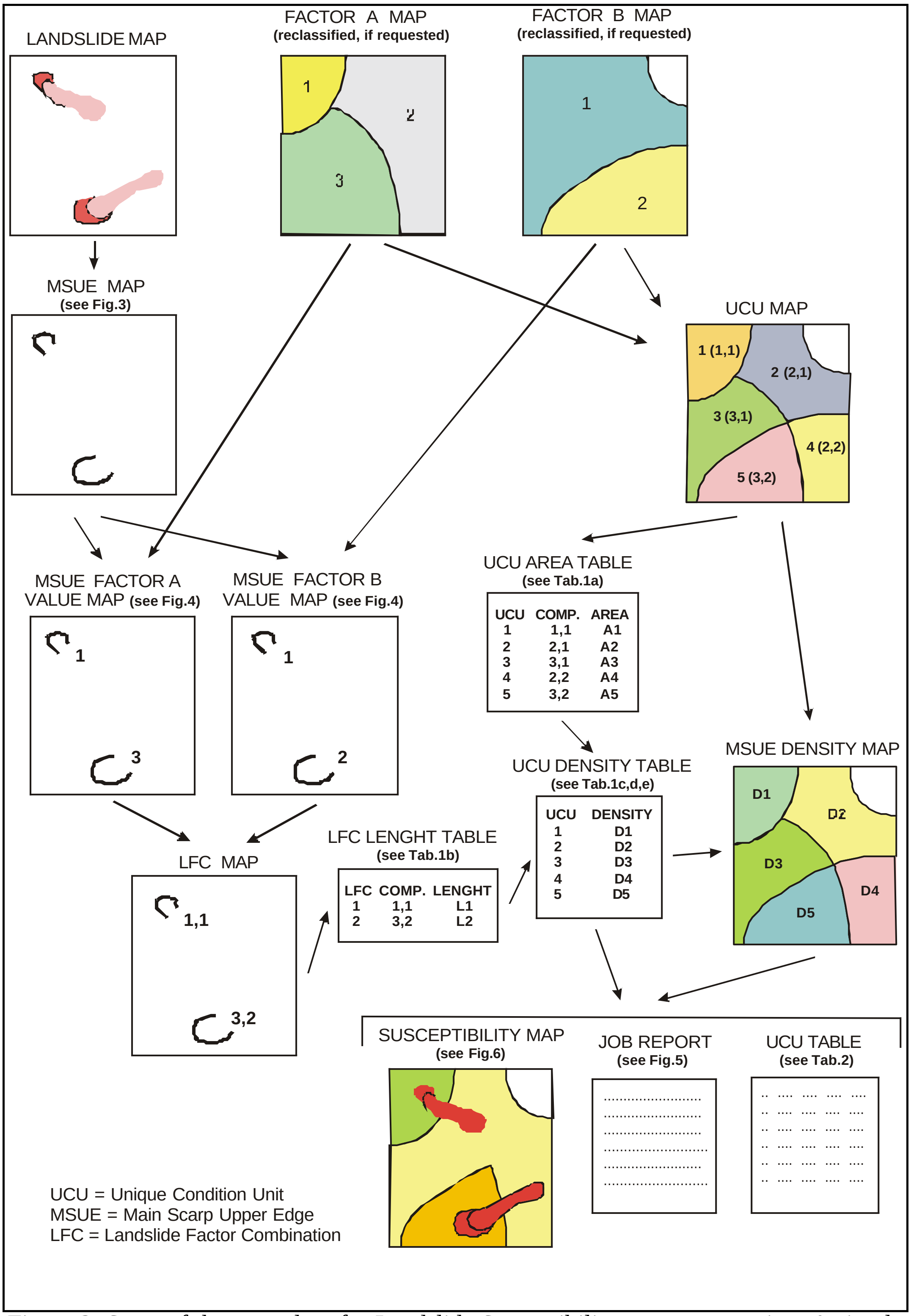

Figure 2: Steps of the procedure for Landslide Susceptibility map construction. A simple case with only two factor maps is considered. See text for a detailed explanation.

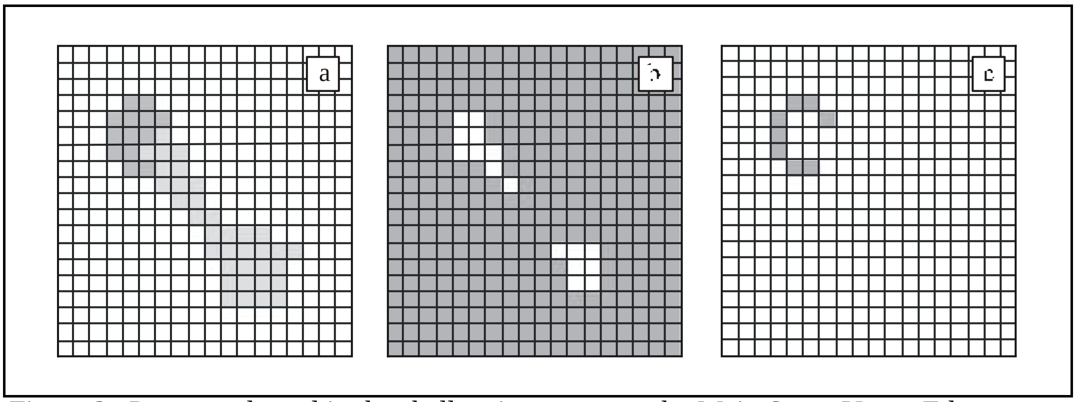

Figure 3: Process adopted in the shell script to extract the Main Scarp Upper Edge (MSUE) from a mapped landslide main scarp. In **a**, a landslide with different category values assigned to the main scarp area (dark gray) and to the accumulatione zone (light gray) is shown. The area external to landslide is grown by one cell along its margins, producing the effect of adding a cell belt along the inner landslide perimeter, as shown in **b**. The MSUE is defined by the eight cells common to the main scarp of map **a** and to the grown area of map **b**, as shown in **c**.

For a distance of two cells (figure 4d), the same criteria are used but the search is carried out on the 16 perimeter cells of a 5x5 moving matrix. Furthermore a map with the landslide areas grown by one cell along the perimeter is adopted to force the exclusion of all the cells immediately adjacent to the MSUE.

For a distance of three cells (figure 4e), a 7x7 matrix with 24 perimeter cells is used and the landslide areas are grown by two cells.

It is worth noting that in some circumstances the same factor cell is selected for more than one MSUE cell, as shown in figures 4d and 4e. Furthermore it can happen, especially if large distance values are adopted, that for some MSUE cells (or, in some cases, for all the cells of a whole MSUE) it is impossible to define the factor value as all the surrounding factor cells are located inside landslide areas or outside the study area. In this case no factor value is assigned to the MSUE cell.

For each factor, a map is thus constructed in which the proper factor value is assigned to each MSUE cell. By crossing these maps, the factor combination is defined for each MSUE cell and a LFC map created. According to the basic principle of the Conditional Analysis method, each LFC should represent the factor combination present in the cell at the failure time.

For the MSUE cells with any factor value not assigned, the LFC is not calculated and the cells are excluded from subsequent processing. Furthermore, if for any factor a distance greater than 0 is adopted, the resulting LFC could be different from any UCU detected in the area by crossing the factor maps. This simply means that the factor combination that caused the slope failure in the past does not exist at present in the investigated area. These cells are also excluded from the analysis. As these last operations can reduce the number of effective MSUE cells, or even delete entire MSUEs, the 'actual' MSUE total length, mean density, number and mean length are also computed and printed on the job report.

Then, as shown in table 1, a file with the MSUE length (in meters) is created for each LFC (table 1b) and joined to the previously created file with UCU extensions (table 1a). The resulting file (table 1c) reports, for each UCU, the identification number, the factor composition, the area and the total length of MSUEs pertaining to that specific UCU.

From these data the MSUE density for each UCU is computed (table 1d). By inserting the ' = ' sign and transforming the density from floating to integer values, a file is created (table 1e) that is in turn used for the reclassification of the original UCU map.

Each cell of the resulting reclassified map contains the MSUE density value expressed in m/km^2.

Then five class intervals are defined for density values. These intervals are designed in such a way that the mean density value represents the midpoint of the middle class (class

3) and the first four classes have the same interval. The upper limit of the fifth and last class is open.
In other terms, calling the mean density Md, the resulting class interval is Ci = Md x 2/5 and the ranges of the five classes are: 0 - Ci, Ci + 1- Ci x 2, Ci x 2 + 1- Ci x 3, Ci x 3 + 1- Ci x 4, > Ci x 4.

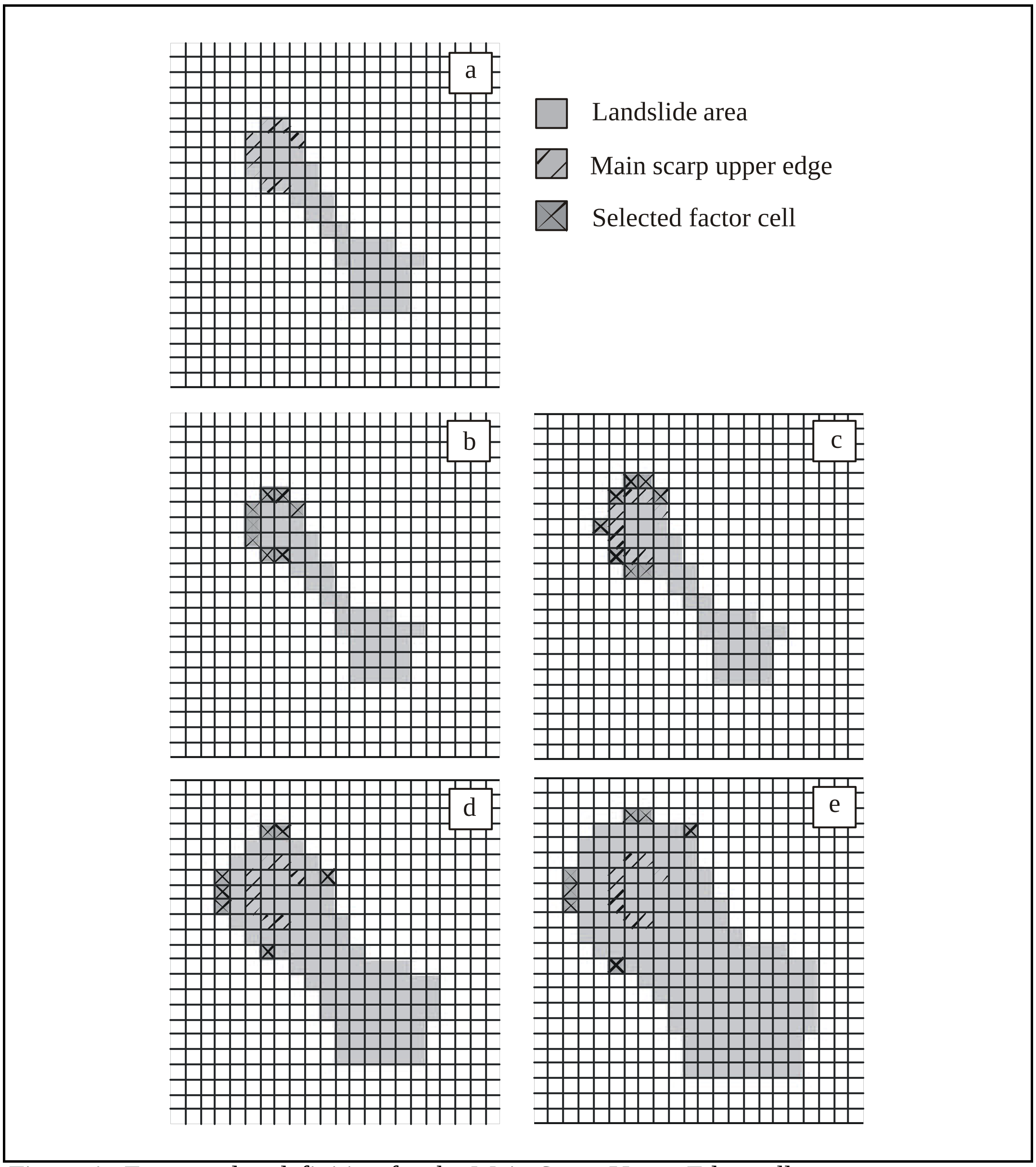

Figure 4: Factor value definition for the Main Scarp Upper Edge cells.
In **a**, a landslide with the main scarp cells highlighted is shown. If dstx = 0, the factor values are defined in coincidence with the MSUE cells, as shown in **b**. For distances greater than 0, a factor cell must be outside the landslide area in order to be selected. In **c**, with a search distance of one cell, a 3x3 cell moving matrix centred on the MSUE cells is used and the factor cells are selected along the external belt immediately adjacent to the MSUE. In **d,** with a distance of two cells**,** the moving matrix has a 5x5 cell size and the landslide area is grown of one cell along the perimeter. In **e,** a 7x7 matrix and a landslide area grown by two cells are used. When a moving matrix is used, like in **c, d** and **e,** the factor cell search proceeds clockwise beginning with the vertical and horizontal cells, then along the diagonal ones**.** Note that in **d** and **e**, the lowest factor cell is selected for both the lowest MSUE cells.

a	b	c	d	e
1 1;1;1 0.1459	1 2;1;1 554.277	1;1;1 1 0.1459	1 0	1=0
2 1;1;2 0.0027	2 2;1;5 174.853	1;1;2 2 0.0027	2 0	2=0
3 1;1;5 0.0548	3 2;2;1 873.939	1;1;5 3 0.0548	3 0	3=0
4 1;2;1 0.5129	4 2;2;2 22.0711	1;2;1 4 0.5129	4 0	4=0
5 1;2;2 0.0113	5 2;2;5 578.523	1;2;2 5 0.0113	5 0	5=0
6 1;2;5 0.1635	6 2;3;1 1264.93	1;2;5 6 0.1635	6 0	6=0
7 1;3;1 0.5195	7 2;3;2 11.0355	1;3;1 7 0.5195	7 0	7=0
8 1;3;2 0.0054	8 2;3;5 771.378	1;3;2 8 0.0054	8 0	8=0
9 1;3;5 0.1468	9 2;4;1 1539.28	1;3;5 9 0.1468	9 0	9=0
10 1;4;1 0.3298	10 2;4;2 15	1;4;1 10 0.3298	10 0	10=0
11 1;4;2 0.0011	11 2;4;5 581.452	1;4;2 11 0.0011	11 0	11=0
12 1;4;5 0.0860	12 2;5;1 1071.37	1;4;5 12 0.0860	12 0	12=0
13 1;4;6 5e-05	13 2;5;2 31.0355	1;4;6 13 5e-05	13 0	13=0
14 1;5;1 0.1560	....	1;5;1 14 0.1560	14 0	14=0
15 1;5;2 0.0001		1;5;2 15 0.0001	15 0	15=0
16 1;5;5 0.0545		1;5;5 16 0.0545	16 0	16=0
17 1;6;1 0.0782		1;6;1 17 0.0782	17 0	17=0
18 1;6;5 0.0518		1;6;5 18 0.0518	18 0	18=0
19 2;1;1 0.3582		2;1;1 19 0.3582 1 554.277	19 1547.29	19=1547
20 2;1;2 0.0534		2;1;2 20 0.0534	20 0	20=0
21 2;1;4 0.0016		2;1;4 21 0.0016	21 0	21=0
22 2;1;5 0.421		2;1;5 22 0.421 2 174.853	22 415.328	22=415
23 2;1;6 0.0007		2;1;6 23 0.0007	23 0	23=0
24 2;2;1 0.9772		2;2;1 24 0.9772 3 873.939	24 894.307	24=894
25 2;2;2 0.0989		2;2;2 25 0.0989 4 22.0711	25 223.166	25=223
26 2;2;4 0.0019		2;2;4 26 0.0019	26 0	26=0
27 2;2;5 0.7983		2;2;5 27 0.7983 5 578.523	27 724.648	27=725
28 2;2;6 0.0025		2;2;6 28 0.0025	28 0	28=0
29 2;3;1 1.4252		2;3;1 29 1.4252 6 1264.93	29 887.514	29=888
30 2;3;2 0.1067		2;3;2 30 0.1067 7 11.0355	30 103.377	30=103
31 2;3;4 0.0010		2;3;4 31 0.0010	31 0	31=0
32 2;3;5 0.7311		2;3;5 32 0.7311 8 771.378	32 1055.06	32=1055
33 2;3;6 0.0027		2;3;6 33 0.0027	33 0	33=0
34 2;4;1 1.3132		2;4;1 34 1.3132 9 1539.28	34 1172.19	34=1172
35 2;4;2 0.0726		2;4;2 35 0.0726 10 15	35 206.612	35=207
36 2;4;4 0.0005		2;4;4 36 0.0005	36 0	36=0
37 2;4;5 0.4846		2;4;5 37 0.4846 11 581.452	37 1199.92	37=1200
38 2;4;6 0.0032		2;4;6 38 0.0032	38 0	38=0
39 2;5;1 1.2216		2;5;1 39 1.2216 12 1071.37	39 877.008	39=877
40 2;5;2 0.0589		2;5;2 40 0.0589 13 31.0355	40 526.472	40=526
....		....	....	....

Table 1: Files used in the landslide density calculation process.
File **a** is a report of the Unique Condition Unit (UCU) map, listing, for each UCU, the identification number, the factor class combination and area (in km^2). File **b** reports the characteristics of the Landslide Factor Combination (LFC) map; more precisely, for each LFC, the identification number, the factor class combination and the Main Scarp Upper Edge (MSUE) length are listed. File **c** is obtained by combining these two files on the basis of the factor combinations common to both; it contains, for each factor combination, the UCU area and the MSUE length. Note that only 13 of the 40 listed UCUs contain MSUEs. By dividing the MSUE length by the UCU area, the landslide density values are computed and stored in a file together with the UCU identification numbers (file **d**). This last file is then transformed into a file (file **e**) to be used for the UCU map reclassification.The files reported in the table are parts of the files produced in the example of the shell script application to the Bratica T. basin.

With the values thus calculated, the previously constructed density map is reclassified, leading to the construction of the final Landslide Susceptibility map to which the 'undefined' areas, assigned class 6, and landslides, class 7, are then superimposed. As already mentioned, the characteristics of this map are printed as a table in the job report (see figure 5). The table contains the five susceptibility class intervals expressed both as density (m/km^2) and as frequency (number of landslides/km^2). The latter values are obtained simply by dividing the density class values by the MSUE mean length. For each of the five classes, for the 'undefined' areas (if any) and for landslides, the extension, both in km^2 and in percent, is also reported.

3.3. Landslide susceptibility map construction

The last part of the script deals with the drawing up of the final Landslide Susceptibility map showing both landslide susceptibility zonation and current landslides. The map is reproduced on the screen and, optionally, a PostScript file is written (see figure 6 as an example). In the latter case, if the width of the map is greater than its height, the drawing is rotated by 90°. If the scale defined in the command line is too big to contain the drawing, it is automatically reduced. The map contains a graphic scale, a north arrow and a legend with the five susceptibility classes. A sixth class is added if undefined areas are present. The units of the graphic scale are automatically defined on the basis of the study area size and may be metres or kilometres. In defining the working region before running the script it is necessary to leave space for the map title at the top and for the legend and graphic scale on the left. The former is 1/6 and the latter approx. 1/20 of the region width. These dimensions can be reduced depending on the shape of the study area.
The script provides for the tracing of the perimeters of the investigated area and of landslides, two vector covers automatically extracted from their relative raster images. It is naturally possible to represent other vector covers, such as streams, roads or contour lines, by introducing the relevant commands in the graphics intructions. Clearly, the corresponding files must be present in the GRASS work directory.
In the script many comments describe various parts of the drawing in order to make it easier to modify map features like font and size for characters, thickness and colours for vector images or colours for raster categories.

4 An example of an application

By way of example, an application of the shell script to a relatively small area is presented. The area, which extension is 31.67 km^2, corresponds to the basin of Bratica torrent, a tributary of the Parma river, in the Northern Italian Apennines about 50 km south-west of Parma town. In order to limit the number of files involved and reduce running time, only three factors, namely geology, slope and land use, were introduced into the analysis. All the data used in the example were extracted from a larger database including the entire basin of the Parma river, for an extension of 332 km^2 [14], and can be downloaded, together with the shell script, from ftp.geo.unipr.it in pub/condan3/.
A 5 m cell side was used and the entire work region includes 2,948,000 cells of which 1,266,924 inside the Bratica T. basin.
As regards the running time, in general it obviously depends on the area extension, cell size, the number of factors and computer used. For this application the running time was less than 5 minutes with a Pentium III 1000MHz with a RAM of 512MB.
In the application the following command line was used:

```
condan3.sh title="LANDSLIDE SUSCEPTIBILITY MAP OF BRATICA T. BASIN"
landsl=br.slides  warea=br.divide  msc=1  fact1=br.geology  refact1=br.geology.rec
fact2=br.slope  refact2=br.slope.rec dst2=3  fact3=br.landuse  dst3=1
report=bratica.report ucutable=bratica.table  psmapname=bratica.ps  psmapsize=a4
psmapscale=1:42000
```

In the following sections, a short description of input and output data is given.

4.1 Input data

Landslide inventory map (br.slides)

This map was created by interpreting aerial photographs of 1996 at a scale of about 1:33000 and field checks. Landslides were mapped at 1:10000 scale and classified according to Varnes [15] and Cruden and Varnes [16], i.e. falls, translational slides,

rotational slides, flows and complex landslides. The maps were subsequently acquired in raster format via a scanner, imported into GRASS, rectified and georeferenced. Digitization of the boundaries of each element of the landslides, i.e. main scarp and accumulation zone, was carried out directly on the screen. A numeric code of eight digits was assigned to each landslide element. The first identifies the kind of element (scarp or accumulation zone), the second the typology (five types), the third the state of activity (relict, inactive and active), while the remaining five digits are the progressive number of the element. Subsequently the vector file was transformed into raster format. To accord the map with the script requirements, in a pre-processing operation the landslides were reclassified assigning category 1 to the main scarps, as specified in the argument msc, and category 2 to accumulation areas.
As the application here described only serves as an illustration, all the landslides present in the area were considered, regardless of typology and state of activity.

Study area boundary (br.divide)
The polygon delimiting Bratica T. basin was digitized directly on the screen from the rasterized and georeferenced topographical map. Category 1 was assigned to the area enclosed in this polygon and the vector file rasterized.

Geological map (br.geology)
In the Geological map of the basin 7 types of detrital cover and 16 lithostratigraphic units are present.
For this factor a reclassification file (br.geology.rec) was introduced in order to reduce the 23 categories to 6 lithological classes with similar or sufficiently homogeneous general lithological features defined on the basis of the prevailing rock composition (clay, sandstone, limestone, etc.), as well as structural characteristics (bedding, pervasive fissility, etc.).
The original class defining the main stream deposits was omitted in the reclassification file. Therefore the pertaining area is not considered in the procedure and classed as 'undefined' in the final map.
As the lithological setting is not modified by the landslide process, this factor can be detected in corrispondence with the MSUE cells. So the dst1 parameter was not specified and the 0 default value assumed.

Slope map (br.slope)
In the area, detailed topographic maps at 1:10000 scale were scanned, rectified and georeferenced. From these maps, spot heights and contours, with contour intervals ranging from 5 to 25m, were digitized directly on the screen and used to produce a Digital Elevation Model (DEM) of the entire basin. The Slope map produced directly from the DEM contains values ranging from 0° to 85° .
An analysis of slope angle frequency reveals a high concentration of cells in the interval from 10° to 30° and a very low number of cells over 30°. In order to avoid the presence of small class areas, the five intervals defined by the following riclassification file (br.slope.rec) were adopted:
0 thru 9 = 1
10 thru 14 = 2
15 thru 19 = 3
20 thru 24 = 4
25 thru 29 = 5
30 thru 90 = 6
Thus, even the class with the smallest extension (class 5) occupies a significant portion of the area (11%).
As the slope angle is computed by a neighbouring method, a distance of three cells (dst2 = 3) was adopted to be sure that no main scarp cell was involved in calculation.

Land use (br.landuse)
From the topographic map at 1:10000 scale, the boundaries of six landuse categories were digitized in the Bratica T. basin.
No reclassification is carried out for this factor; therefore, the relative argument is omitted in the command line.
For this factor a distance of one cell is adopted (dst3 = 1). So the land use class is defined at the cells immediately adjacent to the MSUEs.

4.2 Output data
Job report (bratica.report)
The resulting report is reproduced in figure 5. As previously specified, in the report the input factor maps are listed together with the corresponding reclassification files and the distances for factor evaluation. The characteristics of the PostScript map are also described. The number of potential UCUs in the study area, i.e. the possibile combinations of the different classes of the three factors, is 216 (6x6x6). Of these, only 167 are effectively present in the area.
It should be noted that the 133 landslide main scarps present in the area are reduced to 123 after the LFC definition; 10 are not considered in the analysis as some of the factors are not detectable in the neighbourhood of the MSUE cells. This is due mainly to the factor 'slope' for which a dst = 3 was adopted. Consequently, MSUE length and density also diminish while the mean length increases.
From the table of susceptibility classes, it results that 20.7% of the area is covered by landslides and about 17 % shows high and very high susceptibility. 1.3% of the area classed as 'undefined' corresponds to the area of the not reclassified class (main stream deposits) of the Geological map.

UCU table (bratica.table)
In table 2, the first part of the output table listing the composition of the 167 UCUs present in the area is reproduced. The first column contains the UCU composition, with the factors ordered according to the order number in the input command line, i.e. geology, slope and land use. A number of UCUs are very small, in some cases only a handful of cells. These combinations are clearly of little significance and should be omitted from an eventual statistical analysis.

Landslide susceptibility map (bratica.ps)
The final map, reported in figure 6, was produced in Postcript format on an A4 size paper at 1:42000 scale, as specified by the psmapsize and psmapscale arguments in the command line.
As in the reclassification of the Geological map a category was omitted, undefined areas appear on the map and the 'undefined' class figures in the legend**.**

5 Conclusions

The Conditional Analysis method applied to the Unique Condition Units is a conceptually simple, but operationally complex, technique for performing Landslide Susceptibility zonation. Accepting the assumption that future landslides are more likely to occur under conditions (of factor combinations) that led to landslides in the past, the concept of landslide susceptibility expressed as landslide density in correspondence with different combinations of the instability factors is simple and direct. But the technique requires a long and strict sequence of operations and can be prone to errors. Furthermore, in the previous use of this technique, some aspects, like the definition of landslide dimensions and the assessment of factor values, turned out to be an extremely weak point in the method. For these reasons, a shell script, mainly based on the GRASS GIS commands, was created in which the upper edge of the landslide main scarp is considered as the morphological element representing the landslide and is used to define both the

```
thu sep 20 06:58:49 UTC 2001

Map title:  LANDSLIDE SUSCEPTIBILITY MAP OF BRATICA T. BASIN

Input maps and files
landsl (landslides map):  br.slides
warea (study area boundary):  br.divide
msc (category value for main scarps):  1
factor1:  br.geology
refact1 (reclassification file):  br.geology.rec
factor2:  br.slope
refact2 (reclassification file):  br.slope.rec
dst2:  3
factor3:  br.landuse
dst3:  1

PostScript map characteristics
map size :  a4
map name :  bratica.ps
map scale :  1:42000

File with job report:  bratica.report
File with UCU composition and statistics:  bratica.table

** A total of  3  factors are considered in the analysis

Region limits and cell size
north: 927000 south: 916000 east: 590700 west: 584000
e-w cell size: 5  n-s cell size: 5
cell area (sq meters): 25

Extension of the study area (sq km) =  31.6731

Theoretic number of UCU = 216
Actual number of UCU =  167
Extension of the UCUs area (sq km) =  31.2611

MSUE total lenght (m) =  37771.6
MSUE overall mean density (m/sqkm) =  1208.26
MSUE number =  133
MSUE mean lenght (m) =  283.997

Actual MSUE total lenght (m) =  36341.8
Actual MSUE mean density (m/sqkm) =  1162.52
Actual MSUE number =  123
Actual MSUE mean lenght (m) =  295.462

Class limits and extension
---------------------------------------------------------------------------
|CLASS|    CLASS  |            CLASS VALUE                |    EXTENSION    |
|NUMB.|    NAME   |    (m/sqkm)      |     (N/sqkm)       |  (sq km) |   %  |
|-------------------------------------------------------------------------|
|  1  |  VERY LOW |       0 -    465 |     0.0 -    1.6 |     3.78 |  11.9 |
|  2  |       LOW |     466 -    930 |     1.6 -    3.1 |     5.84 |  18.4 |
|  3  |    MEDIUM |     931 -   1395 |     3.1 -    4.7 |     9.73 |  30.7 |
|  4  |      HIGH |    1396 -   1860 |     4.7 -    6.3 |     3.39 |  10.7 |
|  5  | VERY HIGH |          >   1860 |          >    6.3 |     1.96 |   6.2 |
|-------------------------------------------------------|          |       |
|                                     UNDEFINED AREAS  |     0.41 |   1.3 |
|                                          LANDSLIDES  |     6.56 |  20.7 |
|-------------------------------------------------------------------------|
|                                               TOTAL  |    31.67 | 100.0 |
---------------------------------------------------------------------------

thu sep 20 07:03:09 UTC 2001

****************** E N D ******************
```

Figure 5: Job report resulting from the procedure as applied to the Bratica T. basin.

Unique Cond. Units (factor combinat.)	Numb.of cells	Area (sqkm)	MSUE lenght (m)	MSUE number (N)	MSUE freq. (N/sqkm)	MSUE dens. (m/sq km)
1 1 1	5838	0.1459	0.00	0.00	0.00	0.00
1 1 2	109	0.0027	0.00	0.00	0.00	0.00
1 1 5	2192	0.0548	0.00	0.00	0.00	0.00
1 2 1	20517	0.5129	0.00	0.00	0.00	0.00
1 2 2	453	0.0113	0.00	0.00	0.00	0.00
1 2 5	6539	0.1635	0.00	0.00	0.00	0.00
1 3 1	20781	0.5195	0.00	0.00	0.00	0.00
1 3 2	218	0.0054	0.00	0.00	0.00	0.00
1 3 5	5873	0.1468	0.00	0.00	0.00	0.00
1 4 1	13191	0.3298	0.00	0.00	0.00	0.00
1 4 2	46	0.0011	0.00	0.00	0.00	0.00
1 4 5	3439	0.0860	0.00	0.00	0.00	0.00
1 4 6	2	0.0001	0.00	0.00	0.00	0.00
1 5 1	6242	0.1560	0.00	0.00	0.00	0.00
1 5 2	5	0.0001	0.00	0.00	0.00	0.00
1 5 5	2181	0.0545	0.00	0.00	0.00	0.00
1 6 1	3128	0.0782	0.00	0.00	0.00	0.00
1 6 5	2074	0.0519	0.00	0.00	0.00	0.00
2 1 1	14329	0.3582	554.28	1.95	5.45	1547.29
2 1 2	2136	0.0534	0.00	0.00	0.00	0.00
2 1 4	65	0.0016	0.00	0.00	0.00	0.00
2 1 5	16840	0.4210	174.85	0.62	1.46	415.33
2 1 6	27	0.0007	0.00	0.00	0.00	0.00
2 2 1	39089	0.9772	873.94	3.08	3.15	894.31
2 2 2	3956	0.0989	22.07	0.08	0.79	223.17
2 2 4	78	0.0019	0.00	0.00	0.00	0.00
2 2 5	31934	0.7984	578.52	2.04	2.55	724.65
2 2 6	101	0.0025	0.00	0.00	0.00	0.00
2 3 1	57010	1.4252	1264.93	4.45	3.13	887.51
2 3 2	4270	0.1067	11.04	0.04	0.36	103.38
2 3 4	42	0.0010	0.00	0.00	0.00	0.00
2 3 5	29245	0.7311	771.38	2.72	3.72	1055.06
2 3 6	110	0.0027	0.00	0.00	0.00	0.00
2 4 1	52527	1.3132	1539.28	5.42	4.13	1172.19
2 4 2	2904	0.0726	15.00	0.05	0.73	206.61
2 4 4	19	0.0005	0.00	0.00	0.00	0.00
2 4 5	19383	0.4846	581.45	2.05	4.23	1199.92
...	...	...	...	...	...	...

Table 2: Partial reproduction of the table containing the characteristics of the Unique Condition Units (UCU) in the Bratica T. basin.

factor combination and landslide extension. The script, starting from a Landslide Inventory map and a number of factor maps, automatically carries out the whole procedure resulting in the construction of a Landslide Susceptibility map. All these characteristics make the procedure easy, fast and inexpensive, and therefore accessible to a wide range of users, making Landslide Susceptibility zonation on a regional scale with high spatial resolution possible in a very short time, as is illustrated by its application to the Bratica T. basin.

Acknowledgements

This study was supported partly by funds of the Parma University reasearch project "Analisi della evoluzione gravitativa dei versanti nell'appennino parmense mediante un Sistema Informativo Territoriale" (Coordinator S. Perego), and partly by the "Cofinanziamento MURST 1999. Programma di ricerca nazionale: I movimenti franosi come indicatori di variazioni climatiche dal Tradiglaciale ad oggi. (National coordinator M. Panizza). Programma di ricerca locale: I movimenti franosi come indicatori dei cambiamenti climatici: esempi nell'Appennino emiliano e nelle Alpi centrali. (Local coordinator C. Tellini)".

Thanks to D. Peis for assistence in solving computer problems and P. Sears for help in translation.

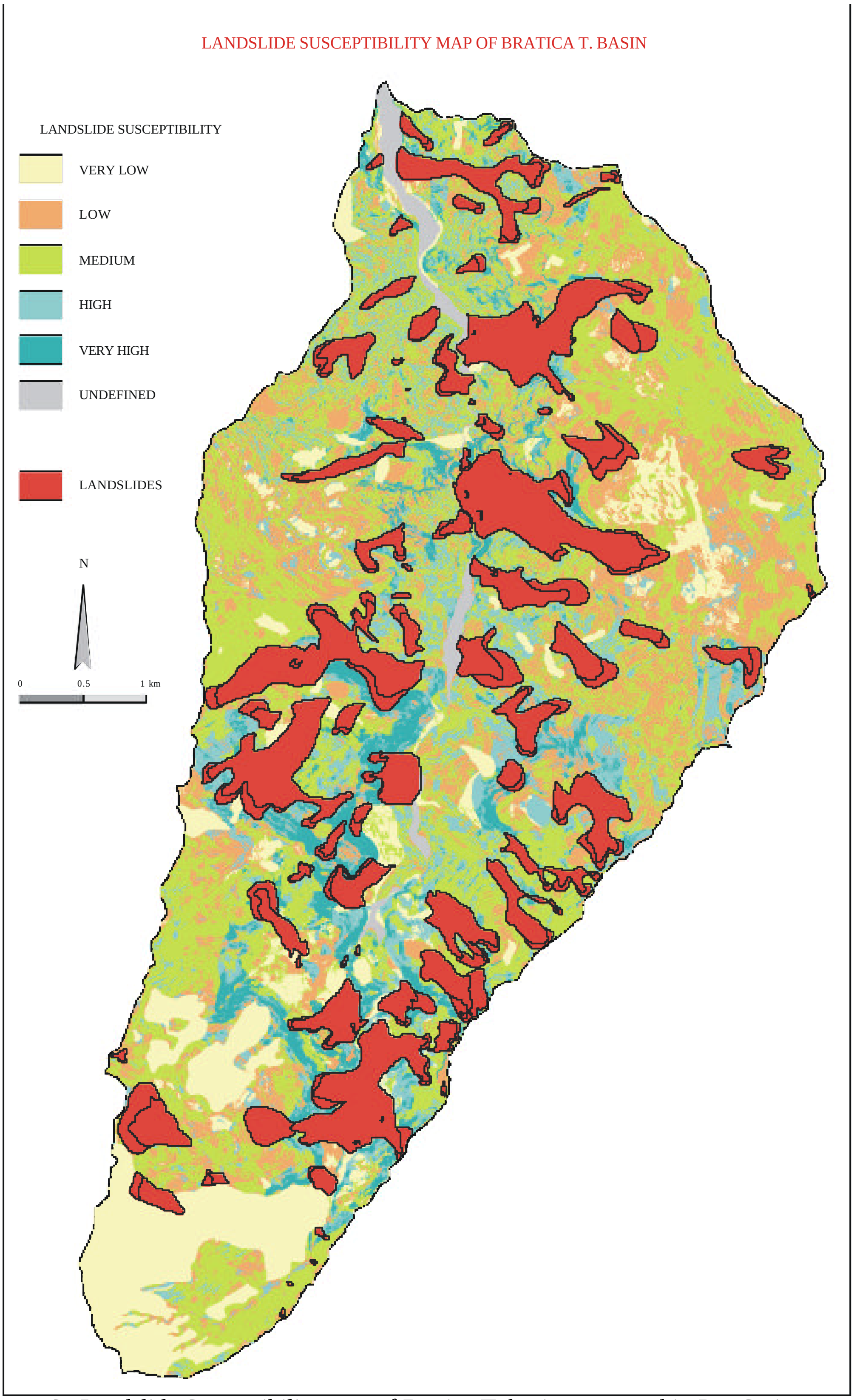

Figure 6: Landslide Susceptibility map of Bratica T. basin generated in PostScript format.

References

[1] Carrara A., Cardinali M., Guzzetti F., Reichenbach P. GIS-based techniques for mapping landslide hazard. (http://deis158.deis.unibo.it), 1995.
[2] Soeters R., van Westen C.J. Slope instability recognition, analysis, and zonation. In Turner A.K. and Schuster R.L., editors, *Landslides: investigation and mitigation.* Transportation Research Board Special Report No. 247, pages129-177, 1996.
[3] Guzzetti F., Carrara A., Cardinali M., Reichenbach P. Lanslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology, 31, pages 181-216, 1999.
[4] Chung Ch., Fabbri A.G., van Westen C.J. Multivariate regression analysis for landslide hazard zonation. In Carrara A. and Guzzetti F., editors, *Geographical Information Systems in Assessing Natural Hazards* (Dordrecht, The Netherlands: Kluwer), pages 107-142, 1995.
[5] Brabb E.E. Innovative approaches to Landslide hazard and risk mapping. In *Proceedings of Fourth International Symposium on Landslides* (Toronto, Canada: Canadian Geotechnical Society), 1, pages 307-324, 1984.
[6] Clerici A., Perego S., Tellini C., Vescovi P. A procedure for landslide susceptibility zonation by the conditional analysis method. Geomorphology (in press), 2002.
[7] Ramey C., Fox B. Bash Reference Manual. (Boston, MA: Free Software Foundation), (http://www.gnu.org/manual/bash/html_mono/bashref.html), 2001.
[8] Cooper M. Advanced Bash-Scripting Guide. An in-depth exploration of the art of shell scripting (Revision 1.3). (http://www.linuxdoc.org/LDP/abs/html/), 2002.
[9] GRASS Development Team. Geographic Resources Analysis and Support System - GRASS. (Waco, Texas: Baylor University, http://www3.baylor.edu/grass/index2.html), (Trento, Italy: Istituto Trentino di Cultura, http://grass.itc.it/index2.html), 1999.
[10] Close D., Robbins A.D., Rubin P.H., Stallman R. The GAWK manual, 0.15 edn.. (Cambridge, MA: Free Software Foundation). (http://www.cco.caltech.edu/cco/texinfo/gawk/gawk_toc.html),1993.
[11] Robbins A.D. Effective AWK Programming. A User's Guide for GNU Awk. Edn. 1.0.3, (Boston, MA: Free Software Foundation), (http://www.gnu.org/manual/gawk/), 1997.
[12] Duarte C. Do It With Sed. (http://www.dbnet.ece.ntua.gr/~george/sed/sedtut_1.html), 1996.
[13] Pizzini K., sed, a stream editor, version 3.02, (Boston, MA: Free Software Foundation), (http://www.gnu.org/manual/sed-3.02/sed.html), 1998.
[14] Clerici A., Perego S., Tellini C., Vescovi P. Il Sistema Informativo Territoriale della Val Parma: modalità costruttive, caratteristiche ed esempi di utilizzo. Geologia Tecnica & Ambientale, 2/02, pages 21-40, 2002.
[15] Varnes D.J. Slope Movement Types and Processes. In Schuster R.L. and Krizek R.J., editors, *Landslides: Analysis and Control.* Transportation Research Board Special Report No. 176, pages 11-33, 1978.
[16] Cruden D.M., Varnes D.J. Landslide Types and Processes. In Turner A.K. and Schuster R.L., editors, *Landslides: investigation and mitigation*, Transportation Research Board Special Report No. 247, pages 36-75, 1996.