

Spot heights generalization: deriving the relief of the Topographic Database of Catalonia at 1:25,000 from the master database

Blanca Baella *, Jesús Palomar-Vázquez **, Josep E. Pardo-Pascual **, Maria Pla *
blanca.baella@icc.cat , jepardo@cgf.upv.es, jpalomav@esp.upv.es, maria.pla@icc.cat

* Institut Cartogràfic de Catalunya, Barcelona, www.icc.cat

** Research Group Geoenvironmental Cartography and Remote Sensing, Department of Cartographic Engineering, Geodesy and Photogrammetry, Universitat Politècnica de València, cgat.webs.upv.es

Abstract

The paper will explain the implementation of the generalization application to select spot heights, developed by the Escola Tècnica Superior d'Enginyeria Geodèsica, Cartogràfica i Topogràfica of the Universitat Politècnica de València (UPV), in the production workflow on the Institut Cartogràfic de Catalunya (ICC) to derive the Topographic Database of Catalonia at scale 1:25.000 from the master database at 1:5.000.

Keywords: Automatic generalization, relief, production workflow, geomorphologic classification, peak, pass, depression, spot height selection.

1.- Introduction

Since 1997, the Institut Cartogràfic de Catalunya (ICC) has been applying automatic generalization techniques in the production workflows of the vector topographic databases and maps, but these techniques never has been used on the objects that characterize the relief. The main reason is that, although there are some research works on this topic (Weibel and Hellerr, 1991), the available commercial software has not offered, until now, any tool specifically designed for relief generalization.

From 2002 the Research Group Geoenvironmental Cartography and Remote Sensing from the Department of Cartographic Engineering, Geodesy and Photogrammetry of the of the Universitat Politècnica de València (UPV) has developed an application, called GENCOTES¹, to generalize spot heights, mainly based on the analysis of the terrain morphology (Palomar and Pardo, 2002; Palomar, 2004; Palomar and Pardo, 2004). Moreover, this application took into account cartographic aspects as significant points or map names.

In 2006 ICC and UPV signed a collaboration agreement to adapt GENCOTES to the requirements defined in the ICC specifications and applied in the cartographic production environment. An additional improvement has been achieved during the project, the geomorphologic analysis of the terrain has allowed the enrichment of the original data with the classification of the spot heights. The

¹ GENCOTES is the acronym of the *generalització de cotes* that means spot heights generalization.

paper will explain in detail the algorithms behind the application and the improvements introduced in the adaptation to the ICC requirements, as well as, the implementation process on the production workflows. Finally, it will be given some aspects of further developments.

2.- The project

The ICC topographic data models represent the relief mainly using contour lines and spot heights. Moreover, the master database at scale 1:5,000, contains additional information to model more accurately the terrain as break lines and contour lines to infer break lines. The spot heights generalization application can be used on the most of the ICC production workflows for deriving large and medium scales vector topographic databases and maps.

The ICC decided starting the implementation of this tool on the generalization workflow to obtain the Topographic Database of Catalonia at 1:25,000 scale from the master database, the Topographic Database of Catalonia at 1:5,000 (Baella and Pla, 2003). Up till then, the generalization operations applied on contour lines have been automatic selection and point filtering, followed by manual revision and editing of the conflicts, and in the case of the spot heights, the generalization has been done totally by manual processes. The main reasons to begin with this production workflow were that at 1:25,000 scale remains yet a significant number of spot heights and that the cost of the manual generalization is quite high.

3.- UPV application: Generalization process of the spot heights (GENCOTES)

The application GENCOTES is a set of informatics tools written in AVENUE language that works on ArcView 3.2. This program automatically generalizes the spot heights following two main steps: hierarchical classification of the spot heights, to give a specific value to each spot height based on parametric criteria, and selection of the spot heights taking into account the hierarchical classification and a balanced distribution of these points on the final data.

3.1.- Hierarchical classification

The criteria for the hierarchical classification were deduced by analyzing the needs of the user: how are spot heights useful to a user, what sort of precision should they have, where should they appear, etc. The answers to these needs were that the height points would be more useful in the areas close to the roads, at characteristic points of the terrain and in reference elements that help in the orientation. Moreover, the density of height points should be slightly higher in elevated areas than in flat areas, since the users are generally more interested in knowing the heights of the peaks rather than the intermediate or low areas.

In terms of hierarchy, three kinds of spots heights are considered:

- a) Special interest points. They are defined by the data producer specifications as significant, and must remain in the generalized data (on the roads, related with map names, etc.).
- b) Geomorphologic interesting points. They identify characteristic terrain points, as peaks, passes or depressions.

- c) Remainder points. Into this category there are included all the other spot heights that have not been considered in the previous categories.

3.1.1.- Categorization of the special interest points

The areas for categorize the special interest points are defined as influence area of roads, inhabited zones and map names. To apply the proximity criteria, the application uses a combination of operations to determine which spot heights are found close to elements of interest, such as a buffer operator, an area of influence or the topological relationship of inclusion. Once these points have been localized, a weight is assigned to each spot height according to the priority, defined by the cartographer, for the elements used to define the areas.

In the case of the map names, GENCOTES has a specific tool that allows to the operators to create a database with all the possible generic names used for geographical landforms (peak, rock, crest, ridge, summit, mountain, etc.) in the language of the cartography. For each map name in the input data, the program searches in the database if there is any generic name partially coincident with it. Then, the application searches, buffering around the map name position, all the spot heights placed around it and selects the closest to the map name.

3.1.2.- Categorization of geomorphologic interesting points

In this phase, the points that are not categorized in the previous phase are analyzed to determine which ones are geomorphologic singular, that is, peaks, mountain passes and depressions. Their magnitude is then analyzed and, finally, some operations are applied to the peaks, which are by far the most abundant, to categorize and classify them according to their shape: conic, asymmetrical (escarpments), open or closed. This categorization will be later used to make a selection, based on objective criteria, of the points that should remain.

The geomorphologic analysis requires an accurate digital elevation model (DEM), if possible generated using the same data to be generalized, as well as other significant altimetry data as break lines. With these elements a DEM in vector format is generated obtaining a triangulation model (TIN). Theoretically, also a DEM in raster format can be used, but it adds some difficulties to this methodology. One is that the centre of a pixel in raster data can not be coincident with the exact position of the spot height and the location of the pixel containing the spot height is unclear. Moreover, in raster format, each pixel of the DEM must be analyzed, while when using a DEM in TIN format, the analysis can be specifically restricted to the spot heights. Therefore, the use of a DEM in TIN format reduces substantially the computation time.

For the ICC data, the TIN model has been created using a standard TIN algorithm included in the ArcView software. The input data used is breaklines (hydrographic network, talwegs, divides and other strong slope lines), spot heights, flat polygons used to represent a flat surface (reservoir areas or lakes) and scan lines. The contour lines have not been used because they are derived automatically from the ICC TIN model, so, they are redundant data.

Once then TIN model has been created, the spot heights that are geomorphologic singular should be determined. Three basic types are differentiated: peaks (summit points), depression points and passes:

- Peaks

- According to the height in a neighborhood. For a given spot height, a fixed circular surrounding area is selected and analyzed. If the point has the maximum height value in the area, then it is considered the peak of that area. The radius of the surrounding area is enlarged several times repeating the analysis for the same point until a higher spot height is found. The peak category is assigned according to the size of the area of which the point is considered peak (Figure 1).

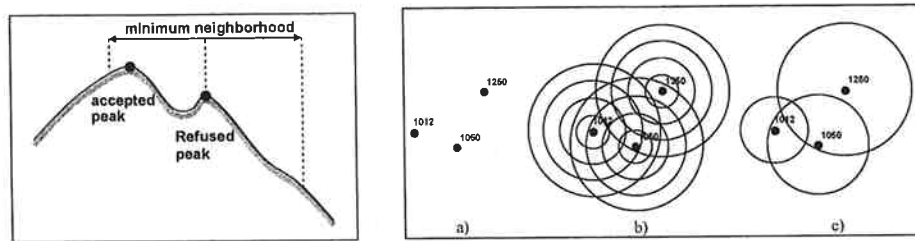


Figure 1. Analysis by surrounding areas.

- According to the shape. For a given peak, a non-circular neighborhood is analyzed. This neighborhood area is defined by the contour of the lowest height that includes only the peak to be analyzed. This curve, called the *Base Contour Line (BCL)*, is defined as the lowest contour line that includes only one peak. Taking it into account, the relationship between the peak and the BCL is studied in order to determine different types of possible shapes, such as conical or asymmetrical shapes respect to the peak, and open or closed shapes according to the ratio between the difference in height between the peak and the BCL.

For topographic data, it is important to keep the points that are located on more abrupt terrain, called critical shapes, as they are the most perceptible in the area. The following parameters are used to detect these critical shapes:

1. Proximity to the edge of the contour: It measures how centered the point is with respect to the curve. A high value shows a tendency towards asymmetrical cliff-like shapes, or escarpments, while a low value detects a tendency towards conical shapes (Figure 2).

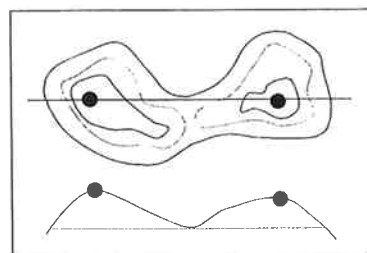


Figure 2. Selection of the analysis curves (red).

2. Average slope: It is the average of the slopes of the straight lines that connect the peak to each of the points of the curve. This parameter denotes the more open or closed peaks depending on whether the values are larger or smaller (Figure 3).

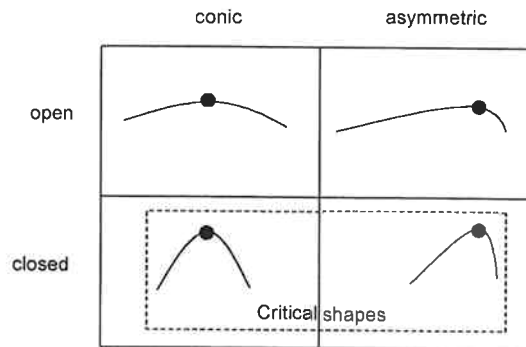


Figure 3. Basic shapes.

The combination of these magnitudes gives information about the type of the peak. In practice, all of the peaks that have critical shapes are placed at the same categorical level as those peaks classified in the maximum category using the analysis of neighborhoods.

- Depressions

In this case, the interest was focused on large depressions, so the same process used in the analysis by neighborhoods for the calculation of peaks can be applied. The surrounding area used is sufficiently large to avoid micro-depressions, and every point is analyzed to determine if it is the one with the minimum height of its area.

- Passes

The location was deduced by the analysis of circular areas or neighborhoods (Figure 4), by studying the relative maxima and minima of the topographic profile resulting from the analyzed circle. It is determined whether or not the selected point is a pass (2 maximums and 2 minimums), applying simplification techniques on the topographical profile to extract the main tendencies (Figure 5).

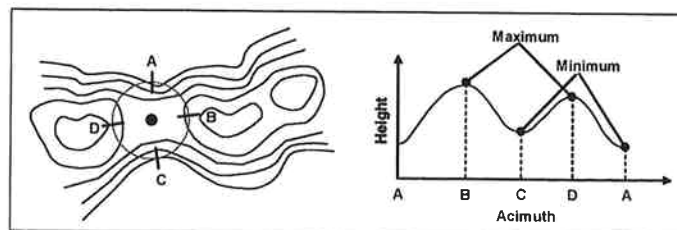


Figure 4. Determination of passes.

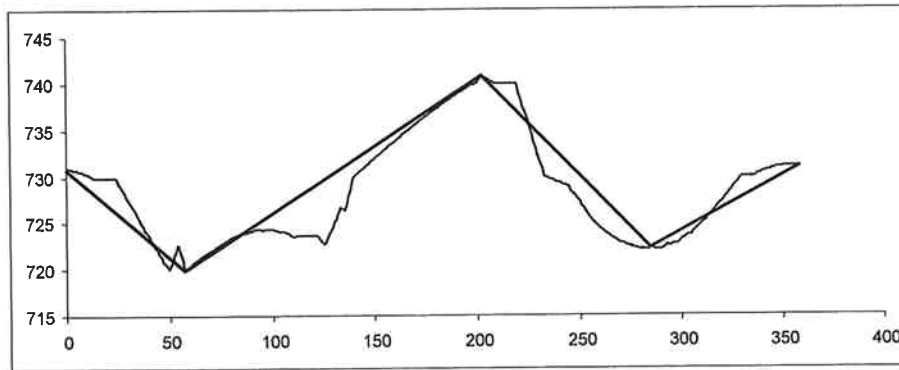


Figure 5. Simplification of the topographical profile.

3.1.3.- Categorization of the remainder points

The value assigned to the remainder points was only related with the height value. This means that the higher points have more importance than the lower ones.

3.1.4.- Weight assignment to the points

The final weight assigned to each point is calculated using a formula which combines the value of the previous categorizations and a specific value assigned by the user to each category. With this specific value, the cartographer can modulate the result of the automatic generalization depending on the final product, for example emphasizing the geomorphologic points or minimizing the interest points.

The formula is:

$$\text{weight} = (w_s_int * k_1) + (w_peak * k_2) + (p_pass * r_pass * k_3) + (p_depression * r_depression * k_4) + \text{height}$$

where:

- w_s_int = weight corresponding to the point of interest categorization
- w_peak = weight corresponding to the peak categorization
- p_pass = 1 or 0, depending on if the point is a pass point or not
- r_pass = weight of the pass point as a function of its altitudinal range
- p_depressions = 1 or 0, depending on if the point is a depression or not
- r_depressions = weight of the depression as a function of its altitudinal range
- height = height of the point
- k_n = specific value previously established by the cartographer

In the ICC test, the priority to assign the specific values (k_n) was, from the most to the least importance:

1. points of interest
 - a. crossroads
 - b. roads
 - c. map names

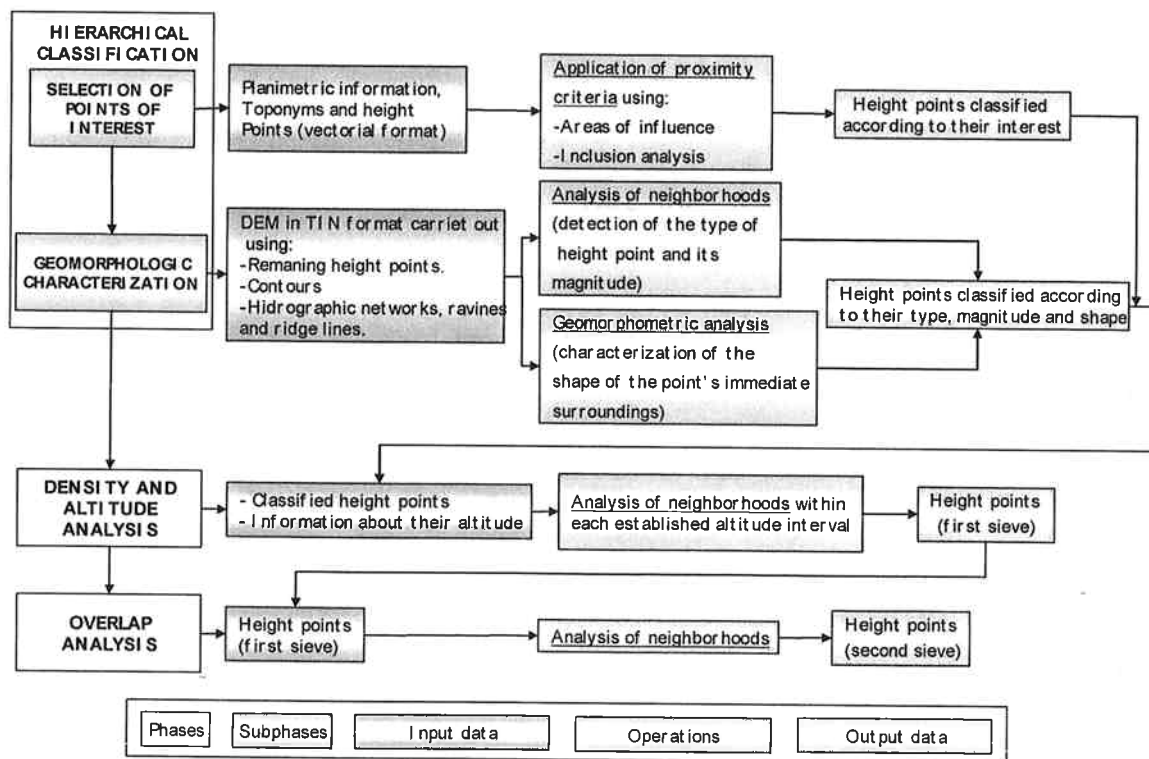
- d. other areas of special interest
- 2. peaks
- 3. mountain passes
- 4. depressions
- 5. rest of the points.

3.2.- Selection process

The selection step begins when each spot height has been evaluated and categorized. One of the input parameters of the application GENCONTES is the final density of the spot heights per square kilometer in the generalized data. The problem arises in deciding which points should be compared to decide which has a lower category and can be eliminated. In principle, it makes no sense to try to compare points that are too far away, because the more separated they are the less possibility there is that they belong to the same geomorphologic unit. To achieve the requirement of homogeneous distribution of the spot heights, the solution adopted was to use a regular space division, with rectangles of the same size that must contain the same number of points. The size of the division and the number of final points per division was decided after several tests.

3.3.- Process summary

The above mentioned methodological process can be summarized in the following diagram:



4.- From the original application of GENCOTES to the ICC specifications implementation

After defining several tests areas, the UPV processed the data and the ICC analyzed the results of the automatic generalization. The conclusion was that, although the results were really promising, some modifications must be introduced by the UPV in order to take into account the complete set of the ICC rules and to achieve results more similar to the obtained using manual generalization. Moreover the modifications, the ICC must prepare the input data to optimize the performance of the application.

The modifications were related to different aspects of the application and original data, as the interest points, the selection criteria, the classification and the workflow optimization:

- Interest points:
 - Road network. The ICC specifications for the Topographic Database at 1:25,000, require that for catalogued road network (roads coded by the Government):
 - There must be a spot height approximately every 1200 meters on the roads.
 - The input spot heights placed on the intersection of roads are considered as preferred spot heights.
 - On a road intersection, with multiple input spot heights, the hierarchy of the roads must be taken into account in order to select the point on the main road.

The input data was specifically processed to facilitate the modifications. For the first requirement, a point every 1200 meters, the roads were chained to eliminate nodes. For the other ones, the intersections were generated as classified point objects according to the hierarchy of the roads (Figure 6).

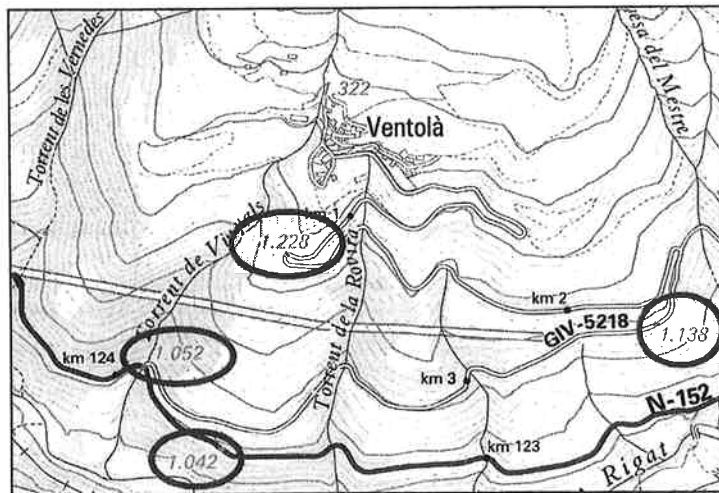


Figure 6. An example of the application of the criteria on the roads and on their intersections.

- Map names. In some cases, the selected point by the original application doesn't correspond to the type of geographical landform described by the map name, for example a spot height classified as peak is selected due its proximity to a map name of a pass. The ICC map names are not attributes of the topographic objects, but they are classified following geographical criteria. To solve the classification problems, the algorithm has been improved taking into account the ICC map name classification (Figure 7).

The input data includes a file with the map names classified as peaks or passes.

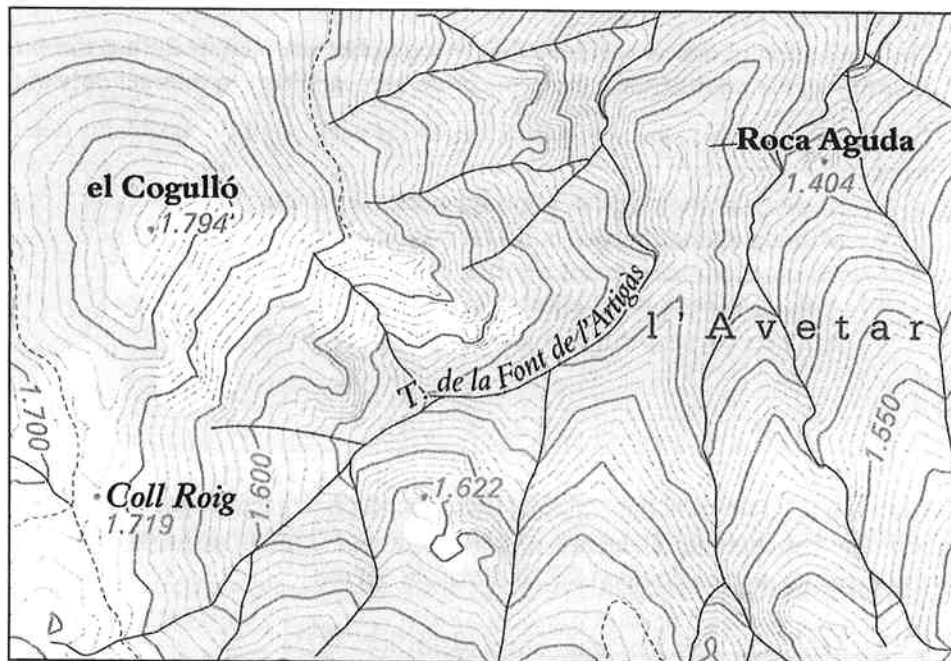


Figure 7. An example of the use of the map names as a criteria on the application.

- Although the application can manage the spot heights in urban areas, by the moment the ICC has decided don't use it, because the database object corresponding to the boundary of urban areas must be modified manually to obtain good results.
 - The application allows preserving spot heights manually selected before the automatic process, but the ICC has decided don't use this facility because the result obtained automatically by the application is good enough.
- Selection criteria
 - An ICC requirement is that an altimetric point with height, in meters, multiple of the contour line interval cannot be selected. In the Topographic Database at 1:25,000, the contour line interval is 10 meters.
 - Spot heights located on intersections between not catalogued roads and paths, are preferred point after interest point, peaks, passes and depressions. An new category of interest points has been added.

- Classification
 - In the original application the spot heights classification from the geomorphologic aspect was not applied to the interest points. The ICC suggested that it would be interesting to enrich the original database introducing the geomorphologic classification of all spot heights.
- Workflow optimization
 - The original application required the cartographer interaction during the process. In order to optimize the productivity, the application was modified to interact only at the start point.
 - To control the errors, the application stores for each point the problems detected during the process, by example:
 - Missing altimetric peak point for peak map name
 - Missing altimetric pass point for pass map name
 - Missing altimetric point for map name
 - Altimetric peak point eliminated by density problems
 - Altimetric pass point eliminated by density problems

5.- Results

In order to evaluate the modified GENCOTES utility with the ICC data, a comparison has been done between the automatic results and the obtained using traditional manual processes. Four sheets of the Topographic Database of Catalonia at 1:25,000 scale corresponding to different geographically areas have been compared (125 km² per sheet):

- A flat area (Sant Pere Pescador, sheet identifier: 7822)
- An urbanized area (Mollet del Vallès, sheet identifier: 7330)
- A middle high mountain area (Planoles, sheet identifier: 7221)
- A high mountain area in the centre of Pyrenees (Isil, sheet identifier: 6616) (Figure 8 and Figure 9).



Figure 8. The result of automatic generalization on the complete sheet, corresponding to a high mountain area in the centre of Pyrenees: in red color, the original spot heights, in green color, the result of the automatic generalization.

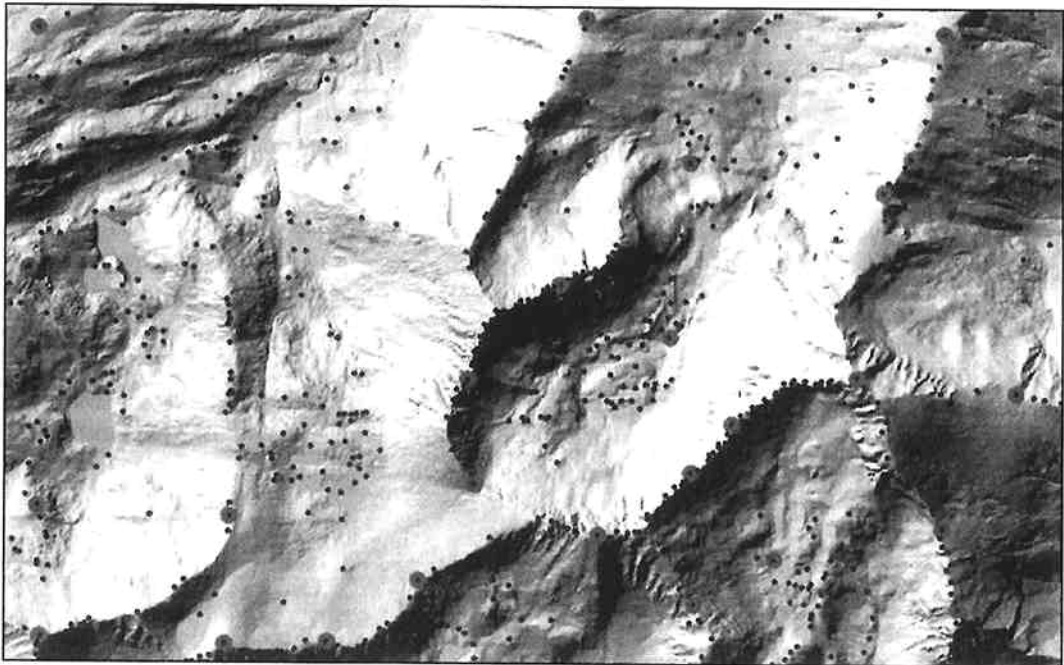


Figure 9. An enlarged view of the results (10 km²).