Generalizing the altimetric information of the Topographic Database of Catalonia at 1:5,000: classification and selection of break lines

Jesús Palomar-Vázquez *
Josep E. Pardo-Pascual *
Blanca Baella **
Maria Pla **

* Research Group Geoenvironmental Cartography and Remote Sensing, Department of Cartographic Engineering, Geodesy and Photogrammetry, Universitat Politècnica de València, cgat.webs.upv.es
** Institut Cartogràfic de Catalunya, Barcelona, www.icc.cat

1.- Introduction

Since 1983 the Institut Cartogràfic de Catalunya (ICC) has been collecting information in order to generate and update topographic databases, digital terrain models, orthoimages and maps. At large scales, until 1:25,000, the data is digitized using photogrammetric systems and stored in 2.5D data models, following, since the digital camera acquisition in 2004, a complete digital workflow.

ICC has been also compiling LIDAR data since 2002. Although the LIDAR systems provide quick acquisition of elevation data at a high level of accuracy, there are applications that require the addition of break lines to improve the finals results.

The Topographic Database at 1:5000 scale (BT-5M) is the most detailed database that covers Catalonia. During the steroplotting process all the vector objects required to generate a digital terrain model (DTM) and a digital surface model (DSM) in triangle format, 1.5 meter accuracy, are also compiled. Mainly the DTM is used for deriving contour lines and shaded relief and the DSM for rectifying orthophotoimages and generating city models. The use of the altimetric information of the BT-5M to derive elevation models, originally designed for scales close to 1:5000, has been extended to smaller scales. The high level of detail and the huge volume of data, especially in the case of the break lines, can produce problems, as costly processes during orthophoto rectification, which should be solved applying terrain generalization processes. These generalization processes must be focussed more in the geomorphologic classification than in the pure geometric simplification. The usefulness of the geomorphometrical studies to the cartographic generalization methods applied to the terrain, has been previously demonstrate in applications focused on the spot height selection (Palomar and Pardo, 2008; Baella *et al.*¸ 2007) and also in several test developed for digital elevation models (DEM) generalization (Palomar *et al.*, 2005).

The paper will explain in detail the developed methodology for break line generalization. After a brief description of the BT-5M break lines typology, a detailed explanation of the classification and selection process will be exposed, including the specific ICC requirements. The results of the first tests performed at the ICC production environment, on different types of terrain, will be resumed. At this moment the methodology is being implemented in the production workflow of the Orthophoto of Catalonia and, although the implementation is not yet completed, some conclusions will be given.

2.- Description of the break lines

In the BT-5M the terrain is modelled collecting scan lines, spot heights, break lines and contour lines to infer break lines. In particular, the break lines are used to represent the following terrain characteristics:

- Man made objects included in the catalogue of the BT-5M, as roads, path, agricultural plots, urban parcels or embankments.
- Natural objects included in the catalogue of the BT-5M, as rivers or creeks.
- Natural objects not included in the catalogue of the BT-5M, as mountain ranges or talwegs.
- Small terrain details not included in the catalogue of the BT-5M and needed to ensure cartographic quality in the automatic derivation of contour lines.

If the break line is coincident with a topographic element, instead of collecting it again the topographic element is attributed as a break line.

Most of the break lines representing small terrain details and some of them representing natural objects are not necessary for some uses of the altimetric information. Moreover, they increase the time and the complexity of the involved processes. Classification and selection of the most significant break lines will help to optimize the use of the altimetric models depending on the purpose.

3.- Classification and selection of break lines

The presented methodology focuses on a suitable characterization of break lines based on geomorphological criteria to decide which part of a break line is enough significant. Two main criteria have been applied: geomorphological analysis and error analysis after generalization process.

Geomorphological criteria analyze mainly the absolute mean curvature parameter (AMC) from a raster DEM. The curvature is the second derivative of the surface and it can be decomposed into two values: profile and plan curvature (figure 1). Although both values indicate either the degree of convexity or concavity, the profile curvature is calculated in the direction of the maximum slope, and the plan curvature is perpendicular to the direction of the maximum slope.

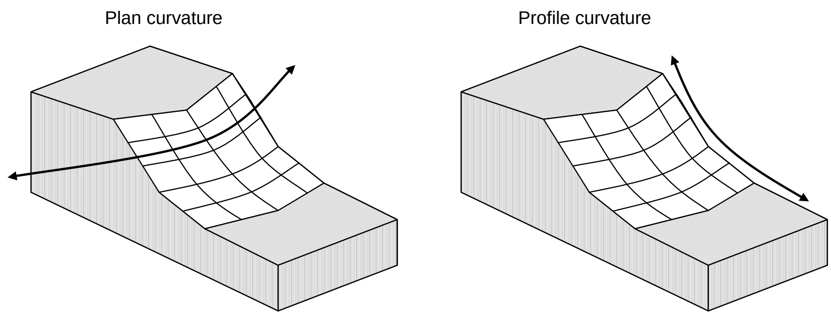

Figure 1. Curvature: plan and profile components

For the project goals, it is not necessary to know the degree of convexity or concavity, but only if this value is significant, so, the absolute mean curvature value (average between plan and profile curvature) will be studied. To obtain this parameter several algorithms (Zevenbergen and Thorne, 1987; Wood, 1996), neighbourhood size (3×3, 5×5) and commercial software (ArcGIS and ENVI) have been tested. Once the AMC has been computed, its value is transferred to all vertices in each break line (figure 2). The average value of AMC parameter from all vertices belonging to a break line is assigned to the break line. After several tests, a suitable average AMC threshold has been obtained which is used to select the most significant break lines. This threshold has been established after several analyses. Initially it was tried to establish a statistical threshold –taking in account the highest and lowest AMC values of the studied area-- but some times they were useful and others not –especially in the flat areas-. So, a lot of selected break lines weren't relevant to define the real terrain. Afterward it was decided to select an absolute threshold and after trying several values, it was selected the 20 value as a good approximation because using it, all selected break lines were significant geomorphologically. The differences between the new DEM, created using the selected break lines, and the original DEM helps to estimate if the selected break lines are able to characterize the terrain. The results showed the importance of a suitable neighbourhood size in the break line selection, small ones work better than higher ones, and, on the contrary, that the type of algorithm used was not determinant. So, it was decided to use the method of Zevenbergen and Thorne, with a 3 x 3 neighbourhood size and ArcGIS software.

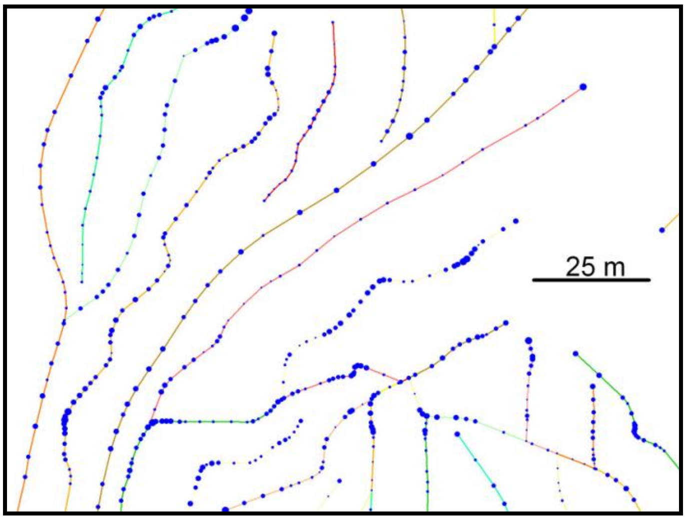

Figure 2. The point size indicates the value of AMC in each vertex.

After the geomorphologic selection, most tested cases showed that over 10% of DEMs surface had error higher than 1 meter. That fact means that the selected break lines can not totally define the landforms. So, it was decided to use an iterative strategy in order to search for the error zones and recover there some parts of the rejected break lines, firstly by converting raster error zones to vector polygons and secondly by splitting and selecting parts of the original break lines using the previous polygons. These break lines are merged with the break lines selected using the geomorphological criteria. This iterative process concludes when the error value reaches the defined value. Table 1 shows some results after applying the process three times. The reduction proportion of break lines ranges from 20 to 80% depending on the topographic roughness of the analyzed area.

Process area	Mean slope (Deg)	Number of BL (%)	BL length (km)	Length of Selected BL (%)	Error > 1 m (%)	Error > 1,5 m (%)
264_119	11.26	373.74	189.38	57.38	0.17	0.06
274_131	2.30	115.77	85.19	27.48	0.13	0.02
253_125	10.73	221.01	135.63	33.85	0.09	0.01
257_118	2.42	115.59	2605.00	28.99	0.32	0.02
280_133	5.19	156.46	103.65	34.03	0.25	0.04
280_129	18.85	269.17	143.69	62.40	0.45	0.12
276_134	14.00	252.67	157.08	49.83	0.32	0.08
262_119	9.36	341.91	151.58	45.89	0.10	0.00
260_75	28.48	117.53	295.93	61.19	0.51	0.09
264_115	3.68	184.00	133.62	38.06	0.13	0.01
261_73	31.69	159.79	289.84	78.68	0.92	0.24
257_73	21.22	139.01	186.23	58.80	0.24	0.04

Table 1. Break lines selected and error evaluation (AMC threshold = 20).

The meaning of each row in this table is the following:

- Process area: Code corresponding to each area to be processed.
- Mean slope: Mean slope in sexagesimal degrees system.
- Number of BL: Number of break lines segments generated by intersecting recursively error zones with original break lines. This value is indicated by a percentage respect the original number of break lines.
- BL length. Total length of all break lines on the original dataset.
- Length of selected BL: Percentage of selected break lines. This parameter indicates the level of database reduction.
- Error > 1m and Error > 1.5 m. They indicate the remaining percentage of error surface after the process.

Among the main preliminary results it can be seen:

- There is a relation between the percentage of selected break lines and the mean slope, as it is showed in Figure 3. This is an expected result, because it can be presupposed that areas with higher slope need higher number of break lines in order to define better the landforms.

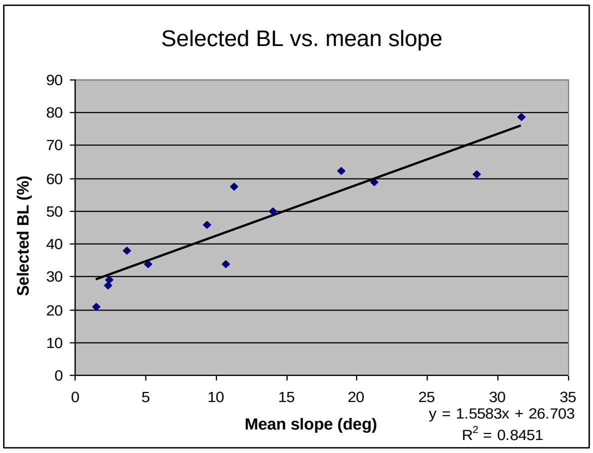

Figure 3. Relation between selected break lines and mean slope.

- After analysing the relation between mean slope and percentage of error higher than 1 meter, it can be deduced that only in areas with higher values of roughness the error caused by the elimination of several break lines will be significant (Figure 4).

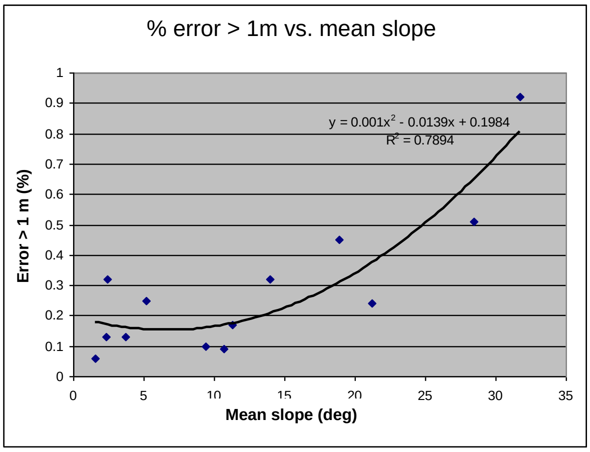

Figure 4. Relation between mean slope and errors higher than 1 meter.

4.- Additional ICC requirements

The previous methodology has been implemented into a Visual Basic application called GenBL (*Gen*eralització de *B*reak *L*ines) that works in the ArcGIS 9.2 environment. After the first version, the application has been improved in order to take into account some specific ICC requirements, mainly related with the break lines representing man made objects and the high level of fragmentation of the break lines after the classification process:

- All break lines representing paved roads must be preserved.
- The minimum length of break lines and gaps should be 25 meters.
- The selection process must ensure continuity between adjacent processed areas.

The first requirement has been solved using the type of the break line in the original data.

The second requirement has been achieved by including segments of break lines that haven't been selected in the previous steps until the break line length is longer than 25 meters or there are no gaps shorter than the same value. In Figure 5 and Table 2 it is showed an example.

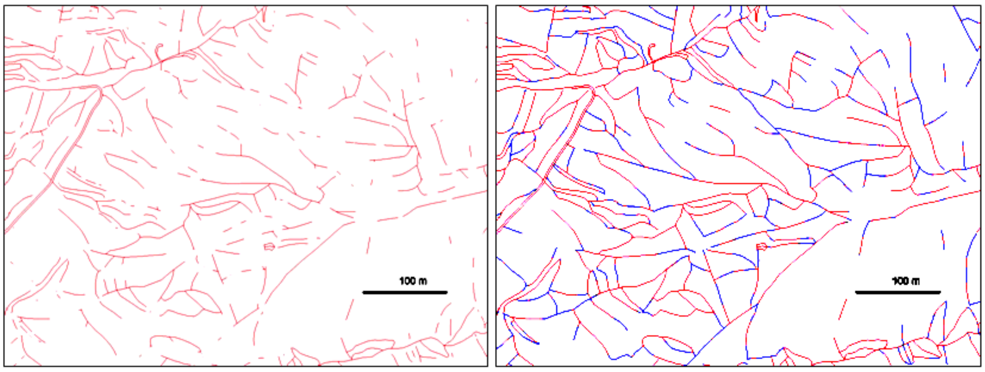

Figure 5. Selected break lines before, at the left, and after, at the right, taking into account the minimum length (process area 256-72).

	Length	
	km	%
Original data	182.85	100.00
Generalized data:	127.53	69.74
Geomorphological criteria	*76.51*	*41.84*
Error selection (1st iteration)	*45.35*	*24.80*
Error selection (2nd iteration)	*5.54*	*3.03*
Error selection (3th iteration)	*0.14*	*0.07*
Generalized data +length threshold (25 m)	145.78	79.73

Table 2. Data reduction in the different steps of the process (process area 256-72).

Third requirement, the continuity in adjacent processed areas, implies analysing a buffer around the area, which contains the information of the adjacent areas. Taking into account that, whatever the dimension of the expanded area, in the terrain model generation there are always errors in the borders, after some proofs it has been decided to use two buffers: one of 200 meters to calculate the curvature analysis and another one of 100 meters to ensure the continuity of the break line classification. To force the same classification in overlapping zones, the break lines have been split in the borders of the 100 meters buffer.

5.- Implementation at the ICC

The last test of the application has been performed at the ICC production environment. The classification and selection of the break lines has been applied to several areas with different characteristics, covering flat, urban and mountainous zones.

The results over a flat area are showed in the following figures:

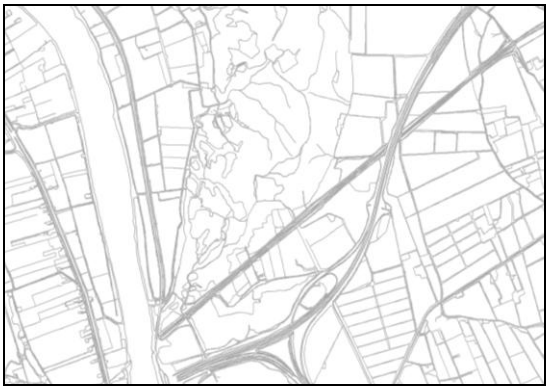
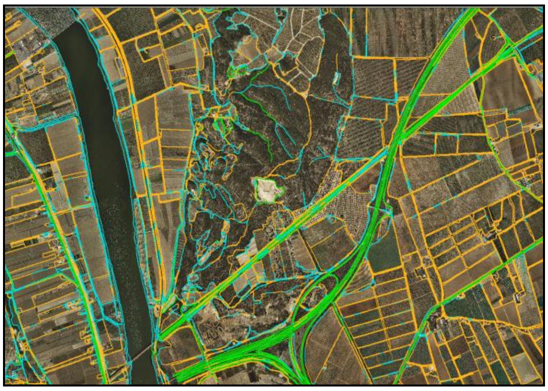

Figure 6. Flat area: on the left, the original break lines; on the right, the classified break lines on top of the image.

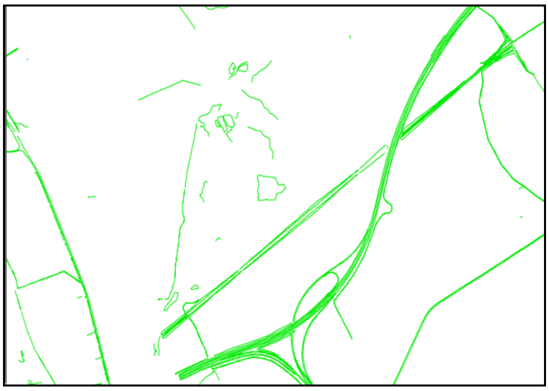
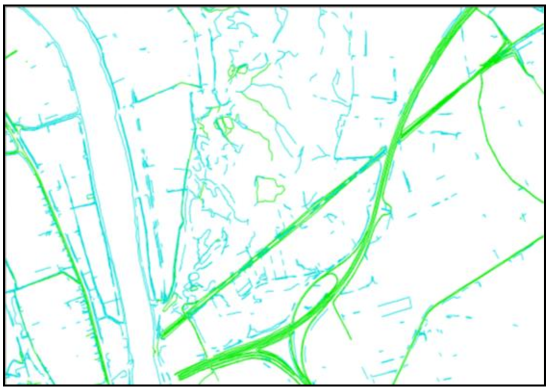

Figure 7. Flat area: on the left, in green colour, the geomorphological break lines; on the right, in blue colour, the break lines selected after the error analysis have been added.

Figure 8. Flat area: in orange colour, the eliminated break lines.

The results over an urban area are showed in the following figures:


Figure 9. Urban area: on the left, the original break lines; on the right, the classified break lines on top of the image.

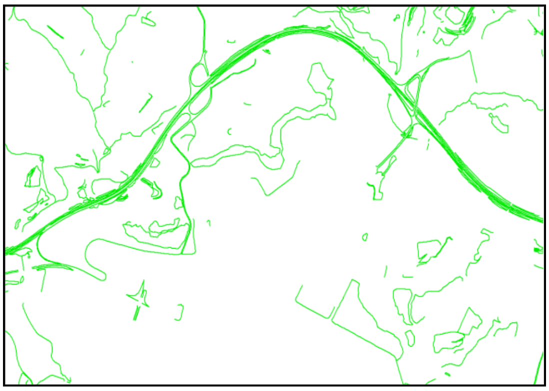
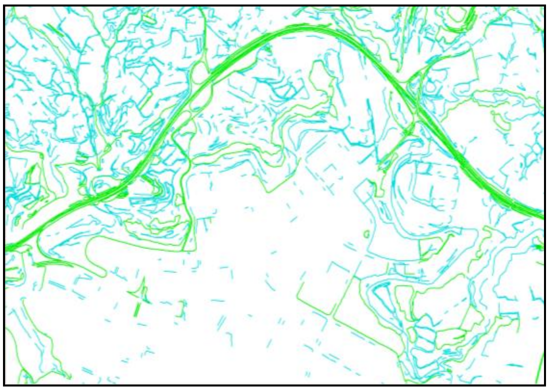

Figure 10. Urban area: on the left, in green colour, the geomorphological break lines; on the right, in blue colour, the break lines selected after the error analysis have been added.

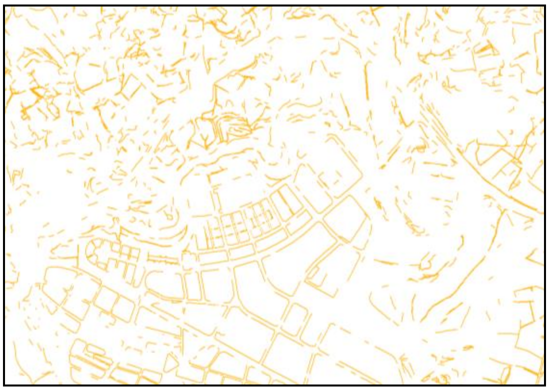

Figure 11. Urban area: in orange colour, the eliminated break lines.

The results over a mountainous area are showed in the following figures:

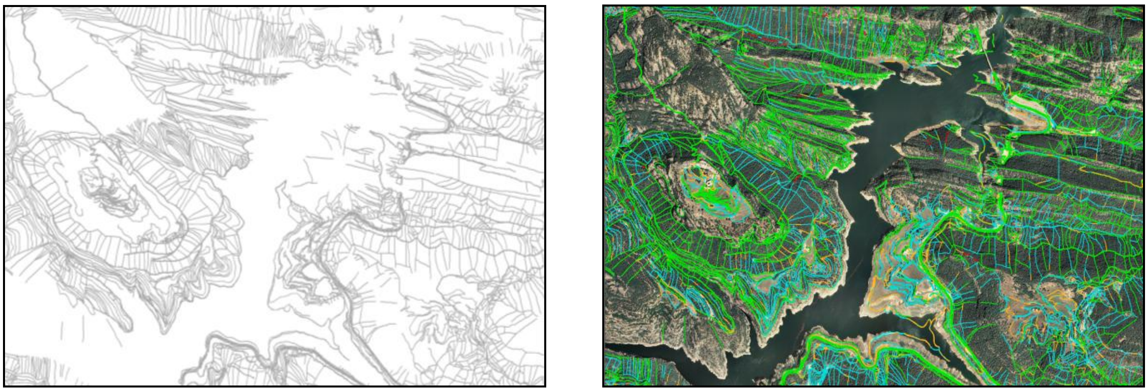

Figure 12. Mountainous area: on the left, the original break lines; on the right, the classified break lines on top of the image.

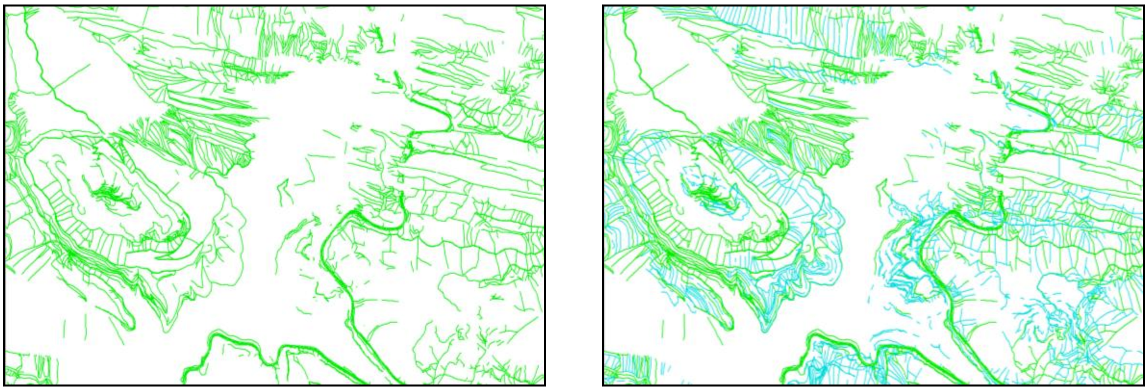

Figure 13. Mountainous area: on the left, in green colour, the geomorphological break lines; on the right, in blue colour, the break lines selected after the error analysis have been added.

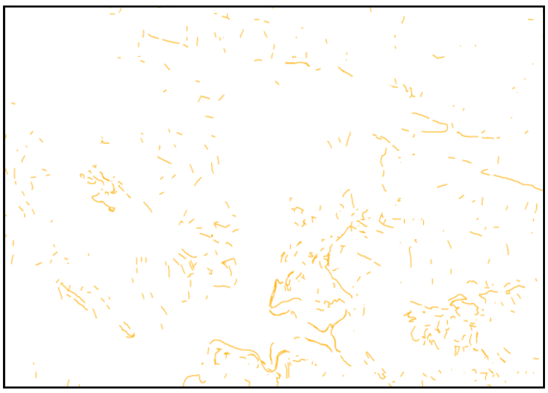

Figure 14. Mountainous area: in orange colour, the eliminated break lines.

The analysis of the results indicates, as can be seen in Table 3 and Table 4, that the method significantly reduces the number and the length of break lines, especially in the flat and urban areas.

	Number of original BL	BL selected by geomorphological criteria		BL selected by error analysis criteria		BL eliminated	
Flat	3229	386	12 %	1410	44 %	1433	44 %
Urban	3934	479	12 %	1932	49 %	1523	39 %
Mountainous	5798	3483	60 %	1626	28 %	689	12 %

Table 3. Classification of break lines: number of break lines.

	Length of original BL	Length of BL selected by geomorphological criteria		Length of BL selected by error analysis criteria		Length of BL eliminated	
Flat	436.4	83.6	19 %	115.0	26 %	237.8	55 %
Urban	400.8	95.8	24 %	161.6	40 %	144.4	36 %
Mountainous	357.3	218.1	61 %	110.9	31 %	28.3	8 %

Table 4. Classification of break lines: length of the break lines in km.

Next step is the implementation for generating the elevation models for the rectification of the Orthophotoimage of Catalonia at 25 cm. pixel, in triangle format, using the generalized break lines instead the original ones.

First tests have been applied on a set of 100 frames and the results have been analyzed in terms of accuracy and in terms of performance. Measurements of the accuracy of the rectified images obtained using generalized elevation models, show that the generalization has no influence on the accuracy of the product. Accuracy is measured comparing the coordinates of points measured on the field and the coordinates of the same point identified on the orthophotoimages. The save on the timing performance is around 12 seconds per frame, which represents an improvement of 3% in the rectification process. Although this reduction seems not very important, it becomes significant due the high number of frames rectified each year, around 80,000.

5.- Conclusions

The terrain generalization must be focussed more in the geomorphologic classification than in the pure geometric simplification. The proposed methodology selects the most significant break lines applying geomorphological analysis followed by an iterative process based on error analysis. The reduction ratio is high dependent of the terrain ruggedness: in flat areas is about 55% of break lines length but the percentage decreases until 8% in mountainous areas.

The methodology applied for the break lines generalization provides also an efficient tool to enrich the original data, adding a classification based on the geomorphological characteristics of the terrain.

The application of generalization techniques to the topographic data, including altimetric information, can go further away of topographic data derivation and benefit the generation of other type of products, as orthophoimages.

6.- References

Baella, B., Palomar Vázquez, J.M., Pardo Pascual, J.E. and Pla, M. (2007): *Spot heights generalization: deriving the relief of the Topographic Database of Catalonia at 1:25,000 from the master database*, ICA Commission on Generalisation and Multiple Representation Workshop, Moscow.

Palomar Vazquez, J.M.; Pardo Pascual, J.E. and Porres de la Haza, M.J. (2005): *Generalization of high-density raster Digital Elevation Models*, Proceedings International Cartographic Conference. Mapping approaches into a changing world, A Coruña, Spain.

Palomar Vázquez, J.M. and Pardo-Pascual, J.E. (2008): *Automated spot heights generalisation in trail maps*, International Journal of Geographic Information Science, 22 (1): 91-110.

Wood, J. (1996): *The geomorphological characterisation of Digital Elevation Models*, PhD Thesis, University of Leicester, UK. Available on http://www.geog.le.ac.uk/jwo/research/dem_char/thesis

Zevenbergen, L.W. and Thorne, C.R. (1987): *Quantitative analysis of land surface topography*, Earth Surface Processes and Landforms 12, pp. 47-56.