

# MORPHOLOGICAL TERRAIN CLASSIFICATION AND ANALYSIS USING GEOSTATISTICAL TECHNIQUES

J.M. Azañón<sup>a</sup>, J. Delgado<sup>b,\*</sup>, A. Gómez<sup>c</sup>

<sup>a</sup> Dpto. de Geodinámica, Facultad de Ciencias, Universidad de Granada.

Avda. Fuentenueva s/n – 18071 Granada, Spain – jazonon@ugr.es

<sup>b</sup> Dpto. de Ingeniería Cartográfica, Geodésica y Fotogrametría, Escuela Politécnica Superior, Universidad de Jaén,  
c/ Virgen de la Cabeza, 2 – 23071 Jaén, Spain – jdelgado@ujaen.es

<sup>c</sup> Dpto. de Ingeniería Cartográfica, Geodésica y Fotogrametría – Expresión Gráfica,  
Escuela Técnica Superior Ingenieros Agrónomos, Universidad Politécnica de Madrid,  
Ciudad Universitaria s/n, 28040 Madrid (Spain) – agomez@stereocarto.com

**KEY WORDS:** Spatial Information Systems, DEM/DTM, Geomorphology, Classification, Statistics

## ABSTRACT:

Nowadays there is a wide range of systems that provide an efficient and fast data capture for the DEM generation and the spread of the computer systems has allowed the use of DEM to a high number of technicians. Unfortunately, the DEM analysis tools do not have similar development. Actually, the analyses are mainly based in the application of local scale operators that provide additional information. In this paper we present a methodology based on the geostatistical techniques applied to the analysis and classification of the different terrain morphologies obtained from the DEM. It is based on the use of several parameters obtained from the spatial autocorrelation functions together with other parameters obtained directly from the DEM following a moving-windows schema. Through this process it is possible to obtain a multivariate digital model that include information relative to the terrain heights, like, for example, variability, continuity level and anisotropies. A non-supervised classification process is applied to the multivariate digital model. This classification allows selecting the more interesting parameters in order to make the final classification. The main advantage of this method is based on its capabilities of the DEM analysis following a multiscale schema. The analysis can include very interesting aspects like the anisotropies and other information derived from the structural analysis. This methodology is illustrated with an application example to the Guadix-Baza basin (S of Spain).

## 1. INTRODUCTION

The application of geostatistical methodology to DEM analysis has been exposed in several papers from different points of view. Thus, for example, the semivariogram has been used in several works in order to characterize the terrain morphology (Mulla, 1988) and the kriging techniques have been utilized for the elevation estimation (for example in Moffat et al., 1986).

The basic tool for the most of the geostatistical analysis is the variogram function. The variogram is a function that establishes the variance per points when the points are considered in pairs separated a distance  $h$  (see Goovaerts, 1997 and Webster and Oliver, 2001 for additional information). The variogram provides very important information. Thus, the variogram provides information about the influence range of the value registered in a point into the surroundings points, the variability at different scales, the anisotropies presence, continuity level, periodicities, among others. This information can be used in order to characterize the phenomenon that we are studying. In this paper, we intend to apply this information on the terrain characterization and automatic classification.

Usually, the terrain classification procedures are based only in the heights of the area contained in a digital elevation model or in the used of derived models such as slopes, orientation, gradients, ... . The main advantage of these procedures is that they work only on local level (neighborhood of a determined

point). Using the geostatistical procedures a multiscale work approach can be made.

## 2. METHODOLOGY

The main objective of the proposed methodology is the terrain classification using the information derived from the statistical and geostatistical treatment of the digital elevation model. The applied classification is a non-supervised classification based in the ISODATA procedure (Tou and Gonzalez, 1974).

The proposed methodology can be divided into the following steps (figure 1):

- determination of the parameters of the moving-windows schema that will be used. These parameters must be defined according to the zone's dimensions and data number available. The schema presents an overlapping windows system.
- basic statistics and directional variograms calculation of the data that are included in the different windows that are considered.
- parameters selection that will be used in the classification process. The parameters can be obtained from the statistic analysis (mean, variance, minimum, maximum, quartiles, range, ...) and the geostatistical analysis (variogram values at different directions and distances, ratio between the variogram values - anisotropy ratios-, ... .

---

\* Corresponding author.

- generation of the Multivariable Digital Terrain Model -MDTM- that includes all the variables that has been extracted from the statistical (and geostatistical) analysis.
- non-supervised classification (ISODATA method) of the MDTM.
- analysis of the detected classes characteristics. Quality control of the classification.
- classification of the experimental data considering the window where each point is included.
- basic statistics and directional variograms of the classified data.

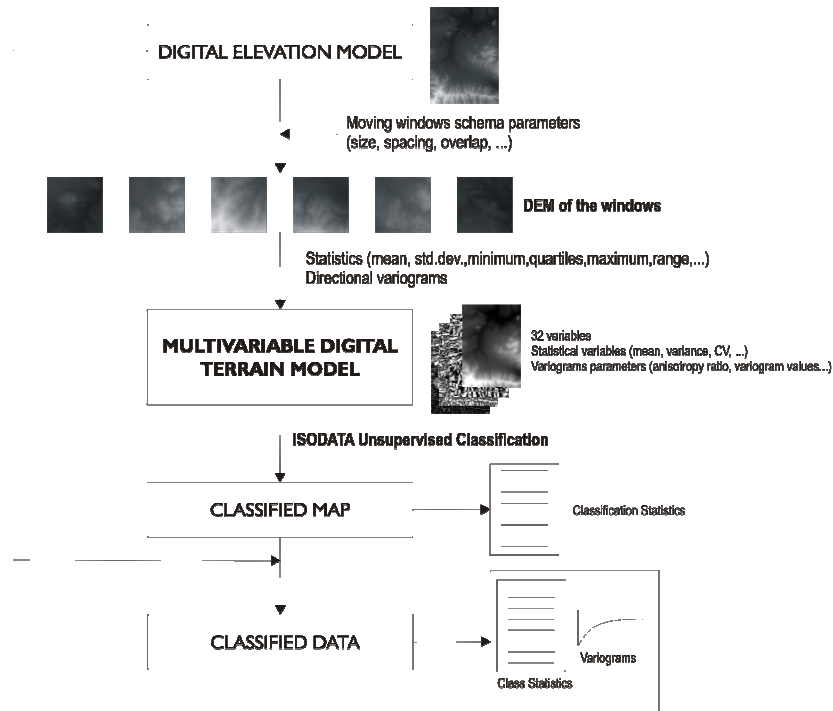


Figure 1. Workflow of the proposed methodology

### 3. EXAMPLE

#### 3.1 Data set

In order to show the capabilities of the proposed methodology, it has been applied to an example area. The selected area is located between the Granada and Almería province (S of Spain) being banked to the S for Sierra Nevada's relief (figure 2).

The used information for this terrain morphological classification has been a digital elevation model in format grid with 20x20m's spacing generated starting from a contour map's digitalization to scale 1:50000.

The selected area was selected for its geological interest. In the area, there is a important diversity of morphological units. In the south, the high summit of Sierra Nevada's foothills with a very important drainage network development is present. In the north the area the terrain presents minor heights (Negratín Dam zone). A very important surface (central area) is occupied by the Guadix Basin that shows up as a flat area surrounded of important relief.

The zone has 50x70 km surface that produces a DEM with 8,756,001 points distributed according to a mesh 2501 samples x 3501 lines. In figure 3 a DEM grayscale

representation is shown. In table 1 the most important basic statistics of the terrain heights are summarized.



Figure 2. Geographical localization

The most important characteristic of the area is the amplitude of the heights range that stretches out from 500 meters on the sea level to the 3480 meters, (Mulhacen peak, Iberian Peninsula maximum height). The mean value is 1261 meters and the median value 1101 meters (that is representative of the heights of the Guadix plain that occupies an important surface of the selected area).

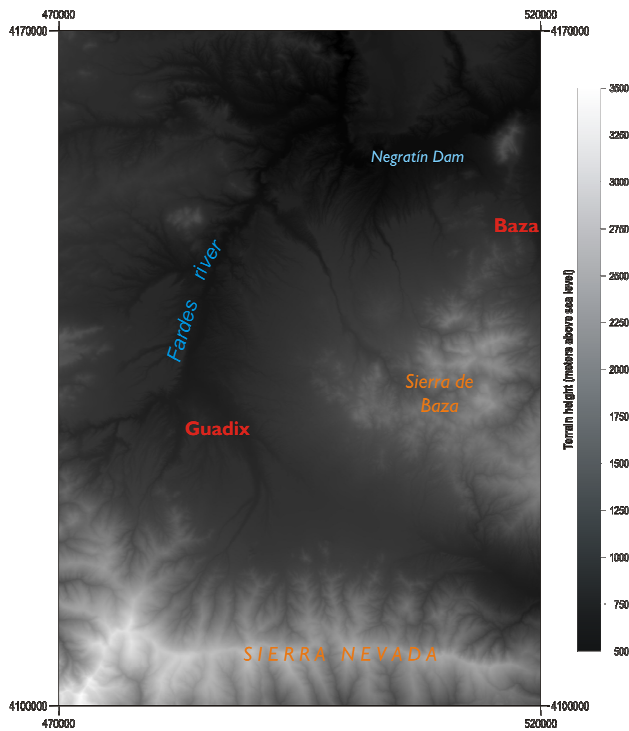


Figure 3. DEM Grayscale representation

No.Data	8756001	Minimum	500
Mean	1261.02	Lower Quartile	933
Variance	241483.79	Median	1101
Range	2980	Upper Quartile	2763
Interquartile range	1830	Maximum	3480

Table 1. DEM Basic statistics

### 3.2 Geological localization

The study area is situated in the central sector of the Betic Cordillera, to the north of Sierra Nevada range. This area comprises most part of the Guadix basin. The Guadix basin overlies the contact between the Internal and the External Zones. This intramontane basin was established as a separate basin in the Late Miocene (Fernández et al., 1996; Soria et al., 1998) after the main tectonic movements that formed the large structures of the Betic Orogen. Sedimentation in the basin began in the Tortonian after the partial closure of the North-Betic Strait. The basin was filled in the Late Pleistocene, concurrent with the termination of the endorheic infilling. Since then, the entire basin area has formed part of the catchment area of the Guadalquivir River, and has been subjected to significant erosion (Calvache and Viseras, 1997).

### 3.3 Classification

Considering the dimensions of the study area and the available information (and its density), a moving windows schema has been designed. The windows size is 2000mx2000m (100x100 data points) and the windows displacement is equal to 200m in both directions (10 data points). This schema produces 241x341 windows.

At each window, ten descriptive statistics have been calculated: mean, variance, coefficient of variation,

minimum, lower quartile, median, upper quartile, maximum, range and interquartile range; and the directional variograms in the four main directions (N0°E, N45°E, N90°E, N135°E) considering 20m lag –at N0°E and N90°E directions- and 28.284m lag –at N45°E and N135°E directions-. From the variogram calculations the following variables have been selected:  $\gamma_{ij}$  value for i-direction (100m, 500m, 1000m and 1500m) at j-direction (N0°E, N45°E, N90°E and N135°E); for each direction; the maximum value, the direction of maximum variability (maximum value) and the anisotropy ratio (R $\gamma$ ) defined as the maximum value divided by the minimum value. A total of 32 statistical parameters have been obtained for each window. These data have been included in a multivariable digital terrain model (figure 4) that has been processed using an ISODATA non-supervised classification.

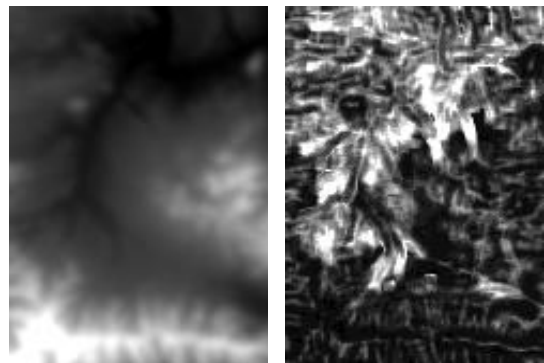


Figure 4. Variables obtained from the statistical analysis used in the classification. Left: Local mean terrain height; Right: Local variogram values at N0°E direction (distance 100m).

### 3.4 Results

A total of seven classes are obtained in the final map (figure 5 and table 2).

**Class A (9.28% of total area):** This class is related with small irregularities that are present in the Guadix basin. The class has a mean height value of 1062m and a coefficient of variation around 16%. A very characteristic statistic is its interquartile range of 146m, that indicate what the terrain is basically a flat area with a reduced variability. The directional variograms suggests that phenomenon is very continuous and isotropic (with anisotropies ratios lower than 1.5 for 1500m distance). This class is composed by the small sills presents in the Guadix basin.

**Class B (47.79% of total area):** This class is related with the most of the Guadix basin area. The class has a mean height value of 983m and a coefficient of variation around 18%. This is a flat area too with a interquartile range of 247m. This is a very continuous and isotropic area with anisotropy ratios lower than 1.16 in all distances.

**Class C (0.98% of total area):** This class presents a higher mean value than Classes A and B (1093m) and greater variability. has similar characteristics than Class A, but present a greater variability. The basic characteristic of this class is its anisotropy. The variability is lower in N45°E direction with ratios that raise around 5.8 at 1500m distance. This class is presented positional linked to the Class A.

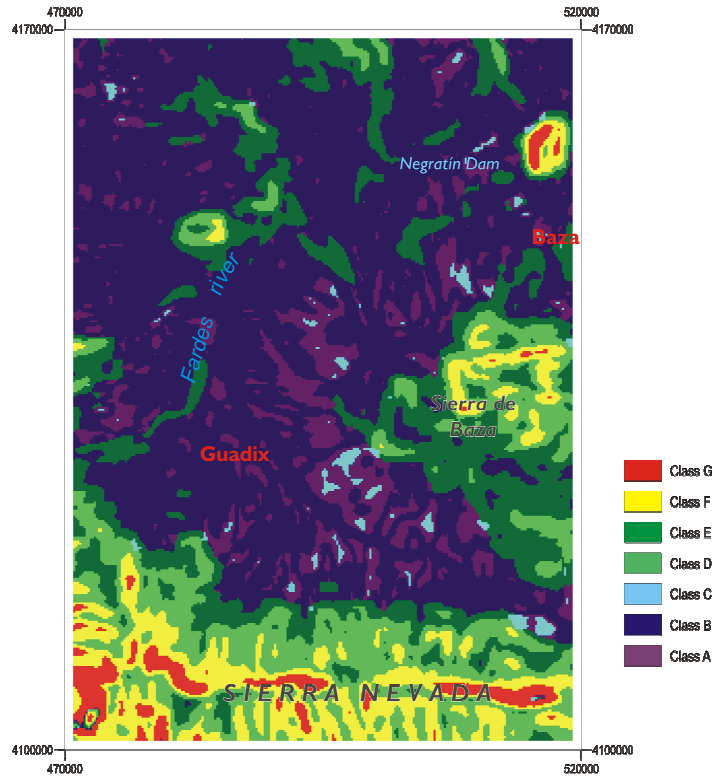


Figure 5. Classified map

	Class A	Class B	Class C	Class D	Class E	Class F	Class G
No.Data	813080	4184286	86200	1506290	1212875	746270	207000
Mean	1062.61	983.20	1092.79	1272.21	1728.53	2018.27	2175.44
Variance	28772.96	37556.74	37418.63	176974.23	210281.30	189515.00	212158.11
CV%	15.96	19.71	17.70	33.07	26.53	21.57	21.17
IQR	146	247	186	606	509	518	514
Lower Quart.	997	853	1000	921	1465	1750	1948
Median	1090	1000	1115	1271	1680	1980	2221
Upper Quart.	1143	1100	1186	1527	1974	2268	2462
$\gamma_{100.0}$	30.86	72.49	24.72	200.02	320.12	436.87	626.90
$\gamma_{100.45}$	28.60	74.51	23.78	199.40	308.30	409.72	605.35
$\gamma_{100.90}$	28.74	74.09	27.24	205.39	302.19	394.64	678.77
$\gamma_{100.135}$	30.07	70.52	26.00	204.25	315.48	430.34	716.74
$\gamma_{500.0}$	173.41	486.45	262.24	2013.31	4086.78	6888.61	9787.60
$\gamma_{500.45}$	162.55	513.36	111.08	1958.45	4142.86	6962.95	10107.81
$\gamma_{500.90}$	181.43	518.73	300.20	2113.38	4397.20	6149.47	12134.25
$\gamma_{500.135}$	182.94	480.28	343.56	2163.30	4523.47	7128.19	13674.94
$\gamma_{1000.0}$	378.92	949.24	827.20	4372.82	9948.51	17186.59	25473.50
$\gamma_{1000.45}$	433.46	1008.22	154.88	4426.16	10426.92	21927.68	31459.38
$\gamma_{1000.90}$	513.50	1041.00	663.36	4760.57	12344.45	17875.52	33310.50
$\gamma_{1000.135}$	443.57	934.81	1127.11	5158.91	12080.86	21735.34	44616.25
$\gamma_{1500.0}$	652.75	1444.43	1474.60	6943.82	18325.97	23798.11	40520.21
$\gamma_{1500.45}$	604.56	1602.13	254.33	6824.80	18092.25	39150.76	55201.54
$\gamma_{1500.90}$	902.04	1676.16	1065.36	7710.14	23775.01	31821.63	50327.83
$\gamma_{1500.135}$	815.39	1453.88	1509.16	8357.93	21686.71	35960.59	69862.50
$R\gamma_{100}$	1.08	1.06	1.15	1.03	1.06	1.11	1.18
$R\gamma_{500}$	1.13	1.08	3.09	1.10	1.11	1.16	1.40
$R\gamma_{1000}$	1.36	1.11	7.28	1.18	1.24	1.28	1.75
$R\gamma_{1500}$	1.49	1.16	5.93	1.22	1.31	1.65	1.72

Table 2. Basic statistics and variograms values of the different classes

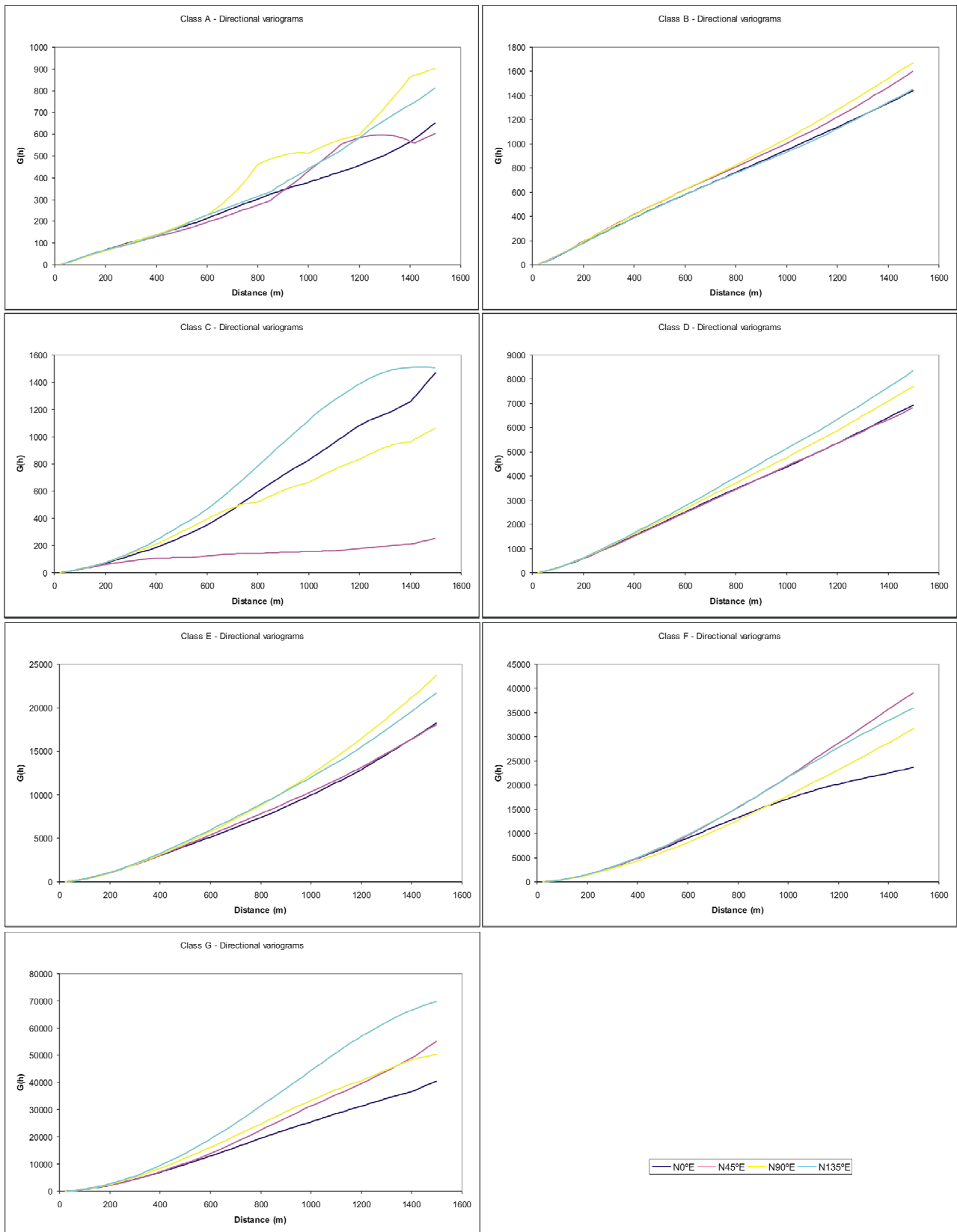


Figure 6. Directional variograms of the classified data

Class D: Represents 17.20% of the total area. This class is related with the surrounding relief to the Guadix basin. The class has a mean height value of 1272m and its variability is greater (coefficient of variation around 33%) due to the well developed hydrological network (gullies). Nevertheless, the directional variograms present a continuous and isotropic behaviour at distances lower than 1.22. The variability is slightly lower in E-W direction due to the foot of Sierra Nevada range.

Class E: Represents 13.85% of the total area. This class is related with the hillside of the surroundings relief. The class has a mean height value of 1728m and present a considerable variability (coefficient of variation around 27%). The variograms present a small variability in N0°E and N45°E directions according to the north hillside of Sierra Nevada.

Class F: Represents 8.52% of the total area. This class includes high-terrain areas of the Sierra Nevada and Sierra de Baza mountains. The class has a mean height value of 2018m and presents a coefficient of variation of 22%. The directional variograms presents a greater continuity in N0°E direction for distances greater than 1000m.

Class G: Represent only 2.36% of the total area and it is related with the highest areas of the zone (Sierra Nevada and Sierra de Baza mountains peaks). The class has a mean height value of 2175m and present a coefficient of variation of 21% (similar that the previous class). The directional variograms have a similar behaviour has the corresponding ones of the Class F.

### 3.5 Geological interpretation

Classes D, E and F (light green, yellow and red colors in Figure 5) correspond with rocks which belong to the basement of the basin (Alboran Domain and Sudiberian Domain). In classes A and B (purple and blue colors) appear sedimentary rocks of the Guadix basin. The present fluvial pattern, corresponding to Guadalquivir river, can be identified in the Classes B and C (blue and green colors). Areas classified as Class A (purple colors) correspond to a glaciais surface which represents the termination of the endorheic infilling.

## 4. CONCLUSIONS

The proposed method for the morphological terrain classification using statistical and geostatistical information provides valuable information about the terrain. This information is obtained from an objective procedure based in the ISODATA non-supervised classification using multivariable digital terrain model. The model is composed with a large number of variables that reflect different characteristics of the terrain (mean terrain height, variability, anisotropies, ...).

The used information in the classification process provides a more robust classification that the usually applied based in a small number of variables. The classical procedures are based in a direct classification using in many times only one variable (for example, heights or slopes).

The presented example shows a clear relationship between the established classes and lithological units of the area, providing additional statistical information very useful for the numerical terrain characterization.

## REFERENCES

- Calvache, M.L. y Viseras, C., 1997. Long-term control mechanisms of stream piracy processes in southeast Spain. *Earth Surface Processes and Landforms*, 22(2), pp. 93-105.
- Fernández, J., Soria, J.M., Viseras, C., (1996). Stratigraphic architecture of Neogene basins in the central sector of the Betic Cordillera (Spain): tectonic control and base level changes. In: *Friend, P.F., Dabrio, C.J., (eds.) Tertiary basins of Spain: the stratigraphic record of crustal kinematic*, Cambridge University Press. Cambridge, pp. 353-365.
- Goovaerts, P., 1997. *Geostatistics for natural resources characterization*. Applied Geostatistics Series, Oxford University Press, New York, pp. 483.
- Leenaers, H.; Okx, J.P. and Burrough, P.A., 1990. Comparison of spatial prediction methods for mapping floodplain soil pollution. *Catena*, vol. 17, pp. 535-550.
- Moffat, A.J.; Catt, J.A.; Webster, R. and Brown, E.H., 1988. A re-examination of the evidence for a Plio-Pleistocene marine transgression on the Chiltern Hills. I. Structures and surfaces. *Earth Surface Processes and Landforms*, vol. 11, pp. 95-106.
- Mulla, D.J., 1988. Using geostatistics and spectral analysis to study spatial patterns in the topography of southeastern Washington State, USA. *Earth Surface Processes and Landforms*, vol. 13, pp. 389-405.
- Soria, J.M., Viseras, C., Fernández, J., 1998. Late Miocene-Pleistocene tectono-sedimentary evolution and subsidence history of the central Betic Cordillera (Spain): a case study in the Guadix intramontane Basin, *Geol. Mag.*, 135, pp. 565-574.
- Webster, R. and Oliver, M.A., 2001. *Geostatistical for Environmental Scientists*, Statistics in Practice Series Edition, John Wiley & Sons, Chichester, United Kingdom, pp. 271.
- Tou, J.T. and González, R.C., 1974. *Pattern recognition principles*, Addison-Wesley Publishing Company, Reading, Massachussets.

## ACKNOWLEDGEMENTS

The present study has been sponsored by the grant REN2001-3378/RIES from the I+D+I program of the spanish Ministerio de Ciencia y Tecnología, partially financed by FEDER funds of the European Union.