METHODS FOR NATURAL LAND MAPPING UNITS DELINEATION FOR AGRICULTURAL LAND EVALUATION

Iuliana Cornelia NICULIȚĂ[1], Mihai NICULIȚĂ[1]
E-mail: iuliananiculita@geomorphologyonline.com

Abstract

Agricultural land evaluation has a key role in the sustainable agriculture. The agricultural land evaluation methodology is applied to land mapping units for computing a suitability index, based on the value of several soil and environment indicators, which characterize these land mapping units. The natural land mapping units, are delineated using various criteria and thematic layers, but most times the approach is subjective. GIS, geomorphometry, remote sensing and geostatistics bring the possibility to objectively delineate most suitable natural land mapping units for applying the agricultural land evaluation methodology. The methods for natural land mapping units delineation can be divided in two classes of methods: supervised and unsupervised. The first, require some knowledge about the area, and can be used to carry the results for a specific purpose of the land evaluation. The last, related especially to cluster analysis and image segmentation, depend on the input data and the number of specified classes or the seed points, so require first the analysis of the input data, to reveal the clusters/seed sampling. Both approaches were used to delineate the natural land mapping units for a DEM covering a test area, and were used to extrapolate the method settings for a DEM covering 15 villages from Iasi county agricultural area. Because reference data concerning the natural land mapping units is almost impossible to derive, we analyzed statistically and conceptually the results along a topographic transect, in order to try to find the most suitable method. Generally, unsupervised segmentation methods gave the best results, and from them the segmentation procedures, although very intensive from a computational point of view, can depict interesting patterns of natural aggregation of natural land mapping units.

Key words: land mapping unit, agricultural land evaluation, GIS

Land evaluation plays an important role in the development of a sustainable agriculture (Davidson, 2002). Using various methodologies, a suitability index is computed for homogenous environmental areas (named in this paper natural land mapping units), as in Romanian Methodology of Land Evaluation (ICPA, 1987), or for land units, as in FAO framework (FAO, 1977). The classical approach for delineating these natural land mapping units, rely on the expert-knowledge opinion applied to topographic contours maps, using slope and exposition (aspect), as the principal criteria for delineation. GIS, geomorphometry, remote sensing and geostatistics recent advances, bring the possibility to derive methods for objectively and automatically delineate the land mapping units for land evaluation. In our approach we considered only a DEM and geomorphometric variables as inputs in the delineation approach, because we considered them stable and error free, in comparison with other interpolated inputs (geology, climate, radiation, vegetation, soil), with considerable uncertainty and also without great mapping resolution especially on small working scales. We performed several supervised and unsupervised methods, used in geomorphometry, remote sensing and earth sciences in general. The supervised methods require knowledge about the data and a previous semantic schema of the target land units. The unsupervised methods don't need previous knowledge of the data, but for better and interpretable results, there is need of a training set or the number of required classes/cluster (which sometimes is computed prior the classification, using several approaches).

MATERIAL AND METHOD

As the base for the delineation approach we used a digital elevation model obtained from contours (from 25k topographic maps) and height points (from 5k topographic maps), obtained with Thin Plate Splines function from SAGA GIS 2.0.7 (www.saga-gis.org), with 11 levels for a smooth generalization and outliers removal. This approach and a pixel size of 12.5 m were used after (Hengl, 2006) guidelines for eliminating as well as possible the interpolation errors. The

[1] University Al.I.Cuza, Iasi, Faculty of Geography and Geology, Departament of Geography

digital elevation model area *(fig. 1)*, cover 15 villages, situated in Iasi county agricultural area, and for computing geomorphometric variables a bigger area was processed (an area bigger with an buffer polygon of 500 m). The landforms of the area are specific for monoclinal structure, with asymmetric hillslopes (cuestas). The floodplains are flat and have wide sections. The climate is temperate-continental. Because of the intense human agriculture practice, the old natural vegetation is replaced.

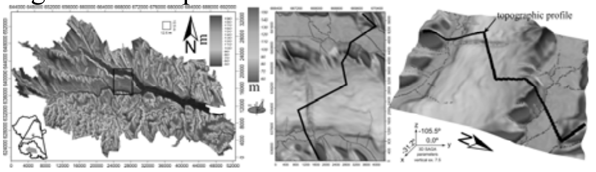

Figure 1 **Geographical position and digital elevation model of the study area**

(a color version is available at www.geomorphologyonline.com/fig1.png)

As supervised methods, the simplest case is a classification tree approach based on *n* classes of slope and *n* classes of orientation, with the result of *n* x *n* final classes of land mapping units. Because there is little knowledge about the real terrain specific classes, and often arbitrary intervals of class are used (in the Romanian methodology (ICPA, 1978) there are used nine classes: ≤ 2.0%, 2.1-5.0%, 5.1-8.0%, 8.1-12.0%, 12.1-18.0%, 18.1-25%, 25.1-35.0%, 35.1-50% and ≥50%), we preferred to analyze the slope and orientation frequency, but to classify we choose the official pedological classification. The slope (percent) and exposition (aspect in degrees, towards north direction) were computed using Horn algorithm in GRASS GIS 6.0.4 (http://grass.fbk.eu/) *r.slope.aspect* function and a histogram for the study area is represented in *fig. 2*. Exposition was classified according to the principal expositions: N, NE, E, SE, S, SW, W and NW. We can obtain by this method 72 possible classes, using the 9 classes of curvature and the 8 classes of exposition. The histogram data show 9 classes but with different threshold values.

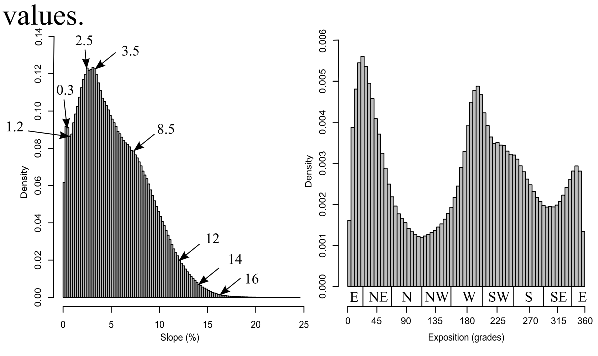

Figure 2 **Histogram of slope and exposition for the study area**

Related to the form of ridge, channels and hillslopes, and quantified using curvatures (profil, tangential, minimum, maximum) is the classification of (Schmidt and Hewitt, 2004) *(fig. 3)*. This classification produce 9 classes of hillslope form (areas with slope bigger than 0.1°: convex-convex - nose, convex-straight, convex-concave – hollow shoulder, straight-convex - spur, straight-straight – planar slope, straight-concave - hollow, concave-convex – spur-foot, concave-straight – foot slope, concave-concave – hollow foot) and 6 classes of ridge/channel form (areas with slope under 0.1°: peak, ridge, flat area, saddle, channel, pit/depression). Profil and tangential curvatures (1/m) were computed using GRASS GIS 6.0.4 *r.slope.aspect* function. Minimum and maximum curvatures (1/m) were computed using (Wood, 1996) formulas and second derivatives computed in GRASS GIS 6.0.4 *r.slope.aspect* function.

As a mixt, form and positional supervised method, the method of (Pennock et. al., 1987) extended by (Reuter et. al., 2006), classify the land in 11 classes *(fig. 4)* (divergent shoulder, planar shoulder, convergent shoulder, divergent backslope, planar backslope, convergent backslope, divergent footslope, planar footslope, convergent footslope, low catchment level, high catchment level) using as input slope (for shoulder, backslope, footslope differentiation), profile curvature, plan curvature (both for divergent, planar, convergent differentiation) and catchment area (for low and high level differentiation). Plan curvature (1/m) was computed using *r.param.scale* function in GRASS GIS 6.4.1, where are implemented [9] equations of curvature computation. The criteria for differentiating planar slopes was ±0.0001 1/m curvature and for separating sloping parameters and flat elements the slope of 0.3°. The criteria for upland/lowland were chosen ≥ 0.85 and ≤ 0.15 normalized height (dimensionless), instead of using catchment area. Normalized height (dimensionless) was computed in SAGA GIS 2.0.7 using Terrain Analysis – Morphometry – Relative Heights and Slope positions functions.

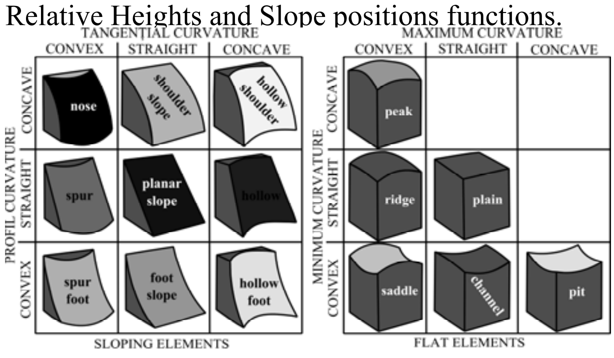

Figure 3 **Schmidt and Hewitt, 2004 curvature based classification (modification after [8])** **(a color version is available at www.geomorphologyonline.com/fig3.png)**

Using extended positional criteria (MacMillan et. al., 2000) extend Pennock-Reuter 11 classes, obtaining 15 classes *(fig. 5)*. The original classification of MacMillan uses fuzzy membership functions and several other geomorphometrical parameters beside slope, profil, plan curvature and catchment area. Instead of this big number of geomorphometrical parameters we used normalized height computed using SAGA GIS 2.0.7 and instead plan curvature we used tangential curvature.

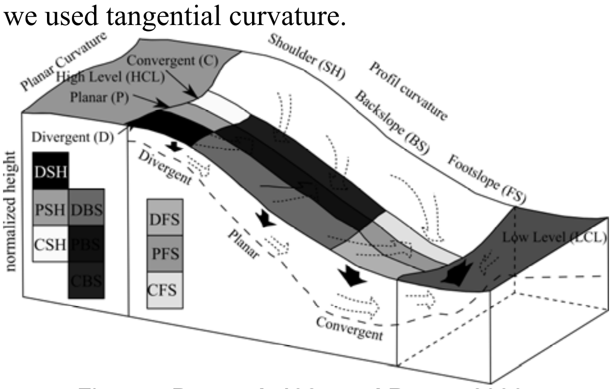

Figure 4 **Pennock 1987 and Reuter 2006 classification (modification after [11])** (a color version is available at www.geomorphologyonline.com/fig4.png)

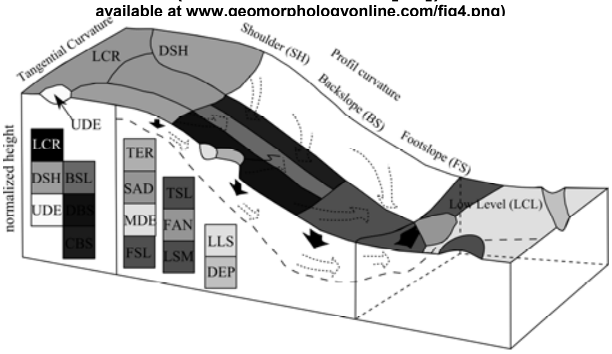

Figure 5 **MacMillan 2000 classification** (a color version is available at www.geomorphologyonline.com/fig5.png)

As unsupervised methods we used several algorithms implemented in SAGA GIS and R software: Cluster analysis, training set Minimum Distance, Region Growth and Image Segmentation. The use of SAGA GIS and R are preferred because the software is open-source and is very easy to use by soil scientists.

Cluster Analysis was applied using R after the code made by (Hengl, 2009) and the code from (Venables and Ripley, 2002) and (http://www.statmethods.net). This unsupervised method of classification tries to search the *n* clusters of *n* supplied by the user classes, which can be spatially assigned to pixels for mapping. An R based PCA of slope, exposition, profile curvature, plan curvature, tangential curvature and normalized height (using Hengl, 2009 code), showed that only slope, aspect and normalized height are uncorrelated. Because the standard method for the optimum number of classes didn't gave the desired results, analyzing the data from Fig. 5 we can conclude that 7-12 classes represent the optimum number of classes, because from this number of classes the sum of the squares between the clusters decrease abruptly and then slowly.

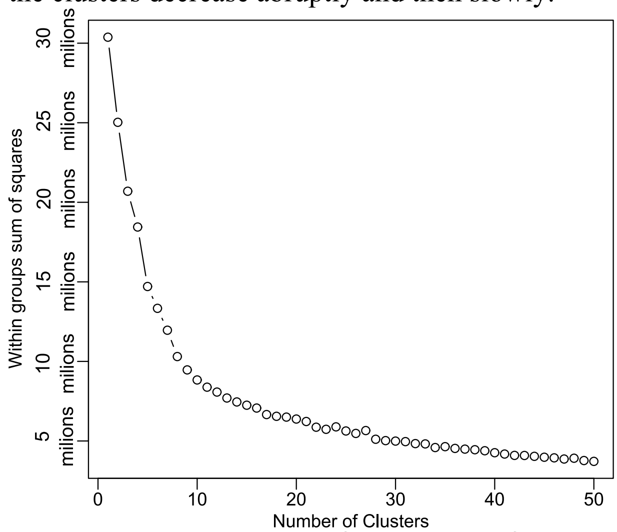

Figure 5 **The decrease of sum squares for increasing number of clusters/classes**

Region Growth method from SAGA GIS 2.0.7 (Bechtel et. al., 2008, Boehner, 2006) is based on the use of a seed point network from which centroid clusters are computed, and then the neighbor pixels are included iteratively or not in a spatial region based on the distance from centroid cluster. The results contain the mean value for the centroid. The most important aspect of this method is the choose of seed points. About this we believe that the use of centroids of MacMillan classification is a good approach.

Watershed segmentation (http://www.orfeo-toolbox.org) is based on a pixel by pixel analysis and performs segmentation on regions using either minima, or maxima to derive ridges in the image, and then merge the "basins" hierarchically, information which can be used further. SAGA GIS contain a watershed segmentation procedure based on maxima or minima, a threshold for watershed joining, outputs seed points of watersheds and seed value or ID for every watershed. The watershed segmentation can be applied only to an image layer, so the results can be mixed with other methods.

The geomorphometrical variables used for the unsupervised methods are slope (in degrees), exposition (aspect, in degrees, towards north direction) and normalized height. We choose these three geomorphometrical variables, because they are also very used in the classical methodology for land mapping unit delineation and the PCA analysis showed uncorrelation between them.

In the literature, the performance of the delineation, (Niekerk, 2010) was tested using aerial recognized points of natural breaks. Because we consider that this approach is not error free, this error can't be estimated, and actually it is hard to consider real reference data, we use a topographic profile (obtained using SAGA GIS 2.0.7 and R (Ihaka and Gentleman, 1996) from the study area and we compare the results of the used methods with the profile interpretation (the position and direction of the topographic profile is figured in *fig. 1*), to obtain some conclusions about the validity of used methods. Also we make a statistically analysis and discussion on the spatial distribution of the land mapping units and on the number and other characteristics of the land mapping units, for the same goal.

RESULTS AND DISCUSSIONS

The problem of land mapping unit delineation in land evaluation it is very complex. This is because these areas are seen as classes or as zones, both with spatial delineation (Pedroso et. al., 2010). Also, the criteria for their delineation are different, as a function of their final use, landforms for digital soil mapping, agro-ecological zones in agro-ecology or agricultural management zones in agricultural management. In our approach we used two types of delineation methods, the supervised classification algorithms which gave spatial classes with a priori landscape knowledge and unsupervised classification algorithms which gave "real" spatial repartition zones present in the data.

In fig. 6 the results for supervised and unsupervised classifications are represented for the Sârca test area. All the algorithms gave results spatially well aggregated and not very diffuse.

Slope x aspect classification has the most reduced spatial complexity, but has the biggest number of classes (24). Slope bigger than 24% do not appear in the study area, so instead of nine classes of slope we have only six classes. Also the landform pattern, with cuesta dipslopes, cuesta scarps and large floodplains without significant terraces, reduces the possibilities of class mixting.

Schmidt and Hewitt classification show a no so predictable repartition of the classes, because they are limited to the form, but there is a separation between floodplain, ridges and hillslopes.

Pennock and Reuter and MacMillan classification gave similar result, and we can consider the last as more complex, containing local types of landforms and performing the better segregation of landforms from ridge, hillslopes and floodplains. MacMillan classification is the most logical and readable classification from a physic-geographical point of view.

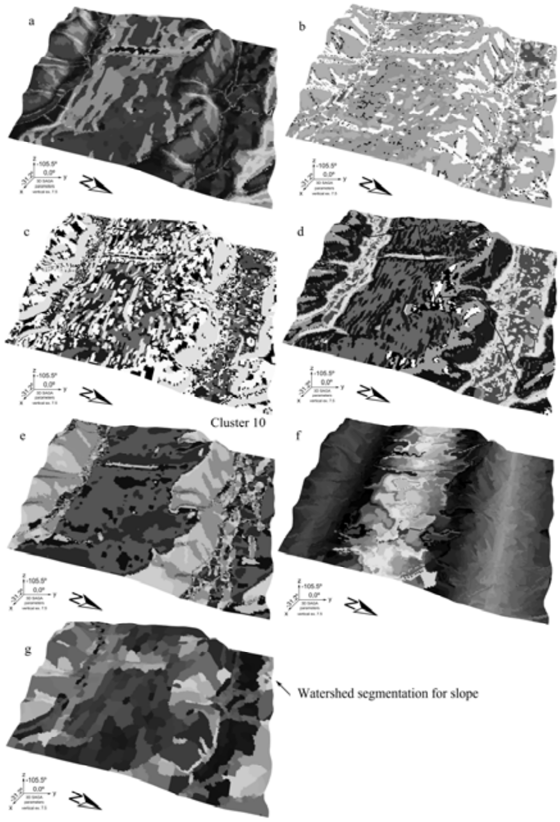

Figure 6 **Results of the supervised classifications: a – slope x aspect, b – Schmidt and Hewitt 2004, c – Pennock and Reuter, d – MacMillan 2000, e – cluster analysis, f – region growing and g – watershed segmentation** (a color version is available at www.geomorphologyonline.com/fig6.png)

The cluster analysis algorithm outputs mean values of the cluster centroids for every class and input layer. As the number of clusters increase, the standard deviation of the distance between centroids decrease, but the mean value of the centroids clusters increase. The majority of the clusters are located around the median of the range, only exposition shows clusters also near the first and fourth quartiles. We can conclude that a big number of classes are not preferable because we need a spatial aggregation of the clusters, but the results of 7 to 12 clusters are not significantly different.

Region growing results are very fragmented and cannot be interpreted.

Watershed segmentation for slope results are very similar with slope x exposition results.

In Fig. 8 and 9 the variation of the values of the geomorphometric parameters and geomorphometric classification are represented along the topographic profile from Sârca test area. The unsupervised classification cannot be represented in this form because every region/segment has a different ID.

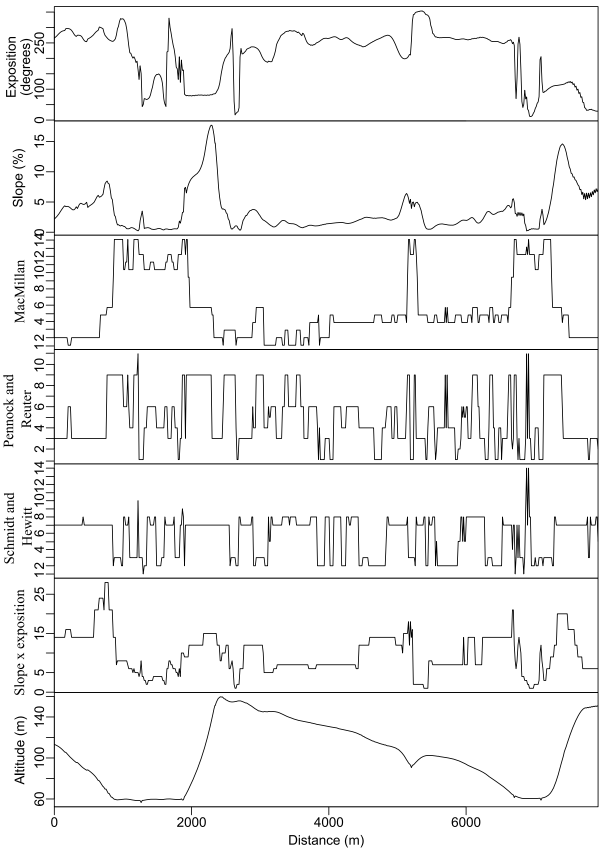

Figure 7 **Topographic profile superimposed with classification results**

CONCLUSIONS

The literature on the studied subject (Malczewski, 2004, Pedroso et. al., 2010, Soto and Pintó, 2010) shows that in general, segmentation algorithm gave best results, in comparison with classification results. Our conclusion is that both types of classification algorithms can be used for delineation of land mapping units for land evaluation inside soil land mapping units, as used in Romanian land evaluation methodology. The supervised classifications can be used also with the incorporated landscape knowledge. Cluster and segmentation results can be very hard to interpret from a pedological and land evaluation point of view.

The possibility of applying fuzzy set theory in the problem of land mapping units delineation is to be considered, especially because we don't have real data for validation.

ACKNOWLEDGMENTS

This work was supported by the European Social Fund in Romania, under the responsibility of the managing Authority for the Sectoral Operational Programme for Human Resources Development 2007-2013 [grant POSDRU/88/1.5/S/47646 and grant POSDRU/89/1.5/S/49944].

BIBLIOGRAPHY

Bohner, J., Selige, Th., Ringeler, A., 2006 - *Image segmentation using representativeness analysis and region growing*, Gottinger Geographische Abhandlungen, 115:29-38.

Bechtel, B., Ringeler, A, Boehner, J., 2008 - *Segmentation for object extraction of trees using MATLAB and SAGA*, in: Boehner J., Blaschke J., Montanarella L. eds. SAGA – Seconds Out. Hamburger Beitrage zur Physischen Geographie und Landschaftsoekologie, 19:59-70.

Davidson, D., 2002 – *The assessment of land resources: achievements and new challenges*, Australian Geographical Studies, 40:109-128.

Hengl, T., 2006 – *Finding the right pixel size*, Computers & Geosciences, 32:1283-1298.

Hengl, T., 2009 – *A practical guide to geostatistical mapping*, http://www.lulu.com, ISBN 978-90-9024981-0, 270 p.

Ihaka, R., Gentleman, R.R., 1996 – *A language for data analysis and graphics*, Journal of Computational and Graphic Statistics, 5:299-314.

MacMillan, R.A., Pettapiece, W.W., Nolan, S.C., Goddard, T.W., 2000 – *A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic*, Fuzzy Sets and Systems 113:81-109.

Malczewski, J., 2004 – *GIS-base land-use suitability analysis: a critical overview*, Progress in Planning, 62:3-65.

Niekerk, van A., 2010 – *A comparison of land unit delineation techniques for land evaluation in the Western Cape*, South Africa, Land Use Policy, 27:937-945.

Pedroso M., Taylor J., Tisseyre B., Charnomordic B., Guillaume S., 2010 – *A segmentation algorithm for the delineation of agricultural management zones*, Computers and Electronics in Agriculture, 70:199-208.

Pennock, D., Zebarth, B., DeJong, E., 1987 – *Landform classification and soil distribution in hummocky terrain*, Geoderma 40:297-315.

Quick, R – *Cluster analysis*, http://www.statmethods.net/advstats/cluster.html

Reuter, H.I., Wendroth, O., Kersebaum, K., 2006 – *Optimisation of relief classification for different levels of generalization*, Geomorphology 77:79-89.

Schmidt, J., Hewitt, A., 2004 – *Fuzzy land element classification from DTMs based on geometry and terrain position*, Geoderma 121:243-256.

Soto, S., Pintó J., 2010 - *Delineation of natural landscapes units for Puerto Rico*, Applied Geography, doi:10.1016/j.apgeog.2010.01.010.

Venables, W.N., Ripley, B.D., 2002 – *Modern applied statistics with S*, Springer, 495 p.

Wood, J., 1996 – *The geomorphological characterization of digital elevation models*, PhD thesis, University of Leicester.

*****, 1977** - *FAO – A framework for land evaluation*, ILRI, Wageningen, 87 p.

*****, 2011** - *GRASS Development Team*, Geographic Resources Analysis Support System (GRASS) Software. Version 6.4.1 Open Source Geospatial Foundation. http://grass.osgeo.org.

*** - **Image segmentation** – http://www.orfeo-toolbox.org/SoftwareGuide/SoftwareGuidech16.html.

*****, 1987 -** *Institute for Soil Science and Agrochemistry***,** *Methodology of pedologic studiea elaboration, II[nd] part – Elaboration of pedological studies for various scopes*, 349 p. (in romanian)

*****, 1978 -** *Institute for Soil Science and Agrochemistry, Work instructions for agricultural land evaluation with elements of pedologic fundamentation and technologic characterization*, Bucharest, 145 p. (in romanian).

*****, 2011** - *SAGA GIS 2.0.7-* http://www.saga-gis.org/