



# Clustering algorithm in ILWIS GIS for classification of Landsat TM scenes: case study of Mecsek Hills region, Hungary

Polina Lemenkova, István Elek

## ► To cite this version:

Polina Lemenkova, István Elek. Clustering algorithm in ILWIS GIS for classification of Landsat TM scenes: case study of Mecsek Hills region, Hungary. Geosciences and Environment. Near-Surface Geophysics, May 2012, Belgrade, Serbia. 10.6084/m9.figshare.7434218.v1 . hal-02000222

**HAL Id: hal-02000222**

**<https://hal.science/hal-02000222>**

Submitted on 31 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC0 - Public Domain Dedication 4.0 International License

## Clustering algorithm in ILWIS GIS for classification of Landsat TM scenes: case study of Mecsek Hills region, Hungary.

Lemenkova Polina (\*), Elek István.

### 1. Summary

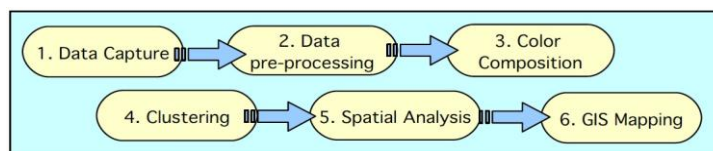
The emphasis of this research is application of spatial analysis using clustering algorithm, and remote sensing data (Landsat TM imagery) for agricultural mapping of land cover types. The study area covers Mecsek Hills region, located in south-western Hungary. This region is characterized by high land heterogeneity and complex landscape structure, caused by intense agricultural land use in the region with mixed vegetation type and high environmental value. The research data consists in Landsat TM scenes taken for years 1992, 1999 and 2006. The methodology is based on cluster classification algorithm available in ILWIS GIS. Based on clustering technique, the agricultural land cover classes were identified by association of pixels on the Landsat TM scenes to thematic clusters. Different land use types were classified, which include natural vegetation coverage, anthropogenic areas and agricultural fields, sub-divided to various crop types. Once classification was complete, agricultural thematic maps have been created. The final research output consists in three independent agricultural thematic maps of land cover types for years 1992, 1999 and 2006.

### 2. Introduction

Combination of the satellite images with GIS techniques is a key method for land patterns identification and classification of agricultural land cover types on the land surface. In this paper an ILWIS GIS approach to perform remote sensing based land cover analysis and mapping is presented. The purpose of the current work is application of the ILWIS GIS method of clustering, which assigns pixels with similar value of Digital Numbers (DNs) to clusters, or thematic classes. The main objective is to perform spatial analysis on distribution of land cover patterns in study area. The research data consist in Landsat TM scenes visualizing agricultural areas in Mecsek Hills region in 1992, 1999 and 2006 years. The choice of the Landsat TM data is explained by their well-known advantages of application in land cover mapping and agricultural cartography, free availability, and suitability for spatial analysis according to temporal and spatial parameters (30m mesh size).

### 3. Methods

The research methodology is based on the spatial analysis tools of the open source ILWIS GIS software aimed at classification of Landsat TM scenes. Current work is organized in several research steps summarized in the research workflow (Fig.1). Methods of data processing include import and pre-processing of the images, enhancement, classification, spatial analysis and interpretation. The research aim is creation of land cover maps, performed by classifying research area into a set of land cover categories, and assigning a land use type label to each of these units.



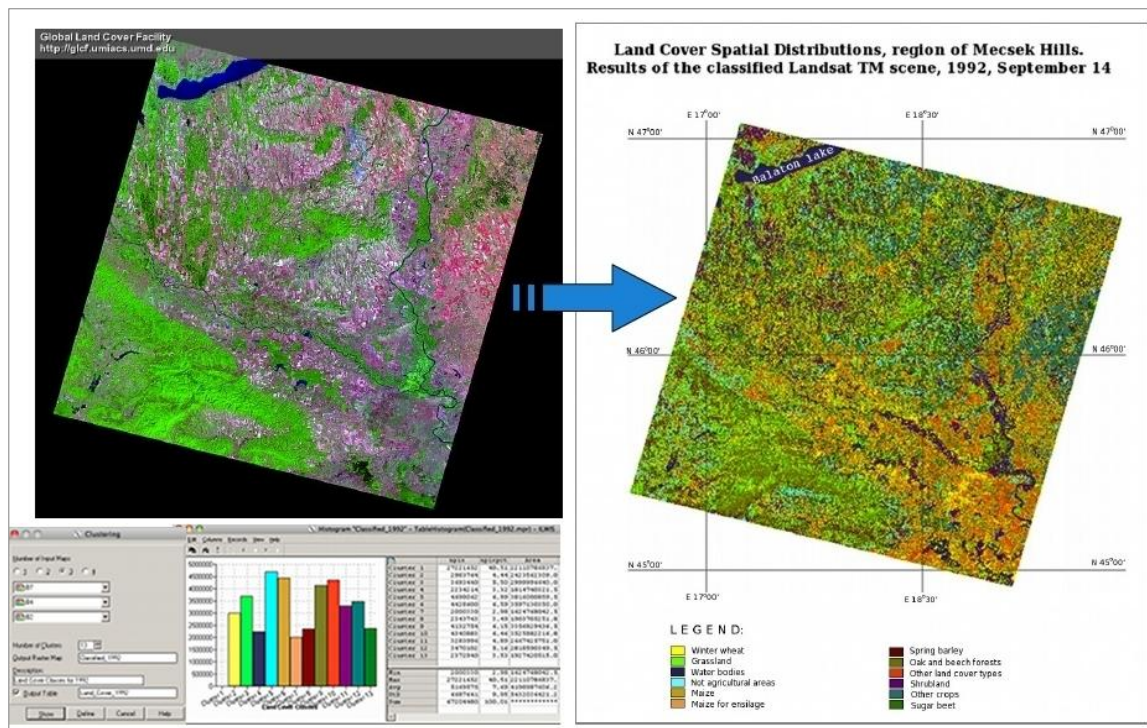
*Figure 1 Research flow algorithm.*

The core method used in the current work for the interpretation of imagery is clustering algorithm, which is based on a ground principle of distinguishability of the spectral signatures of the pixels on the raster image. The clustering was chosen, because it is objective technique, enabling to avoid subjectivity and misclassified pixels, and to ignore external factors while classifying, e.g. atmospheric conditions, solar illumination, etc. This method is based on the remote sensing general principle that

each unique pixel on a multichannel image has spectral signature defined by the reflectance of its DN in each spectral band. The discernible DNs of pixels create unique signatures for various objects, distinguishable from other objects. Multispectral cluster classification is an iteration process which extracts information about values of the pixels DNs, by analyzing their spectral signatures. During the classification process the digital cells are measured according to the similarity of their DNs and assigned into a small number of categories, or clusters. (Jensen, 2007). Existing methods of image processing, described in numerous reports of Landsat application for land cover mapping (e.g. Knorn et al., 2009; Wulder et al. 2008; Julien et al. 2011) were considered and applied in the current work.

#### 4. Examples

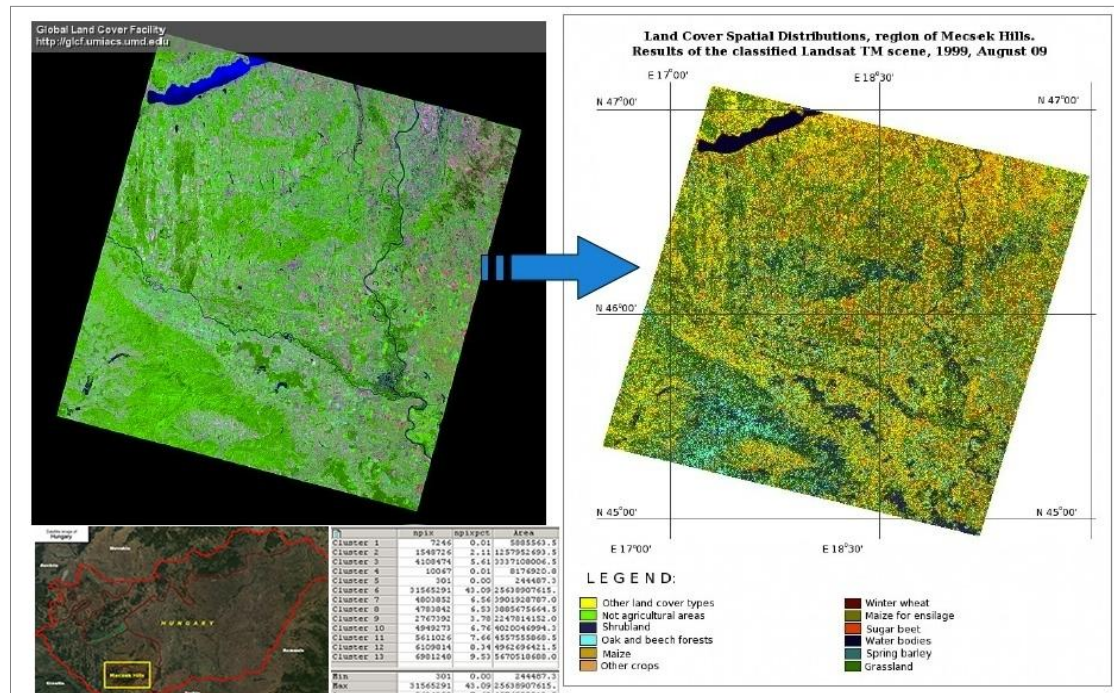
In the scope of current work, the target research task was to identify pixels which belong to different land categories and thus to identify land cover classes, which is done in classification process. The clustering is based on clustering operation in ILWIS, performed in semiautomated regime.



**Figure 2** Original image Landsat TM scene taken on 1992, September 14 (left, above), and the map showing classification results (right), performed using Cluster method in ILWIS: land cover classes in Mecsek area. Below: Clustering techniques in ILWIS GIS; Histogram showing percentage of land cover classes (1992).

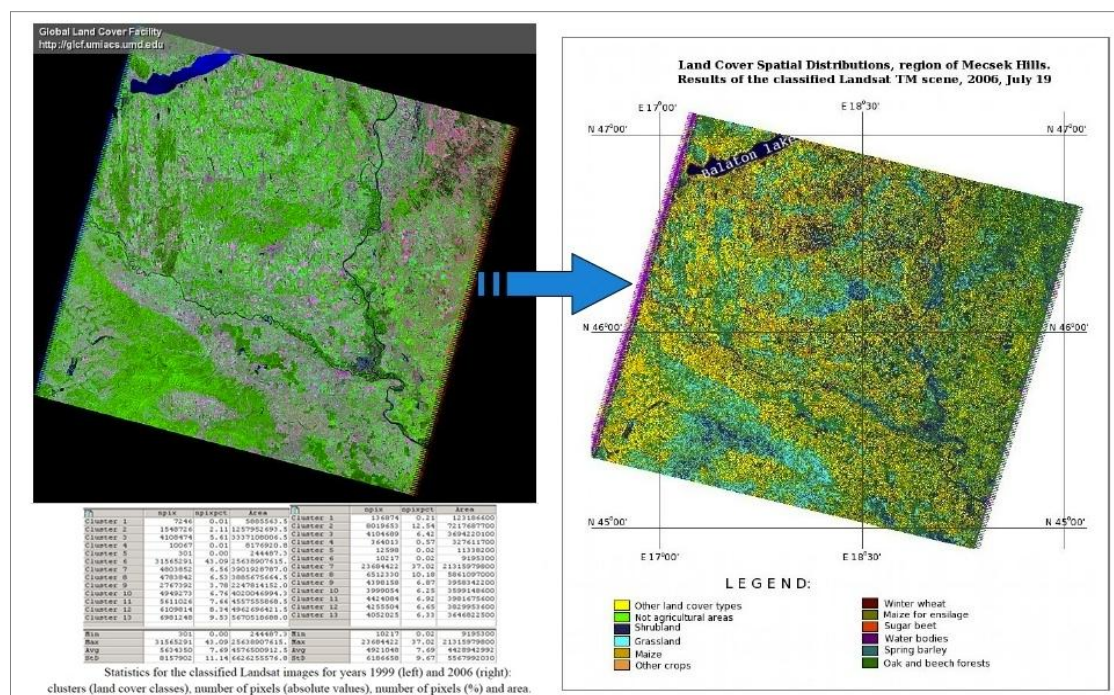
The Landsat data covering Mecsek Hills region (45°00'N–47°00'N; 17°00'E–19°00'E) with time span of 14 years (1992, 1999 and 2006) were acquired in GeoTIFF format from the Earth Science Data Interface website. The images were taken in growing season (summer and early autumn), to enable best distinguishability of spectral signatures in diverse agricultural areas and vegetation (e.g. small leaves, mature leaves or ripe ears). The pre-processing of the imagery included data import and visual settings of colors and contrast enhancement. The images were stored as raster file in .img format, then imported into raster map format (ASCII). After converting, each image contained collection of 7 Landsat raster bands. The image enhancement of the Landsat bands was performed using Edge Enhancement linear filter, in order to improve quality of image and visibility of lines. To sharpen the image and improve color contrast, stretching filter was applied to the bands.





**Figure 3** Landsat TM scene (09 August 1999), left, above, and the agricultural classification map (right). Below: the research area of Mecsek Hills, south-western Hungary; Clustering statistics for scene 1999.

The grouping of the data is done according to the correspondence principle: a similarity of signatures of pixels within a group is larger than among other groups.



**Figure 4** Landsat TM scene (19.07.2006), left, above; agricultural map of land cover classes (right). Below: Statistics on pixels classification during clustering procedure for images dated 1999 and 2006.

As a result, pixels are grouped into cluster classes, which significantly facilitates spatial analysis of the images. For the land cover identification we applied thematic classes developed and accepted by Csornai et al. (2008). The pixels were identified for each of the categories and grouped into following

land categories: 1) winter wheat 2) spring barley 3) maize 4) maize for ensilage 5) oak and beech forests 6) sugar beet 7) 8) other crops 9) shrubland 10) water bodies 11) not agricultural areas 12) grassland 13) other land cover types. The pixels on the scenes were identified for each of the agricultural land cover classes, and assigned into thematic categories, or clusters, according to the statistical values of their spectral signatures. The statistical information on the output clusters, such as average, predominant and extremal value of each cluster in all three bands, is stored in an attribute table, created for the output map. Thus, land cover classes were created, in semi-automated regime based on the assignation of the image pixels to clusters forming the appropriate land cover classes. The outcome consisted of land use classes for each of the Landsat TM scenes (Fig.2, Fig.3 and Fig.4). Crops that have well detected fields, such as “maize”, “winter wheat”, “spring barley”, “sugar beat” and “maize for ensilage” classes have been easily identified and recognized on the images, as crop fields have distinctive colors. However, in some cases it was difficult to distinguish the nature of the crop species (light beige and various shadows of grey color). Such classes we defined as “other crops” and “other agricultural areas”. Since there were no fieldwork available in the study area, a Google Earth aerial imagery was used for visual detailed control inspection of some regions of the study area. The distribution of clusters is shown as a frequency information on pixels' values and classes. Once all clusters were grouped, the mapping legend was created. The image was visualized using Representation Palette, previously defined in the Domain “Land Cover Types”. Finally, the layout was created, which included three thematic agricultural maps (Fig.2, Fig.3 and Fig.4).

## 5. Conclusion

To conclude, it should be underlined that ILWIS GIS is a convenient open source GIS useful for spatial analysis and agricultural mapping due to its functionality, flexible interface and open source availability. Clustering method can be successfully applied for agricultural mapping, since it enables objective identification of the land cover types in regions characterized by high land heterogeneity and complex structure, such as agricultural fields mixed with natural land cover types. The experience of Landsat TM imagery processing by means of ILWIS, described in the current work, is a contribution towards agricultural mapping, and a report of technical functionality of the program. The methods used in this work can be tested on other agricultural areas and regions.

## 6. Acknowledgements

We thankfully acknowledge the financial support of this research, provided by the Balassi Institute, HSB (Hungarian Scholarship Border), Budapest Hungary. Reference No. MÖB/154-2/2011.

## References

- Csornai G, Wirnhardt Cs., Suba Zs., Nádor G., Tikász L., Martinovich L., Kocsis A., Zelei Gy., László I., Bognárornai E. 2008. Cropmon: Hungarian Crop Production Forecast by Remote Sensing. ISPRS Archives XXXVI-8/W48 Workshop proceedings: Remote sensing support to crop yield forecast and area estimates.
- Jensen, J.R. 2007. Remote Sensing of the Environment: An Earth Resource Perspective, 2nd Edition. ISBN-10: 0131889508. Prentice-Hall, Inc.: Upper Saddle River, New York.
- Julien Y., Sobrino J.A., Jiménez-Munoz J.-C. 2011. Land use classification from multitemporal Landsat imagery using the Yearly Land Cover Dynamics (YLCD) method. International Journal of Applied Earth Observation and Geoinformation 13, 711–720.
- Knorn J., Rabe A., Radeloff V.C., Kuemmerle T., Kozak J., Hostert P. 2009. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. Remote Sensing of Environment 113, 957–964.
- Wulder M., White J.C., Goward S.N., Masek J.G. , Irons J.R., Herold M., Cohen W.B. , Loveland T.R., Woodcock C.E. 2008. Landsat continuity: Issues and opportunities for land cover monitoring. Remote Sensing of Environment 112, 955–969.