

## AGRICULTURAL LAND COVER CLASSIFICATION USING RAPIDEYE SATELLITE IMAGERY

**Olguța MORARIU, Florina BURESCU**

**Scientific Coordinators: Assoc. Prof. PhD Eng. Mihai Valentin HERBEI,  
Assoc. Prof. PhD Codruta CHIS**

Banat University of Agricultural Sciences and Veterinary Medicine, “King Michael I of Romania”,  
Timisoara, Calea Aradului 119, 300645, Timișoara, Romania, Phone: +40256 277280,  
Fax: +40256 200296, Email: mihai\_herbei@yahoo.com

Corresponding author email: morariu.olguta@yahoo.com

### **Abstract**

*Remote sensing is of great interest for the study and characterization of the vegetation and of the agricultural crops, in order to monitor them and to develop predictable patterns regarding the evolution of the crops and also for the purpose of the decision making process in real time. This paper aims to perform a qualitative and quantitative analysis between the supervised and unsupervised methods of classification of satellite imagery. In this context, a Rapid Eye satellite image was used at a 5m resolution taken from the www.planet.com platform from May 2016, which underwent complex preprocessing and processing operations to retrieve useful information about the vegetal cover. The studied area is within the Experimental Didactic Station of BUSAMV Timisoara. Supervised classification was based on the Maximum Likelihood algorithm and the unsupervised was based on the ISO DATA algorithm. The Rapid Eye scene was processed with ArcGIS v. 10.5 software.*

**Key words:** Algorithms, Classification, ISODATA, Maximum Likelihood, Remote sensing.

### **INTRODUCTION**

All methods based on satellite techniques are very often used in the analysis and characterization of the terrain and the vegetal cover (Gitelson 2004, Thenkabail et al., 2007, Herbei et al., 2015, Sala, 2011).

The characterization of the vegetal cover can be done both on the basis of the specific absorption, reflectivity and emitting properties (Sala et. al., 2016), as well as on the basis of certain indices calculated from the spectral information contained in the satellite imagery (Thenkabail et al., 2002, Huang et al., 2013, Herbei et al., 2015).

The classification of satellite imagery, based on analysis and group distribution of pixels with similar spectral properties for brightness value, provides high image analysis accuracy and implicitly represented areas (Akgün et al., 2004).

The constellation of five RapidEye satellites differs from other geospatial satellite providers in their unique ability to acquire daily high-resolution, high definition image data (<https://www.satimagingcorp.com/satellite->

[sensors/other-satellite-sensors/rapideye/](https://www.satimagingcorp.com/satellite-sensors/other-satellite-sensors/rapideye/)).

RapidEye imaging capabilities can be applied to a range of industries such as agriculture, forestry, oil and gas exploration, engineering and construction, governments, cartography and mining. The RapidEye system collects 4 million square kilometers of data per day at a nominal ground resolution of 6.5 meters. Each satellite has been designed for a mission life of at least seven years. All five satellites are equipped with identical sensors and are located in the same orbital plane. The satellite images that form the basis for generating action plan maps, if used in the background of intelligent cadastral vector data, can improve the details of thematic maps as well as action plan maps. It also helps to monitor land cover changes (Smuleac et. al., 2013) that can be identified by detailed procedures for detecting changes and implemented within the GIS mapping project (Smuleac et. al., 2016).

Because the RapidEye system was designed to serve the agricultural and forestry markets, the Red Edge tape was chosen as part of the RapidEye Spectral Band Set. Over the years, several studies have confirmed the value of the

Red Edge band. The red band is located spectrally between the red band and the NIR band without overlapping. In a typical spectral response of green vegetation, the red ribbon covers the spectrum portion where the reflection increases drastically from the red portion to the NIR plateau.

The RapidEye multispectral system acquires images in 5 spectral bands, each with a different pixel resolution (or GDS ground sampling distance) of 6.5 m at the top, as shown in the following table:

Table 1: RapidEye Satellites Spectral Bands

Band	wavelength
Blue	440-510nm
Green	520-590nm
Red	630-685nm
Red Edge	690-730nm
Near-Infrared	760-850nm

Many studies have suggested that the Red Edge band is capable of providing additional information to identify plant types, nutrition and health, and to characterize plant coverage and abundance among other characteristics (Kross et. al., 2015, Herbei and Sala, 2016).

## MATERIAL AND METHOD

The study area is located in the western part of Romania, in Timisoara and represents a part of the BUSAMV Timisoara Experimental Station (Figure 1), which we want to study. The land spreads over approximately 1,300 ha.

From the [www.planet.com](http://www.planet.com) database a Rapid Eye scene was downloaded at a 5m resolution from May 2016 and contains the 5 spectral bands: Red, Green, Blue, Red Edge and NIR. This scene was analyzed in a digital environment (ArcGIS 10) (Herbei, 2015), and underwent complex mathematical operations (Herbei and Sala, 2014).

This research has used technology based on satellite imagery for classification of land cover.

By classifying satellite imagery, the analyst tries to classify the features of objects or phenomenon in a satellite image using visual interpretation elements to identify homogeneous pixel groups that represent different features or classes of field coverage in the area of interest.

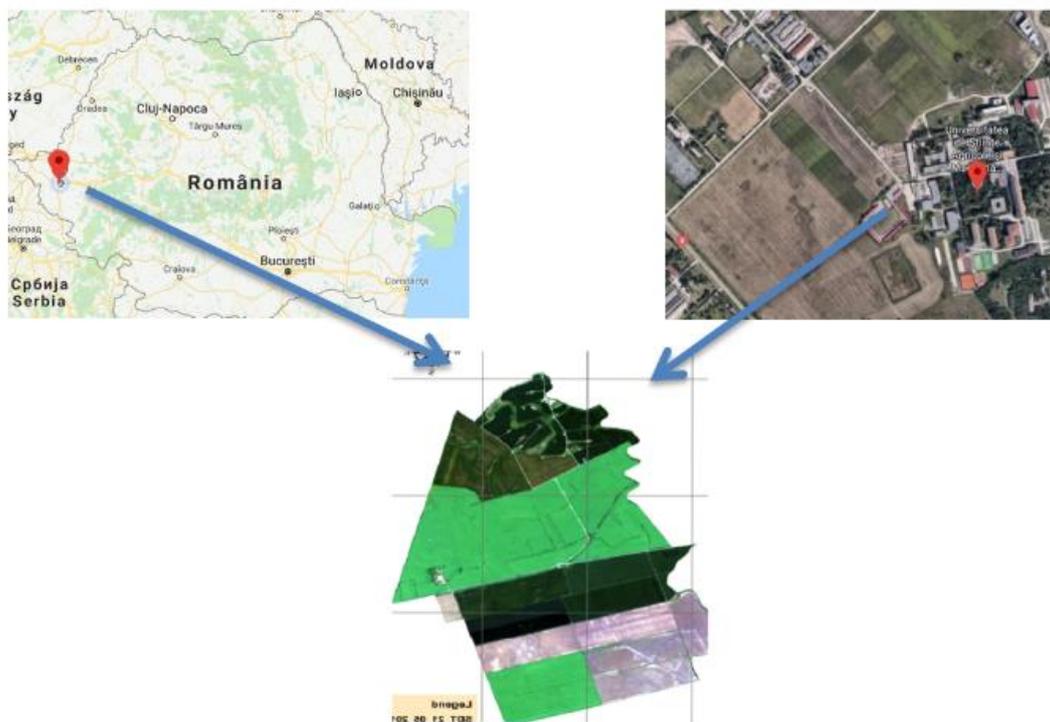


Figure 1. Area of interest

In the classification of digital images (Kokalj and Oštir, 2007), the spectral information represented by digital numbers in one or more spectral bands is used, and each pixel is framed based on this spectral information, the main objective being to assign all the pixels in the image to certain classes or themes, for example: water, coniferous forest, deciduous forest, corn, wheat, etc (Kumar, 2004). The spectral classification aims at objectively mapping areas of the image that have radiometric (reflective and / or emissivity) similar characteristics. Depending on the purpose and the required accuracy, two types of classification are used:

- unsupervised (automated by statistical means)
- supervised (spectral classes are associated with some characteristics identified by photointerpretation on the image) (Iosub, 2012).

In the supervised classification of digital image, field object objects are known in advance on certain restricted areas of the image (areas called test areas or sites). These areas fall into patterns, and then rules are developed to be extended to the unknown portions of the image. In other words, the user identifies some areas on the image that are characteristic of each set of established details.

Maximum likelihood classification, equation (1), assumes that the statistics for each class in each band are normally distributed and calculates the probability that a given pixel belongs to a specific class. Unless you select a probability threshold, all pixels are classified. Each pixel is assigned to the class that has the highest probability (that is, the maximum likelihood). If the highest probability is smaller than a threshold you specify, the pixel remains unclassified (Richards, 1999).

$$g_i(x) = \ln p(\omega_i) - 1/2 \ln |\sum_i| - 1/2(x - m_i)^T \sum_i^{-1} (x - m_i) \quad (1)$$

where:

i = class

x = n-dimensional data (where n is the number of bands)

p(ωi) = probability that class ωi occurs in the image and is assumed the same for all classes

|Σi| = determinant of the covariance matrix of the data in class ωi

Σi-1 = its inverse matrix

mi = mean vector

### 1.1.1. Unsupervised Classification

In this approach, the computer analyzes all the spectral signatures of all the pixels in the image and identifies areas with pixels that have similar values. The user may impose classification criteria (number of classes, number of iterations) to allow for homogeneous and differentiated groups. The algorithms underlying the automatic classification depend primarily on the spectral characteristics of the pixel, unlike the size, texture and shape, which are the characteristics of the objects on which the visual interpretation is based.

The **ISO DATA** algorithm, developed by Geoffrey and Hall (1965), uses the minimum spectral distance formula for cluster formation and clustering, based on the Euclidian distance equation (Swain and Davis, 1978; Melesse and Jordan, 2002).

$$SD_{xye} = \sqrt{\sum_{i=1}^n (\mu_{ci} - X_{xyi})^2} \quad (2)$$

where:

n - number of bands,

i - band number;

c - particular class

$X_{xyi}$  - data file value of pixel x, y in band i;

$\mu_{ci}$  - mean of data file values (digital numbers) in and i for the sample the class c;

$SD_{xye}$  - spectral distance from pixels x, y to the mean of class c.

## RESULTS AND DISCUSSIONS

For the analysis of the territory and the vegetal cover, the classification process involved the grouping of pixels in the composition of digital images in terrain coverage classes in relation to the spectral properties. The analysis of the territory and the structure of the agricultural crops was based on the Rapid Eye satellite images and the ArcGIS 10 software. Maximum Likelihood and ISODATA algorithms were used for this purpose.

The present study used satellite imagery obtained with the RapidEye satellite system. The first map is made by combining three spectral bands, 3-2-1 (Red-Green-Blue).

This color composition is the closest to real colors, those perceived by the human eye. Ensures easy recognition of environmental components, for example, healthy vegetation is green, roads are gray (Figure 2).

It is noticed that there are more shades of green on the plot, indicating that there are several types of crops. Green tones depend on vegetation reflection (Rouse et. al., 1974). This combination of bands is, however, inadequate in the analysis of vegetation due to the absence of an infrared band in which the vegetation has

the maximum reflection.

The second map is derived from the combination of Spectral Strips 5-3-2 (NIR-Red-Green). It's a false-color image which makes the colors on the image no longer match those perceived by the human eye for certain environmental components. For example, vegetation is easily detected, which is identified by different shades of red, depending on the plant type.

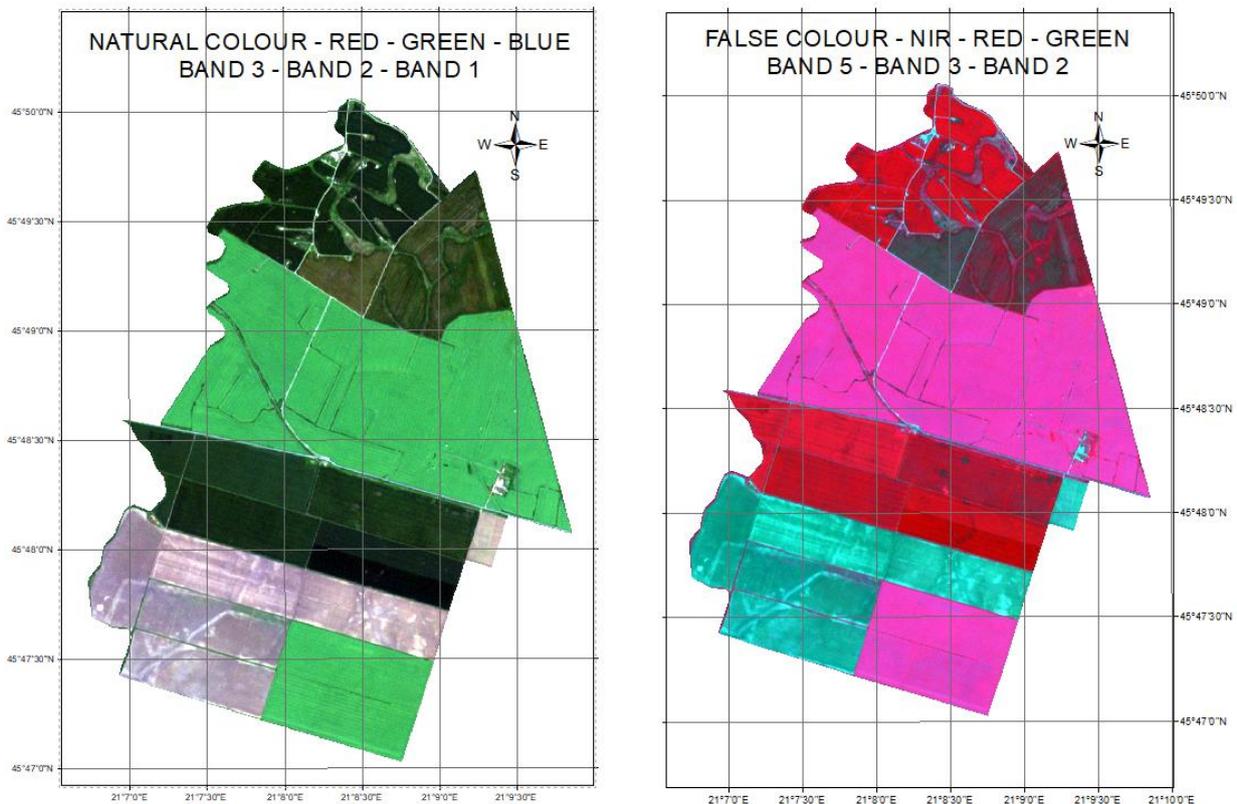


Figure 2. The Study Area in natural colors and false-colors

Because vegetation has maximum reflectivity in the infrared, the description of the area of interest through the two classification methods (supervised and unsupervised) is based on the false-color satellite image obtained by combining the 5-3-2 bands.

For the supervised classification of the satellite image, the Maximum Likelihood algorithm will be used in the paper. This algorithm is based on two principles:

- The cells of each class within the multidimensional space are normally distributed;
- Bayes decision making theorem.

The tool considers both class signature variations and covariances when assigning each cell to one of the classes represented in the signature file. Assuming that the distribution of a class sample is normal, a class can be characterized by the average vector and covariance matrix.

Considering these two characteristics for each class of cells, the statistical probability is calculated for each class to determine cell membership in the class. When specifying the default equal option for a priori weighting,

each cell is assigned to the class at which it is most likely to be a member.

Once class statistics are defined, image samples are ranked according to distance from classroom. Each sample is assigned to the class with the minimum distance. The distance itself is reduced according to the maximum rule of Bayes probability.

Practically, the Maximum Likelihood algorithm:

- It's based on probability
- Assume that all classes have a normal distribution (Gaussian)
- Spectral variance and co-variance is denoted for each class
- Classes can be statistically modeled using an average vector or co-variance matrix
- Each pixel is assigned to the class with the highest probability of membership

Three stages are distinguished in the supervised classification process:

- training stage - in which the selection of the features best describing the object / phenomenon (Figure 3)
- the stage of the actual classification (decision stage) - in which the appropriate method of comparison of the training elements (from the previous stage)
- Conformity assessment of the classification (final stage).

The training stage is the stage in which areas of interest are selected, groups of pixels with homogeneous geographic features

(construction, water, forests, etc.) and a digital number with close values, of great importance in terms of quality classification.

Areas of interest are created by manual digitization, forming polygons of representative pixel groups with a best distribution within the image (to highlight some qualitative differentiations that may occur within a class).

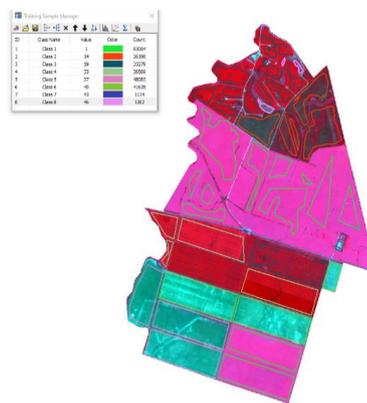


Figure 3. Supervised classification – Training stage

The analyst's experience and analytical capacity are very important, and these traits depend on the final outcome of the classification.

Once areas of interest have been established, classification can be made based on specific algorithms (Figure 4).

The "decision", as this stage is called, refers to the choice of the optimal algorithm, on which the quality of the result obtained depends.

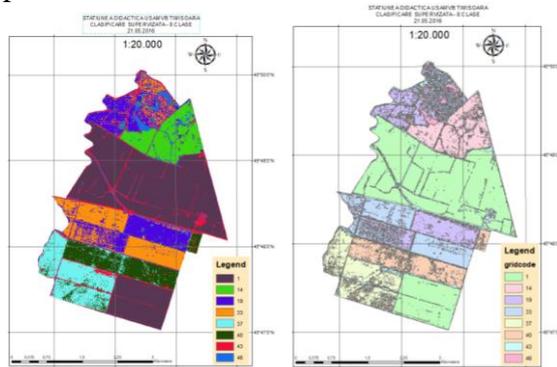


Figure 4. Supervised classification – Raster and vector

In this paper, for the unsupervised classification of maps obtained through the RapidEye satellite system, the ISODATA classification algorithm was used. This ISODATA algorithm does not require prior

knowledge of the area studied by the operator. It also does not need to know the number of clusters because the algorithm divides and merges the clusters according to user-defined threshold values for the parameter, and the

algorithm runs more iterations on the computer until the thresholds are reached.

This algorithm works in the following way:

- Cluster centers are randomly placed and pixels are allocated based on the shortest distance from the center method;
- The standard deviation in each cluster and the distance between the cluster is calculated by the centers
- The next iteration is done with the new cluster centers.
- Other iterations are performed up to:
  - Intermediate mean distance drops below the threshold defined by the user,
  - average change in center the distance between iterations is less than the threshold, or reached

- the maximum number of iterations is reached
- Clusters associated with less than the minimum number of pixels specified by the user are deleted and the single pixels are either reclassified and placed in another class, or ignored and named as unclassifiable.

Practically, when the Iterative Self Organizing Data Analysis algorithm (ISODATA) is used the standard deviation for clusters is calculated and after n iterations:

- Classes can be combined
- Classes can be divided
- Classes can be deleted

After each iteration, the percentage of pixels remaining in the same class is calculated (Figure 5).

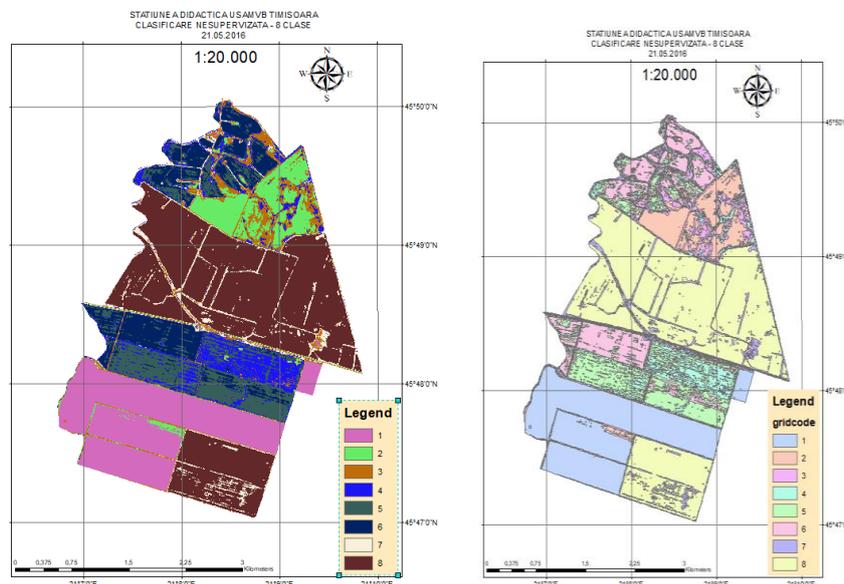


Figure 5. Unsupervised classification – Raster and vector

## CONCLUSIONS

The imaging analysis based on Rapid Eye satellite imagery through supervised and unsupervised methods has made it easier to classify land and crops in the studied area with high precision. Methods can be promoted and used to monitor crops and land use.

From the analysis of the satellite imagery through the methods used, the surfaces of the cultures in the Didactic Resort of the BUSAMV Timisoara were obtained with high precision. The crop structure and cultivated areas are shown in the Figure below.

Following the classifications, there is a difference in the measured area (Figure 6, Figure 7).

Row Labels	Sum of ARIA	Row Labels	Sum of ARIA
1	497,7480015	1	253,5776361
14	109,8061976	2	94,65912859
19	179,4067289	3	80,77551468
33	158,1682415	4	89,43129845
37	128,9554513	5	138,5263703
40	117,4336492	6	124,8246733
43	90,58742947	7	65,48850607
46	47,40246185	8	482,2618484
(blank)		(blank)	
Grand Total	1329,508161	Grand Total	1329,544976

Figure 6. Supervised and Unsupervised areas in hectares

**Unsupervised classification** has the following advantages and disadvantages:

### Advantages:

- Does not require knowledge of the analyzed region
- Human errors are minimized

- Pixels are spectrally separated
- The analyst has control over the number of classes, iterations, etc.

#### Disadvantages:

- The grades obtained do not accurately reflect the categories in the field
- Limited control over class identity
- Spectral properties of classes can change over time

**Supervised classification** has the following advantages and disadvantages:

#### Advantages:

- Class identity control
- Specific classes with certain identities
- The grades obtained reflect the categories in the field
- Classification errors can be detected

#### Disadvantages:

- The analyst requires classification
- Regions of interest are generally related to field information and not to spectral properties
- Interstitial regions may not be representative
- Getting the regions of interest calls for time.

For security, it is necessary to analyze and process digital images to make pixel corrections in line with field reality, with some alterations being generated by the climatic conditions at the time of taking the images.

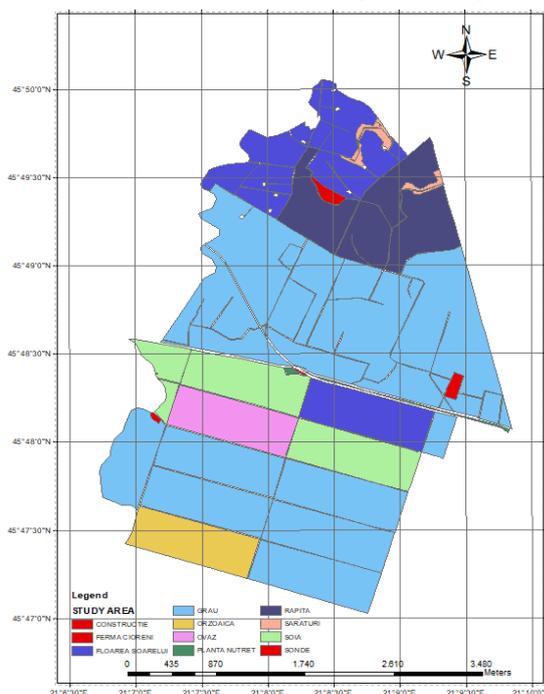


Figure 7. Structure of crops and cultivated areas

## ACKNOWLEDGEMENTS

The authors thank to the staff of the Didactic and Experimental Station of the Banat University of Agricultural Sciences and Veterinary Medicine "King Michael I of Romania" from Timisoara, Romania to facilitate this research.

The authors also thanks to the GEOMATICS RESEARCH LABORATORY and to the GIS Scientific and Remote Sensing Student Circle from BUASMV "King Michael I of Romania" Timisoara.

## REFERENCES

- Akgün A, Eronat AH, Türk N., 2004. Comparing different satellite image classification methods: an application in Ayvalık District, Western Turkey. In: The 4th International Congress for Photogrammetry and Remote Sensing, Istanbul, Turkey.
- Geoffrey H.B. and Hall D.J, 1965 ISODATA, a novel method of data analysis and pattern classification, Stanford Research Institute Menlo Park, California, Technical Report April 1965, for Information Sciences ranch Office of Naval Research Contract Nonr 4918(00) SRI Project 5533.
- Gitelson A.A., 2004. Wide dynamic range vegetation index for remote quantification of crop biophysical characteristics. *J. Plant Physiol* 161(2), p. 165–173.
- Herbei M., Sala F., 2014. Using GIS technology in processing and analyzing satellite images – case study Cheile Nerei Beusnița National Park, Romania. *Journal of Horticulture, Forestry and Biotechnology*, 18(4), p. 113-119.
- Herbei M., Sala F., Boldea M., 2015. Using Mathematical Algorithms for Classification of LANDSAT 8 Satellite Images. *ICNAAM 2014, American Institute of Physics Conf. Proc.* 1648, 670004.
- Herbei, M.V., 2015. GIS si Modelare cartografica, Universitas, p.151-160, Petroșani.
- Herbei, M.V., Sala, F., 2016 - Biomass prediction model in maize based on satellite images. In *AIP Conference Proceedings*, Vol. 1738, No. 1, 350009, AIP Publishing.
- Iosub, F., 2012, Extragerea informației spațiale din imaginis atelitare folosind clasificările supervizate și nesupervizate, [www.geo-spatial.org](http://www.geo-spatial.org).
- Kokalj Ž., Oštir K., 2007. Landcover mapping using Landsat satellite image classification in the Classical Karst- Kras region, *Acta Carsologica*, 36/3 p. 433-440.
- Kross, A., McNairn, H., Lapen, D., Sunohara, M., and Champagne, C., 2015. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops, *International Journal of Applied Earth Observation and Geoinformation*, 34, (0), p. 235-248.

- Kumar R.T.U., 2004. Geographic Resources Decision Support System for land use, land cover dynamics analysis. Proceedings of the FOSS/GRASS Users Conference. 12-14 September 2004, Bangkok, Thailand.
- Melesse A.M., Jordan J.D., 2002. A comparison of Fuzzi vs. Augmented-ISODATA classification algorithms for cloud-shadow discrimination from Landsat images, Photogrammetric Engineering & Remote Sensing. 68(9) p. 905-911.
- Planet Team (2017). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. <https://api.planet.com>.
- Richards JA., 1999. Remote Sensing Digital Image Analysis, Springer-Verlag, Berlin, p. 240.
- Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the great plains with ERTS. In: Proceedings third Earth resources technology satellite-1 symposium, Greenbelt, NASA SP- 351(1), p. 3010-3017.
- Sala F., 2011. Agrochimie, Editura EUROBIT, Timisoara, p. 534.
- Sala F., Rujescu C., Constantinescu C., 2016. Causes and solutions for the remediation of the poor allocation of P and K to wheat crops in Romania. AgroLife Scientific Journal 5(1), p. 184-193.
- Șmuleac A., Popescu C., Imbrea F., Popescu G. and Șmuleac L.I., 2016 Topographic and cadastre works for the establishment of an animal farm with NPRD funds, Measure 12, Varadia, Caras – Severin county, Romania Geo Conference SGEM 2016 SGEM VIENNA Green Extended Scientific Sessions, SGEM 2016 Conference Proceedings 3 685-692 ISBN: 978-619-7105-79-7 2 - 5 November, Book 6.
- Smuleac, L., Oncia, S., Ienciu, A., Bertici, R., Șmuleac, A., Pițiga, C., 2013. A Study On The Possibilities Of Using Groundwater In Rural Communities In South-Western Banat Plain. Research Journal Of Agricultural Science, 45(2), p. 287-293.
- Swain P.H. and Davis S.M., 1978. Remote Sensing: The Quantitative Approach, McGraw Hill Book Company, New York, N.Y., p. 375.
- Thenkabil PS, Gangadhara Rao P, Biggs T, Krishna M, Turrall H., 2007. Spectral matching techniques to determine historical land use/land cover (LULC) and irrigated areas using time-series AVHRR pathfinder datasets in the Krishna River Basin, India. PhotogrammEng Remote Sens 73(9), p. 1029-1040.
- <https://www.satimagingcorp.com/satellite-sensors/other-satellite-sensors/rapideye/>