# EVALUATION OF SUPERVISED CLASSIFICATION ALGORITHMS FOR LANDSAT 5 TM IMAGES

*Iosif VOROVENCII, Assoc. Prof. PhD. Eng. – University "Transilvania" of Braşov, Romania, icatop@yahoo.com*
*Mihaela Denisa MUNTEAN, PhD. Student Eng. - The Office for Cadastre and Real-Estate Publicity Alba, Romania, mihaela11m@gmail.com*

***Abstract***: *Classification of satellite images illustrates the land cover at the time of acquisition. In this paper are evaluated maximum likelihood algorithm, Mahalanobis distance algorithm and minimum distance algorithm in Landsat 5 TM satellite image supervised classification. The satellite images was classified in six land cover classes: cropland, pasture and hay, uncultivated, forest and shrub, urabn and built up and water. In this respect, we have calculated overall accuracy, producer's accuracy, user's accuracy and overall kappa statistics. Among the algorithms applied best accuracy was ensured by maximum likelihood with 86.67% overall accuracy and 0.84 kappa statistics, followed by Mahalanobis distance with 80.83% overall accuracy and 0.77 kappa statistics and minimum distance with 68.33% overall accuracy and 0.62 kappa statistics.*

***Keywords***: *maximum likelihood, Mahalanobis distance, minimum distance*

## 1. Introduction

In order to obtain a land cover map, the study area is divided into a finite set of map units, to each of which a land cover class is assigned. Land cover maps that cover large areas contain hundreds of these units that can be obtained from remote sensing data (Hasmadi et al., 2009). Generally, there are two important ways of gathering information on land cover, and namely field survey and satellite imagery (Comber et al., 2005).

Remote sensing data are widely used for the identification and classification of land cover. Applications that most widely use satellite images are related to environmental monitoring, forestry, agriculture, hydrology, geology, and so on. The data obtained from image classification can be used to assess changes in different ecosystems, to monitor global climate change, to assess natural disaster, to identify and track forest fires etc.

Satellite image classification clusters pixels into an image to form several classes, so that pixels with similar spectral properties are part of the same class. Most classifications of satellite images are based on the spectral response of classes covering the land. Classification depends on distinct signatures of land cover classes from the set of bands used and the ability to clearly distinguish spectral signatures from others that may be present in the same image (Eastman, 2003). There are several methods to classify satellite images, the most common one, which also provides good results, being supervised classification.

Supervised classification requires previous knowledge of land cover classes in the area under study. In order to achieve this, it is necessary to collect spectral signatures from training areas which are then used to train a classification algorithm (Kamaruzaman et al., 2009). The size of training areas is very important in supervised classification so that statistical estimates can be reliable. The size of the sample is mainly related to the number of features whose statistical properties are to be estimated. Usually, it is recommended that the minimum size of a set of samples is 10 to 30 times the number of wavelength per class used in classification (Mather, 1999; Piper, 1992). Once these spectral signatures are collected, the algorithm can be applied in order to classify the entire image.

Accuracy assessment is mandatory in obtaining appropriate thematic maps based on the satellite images classified. This step is very important to gain user confidence and classification quality guarantee for the product (Foody, 2001).

The specific objectives of this study were: (1) to assess the algorithms (maximum likelihood, Mahalanobis distance, minimum distance) applied in supervised classification of Landsat 5 Thematic Mapper (TM) satellite images by comparing the accuracies achieved; and (2) to establish and analyse the areas according to the land cover classes.

## 2. The study area

The coordinates of the area under study are $45^{o}70474125$ – $45^{o}88570725$ Northern latitude and $25^{o}102386$ – $25^{o}482788$ Eastern longitude. The area covers 77,221.35 hectares and it includes cultivated land, uncultivated agricultural land, forest, rivers and lakes, towns and villages, roads, pastures and meadows (Fig. 1). The forest is represented mainly by deciduous trees (beech, oak, hornbeam) and coniferous trees (spruce, fir, larch). The main road linking Brasov to Sibiu is National Road 1 (DN1). The River Olt is located in the northwest of the study area and there are some larger lakes such as the Lake of Dumbravita which is a nature reserve and the Feldioara Plant Lake. Part of the land is cultivated, part is not, and some of it has turned to pasture.

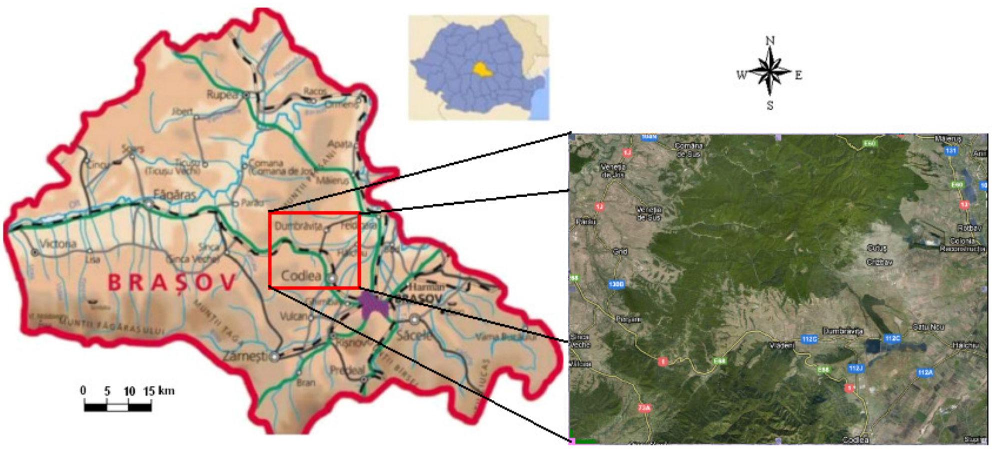



Fig. 1. Localisation of the researchs

Some of the largest towns and villages in the area are Codlea, Hălchiu, Dumbrăvița, Perşani, Sinca Veche, Veneția de Sus and Veneția de Jos.

The altitude specific to the Persani Mountains is between 550 m and 1100 m, ranging from flat agricultural land close to the Magura Codlei Peak (1292 m). The Persani Mountains have an unusual lithological mosaic, reflecting a troubled history. Besides crystalline schist, flysch rock is present (limestone, limestone conglomerates, marls, clays, sandstones), and so is igneous (basalt, andesite, gabbro, serpentinite, porphyry, jasper) and volcanic-sedimentary rock (pyroclastic, tuff). The complex lithology is reflected in the diversity of landforms, soils and vegetation.

The average annual temperature is 8.5$^o$ C (47.3$^o$ F) and the average temperature of July is 19.5$^o$ C (67.1$^o$ F). The annual thermal amplitude ranges between 22.5$^o$ C (72.5$^o$ F) and 23$^o$ C (73.4$^o$ F). The average annual precipitation is about 650 mm.

## 3. Materials and methods

The present paper uses a frame from a Landsat 5 TM satellite image of average spatial resolution taken on 14.07.2011, from path/row 183/28. The image was recorded on seven bands, six of which have a spatial resolution of 30 meters, while band 6 has a spatial resolution of 120 meters. During classification, the content of the satellite image from band 6 was excluded. The radiometric resolution of the image was 8 bits.

The method used in the paper is supervised classification applied to the three algorithms, namely: maximum likelihood, Mahalanobis distance and minimum distance.

*The maximum likelihood algorithm* is one of the most widely used in the classification of satellite imagery (Vorovencii, 2005). The method is based on the likelihood that each pixel belongs to a particular class. The basic theory assumes that these likelihoods are equal for all classes and that input bands are evenly distributed. The method requires extensive computation time and it is based on a normal distribution of data in each band entering the classification. It tends to over-classify signatures with relatively large values in the covariance matrix (Al-Ahmadi and Hames, 2009).

*The Mahalanobis distance algorithm* is similar to the minimum distance algorithm, except that it uses the covariance matrix instead and it takes class variability into account. It can be more useful than minimum distance in cases where statistical criteria are taken into account and weighing factors are not required as with the maximum likelihood algorithm. However, this method tends to over-classify signatures with relatively large covariance matrix values. It also has a slower calculation time than the minimum distance algorithm and it is largely based on a normal distribution of data in each band used as input to classification (Al-Ahmadi and Hames, 2009).

*The minimum distance algorithm* (spectral distance) calculates the spectral distance between the measured vector for the candidate pixel and the average vector for each signature. The equation used in classification is based on the Euclidean distance equation. It requires at least as much computing time as the other supervised classification algorithms. The pixels that may have remained unclassified can now be classified. In addition, it does not consider class variability (Al-Ahmadi and Hames, 2009).

Satellite imaging was performed with Erdas Imagine 9.1 software.

## 4. Experimental

### 4.1. Satellite data pre-processing

Landsat 5 TM image pre-processing used in the present paper involved bringing corrections that included *radiometric, atmospheric and geometric corrections* applied to all bands, except the thermal band. Radiometric corrections consisted in changing 8-bit digital values into radiance and reflectance values (Markham and Barker, 1986). Atmospheric corrections aimed at removing the negative effects produced by the atmosphere (scattering, absorption by aerosols and water vapour) on the reflectance of details in the satellite image (Kaufmann, 1988). These were brought to the image based on Chavez's improved dark object image subtraction approach (Chavez, 1988) due to lack of historical data on the atmosphere.

Geometric corrections were applied using 20 control points clearly identified on field, determined by GPS and identifiable on the satellite image. The root mean square error was less than <0.5 pixels. The resampling method used was *nearest neighbour* that did not degrade the digital values of pixels in the original image. The projection system in which the image was geo-referenced is Universal Transverse Mercator.

### 4.2. Image classification

The image used was classified by using the supervised classification method and the algorithms maximum likelihood, Mahalanobis distance and minimum distance. This was done with the help of all bands found to be most effective in distinguishing each class, excluding the thermal band. The method consisted in choosing the training sites, the actual classification and result evaluation (Lillesand and Kiefer, 1999). During the training phase, 60 training sites were selected by on-screen digitization of specific polygons (Fig. 2). The spectral classes obtained in this way were transferred to "signature editor" of the classification module of ERDAS software. After choosing the spectral signatures and checking them, for the same spectral class, they were merged into a single class. The files obtained were saved and used for image classification. Each training field was assigned a number from 1 to 6 representing land cover classes including: cropland (agricultural land cultivated with different cultures), pasture and hay, uncultivated (uncultivated agricultural land), forest and shrub (deciduous, evergreen, bushes), urban and built up areas (residential, industrial, transportation, communication and utilities, industrial and commercial complexes, mixed urban or built up land, roads) and water (streams, lakes, canals).

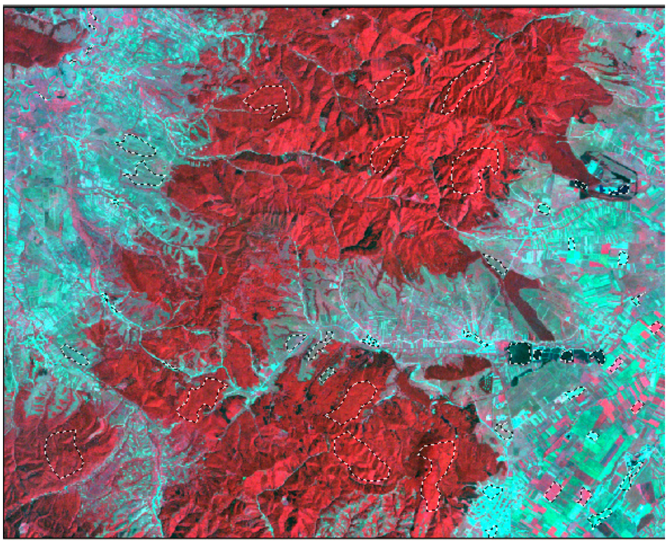


Fig. 2. Collection of spectral signatures

The evaluation of classification accuracy can be defined as the process of comparing the classified image with geographic data considered to be accurate and referential. Typically, the data which the classified image is compared to are ground-truth. In general, a set of reference points are used, which are generated over the classified image and compared with the reference image. The relationship between the two images is expressed in the error matrix, also known as confusion matrix or contingency table. The number of rows and columns in the error matrix must be equal to the number of categories whose precision is being evaluated (Lillesand and Kiefer, 2000).

In the error matrix, the pixels located along the diagonal, from top left to bottom right corner, are pixels correctly classified into the studied categories. Non-diagonal values in columns are omission errors, while non-diagonal values in rows are commission errors. Omission error calculates the likelihood that the pixel is classified accurately (producer's accuracy). It results from dividing the number of pixels correctly classified in each class to the number of training pixels used for that class (column total). This indicates how well the training pixel set of a land set is classified. Commission error determines the probability of a pixel to represent the class to which it was assigned (user's accuracy). It is calculated by dividing the number of pixels correctly classified in each class to the total number of pixels in this class (row total). Overall Accuracy is calculated by dividing the total number of correctly classified pixels (sum of major diagonal) to the total number of tested pixels (Lillesand and Kiefer, 2000). Another characteristic coefficient obtained from the error matrix is the kappa coefficient that takes into account the pixels that have not been correctly classified, outside the main diagonal, with values ranging from 0 (worst) and 1 (best). The calculus relationship for this index is the following (Lillesand and Kiefer, 2008):

$$\hat{k} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}\left(x_{i+} \cdot x_{+i}\right)}{N^2 - \sum_{i=1}^{r}\left(x_{i+} \cdot x_{+i}\right)} \tag{1}$$

în care: $r$ – number of rows in the error matrix;

$x_{ii}$ – number of observation in row $i$ and column $i$ (on the major diagonal);

$x_{i+}$ - total of observation in row i (shown as marginal total to right of the matrix);

$x_{+i}$ – total of observation in column $i$ (shown as marginal total at bottom of the matrix);

$N$ – total number of observation included in matrix.

In the present paper, we used a total of 120 points randomly equalized over the satellite image. These points were used in assessing image classification for all three classification algorithms, thus ensuring their identity.

## 5. Results and discussions

### 5.1. Analysing accuracy classification

After applying the three algorithms, three thematic images were obtained, showing the six land cover categories considered to be representative of the area under study (Fig. 3 a, b, c).

The supervised classification method applied with the help of the three algorithms led to different results highlighted in the error matrix. Thus, in the case of maximum likelihood, the best overall classification accuracy was obtained, of 86.67%, followed by Mahalanobis distance algorithm with 80.83% and then by minimum distance with 68.33% (Table 1, 2, 3, Fig. 4). The overall kappa statistics for maximum likelihood were 0.8400, and 0.7700 for the Mahalanobis distance and of 0.6200 for the minimum distance (Table 4, Fig. 4).

Given the number of pixels correctly classified from the 20 pixels belonging to each class, in the case of maximum likelihood algorithm, the users' accuracy for the land classes was the following: urban and built up areas and water (95%), followed by forest and shrub

(90%) and cropland, pasture and hay and uncultivated (80%). Therefore, for water and urban and built up there were no problems given the different spectral behaviour for these land classes. The following land classes presented difficulties: cropland, pasture and hay, and uncultivated. Due to their close spectral behaviour, they brought some confusion in classification.

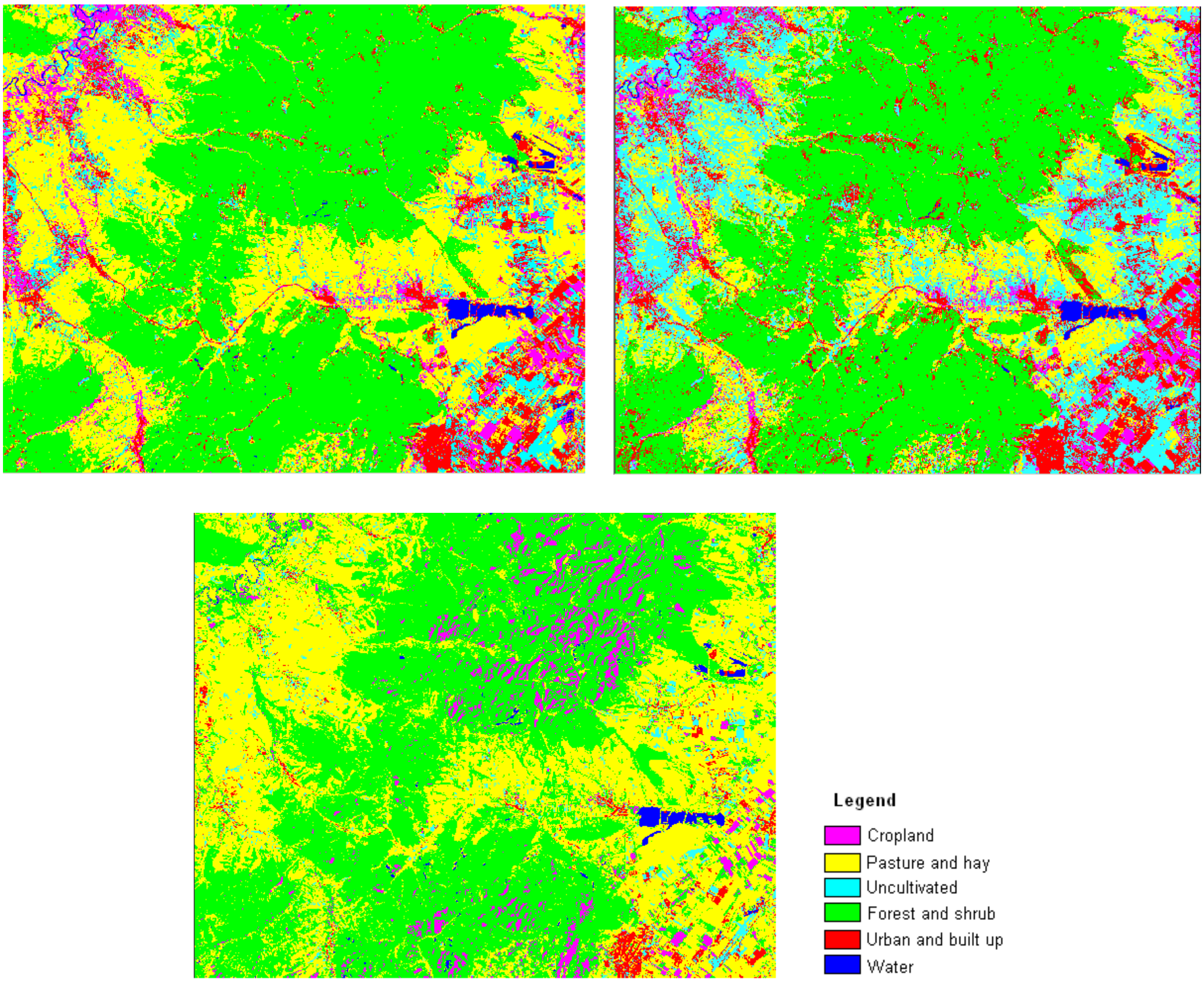



Fig. 3. Classification of Landsat 5 TM (14.07.2011):
a. maximum likelihood, b. Mahalanobis distance, c. minimum distance

Table1. The error matrix for maximum likelihood algorithm

| Class name | Cropland | Pasture and hay | Uncultivated | Forest and shrub | Urban and built up | Water | Total | Users Accuracy |
|---|---|---|---|---|---|---|---|---|
| Cropland | 16 | 0 | 2 | 0 | 2 | 0 | 20 | 80.00% |
| Pasture and hay | 0 | 16 | 0 | 3 | 0 | 1 | 20 | 80.00% |
| Uncultivated | 3 | 0 | 16 | 0 | 1 | 0 | 20 | 80.00% |
| Forest and shrub | 0 | 2 | 0 | 18 | 0 | 0 | 20 | 90.00% |
| Urban and built up | 0 | 0 | 0 | 0 | 19 | 1 | 20 | 95.00% |
| Water | 0 | 1 | 0 | 0 | 0 | 19 | 20 | 95.00% |
| Total | 19 | 19 | 18 | 21 | 22 | 21 | 120 | |
| Producers Accuracy | 84.21% | 84.21% | 88.89% | 85.71% | 86.36% | 90.49% | | |
| **Overall Classification Accuracy = 86.67%** | | | | | | | | |

Table 2. The error matrix for Mahalanobis distance algorithm

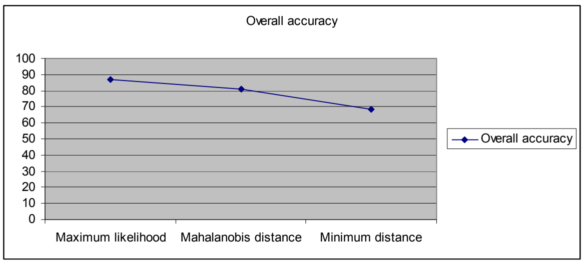
| Class name | Cropland | Pasture and hay | Uncultivated | Forest and shrub | Urban and built up | Water | Total | Users Accuracy |
|---|---|---|---|---|---|---|---|---|
| Cropland | 14 | 0 | 5 | 0 | 0 | 1 | 20 | 70.00% |
| Pasture and hay | 0 | 13 | 7 | 0 | 0 | 0 | 20 | 55.00% |
| Uncultivated | 0 | 1 | 19 | 0 | 0 | 0 | 20 | 95.00% |
| Forest and shrub | 1 | 0 | 1 | 16 | 2 | 0 | 20 | 80.00% |
| Urban and built up | 0 | 0 | 3 | 1 | 16 | 0 | 20 | 80.00% |
| Water | 0 | 0 | 0 | 0 | 1 | 19 | 20 | 95.00% |
| Total | 15 | 14 | 35 | 17 | 19 | 20 | 120 | |
| Producers Accuracy | 93.33% | 92.86% | 54.29% | 94.12% | 84.21% | 95.00% | | |
| **Overall Classification Accuracy = 80.83%** | | | | | | | | |

Table 3. The error matrix for minimum distance algorithm

| Class name | Cropland | Pasture and hay | Uncultivated | Forest and shrub | Urban and built up | Water | Total | Users Accuracy |
|---|---|---|---|---|---|---|---|---|
| Cropland | 9 | 8 | 0 | 1 | 1 | 1 | 20 | 45.00% |
| Pasture and hay | 0 | 19 | 1 | 0 | 0 | 0 | 20 | 95.00% |
| Uncultivated | 0 | 9 | 11 | 0 | 0 | 0 | 20 | 55.00% |
| Forest and shrub | 7 | 1 | 0 | 12 | 0 | 0 | 20 | 60.00% |
| Urban and built up | 0 | 5 | 1 | 2 | 12 | 0 | 20 | 60.00% |
| Water | 0 | 1 | 0 | 0 | 0 | 19 | 20 | 95.00% |
| Total | 16 | 43 | 13 | 15 | 13 | 20 | 120 | |
| Producers Accuracy | 56.25% | 44.19% | 84.62% | 80.00% | 92.31% | 95.00% | | |
| **Overall Classification Accuracy = 68.33%** | | | | | | | | |

Table 4. Kappa statistics for the three algorithms of classification

| Class name | Kappa statistics | | |
|---|---|---|---|
| | Maximum likelihood | Mahalanobis distance | Minimum distance |
| Cropland | 0.7624 | 0.6571 | 0.3654 |
| Pasture and hay | 0.7624 | 0.6038 | 0.9221 |
| Uncultivated | 0.7647 | 0.9294 | 0.4953 |
| Forest and shrub | 0.8788 | 0.7670 | 0.5429 |
| Urban and built up | 0.9388 | 0.7624 | 0.5514 |
| Water | 0.9394 | 0.9400 | 0.9400 |
| | Overall Kappa statistics = 0.8400 | Overall Kappa statistics = 0.7700 | Overall Kappa statistics =0.6200 |



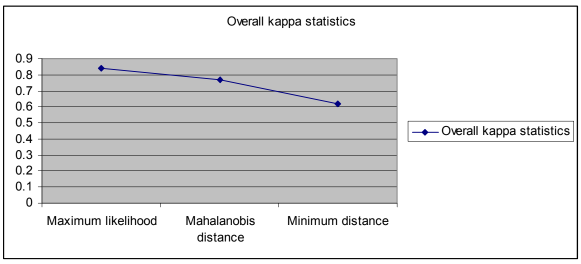




Fig. 4. Reprezentation of overall accuracy and overall kappa statistics

In the case of the Mahalanobis distance algorithm, water was classified best, just as with maximum likelihood algorithm, followed by uncultivated (95%). In descending order, next came forest and shrub, and urban and built up (80%), cropland (70%) and pasture and hay (55%).

In the case of minimum distance algorithm, the ranking order was as follows: water and pasture and hay (95%), forest and shrub and urban and built up (60%), uncultivated (55%) and cropland (45%).

### 5.2. Analysis of the classified areas

Considering the areas obtained by applying the three classification algorithms, certain aspects can be highlighted (Table 5, Fig. 5).

Table 5. Situation of surfaces for the three algorithms (in hectares)

| Class name | Surface (ha) | | |
|---|---|---|---|
| | Maximum likelihood | Mahalanobis distance | Minimum distance |
| Cropland | 4160.16 | 3909.42 | 3665.16 |
| Pasture and hay | 25530.84 | 14984.01 | 29929.77 |
| Uncultivated | 5736.60 | 15126.48 | 1754.37 |
| Forest and shrub | 35413.56 | 33647.76 | 39676.41 |
| Urban and built up | 5724.63 | 9002.61 | 1571.58 |
| Water | 655.56 | 551.07 | 624.06 |
| **Total** | **77221.35** | **77221.35** | **77221.35** |

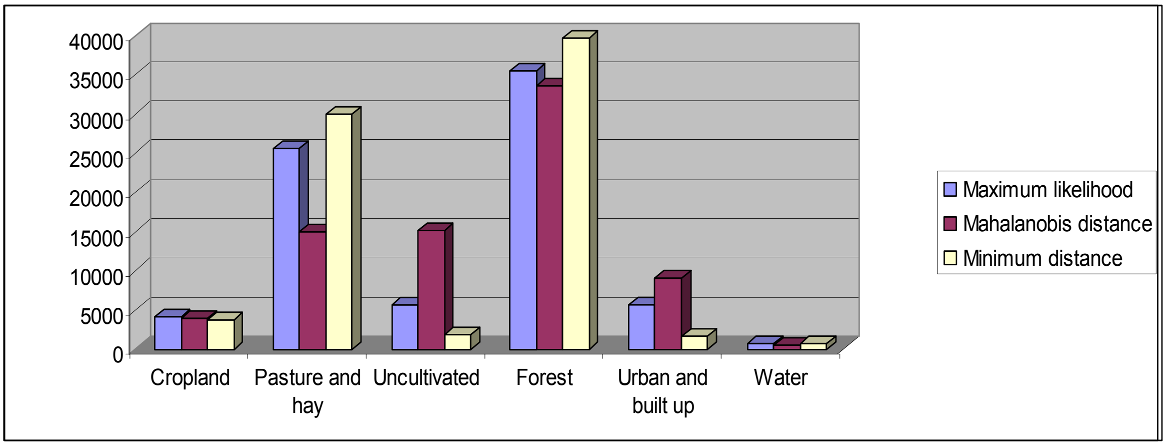



Fig. 5. The surfaces classified by the three algorithms

Speaking about cropland, smaller areas were obtained with the Mahalanobis distance and minimum distance algorithms compared with maximum likelihood. This small decrease can be attributed to the similar spectral behaviour of cropland and pasture and hay, which practically included the mentioned areas. In the case of pasture and hay, the area classified is much smaller (14,984.01 hectares) for Mahalanobis distance compared with maximum likelihood (25,530.84 hectares) and minimum distance (29,929.77 hectares). This reduction in the area pasture and hay can be attributed to the expansion of the area classified as uncultivated. With Mahalanobis distance, it is of 15,126.48 hectares, compared to the maximum likelihood classified area of 5,736.60 hectares. The most significant growth of the area occupied by pasture and hay, namely 29,929.77 hectares, was obtained by applying the minimum distance algorithm. The inclusion of such areas in this class was done at the expense of the area classified as uncultivated (1,754.37 hectares) and urban and built up (1,571.58 hectares), which are quite small compared to maximum likelihood. In fact, this is visible on the classified images, where pasture and hay occupies a large area and uncultivated and built up areas are under-represented. For urban and built up, roads were not classified in that class, while small towns were included in pasture and hay. As for forest and shrub, the Mahalanobis distance algorithm was used to classify a smaller area (33,647.76 hectares) and the minimum distance algorithm to classify a larger area (39,676.41 hectares), both in reference to the surface obtained by maximum likelihood (5724.63 hectares). Although the area classified as cropland through minimum distance (3665.16 hectares) is smaller than in the case of other algorithms, a part of this area comes from areas that should have been classified as forest and shrub. However, the minimum distance algorithm led to a larger area covered by forest and shrub, obviously detrimental to the area occupied by uncultivated and urban and built up.

Related to urban and built up, the areas classified by Mahalanobis distance and minimum distance are quite different (9,002.61 hectares, and 1,571.58 hectares) compared to the one obtained through maximum likelihood (5724.63 hectares). In the case of Mahalanobis distance, a number of uncultivated areas were included in this class because of similar spectral behaviour, due to which the algorithm could not separate them. This was reported in the south-eastern part of the study area (east and north-east of Codlea) and around the towns and villages in the area. It was also found that the valleys in forest and shrub were classified by this algorithm as urban and built up. In exchange, in the case of minimum distance method, many of the smaller villages were classified as pasture and hay, leading to a decrease in the area allocated to this class. The water class was best classified as all three algorithms produced almost identical results.

## 6. Conclusions

The present study analysed land cover by using Landsat 5 TM data with a 30-meter spatial resolution. Classifications were made by using three supervised classification algorithms. The outputs of classification were three thematic maps, accuracy setup for the three algorithms and evaluation of areas belonging to the six land cover classes.

Of the three applied algorithms, maximum likelihood allowed for the best possible accuracy of 86.67%, followed by Mahalanobis distance with 80.83% and by minimum distance with a percentage of 68.33.

Given the high accuracy yielded by maximum likelihood, the areas were evaluated according to it. Thus, the Mahalanobis distance algorithm came second and overestimated the area occupied by uncultivated at the expense of the area covered with pasture and hay as well as the urban and built up area. The minimum distance method was ranked third and it overestimated the area covered by pasture and hay as well as that occupied by forest and shrub at the expense of uncultivated and urban and built up. The extraction of smaller urban and built up areas by using the minimum distance algorithm proved to be difficult, as they were frequently mistaken for pasture and hay.

## Reference

1. *Al-Ahmadi, F. S., Hames, A.S. - Comparison of Four Classification Methods to Extract Land Use and Land Cover from Raw Satellite Images for Some Remote Arid Areas, Kingdom of Saudi Arabia, Earth Sci., Vol. 20 No.1, pp: 167-191, 2009;*

2. *Chavez, P.S. - An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data, Remote Sensing of Environment, **24**, 459-479, 1988;*
3. *Comber, A.J., Fisher, P.F., Wadsworth, R.A. - What is land cover? Environment and planning.Planning and Design **32,** 199-209, 2005*
4. *Eastman, J.R., - Guide to GIS and image processing **14**, 239-247. Clark University manual, USA, 2003;*
5. *Foody, G.M., - Status of Land Cover Classification Accuracy Assessment. Remote Sensing of Environment **80**, 185-201, (2001);*
6. *Kamaruzaman, J., Mohd Hasmadi I., Nurul Hidayah, M.A. - Spectral separability of tropical forest tree species using airborne hyperspectral imager. J. of Environmental Science &Engineering **3**(1), 37-41, 2009;*
7. *Kaufman, Y., - Atmospheric effect on spectral signature – measurement and corrections, IEEE Transactions on Geoscience and Remote Sensing, **26**, 441-450, 1988.*
8. *Lillesand, T.M., Kiefer, R.W. - Remote Sensing and Image Interpretation. John Wiley & Sons, Inc. New Jersey, 1999;*
9. *Lillesand, T.M., Kiefer, R.W. - Remote Sensing and Image Interpretation, 4th ed. Wiley & Sons, 2000;*
10. *Lillesand, T.M., Kiefer, R.W. - Remote Sensing and Image Interpretation, sixth Edition, John. Wiley & Sons, Inc. - 2008;*
11. *Markham, B.L., Barker, J.L. - Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectances and at-satellite temperatures. Earth Observation Satellite Co., Lanham, MD, Landsat Tech. Note 1, 1986;*
12. *Mather, P.M. - Computer Processing of Remotely-Sensed Images: An Introduction. Second Edition, Chichester: John Wiley and Sons, 1999;*
13. *Mohd Hasmadi, I. Pakhriazad, H.Z., Shahrin M.F. - Evaluating supervised and unsupervised techniques for land cover mapping using remote sensing data. Malaysian Journal of Society and Space **5** issue 1 (1 – 10), 2009;*
14. *Piper, J. - Variability and bias in experimentally measured classifier error rates. Pattern Recognition Letters, 13, 685-692, 1992;*
15. *Vorovencii, I. - Researches Concerning the Possibilities of Using Satellite Images in Forest Planning Works. Doctor's degree paper. "Transilvania" University of Brasov, 294 p, 2005.*