**Ivana Petrovska**
MSc, Teaching Associate
Ss. Cyril and Methodius University
Faculty of Civil Engineering – Skopje
email: ivanapetrovska13@gmail.com

**Lazo Dimov**
PhD, Full Professor
Ss. Cyril and Methodius University
Faculty of Civil Engineering – Skopje
email: dimov@gf.ukim.edu.mk

# ACCURACY ASSESSMENT OF UNSUPERVISED LAND COVER CLASSIFICATION

This paper shows the importance of combining remote sensing techniques and GIS tools to quantify the quality of unsupervised classification in addition to determining the land cover classes. In order to quantitative understanding of the allocation of different types of land, it's necessary to perform validation or assessment of the accuracy of the classification.

The validation of the unsupervised land cover classification for the valley of river Treska (SJCE vol 9, issue 1) is performed by comparing the corresponding points placed on the reference surface (satellite image), from which the classification is obtained and the thematic map. The validation results showed an overall accuracy of the classification of 89%, while based on the Kappa coefficient - 85% which is an indicator of high compatibility of the classified thematic map with the actual ground information.

Keywords: remote sensing, GIS tools, validation, reference points, accuracy assessment, unsupervised classification

## 1. INTRODUCTION

Accuracy assessment is an essential and key part of any classification project. The accuracy assessment actually reveals the degree of correspondence between the actual ground data and the classification results and provides the user with more information about where the errors occurred. Depending on the acceptable level of errors, the user will determine if the classified map is useful or needs to be reclassified. Thereby, two data sources are compared: the classified thematic map and ground reference test data, which is considered accurate or contains the "true" values for the land cover classes. Ground truth values can be collected in the field or extracted from the interpretation of high-resolution images or existing classification maps. [10] The level of (mis)match between the two sets of data is a measure of the accuracy of the classified land cover map.

The relationship between the classified map (Figure 1) and the reference data is summarized in a confusion matrix.

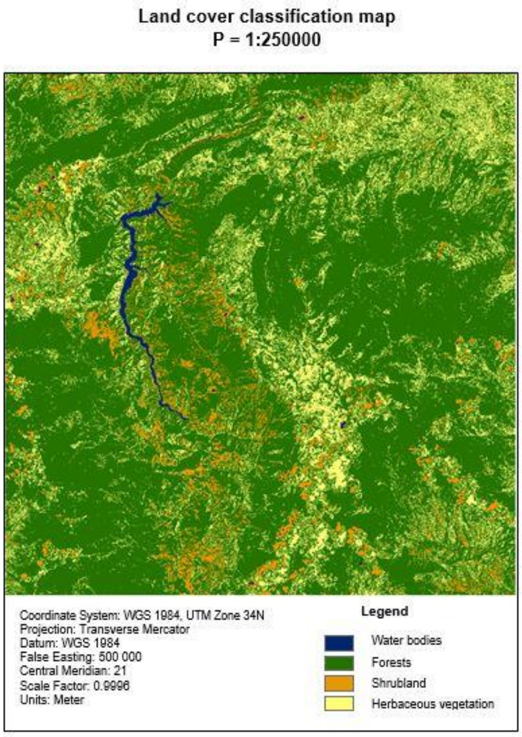


Figure 1: Land cover classification map

## 2. ACCURACY ASSESSMENT

Accuracy assessment is useful for checking the validity of the classification approach for error assessment. The accuracy report includes confusion matrix, commission error, omission error, producer's and user's accuracy for each information class individually, as well as the overall accuracy and statistics of the Kappa coefficient.

The pixels that are correctly assigned to each information class are represented in the diagonal fields of the matrix. Non-diagonal fields show classification errors given the ground reference information.

The commission error determines the probability by class that the test point on the classified map does not belong to the same class on the ground and occurs when the ground cover class is included in an incorrect class category. It is obtained when the sum of incorrectly classified pixels for each class is divided by the number of actual (true) values for each class.

The omission error, in contrast to the commission error, refers to reference locations that have been left out or omitted from the correct class of the classified map, providing the probability by class that the class sample is classified in another class on the map. The right type of land cover is excluded from the class it really belongs to. It is obtained when the number of incorrectly classified pixels in a column is divided by the number of reference pixels of that class - the sum of the column.

Statistics that shows the probability that a reference pixel is correctly classified and is a measure of the omission error, or how well the analyst has classified a particular area, is called producer's accuracy. Producer's accuracy is obtained by dividing the number of correctly classified pixels of a given class by the column of total reference points for the class in question. The low producer's accuracy implies a high error of omission.

User's accuracy, on the other hand, represents the probability that a pixel classified in a given class actually represents that ground class. It is calculated by dividing the number of correctly classified pixels by the total number of pixels of the class and indicating what percentage of a certain type of land cover on the map is really that type of land cover in reality. Low user's accuracy entails a high commission error.

In general, overall classification accuracy is obtained when the number of correctly classified reference points is divided by the total number of reference points. [2]

Another indicator of the overall accuracy of a classified map is the Kappa statistics which also compares two sets of data to see if they differ significantly (classified map and the reference data are considered). Unlike overall accuracy, Kappa statistics is a more reliable indicator of classification accuracy because it uses all the data in the confusion matrix, not just the diagonal ones, and it is calculated by the equation:

$$k = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} * x_{+i})} \qquad (1)$$

where:

k = Kappa value of the coefficient

r = number of inputs (for example, land cover classes) in the matrix

$x_{ii}$ = number of observations in row „*i*“ and column „*i*“

$x_{i+}$ and $x_{+i}$ = marginal amounts for row „*i*“ and column „*i*“

N = total number of observations (test points)

 

The Kappa coefficient can take values from 0 to 1. A value close to 1 indicates a perfect match between the ground land cover and the classified map, while 0 indicates a complete discrepancy between the two sets of data. If the value of Kappa is greater than 0.80, a large match or accuracy can be found between the classification map and the ground reference values, the value of Kappa between 0.40 and 0.80 is considered moderate, while the value of Kappa less than 0.40 indicates a weak correspondence between the two data sets. [2]

## 3. RESULTS OF THE ASSESSMENT

In order to determine to what extent the thematic map corresponds to the current situation on the ground and for what needs and purposes it can be used, a validation process is performed: test points are compared with their corresponding locations on the classified map.

The test points should be evenly distributed on the map because if they are concentrated on only one part of it, the accuracy will only be relevant to that part. Also, the location of the samples must be chosen randomly without bias, as any bias can affect the statistical analysis of the confusion matrix and may result in an error or inaccurate assessment of the actual accuracy of the thematic map. [7]


Figure 2: Allocation of test points

Moreover, in addition to the location, it is important to determine their optimal number in order to further obtain a quantitative indicator of the accuracy of the classification performed. The number of reference points is an important factor in assessing the accuracy of any classified map. A good guide is to use a minimum of 30 test points for each information class to obtain a statistically valid sample. Approximately, more than 250 pixels are needed to estimate the average accuracy of a class within ± 5%. [2] Or, in this case, ten times more test points are used than the number of classes - 40 test points for each class. That is, the validation of the land cover map was performed with a total of 160 test points (Figure 2).

The test points are placed on the reference surface used to perform the classification - satellite imagery. Each test point is compared to the pixel size of the map and it should fall inside the pixel.

After placing and labelling the test points, the next step is to calculate the results of the comparison which are presented in a confusion matrix. The confusion matrix is a central element of the assessment and provides a specific number of individual and total parameters for each of the classes and compares, category by category, the relationship between the known reference data ("true" values) and the corresponding results of the automatic classification procedure [8]. The columns of this matrix represent the information about the reference test points, the rows correspond to the classes generated by the unsupervised classification, and the diagonal fields represent the pixels assigned to the exact class.

However, in order to ensure thematic consistency between the map data and the "true" values, it is useful to define a protocol or document in a recognizable and understandable format. The most common way to express the accuracy of the classification is by means of a contingency table that presents the errors by class and the match between the map classification data and the reference data.

Table 1 summarizes the parameters of the correspondence between the entities on the map and the satellite image, as follows: out of a total of 160 test points in the respective class belong 142 (diagonally) as follows: for class 1 and 2 all test points placed on the reference surface match with the classified values, while for classes 3 and 4, 32 or 30 out of a total of 40 test points are correctly classified. Furthermore, in the next column the total number of points is given, i.e. the "true" values that belong to each class individually. Of the placed test points, 42 belong to the class of water bodies, 40 points

Table 1: Results of the accuracy assessment

| Number and class name | Class 1 | Class 2 | Class 3 | Class 4 | Row total (ground truth values) | Commission error (%) | Omission error (%) |
|---|---|---|---|---|---|---|---|
| Water bodies | **40** | 0 | 2 | 0 | 42 | 4.8 | 0 |
| Forest | 0 | **40** | 4 | 7 | 51 | 21.6 | 0 |
| Shrubland | 0 | 0 | **32** | 3 | 35 | 8.6 | 20 |
| Herbaceous vegetation | 0 | 0 | 2 | **30** | 32 | 6.2 | 25 |
| Column total | 40 | 40 | 40 | 40 | **160** | | |
| Producer's accuracy (%) | *100* | *100* | *80* | *75* | Number of correctly classified pixels 142 | | |
| User's accuracy (%) | *95.2* | *78.4* | *91.4* | *93.8* | | | |
| *Total accuracy 88.8%* | | | | **Kappa 85%** | | | |

from the class itself and 2 which during the validation are assessed that they belong to the class of shrubland, and during the classification are assigned to class 1. Probably, these are the pixels on the river bank or the shadows of clouds whose small parts based on the spectral structure were assigned to class 1. Consequently, in the forest class, instead of only the 40 pixels assessed when positioning the points belonging to that class, there are another 4 points of shrubland and 7 which during the validation process are marked as herbaceous vegetation. Next, the points that are estimated to correspond to shrubland on the ground, compared to the classified map, only 32 correspond. Additionally, in this class are classified 3 points which during the validation are marked as herbaceous vegetation. Finally, the number of "true" values for class 4 is 32, of which 30 belonging to the class itself and two from shrubland.

The next column presents the commission errors for each class individually. This percentage is the highest, or the majority of the test points are placed in the wrong class forest - 21.6%, followed by the classes of shrubland and herbaceous vegetation with 8.6%, and 6.2% respectively and finally the water bodies with 4.8%, which has the lowest number of test points (2) that do not belong to this class.

The last column represents the omission error, which, in contrast to the commission error, shows how well the reference pixels of the particular ground cover are classified. The biggest omission error is present in herbaceous vegetation - 25%. The pixels marked as herbaceous vegetation on the classified map were supposed to belong to forests (7) and shrubland (3). Next is class shrubland with 20% of the pixels excluded from the class to which they really belong. The omitted pixels are distributed in the remaining three classes: 2 belong to water bodies, 4 are assigned to forests and 2 to herbaceous vegetation. While in water bodies and forests there are no pixels that are excluded from the classes and therefore the omission error rate is 0%.

To summarize, the source of errors in land cover classes: forests, shrubland and herbaceous vegetation can be caused by the foliar coverage. The energy reflected and measured by Landsat-8 sensors is based on the interaction of electromagnetic radiation with plant components and bare soil. Density and foliar coverage are particularly affected within forests, where some grasses may affect the reflection response.

Moreover, as already mentioned, one of the prerequisites for validation with the confusion matrix is that each entity on the map and each sample taken are assigned to one class and they represent the same spatial extension. This can be more or less difficult to achieve depending on a number of factors such as: the structure of the scene, the spatial resolution of the pixels, the minimum mapping unit, the positional accuracy of the map and the size of the sample. For example, in a homogeneous scene with huge ground cover samples and mostly "clean" pixels, it is unlikely that positional errors, e.g. some meter displacements between the position of the ground sample and the map, or classification problems will affect the accuracy rate. However, they can have an impact in areas with heterogeneous fragmented areas where "mixed" pixels are common and where a small change in spatial position will transfer the sample to another class of land cover.

When the area of interest is composed of two or more entities that differ significantly in brightness and spectral response, the pixel is composed of several, very different values of digital numbers, so the average of the values of different land cover classes is calculated. That unique value of the digital number that represents the pixel can not accurately represent any of the present categories. "Mixed" pixels are common in data with rough spatial resolution and along the edges of entities and can sometimes lead to misclassifications. [10] Using a satellite image with a spatial resolution of 30x30 meters, it is normal to have "mixed" pixels, whose presence affects the sorting of different types of land cover in the appropriate class.

In addition, the accuracy of the placement and labelling of the test points in the appropriate class should be taken into account. Although Google Earth is used to allocate the reference pixels, in order to identify the ground cover in a certain area, as well as the previous knowledge of the relief and geography of the scene, still the eventual placement of the test points exactly on "mixed" pixels or some oversight in the setting, as well as the influence of the degree of accuracy of the chosen mathematical model and algorithm can not be completely eliminated.

Furthermore, the producer’s and user’s accuracy for each class is also calculated. According to Table 1, the first two classes are absolutely correctly classified, while the percentage of the producer’s error for shrubland is 80% and for herbaceous vegetation this percentage is slightly lower and is 75%. Based on the presented percentages, it can be concluded that all four classes are characterized by a high degree of classification accuracy.

According to the user’s accuracy as well, we can notice a high correspondence, over 91% for class 1, 3 and 4, while this percentage is the lowest for class 2 and is just over 78%, which is to be expected given the largest commission error for this class. Nevertheless, based on this parameter, all four classes are characterized by a high degree of classification accuracy.

After calculating the individual errors for each class, it is necessary to determine the overall accuracy of the thematic map. In this regard, there are two ways to evaluate the confusion matrix of validation: descriptive statistical and analytical-statistical method. [7]

1. The descriptive statistical method evaluates the total accuracy of the classified map, i.e the percentage of correctly classified land cover. Determining the degree of total accuracy of the thematic map can be simplified by dividing the total number of correctly classified pixels (sum of the main diagonal) and the total number of test points (samples) and in this case is high 88.8%.

2. Analytical-statistical methods are used to statistically evaluate the accuracy of classified maps obtained from remote sensors and the confusion matrix. [4] These methods are suitable for analyzing data from remote sensors, because such data are discrete, with binomial distribution. These methods involve a k-coefficient, also called Kappa Analysis.

Unlike total accuracy, which takes into account only the number of total and accurately classified test pixels, the Kappa coefficient, in addition to these parameters, includes in its calculation the number of test points for each class individually, as well as the number of pixels that belong to each class ("true" values) and is therefore a more reliable indicator of the accuracy of the unsupervised classification.

As already mentioned, the values of the Kappa coefficient range from 0 to 1 and if the value of Kappa is greater than 0.80, a large match or accuracy can be found between the classification map and the country reference. For the land cover classified map in this master thesis, based on the calculation expression given in section 3.3, the Kappa coefficient is 0.85 which means high classification accuracy, or expressed as a percentage - 85% (Table 1) for better visual presentation and interpretation.

## 4. CONCLUSION

In order to perceive the quality of the classification, accuracy assessment is inevitable. The accuracy assessment was performed by comparing the values of the corresponding points on the satellite imagery used as a reference for the classification and the classified map. The results are summarized in different groups of errors for each class individually (commission error, omission error, producer's accuracy and user accuracy), but the total accuracy of the map, which is high 89%, is also determined. However, a better indicator of the accuracy of the performed classification is the Kappa coefficient because it takes into account several parameters and is 85%, which is a high degree of coincidence of the map with the current situation in reality.

In this context, it is logical to ask the question of the cause of these errors. The source of errors in the classes: forests, shrubland and herbaceous vegetation, can be found in the density and foliar coverage that particularly affect within forests, where some grasses may affect the reflection response. Also, considering the heterogeneity of the scene and the spatial resolution of the satellite image (30 meters) the presence of "mixed" pixels, which affects the sorting of different types of land cover in the appropriate class is unpreventable.

Of course, the precision of the placement of test points in the appropriate class should also be taken into account. Although an open source such as Google Earth was used to allocate the reference pixels, in order to identify the ground cover of a certain area, as well as the previous knowledge of the relief of the scene, still the possible placement of the test points on "mixed" pixels or some oversight in the placing, as well as the impact of the degree of accuracy of the chosen mathematical model and algorithm should not be completely neglected.

However, although this study discussed about validation of land cover classification, limitations in terms of data sources made this study to not be more specific. This study depends only on satellite imagery to identify different features of the earth. All results are generated by remote sensing products. The accuracy of the methods for classifying the images can be tested more rigorously using terrestrial measurements. Also, land cover and land use change can be studied with increased spatial and temporal resolution.

In future studies, more time and effort should be spent on improving accuracy (including the accuracy of mathematical models and algorithms), as classification accuracy is extremely important for the final output and areas of application of the classification. Improving classification accuracy will improve the quality of land cover detection results, and land resource change statistics will also be more accurate.

## REFERENCES


[1] Cheruto C.M, Kauti K.M, Kisangau D.P, Kariuki P, (2016), „Assessment of Land Use and Land Cover Change Using GIS and Remote Sensing Techniques: A Case Study of Makueni County, Kenya“, Volume 5, Journal of Remote Sensing & GIS.

[2] Congalton R.G, Green K, (1999), „Assessing the Accuracy of Remotely Sensed Data: Principles and Practices“, Lewis CRC Press, Detecting land-cover changes for the tropics, International Journal of Remote Sensing, 23(12): 2457-2474

[3] Gong F, Liu Q, (2013), „Monitoring land use and land cover change: a combining approach of change detection to analyze urbanization in Shijiazhuang“, China, Faculty of Engineering and sustainable development, Department of Industrial Development, IT and Land Management

[4] Hudson W, Ramm C, (1987), „Correct Formula of the Kappa Coefficient of Agreement“, Photogrammetric Engineering and Remote Sensing, 53, 421-422

[5] Humagain K, (2012), „Examining Land Use/Land Cover Change and Potential Causal Factors in the Context of Climate Change in Sagarmatha National Park“, Nepal, Western Kentucky University, Department of Geography and Geology

[6] Jensen R.J, (2005), „Atmospheric Corrections&Geometric Corrections“, Chapter 6, 198-222

[7] Jensen J.R, (2005), „Introductory Digital Image Processing“, Chapter 5, 165-201

[8] Lillesand T.M, Kiefer R.W, (1987), „Remote Sensing and Image Interpretation“, John Wiley & Sons. New York (p 599).

[9] Petrovska I, (2019), „Automatic classification and validation of land cover using multispectral satellite imagery from Landsat-8“, Faculty of Civil Engineering, Skopje.

[10] Sommya D.R, Deepa Shenoy P, Venugopal K.R, (2017), „Remote Sensing Satellite Image Processing techniques for Image Classification: A Comprehensive Survey“, International Journal of Computer Applications (0975-8887), Volume 161-No 11.

[11] Vital A.J, (2008), „Land use/cover change using Remote sensing and geographic Information Systems: Pic Macaya national Park“, Haiti, Michigan Technological University.