European Journal of Geography Volume 10, Number 2:150- 170, June 2019
© Association of European Geographers

VERIFICATION OF COPERNICUS RIPARIAN ZONES LOCAL COMPONENT FOR BULGARIA. SPECIFIC CASES AND TYPICAL PROBLEMS

Yulia G. KROUMOVA
National Institute of Geophysics, Geodesy, and Geography, Bulgarian Academy of Sciences, Sofia, Bulgaria
jkrumova@gmail.com

Radka G. KOLEVA
University of Forestry, Sofia, Bulgaria
rad.koleva@gmail.com

Youlin N. TEPELIEV
University of Forestry, Sofia, Bulgaria
youlin@abv.bg

Ventzeslav N. DIMITROV
Space Research and Technology Institute, Bulgarian Academy of Sciences, Sofia, Bulgaria
ventzeslavd@gmail.com

Abstract

This paper deals with quantitative evaluation of Riparian zones land cover and land use data set for Bulgaria, supported by Copernicus Programme's funds. This verification task is a part of a project of the European Environmental Agency. The methodology is based on visual inspection of sample polygons on top of reference data sets. The web-based LACO-Wiki tool is used for the verification, based on levels 1 to 4 of MAES (Mapping and Assessment of Ecosystems and their Services) nomenclature. Typical class encoding errors and delineation drawbacks are grouped and analysed as several problem types. Detailed scores and accuracy estimates are provided, supported by appropriate visual examples. Class code confusion appears mostly in classes missing or with rare occurrence in Bulgaria. In general, the delineation is correct with some predominance of unnecessary parts inclusion in the sample polygons area. The Riparian zones product provides very good thematic and spatial detail and can be of value for many applications.

Keywords: *Copernicus, local component, thematic accuracy, land cover, sampling, Very High Resolution (VHR) satellite image, visual interpretation.*

1. INTRODUCTION

Digital image data obtained from Earth observation satellites are nowadays a widely used source of information involved in the production of land cover/land use (LC/LU) maps. Accuracy assessment is a mandatory part of the work on deriving LCLU information from remotely sensed data with a strong relation to its further use (Foody, 2002; Caetano et al., 2005).

The Local Component the Copernicus Land Monitoring Service (Copernicus, 2018) delivers thorough LC/LU information on specific areas of interest, so called "hotspots". The exploration of these territories, sensitive to unfavorable natural and anthropogenic impacts, is divided to three main study areas and realised in three project products: urban areas (Local Component *Urban Atlas*), protected areas (Local Component *Natura 2000*) and hydrographic or coastal areas (Local Component *Riparian Zones*). Based on very high resolution (VHR) imagery (2.5 x 2.5 m pixels), Copernicus Local Component products have a high level of spatial and thematic detail (from 1ha to 0.25 ha Minimum Mapping Unit (MMU), applicable for limited specific areas of interest. This distinguishes them from Pan-European component (CLC – CORINE Land Cover), which refers to larger territories at national and regional level but with a lower spatial resolution (25 ha Minimum Mapping Unit (MMU), High Resolution Layers (HRLs) – 1 ha grid).

The Riparian zones LC/LU product provides a detailed thematic and spatial description for areas along a buffer zone of selected rivers in member and cooperating countries of the European Environmental Agency (EEA). The Riparian zones mapping has as main objective to support the Mapping and Assessment of Ecosystems and their Services (MAES), as part of the EU Biodiversity Strategy to 2020. The developed ecosystems typology of MAES aims to harmonize the approaches and methods of inventory and evaluation activities of the ecosystems in the EU countries and Bulgaria joined the MAES process in 2014. (Nedkov et al, 2018).

The Riparian zones LC/LU product, like the other local components has passed independent validation for the Riparian zones coverage for all the 43 Delivery Units at Pan-European level (EEA, 2017). In parallel, member-countries of the EEA have completed verification of the Riparian zones data for their own territory. For Bulgaria, this verification task is a part of a project, managed by the EEA in the frame of Copernicus Programme and realised through Bulgarian Executive Environmental Agency (ExEA).

The main goal of this study is to assess the accuracy, identify and explain the possible reasons for some typical errors, inaccuracies and omissions in Riparian zones LC/LU data set in Bulgaria at level 4 of MAES nomenclature. In addition to accuracy statistics, specific features and problems found during visual photointerpretation are expected to be revealed.

2. METHODOLOGY

In general, the methodology proposed by EEA (Maucha et al, 2017) is followed for the Riparian zones verification fulfilment. As a first step, the input and the reference data sets are prepared. The Riparian zones LC/LU input data set is formed by clipping and merging procedures for the four catchment-based delivery units for Bulgaria. Next, a stratified random sampling of polygons is realized. To have all the classes represented for the complex Riparian Zones LC/LU product, a map layer based stratification is applied (Congalton, 1991).

Table 1. Reference data used for verification

<table>
<tr><th>Dataset</th><th>Riparian Zone status layer 2012</th></tr>
<tr><td rowspan="5">Reference data provided centrally</td><td>GioLand/VeryHighResolution2012 image mosaic</td></tr>
<tr><td>GoogleEarth Imagery</td></tr>
<tr><td>Bing imagery</td></tr>
<tr><td>World Imagery basemap from ArcGIS online</td></tr>
<tr><td>OpenStreetMap</td></tr>
<tr><td rowspan="2">In situ data used</td><td>Resolution: 0.4 m National Ortophoto database
Reference years: 2010-2011 (partial coverages)</td></tr>
<tr><td>Land Parcel Identification System Land Cover data (LPIS LC)
Minimum Mapping Unit (MMU): 0,1 ha or less in case of agricultural areas
Specific nomenclature focused on LPIS applications
Reference years: 2010-2011 (partial coverages)</td></tr>
</table>

	Forest management plans
	Imperviousness High Resolution (HR) layer
	Hydrographic network data
	Topographic maps, 1:50.000 scale, scanned
Software used for verification	LACO-Wiki, ArcGIS10.3,
	GoogleEarth, OpenStreetMap

The sample design is realised by interactively specifying the required parameters in LACO-Wiki tool (LACO-Wiki, 2018), e.g. number of samples per class. LACO-Wiki tool creates two vector layers: Layer of randomly selected sample polygons and Layer of sample points (one sample point inside each sample polygon). These sample data sets are then used for verification through visual interpretation of polygons and points on top of reference imagery (satellite and aerial photos) and other in-situ data (table 1). The so-called enhanced plausibility approach is applied in LACO-Wiki and in a local GIS software. The interpretation is not blind and class codes of polygons are displayed. Experts evaluate several characteristics of sample polygons: Correctness of LC/LU code around the sample point; Correctness of delineation; Comments as free text are added, if necessary. Finally, an evaluation of results of the verification is accomplished, based on the contingency tables and the sample polygon attributes, generated during the sample interpretation. Thus, a scientifically sound estimate of the thematic accuracy and of some geometric characteristics of the Riparian Zones data set is obtained.

3. CASE STUDY

The Riparian Zones in Bulgaria cover 815494.52 ha, which constitutes 7.6% of the country's territory (Figure1.) and include 74 valid classes appearing according to the MAES (Mapping and Assessment of Ecosystems and their Services) nomenclature, level 1–4.

In accordance with the project requirements for Riparian zones local component, 790 representative sample polygons are selected including 12 to 13 samples for each class with some exceptions. Some classes are represented by fewer polygons, due to the fewer number of polygons of the given class, or because the class is not available in Bulgaria. Correctly interpreted samples are 585, which results in 83.5 % weighted overall accuracy with confidence interval of ± 0.0354, while the target accuracy is 85 %.

Figure 1. Overview figure Riparian zone status layer – Bulgaria

The verification process of the incoming database includes accuracy assessment in two main aspects: *1) thematic accuracy* – correctness of LCLU code and *2) correctness of delineation* by 3 criteria – correctness of delineated area, detail of delineation and positional accuracy.

3.1. Thematic accuracy investigations

The thematic accuracy (TA) verification is carried out in accordance with Riparian zones legend (1-4 level), general mapping rules (Minimum Mapping Unit (MMU) – 0.5 ha, and Minimum Mapping Width (MMW) – 10 m) and utilising additional data sources.

The verification results for the 74 classes of the Riparian zones in Bulgaria show 100% user's thematic accuracy for 19 of them (Table 2), for 30 classes this indicator is between 75% and 92% and for the rest 25 classes it is from 0% to 67% (Figure 2).

A comparative analysis of the class code correctness in the verification with another LC/LU component – Natura 2000 in the same project, shows some coincidence of the classes with 100% user's thematic accuracy (Tepeliev et al., 2017). This correspondence refers mainly to urban classes, as well as to some types of water bodies. For the most part, these are linear objects to which another class of the MAES nomenclature can hardly be attributed: Road networks and associated land, Railways and associated land, Lines of trees and scrub and River banks. This group also includes objects, which have interpretation features, specific for the respective class that cannot be found in another class of the nomenclature: Continuous urban fabric (in-situ based or Imperviousness Density (IM.D.) >80-100%), Airports, Non-irrigated arable land, Greenhouses, Mesic grasslands with trees (Tree Cover Density (T.C.D.) ≥ 30%), Permanent natural water bodies, Intensively managed fish ponds and Marine (other).

Somewhat this circumstance confirms the spectral clarity ensuring easy recognisability of the typical land characteristics of these classes in the reference satellite imagery.

Table 2. Riparian zone classes showing 100% user's accuracy (highlighted classes corresponding to Natura 2000 classes with 100% true codes)

Riparian zones classes	Description	Number of samples
1111	Continuous urban fabric (in-situ based or IM.D. >80-100%)	12
1113	Industrial or commercial units	13
1211	Road networks and associated land	12
1212	Railways and associated land	12
1214	Airports	12
1411	Green urban areas T.C.D. ≥ 30%	12
2111	Non-irrigated arable land	13
2121	Greenhouses	12
3000	New Classification 3000 Urban Atlas: Woodland and forest	12
3131	Other natural & semi natural broadleaved forest	13
3412	Lines of trees and scrub	12
6213	River banks	12
6222	Burnt areas (except burnt forest)	2
7121	Inland saline marshes without reeds	1
9000	Urban Atlas: Rivers and lakes	12
9111	Permanent interconnected running water courses	13
9211	Permanent natural water bodies	13
9214	Intensively managed fish ponds	12
10111	Marine (other)	12

Source: Riparian Zones and Natura 2000 Local Component database

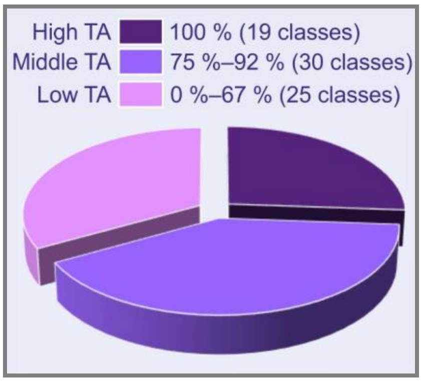

Figure 2. User's thematic accuracy (TA) by Riparian zones classes

During the visual inspection, however, some interpretation specificity and summaries of detected code errors and omissions can be deduced.

In the present study the erroneous thematic encoding of the sample polygons (SP) is exposed and systematized in three main groups, which illustrate the general problem fields and causes of the wrong code interpretation. The examples presented for each of the problem fields demonstrate typical errors in the interpretation of the classes and arguments about the correctness of the changed wrong codes in the verification process.

3.1.1. Particularities of spatial distribution of the Riparian zones classes in Bulgaria

Greater encoding errors with lower percentage of user's thematic accuracy (TA) occur in the classes divided into three main groups:

- ***classes without presence in the country*** (for 2 classes): Olive groves (2 SP – 0.0% TA) and Agroforestry (8 SP – 0.0% TA);
- ***classes with rare distribution in the country*** (for 5 classes): Broadleaved swamp forest (2 SP – 0.0% TA), Broadleaved evergreen forest (12 SP – 0.0% TA), Forest damaged by fire (3 SP – 0.0% TA), Alpine and subalpine grasslands without trees (1 SP – 50.0% TA), Sclerophyllous vegetation (7 SP – 41.7% TA);
- ***mixed classes*** (for 5 classes): Irrigated arable land and rice fields (11 SP – 15.4% TA), Complex patterns of irrigated and non-irrigated arable land (11 SP- 8.3% TA), Agriculture areas with significant areas of natural vegetation (6 SP – 50.0% TA), Low stem fruit trees and berry plantations (10 SP – 16.7% TA), Annual crops associated with permanent crops (5 SP – 16.7% TA).

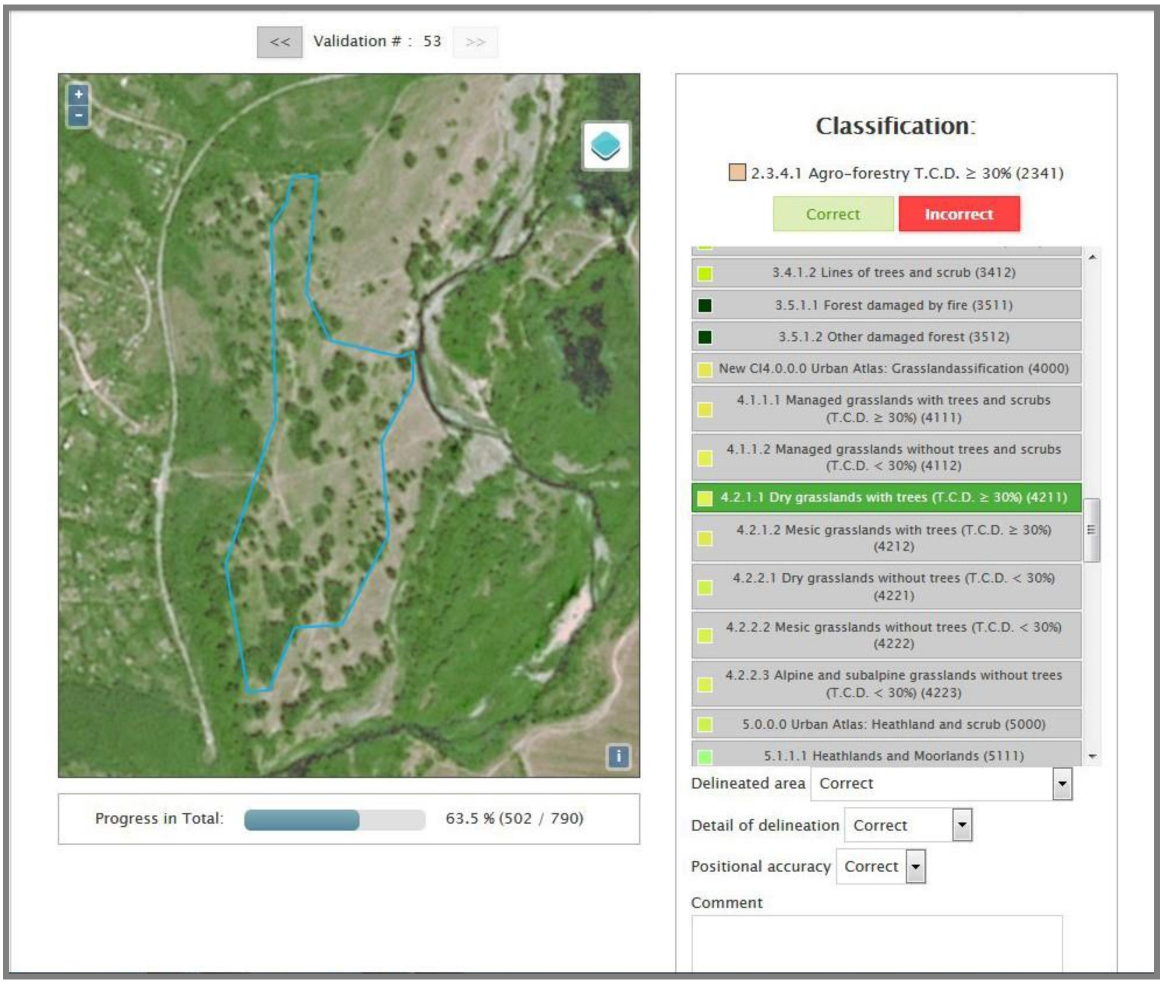

Figure 3. Incorrect classification 2.3.4.1 Agro-forestry T.C.D. ≥ 30%. The correct classes proposed is 4.2.1.1 Dry grassland with trees. (SW Bulgaria, Struma river region, Blagoevgrad district).

It turns out that the two classes without presence in the country are confused with a variety of actual LCLU classes. For example, the land cover of the 8 sample polygons in 2.3.4.1 Agro-forestry T.C.D. ≥ 30% does not correspond with the definition of this class." This category is limited to Mediterranean area" (EEA, 2015). The correct classes proposed are different – 3.4.1.1 Transitional woodland and scrub, 4.2.1.1 Dry grasslands with trees (T.C.D. ≥ 30%), 4.2.2.1 Dry grasslands without trees (T.C.D. < 30%), 3.1.3.1 Other natural & semi natural broadleaved forest and 4.1.1.2 Managed grasslands without trees and scrubs

(T.C.D. < 30%). Figure 3 presents one of the above classes and the interpretation of the visual characteristics in code correction by means of the LACO-Wiki web-based software tool.

One general mistake discovered is assigning to 12 polygons the code 3.1.4.1 Broadleaved evergreen forest (12 SP – 0.0% TA). In fact, small areas of evergreen Mediterranean oak (Quercus coccifera), mostly shrubs and rarely trees, can be found only at the south-west border of Bulgaria with Greece. The problem is that the above mentioned 12 polygons are not located there. In 10 of out of these 12 cases the code 3.1.3.1 Other natural & semi natural mixed forests should be assigned, while for two of them the correct class is 4.2.1.1 Dry grassland with trees.

Another often mistaken class with ***rare spatial distribution*** is 5.2.1.1 Sclerophyllous vegetation. This class is not typical for the country and occupies partially some south-western parts of Bulgaria – the Struma river valley. For 7 out of 12 such polygons the following correct codes are proposed: 3.4.1.1 Transitional woodland and scrub, 4.2.1.1 Dry grasslands with trees (T.C.D. ≥ 30%), 4.2.2.1 Dry grasslands without trees (T.C.D. < 30%) and 3.2.3.1 Other natural & semi natural coniferous forest. An example is given in Figure 4.

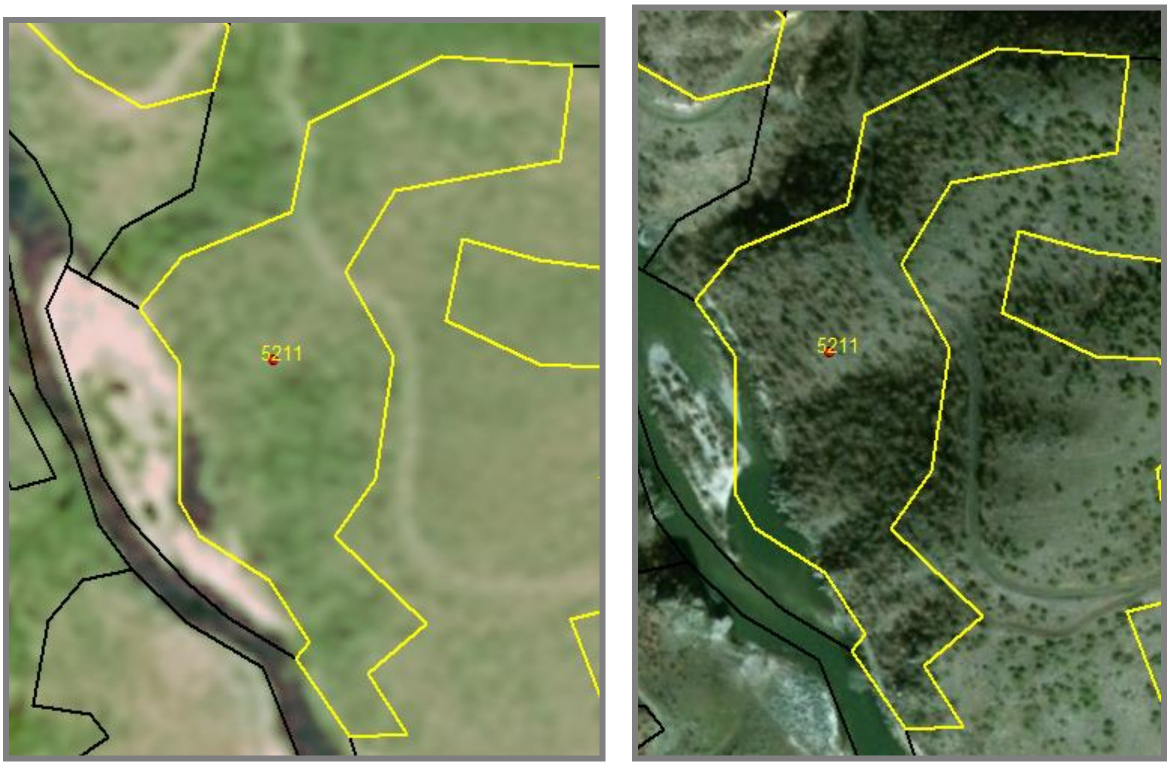

Figure 4. Incorrect classification as 5.2.1.1 Sclerophyllous vegetation. The correct code - 4.2.1.1 Dry grasslands with trees (T.C.D. ≥ 30%) has been proposed (SW Bulgaria, Mesta valley, Blagoevgrad district. Left – GioLand/VeryHighResolution2012, right – World Imagery).

Other typical mistakes are connected with ***mixed classes***. Different spatial combinations of agricultural types have a specific territorial distribution. This is the case with the polygons coded 2.3.3.1 Land principally occupied by agriculture with significant areas of natural vegetation (12 SP – 50.0% TA). This class applies to areas, where parcels of annual or permanent cropland (< 75%) are combined with natural vegetation. The combination of grassland and natural vegetation is excluded from this class. Out of 12 such polygons for 2 a correct code 4.2.1.1 Dry grasslands with trees (T.C.D. ≥ 30%) is proposed, for 1 – 4.1.1.1 Managed grasslands with trees and scrubs (T.C.D. ≥ 30%), for 1 - 4.1.1.2 Managed grasslands without trees and scrubs (T.C.D. < 30%), for 1 – 3.4.1.1 Transitional woodland and scrub and for 1 – 3.2.3.1 Other natural & semi natural coniferous forest.

As a result, in the erroneous coding for this class, misclassification cases with classes of grassland areas are predominant. Example for such a mistake is given in Figure 5. According to Bulgarian LPIS database and the reference imagery, the polygon covers areas with grassland and natural vegetation.

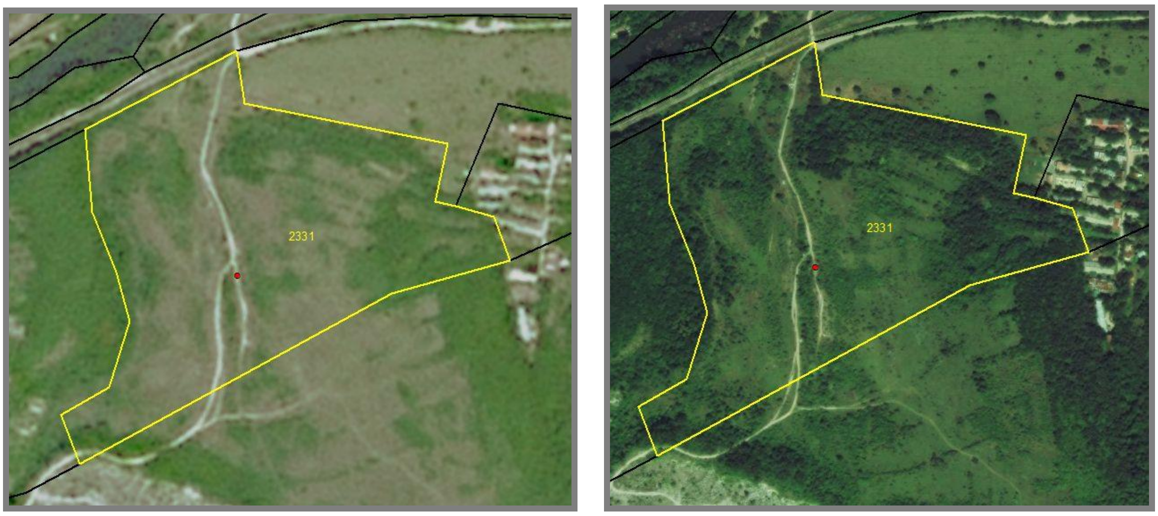

Figure 5. Incorrect class 2.3.3.1. Land principally occupied by agriculture with significant areas of natural vegetation. Correct class is 4.2.1.1 Dry grasslands with trees T.C.D. ≥ 30% (Nord Bulgaria, Samovodene, Veliko Tarnovo district. Left – GioLand/VeryHighResolution2012, right – World Imagery).

Another characteristic of mixed classes causing code confusion is the spatial distribution density of the respective types of objects. This is the case with 2.1.4.1 Complex patterns of irrigated and non-irrigated arable land (11 SP – 8.3% TA).

Figure 6. Incorrect class 2.1.4.1. The mandatory rice fields not present. Correct class is 2.1.1.1 No-irrigated arable land (SW Bulgaria, Hadzhidimovo, Blagoevgrad district. Left – GioLand/VeryHighResolution2012, right – World Imagery).

Permanent irrigated arable land areas in Bulgaria exclusively consist of rice fields, which have limited spatial distribution (principally in the Upper Thracian Plain). In this way, the participation of irrigated lands in this mixed class, as well as in class 2.1.3.1 Irrigated arable land and rice fields comprises irrigated arable lands occupied only by this crop. Figure 6 gives an example of incorrectly classified area as 2.1.4.1 Complex patterns of irrigated and non-irrigated arable land. In the reference aerial images rice fields and elements of an irrigation system and activities are not detected. The proposed correct code is 2.1.1.1 Non-irrigated arable land.

3.1.2. Methodological particularities and omissions in erroneous encoding of Riparian zones

• ***Use of additional sample points in the verification process***

An important specificity of the verification process is reporting encoding correctness around to the so-called *additional sample points* and their location within a single sample polygon.

Sometimes image land characteristics in larger polygons correspond to more than one LCLU class. In these cases, it is necessary code correctness to be checked only in the area (MMU – 5 ha) around the sample point. Correctness of delineation, at the same time, must be checked and referenced according to the entire sample polygon (Maucha et al. 2017). In this regard, often code assigning is considered as wrong. A good example is the case with Sparsely vegetated areas (6.1.1.1 – 58.3% TA) in Figure 7.

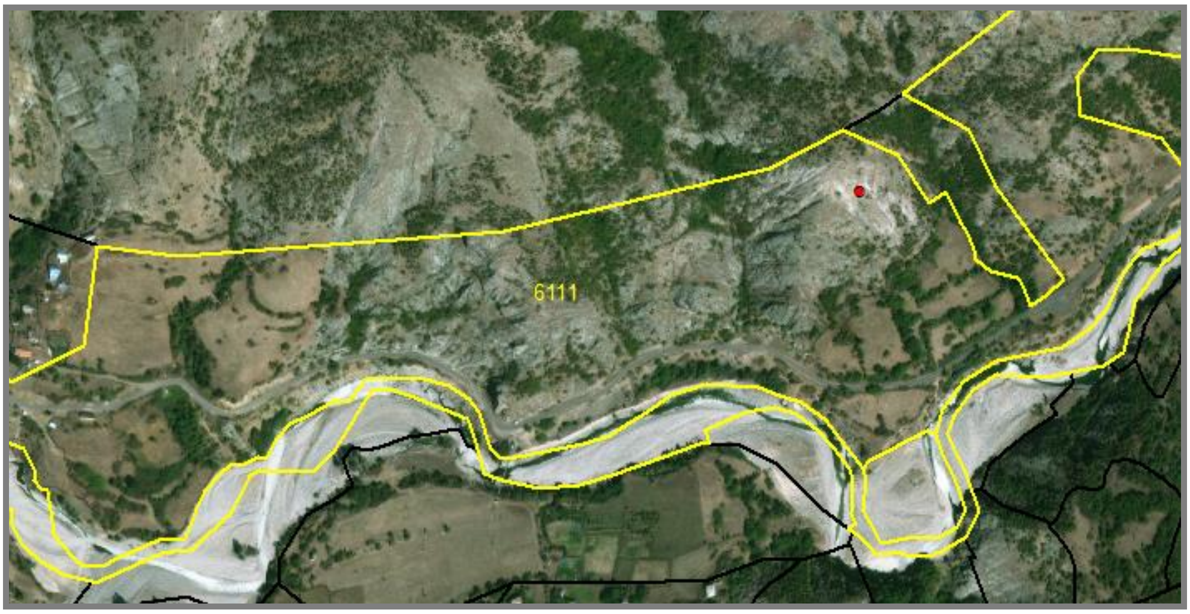

Figure 7. Incorrect class 6.1.1.1 Sparsely vegetated areas. Correct class is 6.2.2.1 Bare rocks and rock debris (South Bulgaria, Borovitsa river region, Kardzhali district. World Imagery).

In Bulgaria, this class includes areas of herbs, grass and scrub with coverage between 10% and 50 %. In many cases, this vegetation is surrounded by rocky surfaces or rocks debris on steep slopes. This class is typical for karstic terrains in the country. In the example the area around the sample point (MMU) represents field vegetation below 10 % so the correct class is Bare rocks and rock debris (6.2.2.1.). The area around the additional point (MMU) represents "Eroded areas with little or no high vegetation (< 10 %)" (EEA, 2015). Separate parts of the sample polygon and one just south of the additional point are grassland and have to be excluded for the correctness of delineation. Other confusions of this class are with Arable land, Broadleaved forest plantations, Grassland and Beaches.

For the same reasons, the sample polygon code 1.4.2.2 Sports and leisure facilities in Figure 8 is considered to be wrong (75.0% thematic accuracy for the class). In the reference images elements of stadium are detected in the western parts of the sample polygon. The MMU area around the additional point falls into an arable land area, which covers the eastern half part of the sample polygon.

Figure 8. Incorrect class 1.4.2.2. Sports and leisure facilities. Arable land occupies a larger part of the polygon area. The correct class is 2.1.1.1 Non-irrigated arable land (SW Bulgaria, Blagoevgrad district).

• ***Use of density layers in the verification process***

The distinction of certain classes at the fourth level of MAES classification is based on the density of specific land cover types, represented by a respective High Resolution Layer (GIO, 2013). Such an indicator is *degree of soil sealing* value which separates urban classes at level 4 into 3 varieties of *Imperviousness Density* (IM.D.): high – 1.1.1.1 Continuous urban fabric (IM.D. >80-100%), middle – 1.1.1.2 Dense urban fabric (IM.D. >30-80%) and low – 1.1.2.1 Low density urban fabric (IM.D. 0-30%). In this respect, an additional source in the incoming database is HR Imperviousness layer, which significantly optimize the coding and verification of these classes. As a result, the accuracy of their coding is higher – 1.1.1.1. – 100.0%, %, 1.1.1.2 – 92.3% and 1.1.2.1 – 91.7%.

Figure 9. Incorrect class 1.1.2.1. Low density urban fabric (IM.D. 0-30%). The correct class is Commercial unit (1.1.1.3) (South Bulgaria, Plovdiv district). GE Street View (to the right) shows commercial activities of the object. To the left - shifted polygon's outline (World Imagery basemap).

Several misclassifications are in other urban classes e.g. commercial unit (1.1.1.3 – 91.7% thematic class accuracy) classified as urban fabric (Figure 9).

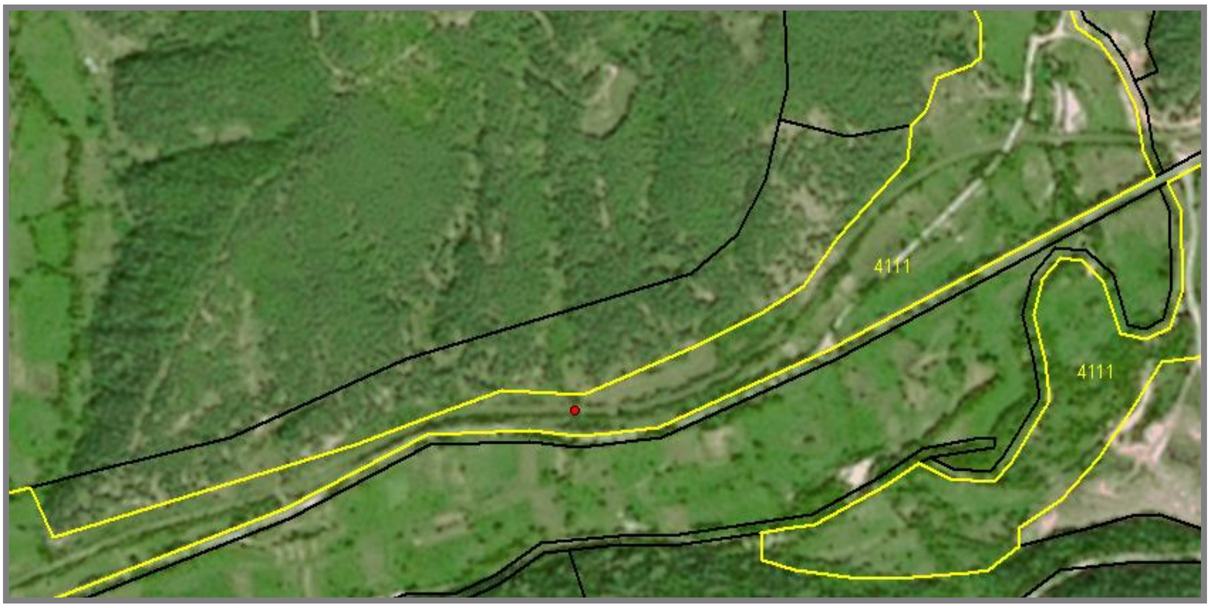

Figure 10. Incorrect class 4.1.1.1. Managed grasslands with trees and scrubs T.C.D. ≥ 30%. Correct class is 4.1.1.2 Managed grasslands without trees and scrubs (T.C.D. < 30%) (SW Bulgaria, Blagoevgrad district. GioLand/VeryHighResolution2012).

Another indicator of coverage density degree is *tree cover density* (T.C.D. ≥ 30% or T.C.D. < 30%). This is typical for defining 2 sub-classes in each of the classes Green urban areas, Sports and leisure facilities. The subdivision of the Managed grasslands, Dry grassland and Mesic grassland classes follows the same principle. Disregarding this percentage results in an erroneous encoding. An example is the incorrect sample polygon coding as Managed grasslands with trees and scrubs T.C.D. ≥ 30% (4.1.1.1 – 75.0% TA) instead of giving to it the neighbouring code in the classification sequence – 4.1.1.2 Managed grasslands without trees and scrubs were the tree cover density is less than 30% (Figure 10).

• *Use of "Potential Riparian Zone" in the verification process*

An important indicator for distinguishing the 4th level of woodland and a forest classes is the "Potential Riparian Zone". This ancillary layer in the input database is a result of the modelling approach and is intended to separate the territories with highest probability of recent riparian features presence (EEA, 2015). This distinguishes "Riparian and fluvial forest" from "Swamp forest" and "Other natural & semi natural forest". According to the methodological mapping approach, the sample polygon area has to be covered by more than 60% by the "Potential riparian zone" product in order to be classified as „Riparian Forest". Ignoring this indicator results in incorrect classification. Relevant example is the class Riparian and fluvial Broadleaved forest (3.1.1.1 – 76.9% TA) in Figure11. In Bulgaria this class comprises broadleaved forests (including artificial broadleaved plantations) that fall into the "Potential Riparian Zone". Most of the sample polygon area in the example is outside of the "Potential Riparian Zone"– more below 60% (the hatching area) which requires a code change to 3.1.3.1 Other natural & semi natural broadleaved forest.

Figure 11. Incorrect class 3.1.1.1. Correct class is 3.1.3.1 Other natural & semi natural broadleaved forest. Partially coarse delineation of dam boundary (North Bulgaria, Al. Stamboliyski lake, Gabrovo district. Left – GioLand/VeryHighResolution2012, right – World Imagery and PRZ layer).

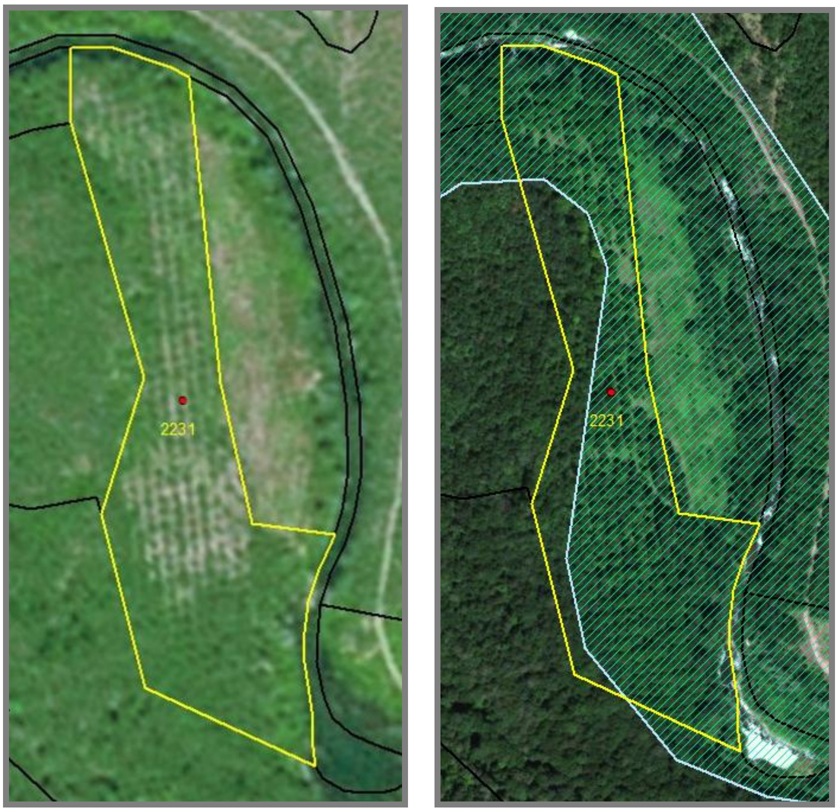

Figure 12. Incorrect class 2.2.3.1. The correct class is 3.1.1.1. Riparian and fluvial broadleaved forest (East Bulgaria, Fakiyska river region, Bourgas district. Left – GioLand/VeryHighResolution2012, right – World Imagery and PRZ layer).

The "Potential Riparian Zone" is also necessary to be taken into account in cases of argumentation for correcting a wrong code. A good example is a code correction of the class 2.2.3.1 Olive groves (0.0 % TA) which has no presence in Bulgaria (Figure 12). The polygon falls into a forest area. In addition, it represents broadleaved plantations that are located inside the "Potential Riparian Zone". This motivates assigning the new correct code 3.1.1.1. Riparian and fluvial broadleaved forest.

3.1.3. Misclassifications due to the Earth objects' characteristics in input imagery

• *Similarities of Earth objects' image characteristics*

In a number of cases the class mismatch is due to the image land characteristics similarity. Such a mismatch is avoided by using layers with additional ancillary information from national map data: vegetation map, topographic map, land cover map and Nature 2000 habitat maps. For example, the separation of forest classes at level 2 of MAES classification is on the base of dominant leaf type of the forest (broadleaved, coniferous or mixed). Often, the land features of Very High Resolution (VHR) images do not show this indicator clearly enough, and the reference only to them throughout the visual interpretation leads to incorrect coding. The most common is the confusion of coniferous with broadleaved forests. In these cases, the use of the ancillary detailed information for different tree species from the national forest management plans is very advantageous. For example, the sample polygon in Figure 13 is coded as Highly artificial coniferous plantations (3.2.4.1 – 75.0 % TA).

Figure 13. Incorrect class 3.2.4.1. Highly artificial coniferous plantations. Correct class is 3.1.5.1 Highly artificial broadleaved plantations (South Bulgaria, Zhrebchevo lake, St. Zagora district. Left – World Imagery and PRZ, right – World Imagery and Forest management plans layer).

In Bulgaria this class comprises coniferous species with artificial planting pattern and visible rows, predominantly intended for wood production. Additionally, representatives of this class have to be outside of the "Potential Riparian Zone". According to the National forest data the polygon area consists of artificial plantation of broadleaved species (Locust) and only about 30% fall in the mentioned zone. In a consequence, the code has been revised to 3.1.5.1 Highly artificial broadleaved plantations.

Detailed soil information from the national forest management plans gives the opportunity to use *secondary indicators* in the verification process. That is the case with the example in Figure 14, concerning the particularities of coding the class Mixed swamp forest (3.3.2.1 – 50,0% thematic class accuracy). In Bulgaria this class comprises forest formations of broadleaved and coniferous trees (including understories of shrub and bush) without the predominance of any of them. Usually they are located on wet terrains or near moors, swamps or marshes. In the sample polygon area broadleaved forest prevails (first indicator). The presence of dry soils (according to Bulgarian Forest data) is a secondary indicator. The third indicator is the "Potential Riparian Zone" which covers more below 60% of the polygon area. This is necessary to distinguish these broadleaved forests from the class of Riparian ones. These three circumstances explain the necessity the wrong code of swamp forest to be corrected to that of Other natural & semi natural broadleaved forest (3.1.3.1).

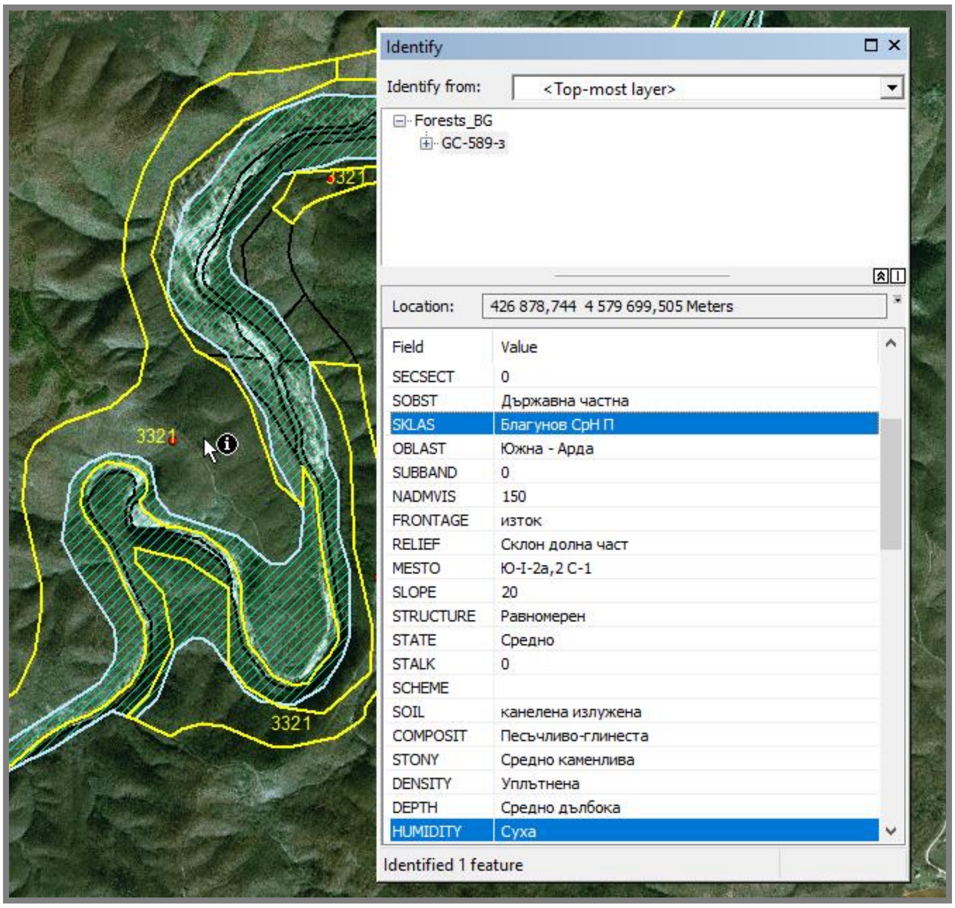

Figure 14. Incorrect class 3.3.2.1. Mixed swamp forest. Correct class is 3.1.3.1 Other natural & semi natural broadleaved forest. (South Bulgaria, South Arda river region, Haskovo district. World Imagery and Forest management plan data).

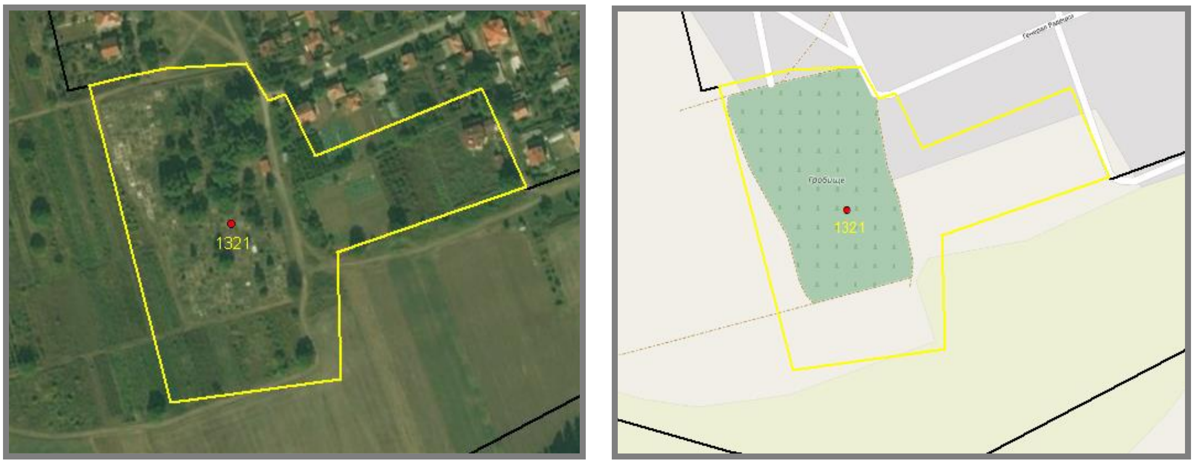

Figure 15. Incorrect classification – cemetery presented as 1.3.2.1 Land without current use. The correct class is 1.1.1.3 Industrial or commercial units (SE Bulgaria, Rousse district. Left – World Imagery, right – OpenStreetMap).

The use of additional information from OpenStreetMap is especially helpful in identifying specific sites, which is necessary to distinguish their belonging to classes with close land characteristics (e.g. class 6.2.1.2 Dunes from class 6.2.1.1 Beaches).

The same source is very useful in specifying the type of the land sites, e.g. cemetery – for code correction from 1.3.2.1 Land without current use to 1.1.1.3 Industrial or commercial units (Figure15). Cemeteries fall to the list of land uses belonging to class 1.1.1.3. Although significant part of the polygon is occupied by complex cultivation pattern, the tested area (5 ha) around the sample point is located within the cemetery territory.

• ***The colour and texture objects' similarities in input imagery***

On the other hand, the similarity of image spectral properties may also be the cause of code mixing. Representative example for this is the class Forest damaged by fire (3.5.1.1 – 0.0% thematic class accuracy) in Figure 16. Two of the three chosen sample polygons represent in reality dry grasslands. This code confusion is due to the ***colour similarity*** (greenish to bluish colour) of these two classes. Although the selected three sample polygons of this class do not reflect the effects of forest fires, this phenomenon is not rare for Bulgaria. Fires occur inside or at the border of forest regions, on south-facing slopes in the dry summer months.

Figure 16. Incorrect class 3.5.1.1 Forest damaged by fire. Correct class is 4.2.2.1 Dry grasslands without trees (T.C.D. < 30%). (South Bulgaria, nearby Galabovo, St. Zagora district. Left – GioLand/VeryHighResolution2012, right – World Imagery).

Figure 17. Incorrect class 4.1.1.2 Managed grasslands without trees and scrubs (T.C.D. < 30%). Correct class is 2.1.1.1 No-irrigated arable land. (South Bulgaria, Haskovo district).

Colour similarities are the reason for the wrong coding of arable land as grassland. In Figure 17 only the north-eastern part of the sample polygon corresponds to the specified class

4.1.1.2 Managed grasslands without trees and scrubs (41.7% TA). The rest greater polygon area is occupied by arable land.

A similar example of such a code confusion is with classes 3.4.1.1 Transitional woodland and scrub and 6.2.2.1 Bare rocks and rock debris were clear-cut area (belonging to classes 3.4.1.1) in broadleaved forest has colour characteristics similar to those of eroded areas (belonging to classes 6.2.2.1).

Another aspect of image characteristics causing thematic errors is the ***similarity of texture characteristics***. Typical example in Figure 18 shows the wrong interpretation of orchard's rows as artificial broadleaved plantations rows (3.1.5.1 – 75.0% TA). In this case, additional information from CORINE Land Cover 2012 (code 222 – Fruit tree and berry plantation) and Land Parcel Identification System (LPIS) Land Cover data concerning the area and the type of agricultural crops (code 020 – permanent crops) gives the solution to such a case.

Figure 18. Incorrect class 3.1.5.1 Highly artificial broadleaved plantations. Correct class is 2.2.2.1 High stem fruit trees (extensively managed). (NW Bulgaria, Vratsa district. Left – GioLand/VeryHighResolution2012, right – World Imagery with CORINE Land Cover 2012 layer).

• ***The ambiguity of Earth objects' appearance***

For some classes the high degree of code confusion is due to the *land characteristics ambiguity*. This is the case with class 9.1.1.2 Intermittently running water courses (12 sample polygons – 25.0% TA). According to the definition (EEA, 2015) this class comprises "Watercourses that cease to flow for part of the year, leaving a dry bed or pools" and corresponds to European Natural Information System (EUNIS) habitat C2.5 – Temporary running waters (Davies, 2004). Class 9.1.1.2 at the same time excludes small rivers with temporary mode caused by summer droughts. In Bulgaria intermittently running water courses comprise all rivers with unstable runoff. This type of rivers is typical for karstic regions (e. g. North-eastern Bulgaria – Dobrudzha region).

The use of national hydrographic data is particularly important in the verification of river regime characteristics (with a riverbed more than 10 m width) and the constancy of the water flow. According to the additional river attribute information most of the selected sample polygons of this class (9.1.1.2) represent rivers with permanent running water course (class

9.1.1.1): Skat, Harmanliyska, Borovitsa, Belenska and Ropotamo river. The frequent reason for this misclassification is the unclear water stream visibility in satellite images. Often the rivers are hidden by riparian vegetation. In Figure 19 the presence of permanent water is concealed by the dense tree crowns.

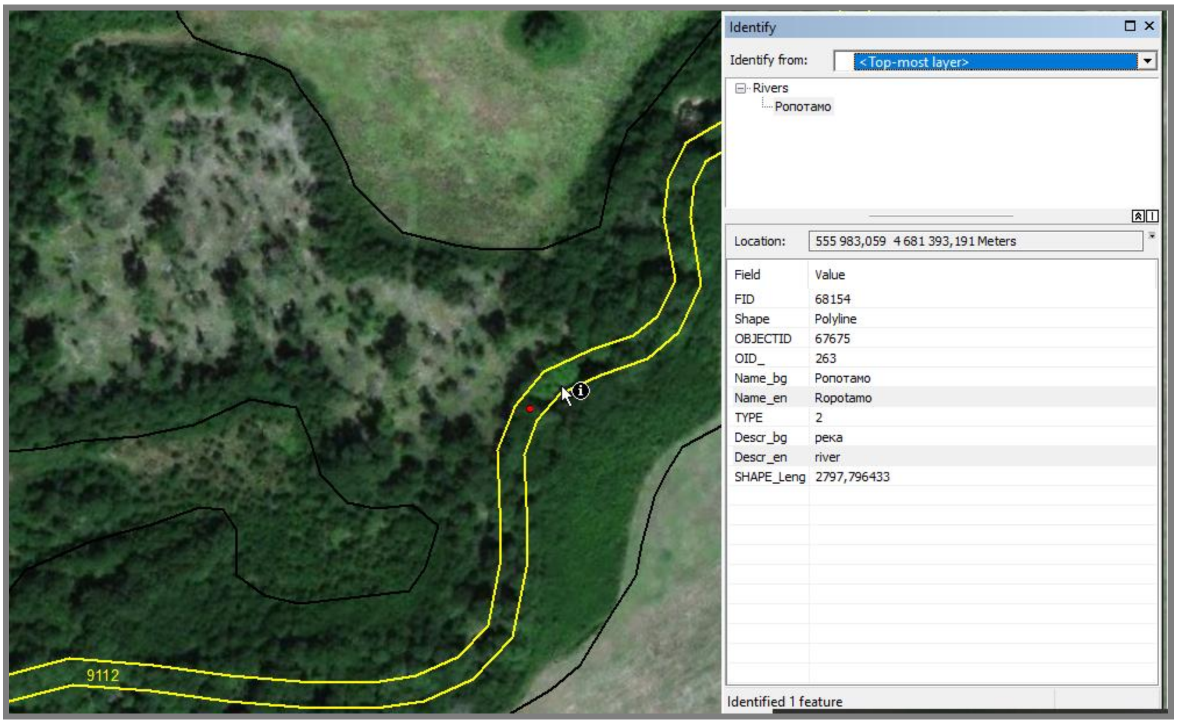

Figure 19. Incorrect classification – 9.1.1.2. Intermittently running water courses. According to Bulgarian Hydrographic data the Ropotamo river is defined as permanent running water course. (SE Bulgaria, Bourgas district. World Imagery with National Hydrographic data layer).

3.2 Correctness of delineation investigations

The second aspect in accuracy assessment - the *Correctness of delineation* - is presented by three components, that are evaluated. The first one, *Delineated area* can take the following values: *Correct*, *Unnecessary parts included* and *Missing parts*. Its score for the whole country is 621 correctly delineated polygons (out of 790 samples), which amounts to the accuracy of 78.6%. The results indicate more common mistakes associated with *unnecessary parts included* (88) than *missing parts* (41) and *both missing parts and unnecessary parts* (40) in the sample polygons.

Some typical examples for unnecessary parts inclusion in delineated areas are, as follows:

- Presence of arable land or grassland in the urban class area (Figure 20);
- Parts of sports facilities included in Green urban areas (Figure 21),
- Included patches from adjacent polygons;
- Presence of patches of natural vegetation in arable land area;
- Grassland inclusion in the forest area;
- River banks included in river areas, etc.

Figure 20. Arable land included in the urban area.

Figure 21. Stadium inclusion in urban park.

The accuracy of the second component – *Detail of delineation* (with possible values *Correct, Too coarse and Too detailed*) is 92.4%. *Too coarse* delineation occurs in fewer cases, according to the obtained results: Correct polygons: 730; Too coarse: 59; Too detailed: 1).

In separate cases omission errors in the sample polygon delineation are due to geometrically incorrect delineation of adjacent polygons. In Figure 22 the road that separates the mineral extraction polygon (class 1.3.1.1 Mineral extraction, dump and construction sites) from its southern neighbour with the same code has a width below 10 m (Minimum Mapping Width (MMW) – 10 m for linear features). Thus, the road together with the adjacent polygon, represents missing parts of the sample polygon.

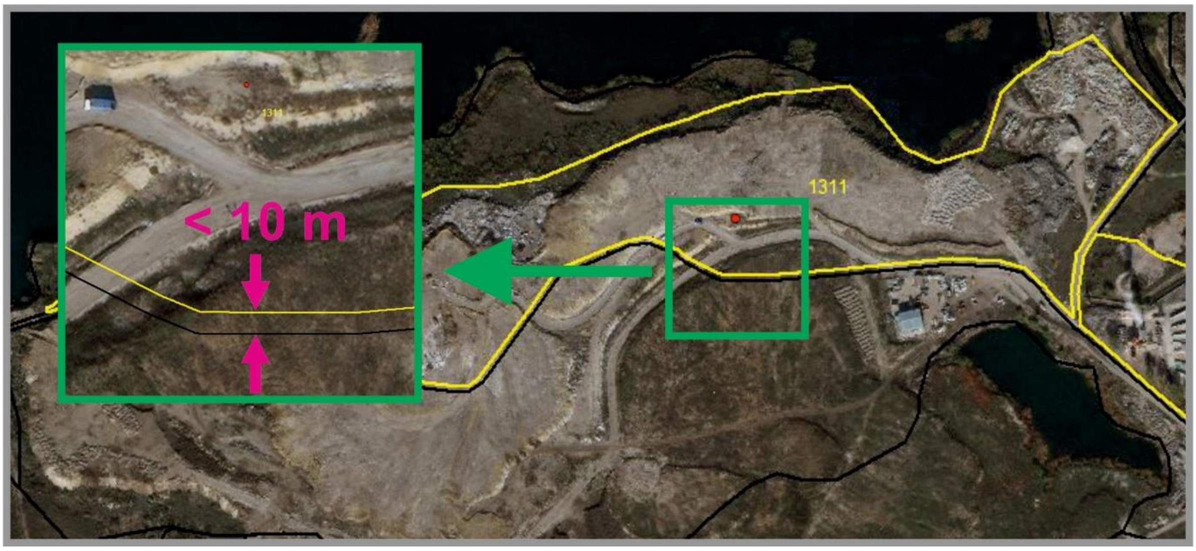

Figure 22. Missing parts and coarse delineation (ballast for inert materials SE from Sofia).

Representative examples for both missing and unnecessary parts presence, coarse delineation and shifted position are shown in Figure 23 and Figure 24.

The evaluation of the third component *Positional accuracy* (taking values *Correct* and *Shifted*) resulted in the value of 97.7%, based on the following scores: Correct polygons: 772; Shifted: 18.

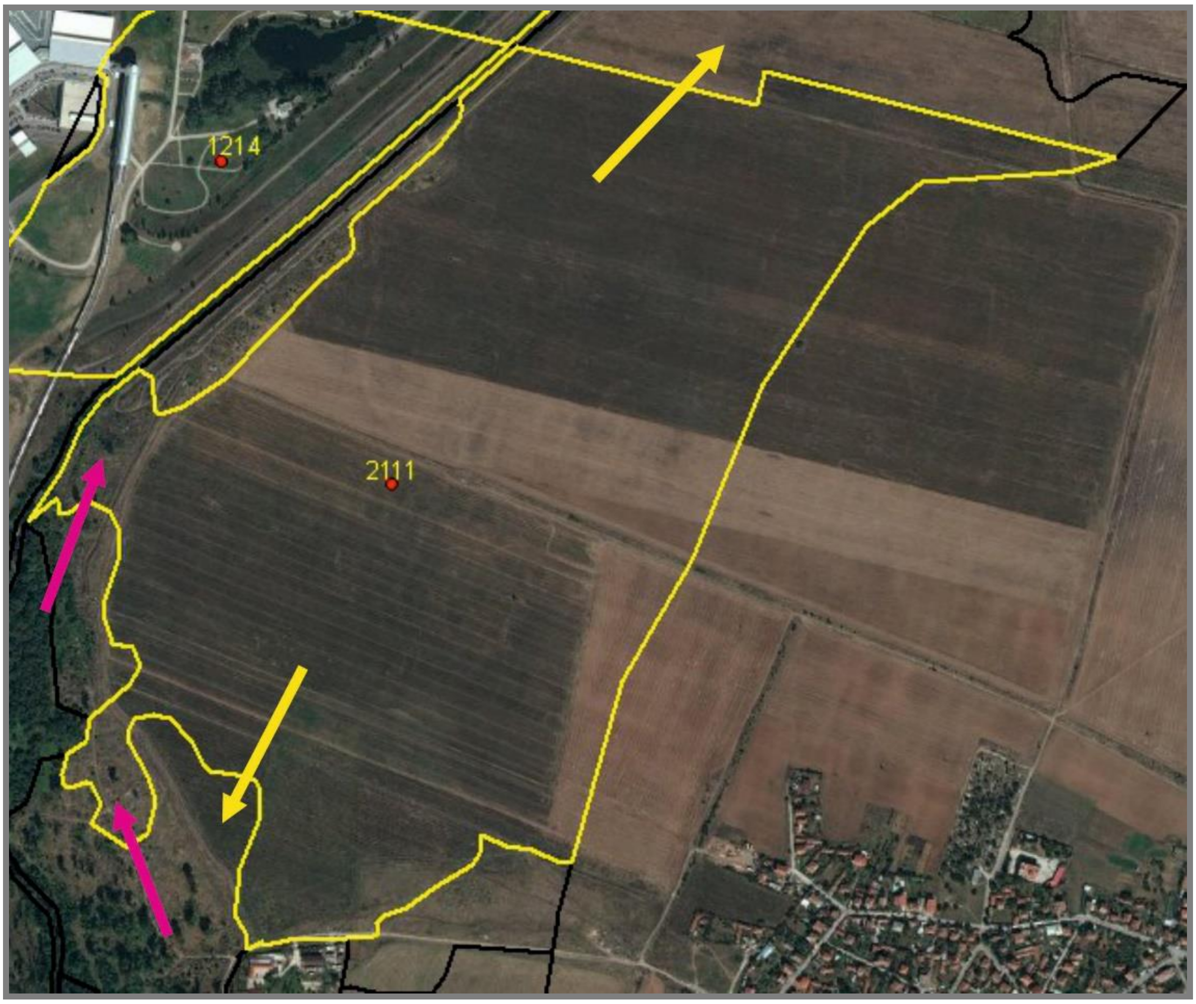

Figure 23. Class 2.1.1.1. Non-irrigated arable land. Both missing parts (yellow arrows) and unnecessary parts (red arrows) included. Coarse delineation.

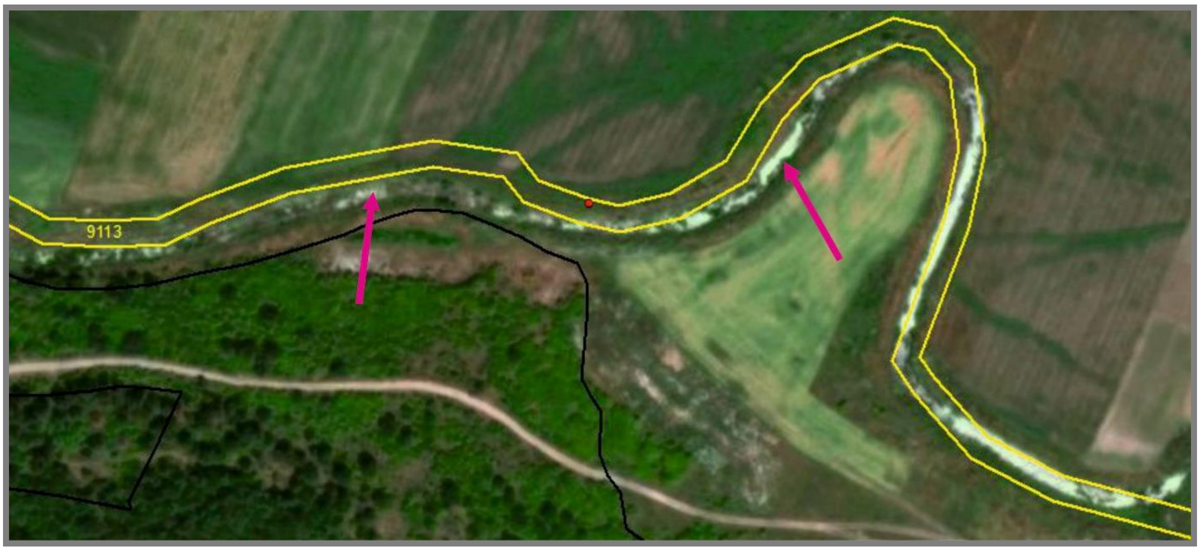

Figure 24. 9.1.1.3 Highly modified natural water courses and canals. The delineation of the polygon is partially shifted.

4. CONCLUSIONS

A quantitative approach to assessing the thematic accuracy of the Riparian zones LC/LU product is applied that is statistically endorsed in view of the objectives pursued. It provides comparability of results between countries. The obtained overall accuracy of 83.50% is a bit lower than the required 85 % in the product specifications. Producer's accuracies for 19 out of 74 classes available in Bulgaria are equal to 100 %. Principally these are linear objects to which another code of the MAES nomenclature can hardly be attributed and also objects with typical and unique interpretation features.

Significant part of the class confusion mistakes appears in classes without or with rare presence in Bulgaria. Another part of the code errors is caused by the lack of specialized, detailed thematic information. A few of them are due to methodological reasons.

The result of the investigations and summaries of the thematic classification errors their species, the frequency and some reasons for admission concludes and confirms the importance of ancillary detailed national information on the land cover/land use in the verification process for achieving high thematic accuracy level.

The assessment of the quality of delineation shows predominance of unnecessary parts inclusion than missing parts and both – unnecessary and missing parts. The boundaries of some polygons are outlined too roughly or shifted. In other cases, the delineation is correct, but polygons can be divided into separate MAES polygons, larger than the Minimum Mapping Unit of 0.5 ha.

The verification of the Riparian zones local component dataset for Bulgaria provides statistical results and quality evaluation based on local expertise, reference data and in situ data used.

The Riparian zones LC/LU dataset provides very good thematic and spatial detail and can be of value for many applications.

ACKNOWLEDGEMENTS

We would like to thank EEA and ExEA for project's funding and supervision! Thanks to Dr. Anton Stoimenov for his contribution and support for the realization of this study!

REFERENCES

Caetano M., F. Mata, S. Freire. 2005. Accuracy assessment of the Portuguese CORINE Land Cover map. *In: 25th EARSeL Symposium on Global Developments in Environmental Earth Observation from Space*, ed. A. Marçal, 459 - 467, Roterdam, Millpress.

Congalton, R. 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sensing of Environment* (37): 35–46.

Copernicus. 2018. Europe's eyes on Earth. Available at: http://land.copernicus.eu/ (accessed 2019-02-08).

Davies C., D. Moss, M. O Hill. 2004. EUNIS Habitat Classification Revised 2004. Report to EEA, 307 p., https://www.eea.europa.eu/data-and-maps/data/eunis-habitat-classification/documentation/eunis-2004-report.pdf (accessed 2019-02-12)

EEA. 2015. Riparian Zones. Final Nomenclature Guideline. Issue 3.0. EEA, 258 p. https://land.copernicus.eu/user-corner/technical-library/RZ_CS3_17_Nomenclature_Guideline_I30.pdf (accessed 2019-02-12).

EEA. 2017. Riparian Zones Land Cover Land Use Validation Report. Issue: 1.3. EEA, 43 p. https://land.copernicus.eu/user-corner/technical-library/rz-lc-lu (accessed 2019-02-12).

Foody G. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment* (80): 185– 201.

GIO. GIO land High Resolution Layers (HRLs) – summary of product specifications - Prepared by: T. Langanke, EEA, 2013, 15 p. https://land.copernicus.eu/user-corner/technical-library/gio-land-high-resolution-layers-hrls-2013-summary-of-product-specifications/at_download/file (accessed 2019-02-12).

LACO-Wiki. 2016. Quick start guide for the LACO-Wiki online tool. Available at: https://LACO-Wiki.net/Content/Docs/LACO-Wiki_quickstart.pdf (accessed 2019-02-07).

Maucha G., R. Pataki, R. Lehoczki, S. Kleeschulte, C. Schröder, D. Abdul Malak, C. Perger. 2017. Guidelines for verification of Local component products 2012, EEA, 43 p.

Nedkov, S, M. Zhiyanski, B. Borisova, M. Nikolova, S. Bratanova-Doncheva, L. Semerdzhieva, I. Ihtimanski, P. Nikolov, Z. Aidarova. 2018. A geospatial approach to the mapping and assessment of ecosystem services in urban landscapes in Bulgaria. *European Journal of Geography*. 9, (4): 34-50.

Tepeliev, Y., R. Koleva, and V. Dimitrov. 2017. Verification of the Natura 2000 local component dataset in Bulgaria. *Forestry Ideas*. 23, 2 (54): 179–192.