High-resolution land cover map 2018

Project deliverable D1.1

Object-based image analysis (OBIA) for mapping Land Cover 2018 in Luxembourg: an approach based on aerial images, LiDAR and intra-annual SENTINEL time series

Version
24.07.2019

space 4 environment

«Using Space to provide Space for the Environment»

DOCUMENT RELEASE SHEET

Book captain:	Karolina Korzeniowska (space4environment)	Sign Date 24/07/2019
Approval:	Stefan Kleeschulte (space4environment)	Sign: Date 24/07/2019

Change Record

Date	Page(s)	Description of Change	Release	Contributing authors
24/07/2019	29	s4e first version including description of high-resolution land cover map 2018 and technical description of the proposed methodological approach based on aerial images.	I1.0	Karolina Korzeniowska, Stefan Kleeschulte

TABLE OF CONTENTS

1 Background of the document 7
1.1 Document content 7
1.2 Related documents 7
2 Introduction 8
2.1 Technical specification of the Land Cover classes 8
3 Study area 9
4 Input datasets 10
5 Methodological approach for land cover mapping with aerial images, LiDAR and SENTINEL 13
5.1 Pre-processing of input data 15
5.2 Segmentation 17
5.3 Classification 18
5.3.1 Ancillary information 18
5.3.2 Time series information 18
5.3.3 Height information 19
5.3.4 Spectral and textural information 19
5.4 Visual inspection & Final refinement 19
5.5 Quality assessment 21
6 Results and Discussion 23
6.1 Remarks while comparing the HRLC'18 with LIS-L LC for 2015: 25
7 Conclusions 26
ANNEX I 27

LIST OF FIGURES

Figure 3-1: Location of test area selected to develop the algorithm for automatic land cover classification of aerial images, LiDAR and time series SENTINEL satellite images. 10

Figure 4-1: Selected scenes and RGB composition of SENTINEL images selected in 2018 for the whole test area of the Grand Duchy of Luxembourg. Area of LiDAR data error. 12

Figure 4-2: Data used for developing the approach for land cover mapping for the year of 2018; a) NIR aerial image for 2018; b) LiDAR digital surface model (DSM); c) LiDAR digital elevation model (DEM); d) ancillary water surfaces data used for classification; e) ancillary transport network (roads and railways) data used for classification; f) ancillary buildings data used for classification; g) ancillary bare soil data used for classification; h) ancillary FLIK data used for classification; and i) time series composition of NDVI gathered from SENTINEL images for 2018. 13

Figure 5-1: Workflow developed in eCognition for mapping the land cover 2018 with aerial images, LiDAR and SENTINEL. The numbers 1, 2, and 3 on figure correspond to the main classification steps. 14

Figure 5-2: Detailed workflow of the LC 2018 classification in the eCognition software. The numbers 5.1-5.4 on image refer to the subsections of this chapter. 16

Figure 5-3: Tiles generated with the eCognition server for classifying the land cover 2018. 17

Figure 5-4: Time series information of the NDVI with regards to different land cover classes gathered based on SENTINEL images. 18

Figure 5-5: Examples of objects updated during visual control of the results of the automatic classification in eCognition. 20

Figure 5-6: Random points across the entire Grand Duchy of Luxembourg for quality check of automatic classification. 21

Figure 5-7: Examples of random points with their ID number that are located close to the border between two classes and were incorrectly classified during automatic classification. 23

Figure 6-1: Land Cover 2018 for the Grand Duchy of Luxembourg with the statistics of the coverage of each LC classes. 24

Figure 6-2: Visual comparison of land cover classification in 2018 (0.2 m spatial resolution) and 2015 (1m spatial resolution) for the Grand Duchy of Luxembourg and an overview of the ortophotomap for 2018. 25

LIST OF TABLES

Table 2-1: Technical specifications of the evaluation land cover product covering the entire region of the Grand Duchy of Luxembourg. 9

Table 4-1: Technical characteristics of input data used for classification. 11

Table 5-1: Accuracy metrics for automatic land cover classification with respect to 2000 randomly selected reference points; Overall Accuracy [%], Producer's Accuracy [%]; Users Accuracy [%], Errors of Commission [%], and Errors of Omission [%]. 22

LIST OF ACRONYMS

CIR	Colour Infrared
DEM	Digital Elevation Model
DSM	Digital Surface Model
EO	Earth Observation
HRLC'18	High-resolution land cover map 2018
LC	Land Cover
LIS-L	Land Information System Luxembourg
LU	Land Use
LiDAR	Light Detection and Ranging
LIS-L	Land Information System Luxembourg
MDDI	Ministère du Développement Durable et des Infrastructures
NDVI	Normalized Difference Vegetation Index
nDSM	normalised Digital Surface Model
NIR	Near Infrared

1 BACKGROUND OF THE DOCUMENT

The high-resolution land cover map 2018 (HRLC'18) project has the objective to develop and implement an object-based image analysis (OBIA) approach to obtain land cover information based on near infrared (NIR) aerial images for the Grand Duchy of Luxembourg.

A major component of the project is to produce an up-to-date land cover map of the Grand Duchy of Luxembourg for 2018 with 0.2 m spatial resolution. In addition, we tested the importance of using 3D topographical data obtained based on Light Detection and Ranging (LiDAR) and intra-annual SENTINEL time series to obtain a high quality land cover map. The project is a follow-up of Land Information System Luxembourg (LIS-L) project aiming at mapping land cover (LC) and land use (LU) in Luxembourg for 2015.

1.1 DOCUMENT CONTENT

This document provides the methodological approach, which was tested for the Luxembourg region of the Grand Duchy of Luxembourg. The goal is to progress on the methodology based on Pleiades images and spot5/take5 which was used to classify land cover of the Grand Duchy of Luxembourg for 2015 and develop a methodological approach for dynamic land cover mapping in Luxembourg based on colour infrared aerial images, LiDAR data, and intra-annual time-series of SENTNEL-2A images. The proposed methodology is developed with applying image segmentation and object classification approach.

The document is structured in seven main sections. Section 1 introduces the content of this document and main goals of the project. Section 2 introduces the most important information about the project and its technical specifications. Sections 3 and 4 provide an overview description of study area and used data, while section 5 offers a detailed description of the methodology used to produce Land Cover map. Remaining sections 6 and 7 provided details on the generated land cover maps and their comparison with Land Cover maps generated for 2015 from Pleiades satellite images and conclusions drawn from the project.

1.2 RELATED DOCUMENTS

Document ID	Descriptor
20181203_Commande_Space4Environment.pdf	Project proposal

2 INTRODUCTION

High-resolution aerial images allow detecting objects with small spatial extent such are single trees, cars, and small building constructions. The use of near infrared (NIR) photos gives the opportunity in applying vegetation and water indices to increase the correctness in classifying objects.

Still, the use of high-resolution data is a challenge, because it also adds noise to the classification when performing it on a pixel level. In addition, shadows of vegetation and high buildings affect the precise determination of boundaries between classes. Moreover, aerial images collected only for one time stamp and having few spectral bands create difficulties to the automatic description of some land cover classes with the same properties for a wider region, such as young forest vs. bushes and densely vegetated grasslands or buildings vs. artificial areas. With the increased availability of satellite data containing more spectral information and higher temporal resolution, and LiDAR data containing 3D information it would be possible to address these issues to improve automatic detection of individual objects.

In this study, we present a follow up of the methodological approach described in Deliverable 2.4 of the LIS-L project for the entire area of Luxembourg. The goal is to progress on the methodology based on aerial images which we tested to classify land cover of the Grand Duchy of Luxembourg for 2017 and develop a methodological approach for land cover mapping 2018 in Luxembourg based on colour infrared aerial images, LiDAR and intra-annual time-series of SENTNEL-2A images. The proposed methodology is developed with applying image segmentation and object classification approach.

The aerial image campaigns are organized for the Grand Duchy of Luxembourg at an annual interval, while SENTINEL images are gathered with 5 days interval for the same area. This means that a new set of very high-resolution RGB and Colour Infrared (CIR) data available once each year, and a large number of lower resolution satellite images can be gathered during the whole year. So far, only one LiDAR campaign has been done that covers the entire country. The time stamp of the LiDAR data is up to October 2017. Combining these three datasets may be an advantage in verifying the yearly LC changes across the whole country.

2.1 TECHNICAL SPECIFICATION OF THE LAND COVER CLASSES

We performed the land cover classification into nine classes described in Table 2-1. In comparison with the LC classification carried out in the LIS-L project, we introduce two additional classes, which are “bushes” and “vineyards”. The “bushes” class has been created due to the availability of LiDAR data for the entire Grand Duchy of Luxembourg, which allows the separation of trees from bushes. The “vineyards” class was introduced because after discussions with MDDI we concluded that this type of land cover does not apply to any of classes used in the LIS-L project and therefore should be treated separately. In order to be consistent with the classification for 2015, we kept the terminologies and individual numbers of classes as it was used in the LIS-L project. For classes “bushes” and “vineyards” we introduced new coding which were 8 and 93 respectively.

For a detailed description of HRLC’18 land cover classes, please see Annex I.

Table 2-1: Technical specifications of the evaluation land cover product covering the entire region of the Grand Duchy of Luxembourg.

Product	**Land cover map of Grand Duchy of Luxembourg**
Content	The product consists of a land cover map of the entire region of Grand Duchy of Luxembourg derived from high-resolution aerial images, LiDAR data, EO images and publicly available ancillary data. It includes the following final land cover classes: • (1) Buildings, • (2) Other constructed areas, • (3) Bare soil, • (6) Water, • (7) Trees • (8) Bushes, • (91) Permanent herbaceous vegetation, • (92) Seasonal herbaceous vegetation, • (93) Vineyards.
Input data	Aerial images: NIR (0.2m spatial resolution) Satellite images: SENTNEL-2A (10m spatial resolution) LiDAR data: DEM and DSM (1m spatial resolution) Ancillary data: building census (combination of topographic and cadastre maps); transport network (LIS-L land Use product for 2015); water surfaces (topographic map), FLIK agriculture data
Temporal Requirement	2018
Spatial coverage	entire region of Grand Duchy of Luxembourg

3 Study area

We selected a test area covering 100km² of the Grand Duchy of Luxembourg to develop an automatic OBIA algorithm for recognising land cover classes. Our test area, presented on Figure 3-1, covers three cantons of Luxembourg (Capellen, Luxembourg and Esch-sur-Alzette) in the southern part of the country. The area is characterised by diverse land cover types, including eight of nine classes described in Table 2-1. The class of vineyards does not apply for the selected area, because this type of vegetation is cultivated only in the eastern part of the country. Therefore, for this class we performed additional tests for an area where this type of land cover occurs. After verifying the accuracy of the classification, we applied the OBIA algorithm for the entire country of Luxembourg.

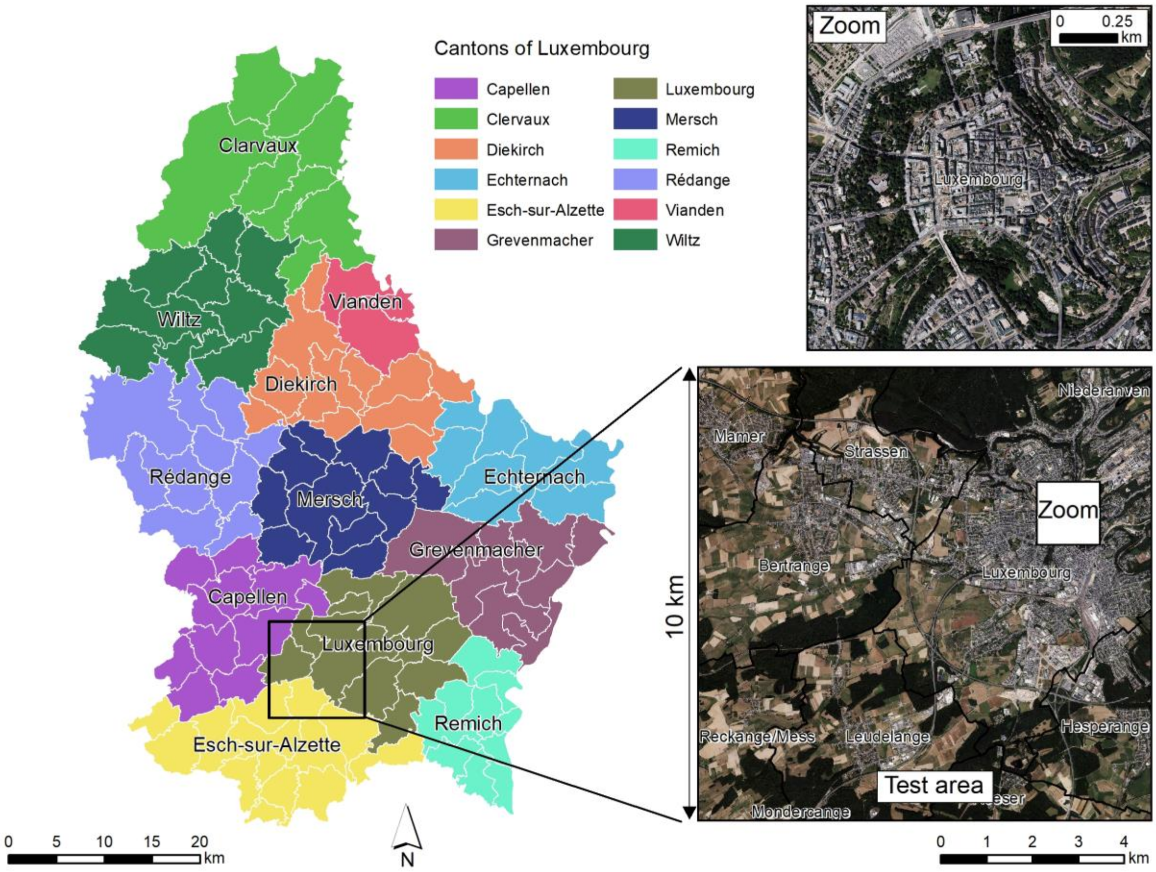


Figure 3-1: Location of test area selected to develop the algorithm for automatic land cover classification of aerial images, LiDAR and time series SENTINEL satellite images.

4 Input datasets

We developed an algorithm for producing the land cover map based on high-resolution NIR aerial images with the spatial resolution of 0.2m gathered in 2018. We used these data as main input dataset to develop the automatic OBIA classification algorithm. From the images, we evaluated normalised difference vegetation index (NDVI) with the equation 1, which allows us to separate biotic and abiotic classes. Other parameter that we used was brightness of images evaluated using the algorithm implemented in the eCognition software.

$$(1) \quad NDVI = NIR\text{-}Red/NIR+Red$$

where:

NIR – near infrared band of image

Red – red band of image

We used a 1 m spatial resolution LiDAR data gathered in October 2017 to classify biotic classes, and to separate bushes from trees. From digital surface model (DSM) and digital elevation model (DEM) we evaluated normalized difference surface model (nDSM) using the equation 2.

$$(2)\quad nDSM = DSM - DEM$$

where:

DSM – digital surface model

DEM – digital elevation model

Additionally, we used time series of Earth observation (EO) SENTINEL images, listed in Figure 4-1, with 10m spatial resolution collected across 2018 as supporting data for classification of the vegetation phenology with the NDVI index (equation 1). To do this, we collected seven SENTINEL images from the Copernicus online portal (https://cophub.copernicus.eu/dhus/#/home) for different dates of the Grand Duchy of Luxembourg. We used three scenes of SENTINEL images to cover the whole area of Luxembourg, which were:

- T31UFR,
- T31UGR,
- T31UGQ.

From all images, we generated a multiband raster dataset that is shown in Figure 4-2.

We also applied ancillary data taken from the LU products of LIS-L project and topographical map of the Grand Duchy of Luxembourg to help classifying buildings, road and railway networks as artificial areas, and water surfaces. See Table 4-1 for more details on input data. We verified visually the ancillary data with aerial images for 2018 and improved them in areas where the major land cover changes occurred, for example new highway or where new large buildings were built.

Table 4-1: Technical characteristics of input data used for classification.

INPUT DATASETS			
Name	**Type**	**Spatial resolution**	**Source**
Water	Vector	-	MDDI
Railways	Vector	-	MDDI/LIS-L
Roads	Vector	-	MDDI/LIS-L
Buildings	Vector	-	MDDI/LIS-L
Bare soil	Vector	-	LIS-L
FLIK data	Vector	-	MDDI
NIR images	Raster	0.2m	MDDI
NDVI (derived from NIR images)	Raster	0.2m	MDDI
Brightness (derived from NIR images)	Raster	0.2m	MDDI
Digital Elevation Model (DEM)	Raster	1m	MDDI
Digital Surface Model (DSM)	Raster	1m	MDDI
Difference Elevation Model (nDSM)	Raster	1m	MDDI
SENTINEL images	Raster	10m	EO- Browser
NDVI (derived from SENTINEL images)	Raster	10m	EO- Browser

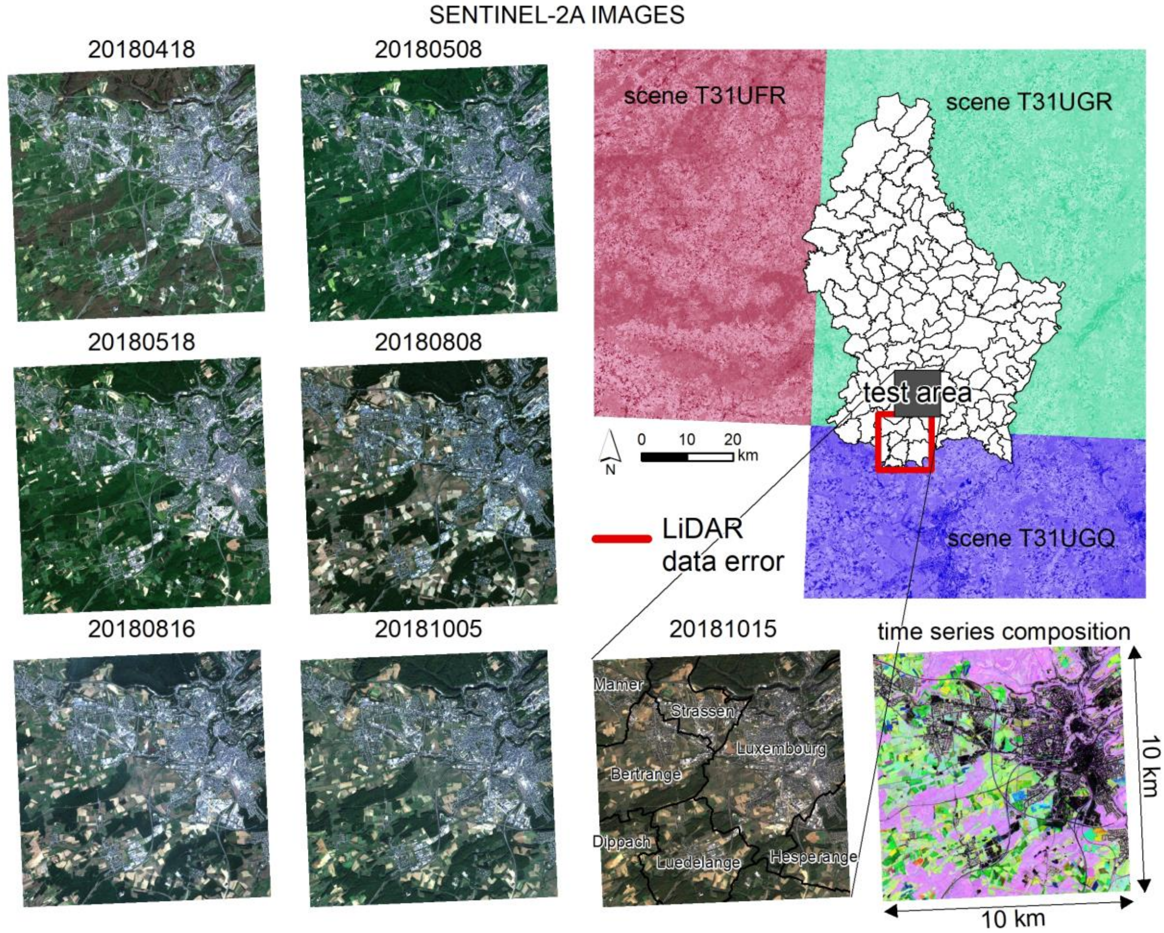


Figure 4-1: Selected scenes and RGB composition of SENTINEL images selected in 2018 for the whole test area of the Grand Duchy of Luxembourg. Area of LiDAR data error.

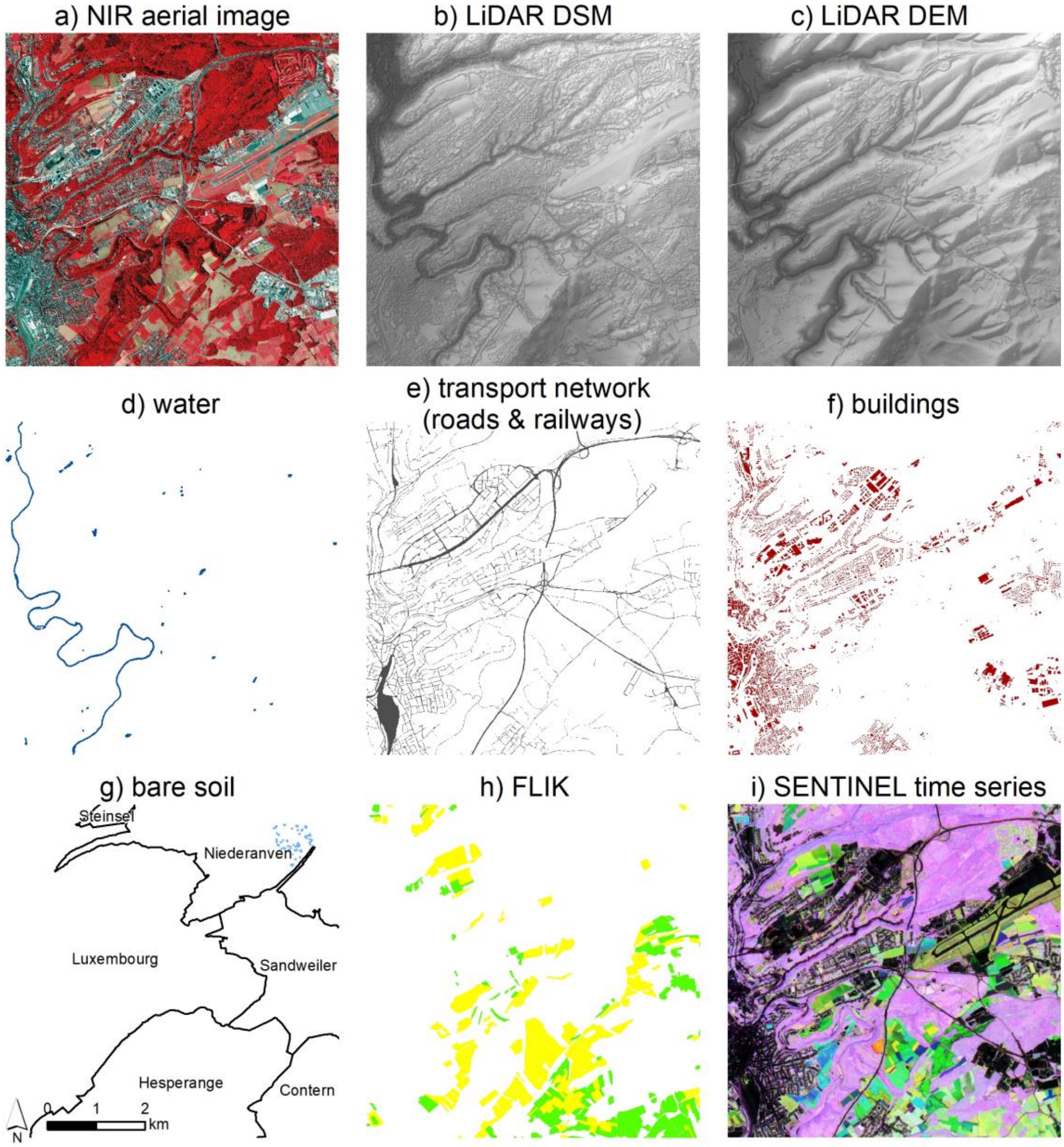


Figure 4-2: Data used for developing the approach for land cover mapping for the year of 2018; a) NIR aerial image for 2018; b) LiDAR digital surface model (DSM); c) LiDAR digital elevation model (DEM); d) ancillary water surfaces data used for classification; e) ancillary transport network (roads and railways) data used for classification; f) ancillary buildings data used for classification; g) ancillary bare soil data used for classification; h) ancillary FLIK data used for classification; and i) time series composition of NDVI gathered from SENTINEL images for 2018.

5 Methodological approach for land cover mapping with aerial images, LiDAR and SENTINEL

We performed the classification in the eCognition software with the server license. We selected one software that allows performing grouping of pixels into homogenous segments representing the same object, and implementing assumptions based on neighbouring relation of segments. In addition, the classification in eCognition can be performed on different scales, which allows using either top-down or bottom-up approach for designing the workflow.

The main two steps of automatic object recognition, which are possible to do in eCognition, are: 1) image segmentation where the pixels are grouped into bigger homogenous objects, and 2) image classification where the objects are assigned into a specific class. Image segmentation can be performed with few segmentation algorithms such as chessboard segmentation, quadtree segmentation, contrast split segmentation, and multiresolution segmentation that can be selected according to the purpose of classification and used dataset. The segmentation methods allow the user to set parameters defining the size of segments. In addition, multiresolution segmentation offers specifying the level of segments homogeneity and the compactness of segment shape what allows advanced delineation of object boundaries comparing to others above-mentioned segmentation algorithms.

Image classification can be performed in few different ways. 1) One can select samples and perform an automatic supervised classification of segments; 2) one can also apply membership functions or thresholding to classify segments; farther 3) one can use geometrical, textural, or thematic information of segments to classify them; or finally 4) one can use the positional and relational assumptions of segments to classify specific objects. All of these approaches can be applied separately as well as in combination with each other to build a complex workflow of classification.

The flowchart in Figure 5-1 shows the main steps of our classification approach developed in the eCognition software. The detailed classification is described in further subsections of this chapter.

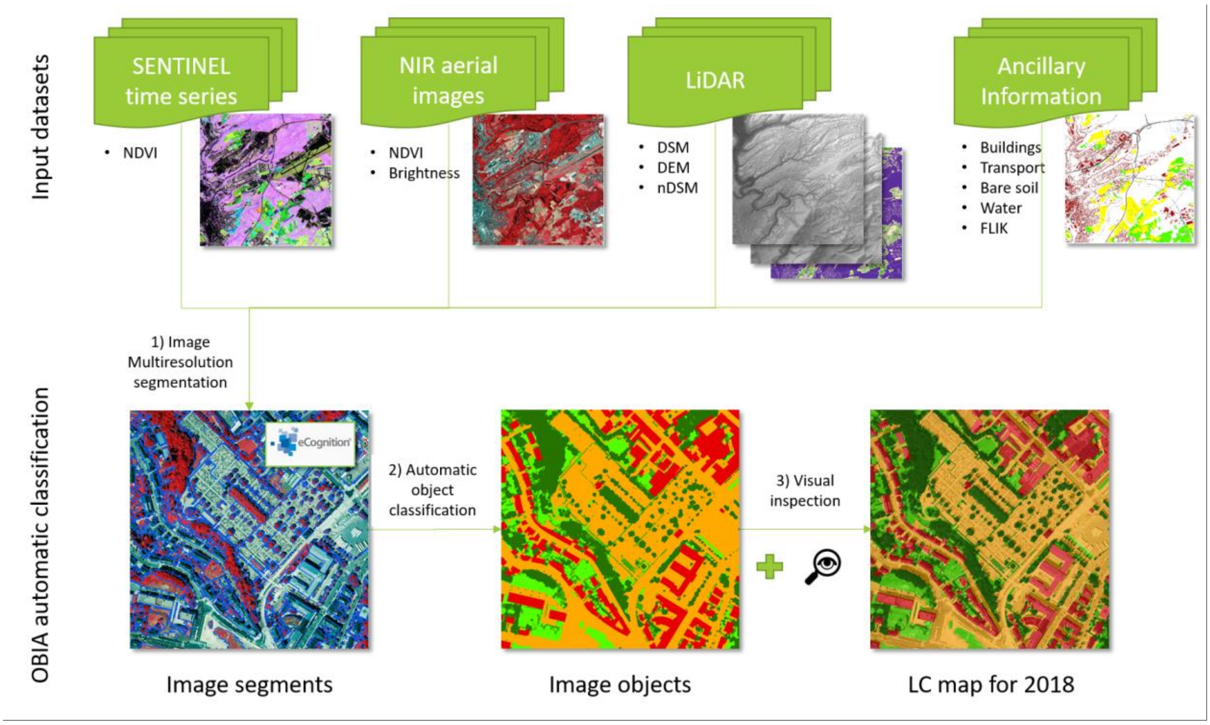


Figure 5-1: Workflow developed in eCognition for mapping the land cover 2018 with aerial images, LiDAR and SENTINEL. The numbers 1, 2, and 3 on figure correspond to the main classification steps.

5.1 PRE-PROCESSING OF INPUT DATA

We performed the classification in the eCognition software with the server license, which allows the processing of large datasets in smaller parts called tiles. We started the classification from creating a new workspace, adding a project, and importing all datasets into the eCognition software (Figure 5-2). The imported layers were as follow:

a) Image layers

- Layer 1 – aerial image – Band NIR;
- Layer 2– aerial image – Band Red;
- Layer 3 – aerial image – Band Green;
- Layer 4 – LiDAR derived nDSM for 2017;
- Layer 5 – SENTINEL derived NDVI for 2018-04-18;
- Layer 6 – SENTINEL derived NDVI for 2018-05-08;
- Layer 7 – SENTINEL derived NDVI for 2018-05-18;
- Layer 8 – SENTINEL derived NDVI for 2018-08-08;
- Layer 9 – SENTINEL derived NDVI for 2018-08-16;
- Layer 10 – SENTINEL derived NDVI for 2018-10-05;
- Layer 11 – SENTINEL derived NDVI for 2018-10-15;

b) Thematic layers

- Layer 1 – buildings from the topo map and LIS-L LC map of 2015;
- Layer 2 – roads and railways from the LIS-L LC map of 2015;
- Layer 3 – water surfaces from the topo map of 2015;
- Layer 4 – bare soil data from the LIS-L LC map of 2015;
- Layer 5 – FLIK data.

Additionally we used the Luxembourgish border data with 10 m buffer which we set as a limit of our calculations and the border of final land cover map.

In the second step, we created the tiles that were used to classify the data in the entire area of Luxembourg. After performing tests we finally chose 18.000 × 18.000 pixels (3.6 × 3.6 km) as the best size that allows smooth processing of data and minimize the number of created tiles. Figure 5-3 shows that in total we had 248 tiles overlapping with aerial images.

In the following step, we set the bands used to evaluate the Brightness value; we selected only the three first Layers which were NIR, red and green of aerial images. We also used as layer arithmetic's the NDVI evaluated based on aerial images with the equation 1.

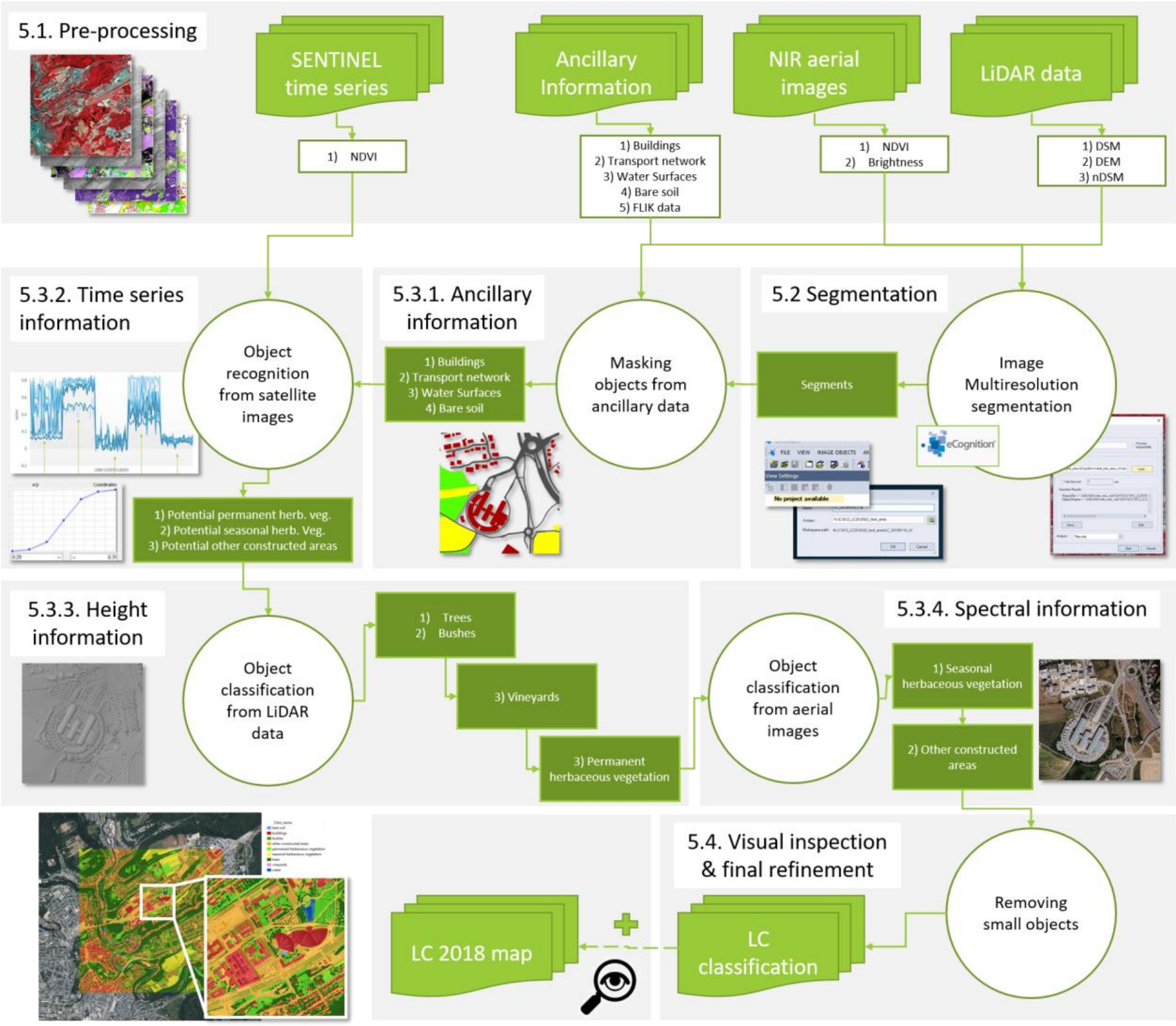


Figure 5-2: Detailed workflow of the LC 2018 classification in the eCognition software. The numbers 5.1-5.4 on image refer to the subsections of this chapter.

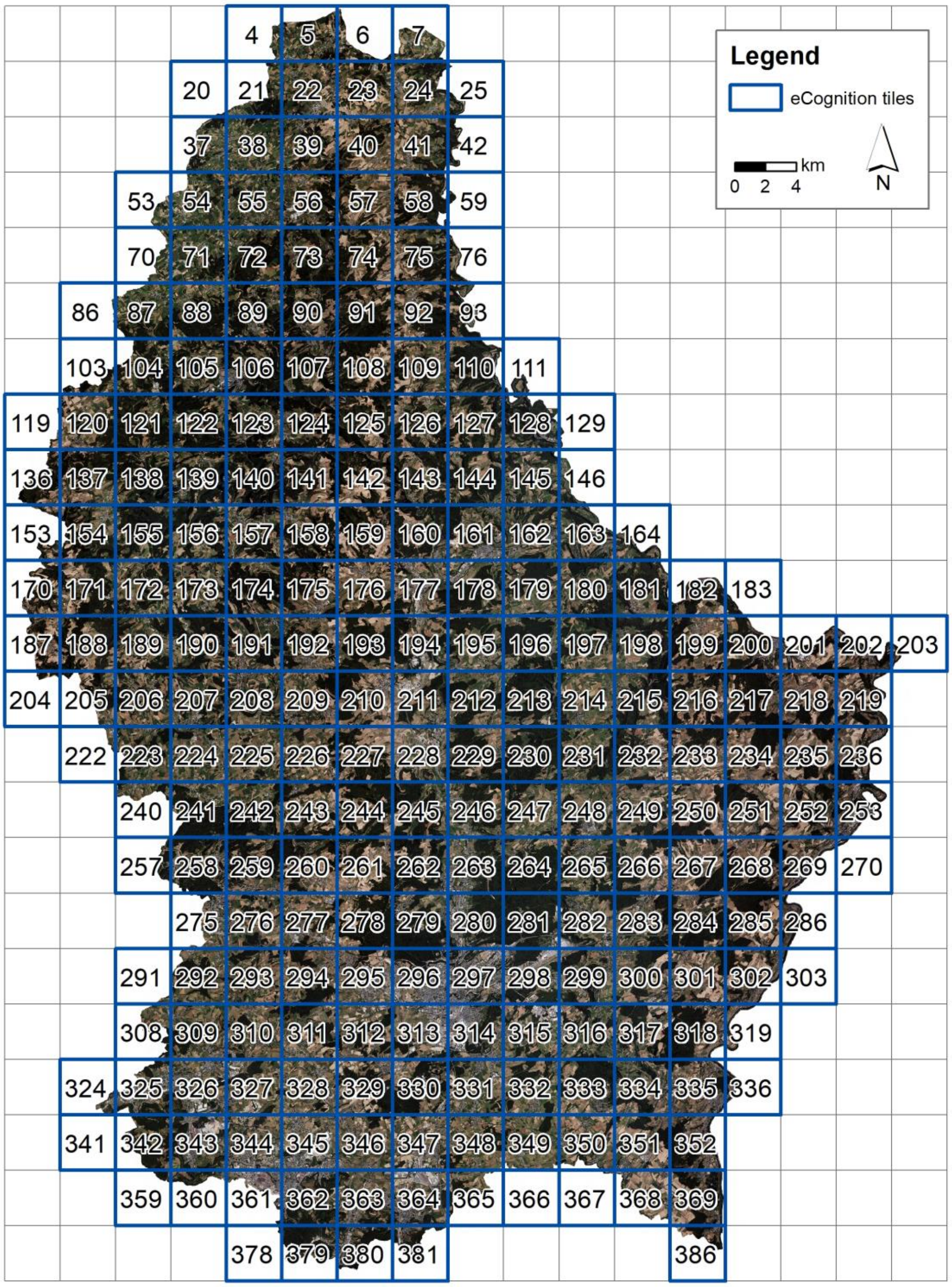


Figure 5-3: Tiles generated with the eCognition server for classifying the land cover 2018.

5.2 SEGMENTATION

In the following, we performed the multiresolution segmentation of aerial images. We applied the segmentation at level 20 with shape of segments equal to 0.3 and the compactness equal to 0.5 as the best math with our input data. The parameters we chosen after performing several tests and visual verifying the results.

In this step, we excluded all layers derived from SENTINEL images – our motivation here was to exclude the influence of the spatial resolution of SENTINEL images on segment shape. The aerial images we used have 0.2m spatial resolution, while SENTINEL only 10m, hence applying SENTINEL in the segmentation process, even with lowering their influence on the segmentation process, would privilege delineating the segment borders along the SENTINEL pixels thus limiting the scope of data that can be obtained from aerial images. We conclude this from tests performed in the developing stage of segmentation, which are presented in Delivery D2.4 of the LIS-L project.

For the vector ancillary data, we used a higher power in evaluating the segments to keep the original boundaries of objects such as buildings, water, transport network, bare soil and agricultural fields.

5.3 CLASSIFICATION

5.3.1 ANCILLARY INFORMATION

In the next step, using the ancillary information on thematic layers 1 – 4 we classified automatically all segments that corresponds to our buildings, roads and railway network, bare soil and water.

5.3.2 TIME SERIES INFORMATION

Further, we analysed the NDVI indices derived from the temporal SENTINEL datasets to classify artificial areas, and seasonal and permanent herbaceous vegetation. Figure 5-4 shows, that different land cover objects have diverse NDVI values over the year what is significant in distinguishing them from each other. Hence, we used the membership functions to separate the segments in a sequential order into the classes:

- Potential artificial areas – other remaining dataset (unclassified);
- Potential seasonal herbaceous vegetation – other remaining dataset (unclassified);
- Potential permanent herbaceous vegetation – other remaining dataset (unclassified).

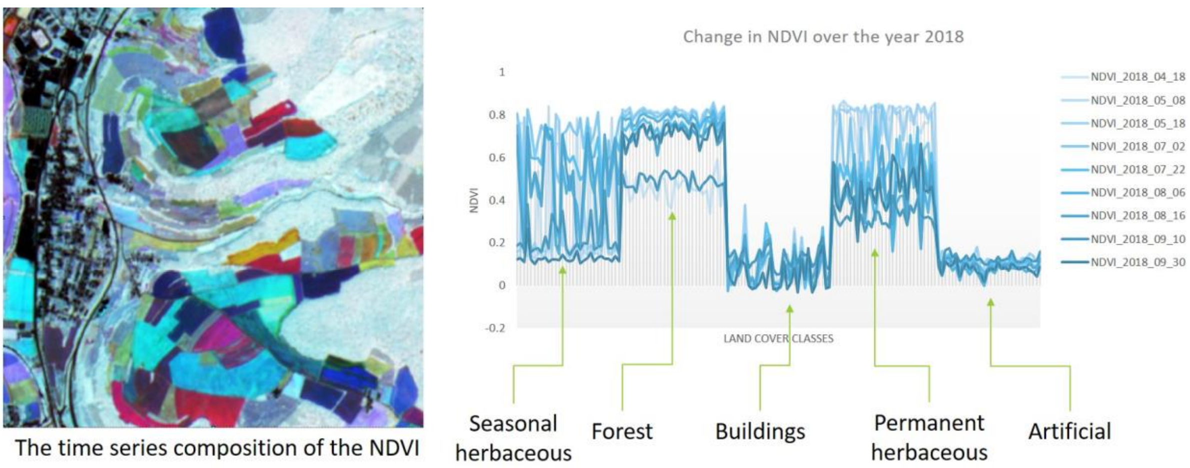


Figure 5-4: Time series information of the NDVI with regards to different land cover classes gathered based on SENTINEL images.

5.3.3 Height information

We conducted the classification of trees, bushes and vineyards using LiDAR data. As a threshold separating trees from bushes, we set the height of 3 m. In the first step, we have reclassified as trees all segments with a height exceeding 5 m and NDVI threshold representing a clear occurrence of vegetation. Similarly, all segments with a height above 0.5 m and NDVI indicating vegetation we reclassified as bushes. Then, we performed the chessboard segmentation for the class of bushes with a segment size equal to one pixel. We did this to finding the exact boundary between the trees and bushes, which was difficult in case of multiresolution segmentation. All segments from the class of bushes, whose height attribute exceeds 3 m, have been reclassified to the trees class. Then we removed the salt-and-pepper noise, which arose by applying the chessboard segmentation, reclassifying the smallest segments to the neighbours. Subsequently, we reclassified into vineyards all bushes that were in the vineyards polygons in the FLIK data, assuming that they represent grapevine plants. Finally, we reclassified all the other areas with the NDVI threshold representing vegetation and the height lower than 0.5 m as a permanent herbaceous vegetation.

5.3.4 Spectral and textural information

In the next step, we classified seasonal herbaceous vegetation. We assigned all remaining segments to this class, which have a low NDVI in the time-series SENTINEL and aerial images, and are in the seasonal herbaceous vegetation class in the FLIK data. We have not simply reclassified all the segments that are in the FLIK data, because we have noticed that in some cases these data are not up to date, for examples some fields exist in areas that are under construction, or buildings already exists, or they include permanent herbaceous vegetation with significant entry of wood vegetation. Then we merged the segments, verified the area of the created objects, and reclassified all objects with a size smaller than 100,000 pixels = 1,000m² =0.1ha (1 pixel = 0.2 × 0.2m) to the neighbouring object, stating that this area is too small to be able and represent a cultivated field.

As the last class, we performed the classification of artificial areas. Here we reclassified all segments from the 'potential artificial' class that have the NDVI value on aerial images corresponding to no vegetated areas (NDVI<0.0). Because the NDVI may vary depending on the used input dataset, we also included segments with the NDVI >0, but adjacent to buildings and roads. In order to avoid incorrect classification of mixed pixels coming from the resolution of the SENTINEL data we also considered the classification of segments located near buildings, even though they are assigned to a 'potential artificial' class or not. Using the neighbourhood assumptions, we grow this class to include additional segments.

5.4 VISUAL INSPECTION & FINAL REFINEMENT

In the last step, we classified all remaining segments using nearest neighbour classification to the largest neighbour. Here we set the hierarchy of the classification, setting the target class as artificial, seasonal herbaceous vegetation, and permanent herbaceous vegetation. Only the remaining segments, which could not be reclassified to the above classes, we reclassified to another class.

During the visual inspection of the results we noticed that in the southern part of the country the input LIDAR data contains errors (shown in Figure 4-1); some parts of forests were missing, and we had to manually draw those missing forest patches and single trees. Therefore in the area of around 140 km² the borders of trees class may visually slightly differ from those in the remaining part of the country.

The final classification we exported as polygon Shapefile and GeoTiff file and verified visually with aerial images. We also compared our classification to the results of the LC for 2015 performed based on 1m spatial resolution Pleiades satellite images (Figure 5-5) – please see Project deliverable D2.1: "Product Delivery Documentation for details on the LC map produced with Pleiades data".

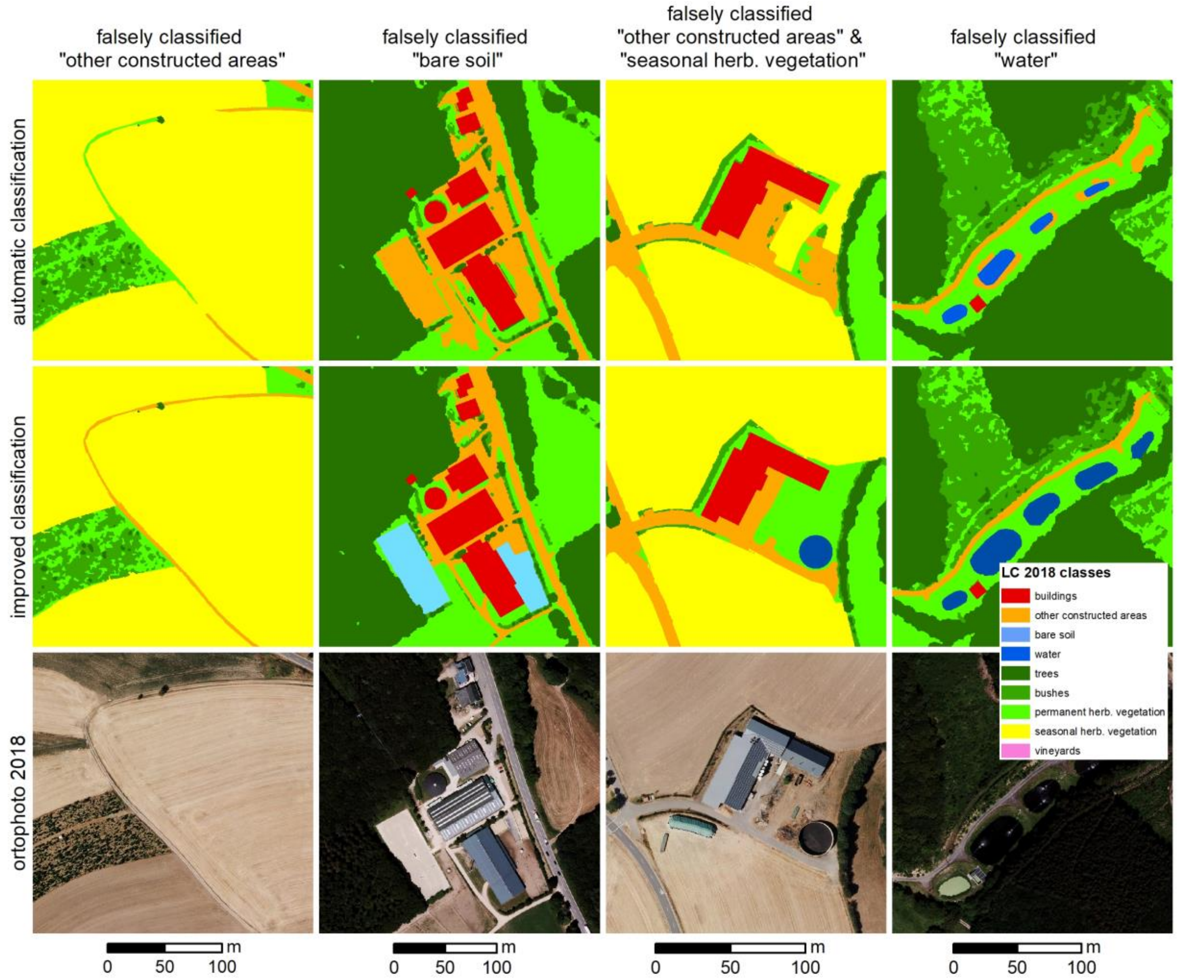


Figure 5-5: Examples of objects updated during visual control of the results of the automatic classification in eCognition.

In general, most of the objects were delineated correctly. After performing the automatic classification and verifying the results visually, we proceeded to improve the results. We did this by manual editing of major polygons that represent a wrong LC class and those that were partly wrongly classified. The class we improved the most was "other constructed area" where many objects located close to buildings were falsely classified into this class. This happened due to weather conditions in 2018; the summer was dry, therefore many areas covered with grasslands were assigned to this class because the NDVI was very low. Areas covered by bare soil were similarly classified as "other constructed areas". Also, some non-asphalted rural roads, shown on Figure 5-5 were not correctly assigned to this class because their width was very small, making it difficult to correctly classify them using an automatic algorithm. In most cases the artificial areas have similar properties as agricultural fields that are not covered by any kind of vegetation; however, applying the SENTINEL time series images allowed us for correct distinguishing between these two classes. The classes, which we improved the least, were "buildings", "trees" and "bushes". We improved the buildings layer before importing it to eCognition, therefore the quality of this class was already high and there was no need in correcting it to a large degree. Trees and bushes we classified mainly with the LiDAR data, what allows us achieving very high accuracy and precision in delineating the boundaries of those classes. The only issue on which we drew attention was new clear cuts, because the LiDAR data were from 2017, and we performed the classification for 2018, what causes the possibility of cutting down trees in this period.

5.5 QUALITY ASSESSMENT

We verified the quality of the LC 2018 map by sampling the random points across the country (Fig 5-6). We selected 200 points per each class (1800 points); additionally to estimate omission errors, a set of 200 random sampling points were collected based on a group of settlement classes from the LIS-L land cover map of 2015, which was used as a generic reference domain. Stratifying map by reference domains, allows to more frequently selected artificial objects on the ground and assess their omission in the final map. In total we performed the quality check based on 2000 randomly selected points across the entire Grand Duchy of Luxembourg.

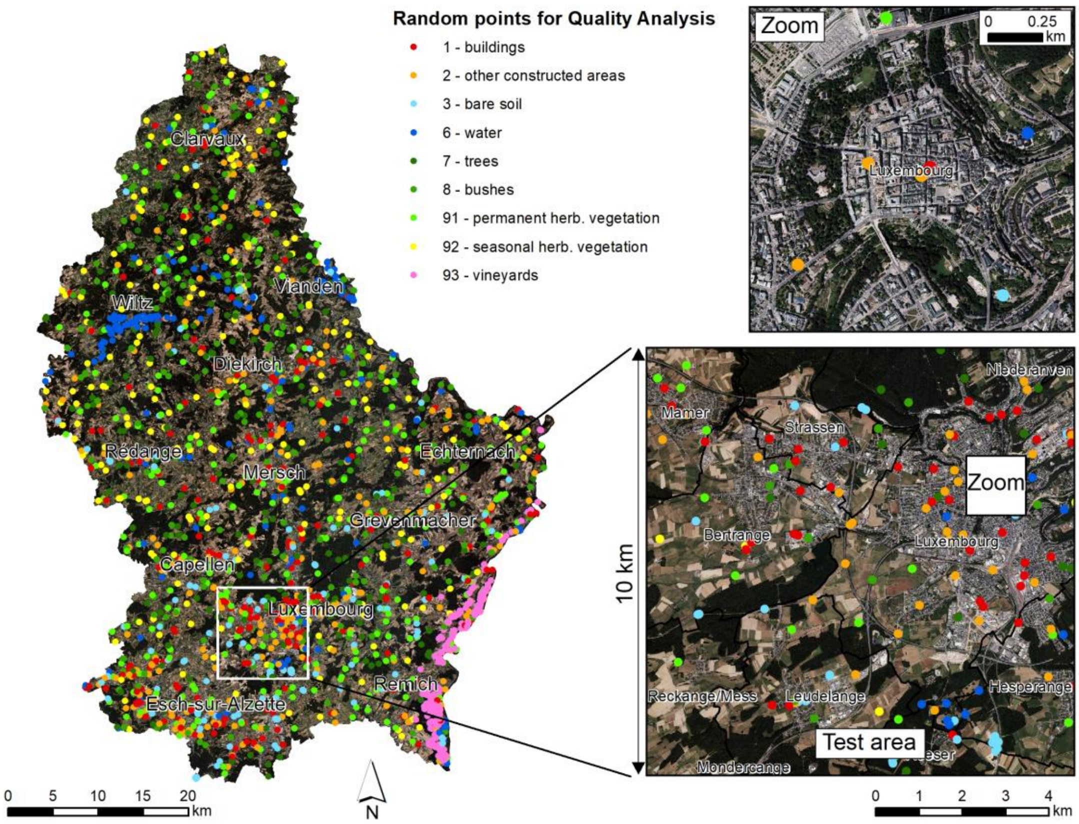


Figure 5-6: Random points across the entire Grand Duchy of Luxembourg for quality check of automatic classification.

The quality check presented on the matrix in Table 5-1, shows, that the commission errors with a standard deviation inferior to 5% at the 95% confidence level do not exceed 8.6% for each individual class, what meets the project requirements. The overall accuracy of the classification for the entire area of Grand Duchy of Luxembourg is equal to 96.4%. More than a half of the classes have the User's and Producers' Accuracies exceeding 95% with the Commission and Omission Errors lower than 5%. In overall, the best metrics of the classification we achieved for seasonal herbaceous vegetation and the class of vineyards.

We also noticed, that some of the randomly selected points that were used for the quality check were located close to the border of different objects (Figure 5-7), thus for example wrongly classified building does not always mean, that the building has not been detected. It also might mean that the building has been correctly detected, but its boundaries are not precise.

Table 5-1: Accuracy metrics for automatic land cover classification with respect to 2000 randomly selected reference points; Overall Accuracy [%], Producer's Accuracy [%]; Users Accuracy [%], Errors of Commission [%], and Errors of Omission [%].

		Reference data									Total	Errors of Commission	User's Accuracy
		1	**2**	**3**	**6**	**7**	**8**	**91**	**92**	**93**			
Classified data	**1**	231	3			1		1			236	**2.1**	**97.9**
	2	8	235	1			2	3			249	**5.8**	**94.4**
	3		1	192		1		5	1		200	**3.3**	**96.0**
	6		1		193	5		1			200	**2.9**	**96.5**
	7					223		4			227	**1.6**	**98.2**
	8	1	4			3	194	12	1		215	**8.6**	**90.2**
	91	3	7		1	1	2	258			272	**5.8**	**94.9**
	92								201		201	**0.0**	**100.0**
	93									200	200	**0.0**	**100.0**
Total		243	251	193	194	234	198	284	203	200	2000		
Errors of Omission		**4.9**	**6.6**	**0.4**	**0.4**	**4.5**	**1.6**	**10.7**	**0.8**	**0.0**		**Overall Accuracy 96.4 %**	
Producer's Accuracy		**95.1**	**93.6**	**99.5**	**99.5**	**95.3**	**98.0**	**90.8**	**99.0**	**100.0**			

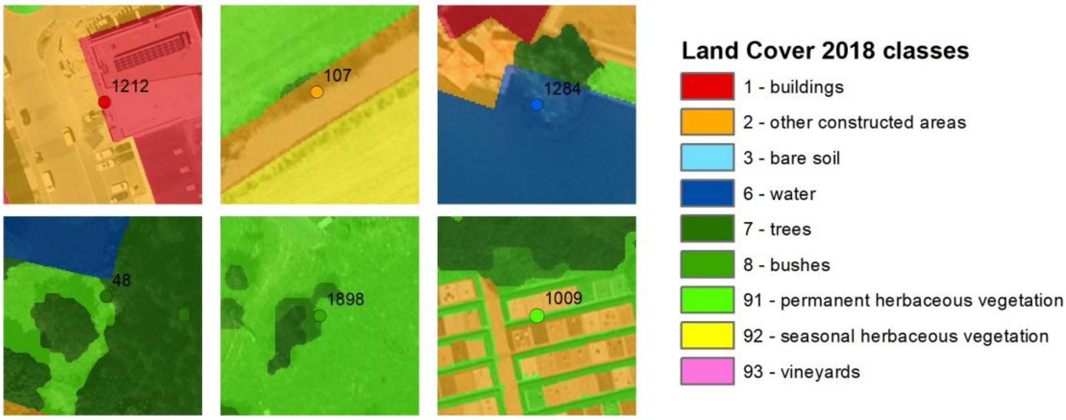


Figure 5-7: Examples of random points with their ID number that are located close to the border between two classes and were incorrectly classified during automatic classification.

6 Results and Discussion

As a result of this project we deliver the Land Cover classification for 2018 of the entire Grand Duchy of Luxembourg with 10 m buffer in vector and raster formats as two products, which are:

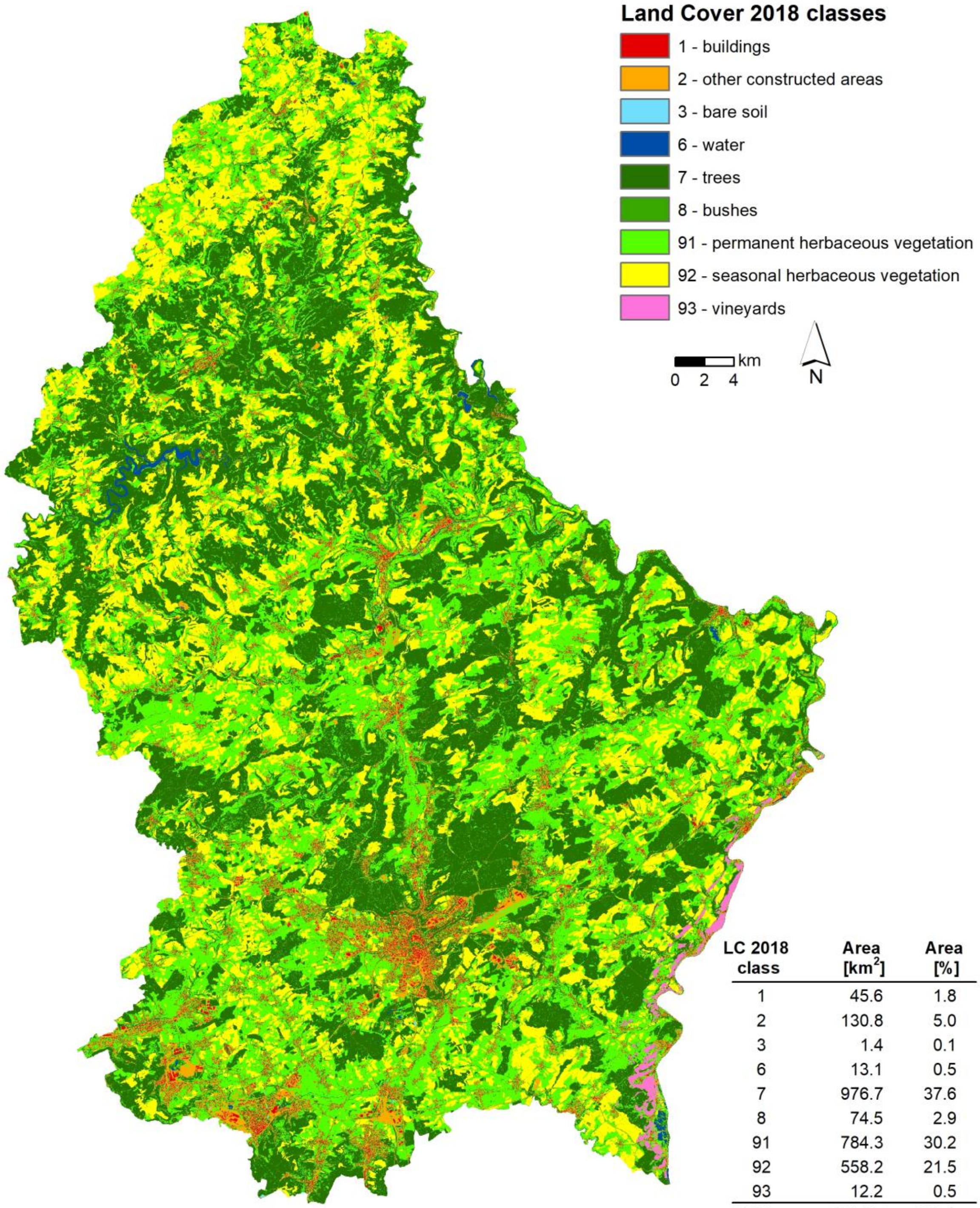
1) Personal Geodatabase: **LC_2018.gdb** containing polygon vectors with the numerical LC class attribute named “LC_class_code” that corresponds to the gridcode of raster dataset saved as “Value”. Additionally, we include the description of classes as text “LC_class_name”.
2) Raster dataset: **LC_2018_20cm.tif** with 0.2 m spatial resolution and the LC classification saved in “Value”.

The final LC map for 2018 shown on Figure 6-1 represent nine land cover classes, from which four are abiotic, and five are biotic classes. The biggest rate of the country constitutes forest covering 976.7 km² of the analysed area (37.6%). Whereas the smallest rate constitutes “bare soil” with only 1.4 km² what constitute 0.05% of the country.

The accuracy assessment shows a good quality of the map. The match of classes with objects represented on the ortophotomap is high.

The results also show that some plots in FLIK dataset represent a false class. There is a lack of new vineyards, there is a mismatch between seasonal and permanent herbaceous vegetation. In addition, some of the currently built-up plots are not removed from the FLIK data. Therefore, for future work, we recommend a short visual verification of the dataset and their update before performing the automatic classification. This shortens the time of final manual refinement, which is more difficult and time-consuming than the correction of input data.

The use of LiDAR data allows obtaining very precise information about individual trees, opening in the forest and vegetation height. The application of nDSM was helpful in separating trees from bushes and young trees, which gives valuable information about vegetation throughout the country and allows analysis and statistics on the structure of vegetation.



LC 2018 class	Area [km^2]	Area [%]
1	45.6	1.8
2	130.8	5.0
3	1.4	0.1
6	13.1	0.5
7	976.7	37.6
8	74.5	2.9
91	784.3	30.2
92	558.2	21.5
93	12.2	0.5
SUM	**2596.7**	**100.0**

Figure 6-1: Land Cover 2018 for the Grand Duchy of Luxembourg with the statistics of the coverage of each LC classes.

A visual comparison of the LC 2018 classification with LC 2015 (Figure 6-2) shows that the greatest improvement occurs in the “trees” class. The boundaries of this class represent more realistically and describe in detail the group of trees and the occurrence of individual trees. In addition, it was possible to separate “trees” from “bushes”. It is also noticeable that the “other constructed areas” class is more detailed. We can observe that the asphalt driveway for buildings are more precisely outlined compared to LC 2015, and that in built-up areas in general there are more facilities in this class – while in 2015 these were most often classified as “permanent herbaceous vegetation”.

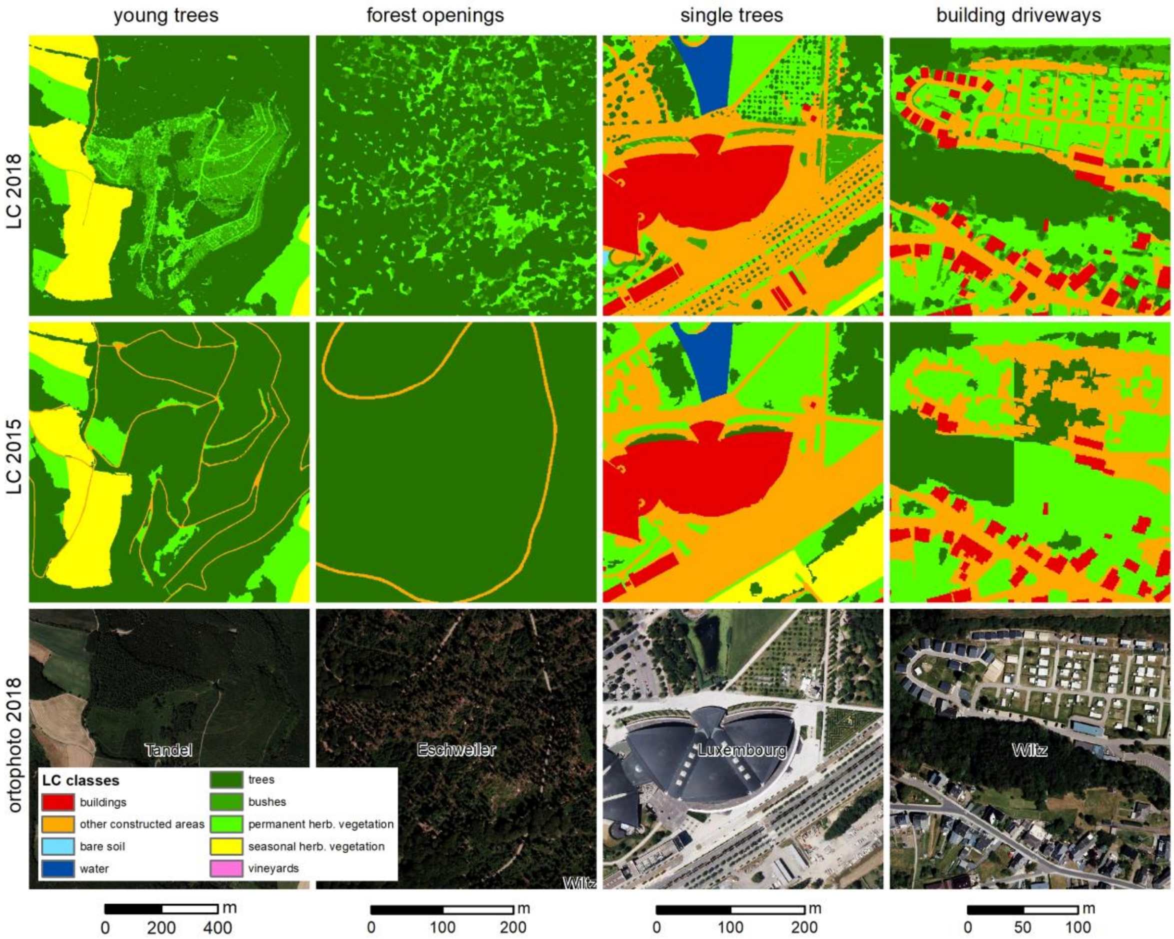


Figure 6-2: Visual comparison of land cover classification in 2018 (0.2 m spatial resolution) and 2015 (1m spatial resolution) for the Grand Duchy of Luxembourg and an overview of the ortophotomap for 2018.

6.1 REMARKS WHILE COMPARING THE HRLC’18 WITH LIS-L LC FOR 2015:

The LC data for 2018 contain two additional classes, which are: “bushes” and “vineyards”, which may cause difficulties in performing the statistics. In addition, the “bare soil” class contains more types of facilities in 2018 comparing to 2015. We therefore recommend:

- not to compare statistics of the class of bare soil, because in 2015 it contains only sand on the golf courses, while in 2018 it additionally includes facilities such as: sand in playgrounds, beach volleyball courts, equestrian facility, areas temporary not covered with vegetation which are close to a construction site (e.g. temporary parking lot) and indoor areas that are not constructed and whose function is uncertain (bare areas between the highway lines that are designed to be grassland);
- when comparing the permanent herbaceous vegetation class, the total area of “vineyards”

classified in 2018 should be subtracted from the permanent herbaceous vegetation in 2015;

- when comparing the 2015 forest, the classes “trees” and “bushes” for 2018 should be combined and treated as one class;
- when comparing the other constructed areas, the total area of non-asphalted agricultural roads (taken from LIS-L LU 2015 dataset) should be subtracted from the total area of other constructed areas for 2015, because in 2018 we did not use these features as ancillary data to carry out the HRLC’18 classification.

7 CONCLUSIONS

In this report, we present an automatic approach for land cover mapping that we performed for the entire area of the Grand Duchy of Luxembourg. This approach is based on combining information from NIR aerial images, LiDAR and time series of SENTINEL images. Our approach uses a segmentation approach to reduce the salt-and-pepper noise of the classification, geometrical and textural information of objects obtained from images, as well as the neighbourhood relation between objects to classify specific land cover classes. The results presented in Figure 5-5 show high reliability in matching information on land cover from aerial images with the support of LiDAR and SENTINEL. The accuracy of the automatic classification does not change depending on the different tiles and parts of the country, which indicates its potential to be used in future LC updates for the entire country of Luxembourg. We have noticed that the input FLIK data in some cases are not up to date and that the LiDAR data contain some errors in the southern part of the country, what may slightly influence the LC classification.

ANNEX I

Detailed description of HRLC'18 land cover classes

Land cover class code	Land cover class name
1	Buildings
2	Other constructed areas
3	Bare soil
6	Water
7	Trees
8	Bushes
91	Permanent herbaceous vegetation
92	Seasonal herbaceous vegetation
93	Vineyards

ABIOTIC (NON-VEGETATED)

The primary separation of land cover elements is based on the vegetation cover. Abiotic includes all landscape elements that are primarily not covered by vegetation. It is further differentiated whether the surface is built-up (i.e. covered by artificial material) or non-built-up and water.

Built-up surfaces

1 Buildings (Class code = 1)

Permanent construction with a roof, providing storage, shelter or residence to people, animals or things. A building is characterised by its extension into the third dimension (contrary to "other constructed area which is "flat"). A minimum height is not defined.

Parking lots with more than one floor, but without a roof are included in this class.

Buildings under construction – where the roof has not been finished – shall be classified as "other constructed areas".

2 Other constructed areas (Class code = 2)

Artificial surfaces, partially or fully covered by impervious material. These artificial surfaces are usually maintained over long periods of time. Surface material includes asphalt / tarmac, concrete, gravel or stones.

This class includes roads, sidewalks, paving's, parking lots and other artificially constructed surfaces like tennis courts, running courses and swimming pools.

Artificial objects having extension into the three dimension, but without a roof or that may be only temporal (e.g. tents, greenhouses) are classified into this class.

Surfaces associated to construction sites, commerce or industry (storage areas, access roads, and mining areas) and permanent agricultural roads are included in this class.

Bare areas

3 Bare soil (Class code = 3)

Naturally open soil permanently not covered by vegetation where the maximum vegetation share is 10%. If more than 10% of bare soil is seasonally covered by natural vegetation, then the area should be classified as "herbaceous vegetation".

This class includes also artificial surfaces that are not covered by vegetation or impervious material such as: sand on golf courses, sand in playgrounds, beach volleyball courts, equestrian facility, areas temporary not covered by vegetation that are located close to a construction site (e.g. temporary parking lot) and areas inside the construction sites that are not constructed and which the future state is not sure (bare areas between road lines on the highway, which are designed to be grassland).

Water

6 Surface water (Class code = 6)

Open surface water of running or standing water, which excludes flooded or other temporary waters. For reservoirs, the actual water level in the input EO data shall be mapped. Also artificial objects e.g. in the wastewater treatment plants, which are filled with water, are included in this class (this does not apply to the swimming pool).

BIOTIC (VEGETATED)

Biotic or vegetated surfaces represent the second main aspect of land cover classes. These surfaces are covered with biomass at the moment of the image acquisition. The amount of biomass can be detected with the help of vegetation indices (e.g. NDVI).

Woody vegetation

Perennial plants with stem(s) and branches from which buds and shoots develop. Distinction between Trees and Bushes is made based on the height information obtained from LiDAR data.

7 Trees (Class code = 7)

A tree is defined as a woody perennial plant with a single, well-defined stem carrying a more-or-less-defined crown, with a height generally greater than 3m at maturity. Exceptions are young growth areas.

The class "trees" encompasses single trees, groups of trees and continuous forest stands. They normally are characterised by a height difference from their neighbouring surfaces (i.e. shadow).

Orchards, plantations and Christmas tree farms are also included under this class.

8 Bushes (Class code = 8)

Bushes are woody perennial plants with persistent woody stems and without any defined main stem in the range of 0.5 – 3m height.

Young trees with the height not exceeding 3m also belong to this class.

Herbaceous vegetation

Plants without persistent stem or shoots above ground and lacking definite firm structure.

9 Herbaceous vegetation (green cover)

Any with low vegetation covered surface (grassy or herbaceous) with an above-ground coverage of more than 10%.

9.1 Seasonal herbaceous vegetation (Class code = 91)

Agricultural areas with seasonal herbaceous vegetation cover are classified into this class.

9.2 Permanent herbaceous vegetation (Class code = 92)

All arable surfaces with permanent herbaceous vegetation (including intensive and extensive grasslands) fall into this class. This class contains also all vegetation-covered, but non-woody surfaces, including natural grasslands, pastures, parks, lawns in residential garden or green areas associated to traffic.

9.3 Vineyards (Class code = 93)

Agricultural areas with permanent vine vegetation cover are classified into this class.