Land Cover Classification of Urban Areas: A Comparison of Object-Based and Pixel-Based Approaches

Master Thesis Cand.Geom.

Aalborg University Copenhagen

23-04-2018

Casper Frederik Stub Trock

AALBORG UNIVERSITY
STUDENT REPORT

Aalborg University – Copenhagen
A.C. Meyers Vænge 15
2450 Copenhagen SW
www.plan.aau.dk/

Title:

Land Cover Classification of Urban Areas: A Comparison of Object-Based and Pixel-Based Approaches

Student:

Casper Trock

Semester:

4th Semester

Project period:

Sep. 2017 – Apr. 2018

Supervisor:

Jamal Jokar Arsanjani

Page count:

68

Abstract:

This thesis contains a study of different classification approaches. Primarily focusing on how segmenting high-resolution satellite imagery into objects before running the classifications, can improve the accuracy. The study area is Copenhagen and the satellite images are from 2015 and 2017.

The classification of two Sentinel 2 images is done with random forest and support vector machine methods. First, the classification is done by only looking at pixels and secondly another classification is done, this time after segmenting the images.

Different segmentation values are explored, in order to discover what variables obtain the best classification.

The produced land cover maps are then used to model the changes in the time period in TerrSet.

LIDAR data is introduced to a smaller area in order to examine if a digital surface model can improve the accuracy scores as well.

The object based classification approach showed significantly improvements on the total accuracy score. Especially the classes of built-up area and grassland improved in accuracy. The LIDAR data showed improvements in some sparse areas.

The change detection revealed a general shift toward more grassland in the study area in the period between 2015 and 2017. While all other classes declined in net area.

Abstract (Danish)

Dette studie undersøger forskellige klassifikationsmetoder. Der fokuseres primært på, hvordan segmentering af high-resolution satellitbilleder i objekter før klassificeringen, kan forbedre nøjagtigheden. Studieområdet er København og satellitbillederne er fra 2015 og 2017.

Klassificeringen af to Sentinel 2 billeder er lavet med ”random forest” og ”support vector machine”. Først foretages klassificeringen ved kun at se på pixels, og derefter foretages en anden klassificering efter segmentering af billederne.

Forskellige segmenteringsværdier undersøges for at opnå den bedste klassificering. De klassificerede kort bruges herefter til at vise udviklingen i tidsperioden.

LIDAR data introduceres i et mindre område for at undersøge om en digital overflademodel forbedrer klassificeringens nøjagtighed.

Den objekt baserede metode viste signifikante forbedringer på nøjagtigheden, specielt for klasserne “bebygget område” og “græs”. Data fra LIDAR viste forbedringer i enkelte områder.

Der har i tidsperioden fra 2015 til 2017 været et generelt skifte til mere græs i det undersøgte område, mens alle andre klasser er reduceret arealmæssigt.

Table of Contents

List of figures

List of Tables

1. Introduction

The first multispectral satellite image became available for civilian remote sensing in the early 1970s. Ever since considerable efforts have been devoted to classifying of satellite images with the aim to detect specific objects, producing high quality thematic maps or establishing accurate land cover maps. Classification is a significant part of remote sensing and most of the published papers on the area of remote sensing, usually concern classification in one way or another (Wilkinson, 2005).

A definition of land cover can be found in the INSPIRE Directive, where it is defined as a physical and biological cover of the Earth's surface including classes such as built-up areas, agricultural areas, forests, wetlands and water bodies (INSPIRE, 2013).
The maps can be used to improve government of the land, both in order to give a better understanding of the current landscape and with change detecting; a way of evaluating past decisions and to estimate the effect of future initiatives. Other important uses of land cover maps and changes in landscape, are to help collecting information on the environment, monitoring natural resources and predicting natural disasters. (Nasa, 2017)

This study will look at multispectral satellite images from the sentinel satellites in order to produce land cover maps. In June 2015, the European Space Agency launched the Sentinel-2A satellite, followed by the Sentinel-2B in March 2017. The satellites are capturing images on a repeat cycle in 786 km height, ten days with one satellite and five days with two satellites. The goal is to provide information and contribute to land observations, emergency response and security services, forestry and agricultural practices as well as to assist in food management (Eurepean Space Agency, 2018).

The multispectral images will then be classified, with help from machine learning, instead of looking only at pixel values that utilizes the concept of multidimensional feature space, like most standard existent GIS programs can do. The techniques have developed since the first low-resolution satellite images were made public and the pixel approach still works with the modern high-resolution images. Nevertheless, there has been an increasing demand to improve both the processing speed and accuracy of images, both from the growing number of potential users with different needs and from the new sensors that significantly increase the spectral variability within each class (Blaschke, 2013). This variability can decrease the potential accuracy of a pixel classification approach (Hay, et al., 1996).

This study will therefor compare the results of both pixel and object based approaches. The object based image classification has had an increasing focus over the past decades and studies have investigated several different sensors, feature selectors, classifiers etc. A comparison study shows that best performing algorithm for object based classification is Random Forest (RF) and that there is strong correlation between accuracy and training set size, unlike the accuracy of the pixel based approaches where increased training points can make the class signature clouded (Lei & Manchun, 2017).

The pixel classification will be compared to the object based classification approach with the expectation that the segmentation of the images, done by looking for clusters and similarities between neighboring cells in one or more dimensions, will lead to better performances.

The above approaches will be used on images from both 2015 and 2017 to include most of the time period where the sentinel 2 images are available. This will allow the study to include a comparison study as well, in order to detect changes. This comparison will focus on changes in built-up area and will be done by comparing the spectral value of the high-resolution images, as well as with produced land cover maps in order to illustrate the development in Copenhagen in the last couple of years.

Lastly, implementing laser-scanning data through a digital elevation model will be tested on a smaller area. This will add another data source, potentially leading to improved classification and if so, be revealed as something to explore further.

1.1 Problem statement

To what degree can an object based approach improve the land cover classification of an urban area compared to pixel-based approaches?

Additional research questions:

- How will additional data from a digital elevation model affect the accuracy of the land cover classification?
- The land cover maps and the images will be used to detect changes: How has the development in the study area been e.g., patterns and magnitude? And how can we predict the future patterns by using the historical observations?

1.2 Structure of study

An overview of the study is shown in Figure 1

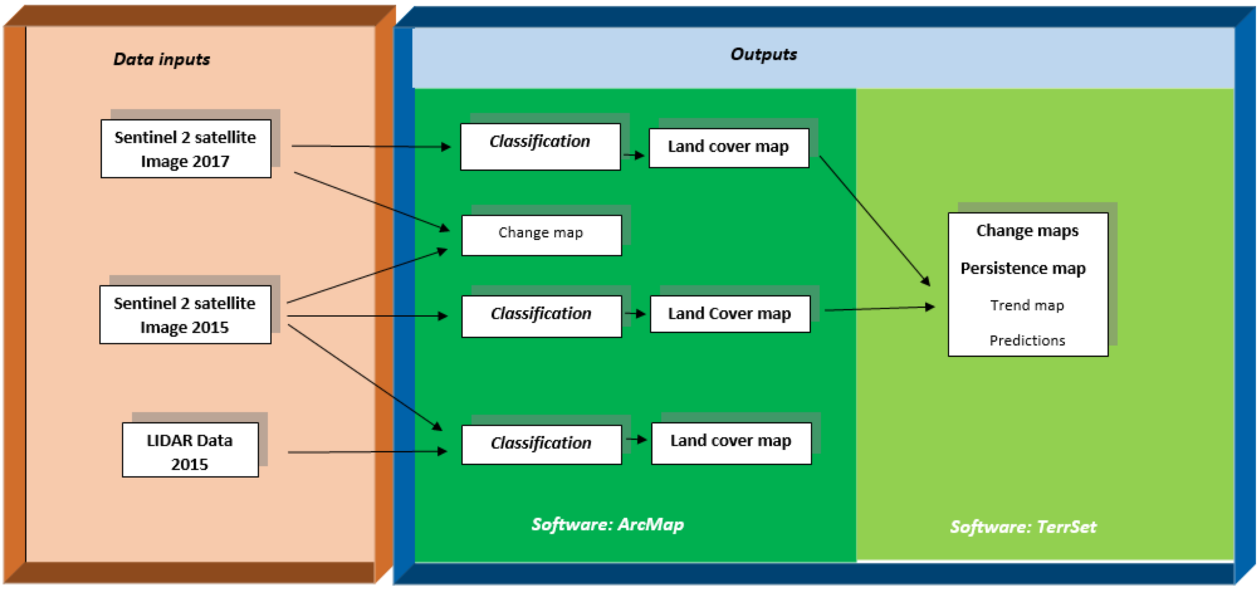


Figure 1 Study overview - bold text in output box indicates a focus area

1.3 Theory

1.3.1 Classification Methods

As described in the introduction, this study will focus on "Random Forest" (RF) and "Support Vector Machine" (SVM), because other studies indicate that they are most likely to produce the best results (Rosca, et al., 2017).

Random forest is well established in numerous disciplines as a good statistical classifier (Cutler, et al., 2007) and can be considered among the best options for classification of remote sensing and other forms of geographic data from multiple sources, as it combines accuracy and efficiency (Gislason, et al., 2006).

Support vector machine is another successful machine learning algorithm for classification of satellite imagery. Support vector machine are found to be more competitive with other classifying algorithms, both with regards to the accuracy of the classification and the processing time (Melgani & Bruzzone, 2004) (Mountrakis, et al., 2011).

1.3.1.1 Random Forest

This supervised learning technique got its name because of the forest structure it forms with the generation of multiple decision trees, that are made with random sampling with replacement included (Breiman, 2001) (Hastie, et al., 2009). The decision tree is a well-established method; it is a rule-based technique that builds binary trees (Kotsiantis, 2007). It splits the data domain in nodes into two and orders them in subdomains to increase the information. Finding the best split is the goal for an optimization algorithm in order to gain the maximum information and therefore ease the classification. The benefit of random forest is that it supplies several trained decision tree classifiers that can be used in the testing phase, compared with the single decision tree (Suthaharan, 2016). See Figure 2.

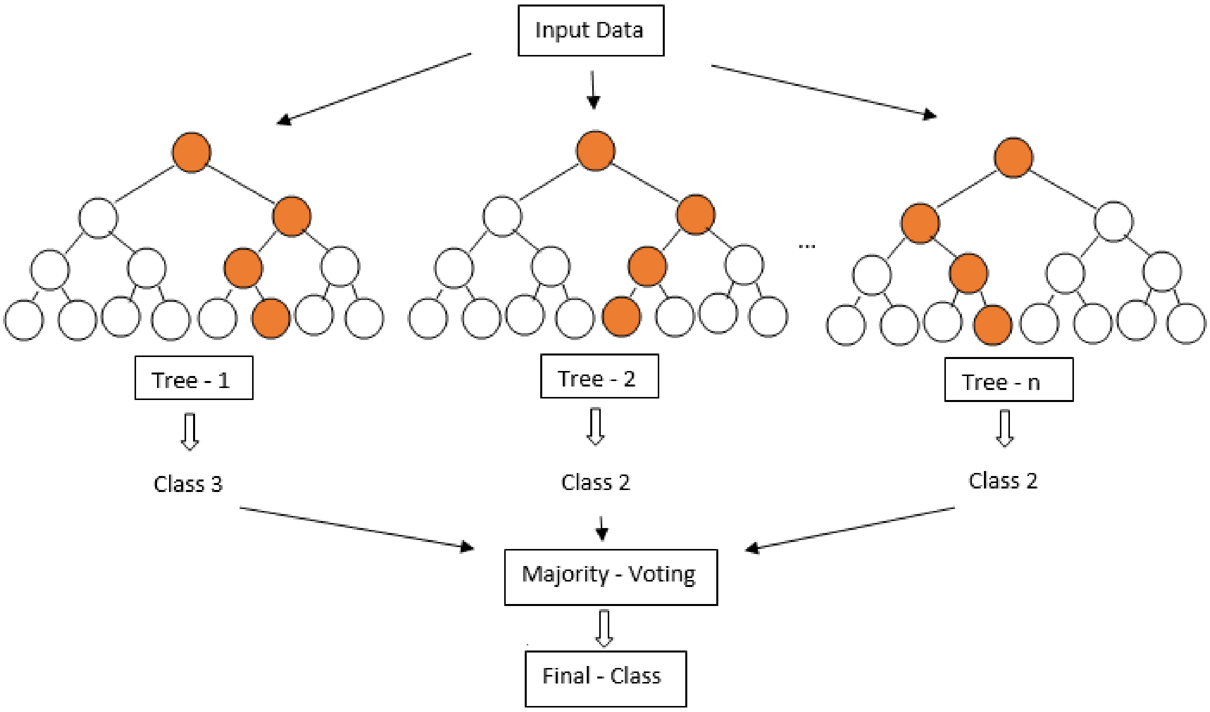


Figure 2 Illustration of random forest (Liaw & Wiener, 2002)

There are two statistical measures that optimize the classification objectives for the random forest; bootstrapping and bagging. Bootstrapping is a randomization technique that has a big impact on supervised learning algorithms like random forest. It generates multiple subsets from the data set, by picking the same number of observations, but with replacement. This allows the subset to have the same number of total observations, but with some observations picked several times (Suthaharan, 2016).

Bagging is another way of saying bootstrap aggregating (Breiman, 1996). It is done by averaging the predictions of the bootstraps models of the final classification result (Suthaharan, 2016). The trees are not dependent on the other trees, as they are constructed on their own by a bootstrap sample from the original data set (Liaw & Wiener, 2002).

Random forest only has two parameters, the number of trees it generates and the number of variables each subset node has. It is often not sensitive to the value of the parameters (Liaw & Wiener, 2002).

1.3.1.2 Support Vector Machine

Support vector machine provides both a classification learning model and a mathematical algorithm. With a model of y = wx' + ϒ that also manipulates it, in order to allow linear domain division. Support vector machines are either linear or nonlinear models. If linear the data domain can be separated linearly to divide the classes. If nonlinear the data needs to be transformed into a feature space, where the data then can be divided linearly. This is done by creating a new space using a polynomial kernel (Suthaharan, 2016).

The concept is to calculate the optimal hyperplane that separates the data into classes that fits with the values of the training samples. This hyperplane delineates the decision boundary and aim to reduce the margin width between the classes as shown in Figure 3 (Vapnik, 1982).

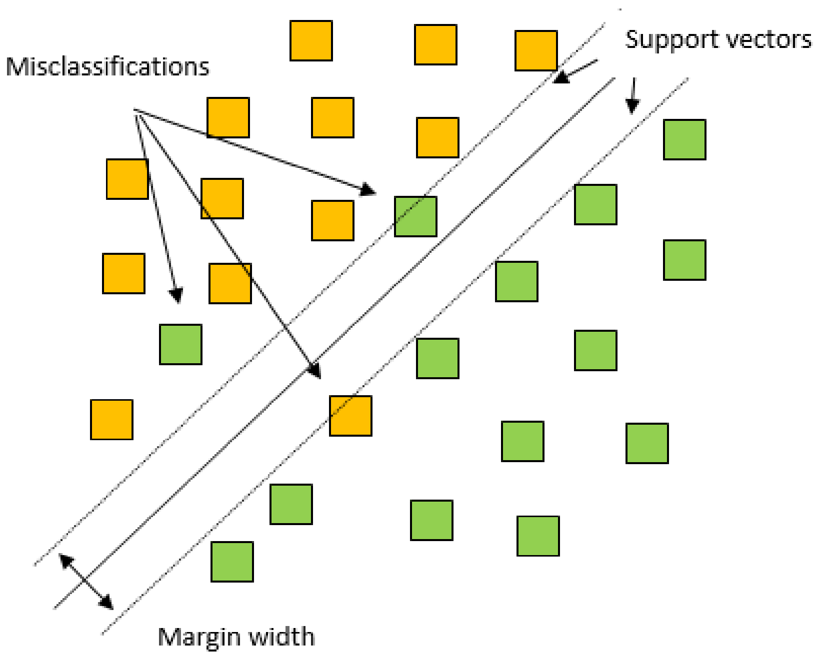


Figure 3 Illustration of support vector machine (Burges, 1998)

When the support vector machine algorithm is used for classifying satellite images, each pixel for every band is presented as a vector. This often allows support vector machine classification to achieve high accuracy results even with few training samples (Mountrakis, et al., 2011).

When the classifier is trained it is an iterative process, with refining of the decision boundary to find the best fit of the hyperplanes separation of the training samples. The better the training samples are the better the accuracy will be. Measurement errors or distortion from the topography and atmosphere can cause data impurities. Misclassified samples can have a huge impact on the classification process and the runtime of the classifier (Mountrakis, et al., 2011).

1.3.2 Classifying pixels or objects

The question of when to classify pixels versus using an object based approach is not new and has been discussed since some researchers pushed for techniques that moved beyond pixels.

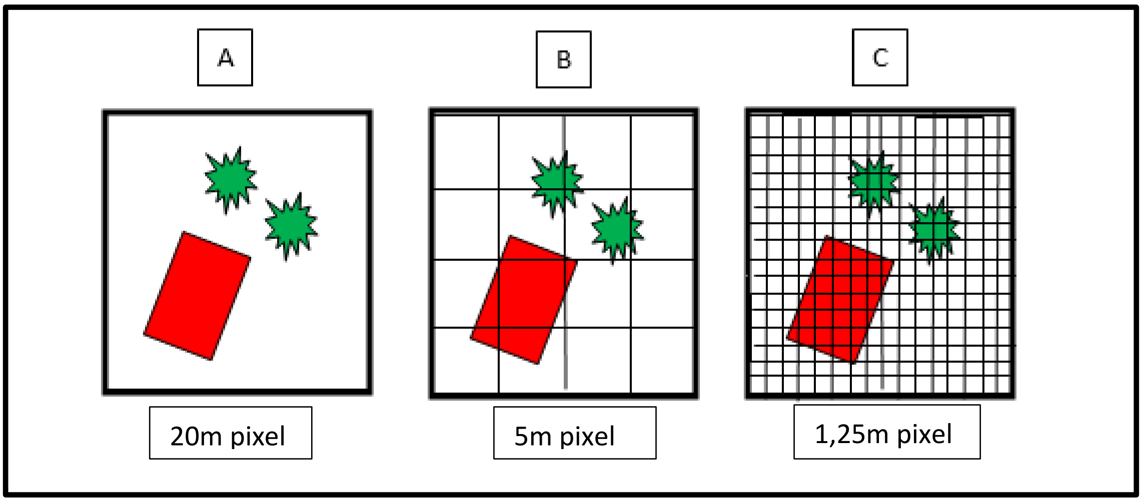


Figure 4 Relationship between objects and the spatial resolution (Blaschke, 2009)

Figure 4 above illustrates the relationship between the object you are trying to identify and the spatial resolution of the images. According to the “Shannon sampling theorem”, the object should be at least the size of 10 pixels in order to become completely independent of its orientation and random position in regard to the classifier. The three examples a, b and c require different approaches; a – needs a pure pixel approach, b – also looks like pixel but some studies utilize objects, c – is where the object approach can give several advantages (Blaschke, 2009).

.

“The results show that object-oriented classification techniques enhance quantitative analysis of traditional pixel-based change detection applied to very high-resolution satellite data and facilitate the interpretation of changes in urban features'' (al Khudairy, et al., 2005).

This or similar quotes can be found in many studies. The advantages of object based image classification are well established, for example can the segmentation of images into objects help when identifying

forest gaps, vegetation patches or the complexity of landscapes (Blaschke, 2009). But some studies have also identified certain problem areas in high-resolution images. For example, when classifying vegetation, as it is by nature distinctly non-uniform as a result of shadow and shade, each pixel is not closely related to the complete physical appearance of the vegetation (Ehlers, et al., 2003).
Developments in the image classification techniques, starting from around year 2000, can document a sharp increase in the usage of image segmentation and an increasing use of object based image analysis (Blaschke, 2009).

Another factor is the classification accuracy related to the size of the training samples. Pixel based approaches has only showed a small improvement in accuracy when the training points increase, while the object based method show a clear positive correlation between the number of training points and the accuracy (Lei & Manchun, 2017).

1.3.2.1 Segmentation

Image segmentation has been done since the late 1970s and became more established in the 1980s, but was mostly used in the computer science community for either pattern analysis or quality control of products. Only few people used image segmentation in order to classify earth observation data (Blaschke, et al., 2004).

Basically, segmentation is dividing image pixels into groups of pixels that are alike in terms of what values the sensor has registered. It is unlikely that there is a one to one relationship between the real-world objects and the image objects that models them, which is the main issue when segmenting. Image objects are a collection of measurements at a time and place, from a sensor that is arrayed systematically.
An issue with segmentation can be that a group object is not meaningful for future use of the results together. "For example, through segmentation, two adjacent forest stands could end up in a single (image-) object even though they are managed differently or are owned by different proprietors." (Blaschke, 2009)

There are several ways of segmenting an image. The algorithms can roughly be classified into belonging to one of these five groups:

- Thresholding

- Template matching
- Clustering
- Edge detection
- Region growing

Each group has been proven successful under some circumstances, but none of them can suit every need and they all contain some drawback. Thresholding is the simplest as it is based on histograms, but it fails to separate pixels with the same grey value from different objects. Template matching becomes very inefficient on larger images. The clustering method has a danger of segmenting the images to much and likewise has the edge detection that can also struggle with images containing many edges (Xiang-Yang, et al., 2010).

No matter which of the methods is applied, segmentation provides the building blocks of object based image analysis (Hay & Castilla, 2008). Segments have more spectral information compared to any single pixel, for example:

- Mean values per band
- Median
- Minimum values
- Maximum values
- Mean ratios
- Variance

Another advantage of segments is the expansion of spectral value descriptions of objects with the additional spatial information. It has been frequently claimed that this spatial dimension (distances, neighborhood, topologies, etc.) is crucial to "Object Based Image Analysis" (OBIA) methods (Blaschke, 2009).

1.3.3 Training points

In order to successfully classify a satellite image a sufficient number of training samples are needed. Training samples are often collected from high-resolution aerial photographs, fieldwork or other data sources (Tempfli, et al., 2009). Different strategies can be used to define the training points, such as pixel, seed or polygons (Chen & Stow, 2002). If the study area has a complex landscape or the image resolution results in many mixed pixels, as they contain several classes, selecting enough can be difficult (LU & Weng, 2007).

1.3.4 Accuracy Assessment of Classification

All possible sources of errors should be identified, when validating a classification (Powell, et al., 2004). In addition to errors directly from the classification, interpretation errors and errors from the training/test data also affect the accuracy of the classification. It is often assumed that the differences between the result of the image classification and the ground truth data are caused by classification errors. However, in truly assessing the accuracy of a classification, non-image errors should also be considered (LU & Weng, 2007).

The error matrix (Foody, 2002) to determine the classification accuracy is the most commonly used approach. When creating an error matrix several factors should be considered: reference data collection, classification scheme, sampling scheme, spatial autocorrelation and sample size (Congalton & Plourde, 2002).

When the error matrix is completed other important values can be derived (Foody, 2002):

- Overall accuracy: Basic measure for how many elements that are classified correctly.
- Omission errors (Producer accuracy): Show the accuracy of each class by comparing the total number of a given class correctly classified as that class and the number of points of the given class in the ground truth data.
- Commission errors (User accuracy): Show the proportion between correctly classified points from a given class and the total number of points that was classified as that class.
- Kappa coefficient: Gives a value for the overall statistical agreement of the error matrix, containing the non-diagonal elements. It basically gives a value for the accuracy of the method compared to random guessing (Anon., 2017).

The Kappa coefficient is considered a strong method of evaluating a single error matrix as well as comparing results of multiple error matrices (Foody, 2004). The error matrix approach for accuracy assessment is only suitable when the map classes are mutually exclusive and when each pixel belongs to a single category (LU & Weng, 2007).

1.3.5 Change detection

Change detection analyses can show how the land cover has evolved in a time period, in quantifiable differences and help identify the changes in a given class. This knowledge can be used to analyze the percentage of built-up area in a region, for uses in for example city planning. Change detection is often defined as "… the process of identifying differences in the state of an object or phenomenon by observing it at different times" (Singh, 1989)

There are two general methods of creating change maps, both using differencing functions (Kennedy, et al., 2009). One method involves computing the spectral difference between the two images directly, while the other method will classify the images separately first and then compare them resulting in a map of the changes in the time period. The first method results in the simplest form of change detection maps where pixels that have changed are mapped in binary discrimination, while the second method can be used for more complex procedures that can generate maps of land cover change from/to classes. Another use is for analysis of increase or decrease of the classes in a given time period (Hecheltjen, et al., 2014).

1.3.5.1 Change detection with satellite images

There have been many change detecting techniques since the launch of the first satellites. In change detection there is a basic premise that a change in land cover also result in a change in the radiance value, and that the change must be large (Singh, 1989). No approach is optimal, but should be determined based on image quality, data sources, the characteristics of the area and the nature of the changes you aim to identify. The latest research is trying to develop new techniques and adapt existing methods to better utilize the high-resolution images that are now available (Shrestha, et al., 2017).

The pixel based method compares all pixels in one image to the corresponding pixels in another image of the same area from a different time. The change map generated will show the changes in reflection in the spectral bands and the assumption is that the significantly change in values are from the change in objects on the earth and not from interference from other factors, such as different atmospheric conditions or illumination. For examples can rooftop values be affected by shadowing effects, as can vegetation that also change values according to season (Shrestha, et al., 2017).

Clouds are a big challenge in this method of change detection using optical imagery. The best case is cloud free images, but to achieve this, analysts can be forced to use images out of season. These images then do not have the same acquisition conditions, the angle of the sun can be different and the phenology of the vegetation is most likely different (Hecheltjen, et al., 2014). For some parts of the world it is almost impossible to create cloud free time series, which can result in each pixel almost having its own time series, because different pixels will be removed by a mask in different images to remove the clouds (Zhu & Woodcock, 2012).

1.3.5.2 Change detection with Land cover maps

Change maps from comparison of classified land cover maps have been used for a long time. One of the most widely used is presumably post-classification comparison (Jensen, et al., 1987) that is based on the classification of two images of the same area from different times. The primary benefit of this method is that it is independent of the data source used as input and there is no need for radiometric preprocessing or adjustment of the images. A weakness of this method can be that none of the classifications have 100 percent accuracy and the errors carry over in the change analysis. In general, this leads to inclusion errors where change is falsely detected (Hecheltjen, et al., 2014).

There has in the last decade been developed a hybrid, known as change-restricted post-classification comparison. The process here is to do change detection on multi imagery using the method of comparing reflectance. A conservative threshold is then applied to identify changed and unchanged areas and lastly the images are classified in order to be able to reclassify and describe the class changes (Xian, et al., 2009) (Hecheltjen, et al., 2014).

2.Materials and methods

2.1. Study area

The area for this study mainly consists of the Danish capital Copenhagen and its suburbs. There is in the study area a large portion of built-up area that varies from continuous urban fabric to discontinuous low density, forests, grassland, cropland, lakes and ocean shores.

Figure 5 Image of the study area in 2015: Copenhagen and its suburbs.

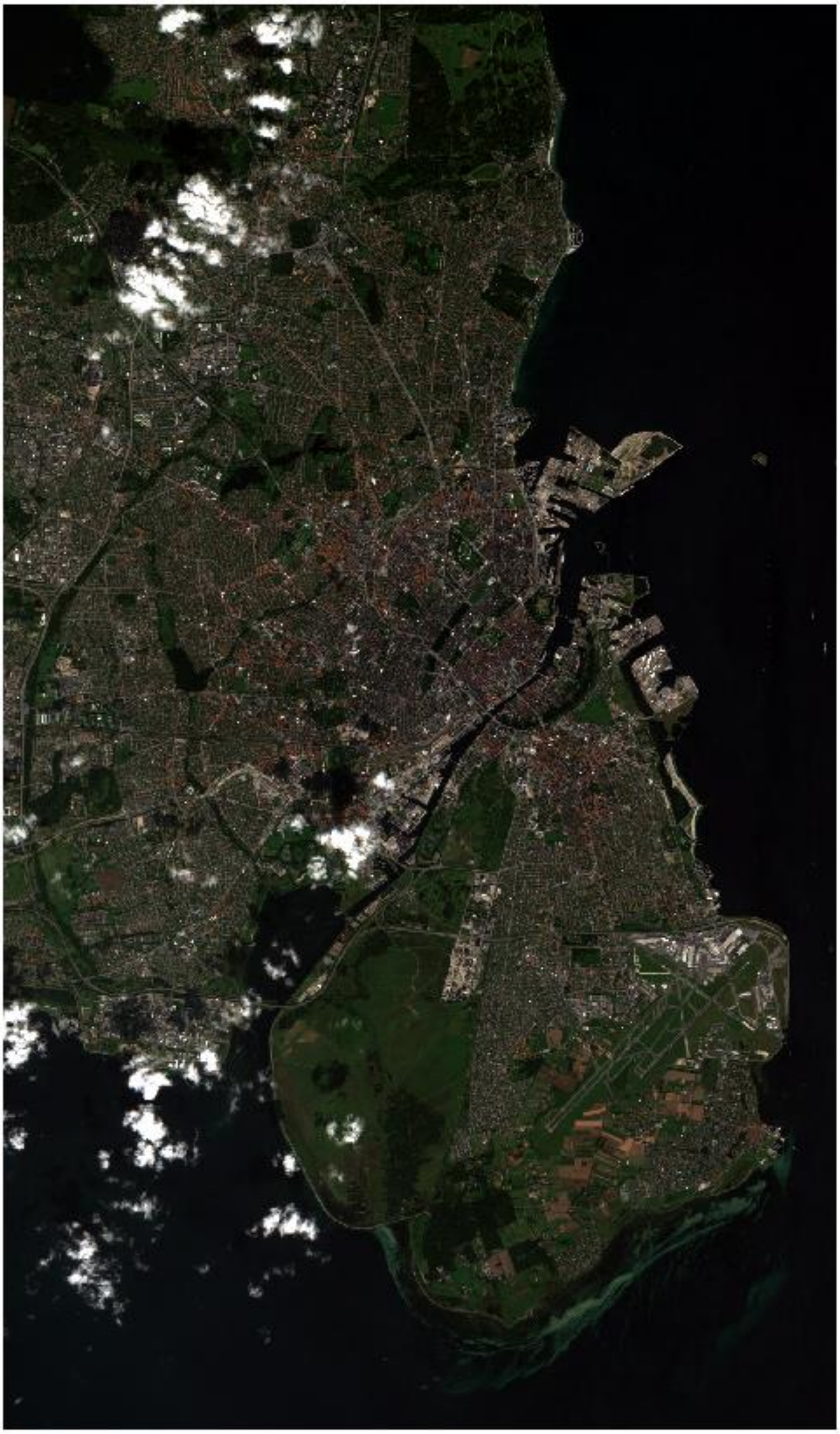


Figure 6 The 2017 Sentinel 2 satellite image from 2017-08-23

2.2 Data

Satellite image:

Sentinel 2 imagery was chosen due to the images being freely available and having a high temporal and spatial resolution. The selected Sentinel 2 tile 33UUB was captured on 2015-08-19. The 2015 image, Figure 5, has 0% cloud cover and 100% of the study area has been correctly captured by the satellite. The second image from 2017-08-23, Figure 6, was chosen so that it was same season and the longest possible time period with sentinel 2. The image does have some clouds that also cover part of the study area.

Vector data:

The majority of input datasets have been acquired from "Kort10". Kort10 is a nationwide topographic object-oriented map in vector format and scale 1:10.000. The data is supplied and maintained by the Danish Agency for Data Supply and Efficiency. (Agency for Data Supply and Efficiency, 2017)

The rest of the vector data is from The Danish Agrifish Agency (The Danish Agrifish Agency, 2017). This agency had agricultural vector data available based on what landowners who apply for agricultural subsidies report to be using the land for.

2.3 Methods

2.3.1 Classifying method

2.3.1.1 Pre-processing:

Pre-processing operations have been implemented by using the Semi-Automatic Classification QGIS plugin (SCP). This open source plugin for QGIS provides a range of tools and uses for downloading, pre-processing, post-processing, and classifying various satellite images.

The images were downloaded with the tool to QGIS, where the bands were converted to surface reflectance. The surface reflectance conversion tool attempts to decrease the effects of the atmosphere in the image, and thus providing a true surface reflectance at ground level (Congedo, 2016).

A basic DOS1 atmospheric correction was also applied to the image. Dark Object Subtraction (DOS) is a family of image-based atmospheric corrections (Chavez, 1996) (Congedo, 2016). Chavez' explanation is that some objects are in complete darkness due to shadows or cloud cover, and the reflectance received at the satellite is due to atmospheric scattering.

The input image was then created, with a ‘band set’. The band set includes all bands recommended for classification; bands 1, 9 and 10 are left out (60m bands). The center wavelength of each band was also calculated before the image was clipped to only contain data for the study area and then saved as a TIF file.

2.3.1.2 Ground Truth map

A land cover map of the study area was generated using the gathered vector data, this was done in ArcMap. The map can then be used as a ground truth map (GTM), which can both be used to selecting random training points and later in the process to validate the results of the different classifiers .

The GTM was divided into a large number of classes where some only covered below 1% of the study area. In order to improve the results of the classifiers, the classes were merged into five classes that best illustrated the land cover in the region.

2.3.1.3 Selection of classes

The five classes going forward will be:

- Class 1: Built-up area – This class will contain all artificial surfaces that represent the urban area. Most of the area is covered by different kinds of buildings, from small huts to large modern buildings. Besides buildings, roads and railroads contribute to the class.
- Class 2: Forest – Besides classified forest, parks and patches of trees
- Class 3: Grassland – All areas with grass, low vegetation, sports areas
- Class 4: Sand – Besides the sand, this class includes construction sites and landfill areas
- Class 5: Water - Ocean, lakes and other water bodies

Before the five classes where merged together to a single layer with GTM, a grid was created with the same dimensions as the satellite image and where the grid cells aligned precisely with the image pixels. This grid was used to intersect with the five vector layers and then only cells with at least 90% of their area in a given class were selected to go into the final GTM.

In the classification from 2017 there was created a 6th class, as there are clouds in this image. This class 6 is treated like the other classes.

See Figure 7 for pixels examples of the five classes.

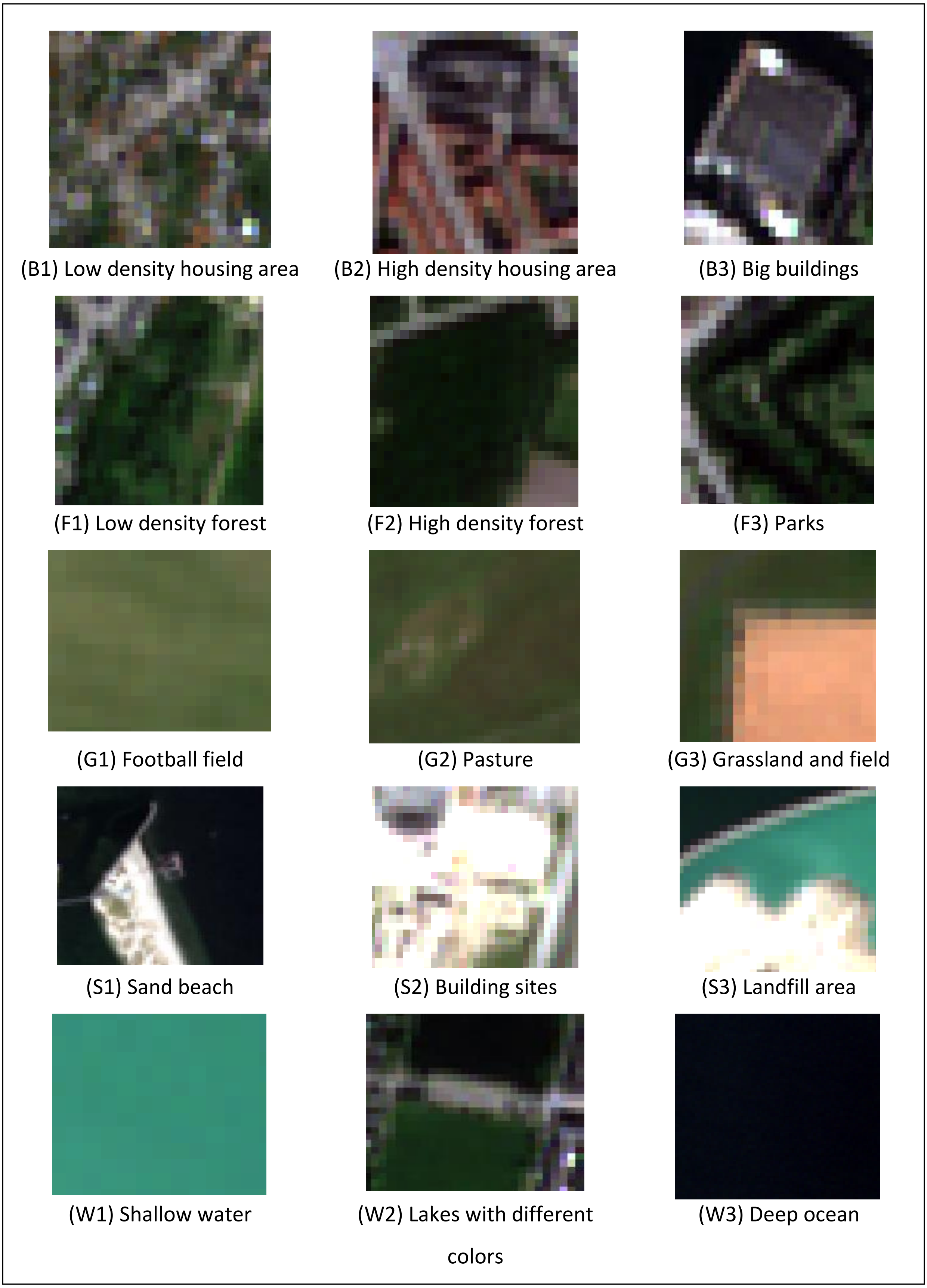


Figure 7 Examples of pixels in the classes: Built-up (B), forest (F), grassland (G), sand (S) and water (W)

2.3.1.4 Sample Point Selection

In order to train the chosen algorithms and to test their accuracy, a number of randomly selected training and testing pixels had to be selected. Comparative studies have employed a minimum number of 10-30 training points per band for each class (Foody, 2002), but as more resent research has shown a positive correlation between classification accuracy and the size of the training sample (Lei & Manchun, 2017), more random training pixels will be used to train the algorithms. For class 1, 2, 3 and 5 - 1000 points were selected and for class 4 – 500 points were selected, because the sand class only covers a small area compared to the other classes. Additional 250 distinct pixels were used to test the algorithm accuracy (125 for class 4).

The GTM created as described in the previous section was used to generate the random training points. Random cells were selected for training and testing purposes. This sampling technique selected 1250 random cells from each class. By using an equalized stratified sampling technique, it was ensured that each class is represented by an unbiased, spatially and spectrally diverse selection of training pixels.
As a double verification process, the randomly selected cells for testing were then visually inspected for accuracy using high-resolution aerial imagery. The high-resolution imagery consists of orthophotographs with 12.5 cm resolution, captured by Geodanmark (Agency for Data Supply and Efficiency, 2015). The acquisition date of the high-resolution imagery was spring 2015. This date does not coincide with the capture date of the Sentinel-2 imagery, but was the best option available.

The training sample file was created of 1000 training points for each class, the points were then loaded into the "Training Sample Manager" in ArcMap, one class at a time and merged. When all classes were in the manager, the sample was saved and could be used for training the classifiers together with the image.

ArcMap has a tool to create random accuracy points, which can be used to contain both a class from the generated classified map and a class from the GTM. The test points were therefore formatted to have the same attributes as the accuracy points, in order to use them with the ArcMap tools. Lastly a confusion matrix was calculated with the results from the testing points.

2.3.2 Segmentation method

Segmentation is needed to switch from pixel classifying into identifying objects first. The 2015 image was segmented with two methods; edge detection was done in QGIS with help from the SCP plugin and the Orfeo toolbox and pixel approach was done in ArcMap, both with standard values in the first test. The results were close, but as ArcMap was faster and the computer was barely able to handle the task in QGIS, the rest of testing was done in ArcMap, in order to find the best possible combination of variables. The three variables for this segmentation are spectral and spatial values, that both can range between 1 and 20, the higher the value the higher the threshold. The last value is an option to force a minimum segment size of objects.

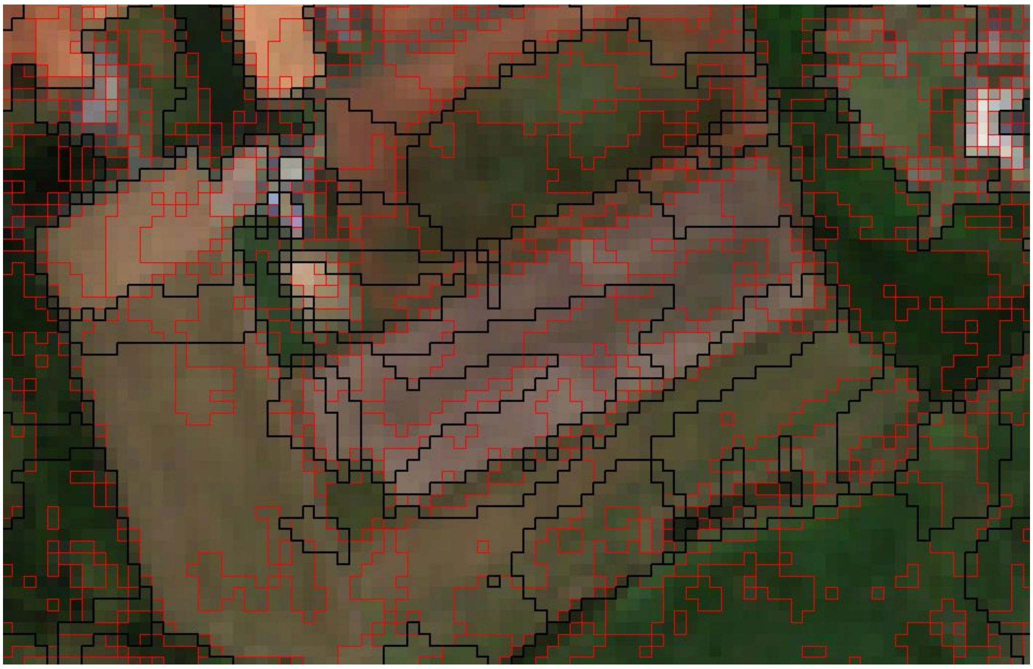


Figure 8 Segmented image showing the influence of parameter settings

Figure 8 shows two of the segmented images, the black segments are an example of a reduced spectral value in order for the pixels to be grouped together (factor 15). While the red segments are done with maximum value, this creates a lot more objects and leaves as seen several pixels by themselves. Both of these segmentations were done with the option of pixels left alone.

2.3.3 Classifying with digital surface model method

A smaller area was selected to test the addition of LIDAR data to the classification. The reduction in the study area was primarily done to be able to handle the size of the data.

The LIDAR data was acquired from the Danish Agency for Data Supply and Efficiency, through their digital surface model (DSM). The Data was gathered in 2015 and the 2015 satellite image was therefore used. The DSM comes in raster with a cell size of 0.4 m. The first step was therefor to create new raster with matching cell sizes to the image. (Figure 10)

The classifiers were tested on the new study area, with and without the addition of the created DSM. ArcMap cannot run the classification with both a segmented image and a DSM as additional inputs to the training points, so classification was done with the pixel approach. It was possible to segment the DSM though and use that as an input.

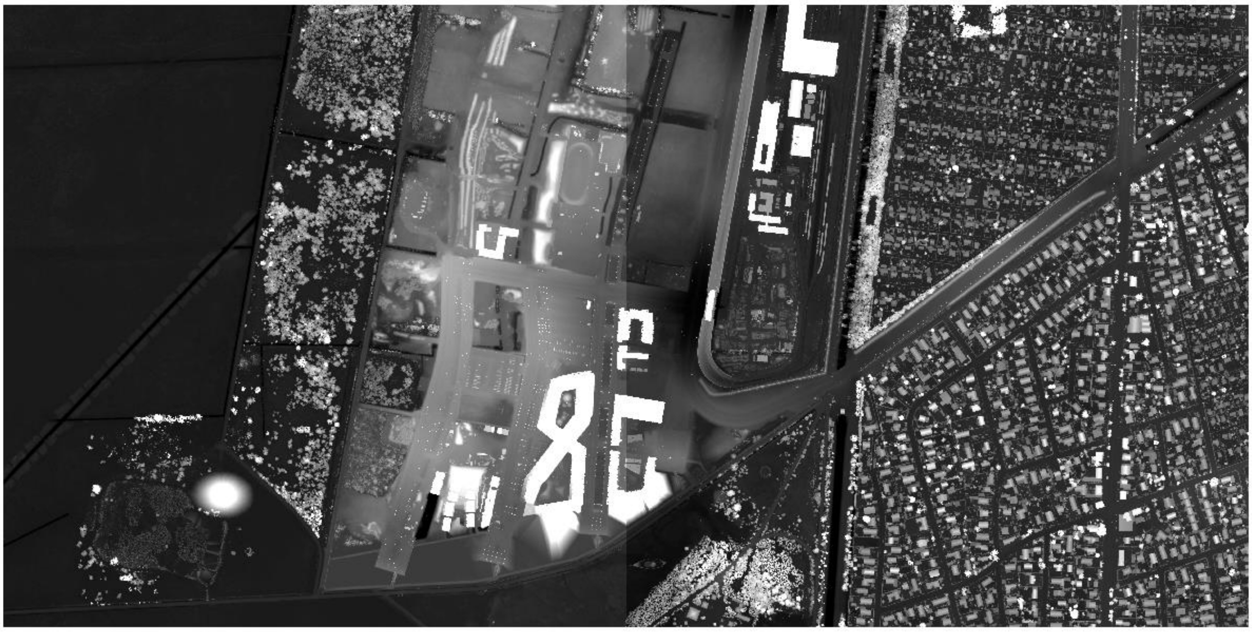
2.3.3.1 Smaller study area

An area in the middle of Amager was selected, see Figure 9. The area contains all the classes from the original area. Furthermore, it has different kinds of build-up area including tall buildings and lower density areas with houses and roads in different sizes. There is also vegetation in varied heights.

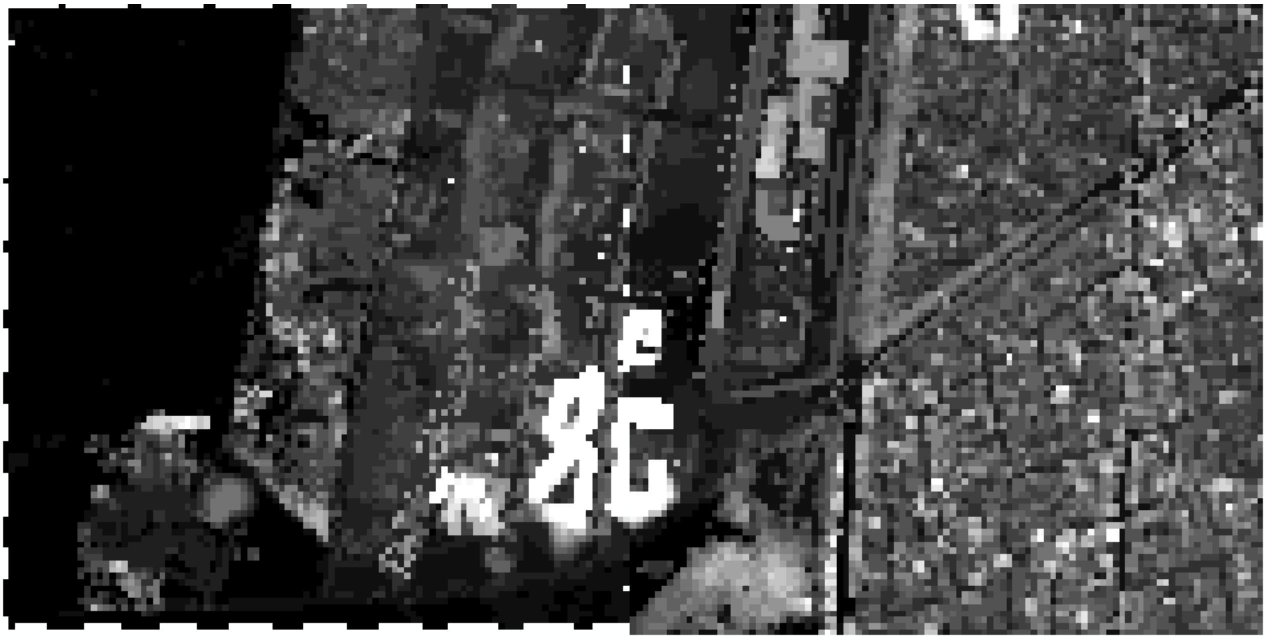
Figure 9 Overview and close up image showing the smaller study area

2.3.3.2 Selection of classes

The classifiers will use the same classes as described for the original study area, with the exception of roads that will be a new class instead of being contained in the built-up area.



(A) Digital surface model from the Danish Agency of Data Supply and Efficiency



(B) “Resampled” to pixels of same size as satellite image

Figure 10 Digital surface model in different resolutions

2.3.4 Change detection method

Change detection will here be done in two ways. First; the two best classified land cover maps produced, will be used in the program TerrSet (Idrisi), to analyze where changes appear as well as how the classes change. Secondly; simply by comparing all pixels in the 2015 image to the corresponding pixels in the 2017 image. The change map generated will show the changes in reflection in the spectral bands (Shrestha, et al., 2017). This method will be done in ArcMap that has a tool for this analysis.

As there are clouds in one of the images they stand out, but aside from this, other areas with major change can be identified by their high color value. The two areas shown in Figure 11 are picked out in order to have a closer look at the changes. There is no doubt that the weakness of this method is shown as well: the change maps show differences around all the areas where there is built-up and several other regions could be picked for examination. For example, the east coast of Amager.

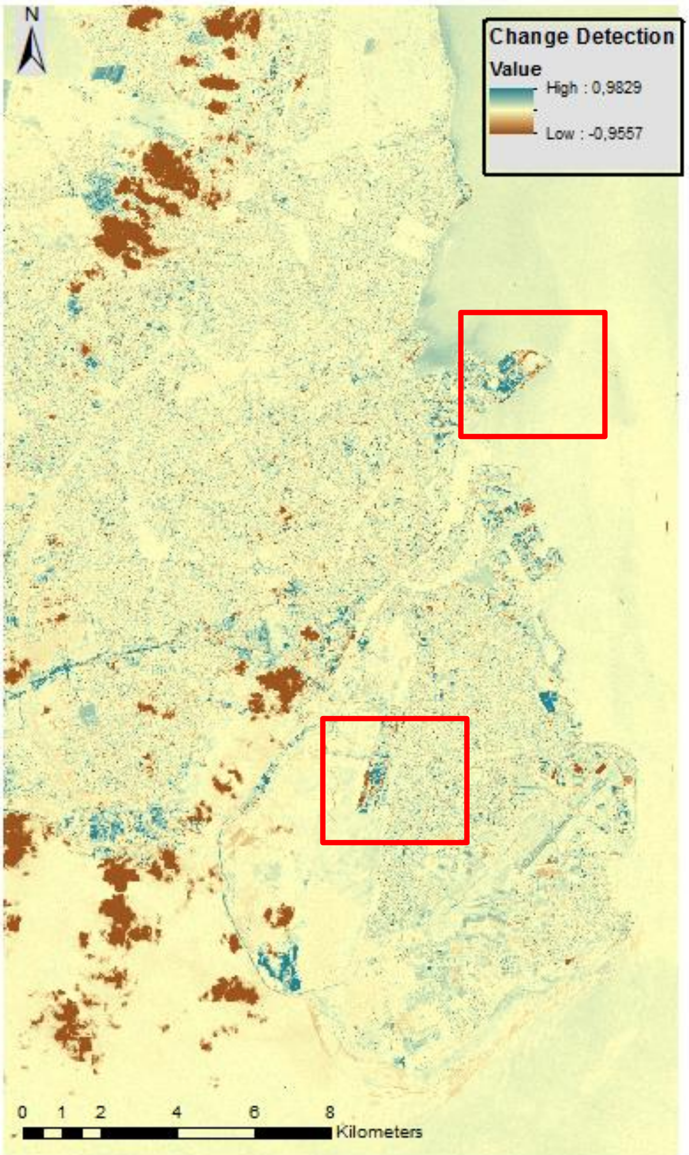


Figure 11 Change detection map between the two images

3. Results and analysis

3.1 Results of pixel based classification

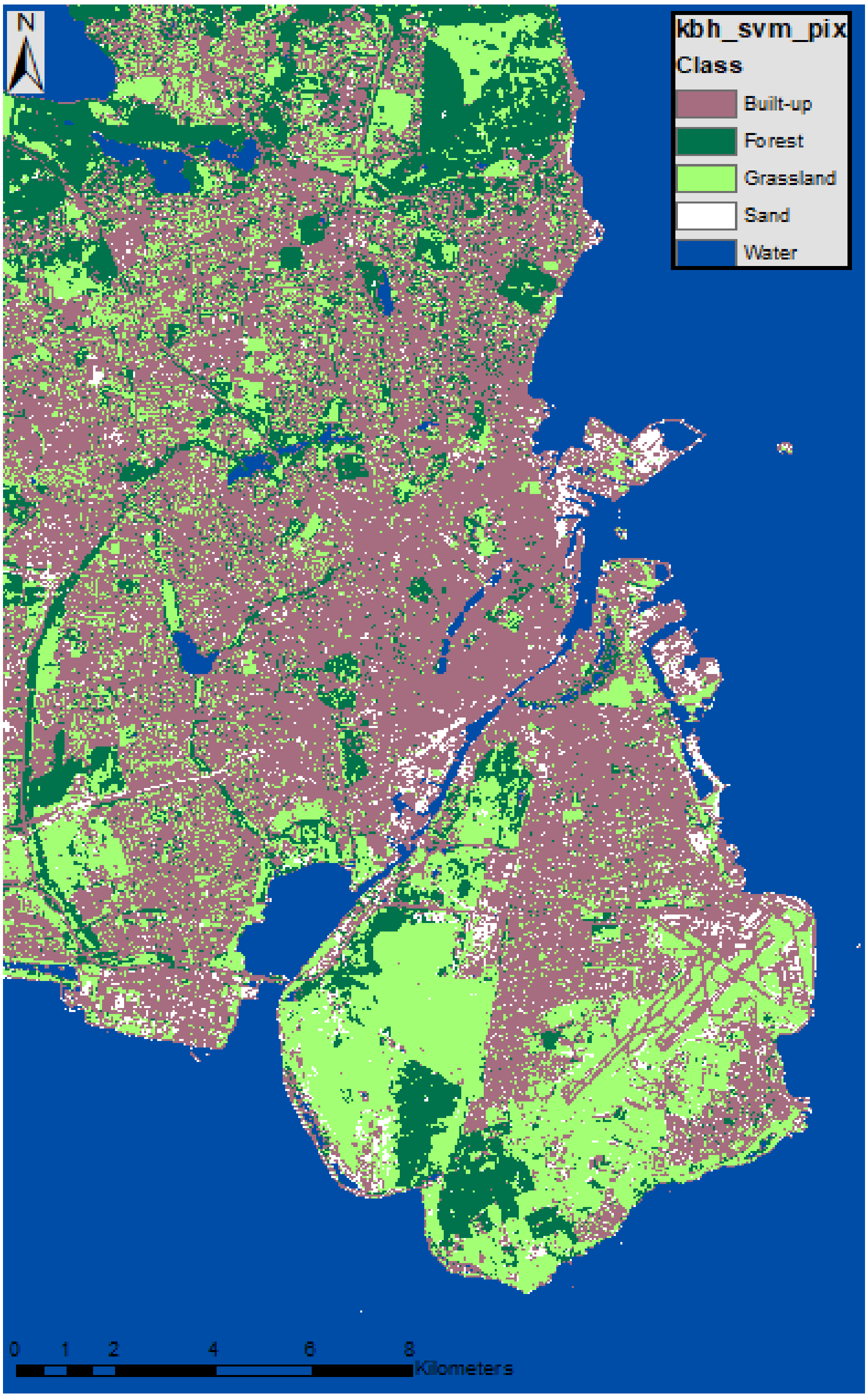


Figure 12 Land cover map made with support vector machine classification of pixels in year 2015

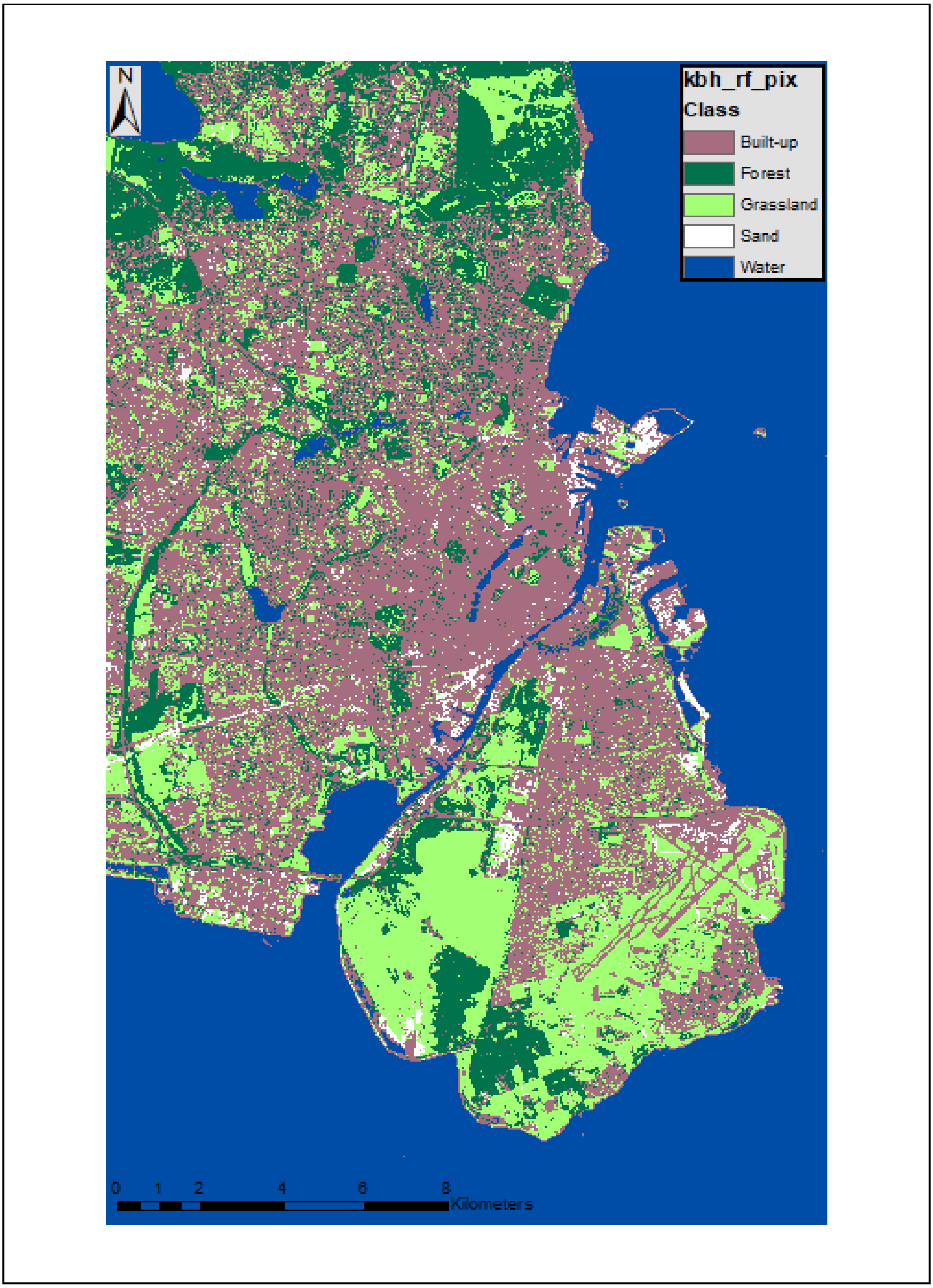


Figure 13 Land cover map made with random forest classification of pixels in year 2015

Table 1 Accuracy for the support vector machine classification of pixels in year 2015 (Figure 12)

Class	Built-up	Forest	Grassland	Sand	Water	Total	User accuracy
Built-up	173	5	39	17	9	243	71,2 %
Forest	27	234	31	0	7	299	78,3 %
Grassland	42	10	175	7	5	239	73,2 %
Sand	8	1	5	101	0	115	87,8 %
Water	0	0	0	0	229	229	100,0 %
Total	250	250	250	125	250	1125	
Producer accuracy	69,2 %	93,6 %	70 %	80,8 %	91,6 %		81,1 %
Kappa							76,0 %

The SVM accuracy of 81,1 % shown in Table 1 is the better of two classifications only looking at pixel value. Looking at the producers' accuracy the class with highest accuracy is forest, followed by water that both score above 90 %. The wrongly classified water pixels are in small lakes close to either buildings or tall threes. The worst classes for the SVM are buildings and grassland. The majority of the errors happen in areas where houses and private gardens are hard to separate from each other. This could be a carryover from the training pixels that contain both grassland and built-up in the same class, as the vector data used do not distinguish in residential areas.

The sand class errors are mostly showing where there are construction sites. The beaches are correctly identified. The user accuracy tells more or less the same story, but it is worth noting that no land pixel has been wrongly classified as water.

Table 2 Accuracy for the random forest classification of pixels in year 2015 (Figure 13)

Class	Built-up	Forest	Grassland	Sand	Water	Total	User accuracy
Built-up	181	11	39	15	9	250	71,0 %
Forest	32	222	35	0	8	314	74,7 %
Grassland	35	16	172	7	3	221	73,8 %
Sand	2	1	4	103	2	112	92,0 %
Water	0	0	0	0	228	228	100,0 %
Total	250	250	250	125	250	1125	
Producer accuracy	72,4 %	88,8 %	68,8 %	82,4 %	91,2 %		80,5 %
Kappa							75,3 %

The results for RF shown in Table 2 are close to performing just as good as the SVM. It has the same best, middle and worst classes, but doing slightly better when identifying built-up and sand and slightly worse with the vegetation.

3.2 Segmentation results

Figure 14 Segmented image showing areas of interest

Figure 14 is an example of the image segmented into groups estimated to belong to the same object; Figure 15 is a closer look at some of the edges between the segments.

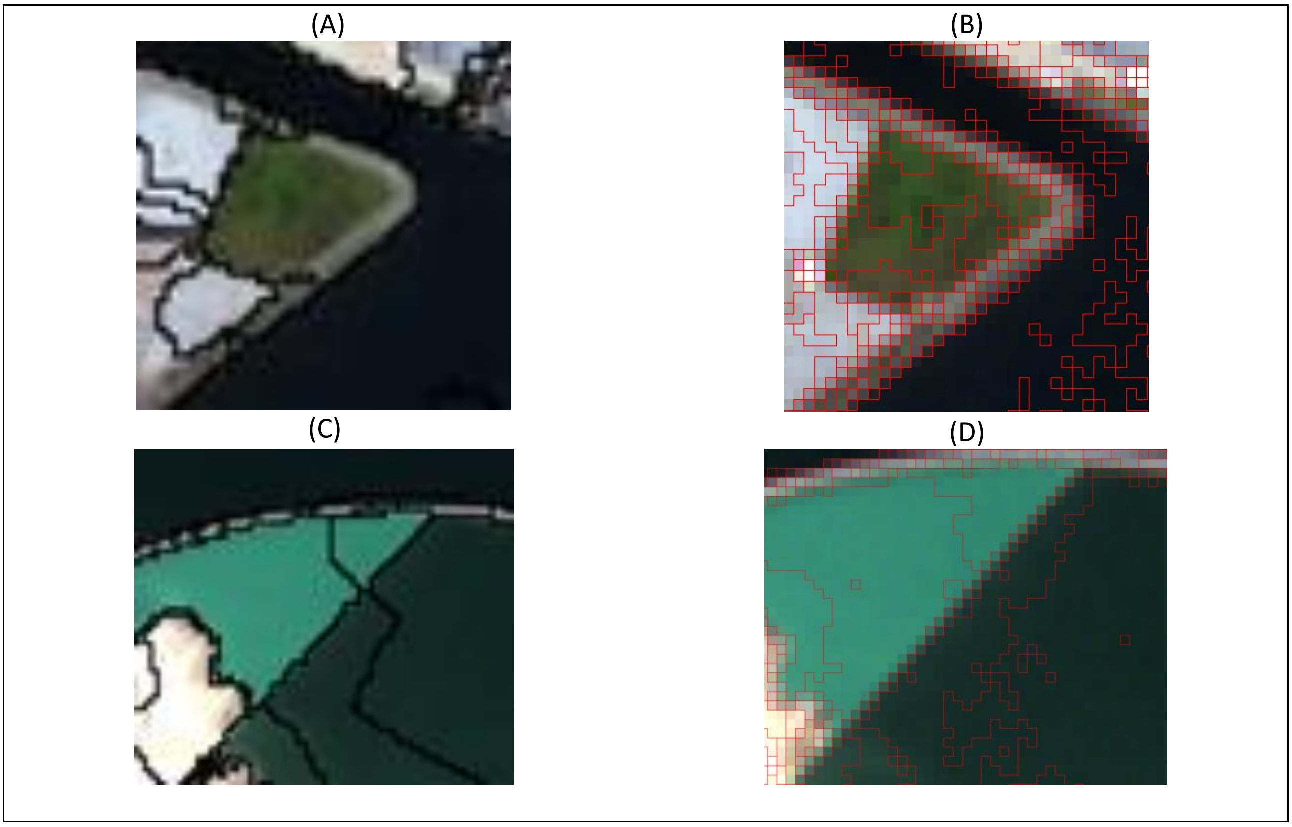


Figure 15 Segmented image examples, (A) (B) (C) (D)

The tests for segmentations were made with random forest classification, as it is expected that this will handle the segmentation best (Lei & Manchun, 2017). The results are shown in Table 3:

Table 3 Results of segmentation in ArcMap with different parameter value

Classifier	Min pixel	Spectral value	Spatial value	Accuracy %	Kappa %
Pixel	1	10	15	71,9	65,6
Pixel	1	15	10	72,5	67,3
Pixel	2	20	20	73,7	68,4
Pixel	1	15	15	81,1	76,0
Pixel	1	15	20	82,4	76,9
Pixel	1	20	15	83,2	77,2
Pixel	1	20	20	85,6	81,7

The standard values for spectral and spatial are 15 in ArcMap. A reduction in either value resulted in worse accuracy, with scores just above 70 %. This is a significant drop in accuracy and also below the results of the pixel classification approaches. The segmentation with both values at 15, giving an accuracy score at 81.1 % is only a 0.6 percentage points improvement compared to no segmentation. Increasing any of the two values leads to an increase in the accuracy. The spectral value had most effect with an accuracy improvement on 2.1 percentage points.

The option of increasing the minimum object size of a segment by forcing pixels into a group containing more than one pixel did not result in better performance. Like the other options that allowed greater size segments, the accuracy declined to be lower than without segments.

The results indicate that there are areas of the image classification that benefit from the segmentation as expected, but other areas clearly need to be evaluated as pixels because any parameter that dissolves the pixels in these areas, leads to a worse total accuracy. The following analyses with segmented images are therefor done with the highest spectral and spatial values.

3.3 Results of object based classification

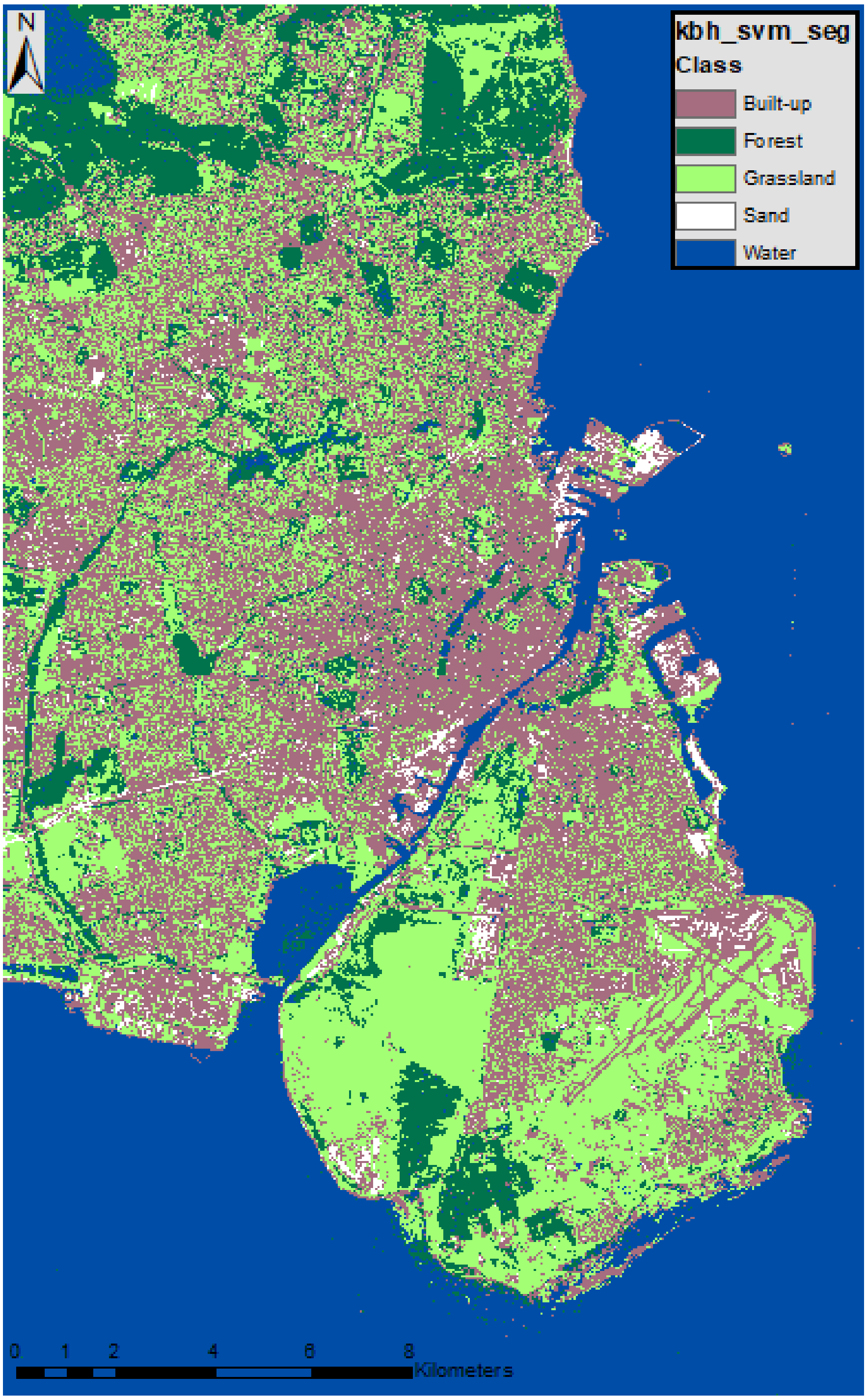


Figure 16 Land cover map made with support vector machine classification of segments in year 2015

Table 4 Accuracy for the support vector machine classification of segments in year 2015 (Figure 16)

Class	Built-up	Forest	Grassland	Sand	Water	Total	Users accuracy
Built-up	192	6	21	8	4	231	83,1 %
Forest	17	226	22	0	10	275	82,2 %
Grassland	37	11	203	7	9	267	76,0 %
Sand	3	1	2	110	0	116	94,8 %
Water	1	6	2	0	227	236	96,2 %
Total	250	250	250	125	250	1125	
Producers accuracy	76,8 %	90,4 %	81,2 %	88 %	90,8 %		85,2 %
Kappa							81,2 %

Table 4 shows that Introducing the segmentation improved the overall accuracy, as the classes of built-up, grassland and sand all improved with more than 7 percentage points on producer accuracy. Forest and water accuracy has fallen though. The water pixel errors are easy to spot, as the shallow water south of Amager now gets classified as land. See Figure 17 below.

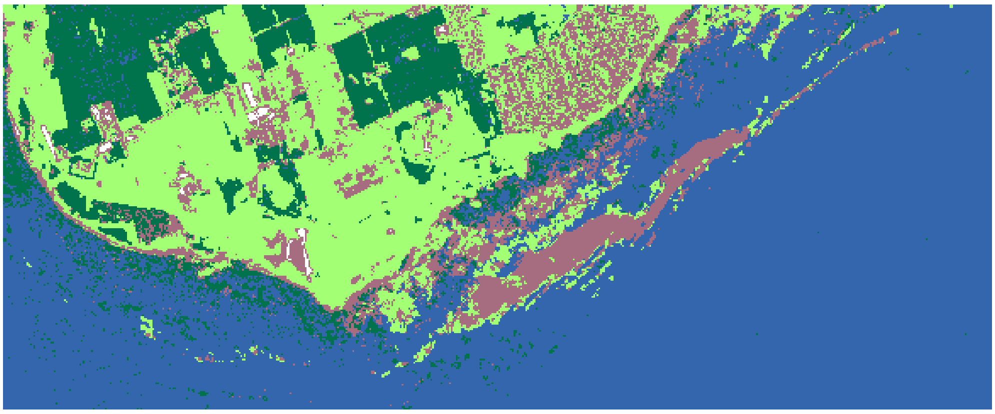


Figure 17 Land cover map of the southern part of Amager

The issues with built-up and grassland are still most common errors, as the majority of them happens in the low-density urban areas where buildings and grassland are mixed together. It is not a surprise that the classifier even with segmentation cannot clearly distinguish in these areas.

The user accuracies have all but water improved. The most common class that gets wrongly classified as water can also be seen on Figure 17, as the forests now show blue spots. Generally, the south of Amager seems to contain most of the errors.

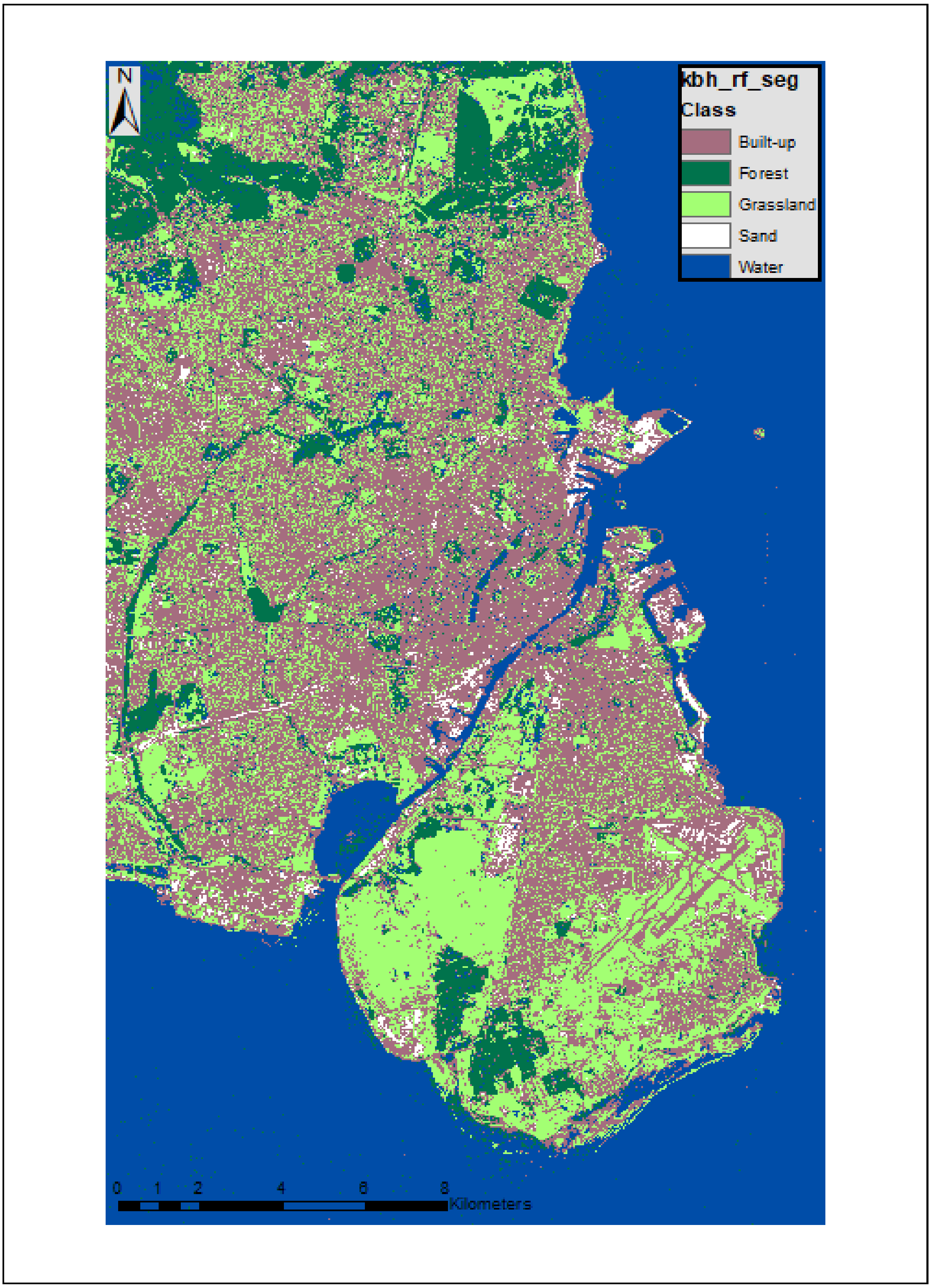


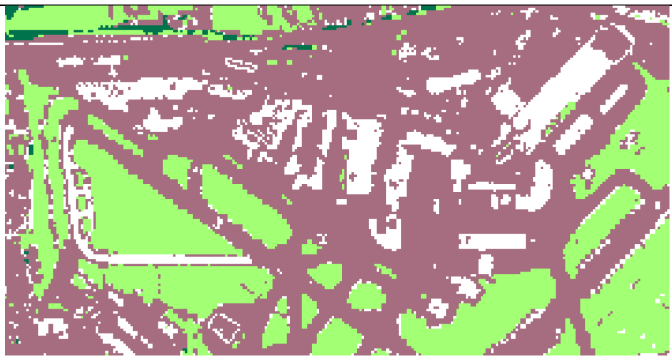
Figure 18 Land cover map made with random forest classification of segments in year 2015

Table 5 Accuracy for the random forest classification of segments in year 2015 (Figure 18)

Class	Built-up	Forest	Grassland	Sand	Water	Total	Users accuracy
Built-up	208	16	20	12	6	262	79,4 %
Forest	8	219	17	0	12	256	85,5 %
Grassland	27	8	206	8	7	256	80,5 %
Sand	3	1	2	105	0	111	94,6 %
Water	4	6	5	0	225	240	93,8 %
Total	250	250	250	125	250	1125	
Producers accuracy	83,2 %	87,6 %	82,4 %	84 %	90 %		85,6 %
Kappa							81,7 %

As shown in Table 5, 85.6 % was the highest accuracy score achieved in the study area. All producer accuracies are above 80 % with the highest score again from the water class, even though RF also has had some issues with shallow water in the south. The user accuracies are also around 80 % or above, with the highest score being sand. The errors in this class often occur around rooftops or buildings with distinct materials.

Another area that is hard for the algorithms to classify correct is the concrete around the buildings of Copenhagen Airport, as seen below in Figure 19.



(A)

(B)

Figure 19 Land cover map (A) and image (B) both of Copenhagen Airport

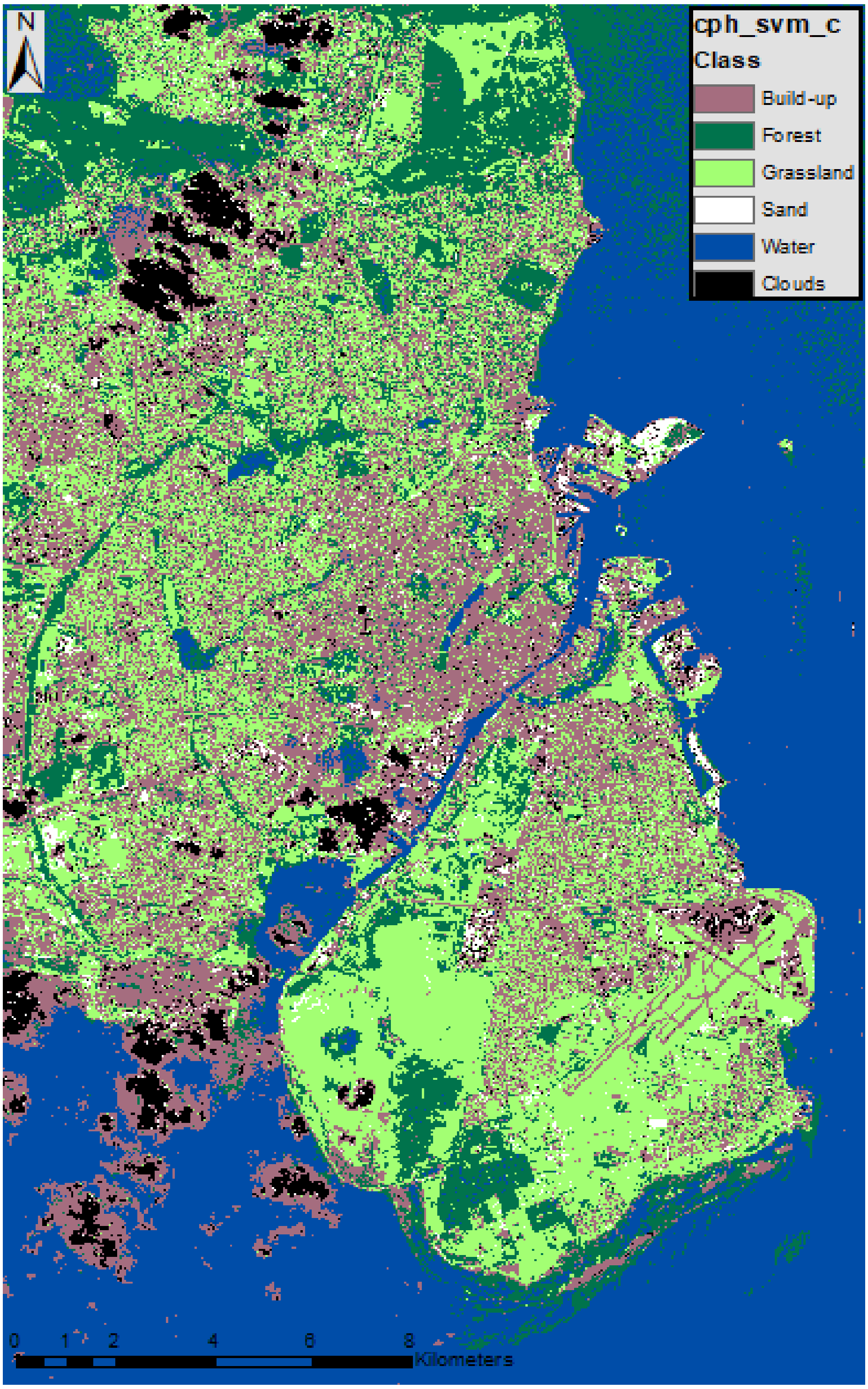


Figure 20 Land cover map made with support vector machine classification of segments in year 2017

Table 6 Accuracy for the support vector machine classification of segments in year 2017 (Figure 20)

Class	Buildings	Forest	Grassland	Sand	Water	Clouds	Total	Users accuracy
Buildings	126	13	13	13	17	86	268	47,0 %
Forest	7	215	15	0	42	0	279	77,1 %
Grassland	61	12	188	4	7	0	272	69,1 %
Sand	4	0	0	108	0	10	122	88,5 %
Water	36	10	34	0	184	0	264	69,7 %
Clouds	16	0	0	0	0	154	170	90,6 %
Total	250	250	250	125	250	250	1375	
Producers accuracy	50,4 %	86 %	75,2 %	86,4 %	73,6 %	61,6 %		70,9 %
Kappa								64,8 %

The SVM results in Table 6 showed the worst result of this study, with an accuracy at 70,9 %. There are clearly several issues, especially with the class built-up area where both producer and user accuracy is around 50 %. The pixels containing buildings are often classified as grassland and on the other hand both large parts of the clouds and the areas where the clouds cast shadows are classified as built-up.

Other parts of the shadow area from clouds get assigned to the water class (Figure 21). Another contributor to the low accuracy is the classification of water where the class accuracy is around 70 % and like the 2015 image classification with segmentation, the shallow water areas south of Amager is an issue as well as both lakes and ocean in the northern parts of the map.

The airport is also an interesting area where the surface around the buildings now gets classified as clouds.

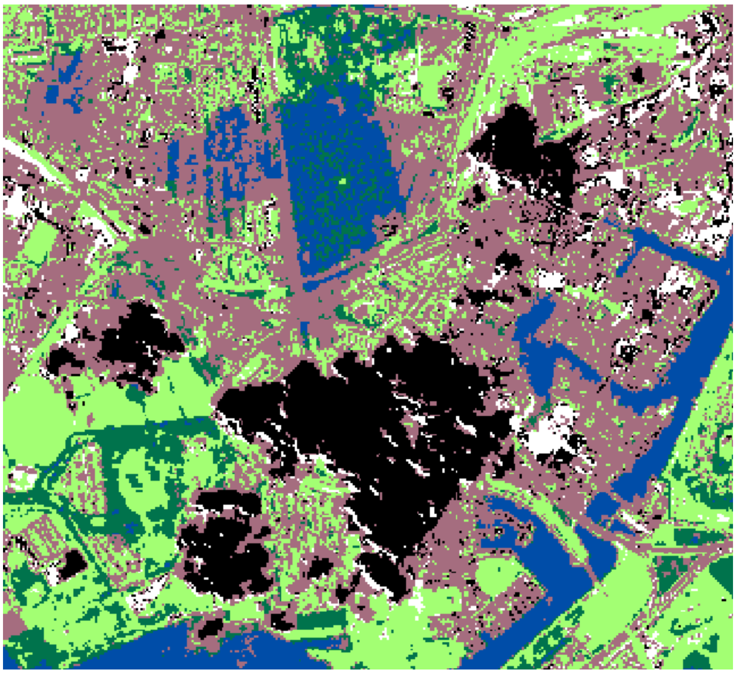


Figure 21 Land cover map from 2017 image showing the effect of clouds

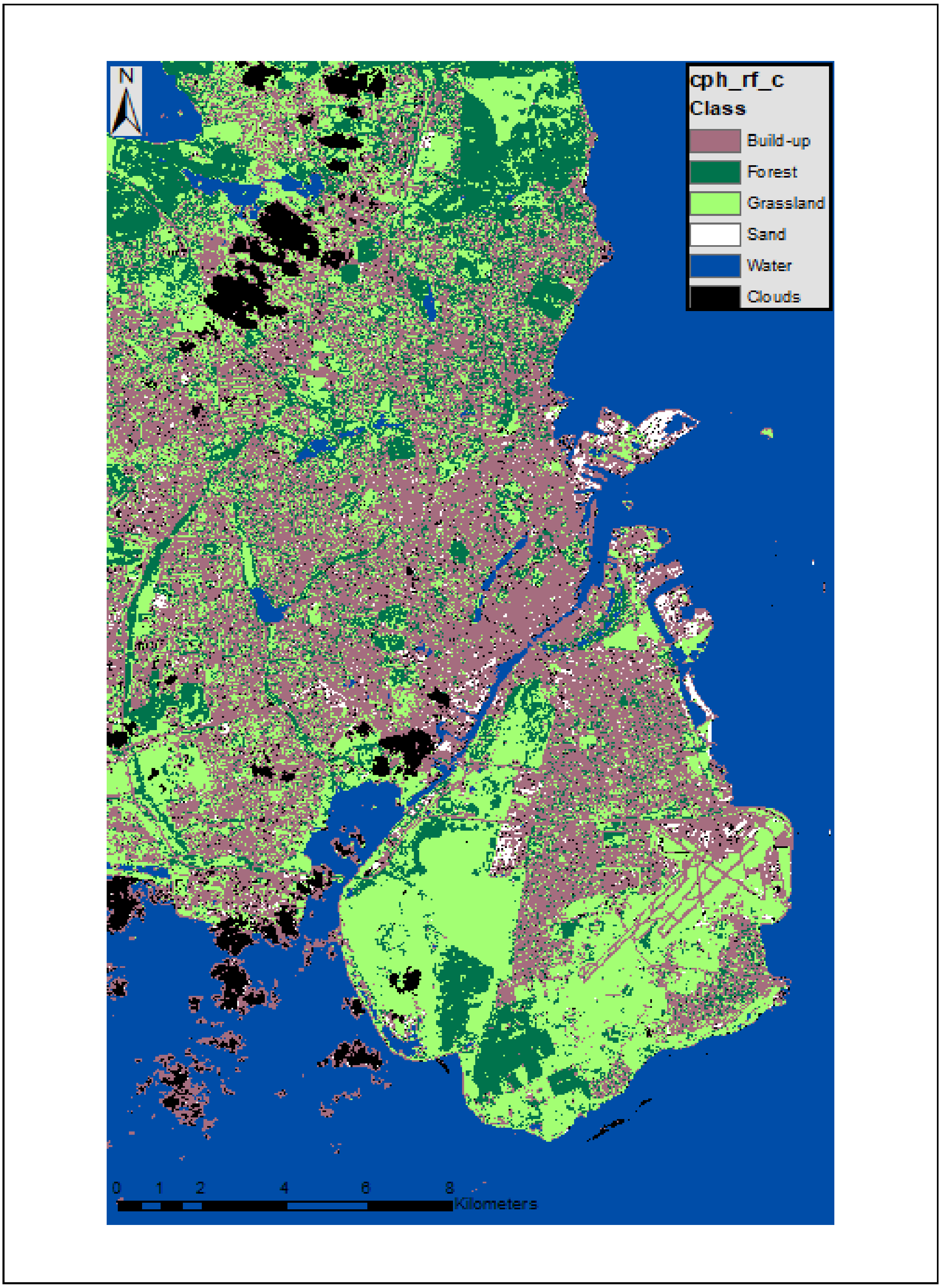


Figure 22 Land cover map made with random forest classification of segments in year 2017

Table 7 Accuracy for the random forest classification of segments in year 2017 (Figure 22)

Class	Buildings	Forest	Grassland	Sand	Water	Clouds	Total	Users accuracy
Buildings	182	21	24	13	2	26	268	67,9 %
Forest	11	216	21	0	0	0	248	87,1 %
Grassland	34	12	201	4	0	0	251	80,1 %
Sand	4	0	3	108	0	3	118	91,5 %
Water	2	1	1	0	244	0	248	98,4 %
Clouds	17	0	0	0	4	221	242	91,3 %
Total	250	250	250	125	250	250	1375	
Producers accuracy	72,8 %	86,4 %	80,4 %	86,4 %	97,6 %	88,4 %		85,2 %
Kappa								82,1 %

Shown in Table 7 are the best classification of the 2017 image, with an accuracy score of 85 % even with clouds and their shadows. Built-up is again the class that has the lowest accuracy score, where common errors are around the clouds and the shadow regions.

The producer accuracy of the water class on 97.6 % stands out. This is the first of the classifiers using a segmented image that does not struggle with the ocean pixels. There are still some mistakes south of Amager and in the airport area (Figure 23), but the rest is handled correctly.

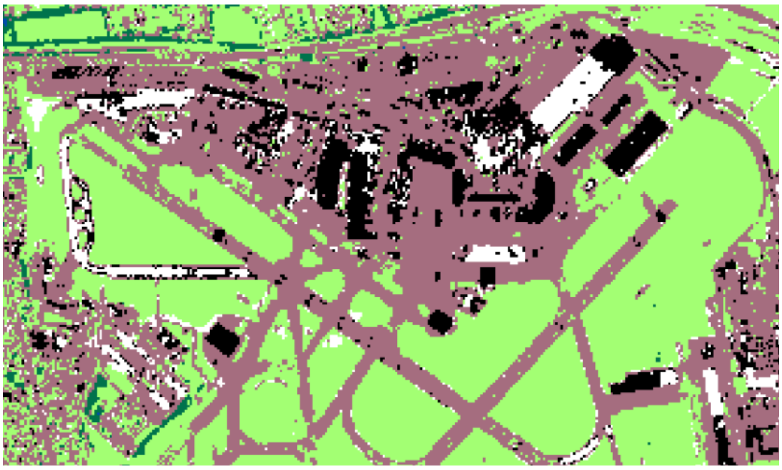


Figure 23 Misclassification in the airport area do to clouds in the image

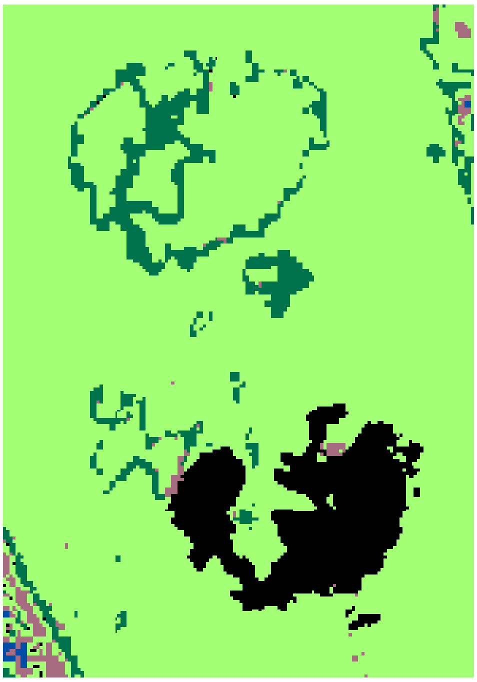


Figure 24 Cloud shadow effect leading to wrong classification

Just like in the SVM classification the shadow area of the clouds is an issue. Here the center of the area is fine, but the edge of the area is classified as forest instead of grassland, see Figure 24.

Table 8 Accuracy for the support vector machine classification of segments in year 2017 (Figure 25)

Class	Built-up	Forest	Grassland	Sand	Water	Total	Users accuracy
Built-up	180	6	21	9	1	217	82,9 %
Forest	17	216	22	0	2	257	84,0 %
Grassland	49	21	201	7	2	280	71,8 %
Sand	3	1	3	109	0	116	94,0 %
Water	1	6	3	0	245	255	96,1 %
Total	250	250	250	125	250	1125	
Producers accuracy	72 %	86,4 %	80,4 %	87,2 %	98 %		84,5 %
Kappa							80,4 %

Table 8 shows an accuracy at 84,5 %, the best class accuracy is water with a 98 % score. This is also visible by the southern coastline, where the shallow waters are correctly classified. There are still some errors in the small lakes on land though; both can be seen on Figure 25.

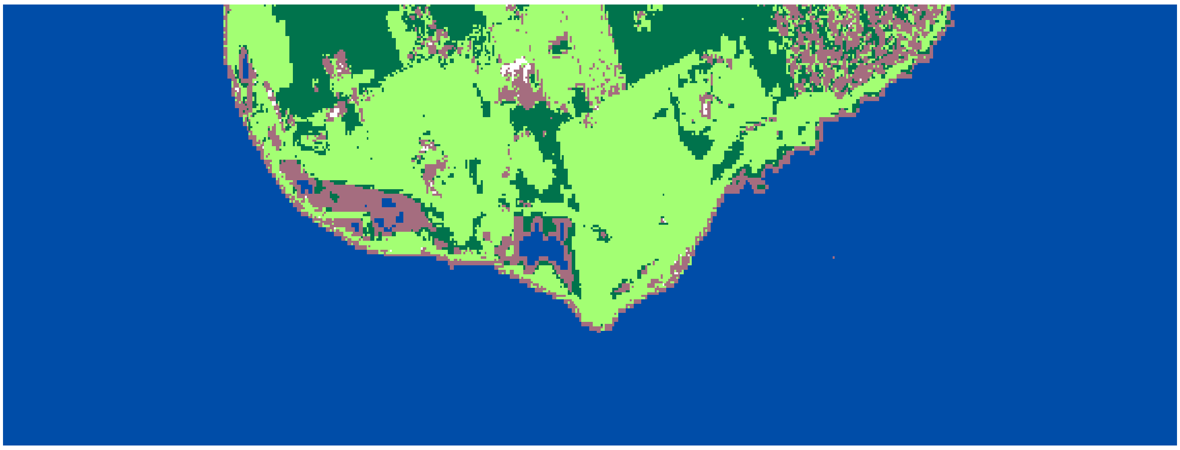


Figure 25 Southern coast classified by SVM

The classifiers worst accuracy is identifying grassland and built-up area. With the lowest score being 72 % producer accuracy for the built-up class. The full map is shown on the next page in Figure 26.

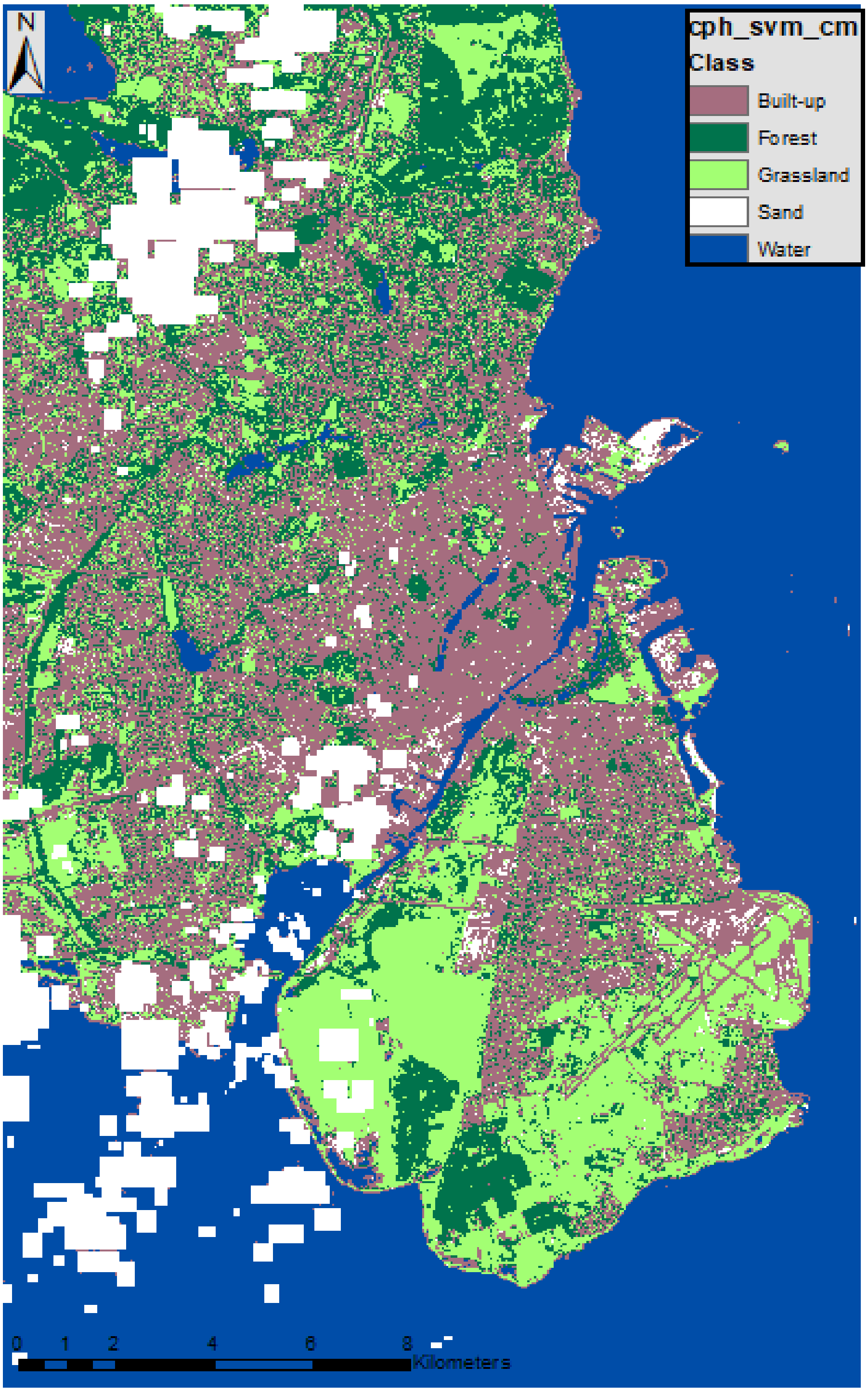


Figure 26 Support vector machine classification of segments in year 2017 (clouds removed)

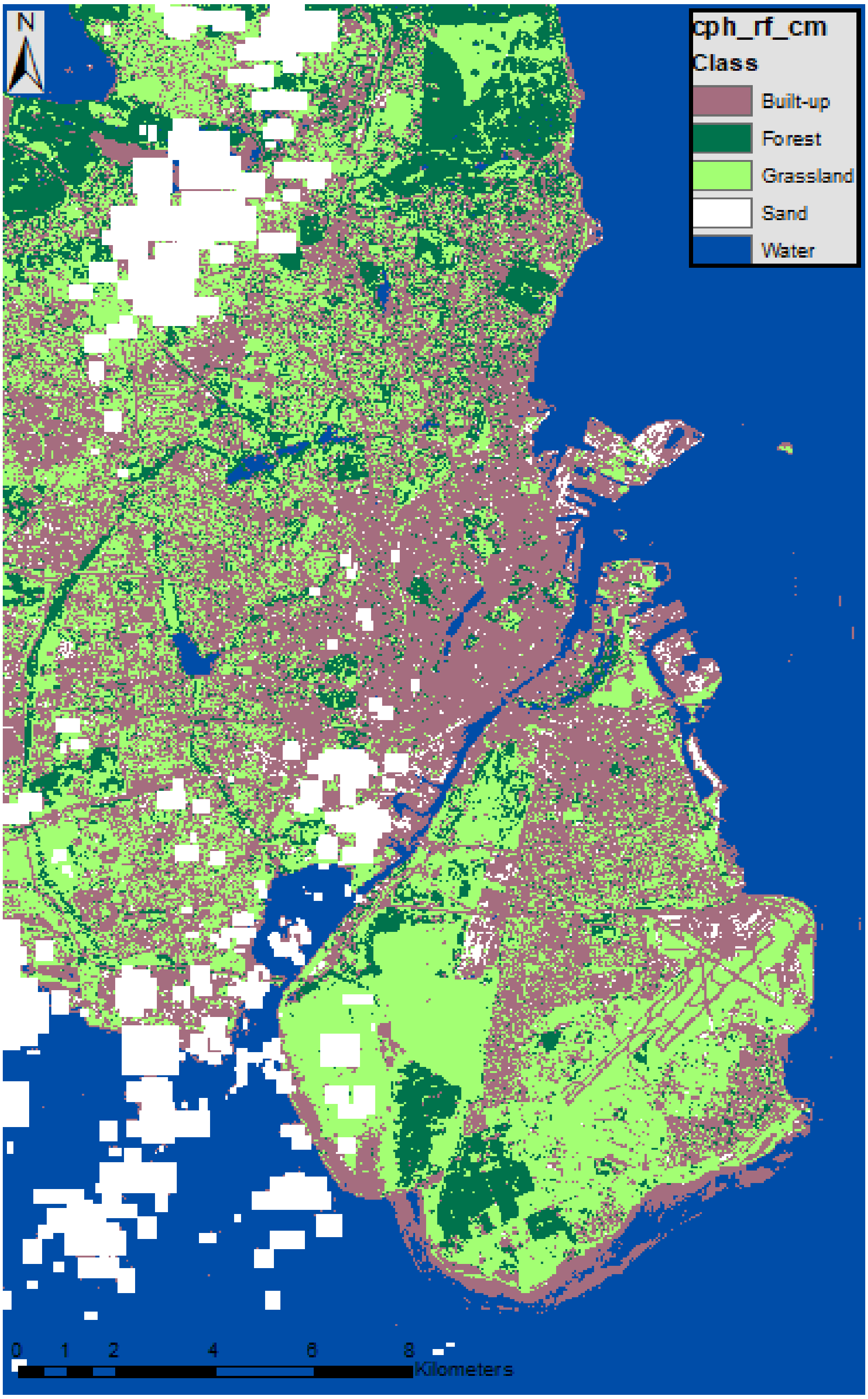


Figure 27 Random forest classification of segments in year 2017 (clouds removed)

Table 9 Accuracy for the random forest classification of segments in year 2017 (Figure 27)

Class	Built-up	Forest	Grassland	Sand	Water	Total	Users accuracy
Built-up	205	16	18	11	14	262	77,7 %
Forest	8	218	18	0	4	256	87,9 %
Grassland	30	9	206	8	5	256	79,8 %
Sand	3	1	3	106	0	111	93,8 %
Water	4	6	5	0	227	240	93,8 %
Total	250	250	250	125	250	1125	
Producers accuracy	82 %	87,2 %	82,4 %	84,8 %	90,8 %		85,5 %
Kappa							81,6 %

The total accuracy of 85,5 % is in the high end (Table 9), but surprisingly when compared to the same 2017 image classified with random forest and adding the clouds as a class, the high accuracy on identifying water has dropped from 97,6 to 90,8 %. Again, the south coast clearly shows some of the issues, both on land and in the sea. (Figure 28)

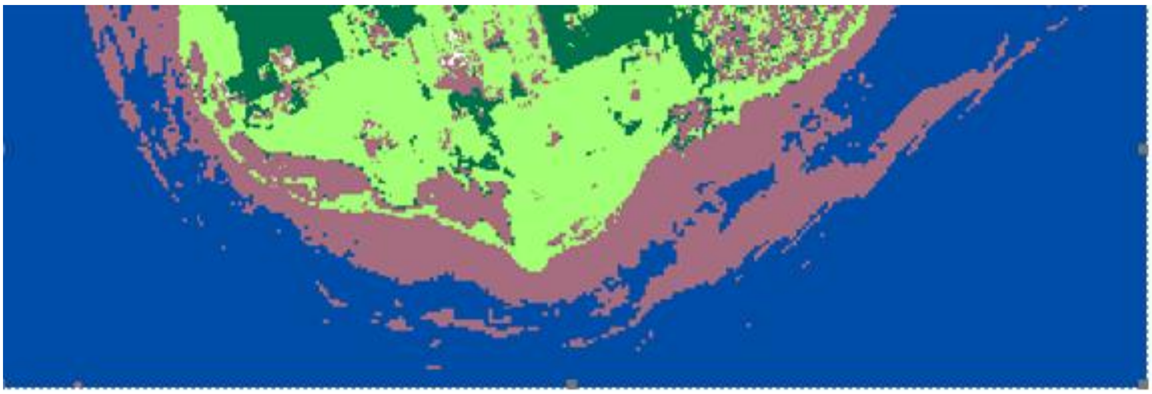


Figure 28 Southern coast 2017 random forest and clouds masked

The improvement compared with the RF classification including clouds comes primarily in the built-up class, where the producer accuracy is up to 82 %. Forest accuracy also shows a slight improvement increasing 0,8 %, while grassland accuracy increases with 2 %, from 80,4 to 82,4.

3.4 Comparison of the classifiers performance

Table 10 Results of classifications

Classifier	Year	Object/pixel	Accuracy %	Kappa %
SVM	2017	Object	70,9	64,8
RF	2015	Pixel	80,5	75,3
SVM	2015	Pixel	81,1	76,0
SVM*	2017	Object	84,5	80,4
SVM	2015	Object	85,2	81,2
RF	2017	Object	85,2	82,1
RF*	2017	Object	85,5	81,6
RF	2015	Object	85,6	81,7

** Clouds are masked out of the image*

Clearly RF classification done with an object based approach performed best, as shown in Table 10, while SVM also performed well on the 2015 image. This result fits fine with other studies, as already introduced (Lei & Manchun, 2017).

The introduction of the clouds in the 2017 image led to the worst result with the SVM classification, but besides this, the classifications on segmented images performed very similar with an accuracy around 85 %. This seems to be what should be expected from classifying a sentinel 2 image using the method introduced in this study.

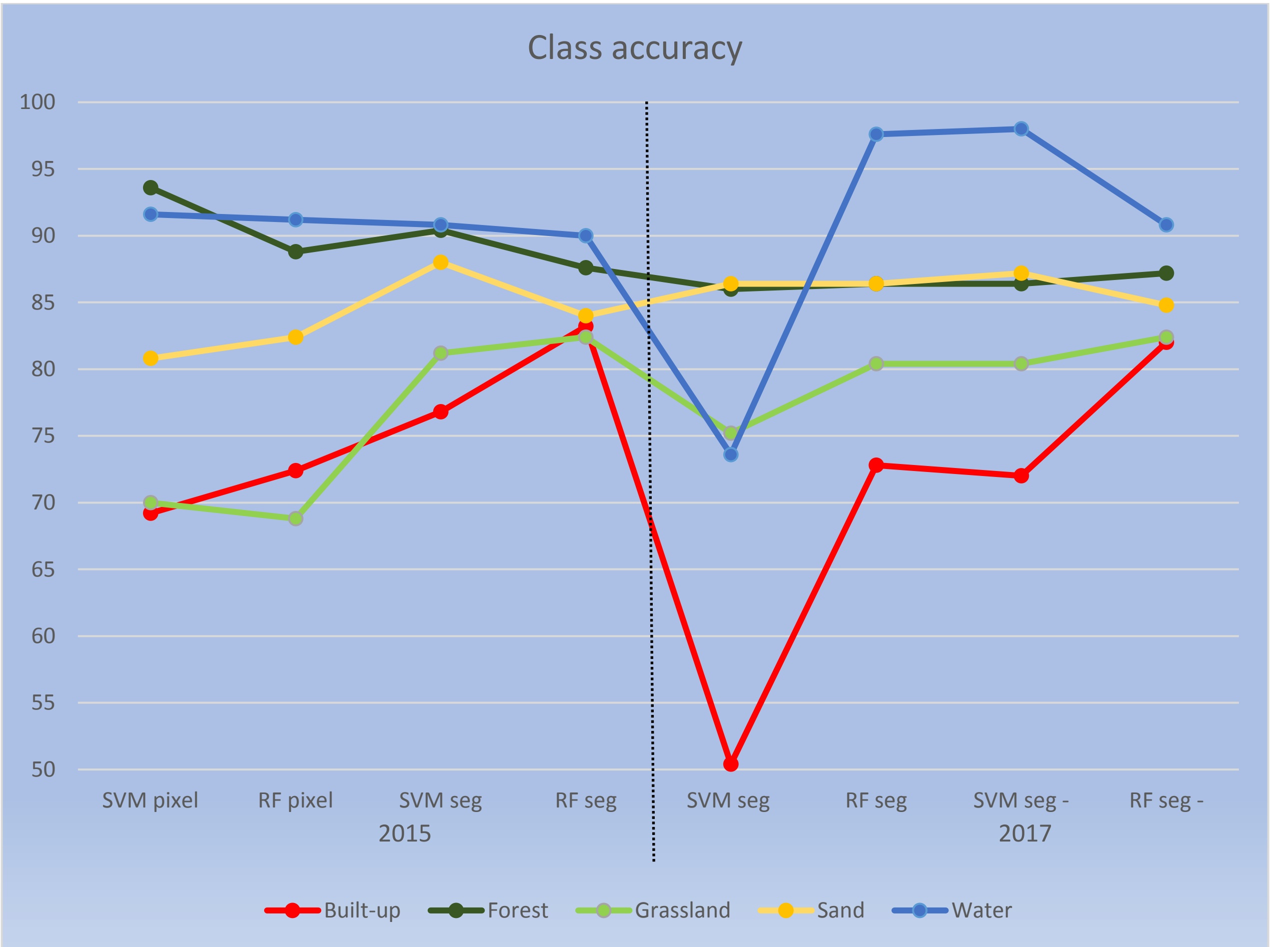


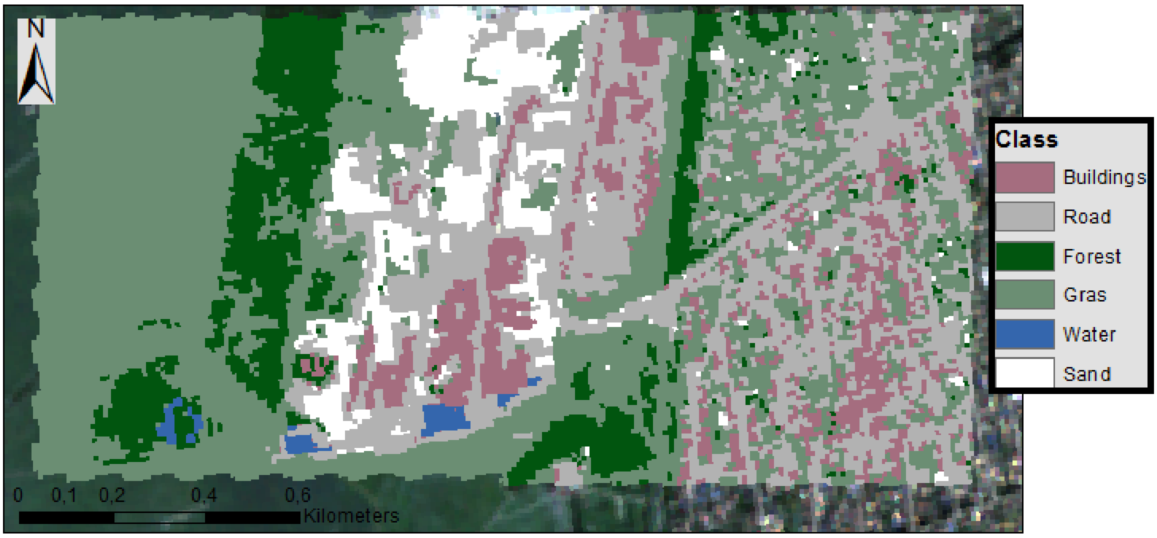
Figure 29 Class producer accuracy in % for each classification

Figure 29 includes all the producer accuracy scores for the different classifications, and illustrates how the class accuracy can be up to 25 percentage points apart in the random forest pixel classification from 2015, while when segmenting the same data it only differs with 10 percentage points. The 2017 random forest classification without clouds (RF seg -), shows the same pattern with class accuracy scores close to each other.

Another observation is that the accuracy for the sand and forest classes is almost consistant inside the 85-90 % bracket with the segmented images. They show no effect from classifing clouds as a class.

3.5 Classifying with addition of a Lidar data

(A) Land cover map classified with support vector machine



(B) Land cover map classified with support vector machine and a digital surface model added

Figure 30 Land cover maps of the small study area classified with and without the digital surface model.

Table 11 Accuracy for the support vector machine classification without the digital surface model

Class	Buildings	Roads	Forest	Grassland	Sand	Water	Total	Users accuracy
Buildings	36	7	2	2	1	2	50	72,0 %
Roads	6	36	2	2	3	0	49	73,5 %
Forest	1	2	41	4	0	2	50	82,0 %
Grassland	6	5	3	40	2	0	56	71,4 %
Sand	0	0	0	2	44	0	46	95,7 %
Water	1	0	2	0	0	46	49	93,9 %
Total	50	50	50	50	50	50	300	
Producers accuracy	72 %	72 %	82 %	80 %	88 %	92 %		81,0 %
Kappa								77,2 %

Table 12 Accuracy for the support vector machine classification with the digital surface model

Class	Buildings	Roads	Forest	Grassland	Sand	Water	Total	Users accuracy
Buildings	38	7	2	1	1	0	49	77,6 %
Roads	5	36	2	2	3	0	48	75,0 %
Forest	1	2	43	4	0	0	50	86,0 %
Grassland	6	5	3	41	2	0	57	71,9 %
Sand	0	0	0	2	44	0	46	95,7 %
Water	0	0	0	0	0	50	50	100,0 %
Total	50	50	50	50	50	50	300	
Producers accuracy	76 %	72 %	86 %	82 %	88 %	100 %		84,0 %
Kappa								80,8 %

There is a clear improvement with the addition of a DSM in the study area, shown in Table 11 and Table 12. The accuracy increases 3 percentage points and there is either no effect or an improvement in the individual class accuracy. The class that improves the most is water that achieves 100 % in both producer and user accuracy. There seems to be more changes in the western part of the map (Figure 30). As that area consists of high buildings and areas with mixed vegetation, this is the part where the DSM should have an effect.

A visual inspection for differences shows that the western part of the area clearly improves; both around the water, but also roads between the tall buildings become easier to identify. The eastern part of the map shows no signs of improvement, but that was expected as it is low-density buildings. The only eastern area that shows any height difference on the DSM is some trees in the north, but they are already rightly classified without the DSM.

Visual inpspection of the DSM shows that it contains errors as well. Below in Figure 31 are an example; The southern part of the building “Ottetallet”, located in the center of the area. The DSM has wrongly been triangulated between water and building points.

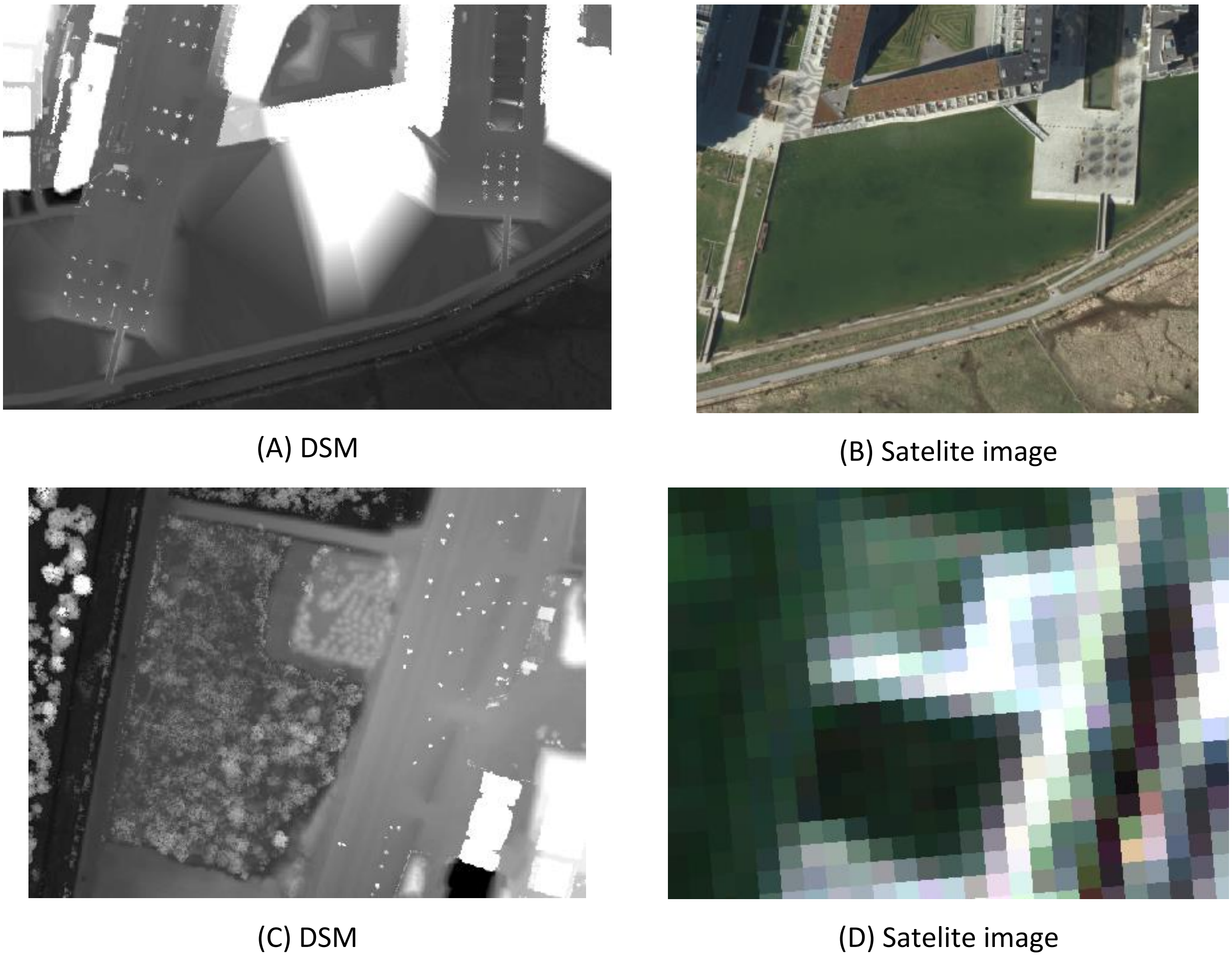


Figure 31 Areas where the digital surface model contains errors

Another issue discovered in the western part, as shown in Figure 31, is an issue with data as it has not been collected at the same time. The DSM shows vegetation, while the image shows that the construction site has expanded its territory.

3.6 Change detection in the study area

3.6.1 Change with TerrSet (Idrisi)

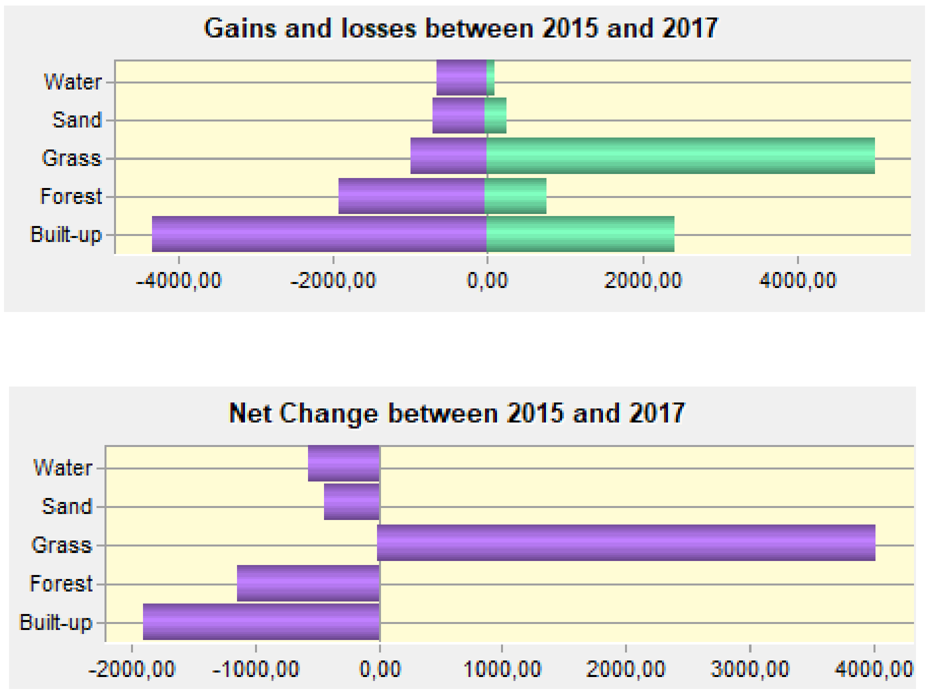


Figure 32 Changes in square kilometres between the classes in 2015 and 2017

TerrSet includes a tool for comparison of the changes in the area of each class in the modeled time period. In Figure 32 it is shown in square kilometers and every class both have gains and losses. The biggest trend seems to be from built-up into grassland.

The net changes look very one sided, all classes except grassland has dropped in area. The class with the greatest loss is built-up, where 2000 square kilometers are now classified as something else. In a growing city this seems counter intuitive, even with some greening of the city taking place, as for example new parks or recreative areas. The area classified as grassland has experienced a net gain of around 4000 square kilometers. This could indicate a problem in identifying the class in the 2015 image, that was not experienced in the 2017 image.

The persistent map shown in Figure 33, shows where the best classification from 2015 and 2017 have the same class for a pixel. If there is difference between them the pixel will be blank.

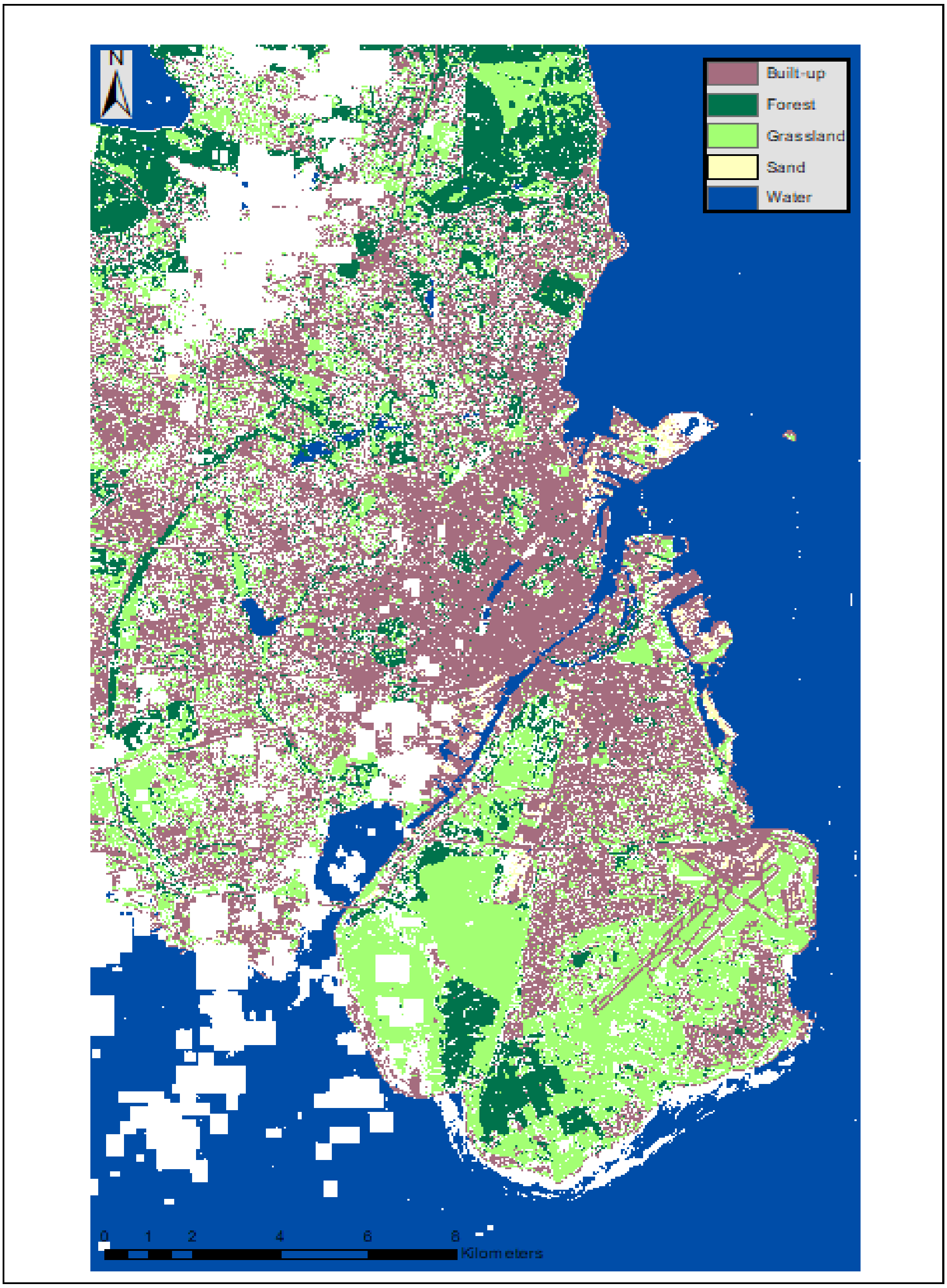


Figure 33 Map of persistent classification in 2015 and 2017

Taking a look at the class by class changes, they are as follows:

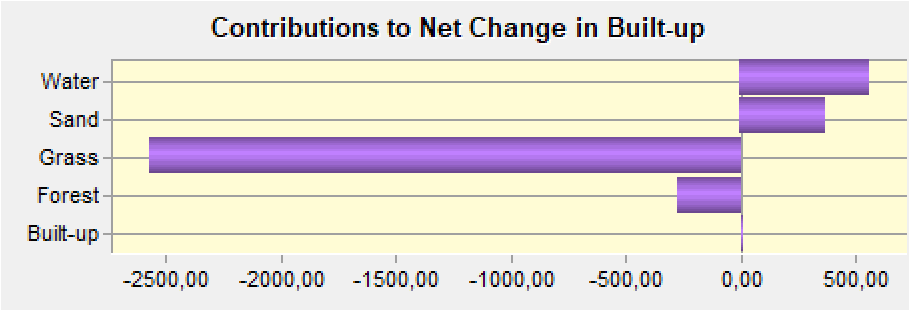


Figure 34 Changes in Built-up area

As seen above in Figure 34 the biggest change is from built-up to grassland. Between the other classes the net changes have been subtler in the range 250-550 square kilometers. Increase in built-up area on the expense of sand and water is as expected; as sand mostly covers building sites in Copenhagen, they should change to built-up when construction is done. The change from water to built-up, can besides misclassification be explained with the landfill happening along the harbor where high-rises are constructed. The change from built-up to forest, could be due to initiatives in Copenhagen trying to increase the number of trees, but also as result of classification parameters being different (Astrup, 2017).

The big change from built-up to grassland cannot be explained form a city planning perspective, so most likely it is due to misclassification, with the year 2015 classification identifying too little and the 2017 identifying too much. Even though the images are from the same season of the year, it is clear to see that the year 2017 was wetter, resulting in greener grassland that may have helped the classifier differentiate between the suburban gardens and built-up areas.

Figure 35 on the next page shows where in the study area the changes has happened. It also reveals that most of the area with new built-up is due to misclassified water bodies.

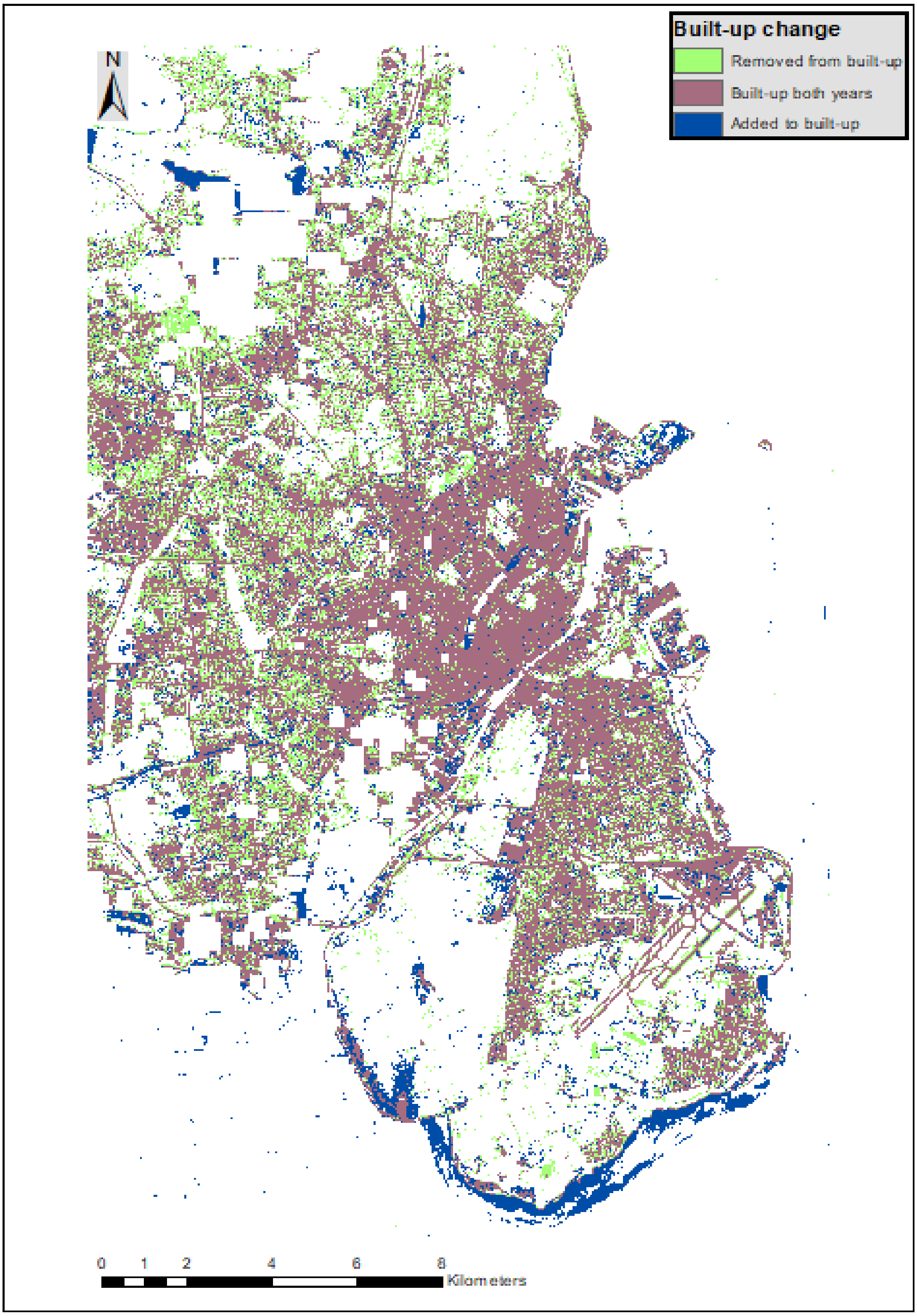


Figure 35 Map of changes in the built-up class between 2015 and 2017

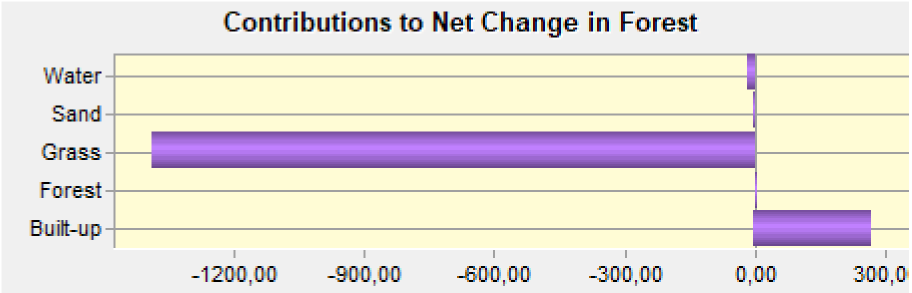


Figure 36 Changes in the forest class

The changes in forest besides the increase from built-up area, is a net loss to grassland (Figure 36). Again, the larger perpetuation may have influenced the ability of the classifiers to differentiate between trees and grassland.

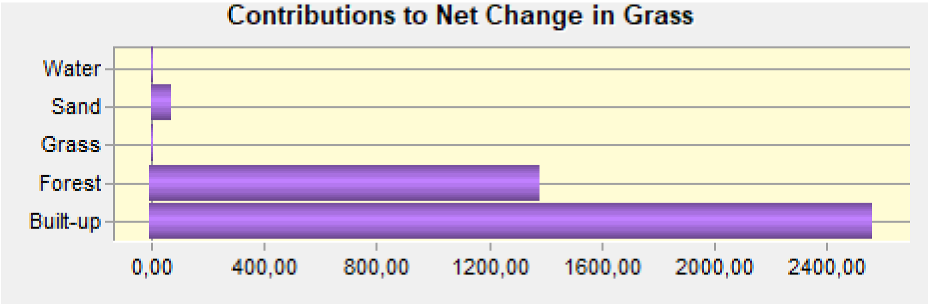


Figure 37 Changes in the grassland class

The change in grassland is, besides illustrated in Figure 37 with the net change, also seen in the trend map in Figure 38. The trend map highlights that the changes mostly have occurred outside the city center, that also is the center of the image.

Figure 39 on the next page clearly shows that the large areas with grassland on Amager is correctly identified in both images, while the suburbs show patches along the roads and rows of houses, where more areas are being classified as grassland.

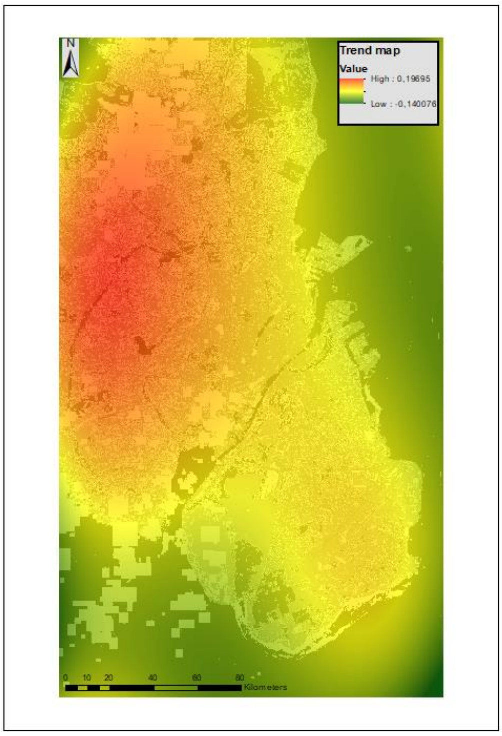


Figure 38 Map of trend for classified built-up change to grassland

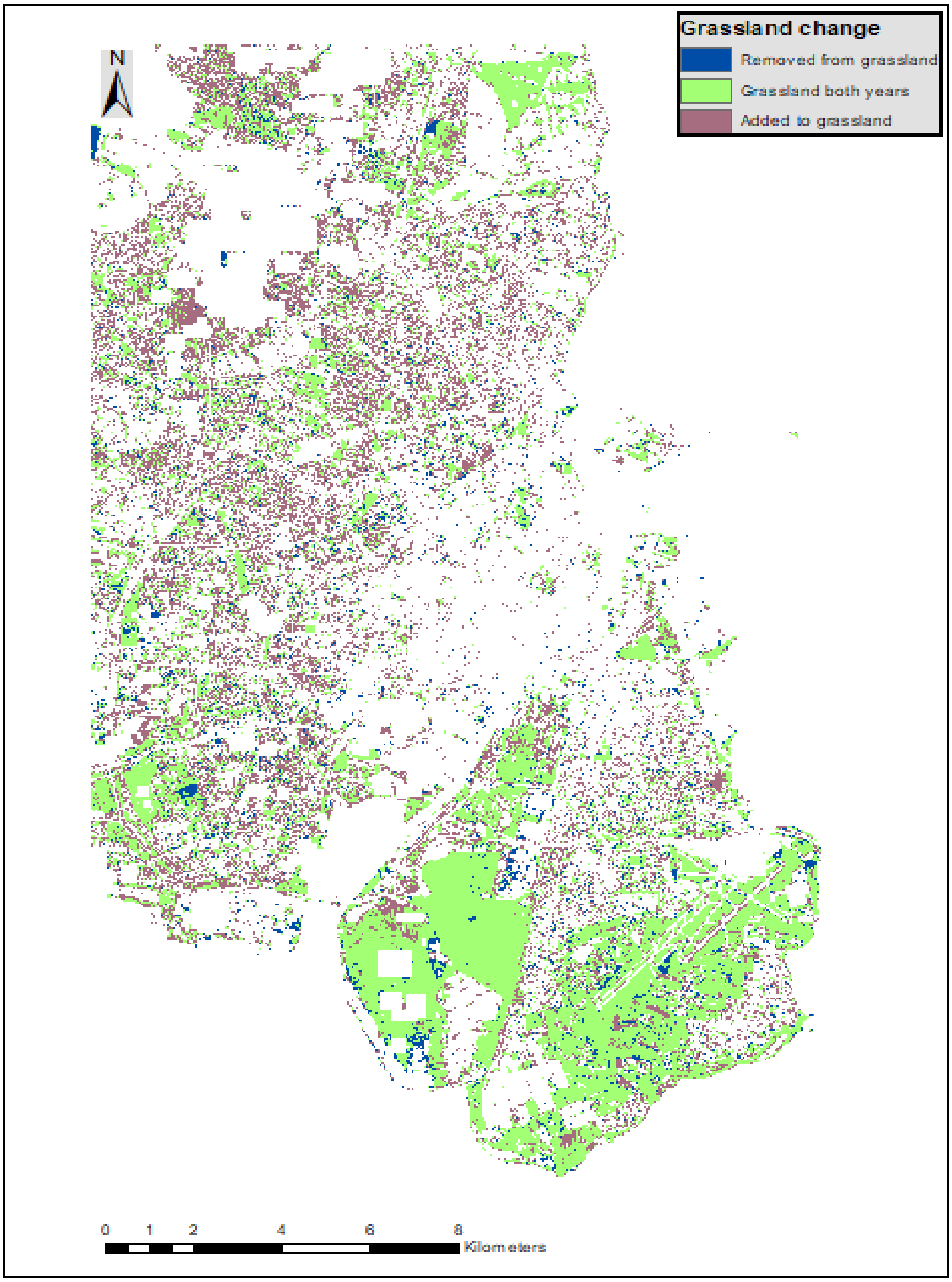


Figure 39 Map of changes in the grassland class between 2015 and 2017

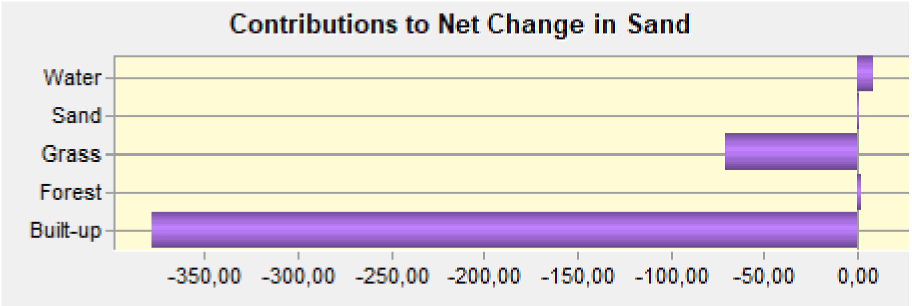


Figure 40 Changes in area classified as sand

The reduction of sand and increase in built-up and grassland is because big construction sites like "Ørestaden" are closing in on being fully developed and that the landfill in "Nordhavn" has changed from the raw material poured into the harbor basin, into green areas waiting to be further developed. (Figure 40)

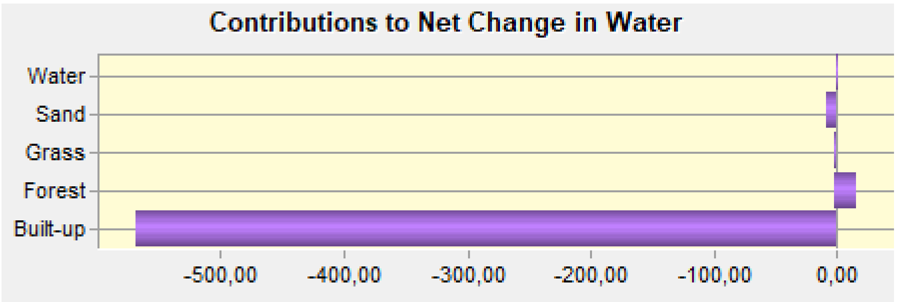


Figure 41 Changes in area classified as water

The change in the area classified as water are mostly due to misclassification of the shallow water south of Amager and a couple of lakes in the northern part of the image. The rest is due to landfill in the area. (Figure 41)

3.6.2 Change with ArcMap

2015 to 2017 in satellite images with focus on built-up area where changes are known to happen.

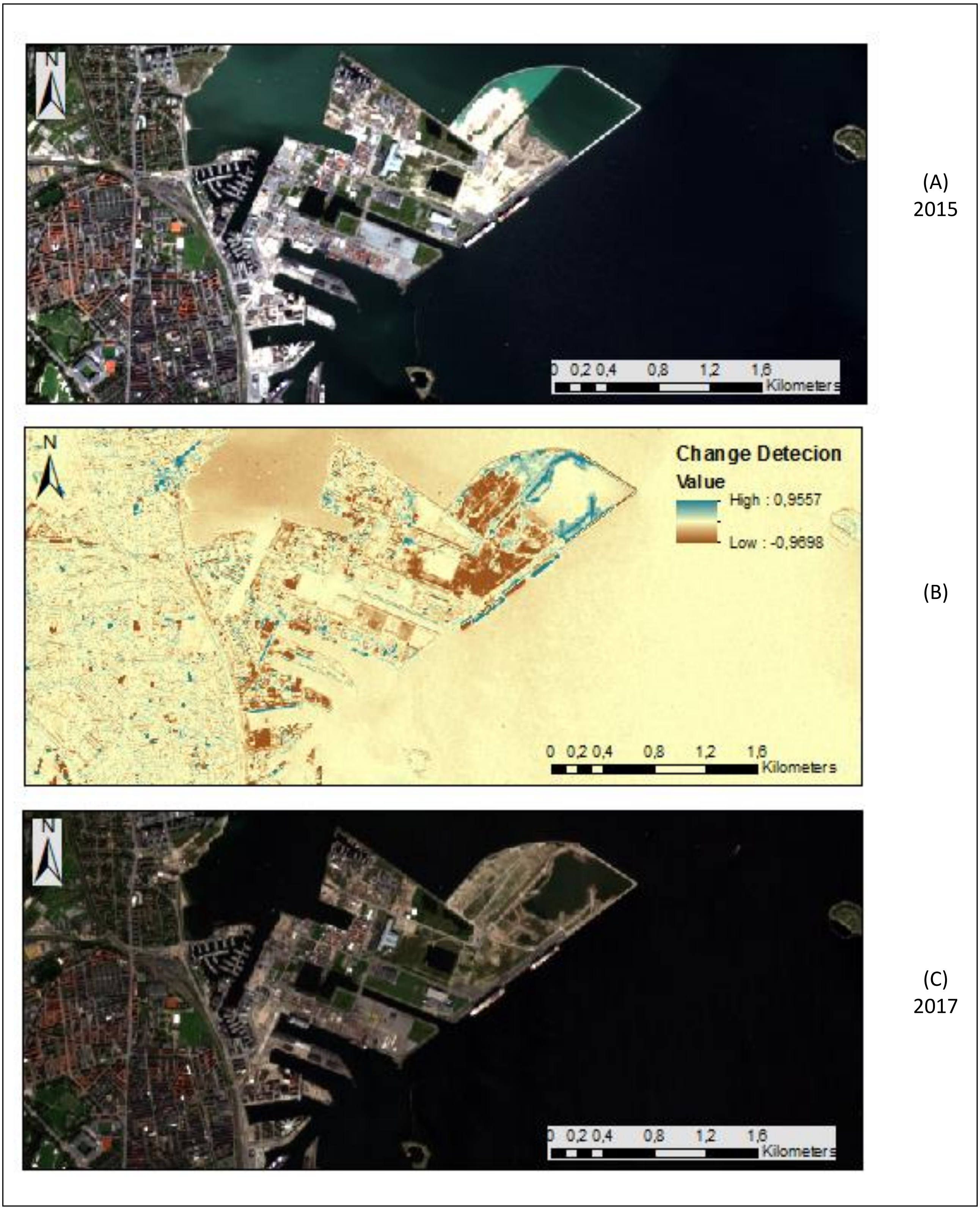
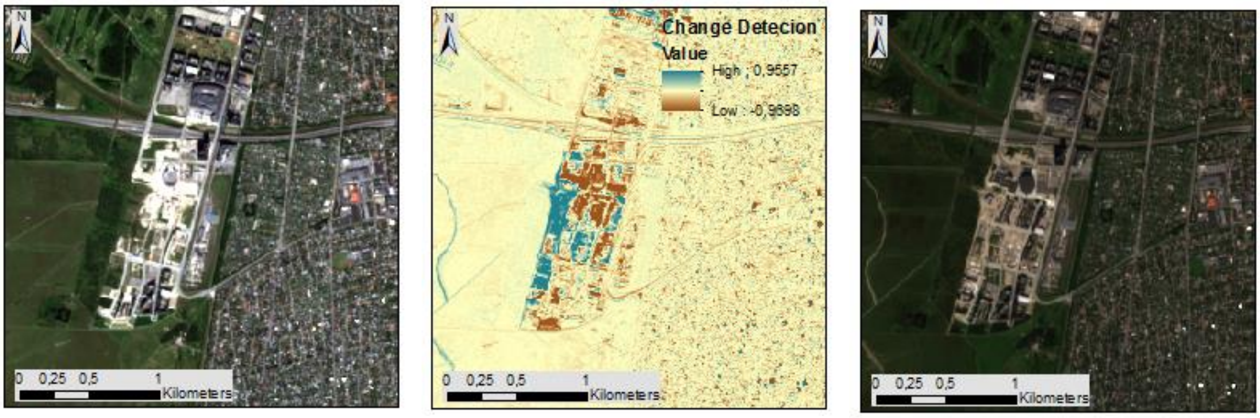


Figure 42 Nordhavn in 2015 (A), 2017 (C) and the change detection value map (B)

The first area is “Nordhavn” as shown above in Figure 42. It is one of the areas that have been developing the last few years. The former commercial harbor is going to have a lot of new residential buildings and is also growing by the rate that the earth dug up from the “Copenhagen metro city ring” construction and put into the basin in the north as landfill. The change detection map shows high values where the landfill has changed the area from water to land. On the other hand, it shows a low value in the area where it was just sand in 2015 and now has reached a point with a darker surface and some low vegetation.



(A) Image from 2015 (B) 2015-2017 (C) Image from 2017

Figure 43 Ørestad syd in 2015 (A), 2017 (C) and the change detection value map (B)

The second area is “Ørestad Syd” as shown in Figure 43. This is another developing area of Copenhagen, mostly with residential buildings. The change maps show low values in the areas where there were building sites in 2015 and finished buildings in 2017, while the areas in the western part of the development area are showing high values. This is a result of the removal of the vegetation and turning it into building sites.

3.6.3 Markov Chain prediction in TerrSet

Shown in Table 13, probabilities of change from a class in 2017 to a class in 2020 in percentage:

Table 13 Prediction matrix for class changes from 2017 to year 2020

Class	Built-up	Forest	Grassland	SAND	Water
Built-up	60,7 %	5,8 %	31,2 %	1,6 %	0,7 %
Forest	9,3 %	55,3 %	34,8 %	0,2 %	0,4 %
Grassland	15,3 %	2,5 %	81,4 %	0,8 %	0,0 %
SAND	66,0 %	1,7 %	19,9 %	12,2 %	0,2 %
Water	4,6 %	0,0 %	0,3 %	0,1 %	95,0 %

The numbers in the prediction matrix are probably overestimating the chances of the other classes becoming grassland and likewise the chance of water being converted to land. It seems likely that if water will change class it will be into built-up, with the current planning of new residential areas along the harbor and also because the last landfill in “Nordhavn” is still not completed.
The relatively small chance (12 %) of sand being sand after three years seems to fit with the notion that the class mostly covers building sites that converts into other classes when completed.

The matrix exposes that even though the net area of built-up and forest went in favor of more trees in Copenhagen, the chance of a forest area being converted to built-up is higher than the other way around.

The accuracy of the change maps produced and the predicting of the future can be derived from the accuracy of the land cover maps used to compute the changes. As the accuracy of the 2015 map was 85,6 % and the 2017 map was 85,2 %, the accuracy of the results, maps and predictions produced in TerrSet is therefore 0,856*0,855 = 0,732 or 73,2 %.

4. Discussion

Remote sensing of the earth surface through high resolution satellite imagery already has a lot concrete applications, from environmental monitoring to tracking urban developments. All needed in order to ensure that planning can be done sustainably and on the basis of knowledge, both of the past and the current conditions. The global trend of urbanization puts pressure on urban development and increases the demand for keeping geographic data updated. Automatic generated land cover maps could be included to achieve current data.

A common issue with developing an effective classification process is that the spectral value of a class is dependent on acquisition conditions, local climate conditions, land cover in the area and building materials used in the area, so it is not possible to transfer from region to region.

This study has tried to compare some of the current available options for this, with a focus on how an object-based approach could increase the accuracy of these maps. The focus on objects instead of pixels did yield better results. However, there is still room for improvements. This study also revealed parts of the area, where the segmentation did not improve results but instead led to misclassification, mainly in areas connected to the water class.

A general weak point for the results was the ability to identify the built-up areas, as a big part of the study area consists of areas like seen in Figure 44. Houses typically have the size of 1.5-4 pixels so segmentation would not help much to tackle this problem as outlined in (Blaschke, 2009).

The classification with segmentation did perform better with regards to built-up and grasslands though, but possibly more in areas with larger buildings and roads.

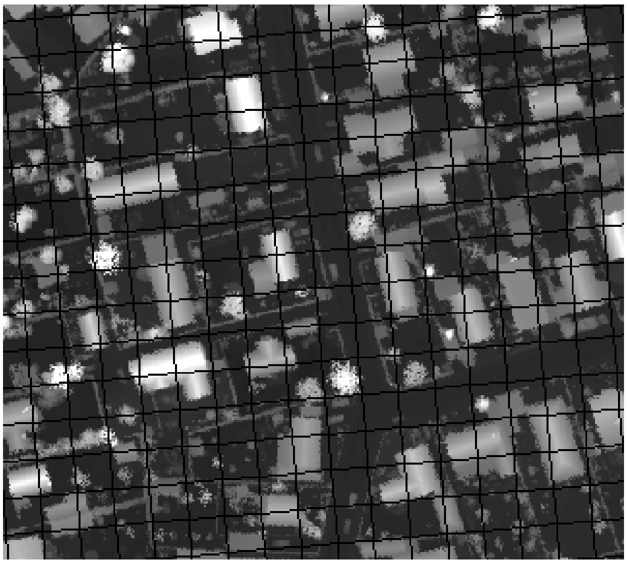


Figure 44 Digital surface model with grid representing the satellite image pixel size

The correlation between built-up and grassland that have been illustrated both in the accuracy scores and in the change maps, could partly be a result of the way the training points for the classifications were selected. Because the vector data did not distinguish between houses and gardens in the suburbs, there could be errors in the training data not detected by the visual inspection.

A method for improving the accuracy of areas where segmentation does not have an effect, could be the inclusion of LIDAR data. This study showed that LIDAR has potential when tested in a small area. It would be interesting to examine if the same improvement could be reached generally. Though to be done effectively, it would require better hardware and software to be able to handle the large amount of data. An arguable element of this DSM approach that only uses the height data, is that over large areas the ground level can differ some. This would muddle the effectiveness to detect road from buildings and low from high vegetation. A possible way around this is to subtract the digital terrain model from the surface model in order to get a net height difference instead. Other uses of LIDAR data could be included as well (Chen, et al., 2018). In this study only the surface height data was used, but a study that also utilizes the intensity might perform even better. LIDAR data is not affected by shadows either, so it could be a way of eliminating some of the errors that could happen in satellite images.

The major trend of more grassland in the change map, could also be affected by the fact that there was more precipitation in the year 2017, leading to more green areas and higher spectral values. Other shifts in classes like the change from built-up to forest areas could be local governments initiatives starting to show, as there is an increasing focus in planning on green areas (Astrup, 2017) and also because of climate changes that contributes to rougher weather, there is a need for more areas to be able to absorb water from the heavy rainfalls experienced in the later years.

The change detection method of measuring the difference in pixel values between the images, can be used to monitor spots with a lot of development, but this method is very limited put up against comparing classified images, as this change detection method does not allow searching for only one kind of change. Another benefit of the method with classified images is that it also will provide other information, like for example when monitoring the progress of a big infrastructure project with drones, it will be possible to identify the number of workers as well and maybe even identify if they are wearing the right protection gear.

A future use of high resolution images is to identify vertical land movement; this can then be used together with land cover maps in order to ensure that urban development, big infrastructure projects and also climate change detection is done on the basis of as much knowledge as possible (Agency for Data Supply and Efficiency, 2017).

5. Conclusion

The comparison of pixel and object based approaches showed that the accuracy of the classifications was improved with 4-4.5 percentage points. It also confirmed the major trend in the research field that the random forest algorithm performed best, with an accuracy score of 85.6 %. The study did reveal parts of the study area, where the segmentation did not improve results and instead led to misclassification, mainly in areas connected to the water class.

The classes that benefited the most from the segmentation was the built-up and grassland with improvements between 7-13 percentage points accuracy, while the forest and water classes experienced a small drop in accuracy with a loss of 1 percentage point.

The introduction of clouds on the 2017 image, reduced the accuracy for the support vector machine with 14 percentage points and especially the built-up class was misclassified. The random forest handled the clouds better and was inside 0,5 percentage points of both the 2015 classification and the classification of the 2017 image with the clouds masked out.

The inclusion of LIDAR data on a smaller study area did lead to improvements in accuracy for part of the area and the total accuracy improved with 3 percentage points. The rest of the area remained the same, with no sign that the added digital surface model would lead to misclassification by itself.

The change analysis revealed that the largest shift between classes was from built-up to grassland and the general trend in the time period was an increase in grassland. Grassland was the only class showing a net gain on 4000 square kilometers, the rest of the classes had a loss in the net change analysis. The increase area in grassland was mainly located in the Copenhagen suburbs in the northwestern part of the map.

The map that showed persistence in the period illustrated several areas with almost no change, like the city center.

References

Agency for Data Supply and Efficiency, 2015. [Online]
Available at: https://kortforsyningen.dk/indhold/data
[Accessed 2018].

Agency for Data Supply and Efficiency, 2017. [Online]
Available at: https://download.kortforsyningen.dk/content/kort10-vektor

Agency for Data Supply and Efficiency, 2017. [Online]
Available at: http://sdfe.dk/data-skaber-vaerdi/nyheder/nyhedsarkiv/2017/aug/nye-satellitdata-kan-fortaelle-hvor-danmark-synker/
[Accessed 2018].

al Khudairy, D., Caravaggi, I. & Glada, S., 2005. Structural damage assessments from Ikonos data using change detection, object-oriented segmentation, and classification techniques. *Photogrammetric Engineering & Remote Sensing 71*.

Anon., 2017. *Cohen's Kappa Statistic for Measuring Agreement.* [Online]
Available at: https://onlinecourses.science.psu.edu/stat509/node/162

Astrup, E., 2017. *https://politiken.dk.* [Online]
Available at: https://politiken.dk/indland/kobenhavn/art5851034/Hvor-skal-vi-plante-100.000-nye-tr%C3%A6er
[Accessed 2018].

Belgiu, M. & Dra˘gut, L., 2016. Random forest in remote sensing: A review of applications and future. *ISPRS Journal of Photogrammetry and Remote Sensing.*

Blaschke, T., 2009. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*.

Blaschke, T., 2013. Geographic Object-Based Image Analysis – Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*.

Blaschke, T., Burnett, C. & Pekkarinen, A., 2004. New contextual approaches using image segmentation for object-based classification. *Remote Sensing Image Analysis: Including the spatial domain.*, pp. Kluver Academic Publishers, Dordrecht 211–236..

Breiman, L., 1996. Bagging predictors. *Machine learning 24(2)*, pp. 123-140.

Breiman, L., 2001. Random forests. *Machine learning*, pp. 5-32.

Burges, C., 1998. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery 2 (2)*, p. 121–167..

Chavez, P. S., 1996. Image-based atmospheric corrections-revisited and improved. *Photogrammetric engineering and remote sensing 62(9)*, pp. 1025-1035.

Chen, D. & Stow, D., 2002. The effect of training strategies on supervised classification at different spatial resolutions. *Photogrammetric Engineering & Remote Sensing*, 68.

Chen, G., Weng, Q., Hay, G. J. & He, Y., 2018. Geographic object-based image analysis (GEOBIA): emerging trends and future opportunities. *GIScience & Remote Sensing*.

Congalton, R. G. & Plourde, L., 2002. Quality assurance and accuracy assessment of information derived from remotely sensed data. In: *Manual of Geospatial Science and Technology.* London: Taylor & Francis, pp. 349-361.

Congedo, L., 2016. *Semi-Automatic Classification Plugin Documentation,* s.l.: s.n.

Cutler, D. R. et al., 2007. Random forests for classification in ecology. *Ecology 88(11)*, pp. 2783-2792.

Ehlers, M., Gähler, M. & Janowsky, R., 2003. Automated analysis of ultra highresolution remote sensing data for biotope type mapping: New possibilities and challenges. *ISPRS Journal of Photogrammetry and Remote Sensing 57 (5-6),*.

Eurepean Space Agency, 2018. *ESA.int.* [Online]
Available at: https://sentinel.esa.int/web/sentinel/missions
[Accessed 2018].

Foody, G. M., 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, April, pp. 185-201.

Foody, G. M., 2004. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering and Remote*.

Gislason, P. O., Benediktsson, J. A. & Sveinsson, J. R., 2006. Random forests for land cover classification. *Pattern Recognition Letters 27(4)*, pp. 294-300.

Hastie, T., Tibshirani, R. & Friedman, J., 2009. *The Elements of Statistical Learning.* New York: Springer.

Hay, G. & Castilla, G., 2008. Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. *Springer, Heidelberg, Berlin, New York*, pp. 93-112.

Hay, G., Niemann, K. & McLean, G., 1996. An Object-Specific Image-Texture Analysis of H-Resolution Forest Imagery. *Remote Sensing of Environment 55*, p. 108–122.

Hecheltjen, A., Thonfeld, F. & Menz, G., 2014. Recent Advances in Remote Sensing Change Detection – A Review. In: *Land Use and Land Cover Mapping in Europe.* New York: Springer.

Haas, J., Ban & Yifang, 2018. Urban Land Cover and Ecosystem Service Changes based on Sentinel-2A MSI and Landsat TM Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

INSPIRE, 2013. *INSPIRE.* [Online]
Available at:

http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_LC_v3.0.pdf
[Accessed 2018].

Jensen, J. R. et al., 1987. Inland wetland change detection using aircraft MSS data. *Photogrammetric Engineering & Remote Sensing*.

Kennedy, et al., 2009. Remote sensing change detection tools for natural resource managers: understanding concepts and tradeoffs in the design of landscapemonitoring projects. *Remote Sensing of Environment*.

Kotsiantis, S. B., 2007. Supervised machine learning: A review of classification techniques. *Informatica*, pp. 249-268.

Lei, M. & Manchun, L., 2017. A review of supervised object-based land-cover image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, june.

Liaw, A. & Wiener, M., 2002. Classification and regression by randomForest. *R news, 2(3)*, pp. 18-22.

LU, D. & Weng, Q., 2007. A Survey of image classifaction methods and techniques for improving classification performance. *International Journal of Remote Sensing*.

Melgani, F. & Bruzzone, L., 2004. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on geoscience and remote sensing 42(8)*, pp. 1778-1790.

Mountrakis, G., Im, J. & Ogole, C., 2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing 66(3)*, pp. 247-259.

Nasa, 2017. [Online]
Available at: https://earthobservatory.nasa.gov/Features/LandCover/

Powell, et al., 2004. Sources of error in accuracy assessment of thematic land-cover maps in the Brazilian Amazon. *Remote Sensing of Environment*.

Rosca, V. M., Takacs, M. T., Vepstas, A. & Trock, C., 2017. *Sentinel Imagery Classification,* Copenhagen: AAU Cph.

Shrestha, S., Smith, C. & Bochenek, Z., 2017. Unsupervised Land Use/Land Cover Change Detection with Integrated Pixel and Object Based Approches for High Resolution Remote Sensing Imagery. *Earth Observation for Land and Emergency Monitering.*

Singh, A., 1989. Review Article Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*.

Suthaharan, S., 2016. *Machine Learning Models and Algorithms for Big Data Classifi cation.* s.l.:Springer.

Tempfli, K., Kerle, N., Huurneman, G. C. & Janssen, L. L., 2009. *Principles of remote sensing. ,* Enschede, Netherlands: The International Institute for Geo-Information Science and Earth Observation (ITC).

The Danish Agrifish Agency, 2017. [Online]
Available at: http://lbst.dk/landbrug/kort-og-markblokke/oversigt-over-al-kortdata-fra-landbrugsstyrelsen/#c30679
[Accessed 2017].

Vapnik, V., 1982. *Estimation of Dependences Based on Empirical Data (English translation).* New York: Springer.

Wilkinson, G. G., 2005. Results and implications of a study of fifteen years of satellite image classification experiments. *IEEE Transactions on Geoscience and remote sensing 43(3)*, pp. 433-440.

Xian, G., Homer, C. & Fry, J., 2009. Updating the 2001 national land cover database land cover. *Remote Sensing of Environment*.

Xiang-Yang, W., Ting, W. & Juan, B., 2010. Color image segmentation using pixel wise support vector machine classification. *Elsevier*.

Zhu, Z. & Woodcock, C. E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*.