

Evaluation of pixel based and object based classification methods for land cover mapping with high spatial resolution satellite imagery, in the Amazonas, Brazil.

Linnea Jonsson

2015

Department of

Physical Geography and Ecosystems Science

Lund University

Sölvegatan 12



Linnea Jonsson (2015). Evaluation of pixel based and object based classification methods for land cover mapping with high spatial resolution satellite imagery, in the Amazonas, Brazil.

Master degree thesis, 30 credits in Physical Geography and Ecosystem Analysis

Department of Physical Geography and Ecosystems Science, Lund University

Level: Master of Science (MSc)

Course duration: September 2014 until January 2015

Disclaimer

This document describes work undertaken as part of a program of study at the University of Lund. All views and opinions expressed herein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Evaluation of pixel based and object based classification
methods for land cover mapping with high spatial
resolution satellite imagery, in the Amazonas, Brazil.

Linnea Jonsson

Master thesis, 30 credits,
in Physical Geography and Ecosystem Analysis

Supervisor 1: Lars Eklundh
Department of physical geography and ecosystem sciences,
Lund University

Supervisor 2: Tobias Edman
Metria AB Stockholm

Exam committee:
David Tenenbaum, Department of Physical Geography
and Ecosystem Sciences, Lund University

Harry Lankreijer, Department of Physical Geography
and Ecosystem Sciences, Lund University

Acknowledgements

Firstly, I would like to thank my supervisor Lars Eklundh, for the support during the process of writing this thesis. Also, a big thanks to Tobias Edman at Metria AB, Stockholm, who gave me this opportunity to take part of the project. I would also like to thank: Julian for his support in Jönköping, Lina for a lot of things, Caroline and Daniel because they are just wonderful, and last but not least, my two sweet cats.

Abstract

In the state of Acre, Brazil, there is ongoing land use change, where inhabitants of this part of the Amazonian rainforest practice shifting agriculture. Practicing this type of agriculture is, according to the SKY Rainforest Rescue organization, damaging to forest ecosystems. This organization aims to educate people in how to maintain sustainable agriculture. By monitoring this shift in agricultural practices with the use of remotely sensed data, the organization can follow the development. In this thesis, an image with high spatial resolution, from the SPOT-5 satellite, is used to evaluate which classification method is most appropriate for monitoring land use change in this specific area. Three methods are tested; two pixels based and one object based. The pixel based methods are the Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel and the Maximum Likelihood Classifier (MLC), and the object based method is segmented with Multi Resolution Segmentation (MRS) and classified with the k-Nearest Neighbor (kNN). The parameters gamma and penalty parameter C in the SVM with an RBF kernel were estimated by a k-fold cross validation and grid search method; and for the MLC, an assumption that each class had an equal probability distribution was made. For the object based approach the first step was segmentation; for the MRS there are three parameters: scale, shape and compactness. The scale parameter was set by using an algorithm that was based on comparing local variance; shape and compactness were defined based on previous studies and visual evaluation of the segments. All three methods will produce two classified maps each; one where the feature space consists of the three original wavebands (green, red and NIR) and one where the feature space is of six dimensions that include the original three wavebands and three texture derivations, one from each original band. The texture is derived from the co-occurrence GLCM method, which can be used to calculate 14 different texture measures. The three most suitable texture derivations were the contrast texture measure from the green and NIR band, and an entropy texture, derived from the red band. When combining these three texture derivations with the original bands, the classes were further separated. The original image was also lowered in resolution, from 2.5m to 25m in pixel size. However, this did not generate neither higher nor similar overall accuracy compared to any of the high spatial resolution classified images. The moderate spatial resolution classifications were only computed with the MLC and SVM due to the inefficiency of an object based image analysis method when used with moderate spatial resolution. Of these six classifications, only two exceeded the 85% threshold of an acceptable overall accuracy. These were the SVM (86.8%) and kNN (86.2%), which included the texture analysis. None of those classifications with only the three original bands exceeded this threshold. In conclusion, the object based method is the most suitable approach for this dataset because: 1) the parameter optimization is less subjective, 2) computational time is relatively lower, 3) the classes in the image are more cohesive and 4) there is less need for post-classification filtering.

Keywords

Classification, SPOT-5, Maximum likelihood, Support vector machine, RBF-kernel, OBIA, Multiresolution segmentation,

Abbreviations

AVIRIS – Airborn Visible/Infrared Imaging Spectrometer

GLCM – Grey level co-occurrence matrix

HR – High Resolution

HRG – High Resolution Geometrical

HRV – High Resolution Visible

HSR – High Spatial Resolution

JM – Jeffrey Matsuita

kNN – k-Nearest Neighbor

LV – Local variance

MLC – Maximum Likelihood Classifier

MSR – Moderate Spatial Resolution

MRS – Multiresolution Segmentation

NIR – Near Infrared

OBIA – Object Based Image Analysis

OB – Object Based

OB-kNN – Object Based k-Nearest Neighbor

ROC – Rate of change

RBF – Radial Basis Function

SVM – Support Vector Machine

SPOT – Satellite Pour l'Observation de la Terre

TM – Thematic Mapper

UV – Ultraviolet

Table of content

1. INTRODUCTION	1
2. BACKGROUND.....	3
2.1 RADIATION	3
2.2 SATELLITE REMOTE SENSING	5
2.3 IMAGE CLASSIFICATION	7
2.3.1 <i>Support vector machine</i>	8
2.3.2 <i>Maximum likelihood classification</i>	12
2.3.3 <i>Object based classification</i>	13
2.4 TEXTURE ANALYSIS	15
2.5 VALIDATION.....	17
3. METHOD.....	18
3.1 DATA AND SOFTWARE.....	18
3.2 TRAINING AND EVALUATION SAMPLES	18
3.3 TEXTURE ANALYSIS	20
3.4 MAXIMUM LIKELIHOOD CLASSIFICATION	22
3.5 SUPPORT VECTOR MACHINE CLASSIFICATION.....	23
3.5.1 <i>Calibrating the SVM algorithm, C & gamma</i>	23
3.6 OBJECT BASED CLASSIFICATION	24
3.6.1 <i>Scale parameter estimation</i>	24
3.6.2 <i>Shape and compactness parameters</i>	25
3.6.3 <i>k-Nearest Neighbor classification</i>	26
3.7 MODERATE SPATIAL RESOLUTION	26
3.8 VALIDATION.....	26
4. RESULTS.....	27
4.1 CLASSIFICATION RESULTS	27
4.1.1 <i>Subsets of classifications</i>	27
4.1.2 <i>The area percentage of each class</i>	30
4.1.3 <i>Overall accuracy</i>	31
4.1.4 <i>Z-test between error matrices</i>	35
5. DISCUSSION	36
5.1 ACCURACY ASSESSMENTS	36
5.1.1 <i>High Spatial Resolution (HSR)</i>	36
5.1.2 <i>Moderate Spatial Resolution (MSR)</i>	37

5.1.3 <i>Misclassified samples</i>	37
5.2 AREA PERCENTAGE OF EACH CLASS AND VISUAL INTERPRETATION.....	38
5.2.1 <i>High Spatial Resolution (HSR)</i>	38
5.2.2 <i>Moderate Spatial Resolution (MSR)</i>	38
5.3 TEXTURE ANALYSIS	39
5.4 ESTIMATION OF SCALE	39
5.5 SVM PARAMETER ESTIMATION.....	40
5.6 COMPUTATIONAL TIME	40
5.7 GOOGLE EARTH AS GROUND TRUTH.....	40
5.8 25M RESOLUTION (MSR)	41
6. CONCLUSION	41
7. REFERENCES.....	43
8. APPENDIX	46
8.1 ERROR MATRIX.....	46
<i>High Spatial Resolution, 2.5m</i>	46
<i>Moderate Spatial Resolution, 25m</i>	49
8.2 MAPS, FULL EXTENT	50
8.3 LOCATION	58

1. Introduction

The earth's surface is constantly changing as a result of direct anthropogenic actions, e.g. land use changes or indirect impacts by a changing climate. But natural processes, such as weathering and tectonic movements also affect the land surface. In comparison to anthropogenic driven changes, however, these processes are usually slower. By gathering information from a distance, i.e. remote sensing, changes can be monitored and mapped (Lillesand et al. 2008).

In the rainforests of the Amazonas, Brazil, there is ongoing land cover change, mainly through deforestation. The forest is cleared for, amongst other reasons, unsustainable agriculture such as shifting cultivation. This method of cultivating is, according to the Sky Rainforest Rescue, harmful to the ecosystems and is in need of observation to prevent further expansion (SKY Rainforest Rescue, 2015). Land cover changes can be monitored by remotely sensed images, which can detect clear cuts and other changes in the landscape.

Common ways to retrieve remotely sensed images are aerial photography with generally high spatial resolution, and sensors on satellites with generally lower spatial resolution. However, with new technology, satellite remote sensing has become more advanced with higher spatial, temporal and spectral resolutions, i.e. high resolution (HR) images (Lillesand et al. 2008).

Since the 1960s, remotely sensed data has been useful for mapping and detecting changes, both on global and local scales. Factors such as the spatial, spectral and temporal resolution of imagery are usually crucial when mapping, e.g. the higher the spatial resolution is, the more objects can be detected and thus, smaller changes can be visualized. The spectral resolution determines the number of bands and the width of each, while temporal resolution defines how often the satellite revisits the same location on earth (Liang, Li and Wang 2012). However, this shift in technology generates a problem of how to handle these highly detailed images (Cartwright et al. 2008).

A common usage of remotely sensed images is for producing thematic maps of land cover and land cover changes. These types of maps are produced by using spatially referenced training areas, that represent various land cover types i.e. class signatures such as forest, water or urban, to recognize a pattern in an image. In the process of recognition, an algorithm assigns each pixel in the image to one of the spatially referenced land cover types depending on various constraints (Tso and Mather 2009).

However, the shift in technology, with satellite images having higher spatial resolution, may lead to necessary changes when processing high spatial resolution remotely sensed data. The SPOT-5 sensor can now offer a 2.5m spatial resolution, and classifiers that rely on some common assumptions of parametric statistics, such as a normal distribution of reflectance, may not be sufficient.

There are different kinds of classification algorithms; some rely on the assumption that the class signatures would be normally distributed, such as the widely used maximum likelihood classifier (MLC), that assigns each pixel to a class depending on the probability of it belonging to that class (Tso and Mather 2009). However, this assumption of a normal distribution can be limiting and may not give a satisfying result in some cases (Huang, Davis and Townshend 2002). Also, there are classifiers that are based on a non-parametric approach, i.e. no assumption of any probability distribution, such as the support vector machine (SVM) classifier, which strives to find the optimal separating boundaries between pixels and classes (Tso and Mather 2009).

Many studies have shown that SVM performs better than the commonly used MLC (Huang et al. 2002). However, pixel based methods, such as MLC and SVM, have started to be questioned, and in particular the MLC method (Rittl et al. 2013, Oommen et al. 2008). The disadvantages of MLC are the strong assumption of a normal distribution in all wavebands, and that it needs a relatively high number of training samples for each class (Oommen et al. 2008).

The high resolution data results in images with pixels being considerably smaller in size than the average size of distinguishable objects. A considerable proportion of the reflected energy can actually be representing a pixel's neighbor. This is the mixed pixel problem and is an issue at all pixel sizes. According to Blaschke and Stroble (2001), merging pixels into segments could reduce this problem. Therefore, another method to classify HR images is to aggregate the pixels into segments that represent meaningful objects (Lang 2008). The size of the generated segments is of high importance, and can be determined by calculating the mean of the local variance between objects and their neighbors (Dragut, Tiede and Levick 2010). This relatively new method of classification is named object based image analysis (OBIA), and has proven to be superior to pixel based methods in some studies (Rahman & Saha 2007).

The area of interest is located in the state of Acre (See Appendix 8.3), northwestern Brazil, which is covered by rainforest with a meandering river system and is a location where humans are practicing shifting agriculture, which is in need of monitoring. To produce a valid map of the changes in land use, an appropriate classification method has to be defined; this method is usually dependent on the image properties such as spatial and spectral resolution.

The aim of this thesis is to both visually and statistically evaluate three different classifiers, two pixel based (MLC and SVM) and one object based k-Nearest Neighbor (OB-kNN), and their appropriateness for analyzing high spatial resolution data in the area of interest. The commonly used maximum likelihood is said to perform poorly in many studies, in comparison to the SVM and Ob-kNN (Huang et al. 2002, Oommen et al. 2008). According to Ouma et al. (2007) & Johansen et al. (2007); including texture layers should increase classification accuracy and therefore such an approach will be included. The data used will be SPOT 2.5 m images, which are considered to be of high spatial resolution. Also, the need for such high resolution images will be discussed. The image will be classified into four classes: Forest, Open Field, Scrubland and Water.

2. Background

2.1 Radiation

All objects that have a temperature above absolute zero, 0 Kelvin (-273°C), radiate electromagnetic energy and depending on the object temperature the radiation differs (Lillesand et al. 2008). The electromagnetic spectrum ranges from low to high energy wavelengths, where radio and microwaves being low energy radiation, and gamma rays and cosmic rays having high energy, see Figure 1.

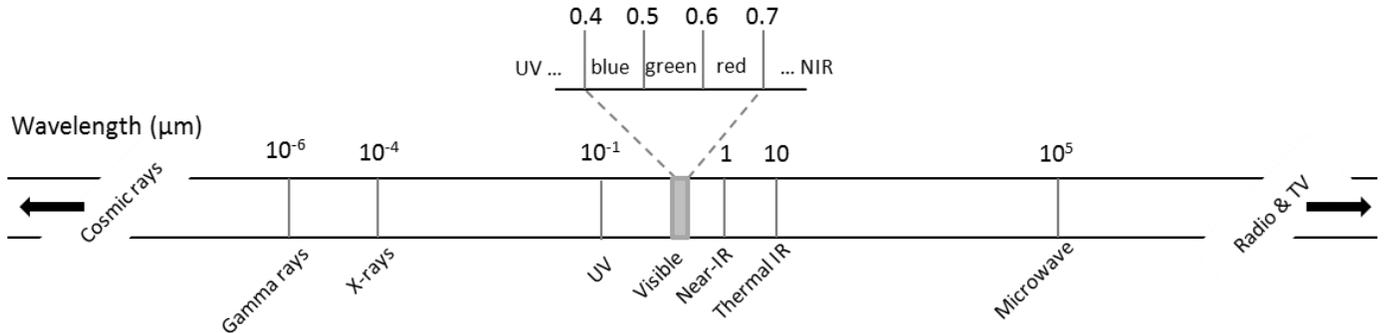


Figure 1. A portion of the electromagnetic spectrum. Image is derived from Lillesand, Kiefer and Chipman (2008).

When electromagnetic radiation passes through the atmosphere and reaching the surface, a photon can be absorbed, transmitted or reflected. Though, this is dependent on the characteristics of the particles and objects in its path, and the amount of energy in the photon itself. On the path from earth to sensor, reflected photons may be altered in intensity and direction due to absorption and scattering. This effects the whole spectrum, though some regions are more affected than other.

The absorption of energy is occurring at various wavelengths and in different levels of the atmosphere where, for example, wavelengths between 15 and $10^3 \mu\text{m}$ are totally absorbed in the atmosphere and cannot be used for remotely sensing surface properties; however, those regions that have low energy absorption are called atmosphere windows, which are suitable for remote sensing (Lillesand et al. 2008).

The wavelengths can be divided into wavebands where each waveband represents a discrete set of continuous wavelengths, such as the visible band from 0.4 to $0.7\mu\text{m}$ and infrared (IR) band from 0.7 to $10\mu\text{m}$. The visible band can be further subdivided into the blue, green and red bands; and the infrared can be divided into the near-IR, mid-IR and thermal-IR bands (Lillesand et al. 2008).

Visible light ranges from the colors blue to red; blue at wavelengths between 0.4 and $0.5\mu\text{m}$, green 0.5 and $0.6\mu\text{m}$, and red between 0.6 and $0.7\mu\text{m}$. Depending on object characteristics, e.g. pigment and surface roughness, different parts of the visible spectrum is absorbed, transmitted or reflected. For example, a green leaf absorbs both blue and red wavelengths but reflect the green, the reflected radiation represents the color detected by the human eye.

Depending on surface roughness, radiation behaves differently. There are in theory two surfaces that represent one perfectly smooth surface and one perfectly rough. A perfectly smooth surface will reflect radiation in the same angle as the incident angle; a perfectly rough surface will reflect radiation in all directions. However, roughness of objects usually lies somewhere between these two extremes (Tso and Mather 2009).

The non-visible spectrum of near-infrared (NIR), from 0.7 to $1.4\mu\text{m}$, makes vegetation easy to distinguish, due to almost 50% of the wavelets get reflected by the leaves. It can also be possible to separate vegetation types from each other, since the reflectance in IR spectrum also is dependent on the inner structure of the leaves and how thick the layer of leaves is (Lillesand et al. 2008).

2.2 Satellite remote sensing

Remote sensing satellites are used, among other wavelength intervals, in the optical spectrum, from 0.3 to 14 micrometers; which includes ultraviolet (UV), visible, near-, mid- and thermal-IR wavelengths. There are a significant number of satellites in space; however, not all of their imagery is suitable for image classification. Optical remote sensing systems such as Landsat, SPOT and NOAA AVHRR, have been widely used during the three previous decades for earth observation. These sensors are multispectral, which denotes that reflected radiation is registered in a small number of broad wavebands; in contrast, hyperspectral sensors collect reflected radiation in a large number of narrow wavebands e.g. AVIRIS (Tso and Mather 2009).

A sensor receives the reflected energy as photons, and collects the signal in the form of a digital number in each waveband. The digital numbers are stored in a raster format where each element is a pixel in a two-dimensional matrix. Depending on the storage space of each pixel, the digital number can range within an interval of varying size, from e.g. 0 to 255, which would be 8-bit precision, or 0 to 1023, 10-bit precision. Both Landsat TM and SPOT HRV have 8-bit precision, but satellites such as NOAA AVHRR, use 10-bits to store more detailed data. The digital numbers in the pixels are a label or a quantized count, rather than a physical value; where, for some applications the number has to be converted into a physical value (Tso and Mather 2009).

A multispectral image has four descriptive resolutions; spatial, spectral, radiometric and temporal resolution. Spatial resolution i.e. pixel size, describes the reflected radiance from a specific area. This area represents a location on the ground that has a particular spectral value; it defines what can be seen in the image. Depending on the number of bits representing a pixel, the radiometric resolution differs, it can be described as the number of tones an image can have. Radiometric resolution determines ability to distinguish features from each other in one waveband. Spectral resolution represents the energy interval in the electromagnetic spectrum that a waveband is collected in. It defines the width of the waveband; hyperspectral images have spectrally narrow bands, compared to multispectral that collects wider bands.

The temporal resolution describes how often the satellite revisits for image data collection at the same location; this depends on type of orbit and sensor characteristics. Temporal resolution often varies with spatial resolution; high spatial resolution usually is associated with low temporal resolution and vice versa. An exception, though, is in the case of geostationary satellites (Shekhar, Shekhar and Xiong 2008)

Most earth observation satellites orbit the earth in an elliptical orbit in an altitude of at least 400 km; the time to complete one orbit depends on this altitude. Satellites usually have different orbiting routes, which are often dependent on the inclination angle relative

to the equator. If the angle is near 90° , the satellite would be defined as polar-orbiting. There is also a type of orbit called sun-synchronous; such a satellite moves westward at a certain angle and always collecting images in daylight. A geostationary satellite has the same orbital period as the earth and thus, it produces images of the same spot at all times. A geostationary satellite orbits at a much higher altitude, approximately 36000km; this compensates for the earth's rotation (Tso and Mather 2009). An example of the geostationary satellite is the Geostationary Operational Environmental Satellite, GOES, which has five bands including water vapor. It is mostly used for meteorological measures; and has a rather coarse spatial resolution of 1km.

The U.S. operated Landsat (Landsat handbook, 2015) and the French SPOT (SPOT, 2015) satellites are some of the oldest programs in remote sensing; which are considered to be very robust monitoring systems; many new systems make use of their principles. The first Landsat was launched in 1972 by NASA and the program has the longest record of earth observation data (Landsat handbook, 2015). The latest Landsat was launched in 1999, the Landsat 7 ETM+ sensor, which has seven bands ranging from visible to thermal-IR with a 30m spatial resolution and a 15m panchromatic band (black and white image). The swath width is 185km and is able to create full views of the earth approximately four times each year (Lillesand et al. 2008).

The French satellite program SPOT was first launched their first satellite in 1986; and in 2002, they launched their latest, SPOT-5, which includes three sensors, one for imaging vegetation, HRS (High Resolution Stereoscopic) and HRG (High Resolution Geometric). The HRG sensor has three bands with 10m spatial resolution, ranging from green to near-IR with an additional panchromatic band with 2.5m resolution. The swath width is 60-80 km and it is a polar orbiting sun-synchronous satellite.

The images can be used in various fields, such as oceanography, meteorology and terrestrial studies. The image data can be applied in environmental modeling, monitoring and land cover mapping. Though, in pattern recognition and land cover mapping for one certain occasion, the digital numbers does not need any conversion due to the interest of inter-pixel relations (Tso and Mather 2009).

In this study, Systeme Pour l'Observation de la Terre, SPOT, is used, which is an earth observation satellite program started by the French government in 1978, but shortly both Sweden and Belgium joined the program. The satellite is designed, and created, by the French Centre National d'Etudes Spatiales, CNES; which has now progressed into a large-scale international program with ground receiving stations in several countries. The SPOT-5 satellite includes high resolution geometric (HRG) instruments (Lillesand et al. 2008).

2.3 Image classification

Multispectral images have a number of wavebands that contains pixels with values representing the reflectance of features in the landscape. Due to features having varying reflectance behavior, the potential of identifying features for image classification increases. The wavebands together create a feature space with the same dimensionality as the number of bands. For example, the reflectance characteristics of vegetation, bare soil and water differ over the electromagnetic spectrum; water absorbs almost all radiation above 0.8 μm , bare soil a rather steady increase with longer wavelengths, and the vegetation curve varies over the wavelengths with a relatively high reflectance in the IR bands (Tso and Mather 2009).

There are two main approaches to classify an image; either by a supervised or unsupervised classification. Unsupervised classification uses a clustering algorithm, which aggregates pixels close to each other in the feature space into classes. These classes are spectral classes that are based on the clustering of the spectral values. The spectral classes can now be assigned to informational classes. In contrast, supervised classification begins with defining informational classes. By collecting training samples representing the classes, one can form a numerical description of the spectral attributes for each informational class, derived from the feature space. With these descriptions, each pixel can be assigned to the class to which it is spectrally closest (Lillesand et al. 2008). The accuracy of unsupervised classifications is usually lower than for supervised (Tso and Mather 2009).

In supervised image classification, an algorithm separates pixels from each other depending on training samples that represents a location on the ground. Training samples are usually collected in field, derived from aerial photography or other sources with high spatial resolution, and work as an interpretation key that defines the spectral attributes for each class. These samples represent the real world, and are used to train the algorithm to separate pixels in the feature space (Lillesand et al. 2008).

The feature space used in the classification can be expanded beyond the original wavebands; information such as elevation or texture can enhance separation of pixels. It is not possible to derive elevation from spectral bands, but texture can be generated. A texture analysis can enhance the separation in feature space between e.g. vegetation types and between urban and vegetation classes. However, when the dimensionality of the feature space increase, more training samples are needed. For example, this is the case for hyperspectral images, where the number of training samples has to be very high. This could lead to difficulties when estimating statistical parameters, such as, mean and covariance (Tso and Mather 2009).

An issue in image classification is the mixed-pixel problem; this means that the registered reflection in a pixel may be influenced by neighboring pixels. The influence from neighbors is often due to atmospheric effects, such as scattering, but can also be due to sensor properties.

Early classification methods usually based on conventional statistical techniques, such as the maximum likelihood classifier; later in the mid-90s, new approaches with alternative strategies, such as the artificial neural network, decision tree and support vector machine, became available (Tso and Mather 2009).

Recently, sensors started to acquire high resolution images. A pixel in these high resolution images covers a much smaller area on the surface. This generates new possibilities and problems in classification. One pixel is no longer a generalization of features, but a part of one. In some cases this can be seen as an overload of information and new methods have to be used, such as object-based methods of classification (Cartwright et al. 2008).

2.3.1 Support vector machine

The Support Vector Machine (SVM) is a group of non-parametric machine learning algorithms used for e.g. pattern recognition, handwriting recognition etc. These algorithms first emerged in the 1970s, but did not receive great attention in remote sensing until the 1990s (Burges 1998).

One of the appealing characteristics of the SVM is the risk of error minimization. This is unlike other classifiers, such as MLC, which deals with errors in an empirical way; that is, the errors are directly determined by the distribution of training samples. The SVM classifier is instead minimizing the risk of misclassification by minimizing the probability of misclassifying a randomly chosen and before unseen point, from a fixed but unknown probability distribution (Tso and Mather 2009). In the simple case it is a linear binary classifier, meaning that only two classes are separated by one single separating line, see Figure 2, where the margin between the support vectors are maximized.

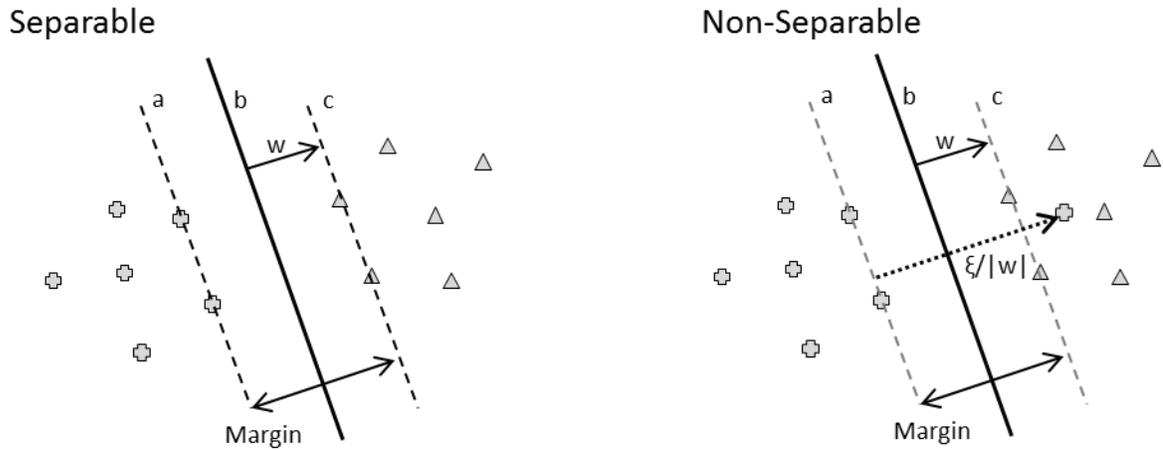


Figure 2. The separable and non-separable case of SVM where a and c are the support vectors, b the optimal hyperplane and w the normal to the hyperplane. In the non-separable case there is also a slack variable ξ (Tso and Mather 2009).

In the separable case, assuming only two classes, where the pair (x_i, y_i) , with $x_i \in \mathbb{R}^d$ (\mathbb{R} is a d-dimensional space) and y_i being either values of 1 or -1. Then the optimal hyperplane would be:

$$w^T x + b = 0 \quad (1)$$

where w is the normal to the hyperplane, x is a point lying on the hyperplane and b is the bias to the origin. Due to the binary case, one can create the constraint of:

$$y_i(w^T x_i + b) - 1 \geq 0 \quad (2)$$

The margin between the two support vectors is $2/\|w\|$, which is to be maximized; and this leads to:

$$\min \left(\frac{\|w\|^2}{2} \right) \quad (3)$$

and by introducing a primal Lagrangian function on Equation 2, it will be easier to handle. See (Tso and Mather 2009) for further derivation. This will generate a hard margin formulation due to the fact that no training errors are allowed:

$$f(x) = \text{sign} \left(\sum_{i=1}^{nsv} \alpha_i y_i (x \times x_i) + b \right) \quad (4)$$

Where nsv denotes the number of support vectors and α_i is a positive Lagrangian multiplier.

However, remotely sensed data can seldom be separated by just a linear boundary, equation 4, nor is usually, only two classes the case. For the non-separable case an additional slack variable ξ is used to represent this situation, where equation 2 becomes:

$$y_i (w^T \times x_i + b) \geq 1 + \xi, \quad \forall i \quad (5)$$

with the optimal hyperplane being:

$$\min \left(\frac{[w]^2}{2} + C \sum_{i=1}^n \xi_i \right) \quad (6)$$

where a penalty parameter C is incorporated to the equation; the hyperplane optimization consists of a combination of margin maximization and error minimization. Parameter C is user-defined, and is penalizing the distance between the wrongly classified point and its support vector. This criterion of error minimization tries to avoid samples being classified on the wrong side of the hyperplane.

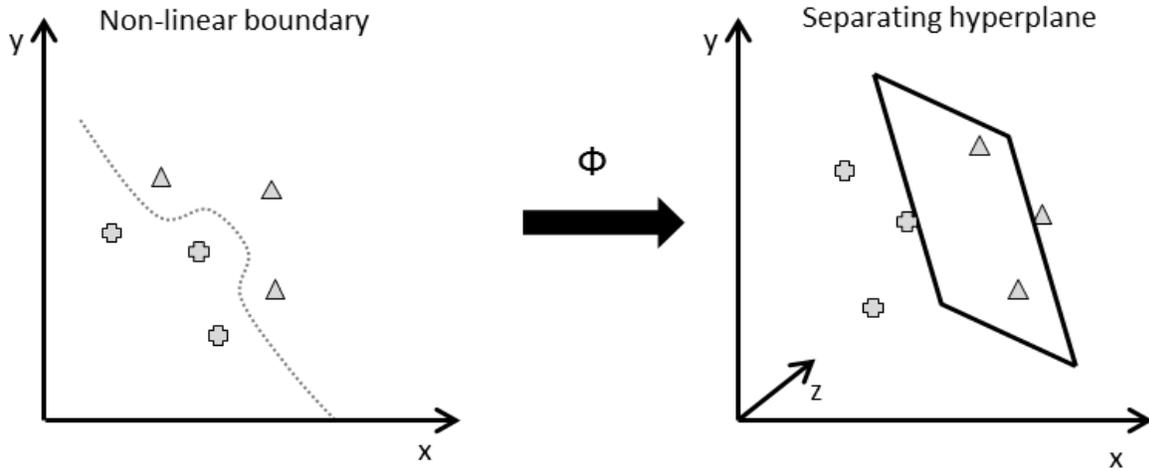


Figure 3. A non-linear case projected into a higher dimension feature space to increase separability (Tso and Mather 2009).

In the case of a non-linear boundary, one can project the data $x \in \mathbb{R}^d$ into a higher dimensional feature space, i.e. a generalization of the Euclidian space called Hilbert space. This will increase separability by spreading the training samples further away from each other, which would make the fitting of a linear hyperplane easier, Figure 3. By projecting into the higher dimensional feature space H , via a nonlinear vector mapping function $\Phi: \mathbb{R}^d \rightarrow H$ and deriving Equation 4 further to:

$$f(x) = \text{sign} \left(\sum_{i=1}^{nsv} \alpha_i y_i \Phi(x_i) \Phi(x) + b \right) \quad (7)$$

In a high dimensional feature space, one can fit a linear hyper-plane that divides the classes. Though, there could be an endless number of hyperplanes, the algorithm always searches for the optimal one, which is the one that has the greatest margin; this is the distance, perpendicular to the support vector, which passes through the points closest to the plane for each class. The two support vectors and the hyperplane are parallel to each other. However, this scenario, where classes are totally separable, is not always the case.

However, when data is projected into higher dimensions, the computational burden would increase significantly. This problem can, according to (Cortes and Vapnik 1995),

be solved by applying a kernel function $K(x_i, x_j) = \Phi(x_i)\Phi(x_j)$, which would also distribute points further apart, but with a lower cost.

There are some different kernel functions, with a varying number of parameters that have to be set by the user. Some however, have more than two parameters, which can be very time consuming to estimate. The Radial Basis Function (RBF), Equation 8, has two user-defined parameters, gamma and the penalty parameter C, and is supposed to perform well in classifying remotely sensed data (Oommen et al. 2008, Hsu, Chang and Lin 2003, Kavzoglu and Colkesen 2009).

Radial Basis Function:

$$K(x_i, x_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right), \quad \gamma > 0 \quad (8)$$

When considering several classes, one can approach the problem in different ways. Either one can compare classes pairwise or one against all other classes. In the case of the latter one, the method is just like a binary classification.

2.3.2 Maximum likelihood classification

The maximum likelihood classifier, Equation 9, is one of the most used classifiers so far; it is a parametric classifier, which assumes that each class is normally distributed in all bands. Every pixel is assigned to a class, depending on a probability of belonging to a certain class. The algorithm is trained with the training samples, creating N number of classes and also N number of ellipsoid probability functions. Also, variance and covariance is evaluated. Not all pixels will fall perfectly within these probability functions, but merely in the periphery. One can avoid pixels being weakly assigned to a class with a very low probability by adding a probability threshold value. Though, if the pixel cannot meet this threshold it would remain unclassified (Richards and Jia 2006).

The equation for MLC is:

$$D = \ln(a_c) - [0.5 \ln(|Cov_c|)] - [0.5 (X - M_c)^T (Cov_c^{-1})(X - M_c)] \quad (9)$$

where D is the weighted distance, C the particular class, \mathbf{X} vector of spectral signature for the test pixel, M_C the mean vector of sample of class C (which is assumed equal probability), Cov_C the covariance matrix of training data belonging to class C , and T is a transposition function. The pixel will be assigned to that class, which has the greatest likelihood (D), C (Oommen et al. 2008).

2.3.3 Object based classification

Multi resolution segmentation

Multiresolution segmentation (MRS) is a bottom up method, where segmentation starts with one pixel and, by loops, merges pixels into pairs and into larger objects. The looping ends when a threshold for the homogeneity within the objects is reached. Starting with one pixel, the algorithm tests if the neighboring pixels can be merged with the starting pixel, and by checking with the homogeneity criteria, the algorithm evaluates if the merge is satisfactory. If not satisfactory, the neighboring pixels will act as new seeds, creating new objects. The MRS algorithm focuses on minimizing the average internal heterogeneity of objects and maximizes the homogeneity, and is according to Rahman and Saha (2008), a very effective segmentation algorithm.

The homogeneity criteria, where heterogeneity is minimized, is based on four sub-criteria; color, shape, smoothness and compactness, where color usually is the most important. Smoothness concerns the borders of the object, and compactness optimizes the objects in regard to the overall shape criterion. For running the algorithm there are three parameters to be set; scale, shape and compactness.

Scale is the parameter that determines the size of the objects, since it decides the maximum accepted heterogeneity within the objects. E.g. for a certain scale, if the data set would be more heterogeneous, the objects would be smaller in size in comparison to a more homogenous set of data. This parameter is a very important parameter in the segmentation process (Dragut et al. 2014, Dragut et al. 2010, Blaschke 2010).

The shape parameter is closely related to the color criteria. Shape can be set between 0.1 and 0.9, and the lower this parameter is the more influence the spectral information has in the segmentation process. However, if the shape would be 1, the segmentation would be optimized for spatial homogeneity with no respect to spectral homogeneity. However, in images, the spectral properties are the primary source of information, thus, a value close to 0.9 could reduce the quality of segmentation.

The shape parameter is actually defined by the third parameter, compactness, which is used for the event of objects having relatively close spectral values. The compactness is

defined as the ratio between the object perimeter and its total area. In this case objects can be differentiated from each other with this parameter (eCognition User Guide 2015).

Scale parameter estimation

Determining the scale parameter that is appropriate for an image can be rather difficult. With a, either too high or too low scale parameter, the segmentation can be quite useless as it will not represent any real objects. A method to estimate scale is to study the local variance (LV) of object heterogeneity in the image.

The LV can be considered as spatial autocorrelation within objects in the image; if several pixels are needed to visualize an object, most pixels will be correlated with their neighbors and therefore the LV will be low. On the other hand, if spatial resolution is coarse, the LV will be high due to the higher spectral difference between pixels (Dragut et al. 2010).

Dragut et al. (2010) developed a tool to implement this concept in eCognition, that is, for a user-defined scale interval and scale step size, segmenting the image through an interval. For each scale, the LV is calculated and exported to a text file. However, LV alone is not a sufficient indicator of scale, and is improved by adding a study of the change of LV over the range of scales, which enhance the detection of these breaks in the curve. Therefore, the rate of change (ROC) is calculated between each scale size step:

$$ROC = \left[\frac{L - (L-1)}{L-1} \right] * 100 \quad (10)$$

Where L is LV at target level and L-1 is representing the LV at the previous lower scale level. By studying the breaks in the ROC curve, one can outline scales that would create objects with characteristics of relatively equal values of homogeneity.

k-Nearest Neighbor classification

The k-nearest neighbor (kNN) is a rather simple, non-parametric and instance-based machine learning algorithm. This classification method uses training samples and assigns each object to its closest training sample in the feature space.

The objects created from segmentation, are classified by their neighbor's majority vote, and the number of neighbors voting is based on the user defined k value. The k value is a positive integer and represents how many of the object's closest neighbors, will contribute in the voting. Often it is a small number but it varies with different data sets.

Therefore, the most common class in the neighborhood, which is represented by the training samples, will be the class assigned to that object.

In eCognition, the distance between a sample object and image object, d , is standardized by the standard deviation of all the features. The distance between sample object s and image object i , d :

$$d = \sqrt{\sum_f \left(\frac{v_f(s) - v_f(o)}{\sigma_f} \right)^2} \quad (11)$$

Where σ_f is the standard deviation of feature values for f , $v_f(s)$ feature value of sample object for f , $v_f(o)$ feature value of image object for f . Based on the distance d (Equation 11), the multidimensional and exponential membership function $z(d)$ (Equation 12), can be calculated and where k determines the decrease of $z(d)$:

$$z(d) = e^{-kd^2} \quad (12)$$

The user defined parameter k is dependent on the data. However, a relatively higher value of k could reduce isolated and, perhaps wrongly classified objects in the image; this would be at the expense of less distinct class boundaries.

2.4 Texture analysis

Texture and tone have a complex relation to each other, where tone is the different shades of grey and texture is defined by the spatial distribution of these shades (Haralick and Shanmuga 1973). Depending on image resolution, a pixel does not only describe the spectral tone, but a part of an object, which is represented by several pixels with varying grey tones. If an image has one or several isolated areas, with tones that range within a narrow interval that also is separable from its surroundings in feature space, a texture analysis could be meaningful. However, spatial resolution and the size of relatively homogenous areas could be crucial, since texture is derived from several pixels creating a pattern rather than a single point (Tso & Mather, 2009).

A method to extract information about texture in images is the grey level co-occurrence matrix (GLCM), where the principal idea is that texture is outlined by the adjacency relationship between the various grey tones in the image. The matrix represents the frequency of grey tones occurring in a spatial relation to each other. These relationships can be considered in various directions, but most often at the angles of 0, 45, 90 and 135 degrees. To minimize any scaling effect, the matrix is normalized to a value between 0

and 1. If the image has a homogenous pattern, with larger areas of cohesive grey tones, the matrix will have larger values along the diagonal. On the contrary, if the image is heterogeneous, with a non-uniform pattern, the GLCM would have larger values away from the diagonal. This spatial relationship information of grey tones from the image can then be used to derive texture by applying different equations (Tso & Mather, 2009).

Haralick et al. (1973) suggested various methods to derive these texture features from the GLCM. Two of these techniques are entropy (Equation 13) and contrast (Equation 14), which also have been used in this thesis. Contrast represents the measure of difference between neighboring pixels gray tones. Entropy is a measure of disorder e.g. if the reflectance of pixels are randomly organized or are organized in a clustered way.

Entropy texture measures have higher output values for homogenous areas, and lower for heterogeneous ones. Contrast, on the other hand, generates higher values in a heterogeneous area and lower in homogenous (Tso & Mather, 2009).

$$Entropy = - \sum_i \sum_j P(i,j) \log(P(i,j)) \quad (13)$$

$$Contrast = \sum_{n=0}^{N_g-1} n^2 \left(\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j) \right) \quad (14)$$

Were $P(i,j)$ is the entry in the normalized GLCM and N_g the number of grey tones in the image. It has been shown in some studies that classification accuracy increases when including texture layers (Ouma, Tetuko and Tateishi 2008, Johansen et al. 2007).

2.5 Validation

The validation of the classified images is based on the widely used error matrix. By collecting ground truth samples and comparing these with the classified image, an approximation of the validity of the classification is possible. From the error matrix, several different estimates can be derived e.g. overall accuracy, kappa analysis, user accuracy and producer accuracy (Congalton and Green 2009).

The overall accuracy is based on the collected ground truth samples in comparison to the same location in the classified image. The user's accuracy is an indication and a measure of the probability that a pixel in the classified map represents the same class in the real world. Producer's accuracy, on the other hand, is an indication on how well training samples are classified for a given class (Lillesand et al. 2008).

The kappa analysis is a discrete multivariate method, which enables statistical comparison between different classifications. In this analysis, a KHAT value (Equation 15) and Z test (Equation 16) will be calculated. KHAT value (K) is an index that accommodates for the effects of chance agreement, i.e. it takes into account that some cases could be correctly classified purely by chance, and indicates the difference between the actual agreement in the error matrix and this chance agreement. If the KHAT value is significantly greater than zero, one can assume that the classification is better than a random classification (Congalton and Green 2009).

$$K = \frac{\text{observed accuracy} - \text{chance agreement}}{1 - \text{chance agreement}} \quad (15)$$

The Z-test (Equation 16) for a single error matrix is a way of indicating the significance of that accuracy. If $Z_{single} \geq Z_{\alpha/2}$, which is the confidence level of a two-tailed Z-test with an infinite number of degrees of freedom, the null hypothesis of KHAT being close to zero i.e. random classification, can be rejected.

$$Z_{single} = \frac{K}{\sqrt{\text{var}(K)}} \quad (16)$$

For the comparison between two error matrixes, one can use the KHAT value and the variance of the two matrixes to be tested to perform a Z-test (Equation 17). This test will tell if the matrixes are significantly different from each other. If $Z_{comp} \geq Z_{\alpha/2}$, one can reject the null hypothesis of the two error matrixes being equal (Congalton and Green 2009).

$$Z_{comp} = \frac{|K_1 - K_2|}{\sqrt{var(K_1) + var(K_2)}} \quad (17)$$

To test the individual class agreement, the conditional Kappa coefficient is calculated; this is done in the same fashion as for the common KHAT value but for each class (Congalton and Green 2009).

3. Method

3.1 Data and software

The image used in this thesis was acquired by the SPOT-5 HRG instrument with three pan-sharpened bands, including green, red and NIR with spatial resolution of 2.5m. The image was acquired on the 4th of august, 2012 (SPOT, 2015).

For the object based approach, eCognition 9.0 was used. This software package supplies a method which can quickly extract geo-data from remotely sensed images by segmentation. eCognition consists of three different components; Developer, Architect and Server, which all can be used for different analysis tasks. For this purpose, the eCognition Developer is used, which is an environment for object based image analysis.

ENVI 5.1 is a software package distributed by ITT Exelis, which combines advanced image processing technology with geospatial analysis for extracting meaningful information. This software is used for classifying the image with the SVM and MLC algorithms.

Matlab, which is a numerical computing environment that allows matrix manipulation, plotting etc., was used for all the classification evaluations. For the visualization of the classified images, ArcGIS 10.2.2 was used.

3.2 Training and evaluation samples

For both training data and evaluation data, visual interpretation of high-spatial resolution images available in Google Earth was used. These images were acquired in July, 2013. Google Earth have been used in previous studies and appear to be a good source for sample collection (Dorais and Cardille 2011).

The training samples were distributed over the area by creating a shape file with randomly located points, the tool Create Random Points in ArcMap was used for this purpose. However, the training samples were not entirely randomly positioned in the shape file, but were conditioned by a constraint of not being too close to each other. As a second step, the shape file was converted to a KML file to be imported to Google Earth, and therein manually determining in which land cover class each point was located. Points that was located in areas that were hard to determine, were removed; if relatively close to an identifiable area, the points were moved to that location. To acquire larger training areas, each point was expanded to a circle with a 5m radius.

The same method was used for the evaluation points. However, for this purpose, some of the points were relocated to make the distribution of points between classes more even. If a point was close to another class that had a lower number of points than the class it already belonged to, the point was moved to the class with a lower number of points.

The collection of samples derived from Google Earth resulted in 146 training samples and 494 evaluation samples. The samples, for both training and validation, were distributed between the classes as evenly as possible (Figure 4).

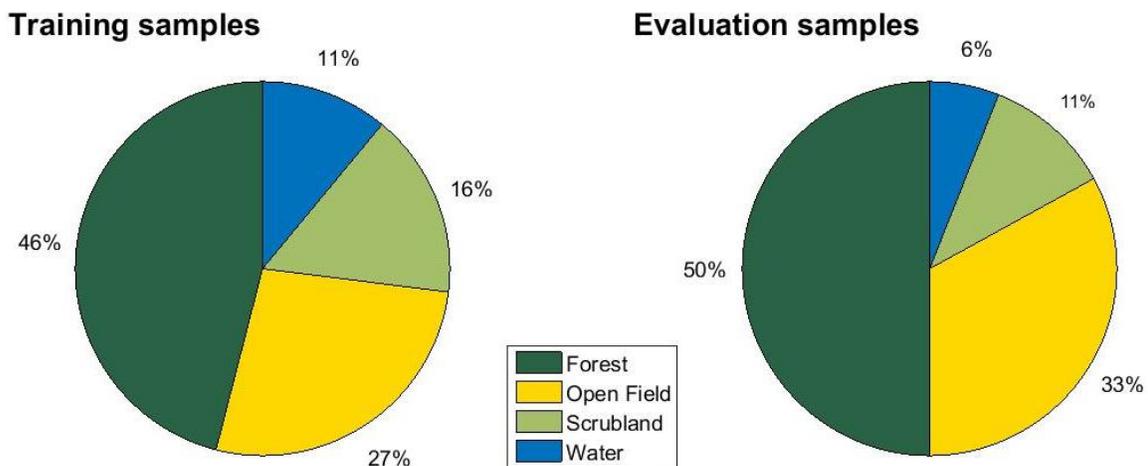


Figure 4. The distribution of the training samples and the evaluation samples between the classes. In total, there were 146 training samples and 494 evaluation samples.

3.3 Texture analysis

To obtain a representable texture layer, the moving window size is of considerable importance when calculating the texture. A moving window is an $N \times N$ pixel matrix which is a subset of the original image; this window moves over the image and calculates local values. Semivariance (Equation 18) is a method to approximate this size which is based on the relationship between pixels with a specific interval or lag along a given transect. More specifically, if the digital number of pixel i is z , and the lag is h , the relation to be studied is between $z(i)$ & $z(i+h)$. At a point in the lag interval, the semivariance curve starts to flatten out and this change in the curve represents the appropriate window size (Ouma et al. 2008).

$$S^2(h) = \frac{1}{2k} \sum [z(i) - z(i+h)]^2 \quad (18)$$

The choice of which texture derivations to use, e.g. entropy and contrast, was based on the training samples. By comparing mean and standard deviation of these samples, one can differentiate between the texture derivations. If the means are separable, and the standard deviation is low, the texture layer was chosen; this choice was based visually from a graph. Also, a test with the Jeffries-Matusita measure was performed to see differences in separability of different waveband and texture combinations (Richards and Jia 2006).

The moving window size is derived from the semivariogram, in which the whole images were used, see figure 5.

The semivariance increases with increasing pixel lag distance, and at about a 9 pixel lag, the curve begins to flatten out which indicates the appropriate window size. By a visual approximation, a window size of 9×9 pixels was chosen.

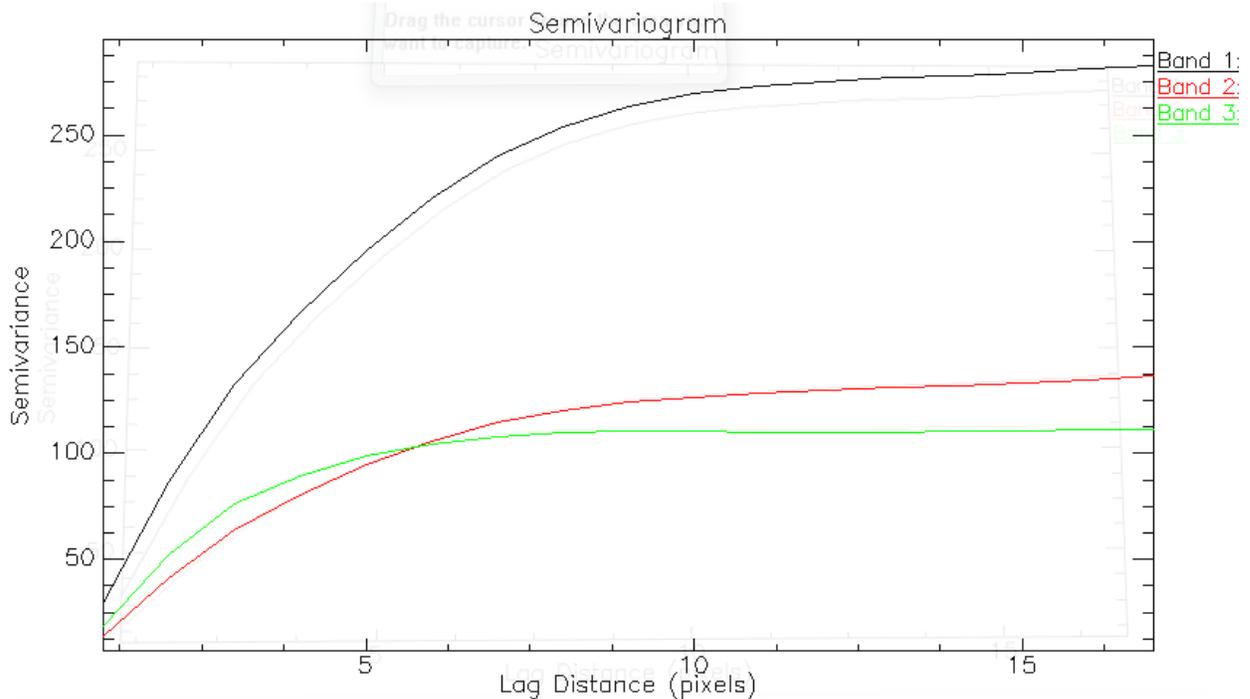


Figure 5. The semivariance, depending on the lag of pixels, for the three HR bands in the SPOT image. Green (Band 1), red (Band 2) and NIR (Band 3).

The choice of which texture measures to use was first based on studying mean and standard deviation for each class and for all texture measures calculated. If the mean of the classes was clearly separated, and had a low standard deviation, the texture was also tested with the Jeffries-Matusita test for separability to examine the separability further.

The Jeffries-Matusita (JM) measure, with possible range of 0 to 2, indicates the separability of the four classes. When combining texture measures with the original image bands from the SPOT-5 HRG, this value increases (see Table 1), and thus, increases the separability between classes. The texture derivations that were chosen to be included in the classifications were contrast texture derived from the green and NIR band, and entropy texture derived from the red band.

Table 1 presents the JM values for separability tests between the classes for the original three wavebands green, red and NIR, and also the test for separability between classes with the green, red, NIR, contrast texture of green and NIR, and entropy texture of red. The JM values are higher when the original image is combined with the three texture derivations i.e. texture increases separability. It is clear that Water is rather well separated from all other classes; however, Open Field and Scrubland are harder to differentiate from each other.

Table 1. The Jeffries-Matusita measure for separability between classes. The number ranges between 0 and 2. All values below a are calculated from the green, red and NIR band. All values below b are calculated with the green, red and NIR band, and also, the contrast texture for green and NIR, and entropy texture for the red band.

Separability	Open field		Scrubland		Water	
	a	b	a	b	a	b
JM						
Forest	1.246	1.511	1.088	1.599	1.476	1.882
Open field			0.554	0.914	1.261	1.949
Scrubland					1.375	1.969

For the moderate resolution image, the texture analysis was not done due to the test with semivariance and pixel lag showed that no appropriate moving window size was identifiable, and therefore, a texture analysis would be invalid. The original image with moderate spatial resolution with green, red and NIR bands was tested for separability, see Table 2. The highest separation is between Water and Scrubland; though, Forest can merely be separated from Open Field and Water. All JM values are considered to be low and the separability is not optimal.

Table 2. The Jeffries-Matusita measure of separability between classes; for the medium spatial resolution (25m) image, with the green, red and NIR band. The number ranges between 0 and 2.

Separability	Open field	Scrubland	Water
Forest	1.22	1.09	1.28
Open field		0.64	1.09
Scrubland			1.33

3.4 Maximum likelihood classification

For the maximum likelihood classification, the probability distribution for each class is assumed to be equal. No threshold is set for the degree of probability that a pixel belongs to a class (Duro, Franklin and Dube 2012).

3.5 Support vector machine classification

3.5.1 Calibrating the SVM algorithm, C & gamma

For calibration of the RBF-kernel, a k-fold cross validation and a grid search was applied. In the k-fold cross validation, the training samples were divided into six random subsets. Six classifications were made with different values of C and gamma for each subset. The choice of interval for the grid searches was based on previous studies (Oomen et al. 2008, Huang et al. 2002). For each classification, a subset of training samples was left out and used as validation samples for the classified image; all six classifications left out different subsets.

Using this approach when estimating the two parameters can be very time consuming and therefore a grid search is done within fixed intervals. Firstly, a coarser interval is applied, to estimate where it could be useful to apply a finer grid interval (Tso and Mather 2009, Hsu et al. 2003). By starting with a coarse grid search, it is easier to find a point to enclose the finer grid to search for suitable values, C and γ were identified to be 5.33 and 4 respectively with the coarse grid search. For the second, and finer grid search, these parameters were approximated to C being 9.6 and γ 7.2, with the highest overall accuracy at 60.6%, see Table 1.

Table 3. The calibration of γ and penalty parameter C, applying a k-fold cross validation and a grid search. The coarse grid search approximated the highest accuracy around parameter values of 4 and 5.33 respectively. The fine grid search narrows down the parameters to 7.2 and 9.6 respectively.

Coarse grid search			Fine grid search		
γ	C	Over. Acc. %	γ	C	Over. Acc. %
0	0.03	55.50	2	2.68	54.66
4	5.33	62.68	2.4	3.20	47.69
8	10.66	52.12	4.8	6.40	48.64
12	15.98	52.30	7.2	9.60	60.60
16	21.31	46.98	9.6	12.80	42.12
20	32	41.03	10	13.32	57.32

3.6 Object based classification

3.6.1 Scale parameter estimation

The scale parameter estimation tool, developed by Dragut et al. (2010), segments an image at different scales. The number of different scales, i.e. start and stop scale, and the step size of this interval is user defined. The scale interval was set to 1-200 with a step size of 2. These settings resulted in about 100 tested scale values.

Figure 6 shows the scale parameter estimation and the ROC-LV curve. Where the ROC-LV curve starts to flatten and vary within a relatively small range, this is where one can find the optimal scale parameter. For this study, the appropriate scale was derived visually, in Figure 6, the vertical lines points out three possible scale values. These scales are 47, 49 and 61; all three were tested by classification, but after evaluating the images visually, the scales 47 and 61, were ruled out due to a non-satisfactory result. And thus, the scale parameter was set to 49 for the classification in this thesis.

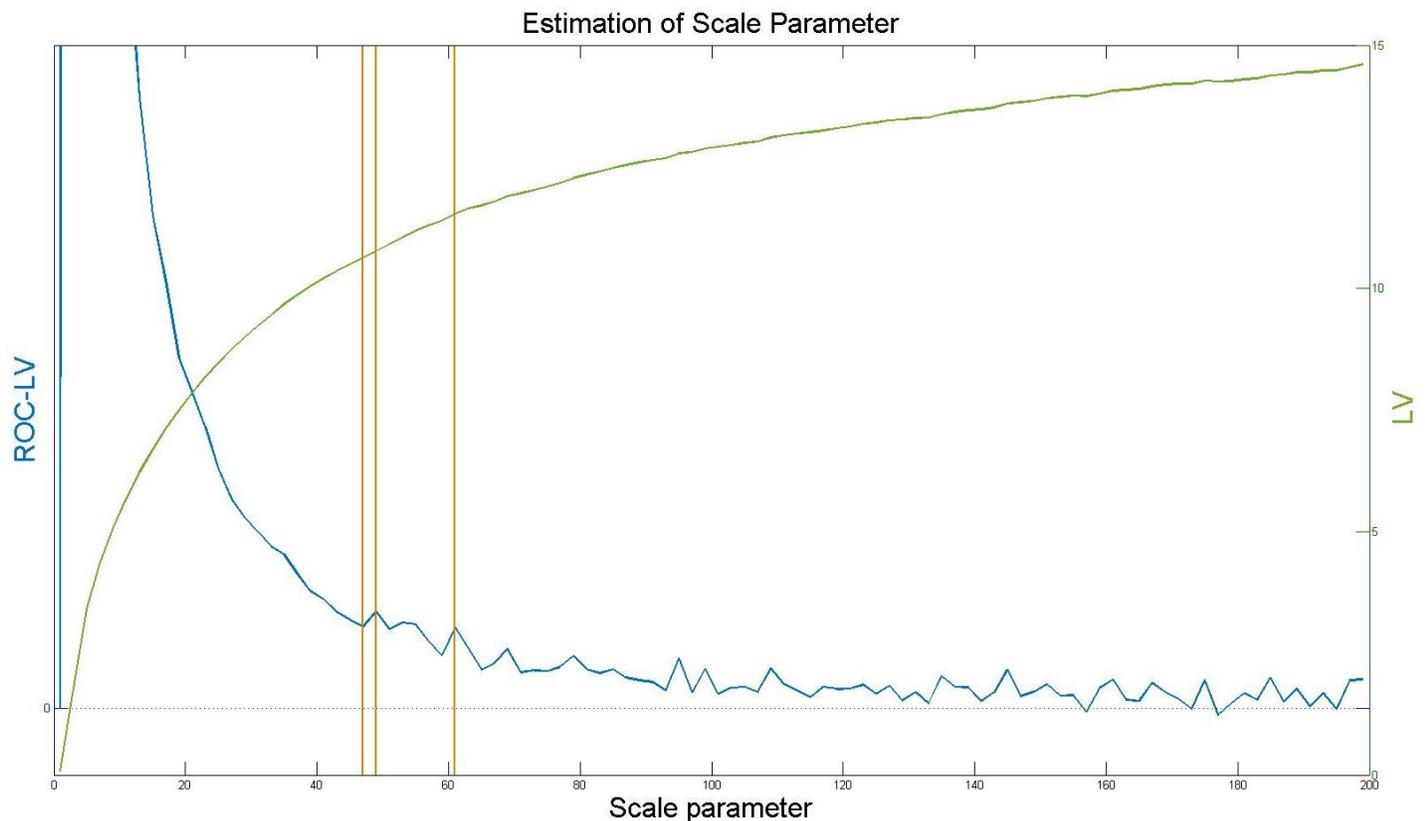


Figure 6. The blue curve represents the ROC-LV curve, green LV and the three vertical orange lines represent three scales being 47, 49 and 61 respectively.

3.6.2 Shape and compactness parameters

The parameters of shape and compactness were determined to be 0.5 and 0.9 respectively. These values were used by Duro et al. (2012), who also classified land cover with SPOT-5 HRG.

To test if these values would be appropriate both lower and higher values, in different combinations, of shape and compactness were tested. However, there was no significant change in the number of segments generated by altering the parameters.

Figure 7 shows a subset of the segmented image with the previously mentioned parameter values. The image is segmented into 7478 different objects.

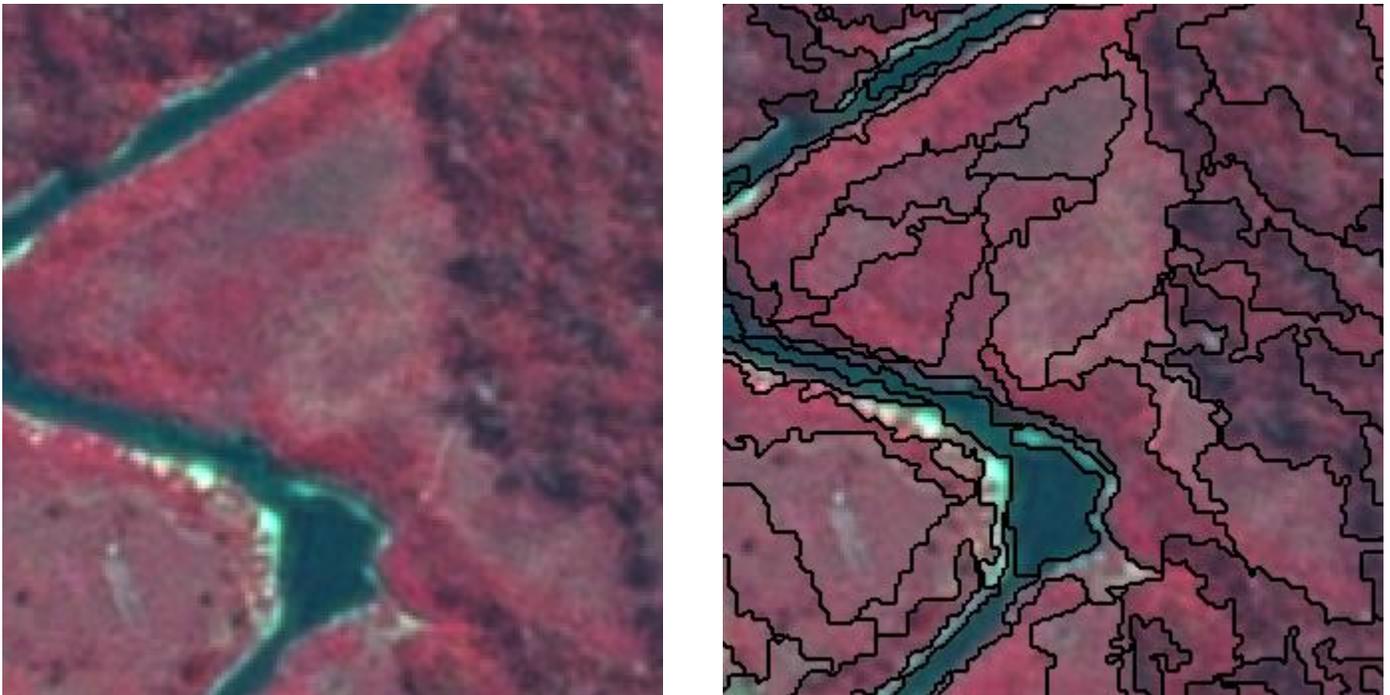


Figure 7. A subset of the resulting segmented image to the right and the original image to the left. The segmentation is done with scale parameter 49, 0.5 and 0.9 for shape and compactness respectively.

3.6.3 k-Nearest Neighbor classification

The kNN method was used for classifying the segmented image. The number of neighbors to be considered (k), when assigning an object to a class, was determined to 3. This decision was based on a visual evaluation, when testing k values ranging from 1 to 10. A k value above 5 was not suitable at all for the segmented image.

3.7 Moderate spatial resolution

The high spatial resolution image was resampled to 25m spatial resolution, i.e. ten times lower resolution which would be a rough comparison to the moderate resolution of Landsat images (30m spatial resolution). The resampling was done with a nearest neighbor approach, which assigns the pixel the value that is most often present in the neighboring pixels.

The moderate spatial resolution images went through the same processing as the high resolution images. No object based classification was done due to pixel size i.e. the potential of this approach is significantly reduced with lower spatial resolution.

A post-classification filter was applied to the MLC and SVM classifications to reduce isolated pixels. The filter applied was a majority filter with equal weights on the four closest pixels.

3.8 Validation

For the validation of the overall accuracy, the kappa coefficient and Z-value were calculated. For the Z-value, two probability thresholds are set. The first was $P < 0.12$, which is lower than the common $P < 0.05$. This extra threshold of $P < 0.12$ is set to show those Z-values that do not reach the high significance value of $P < 0.05$, but still indicates some significance at a lower value.

4. Results

4.1. Subsets of classifications

Figure 8 shows the MLC, SVM and OB-kNN classification results, produced using the three original bands; green, red and NIR. The MLC image is patchier than the two other classifications. Another visible difference is in the lower left corner in Figure 8.; the patch of Scrubland is almost not visible in the MLC and is not visible in the SVM. This is not an issue for the OB-kNN shown in Figure 8.

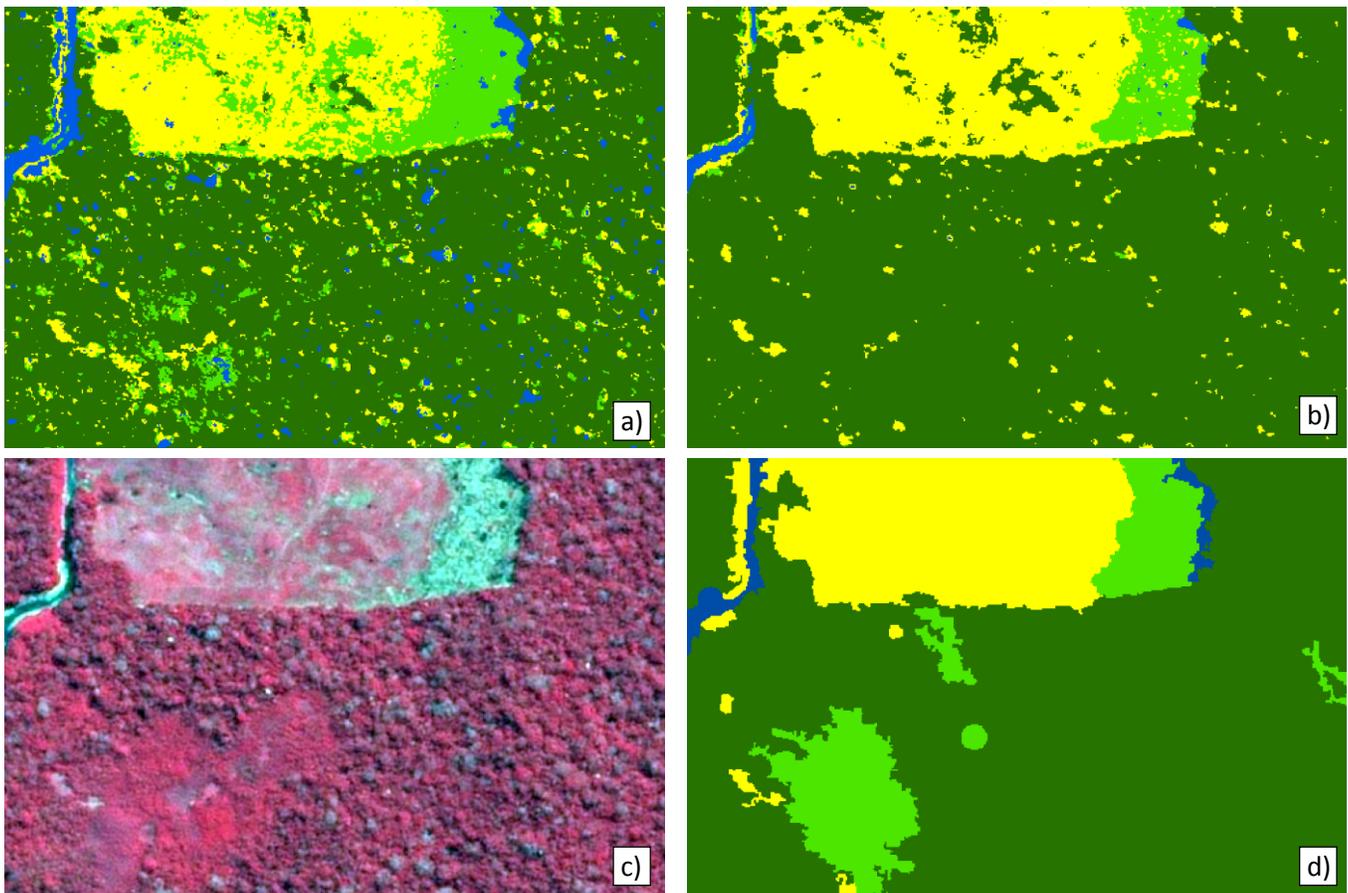


Figure 8. Subsets of the classified, high spatial resolution, images with green, red and NIR band. a) MLC b) SVM c) original SPOT image d) KNN. For the classified images, the colors represent: **dark green**-Forest, **light green**- Scrubland, **yellow**-Open Field and **blue**-Water. Scale 1:300. For full images see Appendix 1.

In Figure 9, a subset of classifications with MLC, SVM and OB-kNN, with the three original bands, green red and NIR, and contrast texture derived from green and NIR, and entropy texture is shown. It is visually obvious from Figure 9, that the Scrubland class is better distinguished in the pixel based classifications when including texture layers. The OB-kNN classification with texture layers is also classifying the Scrubland class more correctly than without texture layers.

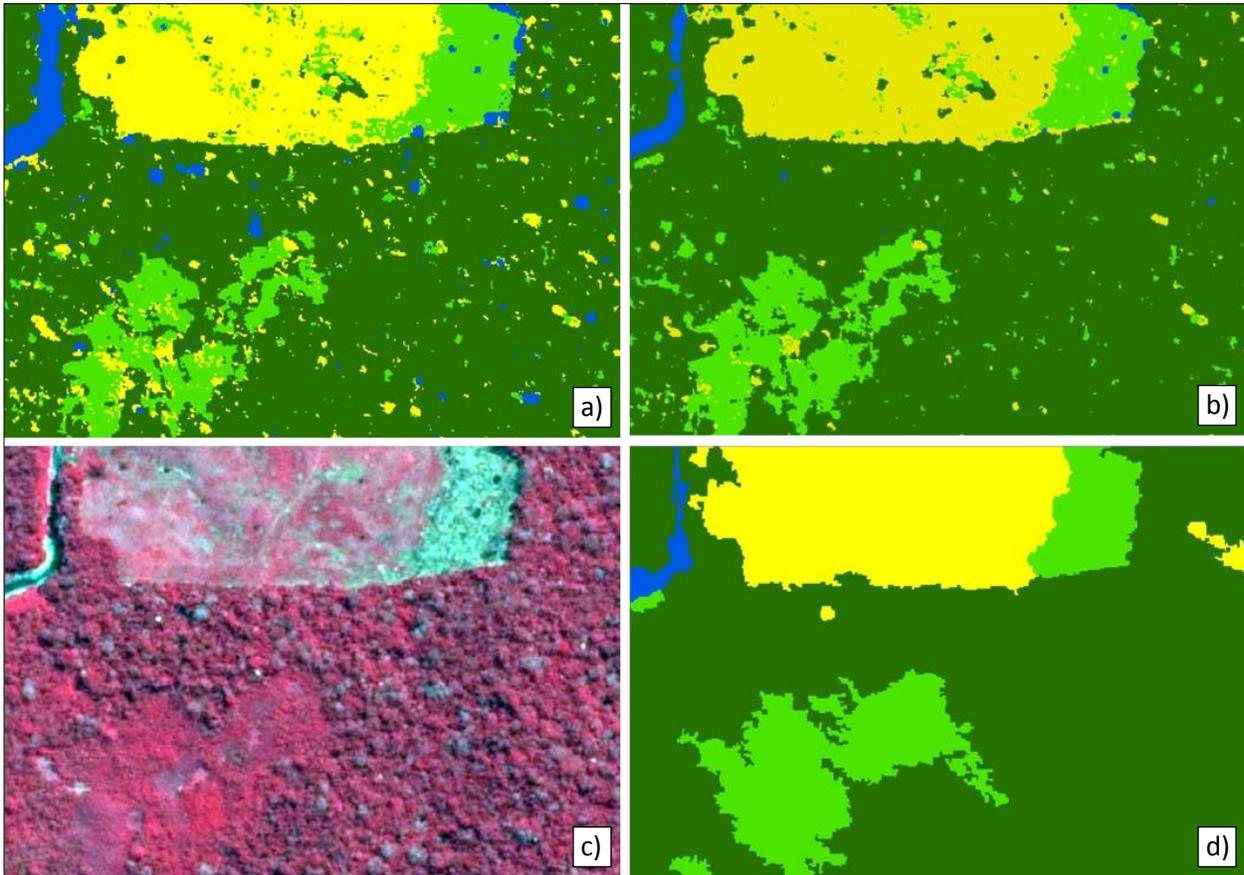


Figure 9. Subsets of the classified, high spatial resolution, images with green, red, NIR, contrast texture of green and NIR, and entropy texture of red. a) MLC b) SVM c) original SPOT image d) OB-kNN. For the classified images, the colors represent: **dark green**-Forest, **light green**- Scrubland, **yellow**-Open Field and **blue**-Water. Scale 1:300. For full image view see Appendix 1.

The subsets in Figure 10, show the MSR images, created by the resampling of the original image. The classification algorithms are MLC and SVM, and the figure also presents the classifications filtered by a majority filter

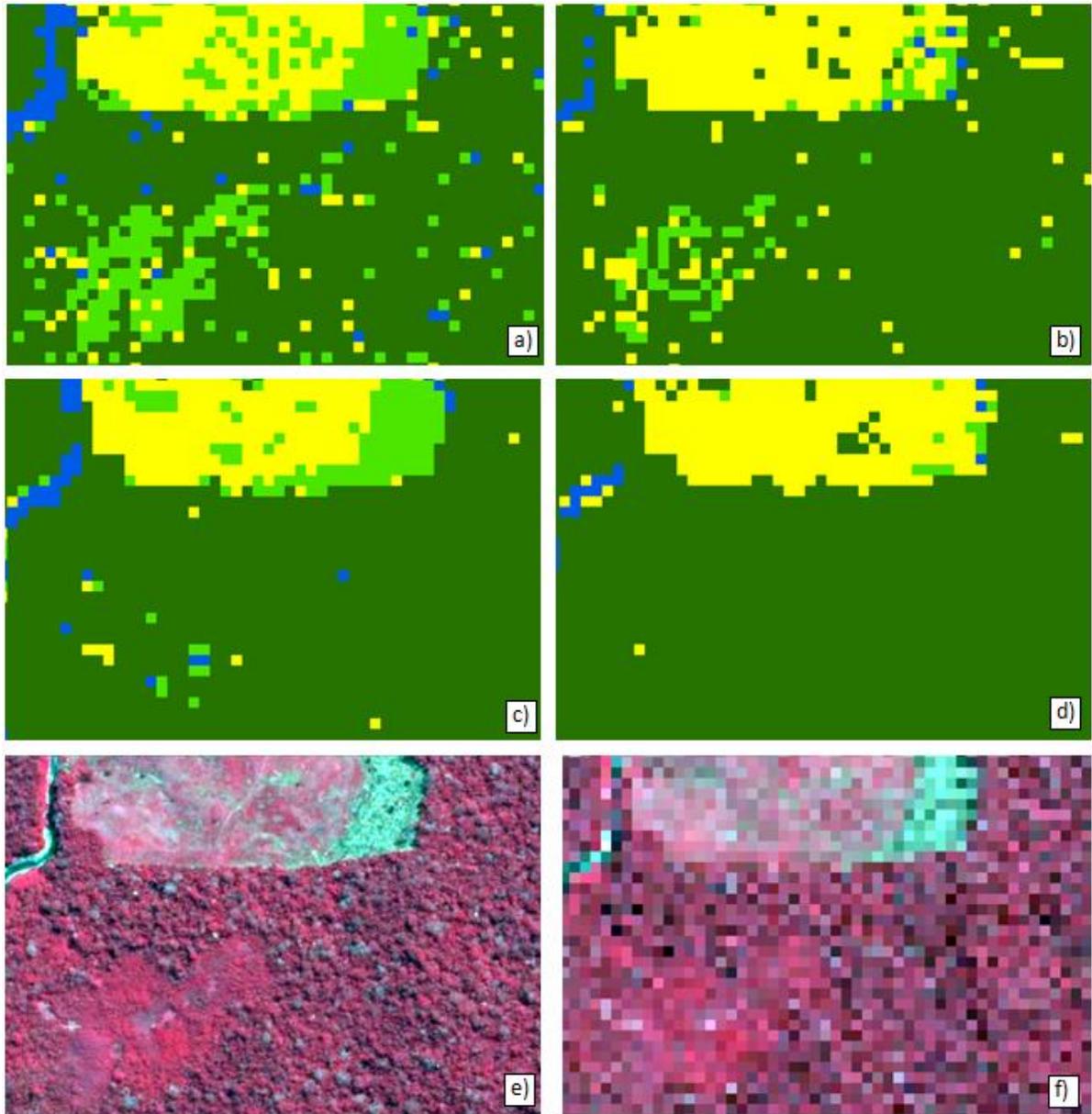


Figure 10. Subsets of the classified MSR images with green, red and NIR band. a) MLC b) SVM c) filtered MLC d) filtered SVM e) original HSR SPOT image f) MSR SPOT image. For the classified images, the colors represent: **dark green**-Forest, **light green**- Scrubland, **yellow**- Open Field and **blue**-Water. Scale 1:300. For full image view see Appendix 1.

4.2 The area percentage of each class

Table 4 displays the total area of each class in the classified images as a percentage. The area percentage distribution is similar for most of the classes when comparing the algorithms, i.e. classifications done with green, red and NIR versus classifications done with green, red, NIR and three texture layers. Between the number of input data layers to the classification algorithm (columns a vs b), the difference is generally small.

There are some classes that diverge from this tendency, and for which the extent of classes varies between classifications with and without texture. Scrubland, with SVM classification, increased by 2.9% but decreased in area by 2.7% with OB-kNN; and the Forest class increased by 3.6% for the OB-kNN when including texture.

The difference between algorithms is however greater; there is a 4.5% difference between MLC and OB-kNN for Forest, and less than 2% for Open Field and Scrubland (texture included).

The Water class in the two MLC classifications has a rather noteworthy high percentage in comparison to the other algorithms. The MLCs is around 2.5% but only around 0.5% for the two others.

Table 4. Percent of each land cover class for classifications with MLC, SVM (HSR and MSR) and object based KNN (HSR). Column labeled a represent classifications with green, red and NIR band; and b, classifications with green, red, NIR, contrast texture of green and NIR, and entropy texture of red. Column labeled c represents the classification that has been filtered by a majority filter.

(%)	HSR						MSR			
	MLC		SVM		OB-kNN		MLC		SVM	
	a	b	a	b	a	b	a	c	a	c
Forest	78.3	79.6	85.5	83.4	80.5	84.1	80.2	84.8	84.3	86.4
Open Field	14.6	14.1	13.1	12.1	12.9	12.2	12.7	11.4	15.0	13.2
Scrubland	4.3	3.8	1.0	3.9	5.9	3.2	3.7	2.0	0.3	0.1
Water	2.7	2.5	0.4	0.5	0.6	0.5	3.3	1.8	0.4	0.3

4.3 Overall accuracy

Table 4 shows the overall classification accuracies for the HSR classifications. All six classifications range within an 8.5% interval, and it is apparent from the table that all classifications perform better than a total random classification (according to the Z-test with confidence level of $P < 0.05$). For the non-texture classifications there is a 3% difference between pixel based SVM and OB-kNN. Furthermore, the overall accuracy is higher for all three algorithms when adding texture layers. However, this adds less accuracy to the OB-kNN, leading to only a small difference of 0.6% between this method and SVM

Table 4. The overall accuracy (%), kappa coefficient (unitless), kappa variance (unitless), and Z value (unitless), for each error matrix of MLC, SVM and OB-kNN. All values below column a, are derived from classifications done with the green, red and NIR band. All values below column b, are derived from classifications done with the green, red and NIR band and the contrast texture for green and NIR, and entropy texture for the red band.

Classification	HSR					
	MLC		SVM		KNN	
	a	b	a	b	a	b
Overall accuracy	78.3	84.4	79.4	86.8	82.4	86.2
Kappa coefficient	0.64	0.75	0.65	0.79	0.71	0.77
Kappa variance	0.0014	0.0009	0.0016	0.0007	0.0011	0.0008
Z-value	17.21	25.52	16.12	28.8	21.3	26.85

Table 5 shows the overall accuracies for the MSR classifications. For both filtered and non-filtered images the accuracies are higher for both MLC and SVM post-filtering, but the largest increase is between the MLC classifications, the difference between these two is 4.9%.

Table 5. The overall accuracy (%), kappa coefficient (unitless), kappa variance (unitless) and Z value (unitless), for each error matrix of MLC and SVM classifications, with the moderate spatial resolution of 25m. All classifications are done with the green, red and NIR band. Values below a represents unfiltered classifications, and b represents classifications filtered by a majority filter.

Classification	MSR			
	MLC		SVM	
	a	b	a	b
Overall accuracy	74.9	79.8	76.7	77.7
Kappa coefficient	0.59	0.65	0.60	0.61
Kappa variance	0.0017	0.0015	0.0019	0.0018
Z value	14.36	17.00	13.81	14.35

Table 6 shows user and producer accuracies for classified HSR images with green, red and NIR band. Generally, Scrubland has a low accuracy for both user and producer accuracy. Except for Scrubland, the OB-kNN has an overall high user and producer accuracy for all the classes.

Table 6. The user and producer accuracy, %, for individual classes of the three classification algorithms, MLC, SVM and OB-KNN, with the HSR green, red and NIR wavebands.

%	HSR					
	MLC		SVM		KNN	
	User	Producer	User	Producer	User	Producer
Forest	77.8	92.7	73.4	98.8	81.1	97.9
Open field	90.3	68.7	93.9	66.3	94.4	73.0
Scrubland	45.9	30.9	70.6	21.8	50.0	34.5
Water	75	100	93.5	96.7	84.8	93.3

Classifications done with the HSR original SPOT bands and texture layers, both the user and producer accuracy are lower for Scrubland than all other classes, for all classifications, see Table 7. However, the OB-kNN has slightly higher user accuracy in most classes, except for Forest where the results appear to be reversed, and the MLC has the higher accuracy. The producer accuracy however, does not follow the same pattern. Forest has very high producer accuracy in comparison to Open Field and Scrubland. Water on the other hand, has 100% producer accuracy in both MLC and SVM, and around 73% in OB-kNN.

Table 7. The user and producer accuracy (%) for the three classification algorithms, MLC, SVM and OB-kNN, with the HSR green, red and NIR wavebands and contrast texture derived from green and NIR, and entropy texture derived from the red waveband.

%	HSR					
	MLC		SVM		OB-kNN	
	User	Producer	User	Producer	User	Producer
Forest	87.6	95.1	87.1	98.4	84.2	100
Open field	90.1	72.4	95.2	73.6	97.6	75.46
Scrubland	58.3	63.6	63.8	67.3	66.0	63.6
Water	83.3	100	93.8	100	95.7	73.3

Table 8 shows the user and producer accuracy as a percentage for the classified images with MSR. These classifications are done with the original bands with the MLC and SVM algorithm. The producer accuracy of Scrubland is very low for both MLC and SVM and is even lower when filtering the classified image with a majority filter.

Table 8. The user and producer accuracy (%) for the MLC and SVM classifications with the green, red and NIR wavebands, MSR. Both for the non-filtered images and filtered by a majority filter.

	MSR							
	MLC				SVM			
	Non-filtered		Filtered		Non-filtered		Filtered	
%	User	Producer	User	Producer	User	Producer	User	Producer
Forest	75.3	91.9	75.4	99.6	73.3	95.9	74.3	100
Openfield	89.4	61.9	94.8	66.9	82.9	71.8	85.0	73.0
Scrubland	38.9	25.5	56.5	23.6	71.4	9.1	66.7	3.6
Water	65.9	96.7	87.1	90	87.5	70	85.0	56.6

Table 9 shows the number of evaluation samples that are misclassified, also how each class from the classified image is distributed among the evaluation samples. Below the subtitle GrTr (ground truth) is shown which evaluation class the samples from the classification correspond. Neither the Open Field nor Scrubland class agrees well with the evaluation samples. Samples that should have been classified as Open Field have been distributed within Forest and Scrubland, and samples that should have been Scrubland are generally classified as Forest. However, misclassification is reduced in classifications including texture analysis.

Classified samples (pixels) that correspond to the same location as the evaluation samples do not always agree with the correct class within the evaluation samples. Table 9, in column “Class”, shows that areas classified as Scrubland are not that well classified; many of these classified samples have other corresponding evaluation classes, in most cases Open Field. Forest, Open Field and Water are, however, fairly well classified compared to the evaluation samples.

Table 9. The table is firstly divided depending on classification algorithm; MLC, SVM and OB-kNN, further, which data included in the classification i.e. a) represents classifications done with green, red and NIR band, and b) represents classifications with green, red, NIR, contrast texture of green and NIR, and entropy texture derived from the red band. The next subcolumn is GrTr/Class; where GrTr columns represent how the evaluation samples (Ground Truth) have been classified, and the Class columns represent the distribution of each classified class within the valuation samples.

	MLC				SVM				kNN			
	a)		b)		a)		b)		a)		b)	
	GrTr	Class	GrTr	Class	GrTr	Class	GrTr	Class	GrTr	Class	GrTr	Class
	Forest				Forest				Forest			
Forest	228	228	234	234	243	243	242	242	241	241	246	246
OpenField	9	32	5	23	3	49	1	24	0	27	0	22
Scrubland	6	33	4	10	0	38	3	12	5	29	0	16
Water	3	0	3	0	0	1	0	0	0	0	0	8
	Open Field				Open Field				Open Field			
Forest	32	9	23	5	49	3	24	1	27	0	22	0
OpenField	12	12	118	118	108	108	120	120	119	119	123	123
Scrubland	14	3	21	8	5	4	18	5	13	6	18	3
Water	5	0	1	0	1	0	1	0	4	1	0	0
	Scrubland				Scrubland				Scrubland			
Forest	33	6	10	4	38	0	12	3	29	5	16	0
OpenField	3	14	8	21	4	5	5	18	6	13	3	18
Scrubland	17	17	35	35	12	12	37	37	19	19	35	35
Water	2	0	2	0	1	0	1	0	1	1	1	0
	Water				Water				Water			
Forest	0	3	0	3	1	0	0	0	0	0	8	0
OpenField	0	5	0	1	0	1	0	1	1	4	0	0
Scrubland	0	2	0	2	0	1	0	1	1	1	0	1
Water	30	30	30	30	29	29	30	30	28	28	22	22

4.4 Z-test between error matrices

According to the Z-test in Table 10, none of the error matrices significantly differ from each other when comparing the algorithms. However, comparing the matrices of the same algorithm, but with and without texture, a significant difference is evident for MLC and SVM.

Table 10. Z-test between the kappa values derived from the error matrices (see Appendix). a) P-values for classifications done with green, red and NIR bands. b) P-values for classifications done with green, red, NIR, contrast texture of green and NIR, and entropy texture of red, and c) test between classifications with original band and texture derivations (as in b)) .

	a)			b)			c)
	MLC	SVM	kNN	MLC	SVM	kNN	
MLC		-	P<0.12		-	-	P<0.05
SVM			P<0.12			-	P<0.05
KNN							P<0.12

The Z-test results in Table 11, analyses the differences in agreement between classes for the classifications, with only the Forest and Water class being significantly different from each other.

Table 11. Z-test results between individual class agreements for classification with green, red, NIR, contrast texture of green and NIR, and entropy texture of red.

	High			Moderate
Z-test	MLC vs SVM	KNN vs MLC	SVM vs KNN	MLC vs SVM
Forest	P<0.05	P<0.05	P<0.05	P<0.05
Open Field	NS	NS	NS	P<0.12
Scrubland	NS	NS	NS	P<0.12
Water	-	P<0.05	P<0.05	P<0.05

5. Discussion

When using only the three original image bands as input to the classification algorithm, SVM and OB-kNN were superior to MLC, with an overall accuracy of 79.4%, 82.4% and 78.3% for the SVM, OB-kNN and MLC respectively. Expanding the original feature space with three texture layers, increases the overall accuracy and kappa value for all three classification algorithms. Also when including texture layers, the SVM and OB-kNN performs better than the MLC, with overall accuracies of 86.8%, 86.2% and 84.4% for SVM, OB-kNN and MLC respectively.

This study of classification methods has given results that both correspond to previous studies and also some results that show the opposite. Most studies show that the parametric MLC is insufficient when using high spatial resolution images and that it generate noisy classified images with single or patches of pixels scattered throughout the image.

In this study, some of the results are similar to what previous studies concluded, where e.g. the SVM often outperform the MLC algorithm (Huang et al. 2002, Oommen et al. 2008, Sanchez-Hernandez, Boyd and Foody 2007). The studies provide various reasons to explain this superior behavior of SVM, e.g. according to Huang et al. (2002), this is due to its ability to find the optimal separating hyperplane, and according to Sanchez-Hernandez et al. (2007), the reason is that less training samples are needed for good results than for the MLC approach.

For high spatial resolution images, the object based classifications often outperform the pixel based ones (Rittl et al. 2013, Platt and Rapoza 2008, Rahman and Saha 2008, Duro et al. 2012). However, this is not the case in this study, where there is no significant difference between the best performing pixel based method and the object based method. Jebur et al. (2014), who also used SPOT-5 data for land cover mapping, concluded that the pixel based approach was outperformed by the object based one when using the original image bands.

5.1 Accuracy assessments

5.1.1 High Spatial Resolution (HSR)

The HSR classified images' overall accuracies vary within the narrow interval of 4% for classifications with the original bands; and only within 2%, when including the three texture layers. The overall accuracy and kappa coefficient for all three classification algorithms with the high HSR data, increases when including the texture derivations of entropy and contrast to the original wavebands. However, from a visual interpretation of

the classified images, and of all three algorithms, it is clear that the classification methods produce results that differ to a great extent.

According to Foody (2002), the overall accuracy for a classified image should be at least 85% to be accepted as valid. The pixel based SVM and the OB-kNN with texture derivations, both have an overall accuracy above this acceptable level, 86.8% and 86.2% respectively. The MLC with texture derivations, has an 84.4% overall accuracy, and thus, does not meet this threshold; neither do any of the classified images that did not include texture layers which had overall accuracies of 78.3%, 79.4% and 82.4% for MLC, SVM and OB-kNN respectively.

Even though the classifications without texture layers did not meet the threshold of 85% there is a significant difference between the object based and the two pixel based algorithms ($P < 0.12$), indicating that the object based method is superior. The reason for this could be due to both texture analysis and the fact that the object based segmentation considers spatial information from neighboring pixels. When comparing the three non-texture classifications, this is seen when the OB-kNN is performing significantly better than the MLC and SVM without any spatial information (Li et al. 2014).

5.1.2 Moderate Spatial Resolution (MSR)

None of the classified images of MLC and SVM with MSR, meet the overall accuracy threshold of 85% (Foody 2002). The overall accuracy and kappa coefficient increased when filtering the classification with a majority filter, where the overall accuracy increased by 4.9% for MLC and 1% for SVM. This difference of increasing overall accuracy between the two classifications could be a function of the more speckled pre-filter image for the MLC than SVM; e.g. a higher number of isolated pixels in the MLC image that are reduced by filtering.

5.1.3 Misclassified samples

Misclassified samples are most noticeable in the Scrubland and Open Field classes, many of the evaluation samples do not correspond to the same class in the classified images. Misclassified samples often correlate with the boundaries of classes but none of this information is provided in the error matrix. Therefore, the method of selecting evaluation samples is of great importance to the error matrix (Foody 2002, Foody 2005).

The standard approach to accuracy assessment (Congalton and Green 2009), which is comparing samples of ground truth with spatially corresponding samples from the classified image, has some limitations, e.g. the spatial distribution of misclassified samples. It is important to note which class the misclassified sample actually belongs to, as the spatial distribution of errors in the image is of great importance and should be investigated further. Leibovici (2009), developed a method to describe the spatial entropy, which could be considered as a measure of disorder, and this measure could

explain patterns of misclassified samples. Leibovici (2009), used the distribution of multiple co-occurrences and the spatial entropy of the classification errors that describe the errors spatially. This is, however, not done in this study, but could be an extension to the error matrix in future studies.

5.2 Area percentage of each class and visual interpretation

5.2.1 High Spatial Resolution (HSR)

It is clear that the Water class has been assigned to far more pixels in the MLC classifications, both with and without texture layers. When visually observing both of the MLC classifications, the Water class is relatively evenly distributed in patches of approximately 7x7 pixels over the entire image. Comparing these patches to the original SPOT image, the reason seems to be shadows and deciduous trees being leafless, i.e. the MLC algorithm cannot differentiate between these spectral signals properly. This could be due to the assumption of normal distribution of each class, and the training samples of Water are not normally distributed. This issue for the Water class only pertains to the MLC classifications.

The Scrubland class in the images, classified with only original bands by MLC and SVM, is speckled and evenly distributed in patches over the images, and in the OB-kNN image there are whole segments assigned to Scrubland, also scattered throughout the image. For all three algorithms, this issue of patchiness is reduced by including texture layers. Furthermore, the areas with Scrubland are better defined with texture derivations, and this increase in differentiation between vegetation classes has also been seen in other studies (Johansen et al. 2007, Kayitakire, Hamel and Defourny 2006, Ouma et al. 2008).

There is an important visual difference between the MLC and SVM algorithms when including texture layers; in the MLC image, both the classes of Open Field and Scrubland are distributed in patches. However, the SVM algorithm does not have this problem to the same extent. When only applying the original bands with the SVM algorithm, very few pixels are assigned to Scrubland; by including texture derivations this class increases from only covering 1% to almost 4% of the image.

5.2.2 Moderate Spatial Resolution (MSR)

In the classified images with MSR, the Scrubland class is poorly distinguished, especially in the SVM classification. This is an issue for both high and moderate spatial resolutions, which could indicate that it is difficult for the algorithm to discriminate between some of the classes when only using the original bands as input, since this problem is reduced with texture layers in HSR images. The filter applied to MLC and SVM, at MSR, seems to reduce the speckled pattern. However, this seems to come at the cost of accurately

representing the Scrubland class, which is severely reduced when applying the majority filter.

5.3 Texture analysis

From this study, it is apparent that a texture analysis increases the separation between classes in the feature space; especially between vegetation classes, where the Scrubland class is better distinguished for all algorithms with texture layers. This could indicate that even with HSR data it is hard to discriminate between vegetation classes when using a pixel based method (Johansen et al. 2007).

According to previous studies when using contextual information applied in pixel based classifications outperformed object based methods, that pixel based method with contextual information showed higher flexibility in predicting different sized objects than e.g. segmentation (Cai and Liu 2013).

Also, the question of how to collect training samples i.e. should a sample be collected representing the center of an object or close to the border (due to within-class spectral variability), can be reduced by including spatial information such as texture (Pacifi, Chini and Emery 2009).

5.4 Estimation of scale

Previous studies have concluded that for HSR images, the object based classification outperform the pixel based methods (Rittl et al. 2013). The merging of pixels into objects has proven to be efficient, especially for high and very high spatial resolution images; this is highly dependent on the estimation of the scale parameter that determines the size and number of objects. Some of the previous studies have done segmentation on multiple levels (Benz et al. 2004), and received high overall accuracies and kappa coefficients. In this thesis, only one level of segmentation is used. The accuracy assessment for the OB-kNN in this thesis, resulted in a valid thematic map with an overall accuracy above 85% (Foody 2002), but it did not exceed the overall accuracy of the pixel based SVM, as was the case in some other studies. Due to the importance of the scale parameter in segmentation, one can assume that this parameter could have been better estimated. Also, the appropriate scale chosen in this thesis is partly subjective when deriving the value visually from a graph.

However, very recently an automated parameterization was developed for segmenting multiple layers, which would generate a segmented image in multiple layers. Using this method would result in an objective scale determination (Dragut et al. 2014).

5.5 SVM parameter estimation

The SVM algorithm with an RBF kernel has two user-defined variables, gamma and Penalty parameter C. The algorithm is according to Kavzoglu and Colkesen (2009), very sensitive to these parameters. The cross validation and grid search method could be seen as a rather rough parameter estimation method; a less subjective and more precise method would be preferable. However, the fact that the RBF kernel only has two parameters makes it less sensitive than e.g. a polynomial that has three parameters. Finding an objective means to determine which kernel to use and how to calibrate parameters correctly and efficiently is a non-trivial task (Kavzoglu and Colkesen 2009, Oommen et al. 2008).

5.6 Computational time

The computation of the pixel based SVM algorithm, with HSR, was rather time consuming in ENVI, and using MLC was only slightly faster. The multi-resolution segmentation process and the classification of the segments in eCognition, was fairly fast in comparison to the pixel based approaches in ENVI. However, running the scale estimation algorithm was also rather time consuming. By reducing the spatial resolution by a factor of 10, one MSR pixel contained 100 HSR pixels and this definitely decreased the computational time. According to the results of this study, this gain was at the cost of accuracy.

5.7 Google Earth as ground truth

The method of acquiring and selecting both training and evaluation sample is of great importance in the error matrix (Foody 2002). Using Google Earth as a source, and its high resolution images for sampling both training and evaluation samples, seems to be an adequate method, which has also been shown in previous studies (Dorais and Cardille 2011). Creating a random point layer can cause biases if points are moved to more satisfying locations. However, using Google Earth as a source for ground truth sampling also generates lower costs and can be useful in studies with financial limitations. However, there are also limitations with this approach, as the Earth's entire surface is not covered with sufficiently HSR images in Google Earth.

However, the traditional concept of collecting training samples is to collect samples where the class seems most representative, i.e. where the spectral signal is as true as possible to the class it is to represent. However, it could be argued that due to the higher spatial resolution, the within-class spectral variability increases (Pacifici et al. 2009), and according to Foody and Mathur (2006), training samples should be collected differently depending on classifier. They argued that at least for training samples to be applied in an

SVM classifier, that samples should be collected where pixels are spectrally mixed, i.e. close to borders between classes.

5.8 25m resolution (MSR)

The lowering of resolution did not generate results similar to those from HSR classified images. The MSR classifications produced by both MLC and SVM, performed significantly more poorly than all HSR classifications. This has also been shown in studies by Yang et al. (2011), although, they were more specific in their classification classes than in this study. Kavzoglu and Colkesen (2009) found that the SVM with RBF kernel is better than MLC for 30m spatial resolution images, which is also confirmed by this study. However, there is not any significant difference, and with the majority filter, the MLC performs better than the SVM.

The nearest neighbor resampling method may not be the best choice; the image should perhaps have been resampled with cubic convolution due to its better ability to handle continuous data. Nearest neighbor method is better suited for categorical data.

6. Conclusion

Of all three classification algorithms, SVM, MLC and the OB-kNN, only two had an overall accuracy of 85%. Both of these two classified images was classified with the green, red and NIR band, with contrast texture derived from the green and NIR band, and entropy texture derived from the red band. Texture analyses with the co-occurrence GLCM separated the classes further in feature space, as compared to when only using the original green, red and NIR bands.

For classified images without texture derivation i.e. only green, red and NIR band, the overall accuracy was below the 85% threshold. However, there was a significant difference between the two pixel based and object based classifications. This is probably due to the fact that the object based approach considers the spatial information in the segmentation process; this is not considered in any of the pixel based approaches when not including texture derivations. The lowering of the resolution of the original SPOT-5 image did not give a satisfying result; this leads to the conclusion that high spatial resolution is appropriate for this purpose.

When considering which classifier that is the most appropriate for this area and spatial resolution, the object based is the best suited approach. This decision is based on 1) the parameter estimation is less subjective compared to the SVM parameter estimation, 2) computational time is relatively lower, and 3) the classes in the image are more cohesive, and thus the need for post-classification filtering is lower.

7. References

- Benz, U. C., P. Hofmann, G. Willhauck, I. Lingenfelder & M. Heynen (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *Isprs Journal of Photogrammetry and Remote Sensing*, 58, 239-258.
- Blaschke, T. (2010) Object based image analysis for remote sensing. *Isprs Journal of Photogrammetry and Remote Sensing*, 65, 2-16.
- Blaschke, T. & J. Stroble (2001) What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. *Interfacing Remote Sensing and GIS, eCognition*, 6.
- Burges, C. J. C. (1998) A tutorial on Support Vector Machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2, 121-167.
- Cai, S. & D. Liu (2013) A comparison of object-based and contextual pixel-based classifications using high and medium spatial resolution images. *Remote Sensing Letters*, 4, 998-1007.
- Cartwright, W., T. Blaschke, G. Gartner, G. J. Hay, S. Lang, L. Meng & M. P. Peterson. 2008. *Object-Based Image Analysis. [Elektronisk resurs] : Spatial Concepts for Knowledge-Driven Remote Sensing Applications*. Berlin, Heidelberg : Springer-Verlag Berlin Heidelberg, 2008.
- Congalton, R. G. & K. Green. 2009. *Assessing the accuracy of remotely sensed data. [Elektronisk resurs] : principles and practices*. Boca Raton : CRC Press, 2009.
2. ed.
- Cortes, C. & V. Vapnik (1995) Support-Vector Networks. *Machine Learning*, 20, 273-297.
- Dorais, A. & J. Cardille (2011) Strategies for Incorporating High-Resolution Google Earth Databases to Guide and Validate Classifications: Understanding Deforestation in Borneo. *Remote Sensing*, 3, 1157-1176.
- Dragut, L., O. Csillik, C. Eisank & D. Tiede (2014) Automated parameterisation for multi-scale image segmentation on multiple layers. *Isprs Journal of Photogrammetry and Remote Sensing*, 88, 119-127.
- Dragut, L., D. Tiede & S. R. Levick (2010) ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science*, 24, 859-871.
- Duro, D. C., S. E. Franklin & M. G. Dube (2012) A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment*, 118, 259-272.
- Foody, G. M. (2002) Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80, 185-201.
- Foody, G. M. (2005) Local characterization of thematic classification accuracy through spatially constrained confusion matrices. *International Journal of Remote Sensing*, 26, 1217-1228.
- Foody, G. M. & A. Mathur (2006) The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment*, 103, 179-189.
- Haralick, R. M. & K. Shanmuga (1973) COMPUTATIONALLY SIMPLE PROCEDURE FOR IMAGERY DATA COMPRESSION BY KARHUNEN-LOEVE METHOD. *Ieee Transactions on Aerospace and Electronic Systems*, AES9, 813-813.
- Hsu, C.-W., C.-C. Chang & C.-J. Lin. 2003. A practical guide to support vector classification.
- Huang, C., L. S. Davis & J. R. G. Townshend (2002) An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, 23, 725-749.

- Jebur, M. N., H. Z. M. Shafri, B. Pradhan & M. S. Tehrany (2014) Per-pixel and object-oriented classification methods for mapping urban land cover extraction using SPOT 5 imagery. *Geocarto International*, 29, 792-806.
- Johansen, K., N. C. Coops, S. E. Gergel & Y. Stange (2007) Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. *Remote Sensing of Environment*, 110, 29-44.
- Kavzoglu, T. & I. Colkesen (2009) A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, 11, 352-359.
- Kayitakire, F., C. Hamel & P. Defourny (2006) Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sensing of Environment*, 102, 390-401.
- Leibovici, D. G. 2009. Defining Spatial Entropy from Multivariate Distributions of Co-occurrences. In *Spatial Information Theory, Proceedings*, eds. K. S. Hornsby, C. Claramunt, M. Denis & G. Ligozat, 392-404.
- Li, M., S. Zang, B. Zhang, S. Li & C. Wu (2014) A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information. *European Journal of Remote Sensing*, 47, 389-411.
- Liang, S., X. Li & J. Wang. 2012. *Advanced remote sensing. [Elektronisk resurs] : terrestrial information extraction and applications*. Oxford : Academic Press, 2012.
- Lillesand, T. M., R. W. Kiefer & J. W. Chipman. 2008. *Remote sensing and image interpretation*. New York : Wiley, cop. 2008

6. ed.

- Oommen, T., D. Misra, N. K. C. Twarakavi, A. Prakash, B. Sahoo & S. Bandopadhyay (2008) An objective analysis of Support Vector Machine based classification for remote sensing. *Mathematical Geosciences*, 40, 409-424.
- Ouma, Y. O., J. Tetuko & R. Tateishi (2008) Analysis of co-occurrence and discrete wavelet transform textures for differentiation of forest and non-forest vegetation in very-high-resolution optical-sensor imagery. *International Journal of Remote Sensing*, 29, 3417-3456.
- Pacifici, F., M. Chini & W. J. Emery (2009) A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sensing of Environment*, 113, 1276-1292.
- Platt, R. V. & L. Rapoza (2008) An evaluation of an object-oriented paradigm for land use/land cover classification. *Professional Geographer*, 60, 87-100.
- Rahman, R. & S. K. Saha (2008) Multi-resolution Segmentation for Object-based Classification and Accuracy Assessment of Land Use/Land Cover

Classification using Remotely Sensed Data. *J. Indian Soc. Remote Sens.*, 189-201.

Richards, J. A. & X. Jia. 2006. *Remote Sensing Digital Image Analysis. [electronic resource] : An Introduction*. Berlin, Heidelberg : Springer-Verlag Berlin Heidelberg, 2006.

4th Edition.

- Rittl, T., M. Cooper, R. J. Heck & M. V. R. Ballester (2013) Object-Based Method Outperforms Per-Pixel Method for Land Cover Classification in a Protected Area of the Brazilian Atlantic Rainforest Region. *Pedosphere*, 23, 290-297.

Sanchez-Hernandez, C., D. S. Boyd & G. M. Foody (2007) Mapping specific habitats from remotely sensed imagery: Support vector machine and support vector data description based classification of coastal saltmarsh habitats. *Ecological Informatics*, 2, 83-88.

Shekhar, S., S. Shekhar & H. Xiong. 2008. *Encyclopedia of GIS. [Elektronisk resurs]*. New York : Springer, c2008.

Stein, A., F. v. d. Meer & B. Gorte. 1999. *Spatial statistics for remote sensing*. Dordrecht : Kluwer, cop. 1999.

Tso, B. & P. M. Mather. 2009. *Classification methods for remotely sensed data. [Elektronisk resurs]*. Boca Raton : CRC Press, c2009.

2nd ed.

Yang, C., J. H. Everitt & D. Murden (2011) Evaluating high resolution SPOT 5 satellite imagery for crop identification. *Computers and Electronics in Agriculture*, 75, 347-354.

SKY Rainforest Alliance (2015). Available from:

http://www.wwf.org.uk/where_we_work/south_america/amazon/sky_rainforest_rescue.cfm
[2015-03-11]

SPOT (2015). Available from: <http://www.geo-airbusds.com/en/4388-spot-1-to-spot-5-satellite-images> [2015-03-11]

Landsat (2015). Available from: <http://landsathandbook.gsfc.nasa.gov/program/> [2015-03-11]

eCognition User Guide (2012). eCognition Developer 8.8 User Guide, Trimble Germany GmbH.

8. Appendix

8.1 Error matrix

High Spatial Resolution, 2.5m

Maximum likelihood classificaton

Table 12. The error matrix of MLC classification with green, red and NIR band; including user and producer accuracy.

HSR	Classified data						Accuracy (%)	
	Forest	Openfield	Scrubland	Water	Total	User	Prod.	
Ground truth data	Forest	228	9	6	3	246	77.8	92.7
	OpenField	32	12	14	5	163	90.3	68.7
	Scrubland	33	3	17	2	55	45.9	30.9
	Water	0	0	0	30	30	75	100
	Total	293	124	37	40	494		

Table 13. The error matrix of MLC classification with green, red and NIR band, and contrast texture derivation of green and NIR band, and entropy texture derived from the red band; including user and producer accuracy.

HSR	Classified data						Accuracy (%)	
	Forest	Openfield	Scrubland	Water	Total	User	Prod.	
Ground truth data	Forest	234	5	4	3	246	87.6	95.1
	OpenField	23	118	21	1	163	90.1	72.4
	Scrubland	10	8	35	2	55	58.3	63.6
	Water	0	0	0	30	30	83.3	100
	Total	267	131	60	36	494		

Support vector machine

Table 14. The error matrix of SVM classification with green, red and NIR band; including user and producer accuracy.

HSR	Classified data						Accuracy (%)	
Ground truth data		Forest	Openfield	Scrubland	Water	Total	User	Prod.
	Forest	243	3	0	0	246	73.4	98.8
	OpenField	49	108	5	1	163	93.9	66.3
	Scrubland	38	4	12	1	55	70.6	21.8
	Water	1	0	0	29	30	93.5	96.7
	Total	331	115	17	31	494		

Table 15. The error matrix of SVM classification with green, red and NIR band, and contrast texture derivation of green and NIR band, and entropy texture derived from the red band; including user and producer accuracy.

HSR	Classified data						Accuracy (%)	
Ground truth data		Forest	Openfield	Scrubland	Water	Total	User	Prod.
	Forest	242	1	3	0	246	87.1	98.4
	OpenField	24	120	18	1	163	95.2	73.6
	Scrubland	12	5	37	1	55	63.8	67.3
	Water	0	0	0	30	30	93.8	100
	Total	278	126	58	32	494		

Object based k-nearest neighbor

Table 16. The error matrix of MLC classification with green, red and NIR band; including user and producer accuracy.

HSR	Classified data						Accuracy (%)	
		Forest	Openfield	Scrubland	Water	Total	User	Prod.
Ground truth data	Forest	241	0	5	0	246	81.1	97.9
	OpenField	27	119	13	4	163	94.4	73.0
	Scrubland	29	6	19	1	55	50.0	34.5
	Water	0	1	1	28	30	84.8	93.3
	Total	297	126	38	33	494		

Table 17. The error matrix of MLC classification with green, red and NIR band, and contrast texture derivation of green and NIR band, and entropy texture derived from the red band; including user and producer accuracy.

HSR	Classified data						Accuracy (%)	
		Forest	Openfield	Scrubland	Water	Total	User	Prod.
Ground truth data	Forest	246	0	0	0	246	84.2	100
	OpenField	22	123	18	0	163	97.6	75.46
	Scrubland	16	3	35	1	55	66.0	63.6
	Water	8	0	0	22	30	95.7	73.3
	Total	292	126	53	23	494		

Moderate Spatial Resolution, 25m

Maximum likelihood

Table 18. The error matrix of MLC classification with green, red and NIR band; including user and producer accuracy.

MSR	Classified data						Accuracy (%)	
		Forest	Openfield	Scrubland	Water	Total	User	Prod.
Ground truth data	Forest	226	8	7	5	246	75.3	91.9
	OpenField	40	101	15	7	163	89.4	61.9
	Scrubland	34	4	14	3	55	38.9	25.5
	Water	0	0	1	29	30	65.9	96.7
	Total	300	113	37	44	494		

Support vector machine

Table 19. The error matrix of MLC classification with green, red and NIR band; including user and producer accuracy.

MSR	Classified data						Accuracy (%)	
		Forest	Openfield	Scrubland	Water	Total	User	Prod.
Ground truth data	Forest	236	10	0	0	246	73.3	95.9
	OpenField	43	117	2	1	163	82.9	71.8
	Scrubland	36	12	5	2	55	71.4	9.1
	Water	7	2	0	21	30	87.5	70
	Total	322	141	7	24	494		

8.2 Maps, full extent

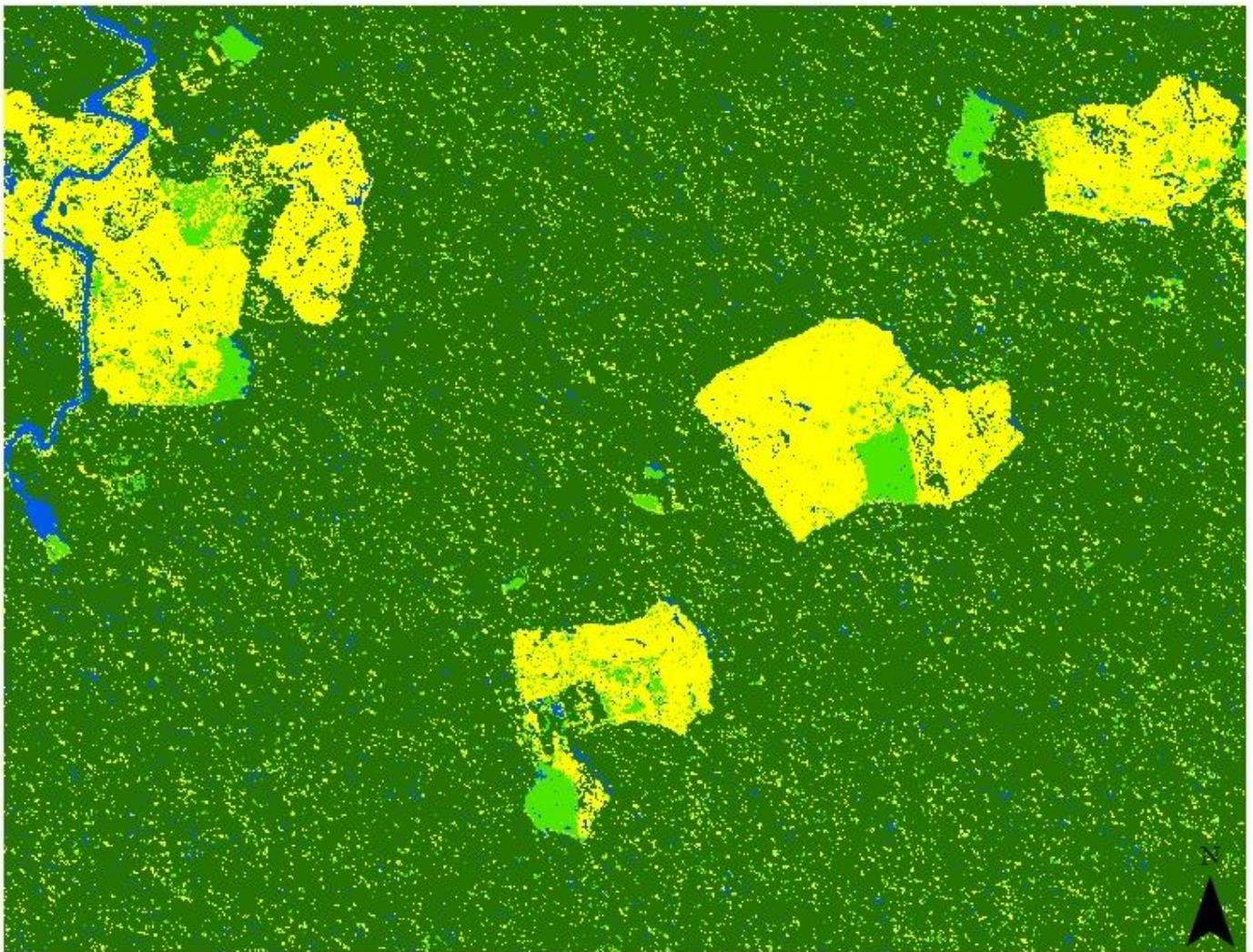
MLC

Legend

-  Forest
-  Open field
-  Scrubland
-  Water

2.5m resolution

Original bands: Green, red, NIR



0 0.4 0.8 1.6 Kilometers

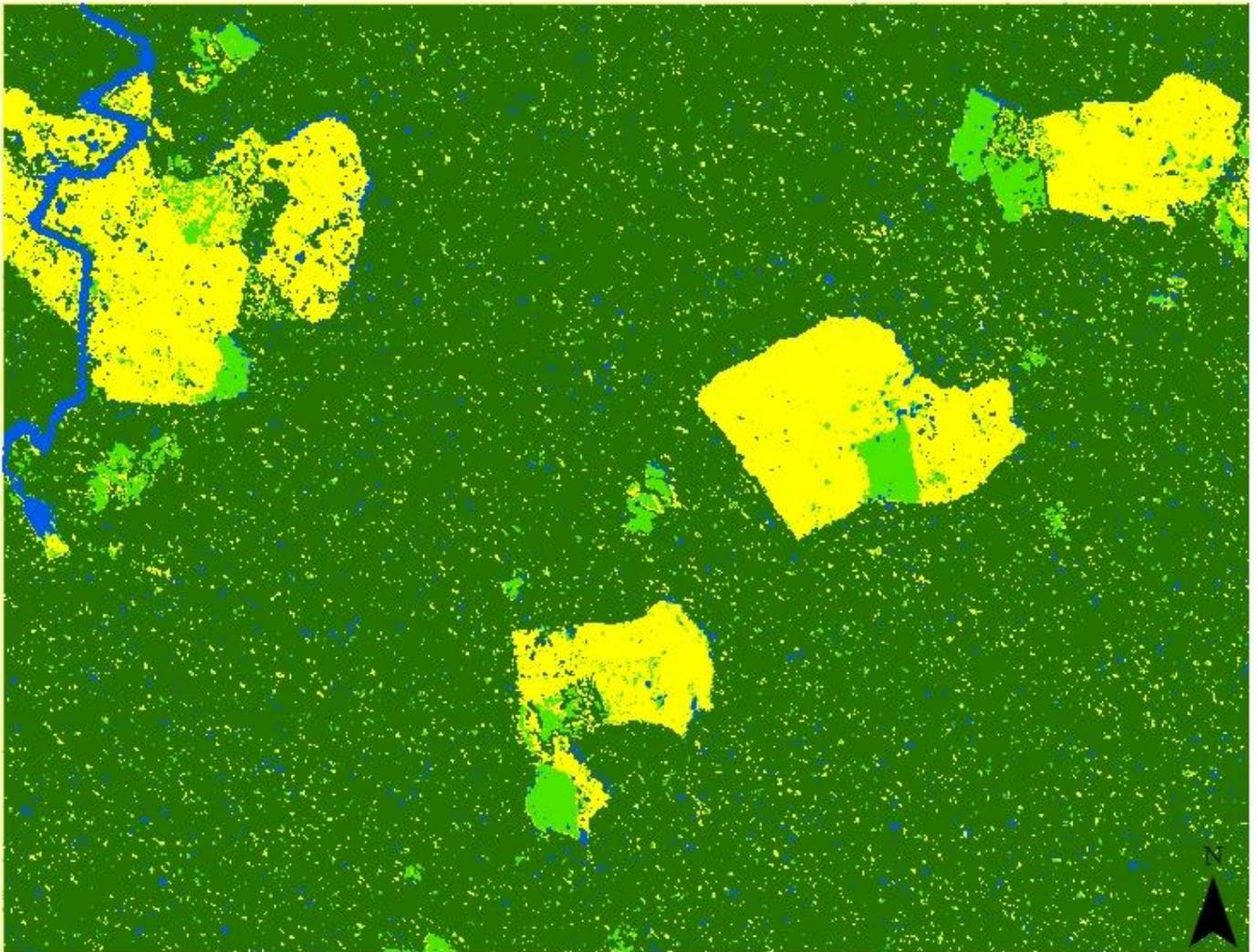
MLC

2.5m resolution

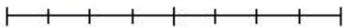
Legend

- Forest
- Open field
- Scrubland
- Water

Original bands: Green, red, NIR
Texture: Contrast measure of green and NIR,
entropy measure of red



0 0.4 0.8 1.6 Kilometers



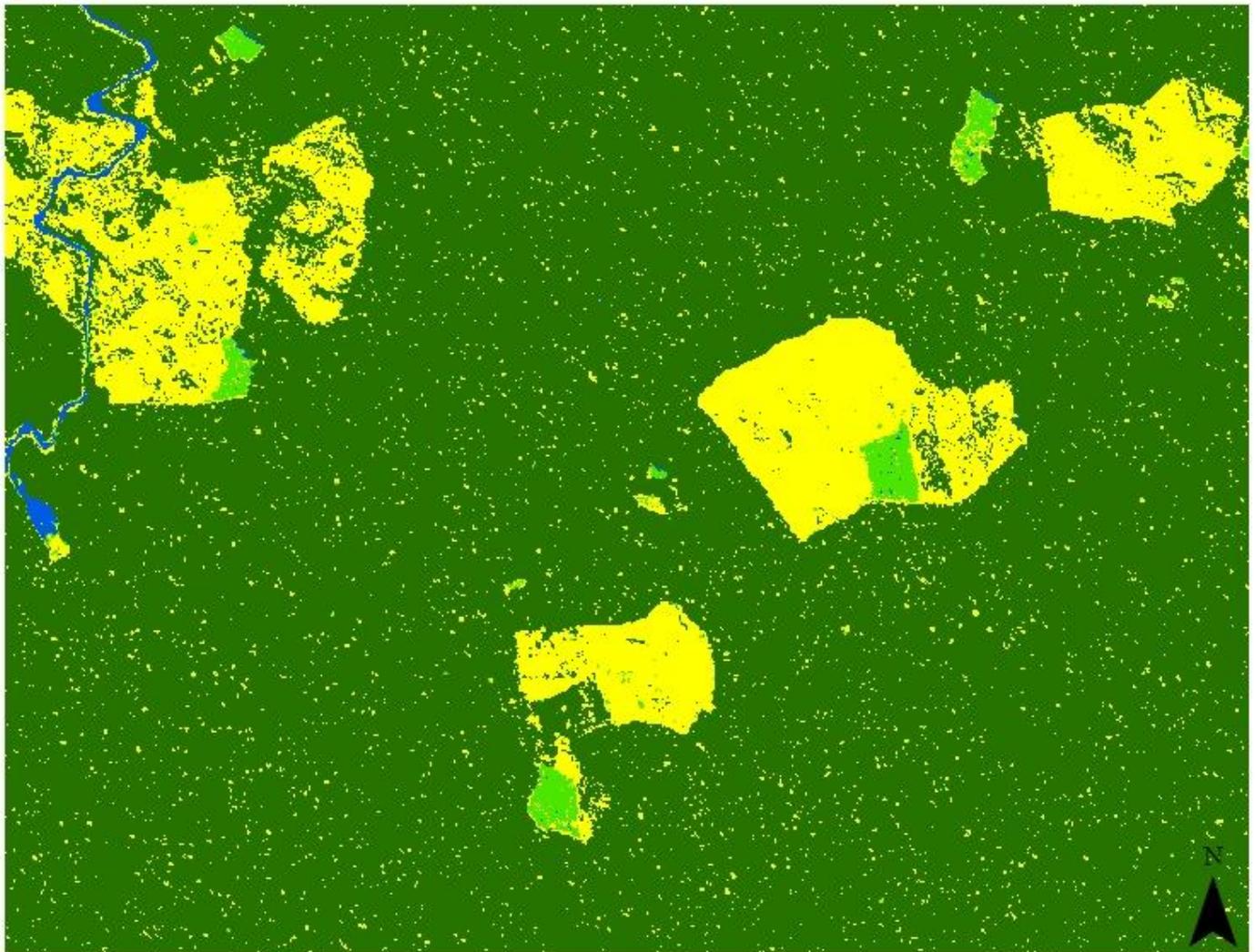
SVM

Legend

- Forest
- Open field
- Scrubland
- Water

2.5m resolution

Original bands: Green, red, NIR



0 0.4 0.8 1.6 Kilometers



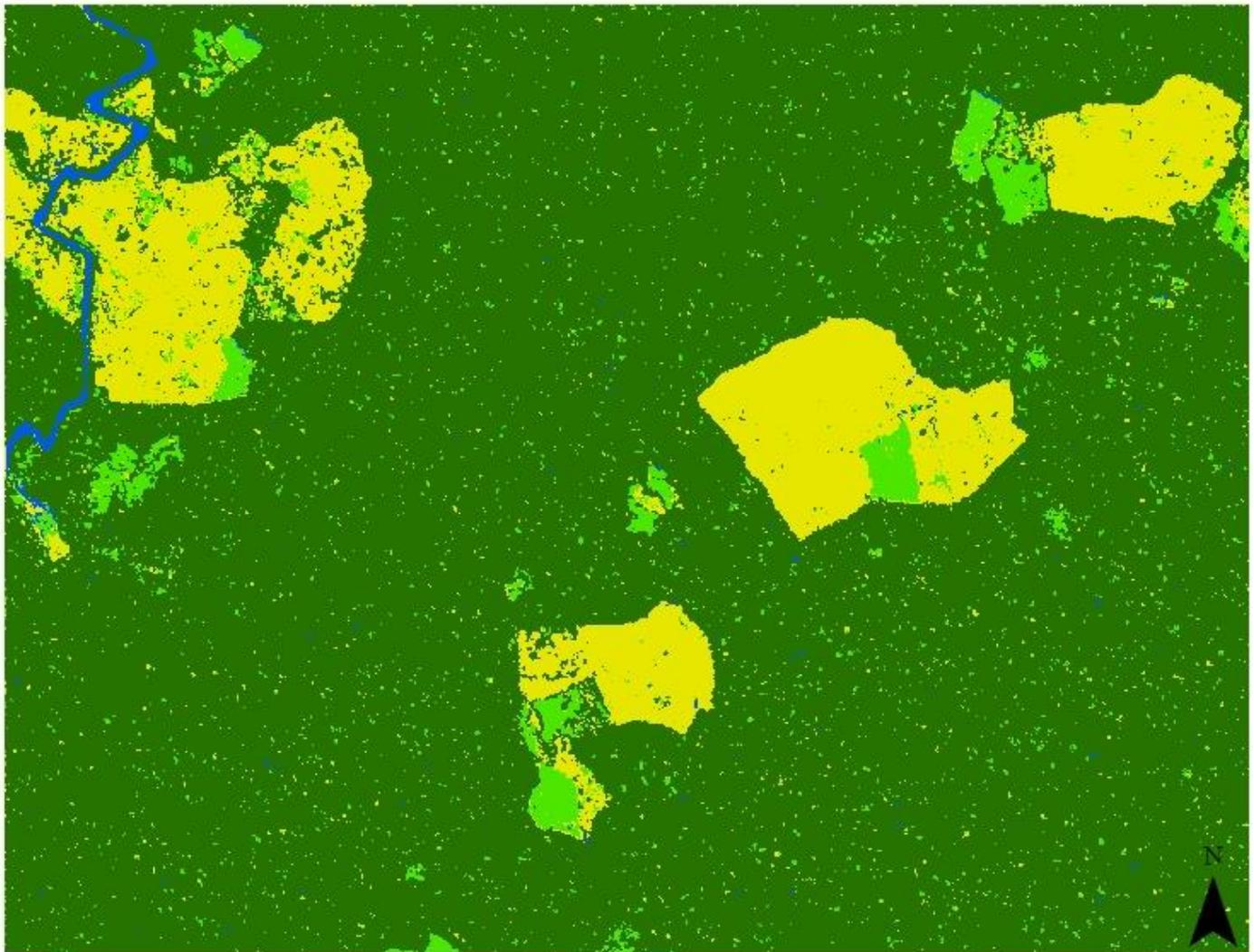
SVM

Legend

- Forest
- Open field
- Scrubland
- Water

2.5m resolution

Original bands: Green, red, NIR
Texture: Contrast measure of green and NIR,
entropy measure of red



0 0.4 0.8 1.6 Kilometers



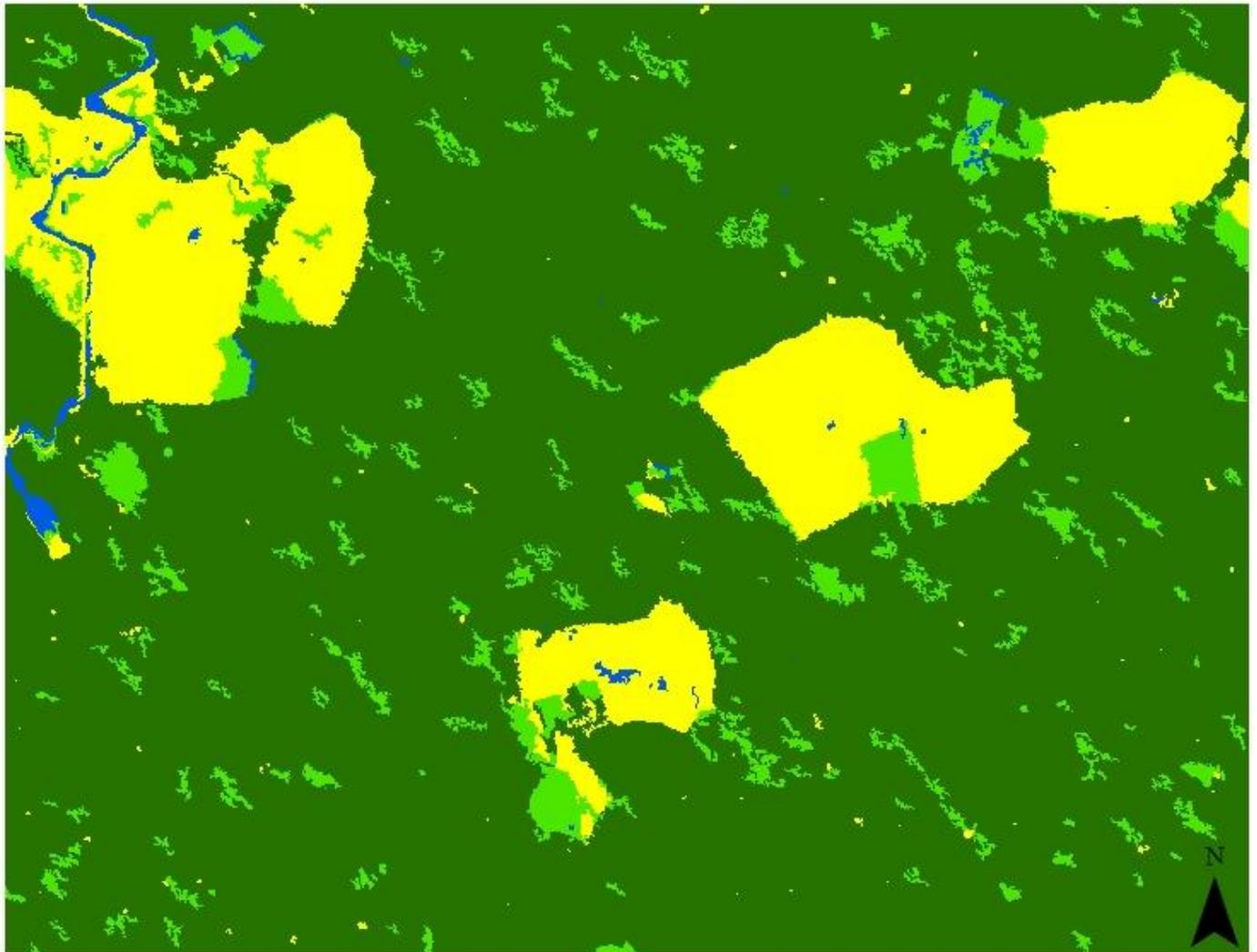
KNN

2.5m resolution

Original bands: Green, red, NIR

Legend

- Forest
- Open field
- Scrubland
- Water



0 0.4 0.8 1.6 Kilometers



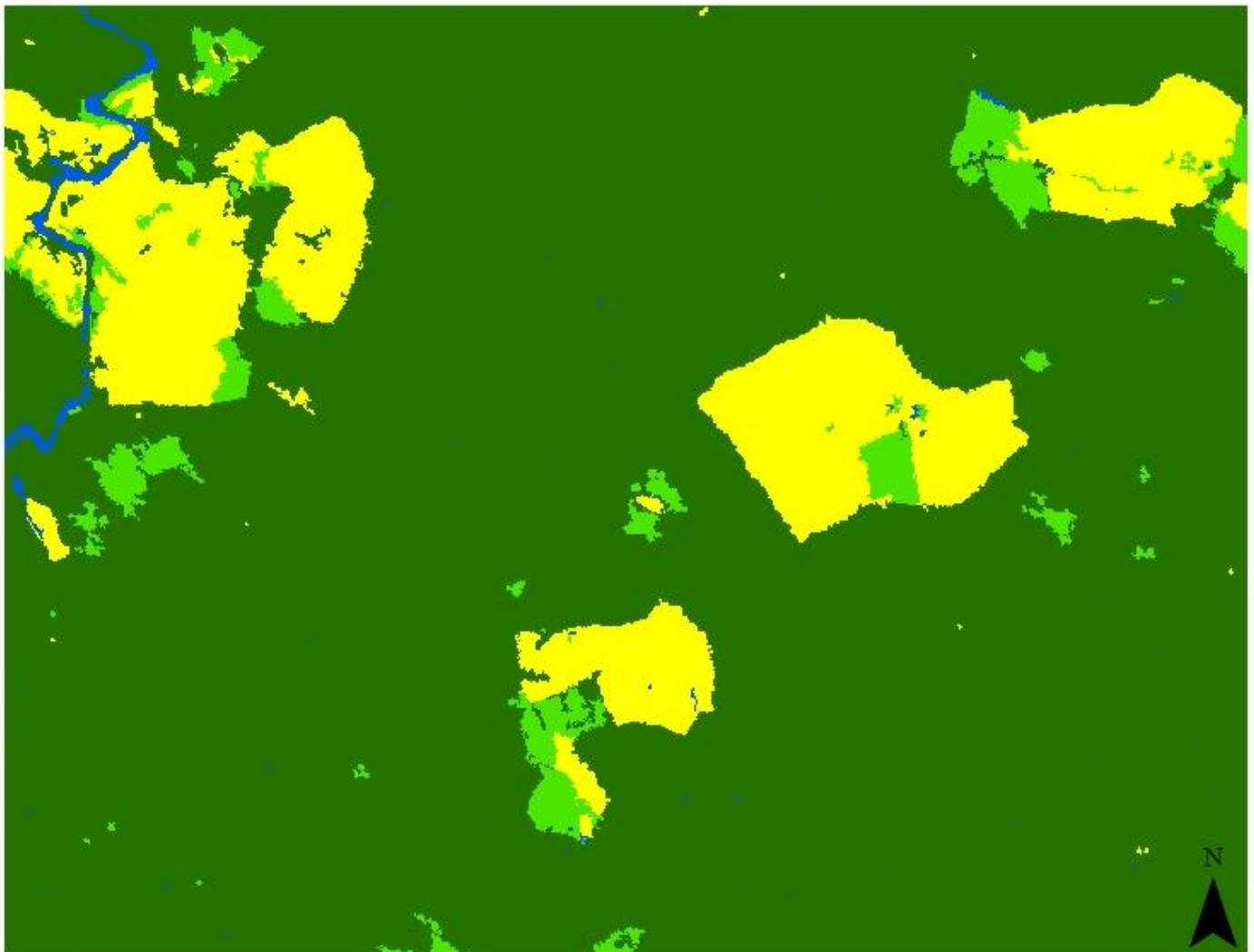
KNN

2.5m resolution

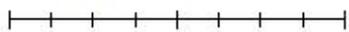
Original bands: Green, red, NIR
Texture: Contrast measure of green and NIR,
entropy measure of red

Legend

- Forest
- Open field
- Scrubland
- Water



0 0.4 0.8 1.6 Kilometers



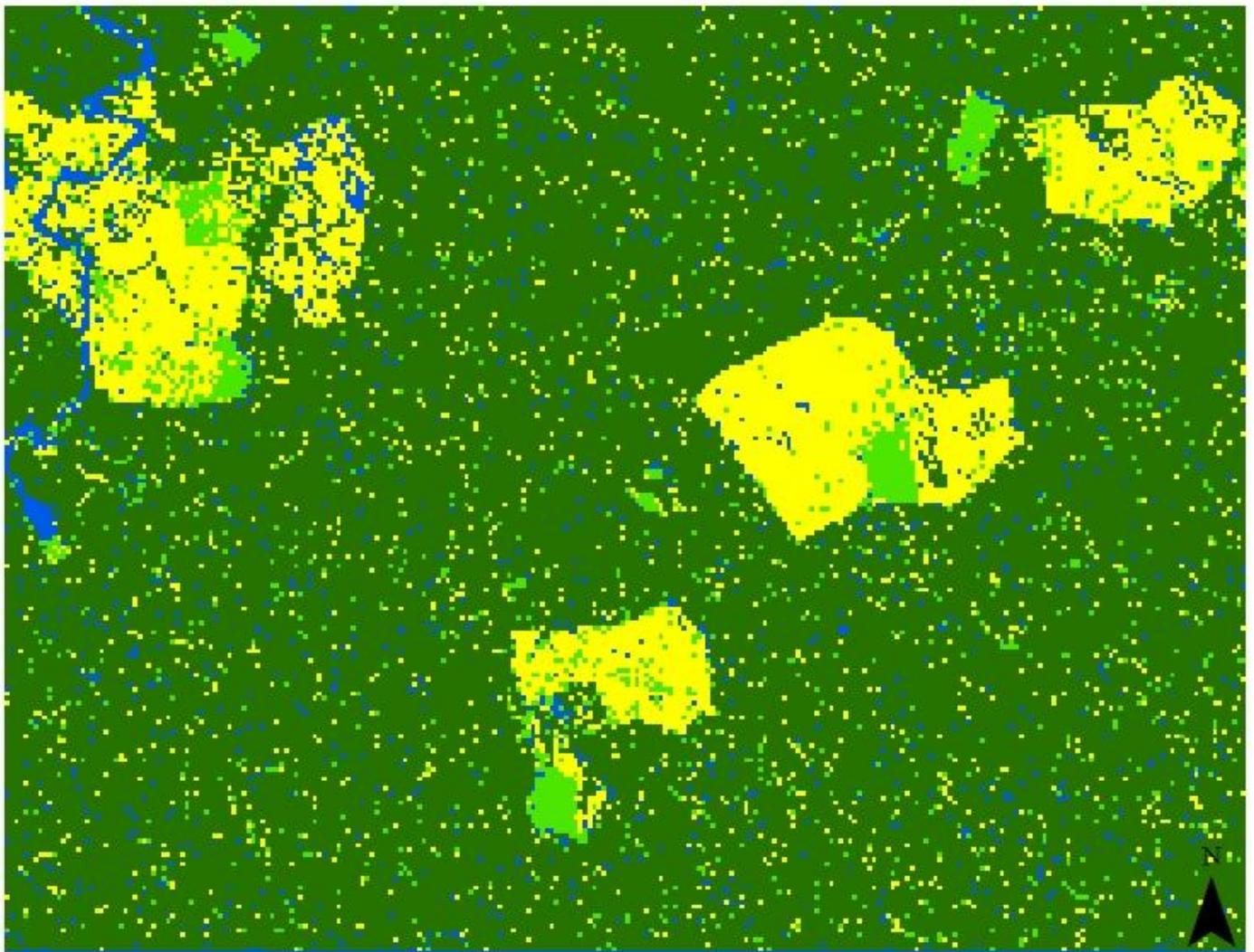
MLC

25m resolution

Original bands: Green, red, NIR

Legend

- Forest
- Open field
- Scrubland
- Water



0 0.4 0.8 1.6 Kilometers



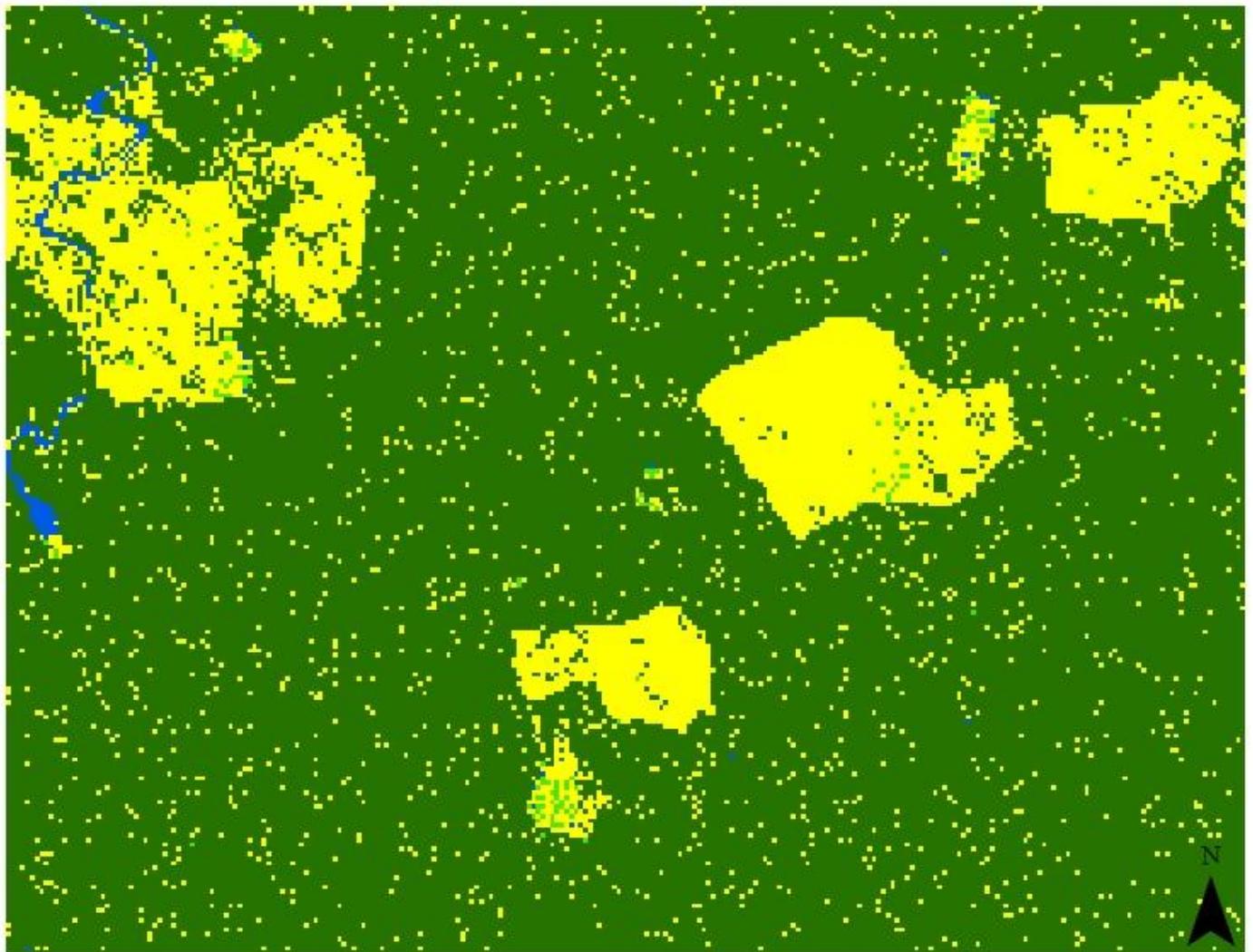
SVM

25m resolution

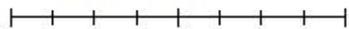
Original bands: Green, red, NIR

Legend

- Forest
- Open field
- Scrubland
- Water



0 0.4 0.8 1.6 Kilometers



8.3 Location



Figure 11. Site location.

Institutionen för naturgeografi och ekosystemvetenskap, Lunds Universitet.

Student examensarbete (Seminarieuppsatser). Uppsatserna finns tillgängliga på institutionens geobibliotek, Sölvegatan 12, 223 62 LUND. Serien startade 1985. Hela listan och själva uppsatserna är även tillgängliga på LUP student papers (<https://lup.lub.lu.se/student-papers/search/>) och via Geobiblioteket (www.geobib.lu.se)

The student thesis reports are available at the Geo-Library, Department of Physical Geography and Ecosystem Science, University of Lund, Sölvegatan 12, S-223 62 Lund, Sweden. Report series started 1985. The complete list and electronic versions are also electronic available at the LUP student papers (<https://lup.lub.lu.se/student-papers/search/>) and through the Geo-library (www.geobib.lu.se)

- 335 Fei Lu (2015) Compute a Crowdedness Index on Historical GIS Data- A Case Study of Hög Parish, Sweden, 1812-1920
- 336 Lina Allesson (2015) Impact of photo-chemical processing of dissolved organic carbon on the bacterial respiratory quotient in aquatic ecosystems
- 337 Andreas Kiik (2015) Cartographic design of thematic polygons: a comparison using eye-movement metrics analysis
- 338 Iain Lednor (2015) Testing the robustness of the Plant Phenology Index to changes in temperature
- 339 Louise Bradshaw (2015) Submerged Landscapes - Locating Mesolithic settlements in Blekinge, Sweden
- 340 Elisabeth Maria Farrington (2015) The water crisis in Gaborone: Investigating the underlying factors resulting in the 'failure' of the Gaborone Dam, Botswana
- 341 Annie Forssblad (2015) Utvärdering av miljöersättning för odlingslandskapets värdefulla träd
- 342 Iris Behrens, Linn Gardell (2015) Water quality in Apac-, Mbale- & Lira district, Uganda - A field study evaluating problems and suitable solutions
- 343 Linnéa Larsson (2015) Analys av framtida översvämningsrisker i Malmö - En fallstudie av Castellums fastigheter
- 344 Ida Pettersson (2015) Comparing *Ips typographus* and *Dendroctonus ponderosus* response to climate change with the use of phenology models
- 345 Frida Ulfves (2015) Classifying and Localizing Areas of Forest at Risk of Storm Damage in Kronoberg County
- 346 Alexander Nordström (2015) Förslag på dammar och skyddsområde med hjälp av GIS: En studie om löv- och klockgroda i Ystad kommun, Skåne
- 347 Samanah Seyedi-Shandiz (2015) Automatic Creation of Schematic Maps - A Case Study of the Railway Network at the Swedish Transport Administration
- 348 Johanna Andersson (2015) Heat Waves and their Impacts on Outdoor Workers – A Case Study in Northern and Eastern Uganda

- 349 Jimmie Carpman (2015) Spatially varying parameters in observed new particle formation events
- 350 Mihaela – Mariana Tudoran (2015) Occurrences of insect outbreaks in Sweden in relation to climatic parameters since 1850
- 351 Maria Gatzouras (2015) Assessment of trampling impact in Icelandic natural areas in experimental plots with focus on image analysis of digital photographs
- 352 Gustav Wallner (2015) Estimating and evaluating GPP in the Sahel using MSG/SEVIRI and MODIS satellite data
- 353 Luisa Teixeira (2015) Exploring the relationships between biodiversity and benthic habitat in the Primeiras and Segundas Protected Area, Mozambique
- 354 Iris Behrens & Linn Gardell (2015) Water quality in Apac-, Mbale- & Lira district, Uganda - A field study evaluating problems and suitable solutions
- 355 Viktoria Björklund (2015) Water quality in rivers affected by urbanization: A Case Study in Minas Gerais, Brazil
- 356 Tara Mellquist (2015) Hållbar dagvattenhantering i Stockholms stad - En riskhanteringsanalys med avseende på långsiktig hållbarhet av Stockholms stads dagvattenhantering i urban miljö
- 357 Jenny Hansson (2015) Trafikrelaterade luftföroreningar vid förskolor – En studie om kvävedioxidhalter vid förskolor i Malmö
- 358 Laura Reinelt (2015) Modelling vegetation dynamics and carbon fluxes in a high Arctic mire
- 359 Emelie Linnéa Graham (2015) Atmospheric reactivity of cyclic ethers of relevance to biofuel combustion
- 360 Filippo Gualla (2015) Sun position and PV panels: a model to determine the best orientation
- 361 Joakim Lindberg (2015) Locating potential flood areas in an urban environment using remote sensing and GIS, case study Lund, Sweden
- 362 Georgios-Konstantinos Lagkas (2015) Analysis of NDVI variation and snowmelt around Zackenberg station, Greenland with comparison of ground data and remote sensing.
- 363 Carlos Arellano (2015) Production and Biodegradability of Dissolved Organic Carbon from Different Litter Sources
- 364 Sofia Valentin (2015) Do-It-Yourself Helium Balloon Aerial Photography - Developing a method in an agroforestry plantation, Lao PDR
- 365 Shirin Danephash (2015) Evaluation of Standards and Techniques for Retrieval of Geospatial Raster Data - A study for the ICOS Carbon Portal
- 366 Linnea Jonsson (2015) Evaluation of pixel based and object based classification methods for land cover mapping with high spatial resolution satellite imagery, in the Amazonas, Brazil.