

Thesis in geographical information technics nr 34

Looking for shrubs in an alvar

- Investigating classification of orthophotos as a way of mapping shrub species *Juniperus communis* and *Dasiphora fruticosa* on Stora Alvaret, Öland

Jorun Westman

Program in environmental engineering
Faculty of Engineering

Department of Physical Geography and Ecosystem Science
Lund University



**Jorun Westman (2022). Looking for shrubs in an alvar –
Investigating classification of orthophotos as a way of mapping shrub species *Juniperus communis* and *Dasiphora fruticosa* on Stora Alvaret, Öland**

Master degree thesis, 30 credits in Subject of degree Department of Earth and Ecosystem Sciences, Physical Geography and Ecosystem Science, Lund University.



Lund University
Faculty of Engineering

Looking for shrubs in an alvar

- Investigating classification of orthophotos as a way of mapping shrub species
Juniperus communis and Dasiphora fruticose on Stora Alvaret, Öland

EXTM05 Master thesis, 30 ECTS
Program in environmental engineering
Jorun Westman

Supervisors
Veiko Lehsten
Department of Physical Geography and Ecosystem Science

Virginia García Millán

June 22, 2022

Opponent: Marion Karlsson Faudot
Examiner: Per-Ola Olsson

Copyright © Jorun Westman, LTH

Department of Physical Geography and Ecosystem Science
Lund University
Sölvegatan 12
223 62 Lund

Telephone: 046-222 30 30
Fax: 046-222 03 21
Web: <http://www.nateko.lu.se>

Examensarbete i geografisk informationsteknik nr 34
Printed by E-tryck, E-huset, 2022

Preface

This master thesis is the result of a degree project in geographical information technology for the program environmental engineering, Faculty of Engineering, at Lund University of Technology. The degree project is written and published for the Department of Physical Geography and Ecosystem Science at Lund University.

Finally, after a lot of up and downs, hard work and breaks, and struggling with what felt like endless gigabytes of data to download, this thesis is finished! I had the opportunity to explore a for me mostly unknown subject, and I have learnt a lot. About remote sensing in general, classification of orthophotos in particular, as well as about Stora Alvaret, its ecology and history. I have spent countless hours with ArcGIS, where the most important lesson might be to restart the computer from time to time.

I want to thank my supervisor Veiko Lehsten for letting me take on this project, providing valuable guidance during the process and maybe most importantly, teaching me to think more as a scientist and a little bit less like an engineer.

This thesis could not have been made without the great support of my co-supervisor Virginia Garcia. Thank you for dedicating your time to this project and all the help with everything from technical issues to project management and thesis structure.

A lot of valuable insight on Stora Alvaret as well as additional sources were provided by Michael Tholin from Länsstyrelsen in Kalmar län. Thank you for sharing your knowledge and for answering emails so fast.

Lastly, thank you Juan for joining the field trip and holding the papers in the strong winds. And a big appreciation to my whole family and friends for the support, both during the thesis writing as well as throughout my 5 years of engineering studies coming up to this.

Jorun Westman

Malmö, June 2022

Abstract

Loss of biodiversity is an immediate threat to the planet and if no actions are taken it is expected to accelerate. The highest species richness on smaller scale is found in semi-natural temperate grasslands, with one of the documented records from Stora Alvaret on Öland, Sweden. The area is protected under the Natura 2000-network and Länstyrelsen Kalmar are responsible for monitoring while local farmers execute most of the management with support. One of the main threats to the biodiversity on Stora Alvaret is encroachment by shrubs when the grazing pressure is decreased.

The study investigates the possibility to use classification of high resolution orthophotos as an aid in monitoring if and where shrub encroachment is happening. The studies main aim is to develop a method to successfully classify the species of interest *Juniperus communis* and *Dasiphora fruticosa* with the software ArcGIS pro.

The result shows that it is possible to successfully classify the species of interest if following the proposed method. The method consists of a first classification on 5 m resolution with removed agricultural fields, to extract the forest class which then is removed from the 1 m resolution. The 1 m resolution without forest is then used as a base for the final classification.

The final result had a kappa value of 0,92 which translates to an almost perfect classification. As the study is mainly focused on the shrub species it is important to also note the producer's and user's accuracy for their classes. *Juniperus communis* had a producer's accuracy of 0,86 and user's accuracy of 0,38 while the corresponding values for *Dasiphora fruticosa* were 0,92 and 0,86. The weakness in the classification thus lies with other classes being included in the juniper pixels. The removal of forest also introduces uncertainty and the producer's and user's accuracy of that class should be considered when interpreting and using the results.

Further research should focus on improving the user's accuracy for juniper and investigating the use of change detection between classified results from different years.

Keywords: Physical Geography, Orthophoto, Stora Alvaret, Shrub, Monitoring

Table of contents

Preface.....	i
Abstract	ii
Abbreviations & definitions.....	v
List of figures	vi
List of tables	vii
1 Introduction.....	1
1.1 Aim.....	2
2 Background.....	2
2.1 Remote sensing and digital images for vegetation mapping	2
2.1.1 Resolution.....	2
2.1.2 Image classification	3
2.1.3 Related studies	4
2.2 Species of interest	5
2.2.1 Species of interest for classification	5
2.2.2 Species of interest on Stora Alvaret	6
3 Materials and method.....	7
3.1 Study area.....	7
3.2 Data	8
3.2.1 Orthophotos	8
3.2.2 Ancillary data.....	8
3.2.3 Field data collection	9
3.3 Methodology	10
3.3.1 Pre-processing	12
3.3.2 Classification.....	14
3.3.3 Post-processing	15
3.3.4 Validation.....	15
4 Result.....	17
4.1 Shrub and tree species	17
4.2 Classification.....	17
4.2.1 Classification at 0.16 m resolution	17
4.2.2 Classification at 4.8 m resolution	21
4.2.3 Classification at 0.96 m resolution	22
4.2.4 Forest removal.....	24
4.2.5 Classifying without forest.....	24

4.2.6	Summary of results.....	26
4.3	Proposed method.....	26
4.3.1	Overview of final classification.....	28
5	Discussion.....	29
6	Conclusion.....	33
7	References.....	34
	Appendix.....	38

Abbreviations & definitions

EMR – Electromagnetic radiation

GIS – Geographic Information System

LiDAR – Light detection and Ranging (*Remote sensing technique*)

ML – Maximum likelihood (*Algorithm*)

nDSM – normalised Digital Surface Model

RGB – Red Green Blue

RF – Random Forest (*Algorithm*)

Geoprocessing tool – function within a GIS software that preforms an operation on a dataset. In the report referenced as *tool*.

List of figures

Figure 2-1 Pictures of the shrub species of interest. (a) Closeup of <i>Dasiphora fruticosa</i> (b) <i>Juniperus communis</i> (c) Landscape view of the two species. Images: Jorun Westman.....	5
Figure 3-1 Overview of location of study area (a) Study area on Öland marked by black box (b) Location of the study area within the black box marked with dashed line	7
Figure 3-2 Location of field study areas on Stora Alvaret.	10
Figure 3-3 Illustration of the iterative approach to classification used in the study.	10
Figure 3-4 Illustration of the order of the classifications and changes performed to the prepared area in the study.....	11
Figure 3-5 Illustration of the order of the classifications and changes performed to the prepared area in the study.....	12
Figure 3-6 Small part of Stora Alvaret displayed with four different pixel resolutions to visualise the change. (a) 0.16 m (b) 0.96 m (c) 4.8 m (d) 9.6 m. Scale 1:2000.....	12
Figure 3-7 Location of tiles used for training and validation samples in the study.	14
Figure 4-1 Display of orthophoto and classification result. (a) Orthophoto 0.16 m (b) Classification 0.16 m. Scale: 1:20 000. Orthophoto © Lantmäteriet.	18
Figure 4-2 (a) Detail of orthophoto 0.16 m (b) Detail of classification 0.16 m. Scale: 1:1000. Orthophoto © Lantmäteriet.	18
Figure 4-3 Display of orthophoto and classification result. (a) Orthophoto 9.6 m (b) Classification 9.6 m. Scale 1:20 000. Orthophoto © Lantmäteriet	19
Figure 4-4 Display of orthophoto and classification result. (a) Orthophoto 0.16 m (b) Classification 0.16 m w/o fields. Scale 1:20 000. Orthophoto © Lantmäteriet	20
Figure 4-5 Display of orthophoto and classification result. (a) Orthophoto 4.8 m (b) Classification 4.8 m w/o fields Scale 1:20 000. Orthophoto © Lantmäteriet	21
Figure 4-6 Display of orthophoto and classification result. (a) Orthophoto 0.96 m (b) Classification 0.96 m w/o fields. Resolution: 1:20 000. Orthophoto ©Lantmäteriet	23
Figure 4-7 Display of results from post processing with generalization method. (a) 4.8 m classification (b) generalization of 1 pixel (c) generalization of 10 pixels (d) generalization of 20 pixels. Resolution 1:10 000. Orthophoto ©Lantmäteriet	24
Figure 4-8 Display of orthophoto and classification result. (a) Orthophoto 0.96 m (b) Classification 0.96 m w/o fields and forest. Resolution: 1:20 000. Orthophoto ©Lantmäteriet.....	25
Figure 4-9 Illustration of workflow of proposed method for classification in the study.	27
Figure 4-10 Classification of the five tiles at 0.96 m resolution w/o forest together with the forest class from the 4.8 m resolution classification. Scale 1:180 000	28

List of tables

Table 3-1: Technical details of orthophotos used in this study produced by Lantmäteriet. (Lantmäteriet, 2019)	8
Table 3-2 The land cover types used in the study within the green code system. (Jordbruksverket, 2021b).	9
Table 3-3 Interpretation of Kappa index (Sim & Wright, 2005).	16
Table 4-1 Shrub and tree species identified in the field study.....	17
Table 4-2 Training sample set for classification of 0.16 m pixel resolution orthophoto with presented with number of polygons and corresponding area for each class.	17
Table 4-3 Accuracy assessment for 0.16 m	18
Table 4-4 Training sample set for 9.6 m with number of polygons and corresponding area.	19
Table 4-5 Accuracy assessment for 9.6 m	19
Table 4-6 Training sample set for 9.6 m w/o field with number of polygons and corresponding area.	20
Table 4-7 Accuracy assessment for 9.6 m w/o field.....	20
Table 4-8 Training sample set for 4.8 m with number of polygons and corresponding area.	21
Table 4-9 Accuracy assessment for 4.8 m	22
Table 4-10 Training sample set for 0.96 m with number of polygons and corresponding area.	22
Table 4-11 Accuracy assessment 0.96 m.....	23
Table 4-12 Training sample set 0.96 m w/o forest with number of polygons and corresponding area.	25
Table 4-13 Accuracy assessment for 0.96 m w/o forest	25
Table 4-14 Summary of result of classification	26
Table 4-15 Classes in the final map with their total area and the corresponding percentage of the classified area.	28

1 Introduction

Biosphere integrity, or the loss of biodiversity and extinctions, is one of the nine planetary boundaries defined by Rockström et al. (2009) and one of the two boundaries that are already assessed as beyond zone of uncertainty or high risk (Stockholm Resilience Center, n.d.).

The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) estimate that around 25 % of all species in assessed groups are threatened and without actions to reduce the drivers of biodiversity loss, it will likely further accelerate. Noticeably, the rate of extinction is already between ten to a hundred times the average during the last 10 million years. But the report also notes that the local scale of biodiversity is disappearing, with the loss of endemic species and local ecosystems, biological communities look more and more the same (IPBES, 2019).

Action to halt these processes exist on different levels. On EU level there is the EU Biodiversity Strategy, the current effective until 2030, which is implemented on a national level in Sweden (Naturvårdsverket, n.d.). A central part of the work is derived from the habitats directive together with the Natura 2000 network (Naturvårdsverket, n.d.). The network consists of protected areas, currently covering 18 % of EUs land cover, aiming at protecting the species and habitats listed in the directives (Directorate-General for Environment, n.d.).

A study that looked at the highest species richness in the world found that semi-natural temperate grasslands had the highest values in the smaller scale of areas less than 100 m² and one of the records were found on limestone grassland in Sweden with 12 species on an area of 10 cm² (Wilson, Peet, Dengler, & Pärtel, 2012). The study was performed in Stora Alvaret (van der Maarel; Sykes, 1993), one of the Natura 2000 areas in Sweden. It is an area of around 26 000 ha of limestone grassland, alvar, habitats situated on a plateau on the southern part of the island of Öland. It is the world's largest continuous alvar constituting 27 % of the total alvar lands in the world (Ekstam & Forshed, 2002 see Rosén & Bakker, 2005).

Alvar ecosystems are known to have a high plant species richness that often include both endemic and rare species (Rosén & van der Maarel, 2000). In 2016, around 580 species found on Stora Alvaret were on the Swedish national red list for threatened species which constitutes 14 % of the total amount on the list (Tholin, 2016; SLU Artdatabanken, 2015). The area is home to species that normally are found in such different habitats as mountain regions, the steppe or taiga and since the end of the latest ice age several new and sub-species have developed there (Tholin, 2016).

Most of the area has been kept open for thousands of years due to the special climate and soil conditions in combination with human influence of wood cutting and grazing (Rosén & van der Maarel, 2000). However, during the 20th century the land use on Stora Alvaret changed, mainly due to changes in agricultural practices, and most parts were left abandoned which lead to shrub encroachment by *Juniperus communis* (juniper) and *Dasiphora fruticosa* (shrubby cinquefoil) (Maad & Rosén, 2010). Shrub encroachment cannot universally be seen as a degradation of ecosystems as it can have positive, negative and neutral effects but studies in the area showed a monotonic decrease of alvar species with Juniper cover (Eldridge et al., 2011; Rejmánek & Rosén, 1992). After shrubs have established it is possible for trees to establish in the area as well. The beginning of the encroachment process is slow, but after the first trees the speed accelerates, and open limestone grasslands can become completely overgrown within 100 years (Forslund & Lager, 2000).

Within the Natura 2000 area, 32 different species and habitat types are specified to be protected and the responsibility to create and maintain a plan for conservation falls upon the county administrative board (Länsstyrelsen). In the conservation plan from 2016 it is stated that the protection of the area

relies on management of the grasslands. The prioritized measure is to maintain a yearly grazing as well as clearing up overgrown areas in order to conserve and recreate a landscape which is dominated of open fields free from trees and bushes (Tholin, 2016).

Most of the land is managed by farmers and is funded through EUs agricultural aid. If additional clearing is needed, the county administrative board funds and executes it. To conserve the area poses many challenges where one is to balance the grazing at sufficient level to prevent shrub encroachment while not putting a too high strain on the flora and insect populations. Other factors such as the latest years of summer droughts in the area, which affects the agriculture as well as plant communities, has added to the complexity (M. Tholin, personal communication, 16 February, 2022).

To monitor the spread of tree and shrub populations, the county administrative board compare orthophotos, orthogonal projected aerial photos, over time, from the 1960s up to today. Field visits are then made to places of interest to make detailed management decisions. Inventories are made for typical species in all Natura 2000 areas, as well as more directed efforts for endangered species (M. Tholin, personal communication, 16 February, 2022).

Several studies show that classifying vegetation by machine learning algorithms on orthophotos is possible, and that the outcome can have a substantial or higher accuracy according to the scale defined by Sim & Wright (2005) (Ayhan & Kwan, 2020; Hellesen & Matikainen, 2013; Mora et al., 2021). Additionally, Boswell et al. (2017) shows that digital classification can be just as accurate as field observations.

With the objective of investigating possible ways of monitoring shrub encroachment over large areas this study continues the research on digital classification of vegetation and land cover on orthophotos.

1.1 Aim

The general aim of the thesis is to investigate the applicability of classification of high resolution orthophotos as a method of mapping vegetation on alvar habitats on Öland, Sweden.

The study looks at the feasibility of identifying the shrub species *Juniperus communis* (juniper) and *Dasiphora fruticosa* (shrubby cinquefoil) from other vegetation and ground layers.

2 Background

2.1 Remote sensing and digital images for vegetation mapping

Remote sensing in environmental science commonly refers to recording of electromagnetic radiation (EMR) that is emitted or reflected from objects on the Earth surface, via a sensor on a satellite or other air born devices (Khorram, Koch, van der Wiele, & Nelson, 2012). All material and objects emit and reflect EMR in specific patterns that are called spectral signatures or profiles (Khorram et al., 2012). Digital images are a common way in which remotely sensed data is produced.

2.1.1 Resolution

Digital images can be described by its *spatial*, *spectral*, *radiometric*, and *temporal* resolution. *Spatial* resolution refers to the detail level of the image, described in the ground area captured in each pixel. *Spectral* resolution is the width of interval and/or the number of different bands the sensor has. The *radiometric* resolution is a measure of the ability to detect difference in intensity of the EMR. And lastly, the interval in which a sensor record data on the same location is the *temporal* resolution (Khorram et al., 2012).

Spectral resolution can be divided into single-band, multispectral and hyperspectral images based on the number of bands included (Khorram et al., 2012). The exact boundaries varies but some definitions of multispectral are “3 to 10”, “less than 15” or simply just “multiple” bands. Hyperspectral images can include up to 100s of bands where each band also typically has a more narrow spectral interval (GISGeography, 2021; Khorram et al., 2012; STARS, n.d.).

The spatial resolution is measured as the length of the pixel side on the ground and it can vary between a kilometre down to a few centimetres (SEOS, n.d.). The appropriate resolution for classification differs between the target classes. If the pixels are smaller than the structure to be classified, it will increase the variability (Curran & Atkinson, 2002). Woodcock & Strahler (1987) found that for forest the best classification would be achieved above 60-80 m as a lower resolution increases the local variance, or at a high resolution, below 3 m, which instead would classify individual trees. For agriculture the opposite relationship could be expected, where the best classification occurs on a higher resolution. On a smaller scale however, Ayhan & Kwan (2020) got similar classifications accuracies for the same classes on both 1 and 10 m.

2.1.2 Image classification

One way of processing the information in digital images is classification, where the spectral value of a pixel or an object (multiple pixels) is used to pair it to a class. Classifying images by land use and land cover is one application in environmental remote sensing and is commonly used in research to study change in vegetation (Khorram et al., 2012).

Digital classification can be divided into supervised and unsupervised methods, based on the way the classes are created. In unsupervised classification the algorithm clusters homogeneous values into classes. Supervised methods however require previous knowledge of the area to select categories and samples of them (Chuvieco, 2016). In this study supervised classification was used. This because the aim of the study is to sufficiently map two specific classes, which then can be defined in supervised classification.

For supervised classification training samples are created that should represent each class identified in the area. This can be done with the help of either field work, previous knowledge of the area or external sources. These samples are used to train the algorithm to recognize the different classes to be able to classify the training data (Chuvieco, 2016).

The classes in the classification scheme should be distinguishable from each other and therefore chosen with respect to the spectral and spatial resolution of the image. They should also be relevant to the end user (Khorram et al., 2012).

Choosing and creating the training samples should be done in a systematic way. Khorram et al. (2012) lists four rules to follow. (1) The number of samples should be at least three times the number of classes, (2) they should include the spectral distribution of each class, (3) each site should be as homogenous as possible, e.g. only include pixels with similar spectral values, and (4) they should not be clustered in a part of the area, and therefore either randomly distributed or systematically placed over the whole area.

Supervised classification algorithms

There are several algorithms that can be used in supervised classification, of which some are: Support Vector Machine (SVM), Maximum likelihood (ML) and Random forest (RF). Mountrakis, Im, & Ogole (2011) found that SVMs have a great ability to generalize from a limited set of training data, which is common in remote sensing analysis, compared to ML, in their review article including high to medium spatial resolution from both satellite and airborne sensors. Random forest was in a study

by Pal (2005), that was performed on satellite images with 30 m resolution, found to have the same level of classification accuracy and equal training time to SVMs and in the same time easier to define as it requires fewer user-defined parameters. In a study about training set size effects on land cover classification Ramezan et al. (2021) found that RF had the highest overall accuracy as well as the smallest decrease of accuracy with the training set size in a study performed with 1 m spatial resolution orthophotos. Ramezan et al. (2021) highlights that good practice should include testing different machine-learning methods since the accuracy vary but if this is not feasible Random Forest showed the best overall result. In a study comparing SVM and RF algorithms for land classification of hyperspectral images with 20 m spatial resolution the result for the SVM showed a slightly higher overall accuracy but instead had much longer computational time (Abe, Olugbara, & Marwala, 2014). This result support the findings of Bosch, Zisserman, & Muñoz (2007) in a study about image classification, that RFs have comparable performance to SVMs while reducing the computational cost. For these reasons, Random Forest was selected for classification in this study.

2.1.3 Related studies

To get an understanding of the possibilities and constraints of shrub classification several studies are presented in this section. All studies use supervised classification, however the spatial and spectral resolution, type of algorithms used, number of classes, vegetation and geographical location differ.

In a study by Boswell et al (2017) on rangeland monitoring they found that vegetation cover estimation produced from high resolution (0.25 m pixel resolution) Red-Green-Blue (RGB) images did not differ from estimations based on field observations. Using maximum likelihood, it had an overall accuracy of 91 % and a kappa value of 0.88, and the authors argue that spectral classifications techniques can be used for vegetation classification with satisfactory results when the classes are broad. In the study tree, shrub, herbaceous and surface was used as functional groups.

Hellesen & Matikainen (2013) highlights that classifying trees and shrubs in grassland is a challenging task as the spectral reflection is similar between the classes. Although, using object-based classification with a classification and regression trees algorithm on colour-infrared high resolution (0.5 m) orthophotos, they got an overall accuracy of 82 % which increased to 97 % when including LiDAR data in the form of normalised digital surface model (nDSM) as supplementary data. The study used few and broad classes with one for shrubs and trees, one for buildings and one for ground, covering both vegetated and bare areas.

A study by Ayhan & Kwan (2020) investigated classification of three, shrub and grass from RGB images with the deep learning algorithm DeepLabV3+. The study included two different sites, one with 10 m resolution images and one with 1 m. For the 10 m resolution the average accuracy for the three vegetation classes were only 59 % when classifying the whole area containing 8 classes including other land cover classes such as wetlands and cultivated land. However, when isolating the vegetation classes, the accuracy went up to 78 %. For the 1 m resolution classification, the difference was smaller with an average accuracy for the vegetation of 75 % with all 6 classes included and 79 % for the isolated case.

A study by Mora et al. (2021) conducted on Svalbard showed that high resolution aerial photos (0.1 m and 0.2 m) with visible spectre (RGB) and near infrared bands combined worked well for identifying main vegetation groups, but would not be suitable to distinguish between separate species. The classification was made on 9 classes with 4 vegetation classes based on their dominating species. They tested four algorithms where maximum likelihood showed the best result. At 0.1 m resolution the overall accuracy was 84 % and kappa value was 0.82 while a resolution of 0.2 m slightly decreased the accuracy to 83.1 %. The main limitation in the classification was mixing

between a grass-dominated and a moss-dominated class, which lowered the overall accuracy. However, it was not seen as problematic since these communities were similar.

Two studies looked at the influence of temporal resolution. Tree species classification with hyperspectral data performed in a study by Persson, Lindberg, & Reese (2018) gave the highest accuracy with multi-temporal data from all four seasons, but the highest accuracy for a single image was produced by the one in late spring in the month of May. Similarly, investigating the variable importance in multi-temporal hyperspectral data, Abdi (2020) found that among the top 20 bands most were acquired during May and June and over half were from the near red or infrared part of the spectrum.

2.2 Species of interest

2.2.1 Species of interest for classification

The two shrub-species of interest in this study is the common juniper and shrubby cinqfoile, as these two species are identified as the main cause of shrub encroachment on Stora Alvaret (Forslund & Lager, 2000).

Juniperus communis

The common juniper, *Juniperus communis*, is a shrub in the cypress family that is spread all over Sweden. It can grow between 0.5-15 m tall and is normally found on dry open grounds (Mossberg, Stenberg, 2010).

Dasiphora fruticose

Shrubby cinqfoile, *Dasiphora fruticose*, from the rose family, is in Sweden only found wild on the islands of Öland and Gotland. It grows on moist, calcareous ground and becomes between 0.5-1 m tall (Mossberg, Stenberg, 2010).

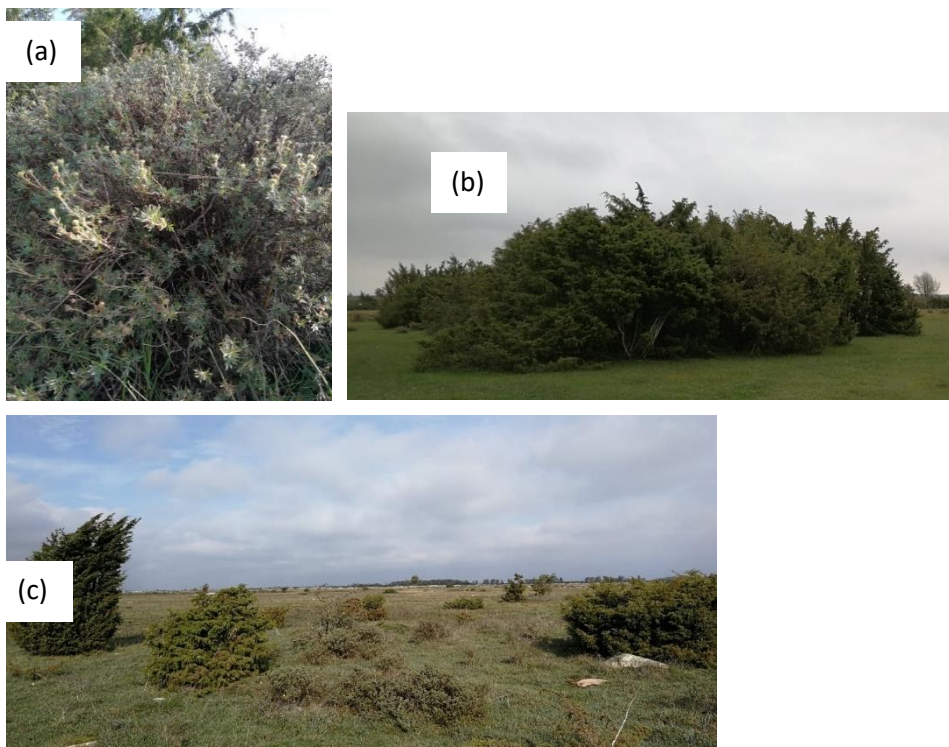


Figure 2-1 Pictures of the shrub species of interest. (a) Closeup of *Dasiphora fruticose* (b) *Juniperus communis* (c) Landscape view of the two species. Images: Jorun Westman

2.2.2 Species of interest on Stora Alvaret

32 habitats and species are specified to be protected under the Natura 2000-network. Five are plants, fifteen are birds, and there is one salamander, one snail and one butterfly specie respectively while the last nine are habitats (Tholin, 2016). The protected plant species are presented below together with their known geographical spread in Sweden.

- *Sisymbrium supinum*, a plant in the cabbage family. Mostly found on the islands of Öland and Gotland (Artfakta, n.d.-d).
- *Artemisia oelandica*, a plant in the sunflower family, endemic to Öland (Artfakta, n.d.-b).
- *Senecio jacobea ssp. gotlandicus*, a plant in the sunflower family, found on Öland and Gotland (Artfakta, n.d.-c).
- *Encalypta mutica*, a moss in the family *Encalyptaceae* scarcely found on lime-rich grounds, but more common in the mountain regions (Artfakta, n.d.-g).
- *Tortella rigens*, a moss in the family Pottiaceae, mostly found on Öland and Gotland but also occurring at other alvar habitats in Sweden (Artfakta, n.d.-f).

In the evaluation of the EU project for restoration of Stora Alvaret *Artemisia oelandica* was monitored together with six other plant species, listed below (Forslund & Lager, 2000).

- *Globularia vulgaris*, a plant in the plantain family, occurring on Öland and Gotland.
- *Anthericum ramosum*, a plant in the asparagus family, found on Öland, Gotland and in Scania.
- *Ranunculus illyricus*, belonging to the buttercup family, only found on Öland.
- *Calluna vulgaris*, in the heath family, on the mainland widespread in limestone-poor areas, but individual plants found on Öland.
- *Orchis maculate*, belonging to the orchid family, found in the South of Sweden (Artfakta, n.d.-e).
- *Dactylorhiza latifolia*, also in the orchid family, found on the Southern East coast of Sweden (Artfakta, n.d.-a).

3 Materials and method

3.1 Study area

The study area is located on the island of Öland, which is situated on the southeast coast of Sweden, in the Baltic Sea, see figure 3-1. Stora Alvaret stretches over 260 km² covering a large part of the southern part of the island (Tholin, 2016).

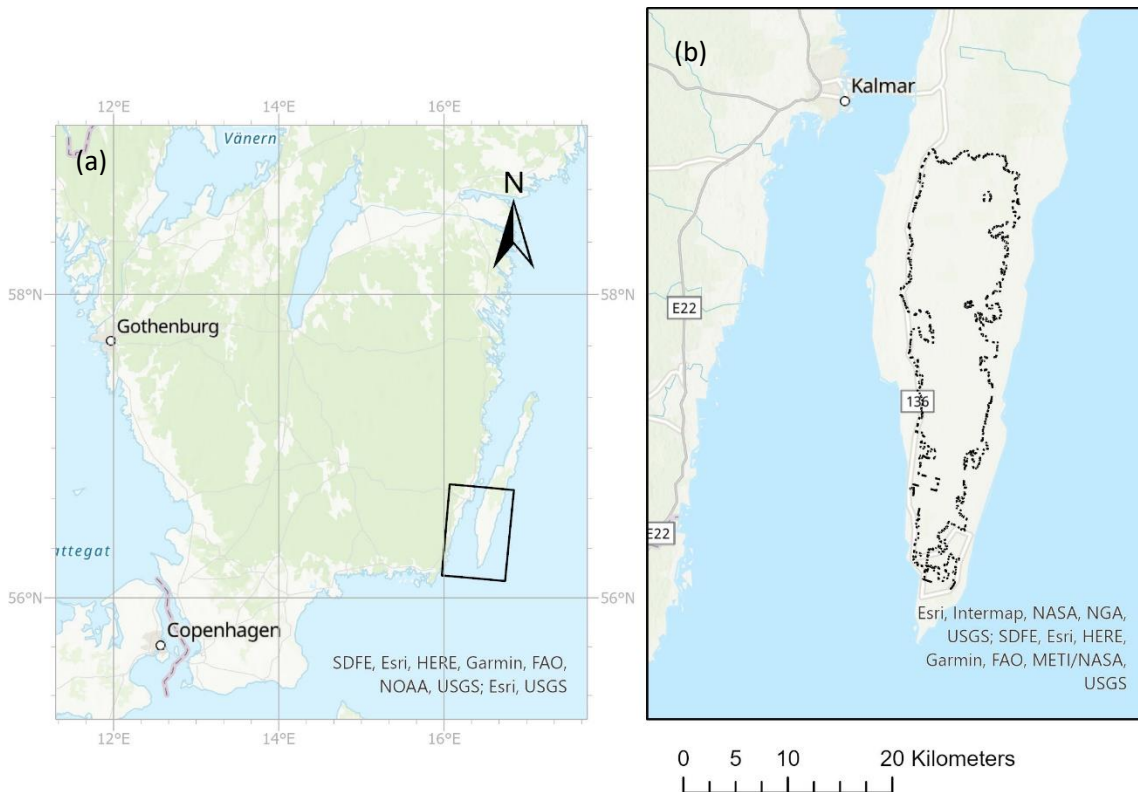


Figure 3-1 Overview of location of study area (a) Study area on Öland marked by black box (b) Location of the study area within the black box marked with dashed line

Öland has a flat landscape with the highest point only 50 meters above sea level and since the widest part of the island is no more than 20 km the climate does not vary substantially over the area (SMHI, 2021). The surrounding sea functions as a heat and cool reserve that equalizes the land temperature over the year, which results in a smaller difference in maximum and minimum temperature than the rest of Sweden ((M. Forslund, 2001). The mean temperature is -1° to -2° C in February, the coldest month, and lays between 16° to 17° C in July which is the warmest during the year (SMHI, 2021).

It is the driest region of Sweden with a mean annual precipitation of 500 mm (SMHI, 2021) but it also has one of the highest difference in annual precipitation, varying between dry and wet years (M. Forslund, 2001). With a high amount of sun hours, the evaporation rate is high which during dry years can give a desert-like climate. However, with thin or no soil layers large parts also easily get flooded and these variations in climate put a high strain on the vegetation (M. Forslund, 2001).

Stora Alvaret contains different habitat types mixed in a mosaiced structure where the ground layer is one of the influential factors. 2 % of the area consists of rocky ground while around 58 % is covered with a thin soil layer of weathered limestone. Another 20 % has a thicker, but dry, soil layer where grasses, trees and juniper can grow while the remaining 20 % consist of depressions with finer sediments, keeping the area moist and habituating grass and birch groves (M. Forslund, 2001).

Apart from abiotic factors, the human impact from cattle grazing and wood cutting has had a significant part of the ecosystem development and the resulting high biodiversity. With a decreased grazing pressure especially two species can spread and outcompete others: shrubby cinquefoil on the moist grasslands and juniper in the drier grasslands as well as in cracks on the rocky grounds (M. Forslund, 2001).

Öland, and Stora Alvaret in particular, are known for its orchids which attracts tourists, but also the grazing animals, during springtime. They grow on the alvar's dry meadows, which is where a good deal of the high biodiversity in the area stems from (Forslund & Lager, 2000).

As presented in the background (section 2.2.2) five plant species are protected on Stora Alvaret, which are directly threatened by shrub encroachment. As described in the introduction, shrub encroachment transforms habitats from alvar grounds possibly all the way to a forest which threatens some of the protected habitats. Animal species are then indirectly threatened via the transformation of habitat and loss of shelter or food. Apart from shrub encroachment, the flora and fauna on Stora Alvaret can also be threatened by too high grazing pressure, misdirected clearing and high predation (Tholin, 2016).

3.2 Data

3.2.1 Orthophotos

Orthogonal projected aerial photos, orthophotos, over Sweden are produced by Lantmäteriet, the Swedish mapping, cadastral and land registration authority and can be retrieved via the Geodata Extraction tool (GET), for use in research, provided by The Swedish University of Agricultural Sciences (Svenska lantbruksuniversitetet, n.d.).

Over the area of investigation, the island of Öland, the photos are produced every second year with the resolution, band composites and file coverage presented in Table 3-1 (Lantmäteriet, 2019). Date of acquisition and flight altitude is also included in table 3-1 which has been retrieved from the metadata of the orthophotos.

Table 3-1: Technical details of orthophotos used in this study produced by Lantmäteriet. (Lantmäteriet, 2019)

DATE AND TIME OF ACQUISITION	PIXEL RESOLUTION	PIXEL DEPTH AND TYPE	BAND COMPOSITES	FILE COVERAGE	FLIGHT ALTITUDE
2019-05-29 8.30-15.45	0.16 m/pixel	8 bit unsigned integer	Monochrome, RGB, IRF, RGBIR	2,5x2,5 km	3000 m

3.2.2 Ancillary data

To define the area and to simplify the classification two other data sources were used apart from the orthophotos: a boundary shapefile and a cadastral map of agricultural land.

Study area boundary

A shapefile of Natura 2000 areas, retrieved from the European environment agency, was used to determine the boundaries of the study area (EEA, 2021). The polygon of *Stora Alvaret* was saved as a separate file.

Agricultural fields

Some agricultural fields are situated within the Natura 2000 area and these were identified through data from the Swedish Board of Agriculture. All agricultural land is divided into blocks which are the basis of financial support to the farmers (Jordbruksverket, 2021a). The information is available in the

form of a map layer with shapefiles and can be downloaded from the agency's website (Jordbruksverket, 2021c).

The land cover type is defined by 'green codes' (Swedish: *grönkoder*) connected to different land use and crops. This data can be retrieved in the form of an excel sheet available at the Swedish Board of Agriculture's website (Jordbruksverket, 2021b).

Land cover types represented within the Natura 2000 area that were included in the analysis were *pasture*, *alvar pasture* and *pasture and hay meadow under construction* while the rest were excluded. The codes with corresponding classes in Swedish and the English translation are presented in table 3-2.

Table 3-2 The land cover types used in the study within the green code system. (Jordbruksverket, 2021b).

CODE	SWEDISH DEFINITION	ENGLISH TRANSLATION
52	Betesmark (ej åker)	Pasture (not fields)
56	Alvarbete	Alvar pasture
95	Betesmark och slåtteräng under restaurering	Pasture and hay meadow under restoration

The three aforementioned classes were excluded from the shapefile in order to clip out the remaining classes from the orthophoto, explained in section 3.3.1 *Extracting orthophotos*. It was done by selecting and saving the wanted classes as a new file.

3.2.3 Field data collection

In order to collect training and validation samples for the digital classification of the alvar, a field study was performed. The goal was to record the coordinates of shrubs of interest and other map classes. The area was visited on two subsequent days during early autumn 2021, 1/10 and 2/10.

The field observations took place on three different sites spread across the area of Stora Alvaret. The sites were chosen with three criteria: (1) to represent the diversity of vegetation and land cover present over Stora Alvaret, (2) to be situated in different geographical parts and (3) to be located next to a road for easy access.

The main areas are three 2.5x2.5 km squares corresponding to the extent of the retrieved orthophotos and are shown in figure 3-2. The indices for the three sites with the index by Lantmäteriet are 624_59_5000, 626_50_5000 and 626_59_5025 from south to north.

The samples were collected to represent the variety of ground structures and vegetation at the site. A few samples of each structure or specie found were marked at each site. They were then used as the base to create training and validation samples, explained further in section 3.3.1 *Preparation of training and validation data*. The amount and size of training and validation samples differed for the different spatial resolutions and are presented at each sub section in chapter 4.2.

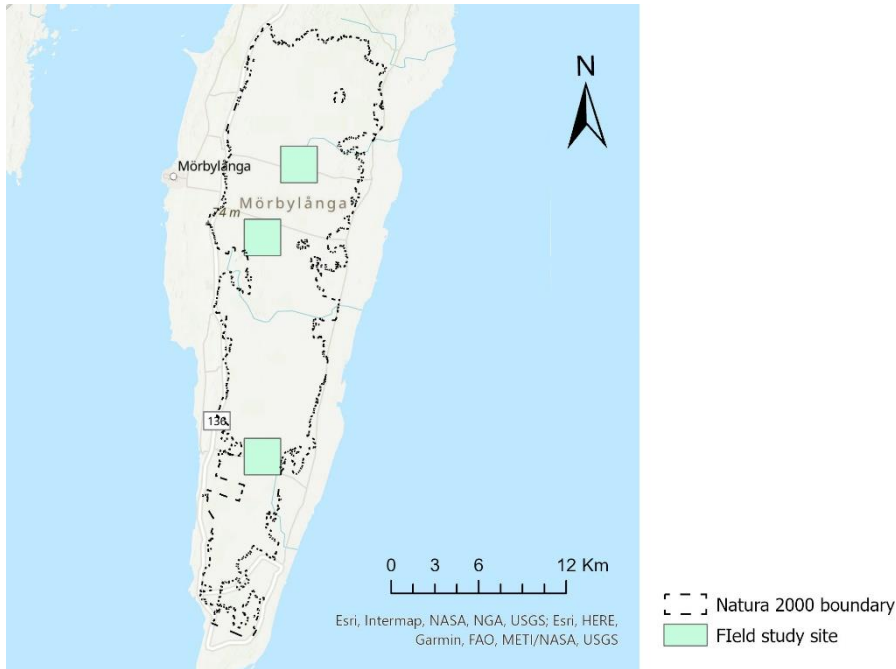


Figure 3-2 Location of field study areas on Stora Alvaret.

The identification of species was done with the aid of a flora (Mossberg, Stenberg, 2010) and the mobile app *PlantNet* (PlantNet, 2021) which identifies species based on pictures uploaded by the user.

With the help of *GPS waypoint finder* (keuwsoft, 2021), a mobile app to locate GPS points, and the coordinates on orthophotos of the study site, structures on the orthophotos were compared to the vegetation or land cover type on site. The identified types were then noted on the printed orthophotos with the associated area covered.

3.3 Methodology

The thesis has an iterative approach to the classification process, visualised in figure 3-3, with consecutive tests done over different number of tiles, different spatial resolution and with extraction of specific land covers.

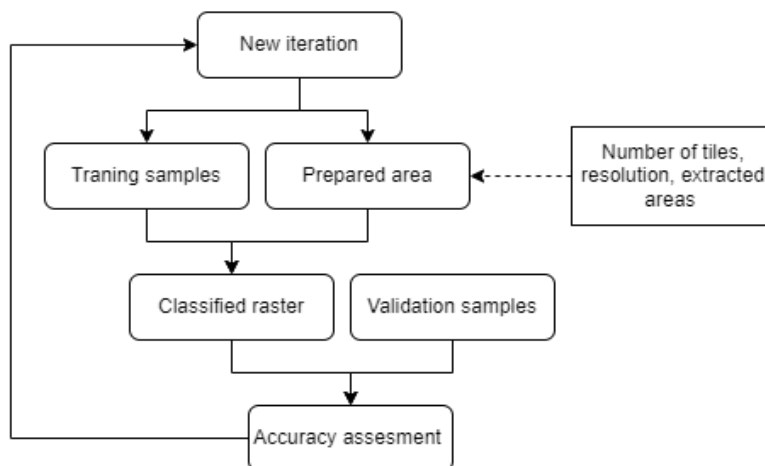


Figure 3-3 Illustration of the iterative approach to classification used in the study.

A first classification test was made on the original resolution (0.16 m). The preliminary results showed visually that small shrubs are identified, but mostly classified as forest. However, large shrubs and forests were being classified as a mix of different classes, as at original resolution different colours and shadows within the same shrubs are observed. The hypothesis was that a lower resolution might blend shadows and colours into a homogeneous hue and a shrub/forest patch could be classified as a single feature.

Consecutive tests were then performed over different resolutions to develop the methodologic workflow. The orthophoto pixel size was increased to 0.96 m, 4.8 m, and 9.6 m to investigate if coarser pixels would decrease the salt and pepper structure by averaging the spectral value and to evaluate the most appropriate scale for forest classification, since that is the class with most diverse spectral profile. The classification on 9.6 m resolution showed a substantial confusion between one of the agricultural classes and other vegetation types (shrub and forest), therefore fields within Alvaret were removed using the agriculture shapefile layer from the Swedish board of Agriculture. Following the classification of 0.96 m, 4.8 m and 9.6 m resolution the 4.8 m resolution were chosen for extraction of the forest class. Further classification of the smaller shrubs was performed on 0.96 m resolution to be able to classify the whole area, as the coarser resolution substantially reduces the processing power needed.

The main steps in the process are illustrated in figure 3-4 below.

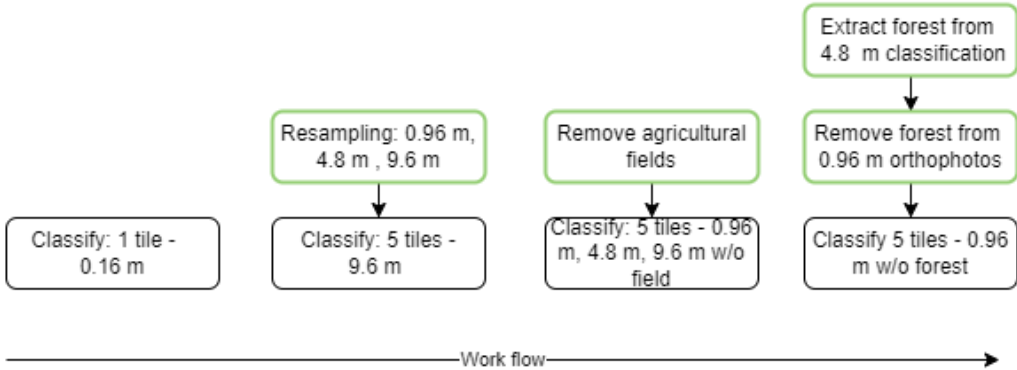


Figure 3-4 Illustration of the order of the classifications and changes performed to the prepared area in the study.

Software

The data processing and analysis was made with the GIS software ArcGIS Pro version 2.7 (ESRI, n.d.-a)

If nothing else is mentioned, the default settings for tools and functions were used.

3.3.1 Pre-processing

Overview

An overview of the preparation of the orthophotos is displayed in figure 3-5.

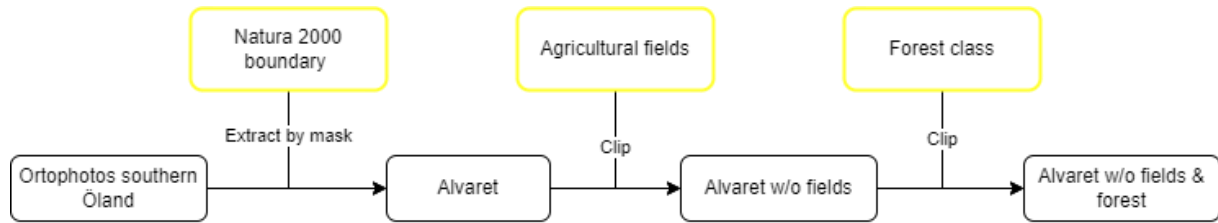


Figure 3-5 Illustration of the order of the classifications and changes performed to the prepared area in the study.

Extracting orthophotos

First the tool *Extract by mask* was used to remove superfluous raster data, clipping the orthophotos with the boundary shapefile. To automate the process over all the downloaded files, the model builder was used to iterate the extract tool over all files in the downloaded folder.

Resampling

The orthophotos prepared in the original resolution (0.16 m) was resampled to three different resolutions: 0.96 m, 4.8 m and 9.6 meters corresponding to 6, 30 and 60 of the original size pixels.

This was done with the raster function *Resampling* with the method *average*. A sample of the results, together with the original resolution are presented in figure 3-6 below, to visualise the change of distinguishability between the different structures.

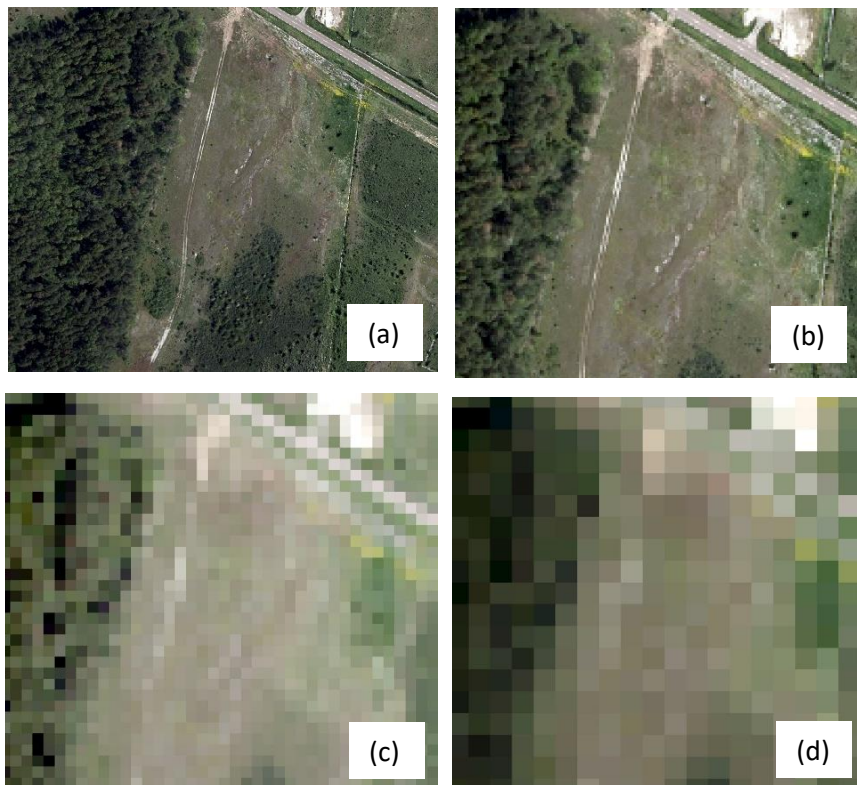


Figure 3-6 Section of Stora Alvaret displayed with four different pixel resolutions to visualise the change. (a) 0.16 m (b) 0.96 m (c) 4.8 m (d) 9.6 m. Scale 1:2000

Remove fields and forest

To remove the agricultural fields the raster function *Clip* was used with the clipping extent inside, and the land use shapefile as clipping extent. As raster functions create virtual raster layers the new layer was exported as a raster to be saved and possible to use in other project files.

Further in the project the clip function ceased to work. Therefore, the extracted forest pixels were removed with another method.

The raster was reclassified with the tool *reclassify* so all pixels representing forest had the value 0 and all others had the value 1. For the orthophoto each band was extracted with the raster function *extract bands*. Each band was then multiplied with the reclassified forest raster, so all pixels within the forest area got a value of 0 and all others remained the same. The four bands were then put together into one raster again with the raster function *composite bands*.

Extract single class from image

With the tool *Extract by attribute* one class could be extracted from the classified image and saved as a separate raster. In order to use it as a clipping extent on other files the raster was converted to polygons with the tool *Raster to Polygon*. This was used to extract the forest class from the lower resolution classification (4.8 m) to be able to remove the corresponding area from the higher resolution orthophoto (0.96 m) to eliminate the mixing between forest and shrub classes in the classification.

NoData

Several times during the processing NoData pixel, e.g. outside of the boundaries of the alvar or within the range of the fields that were removed, were updated to a value of 0 (RGB 0,0,0) which corresponds to a black colour. As this would affect the result of the classification they had to be removed. For this the tool *Set raster properties* was used to change the 0 value back to NoData.

Mosaicking

The mosaics of the 5 tiles as well as the full area were created with the tool *Mosaic to new raster*. It uses the mosaic operator *last* and the mosaic colormap *first* as a default, meaning that the value of the cell in the last raster dataset in an overlapping will be used while the colour scheme from the first raster will be applied to all.

Preparation of training and validation data

The training and validation samples were created in ArcGIS based on the observations during the field study in two ways. First, specific samples noted in the field study were digitalized by drawing the noted areas on the digital orthophotos with the training samples manager pane in the program. Second, samples were photo-interpreted and digitalized outside of the areas of the field study based on the knowledge obtained in the field work. This method enables more samples to be added to classes that do not have enough representation at the chosen study sites and spread samples along the full Alvaret.

The training and validation samples were spread over an area covered by 5 orthophoto tiles, as shown in Figure 3-7. The 5 tiles cover around 30 km² which represent around 12 % percent of full area (including fields). The second, fourth and fifth orthophotos are where the field study was conducted, while the first and third orthophoto were selected to include structures that were not, or only scarcely, represented before.

Their index numbers are from top down, with a separate numbering in parenthesis: (1) 627_59_2500, (2) 626_59_5025, (3) 626_59_2550, (4) 626_59_0000, (5) 624_59_5000. Tile 1 was selected based on the large presence of tree structures with a white colour, identified as willow, of which only few examples were found at the field study sites. Tile 3 was selected due to the many agricultural fields situated on the border.

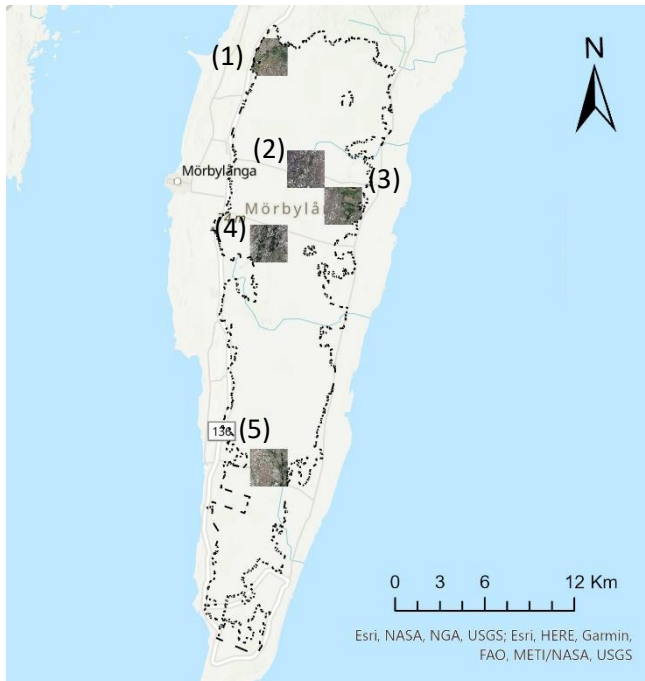


Figure 3-7 Location of tiles used for training and validation samples in the study.

Separate training data sets, with different classes, were created for the respective resolutions as the possibility to differentiate between structures changes with the pixel size, see figure 3-3.

To compute the classification and validation the training samples need to be split into two parts. For the first tests on the single orthophotos the classification data and validation data were on separate sites, e.g. the training of the model was done on one orthophoto and the validation was done on another orthophoto. For the other classifications the training and validation was done on all five tiles. The split between the training and validation samples was done by selecting, at random, 30% of the samples from each class in the attribute table and saving them to a new layer as the validation data. The rest of samples, 70%, were used for training.

3.3.2 Classification

For the classification the supervised classification method Random Forest was used. Random Forest is a classification algorithm based on multiple decision trees. A decision tree classifier discriminates between classes with binary rules, on for example spectral values, applied consecutively. In Random Forest each decision tree is built from a random set of input variables and the end classification result is the average of all the trees, which minimizes generalization and overfitting of the algorithm (Chuvieco, 2016).

In ArcGIS pro the tool *Train random tree* utilizes the random forest method. The default settings are: Maximum number of trees: 50, Maximum tree depth: 30, Maximum sample class: 1000 (Esri, n.d.-b).

3.3.3 Post-processing

Generalization

After the classification, the 4.8 m resolution was processed to remove noise and small miss-classified regions in the forest class. This was done with a process of three steps in ArcGIS called generalizing. The process uses the tools *Region Group*, *Set Null* and *Nibble*. The region group tool identifies regions for each cell, it uses *four neighbours* and connectivity test *within* as default. The *Set Null* tool then replaces all regions smaller than the set value with NoData. Three different values were tested, 1 pixel, 10 pixels and 20 pixels and the choice of maximum 10 pixels was based on a visual interpretation. As the last step the nibble tool replaces the NoData values with the nearest neighbour (ESRI, n.d.-b).

Calculation of area

The area for the training and validation polygons were calculated within the attribute table. For each file a new field was created, and the geometry was calculated for each polygon.

The area for the classes in the full area classification were calculated with the tool *Tabulate area*. It calculates the area for each class within a zone, in this case the index file for the orthophotos was used.

The statistics were then summarized in Microsoft Excel.

3.3.4 Validation

Validation of data can be done by evaluating the accuracy. Accuracy assessment is used both to test the significance of the result and to evaluate the methodology in relation to others. It can be done in several ways but one of the most reliable methods is to have an independent sample set for validation (Chuvieco, 2016).

Confusion matrix

Confusion matrixes or error matrixes have been used since the 70s to evaluate remote sensing data with respect to mapping accuracy. It is constructed with points sampled from the map containing reference point or ground truth values together with the classified value (Stein, Van der Meer, & Gorte, 2013).

The validation of the classification in this study was done with a confusion matrix computed by accuracy assessment points. The tool *Create accuracy assessment points* was used to create points within the validation samples with the ground truth value. It uses *stratified random* as a default sampling strategy, where the number of random points were set to 5000. With the tool *Update accuracy assessment points* the previously created points were updated with the class value from the classified raster layer. The tool *Compute confusion matrix* compares the two values for the same point and creates a confusion matrix.

In ArcGIS pro the confusion matrix includes overall accuracy, kappa index of agreement, producer's accuracy and user's accuracy, (Esri, n.d.-a).

Overall accuracy

The overall accuracy is a percentage measure calculated as the number of correctly classified samples or pixels divided by the total amount of samples or pixels (Sim & Wright, 2005). It is calculated according to equation 1, where P_{CC} is correctly classified pixels and P_T is the total number of pixels.

$$OA = \frac{P_{CC}}{P_T} \quad (1)$$

Kappa index

The kappa index of agreement or Cohen's kappa coefficient is a measure of agreement taking into consideration the possible agreement by chance. Thus, it can be referred to as a "true" agreement or agreement beyond chance. It is calculated as observed agreement minus chance agreement divided with 1 minus chance agreement. The possible result lay between -1 and 1 although values below 0 are rare as it would mean an agreement worse than chance. The kappa value is calculated according to equation 2, where P_O is proportion of observed agreements and P_C is proportion of agreements expected by chance (Sim & Wright, 2005).

$$\kappa = \frac{P_O - P_C}{1 - P_C} \quad (2)$$

An interpretation of values of kappa index are presented in table 3-3 (Sim & Wright, 2005).

Table 3-3 Interpretation of Kappa index (Sim & Wright, 2005).

KAPPA VALUE	STRENGTH OF AGREEMENT
≤0	Poor
0.01 – 0.20	Slight
0.21 – 0.40	Fair
0.41 - 0.60	Moderate
0.61 – 0.80	Substantial
0.81 - 1	Almost perfect

User's and producer's accuracy

User's and producer's accuracy might be more well-known under the names type 1 and type 2 error or the related error of commission and omission (Esri, n.d.-a).

User's accuracy is a measure of false positive, counting the number of samples or pixels incorrectly being classified to a certain class (Esri, n.d.-a). One minus the user's accuracy corresponds to the error of commission. User's accuracy is calculated according to equation 3 where P_{CW} is wrongly classified pixels and P_T is total classified pixels the specified class.

$$UA = \frac{P_{CW}}{P_T} \quad (3)$$

Producer's accuracy is instead a measure of pixels not being classified as the class they belong, called a false negative (Esri, n.d.-a). One minus the producer's accuracy corresponds to the error of omission. Producer's accuracy is calculated according to equation 4 where P_{CO} is pixels classified as another class and P_T is total classified pixels in the specified class.

$$PA = \frac{P_{CO}}{P_T} \quad (4)$$

4 Result

4.1 Shrub and tree species

The shrub and tree species that were identified during the field study are presented, with their English and Latin name, in table 4-1 below.

Table 4-1 Shrub and tree species identified in the field study.

English name	Latin name
Common juniper	<i>Juniperus communis</i>
Shrubby cinquefoil	<i>Dasiphora fruticosa</i>
European ash	<i>Franxinus exelsior</i>
Silver Birch	<i>Betula pendula</i>
Scots Pine	<i>Pinus sylvestris</i>
Azerole	<i>Cratageus azaralus</i>
Dog rose	<i>Rosa canina</i>
Rockspray cotoneaster	<i>Cotoneaster horizontalis</i>
Barberry	<i>Berberis vulgaris</i>
Common sallow /willow	<i>Salix cinerea</i>
Blackthorn	<i>Prunus spinosa</i>

Several of the shrub species such as azerole, rockspray cotoneaster, barberry and blackthorn were only found in single or a few specimens and often growing in close proximity or intertwined with juniper. This made it difficult to produce a sufficient amount of training and validation sample and therefore it was not possible to perform a classification on the species separately.

Most trees were found growing in groves or forest patches. They were therefore classified together as a forest class with the exemption of willow, which has a very contrasting, white colour.

4.2 Classification

4.2.1 Classification at 0.16 m resolution

The first classification was validated on one tile. There were 62 training polygons and 45 validation polygons covering an area of 13 000 m² and 4500 m², respectively. The separate classes and corresponding number of polygons and area are presented in table 4-2.

Table 4-2 Training sample set for classification of 0.16 m pixel resolution orthophoto with presented with number of polygons and corresponding area for each class.

Class	Training		Validation	
	Polygons	Area / (m ²)	Polygons	Area / (m ²)
Juniper	16	329	12	667
Shrubby cinquefoil	18	2 456	4	169
Trees	2	3 592	5	2 897
Rosehip	2	59	1	11
Stone	3	460	3	106
Grass	9	1 236	4	290
Soil	2	964	4	49
Soil2	2	2 519	3	126
Soil3	3	1 287	4	62
Shadow	5	183	5	122
Total	62	13 085	45	4 499

The result of the first classification is displayed in figure 4-1. As seen, the result shows mixing of the classes known as a ‘salt and pepper’. This pattern is especially prominent within the forest structure.

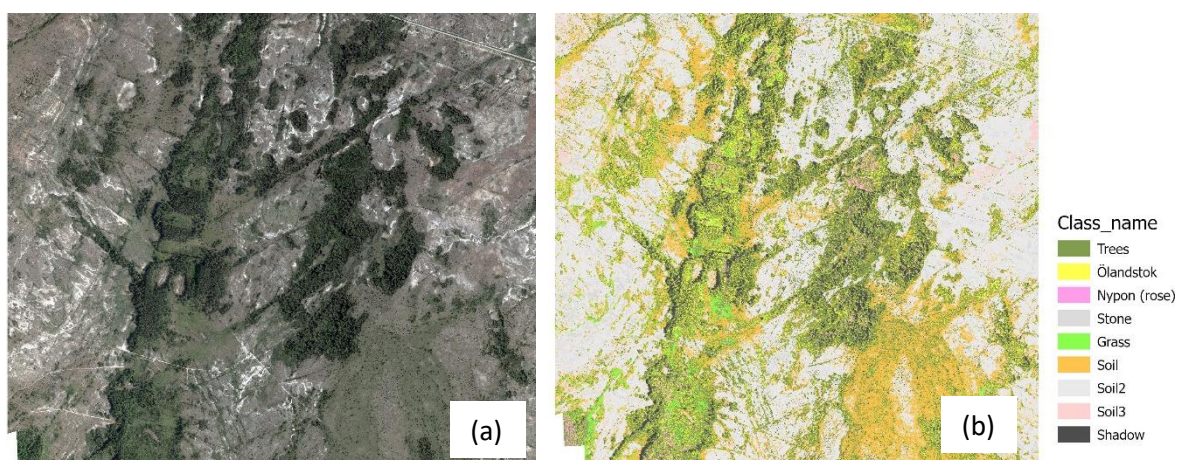


Figure 4-1 Display of orthophoto and classification result. (a) Orthophoto 0.16 m (b) Classification 0.16 m. Scale: 1:20 000. Orthophoto © Lantmäteriet.

Separate from the visual judgment the confusion matrix also demonstrated an overall weak classification with an overall accuracy of 0.21 and a kappa value of 0.10. The results within the classes however vary with e.g. a producer’s accuracy of 0.86 for the class *shadow* but 0.00 for *rosehip*, see table 4-3.

Table 4-3 Accuracy assessment for 0.16 m presented with producer’s and user’s accuracy for each class.

	Shrubby									
	Juniper	cinqfoile	Trees	Rosehip	Stone	Grass	Soil	Soil2	Soil3	Shadow
Producer’s accuracy	0.61	0.52	0.04	0.00	0.65	0.32	0.00	0.00	0.64	0.86
User’s accuracy	0.21	0.26	0.42	0.00	0.35	0.99	0.00	0.00	0.86	0.09

As can be seen in figure 4-2, the smaller structures are clearly distinguished from the ground layer but are classified as mostly trees (dark green) with a little bit of shrubby cinquefoil (yellow). Classification at 9.6 m resolution

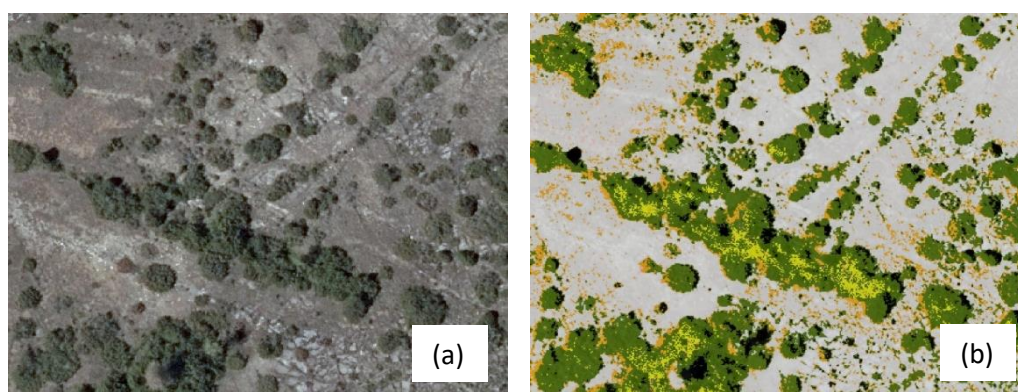


Figure 4-2 (a) Detail of orthophoto 0.16 m (b) Detail of classification 0.16 m. Scale: 1:1000. Orthophoto © Lantmäteriet.

The 9.6 m resampled mosaic was classified and validated with the help of 39 training and 18 validation polygons over 8 classes covering a total of 330 000 m² and 140 000 m² respectively. The separate classes and corresponding number of polygons and area are presented in table 4-4.

Table 4-4 Training sample set for 9.6 m with number of polygons and corresponding area.

Class	Training		Validation	
	Polygons	Total area / (m ²)	Polygons	Total area / (m ²)
Forest	10	132 964	5	86 611
Soil2	5	3 898	2	2 251
Soil3	5	2 908	2	1 832
Grass	6	10 946	3	10 377
Water	4	6 366	2	5 501
Field	6	94 964	3	35 188
Field_bare	3	76 645	1	6 143
Total	39	328 691	18	141 760

For this classification the vegetation classes were reduced, only differing between *forest*, *grass* and *field* (agricultural), since the shrub structures were difficult to distinguish from the ground layer (e.g. *grass*, *soil*).

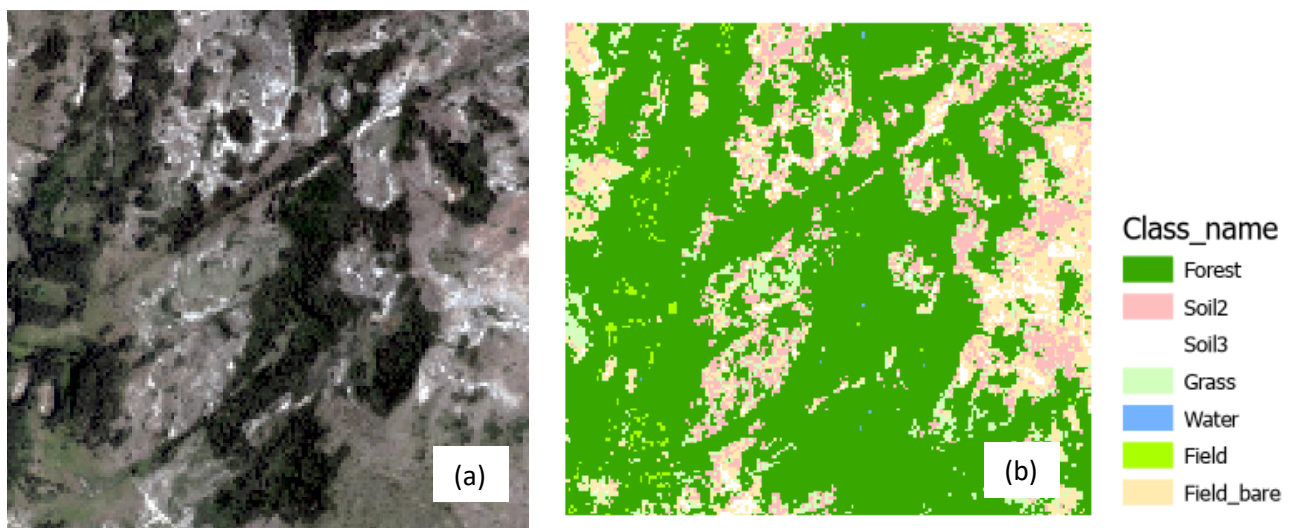


Figure 4-3 Display of orthophoto and classification result. (a) Orthophoto 9.6 m (b) Classification 9.6 m. Scale 1:20 000. Orthophoto © Lantmäteriet

The kappa value for the classification was 0.79 which, compared to the first classification is a considerable increase. However, as can be seen in figure 4-3, the classes *field* and *field bare* cover a lot of the area which does not include any farmland. This error is represented in the producer accuracy (errors of omission) for the classes *Forest* and *Soil2* with a value of 0.84 and 0.28 respectively, see producer's and user's accuracies in table 4-5.

Table 4-5 Accuracy assessment for 9.6 m presented with producer's and user's accuracy for each class.

	Forest	Soil2	Soil3	Grass	Water	Field	Field_bare
Producer's Accuracy	0.84	0.28	1.00	0.88	0.81	0.96	1.00
User's Accuracy	0.97	0.95	0.72	0.97	1.00	0.71	0.75

To minimize these errors the agricultural fields within the area were removed. The same training sample set as before was used, except the training and validation polygons from the classes *field* and *field_bare*, which were excluded as they lay in areas with NoData. This resulted in 30 training and 14 validation polygons covering a total of 160 000 m² and 110 000 m² respectively.

The separate classes and corresponding number of polygons and area are presented in table 4-6.

Table 4-6 Training sample set for 9.6 m w/o field with number of polygons and corresponding area.

Class	Training		Validation	
	Polygons	Total area / (m ²)	Polygons	Total area / (m ²)
Forest	10	132 964	5	86 611
Soil2	5	3 898	2	2 251
Soil3	5	2 908	2	1 832
Grass	6	10 946	3	10 377
Water	4	6 366	2	5 501
Total	30	157 082	14	106 572

The result of the new classification, without the field classes, is displayed in figure 4-4.

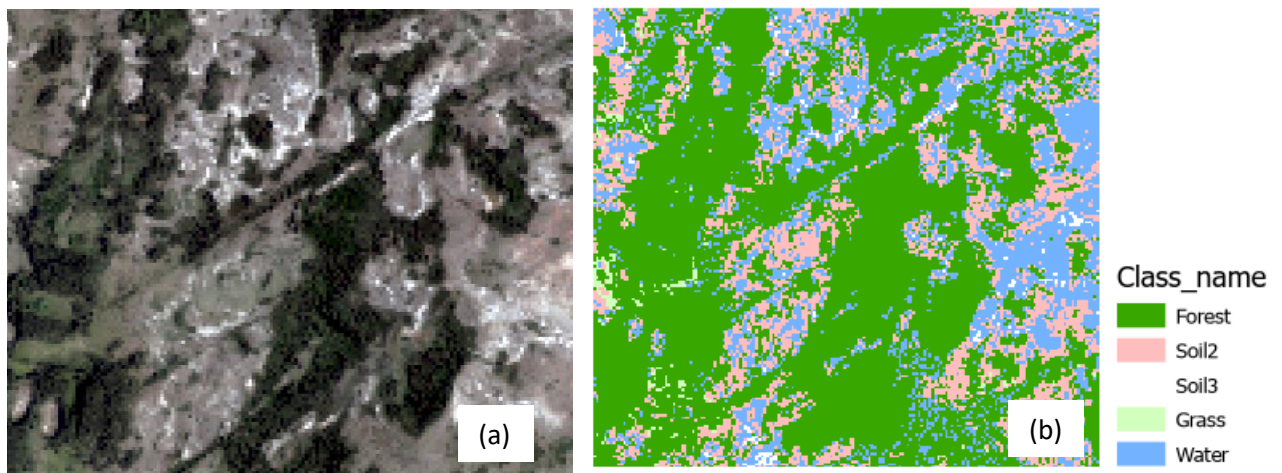


Figure 4-4 Display of orthophoto and classification result. (a) Orthophoto 0.16 m (b) Classification 0.16 m w/o fields. Scale 1:20 000. Orthophoto © Lantmäteriet

With the removal of the fields the kappa value increased to 0.92. But as can be seen in figure 4-4 a lot of ground is classified as water, which is represented by the producer's accuracy of 0.72 for water, see producer's and user's accuracies in table 4-7.

Table 4-7 Accuracy assessment for 9.6 m w/o field presented with producer's and user's accuracy for each class.

	Forest	Soil 2	Soil3	Grass	Water
Producer's accuracy	0.99	0.98	0.94	0.97	0.72
User's accuracy	0.99	1.00	1.00	0.86	0.91

4.2.2 Classification at 4.8 m resolution

For the 4.8 m resolution 62 training samples and 29 validation samples were used over a total area of 170 000 m² and 38 000 m² respectively. Since the removal of the fields in the 9.6 m classification contributed to a classification with a higher overall classification and kappa value, the first classification on 4.8 m was done after the fields were removed. The classification of the 4.8 m resolution orthophotos was made with 8 classes, 3 more than that for 9.6 m resolution, as it was possible to differentiate between more structures at this scale. The separate classes and corresponding number of polygons and area are presented in table 4-8.

Table 4-8 Training sample set for 4.8 m with number of polygons and corresponding area.

	Training		Validation	
	Polygons	Total area (m ²)	Polygons	Total area (m ²)
Shrub	7	4862	3	3969
Willow	8	584	4	248
Forest	16	119184	7	21225
Soil2	8	6504	4	1610
Soil3	6	2187	3	2701
Stone	4	546	2	163
Grass	9	20569	4	6482
Water	4	10773	2	1094
Total	62	165209	29	37492

The orthophoto resampled to 4.8 m resolution and the classification are shown in figure 4-5.

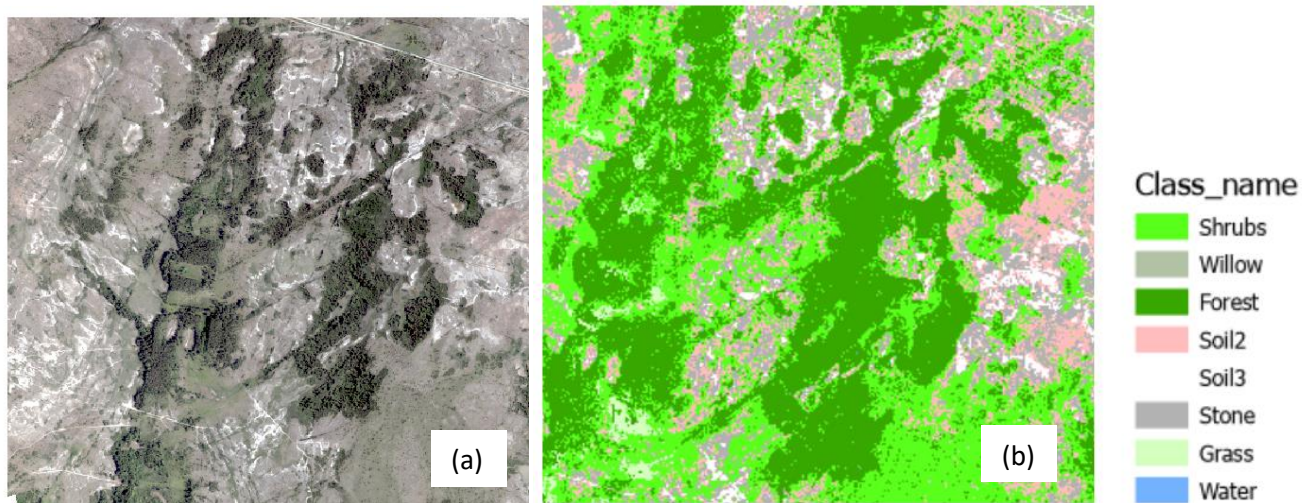


Figure 4-5 Display of orthophoto and classification result. (a) Orthophoto 4.8 m (b) Classification 4.8 m w/o fields Scale 1:20 000. Orthophoto © Lantmäteriet

As can be seen in figure 4-5 the image contains a similar “salt and pepper” structure as the classification on the original resolution, figure 4-1.

The validation gave a kappa value of 0.86 and overall accuracy of 0.91. That is lower than for the 9.6 m classification without fields, but grass and shrub regions are identifiable. See producer’s and user’s accuracies in table 4-9.

Table 4-9 Accuracy assessment for 4.8 m presented with producer's and user's accuracy for each class.

	Shrub	Willow	Forest	Soil2	Soil3	Stone	Grass	Water
Producer's accuracy	0.61	0.48	0.96	0.99	1.00	0.95	0.95	0.57
User's accuracy	0.88	1.00	0.93	0.67	1.00	0.95	0.90	1.00

4.2.3 Classification at 0.96 m resolution

The 0.96 m resolution was, like the 4.8 m resolution, classified without the field classes. At this scale it was possible to differentiate between the two shrubs juniper and shrubby cinquefoil and the classification was done over 10 classes. There was a total of 89 training and 36 validation polygons which covered an area of 160 000 m² and 34 000 m² respectively. The separate classes and corresponding number of polygons and area are presented in table 4-10.

Table 4-10 Training sample set for 0.96 m with number of polygons and corresponding area.

Class	Training		Validation	
	Polygons	Total area / (m ²)	Polygons	Total area / (m ²)
Juniper	18	621	6	99
Shrubby cinquefoil	13	839	6	363
Willow	10	726	4	146
Forest	15	125 964	6	18 126
Shadow	5	168	2	59
Soil2	6	5 439	2	1 031
Soli3	8	3 148	2	1 501
Stone	4	456	2	253
Grass	6	15 592	3	5 731
Water	4	6 665	3	6 512
Total	89	159 618	36	33 821

The orthophoto resampled to 0.96 m resolution and the classification are shown in figure 4-6.

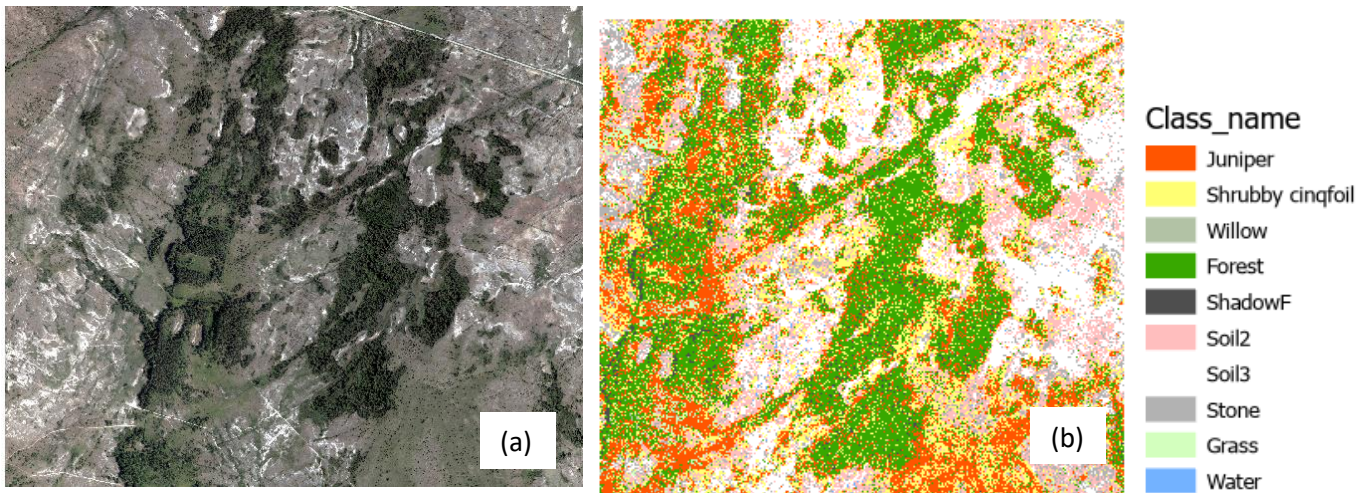


Figure 4-6 Display of orthophoto and classification result. (a) Orthophoto 0.96 m (b) Classification 0.96 m w/o fields. Resolution: 1:20 000. Orthophoto ©Lantmäteriet

This classification shows similarity to the original classification with a lot of salt and pepper structures. It however shows a much higher result on both overall accuracy and kappa value, which in this classification were 0.88 and 0.82, see producer's and user's accuracies in table 4-11.

Table 4-11 Accuracy assessment 0.96 m presented with producer's and user's accuracy for each class.

	Shrubby									
	Juniper	cinqfoil	Willow	Forest	Shadow	Soil2	Soli3	Stone	Grass	Water
Producer's accuracy	0.73	0.67	0.91	0.80	0.90	0.95	1.00	0.95	0.93	1.00
User's accuracy	0.04	0.11	0.42	0.99	0.41	0.99	1.00	0.83	0.99	1.00

4.2.4 Forest removal

An attempt to combine the different resolutions was made, with extracting the forest class from a higher resolution and removing the corresponding area from a lower resolution orthophoto. The 4.8 m resolution was chosen, as the 9.6 m resolution classified a lot of open ground as forest. In order to remove the smaller isolated regions within the forest the classified image was put through pixel generalization by nearest neighbour, as described in section 3.3.3 (*Generalization*). The results are displayed, together with the 4.8 m classification, in figure 4-7 with removal of regions smaller than 1, 10 and 20 pixels respectively.

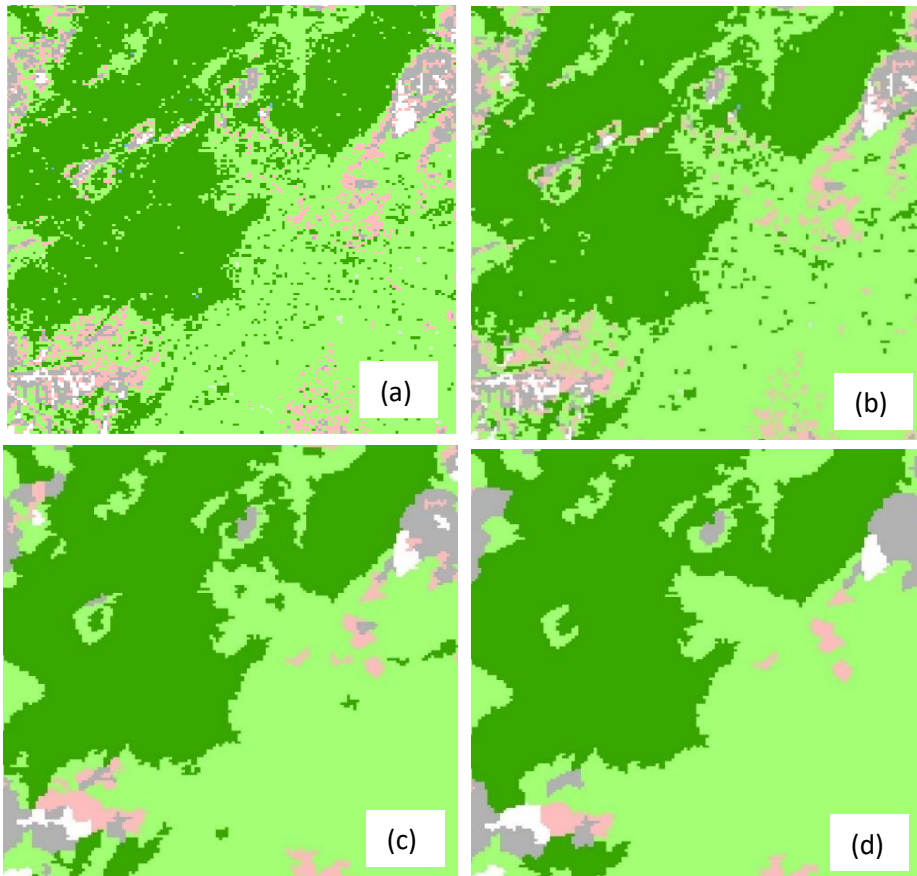


Figure 4-7 Display of results from post processing with generalization method. (a) 4.8 m classification (b) generalization of 1 pixel (c) generalization of 10 pixels (d) generalization of 20 pixels. Resolution 1:10 000. Orthophoto ©Lantmäteriet

The generalization of 10 pixels was chosen as the best, based on visual interpretation over a larger part of the mosaic.

With the generalized image of 10 pixels from the 4.6 m resampling classification, the forest class was extracted using the method described in section 3.4.1 (*Extract single class from image*). That area was then removed from the mosaic of 0.96 m resolution, with the raster clip function, described in section 3.4.1 (*Remove fields and forest*).

4.2.5 Classifying without forest

The 0.96 m resolution was chosen as the lower resolution, as it still was possible to differentiate between the shrub species, but reduces the processing power needed highly with respect to the original 0.16 m resolution. This time with 101 training polygons and 46 validation polygons with a

total area of 31 000 m² and 26 000 m² respectively. The separate classes and corresponding number of polygons and area are presented in table 4-12.

Table 4-12 Training sample set 0.96 m w/o forest with number of polygons and corresponding area.

Class	Training		Validation	
	Polygons	Total area (m ²)	Polygons	Total area (m ²)
Juniper	25	985	11	697
Shrubby cinqfoile	17	1043	10	567
Willow	4	270	3	253
Soil2	15	6165	6	5439
Soil3	11	3084	4	2619
Stone	9	688	4	456
Grass	14	16817	6	15592
Water	6	2061	2	394
Total	101	31113	46	26017

The orthophoto with the remove forest area and the classification result are shown in figure 4-8.

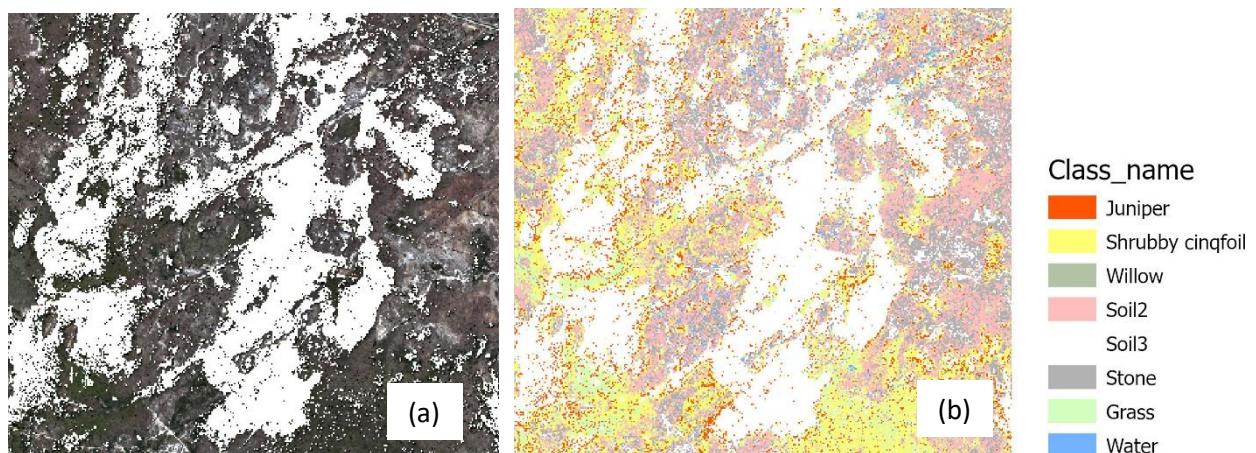


Figure 4-8 Display of orthophoto and classification result. (a) Orthophoto 0.96 m (b) Classification 0.96 m w/o fields and forest. Resolution: 1:20 000. Orthophoto ©Lantmäteriet

It resulted in an overall accuracy of 0.96 and a kappa value of 0.94, see producer's and user's accuracies in table 4-13.

Table 4-13 Accuracy assessment for 0.96 m w/o forest presented with producer's and user's accuracy for each class.

	Shrubby							
	Juniper	Cinqfoil	Willow	Soil2	Soil3	Stone	Grass	Water
Producer's accuracy	0.84	0.80	0.73	0.99	0.95	0.99	0.97	1.00
User's accuracy	0.54	0.69	0.86	0.99	1.00	0.73	1.00	1.00

4.2.6 Summary of results

The number of classes, overall accuracy, and kappa value together with producer's and user's accuracy for the two species of interest, juniper and shrubby cinquefoil, are presented in table 4-14, for all classifications presented.

Table 4-14 Summary of result of classification

Resolution	Number of classes	Overall accuracy	Kappa value	Juniper		Shrubby cinquefoil	
				Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy
0.16 m	10	0.21	0.10	0.61	0.21	0.52	0.26
9.6 m	7	0.87	0.79	-	-	-	-
9.6 m w/o field	5	0.97	0.92	-	-	-	-
4.8 m	8	0.91	0.86	-	-	-	-
0.96 m	10	0.88	0.82	0.73	0.04	0.67	0.11
0.96 m w/o forest	8	0.96	0.94	0.84	0.54	0.80	0.69

4.3 Proposed method

Following the results of the performed tests, a final method is proposed to achieve a sufficient classification of *Juniperus communis* and *Dasiphora fruticosa* on Stora Alvaret. It is based on the following findings:

- The original pixel resolution is too heavy for processing on available resources. Additionally, on a first test the fine resolution contributes to a separation of vegetation and shadows which contribute to the mixing of classes.
- The agriculture fields share similar reflectance with some of the other vegetation classes. As there is accurate and updated geodata available, they can be removed from the classification.
- Forest areas showed to be an additional source of errors for the classification of shrubs. Therefore, they should be masked out if possible. In this study, it was found out that the best resolution to classify forests is 4.8 m.

The steps and workflow of the proposed method is presented in figure 4-9.

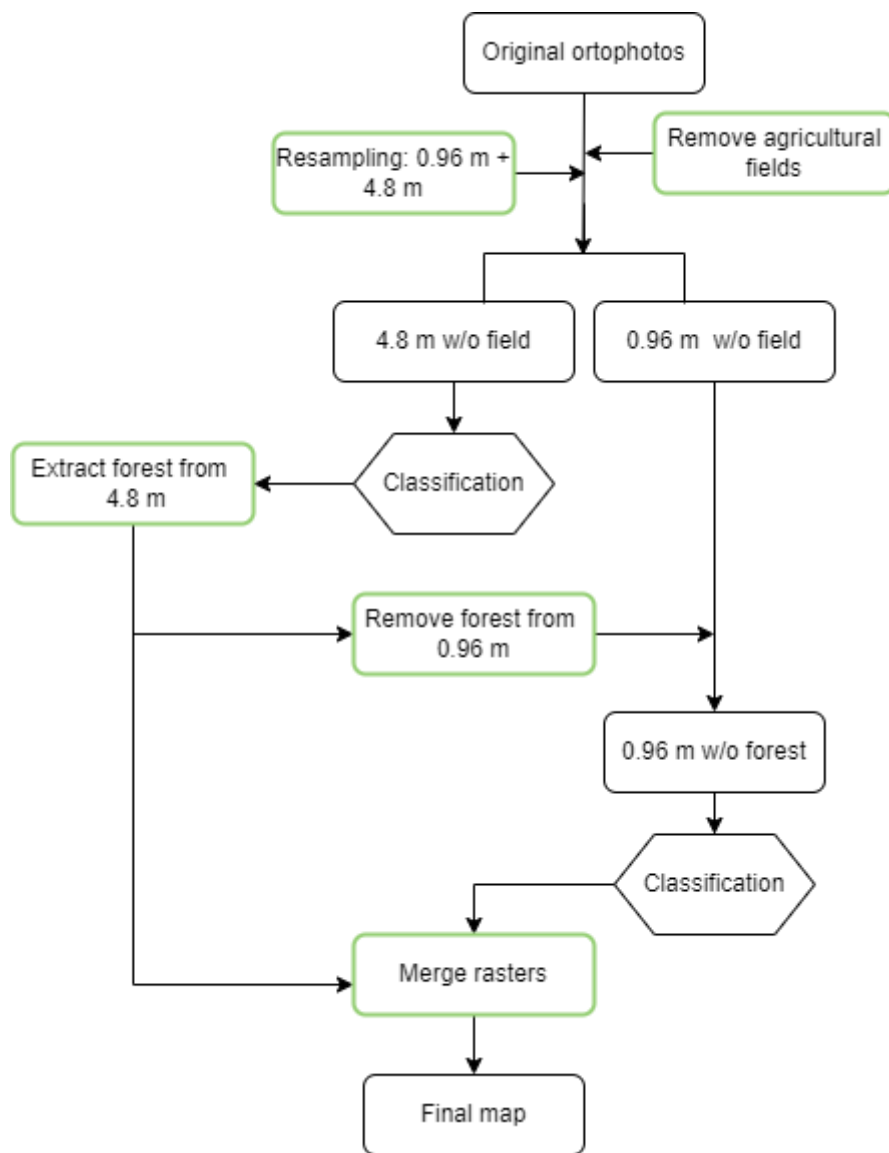


Figure 4-9 Illustration of workflow of proposed method for classification in the study.

4.3.1 Overview of final classification

Finally an overview of the final classification of the five tiles are presented together with the coverage area of the classes.

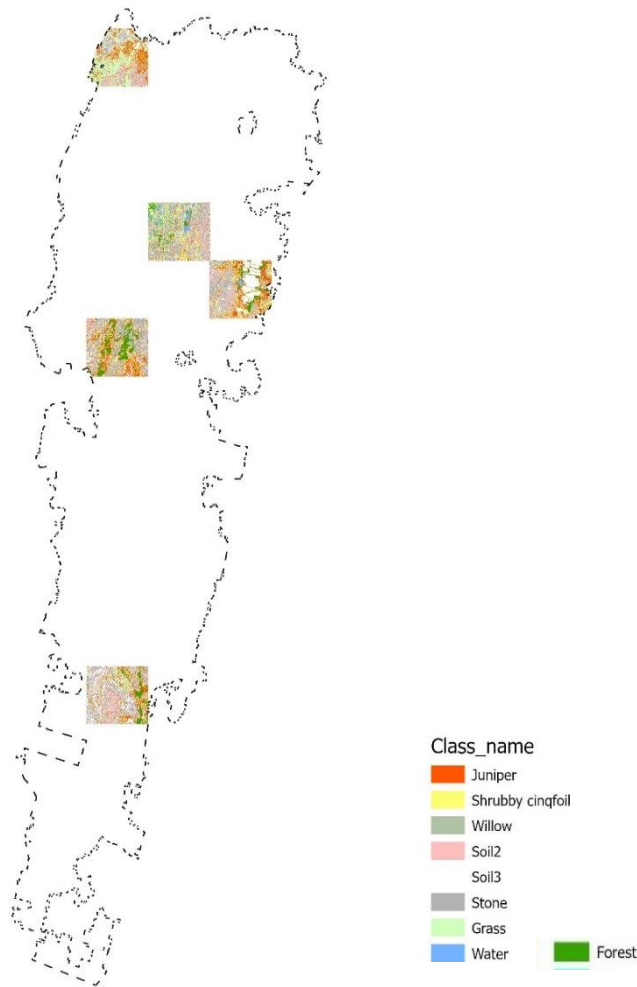


Figure 4-10 Classification of the five tiles at 0.96 m resolution w/o forest together with the forest class from the 4.8 m resolution classification. Scale 1:180 000

The areas of each class and the corresponding percentage of the five tiles are summarized in table 4-15.

Table 4-15 Classes in the final map with their total area and the corresponding percentage of the classified area.

Class	Area (m ²)	Area (km ²)	Percentage of area
Juniper	3482630.6	3.5	8.2%
Shrubby Cinqfoil	2828735.1	2.8	6.7%
Willow	352266.9	0.4	0.8%
Soil2	4162534.5	4.2	9.8%
Soil3	2206817.3	2.2	5.2%
Stone	7313057.3	7.3	17.2%
Grass	3873942.8	3.9	9.1%
Water	321961.9	0.3	0.8%
Forest	17878313.8	17.9	42.1%
Total	42420260	42.4	100.0%

5 Discussion

According to Sim & Wright (2005) a kappa value above 0.61 implies a substantial classification while a kappa value above 0.81 is almost perfect. This suggests that all classifications except the first one, on the original resolution, could be somewhat useful and that the 9.6 m without field and the 0.96 m without forest could be useful results. The two studies presented in the background which included kappa values in their analysis reached a result of 0.88 (Boswell et al., 2017) and 0.82 (Mora et al., 2021) which means that this study's final classification (*0.96 m without forest*) is well comparable with a kappa value of 0.94.

While kappa value and overall accuracy is an easy metric to be able to compare classification results it is also important to examine the producer's and user's accuracy for the classes of interest. In this study a mixing between the different soil types, or between water and grass would not be of great importance since the vegetation is the classes of interest, however they affect the overall metrics. For example, the *water* class showed a very high accuracy on most classifications while the *shrubs*, *trees* and *grass* mixed a lot. Similarly to what Hellesen & Matikainen (2013) stated that separating between vegetation classes is difficult due to similar spectral reflection.

The final classification gave a high producer's accuracy for both species of interest, with 0.84 for juniper and 0.80 for shrubby cinquefoil, meaning that most pixels that are juniper and shrubby cinquefoil are also classified as such. However, the user's accuracy is lower, especially for juniper with 0.54 while shrubby cinquefoil has 0.69. For juniper, the classification mostly includes pixels that were in fact grass but also some that were validated as willow and shrubby cinquefoil (see appendix). To have a misclassification between juniper and shrubby cinquefoil does not necessarily create large implications for the end use of the data, if used for monitoring total shrub encroachment. Misclassification of grass pixels as juniper however introduces an overestimation of the area of shrubs which is important to consider in the interpretation or usage of the final map.

The size of the pixels changes the possibility to differentiate between the structures on the orthophotos, as was shown in this study. Depending on the application different classes can be included, as stated by Khorram et al. (2012) the choice of classes should "make sense" to the end user, and therefore different resolutions could be suitable. In the study by Mora et al. (2021) the final classification was performed on groups of vegetation with dominating species, since it wasn't possible to separate the single species altogether. Boswell et al. (2017) argue that classification of broad vegetation groups is possible, in their study tree, shrub and herbaceous were used.

The previous studies presented in the background are as mentioned differing in a lot of variables, one of which is the pixel size used for classification, with a range from 0.1 m to 10 m. As they are differing in more ways it's not possible to directly compare the results of the classifications connected to the pixel size, but it shows a possibility to obtain substantial results for the range of resolution. This study further shows that it is possible to classify single species, as juniper and shrubby cinquefoil with a substantial result. However, classifying single species requires a finer resolution of the orthophotos, which was shown in this study as it was not possible to differentiate between the shrub species at coarser resolution when creating training samples. In this case the original data had a pixel size of 0.16 m, but going back in time just 1 year to 2018 the available resolution of Alvet was 0.25 m and more remote parts of Sweden 0.5 m (Lantmäteriet, n.d.). While all these resolutions still are finer than the final resolution used in this study, if working with another area, or further back in time it is possible that only coarser resolutions are available, which might change the preferred method.

A possible application in monitoring would be to classify the area with orthophotos from two separate years and then perform a change detection that shows which areas had an encroachment or a reduction of shrubs. Both these methods should be considered as a first part in the screening process and not a final decision basis. However, it fits well with the work structure already implemented at the county administrative board where orthophotos are examined to decided areas that are visited in field.

In the EU project for restoration of Stora Alvaret performed between 1996 and 1999, a big effort was put into monitoring and evaluating the measures made. One of the two applications for monitoring was the creation of a vegetation map with the biggest detail yet over the alvar habitats on Öland. It was made with interpretation of aerial photographs and divided the area into vegetation types with a detail of 2 500 m². 80 different vegetation types were included and additionally a layer of shrub coverage was provided, with indication of less than 10 %, between 10-70% or above 70 % of shrubs. The authors point out that even with the high detail, it was necessary to do generalizations, as not all vegetation types are differentiable with aerial photos (Forslund & Lager, 2000). This indicates that the previously proposed adaptation for monitoring indeed provides a relevant and usable tool. The result of this thesis differs as it classifies the two individual shrub species together with broader vegetation types such as *forest* and *grass*. A possible implication for future research would be to focus the result on producing shrub coverage maps, similar to the suggestion based on Ekstam & Forshed's sub environments.

For monitoring, new classifications would have to be done for each year included, both forward and backwards in time. With every classification, all steps in the method need to be performed again and the training sample set would have to be updated. However, with the knowledge from one year's classification the training and validation samples could be created without doing a field visit, which would reduce the time and effort needed to produce the result. It is also possible, especially if the years of classification are close in time, that some or most of the training and validation samples can be re-used. If implementing the method as a part of a long term monitoring program it would be beneficial to further investigate the usage of the function *model builder* in ArcGIS pro, as a way to automate the process and reduce the effort needed for each new classification.

With removing first the fields and then the forest class from the orthophotos, this study uses a similar method as in Ayhan & Kwan (2020) in isolating the vegetation classes of interest to improve the classification. The two removed areas are however different in how much uncertainty they introduce to the final result. The field polygons are defined by the Swedish board of agriculture and while they are not completely covering, they are not removing any other land cover, which was validated through a visual inspection. The forest area removed is defined by the classification on the 4.8 m resolution. It had a producer's and user's accuracy of 0.96 and 0.93 respectively, which is high but not a perfect classification meaning that 4 % of pixels that are not forest will be included (and 7 % of pixels that are forest will be excluded). For a transparent result this uncertainty should be included in the final results, either by clearly stating it or re-calculating the overall accuracy and kappa value, with the producer's and user's accuracy for the forest class included.

Another possible effect of removing areas from the original image is that training and validation samples fall within areas of NoData. As previously mentioned, pixels of other classes were included in the forest removed and the classes wrongly classified as forest were *shrub*, *willow*, and *water* (see appendix), which means that possibly samples of the corresponding classes were removed. If following the suggested method, and creating the sample set for the finer resolution after the forest is removed, this problem does not emerge, but is important to be aware of when working with an iterative process and reusing sample sets.

Hellesen & Matikainen (2013) showed in their study that classification could improve when including LiDAR data as auxiliary information. However, LiDAR data is not produced in the same frequency as orthophotos by Lantmäteriet. The existing dataset was collected between 2009 and 2019 and there is no existing plan of updating the information (Lantmäteriet, 2020). Therefore, if including LiDAR data to improve the method for monitoring on long term, new data would need to be produced. Additionally it would need further knowledge on processing of LiDAR data or external production of an nDSM, that then could easily be included in the classification tool in ArcGIS pro.

The orthophotos used in this study, from 2019, were acquired in the end of May which according to both Persson et al. (2018) and Abdi (2020) falls within the best time for remote sensing data acquisition if not using multi-temporal data. Lantmäteriet only produces orthophotos every second year, and alternates with dates before and after the budburst (Lantmäteriet, n.d.). So, if only using late spring data it would be possible to perform a new classification every fourth year. Using multi-temporal data could improve the classification, as Persson et al. (2018) had the highest accuracy with all-year data, however it would require switching to satellite data or performing separate drone or airplane acquisition on several occasions.

The split between the training and validation samples of 70 % and 30 % described in the method were not kept consequent during the whole study. During the iteration process the same sample sets were reused and were already divided between training and validation data. The added data was not as accurate in terms of percentage. Also noted is that the split was made based on the number of polygons and not the number of pixels. Although most classes had similarly large polygons within the class, it was not always the case and therefore a 70% - 30 % split in number can correspond to a different percentage in area.

The number of polygons and area covered in the training sample sets also differed between the resolutions. One explanation is the change of number of classes, but it is also not possible to create an equally small sample on the 9.6 m resolution as the 0.16 m, and because of the time frame during the processing, the sample sets were kept at a small number. This resulted in that the first rule by Khorram et al. (2012) in creating training samples was not kept, as the samples of each class did not exceed the number of classes by three. However, it might be questionable if purely the number of samples or polygons should be measured, as previously discussed the number of pixels in each sample can differ a lot. Rule number two, that the training samples should include the spectral distribution of the class was kept, as samples were picked to cover the variation within the classes and all over the area which also covers rule four, to not cluster the samples. Rule three, to keep each sample as homogenous as possible, was kept within the smaller samples and single species classes. While for forest, a lot of mixed pixels were included, as it was a mix between the single trees and shadows in between.

Only one algorithm, *Random Forest*, was used for the classification in this study. It is showed to have high accuracy, to be easy to work with and to have good results on small datasets on land cover classification (Pal, 2005; Ramezan et al., 2021). With the limited time in this study, as Ramezan et al. (2021) recommends, *Random Forest* was chosen, but for further research several machine learning algorithms could be compared with, to ensure the highest accurate classification for the specific area classes.

Alterations can be done in almost all ArcGIS functions, and it would with the time frame of the study be impossible to test and evaluate the optimal values and settings for each. Therefore, the default has been used for all functions and as the classification reached sufficient results it is not evident to include optimisation test on each step.

With the high resolution orthophotos the data set covering the whole of Stora Alvaret amounted to around 40 GB. The big data set made it difficult to work with the whole area in ArcGIS when e.g. creating training samples, therefore the method was developed and validated on five tiles spread over the area. The five tiles were selected to include variation in vegetation and land cover, but for a more secure result, additional training and validation should be done all over the area. In further research working directly with the whole dataset would therefore be preferred, which requires computers with better processor and more physical memory. The trained classification algorithm could also be applied to all tiles in a batch process, which requires less processing powers and will produce classified rasters for the whole area, but will not contribute to any further training and validation material.

6 Conclusion

This study finds that it is possible to successfully classify juniper and shrubby cinquefoil on Stora Alvaret with the method proposed. It could be used as a step in a monitoring program for shrub encroachment but should not be used to make detailed management decisions without complementary field visits.

The kappa value for the final classification was 0.94 which is defined as an “almost perfect” result. However, it is important to examine the metrics for the classes of interest as their accuracy might differ from the overall.

For further research the method should be tested on the full area and then be applied on multiple years to be able to evaluate the usage of change detection as a tool in monitoring shrub encroachment. The creation of the sample set should be held more consequent to ensure comparable results.

7 References

- Abdi, A. M. (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience and Remote Sensing*, 57(1), 1–20. <https://doi.org/10.1080/15481603.2019.1650447>
- Abe, B. T., Olugbara, O. O., & Marwala, T. (2014). Experimental comparison of support vector machines with random forests for hyperspectral image land cover classification. *Journal of Earth System Science*, 123(4), 779–790. <https://doi.org/10.1007/s12040-014-0436-x>
- Artfakta. (n.d.-a). Adam och Eva. Retrieved March 22, 2022, from <https://artfakta.se/naturvard/taxon/dactylorhiza-sambucina-219792>
- Artfakta. (n.d.-b). Alvarmalört. Retrieved March 22, 2022, from <https://artfakta.se/naturvard/taxon/artemisia-oelandica-91>
- Artfakta. (n.d.-c). Alvarstånds. Retrieved March 22, 2022, from <https://artfakta.se/artbestamning/taxon/1489>
- Artfakta. (n.d.-d). Kalkkrassing. Retrieved March 22, 2022, from <https://artfakta.se/artbestamning/taxon/erucastrum-supinum-1939>
- Artfakta. (n.d.-e). Sankt Pers nycklar. Retrieved March 22, 2022, from <https://artfakta.se/naturvard/taxon/orchis-mascula-219864>
- Artfakta. (n.d.-f). Styv kalkmossa. Retrieved March 22, 2022, from <https://artfakta.se/artbestamning/taxon/2083>
- Artfakta. (n.d.-g). Trubbklockmossa. Retrieved March 22, 2022, from <https://artfakta.se/artbestamning/taxon/encalypta mutica-581>
- Ayhan, B., & Kwan, C. (2020). Tree, shrub, and grass classification using only RGB images. *Remote Sensing*, 12(8). <https://doi.org/10.3390/RS12081333>
- Bosch, A., Zisserman, A., & Muñoz, X. (2007). Image classification using random forests and ferns. *Proceedings of the IEEE International Conference on Computer Vision*. <https://doi.org/10.1109/ICCV.2007.4409066>
- Boswell, A., Petersen, S., Roundy, B., Jensen, R., Summers, D., & Hulet, A. (2017). Rangeland monitoring using remote sensing: comparison of cover estimates from field measurements and image analysis. *AIMS Environmental Science*, 4(1), 1–16. <https://doi.org/10.3934/environsci.2017.1.1>
- Chuvieco, E. (2016). *Fundamentals of Satellite Remote Sensing*. Taylor & Francis Group.
- Curran, P. J., & Atkinson, P. M. (2002). Issues of scale and optimal pixel size. In A. Stein, F. der Meer, & B. Gorte (Eds.), *Spatial Statistics for Remote Sensing* (pp. 115–133). https://doi.org/10.1007/0-306-47647-9_7
- Directorate-General for Environment. (n.d.). Natura 2000. Retrieved February 17, 2022, from https://ec.europa.eu/environment/nature/natura2000/index_en.htm
- EEA. (2021). Natura 2000 data - the European network of protected sites. Retrieved from <https://www.eea.europa.eu/data-and-maps/data/natura-12>
- Ekstam, U., & Forshed, N. (2002). *Svenska Alvarmarker*. Naturvårdsverket.
- Eldridge, D. J., Bowker, M. A., Maestre, F. T., Roger, E., Reynolds, J. F., & Whitford, W. G. (2011). Impacts of shrub encroachment on ecosystem structure and functioning: Towards a global

- synthesis. *Ecology Letters*, 14(7), 709–722. <https://doi.org/10.1111/j.1461-0248.2011.01630.x>
- EOS. (2019). Satellite Data: What Spatial Resolution Is Enough? Retrieved February 26, 2022, from <https://eos.com/blog/satellite-data-what-spatial-resolution-is-enough-for-you/>
- Esri. (n.d.-a). Compute Confusion Matrix (Spatial Analyst)—ArcGIS Pro | Documentation. Retrieved January 27, 2022, from <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/compute-confusion-matrix.htm>
- Esri. (n.d.-b). Train Random Trees Classifier (Spatial Analyst). Retrieved January 28, 2022, from <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/train-random-trees-classifier.htm>
- ESRI. (n.d.-a). *ArcGIS Pro*.
- ESRI. (n.d.-b). Generalizing classified output by removing small isolated regions. Retrieved February 1, 2022, from <https://desktop.arcgis.com/en/arcmap/latest/extensions/spatial-analyst/image-classification/generalizing-classified-output-by-removing-small-isolated-regions.htm>
- Forslund, M. (2001). *Natur och Kultur på Öland*. Länsstyrelsen i Kalmar län.
- Forslund, S., & Lager, H. (2000). *Protection and restoration of parts of Stora Alvaret*.
- GISGeography. (2021). Multispectral vs Hyperspectral Imagery Explained. Retrieved February 24, 2022, from <https://gisgeography.com/multispectral-vs-hyperspectral-imagery-explained/#:~:text=The main difference between multispectral,to 3 to 10 bands.&text=Hyperspectral imagery consists of much,hundreds or thousands of bands.>
- Hellesen, T., & Matikainen, L. (2013). An object-based approach for mapping shrub and tree cover on grassland habitats by use of LiDAR and CIR orthoimages. *Remote Sensing*, 5(2), 558–583. <https://doi.org/10.3390/rs5020558>
- IPBES. (2019). Report of the Plenary; Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. *Report of the Plenary*, 7(1), 56.
- Jordbruksverket. (2021a). Block och blockareal. Retrieved November 25, 2021, from <https://jordbruksverket.se/stod/lantbruk-skogsbruk-och-tradgard/sam-ansokan-och-allmant-om-jordbrukarstoden/block-och-blockareal>
- Jordbruksverket. (2021b). Grödkoder. Retrieved from <https://jordbruksverket.se/stod/lantbruk-skogsbruk-och-tradgard/sam-ansokan-och-allmant-om-jordbrukarstoden/grodkoder>
- Jordbruksverket. (2021c). Kartor och Geografiska informationssystem. Retrieved from <https://jordbruksverket.se/e-tjanster-databaser-och-appar/e-tjanster-och-databaser-stod/kartor-och-gis>
- keuwlsoft. (2021). *GPS Waypoint Finder*. Retrieved from <https://www.keuwl.com/apps/gpswaypointfinder/>
- Khorrām, S., Koch, F. H., van der Wiele, C. F., & Nelson, S. A. C. (2012). Remote sensing. In *Progress In Electromagnetics Research C* (Vol. 39). <https://doi.org/10.1007/978-1-4614-3103-9>
- Lantmäteriet. (n.d.). Ortofoto. Retrieved March 17, 2022, from <https://www.lantmateriet.se/sv/Kartor-och-geografisk-information/geodataprodukter/produktlista/ortofoto/>
- Lantmäteriet. (2019). *Orthophoto* (Vol. 1).

- Lantmäteriet. (2020). *Laserdata NH*. Retrieved from https://www.lantmateriet.se/globalassets/kartor-och-geografisk-information/hojddata/laserdata_nh_v2.10.pdf
- Maad, J., & Rosén, E. (2010). *Utvärdering av permanenta provytor på Stora Alvaret*. Retrieved from https://www.lansstyrelsen.se/download/18.7ab1493f1677d97be13c65/1544092797757/Utvärdering_av_permanenta_provytor_på_Alvaret.pdf
- Mora, C., Vieira, G., Pina, P., Lousada, M., Christiansen, H., Geografiska, S., ... Christiansen, H. H. (2021). *LAND COVER CLASSIFICATION USING HIGH-RESOLUTION AERIAL PHOTOGRAPHY IN* Published by : Taylor & Francis , Ltd . on behalf of the Swedish Society for Anthropology and Geography Stable URL : <https://www.jstor.org/stable/43871216> LAND COVER CLASSIFICATION USIN. 97(3).
- Mossberg, B. Stenberg, L. (2010). *Den nya nordiska floran*. Bonnier Fakta.
- Mountrakis, G., Im, J., & Ogole, C. (2011). ISPRS Journal of Photogrammetry and Remote Sensing Support vector machines in remote sensing : A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247–259. <https://doi.org/10.1016/j.isprsjrs.2010.11.001>
- Naturvårdsverket. (n.d.). Biologisk mångfald. Retrieved February 17, 2022, from <https://www.naturvardsverket.se/om-miljoarbetet/miljoarbete-i-eu/biologisk-mangfald>
- Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1), 217–222. <https://doi.org/10.1080/01431160412331269698>
- Persson, M., Lindberg, E., & Reese, H. (2018). Tree species classification with multi-temporal Sentinel-2 data. *Remote Sensing*, 10(11), 1–17. <https://doi.org/10.3390/rs10111794>
- PlantNet. (2021). *PlantNet Plant Identification*. Retrieved from <https://plantnet.org/en/>
- Ramezan, C. A., Warner, T. A., Maxwell, A. E., & Price, B. S. (2021). Effects of training set size on supervised machine-learning land-cover classification of large-area high-resolution remotely sensed data. *Remote Sensing*, 13(3), 1–27. <https://doi.org/10.3390/rs13030368>
- Rejmánek, M., & Rosén, E. (1992). Influence of Colonizing Shrubs on Species-Area Relationships in Alvar Plant Communities Author (s): Marcel Rejmánek and Ejvind Rosén Published by : Wiley Stable URL : <https://www.jstor.org/stable/3235829> REFERENCES Linked references are available on JST. *Journal of Vegetation Science*, 3(5), 625–630.
- Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E., ... Foley, J. (2009). Planetary boundaries: Exploring the safe operating space for humanity. *Ecology and Society*, 14(2). <https://doi.org/10.5751/ES-03180-140232>
- Rosén, E., & Bakker, J. P. (2005). Effects of agri-environment schemes on scrub clearance, livestock grazing and plant diversity in a low-intensity farming system on Öland, Sweden. *Basic and Applied Ecology*, 6(2), 195–204. <https://doi.org/10.1016/j.baae.2005.01.007>
- Rosén, E., & van der Maarel, E. (2000). Restoration of alvar vegetation on Öland, Sweden. *Applied Vegetation Science*, 3(1), 65–72. <https://doi.org/10.2307/1478919>
- SEOS. (n.d.). Spatial resolution. Retrieved February 26, 2022, from <https://seos-project.eu/remotesensing/remotesensing-c03-p02.html>
- Sim, J., & Wright, C. C. (2005). The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy*, 85(3), 257–268. <https://doi.org/10.1093/ptj/85.3.257>

- SLU Artdatabanken. (2015). Rödlista 2015 - Sammanfattning. Retrieved February 16, 2022, from <https://www.artdatabanken.se/var-verksamhet/rodlistning/sammanfattning-rodlista-2015/>
- SMHI. (2021). Ölands klimat. Retrieved February 18, 2022, from <https://www.smhi.se/kunskapsbanken/klimat/klimatet-i-sveriges-landskap/olands-klimat-1.4870#:~:text=Medeltemperaturen varierar exempelvis bara mellan,är det torraste i landet.>
- STARS. (n.d.). Multispectral and panchromatic images. Retrieved February 24, 2022, from <https://www.stars-project.org/en/knowledgeportal/magazine/remote-sensing-technology/introduction/multispectral-and-panchromatic-images/>
- Stein, A., Van der Meer, F., & Gorte, B. (2013). Remote Sensing and Digital Image Processing: Spatial Statistics for Remote Sensing. In *Kluwer Academic* (Vol. 53).
- Stockholm Resilience Center. (n.d.). The nine planetary boundaries. Retrieved February 17, 2022, from <https://www.stockholmresilience.org/research/planetary-boundaries/the-nine-planetary-boundaries.html>
- Svenska lantbruksuniversitetet. (n.d.). GET. Retrieved from 2021 website: <https://zeus.slu.se/get/?drop=get>
- Tholin, M. (2016). *Bevarandeplan för Natura 2000-område SE0330176 Stora Alvaret*. <https://doi.org/511-4969-16>
- van der Maarel; Sykes, M. T. (1993). Small-Scale Plant Species Turnover in a Limestone Grassland : The Carousel Model and Some Comments on the Niche Concept Author (s): Eddy van der Maarel and Martin T . Sykes Published by : Wiley Stable URL : <http://www.jstor.org/stable/3236103> REFERENCES. *Journal of Vegetation Science*, 4(2), 179–188.
- Wilson, J. B., Peet, R. K., Dengler, J., & Pärtel, M. (2012). Plant species richness: The world records. *Journal of Vegetation Science*, 23(4), 796–802. <https://doi.org/10.1111/j.1654-1103.2012.01400.x>
- Woodcock, C. E., & Strahler, A. H. (1987). The Factor of Scale in Remote Sensing. *REMOTE SENSING OF ENVIRONMENT*, (21), 311–332.

Confusion matrix 0,16 m

Shrubby													
ClassValue	Juniper	cingfoile	Trees	Rosehip	Stone	Grass	Soil	Soil2	Soil3	Shadow	Total	U_Accuracy	Kappa
Juniper	456	40	1694	6	0	2	0	0	0	19	2217	0,205683	0
Shrubby													
cingfoile	127	97	123	0	0	24	0	0	0	0	371	0,261456	0
Trees	22	47	136	0	0	120	0	0	0	0	325	0,418462	0
Rosehip	8	2	62	0	0	74	0	0	0	0	146	0	0
Stone	0	0	0	0	77	0	2	139	0	0	218	0,353211	0
Grass	0	0	1	0	0	102	0	0	0	0	103	0,990291	0
Soil	18	1	124	0	0	0	0	0	1	0	144	0	0
Soil2	3	1	38	1	41	0	45	0	24	0	153	0	0
Soil3	0	0	0	0	0	0	7	0	44	0	51	0,862745	0
Shadow	111	0	1041	5	0	0	0	0	0	116	1273	0,091123	0
Total	745	188	3219	12	118	322	54	139	69	135	5001	0	0
P_Accuracy	0,612081	0,515957	0,042249	0	0,652542	0,3168	0	0	0,637681	0,859259	0	0,205559	0
Kappa	0	0	0	0	0	0	0	0	0	0	0	0	0,095934

Confusion matrix 9.6 m

Confusion matrix 9.6 m										
	Forest	Soil2	Soil3	Grass	Water	Field	Field_bare	Total	Users accuracy	
Forest	2450	0	0	36	3	48	0	2537	0,97	0
Soil2	0	21	0	0	0	0	1	22	0,95	0
Soil3	0	0	62	0	24	0	0	86	0,72	0
Grass	8	0	0	310	0	0	0	318	0,97	0
Water	0	0	0	0	150	0	0	150	1,00	0
Field	470	0	0	0	0	1142	0	1612	0,71	0
Field_bare	0	55	0	5	9	0	207	276	0,75	0
Total	2928	76	62	351	186	1190	208	5001	0,00	0
Producers Accuracy	0,84	0,28	1,00	0,88	0,81	0,96	1,00	0	0,87	0
Kappa	0	0	0	0	0	0	0	0	0,00	0,79

Confusion Matrix 9.6 w/o fields

	Users						Total	accuracy	Kappa
	Forest	Soil 2	Soil3	Grass	Water	Total			
Forest	4024	0	0	15	25	4064	0,99	0	
Soil 2	0	104	0	0	0	104	1,00	0	
Soil3	0	0	81	0	0	81	1,00	0	
Grass	27	0	0	472	47	546	0,86	0	
Water	12	2	5	0	186	205	0,91	0	
Total	4063	106	86	487	258	5000	0	0	
P_Accuracy	0,99	0,98	0,94	0,97	0,72	0,00	0,97	0	
Kappa	0	0	0	0	0	0	0	0,92	

Confusion matrix 4.8 m

	Confusion matrix 4.8 m										Total	Users accuracy	Kappa
	Shrub	Willow	Forest	Soil2	Soil3	Stone	Grass	Water	Total	Users accuracy			
Shrub	322	9	33	2	0	0	0	0	0	0	366	0,88	0
Willow	0	16	0	0	0	0	0	0	0	0	16	1,00	0
Forest	207	0	2716	0	0	0	0	0	0	0	2923	0,93	0
Soil2	0	0	0	212	0	1	42	63	318	0	0	0,67	0
Soil3	0	0	0	0	360	0	0	0	0	0	360	1,00	0
Stone	0	0	0	1	0	21	0	0	0	0	22	0,95	0
Grass	0	8	81	0	0	0	822	0	911	0	0	0,90	0
Water	0	0	0	0	0	0	0	83	83	0	0	1,00	0
Total	529	33	2830	215	360	22	864	146	4999	0,00	0	0,00	0
P_Accuracy	0,61	0,48	0,96	0,99	1,00	0,95	0,95	0,57	0,00	0,91	0	0,91	0
Kappa	0	0	0	0	0	0	0	0	0	0	0	0	0,86

Confusion matrix 0.96 m

	Shrubby										Users		
	Juniper	Shrubby cinqfoil	Willow	Forest	Shadow	Soil2	Soil3	Stone	Grass	Water	Total	accuracy	Kappa
Juniper	11	3	0	206	0	0	0	0	36	0	256	0,04	0
Shrubby cinqfoil	1	36	2	295	0	0	0	0	0	0	334	0,11	0
Willow	0	3	20	3	0	0	0	22	0	0	48	0,42	0
Forest	1	12	0	2155	1	0	0	1	0	0	2170	0,99	0
Shadow	0	0	0	13	9	0	0	0	0	0	22	0,41	0
Soil2	0	0	0	0	0	144	0	1	0	0	145	0,99	0
Soil3	0	0	0	0	0	1	222	0	0	0	223	1,00	0
Stone	0	0	0	0	0	7	0	35	0	0	42	0,83	0
Grass	2	0	0	8	0	0	0	1	788	0	799	0,99	0
Water	0	0	0	0	0	0	0	0	0	963	963	1,00	0
Total	15	54	22	2680	10	152	222	37	847	963	5002	0	0
P_Accuracy	0,73	0,67	0,91	0,80	0,90	0,95	1,00	0,95	0,93	1	0	0,88	0
Kappa	0	0	0	0	0	0	0	0	0	0	0	0	0,82

Confusion matrix 0.96 w/o forest

	Shrubby						User's				
	Juniper	Cinqfoil	Willow	Soil2	Soil3	Stone	Grass	Water	Total	accuracy	Kappa
Juniper	112	21	9	0	0	0	64	0	206	0,54	0
Shrubby Cinqfoil	21	87	3	0	2	0	13	0	126	0,69	0
Willow	0	0	36	0	0	0	6	0	42	0,86	0
Soil2	0	0	0	1034	1	1	6	0	1042	0,99	0
Soil3	0	0	0	0	476	0	0	0	476	1,00	0
Stone	0	0	0	10	23	87	0	0	120	0,73	0
Grass	1	1	1	1	1	0	2907	0	2912	1,00	0
Water	0	0	0	0	0	0	0	76	76	1,00	0
Total	134	109	49	1045	503	88	2996	76	5000	0,00	0
Producer's accuracy	0,84	0,80	0,73	0,99	0,95	0,99	0,97	1,00	0,00	0,96	0
Kappa	0	0	0	0	0	0	0	0	0	0	0,94

Institutionen av naturgeografi och ekosystemvetenskap, Lunds Universitet

Student-examensarbete (seminarieuppsatser) i geografisk informationsteknik.

Uppsatserna finns tillgängliga på institutionens geobibliotek, Sölvegatan 12, 223 62 LUND. Serien startade 2010. Hela listan och själva uppsatserna är även tillgängliga på LUP student papers och via Geobiblioteket (www.geobib.lu.se).

Serie examensarbete i geografisk informationsteknik

1. Patrik Carlsson och Ulrik Nilsson (2010) Tredimensionella GIS vid fastighetsförvaltning
2. Karin Ekman och Anna Felleson (2010) Att välja grundläggande karttjänst - Utveckling av jämförelsemodell och testverktyg för utvärdering
3. Jakob Mattsson (2011) Synkronisering av vägdatabaser med KML och GeorSS - En fallstudie i Trafikverkets verksamhet
4. Patrik Andersson and Anders Jürisoo (2011) Effective use of open source GIS in rural planning in South Africa
5. Nariman Emamian och Martin Fredriksson (2012) Visualisering av bygglovsärenden med hjälp av Open Source-verktyg - En undersökning kring hur man kan effektivisera ärendehantering med hjälp av en webbapplikation
6. Gustav Ekstedt and Torkel Endoff (2012) Design and Development of a Mobile GIS Application for Municipal FieldWork
7. Karl Söderberg (2012) Smartphones and 3D Augmented Reality for disaster management - A study of smartphones ability to visualise 3D objects in augmented reality to aid emergency workers in disaster management
8. Viktoria Strömberg (2012) Volymberäkning i samhällsbyggnadsprojekt
9. Daniel Persson (2013) Lagring och webbaserad visualisering av 3D stadsmodeller - En pilotstudie i Kristianstad kommun
10. Lisette Danebjer och Magdalena Nyberg (2013) Utbyte av geodata - studie av leveransstrukturer enligt Sveriges kommuner och landstings objekttypskatalog
11. Alexander Quist (2013) Undersökning och utveckling av ett mobilt GISsystem för kommunal verksamhet
12. Nariman Emamian (2014) Visning av geotekniska provborrningar i en webbmiljö
13. Martin Fredriksson (2014) Integrering av BIM och GIS med spatiala databaser – En prestandaanalys
14. Niklas Krave (2014) Utveckling av en visualiseringsapplikation för solinstrålningsdata
15. Magdalena Nyberg (2015) Designing a generic user interface for distribution of open geodata: based on FME server technology

16. Anna Larsson (2015) Samredovisning av BIM- och GIS-data
17. Anton Lundkvist (2015) Development of a WEB GI System for Disaster Management
18. Ellen Walleij (2015) Mapping in Agricultural Development – Introducing GIS at a smallholders farmers’ cooperative in Malawi
19. Frida Christiansson (2016) Lagring av 3D - geodata - en fallstudie i Malmö Stad
20. Lisette Danebjer (2016) Methodology for creating and modifying distributed topologically structured geographical datasets
21. Jeanette Dunn Ekelund (2016) En jämförelse av algoritmer och resultat för flödesberäkning i QGIS/GRASS och ArcGIS
22. Ebba Gröndahl och Frida Thorman (2016) Verksamhetens optimala läge i staden och hur de är lokaliserade idag
23. Gunnar Rolander (2017) Data transformation using linked data ontologies
24. Måns Andersson och Moa Eklöf (2017) Stilsättning av geografiska data
25. Josefine Axelsson (2018) Automatisering av bygglovsansökningsprocessen med stöd av BIM och GIS
26. Leonard B. O. Berge (2018) Uppdatering och visualisering av stadsmodell med stöd av konverterade BIM-modeller
27. Rickard Ingesson & Gabriella Olsson (2019) Publicering av geografiska data på webben : En utvärdering av programsystem med fokus på öppen källkod
28. Alfred Hildingson & Patrik Sylve (2020) Visualisering av stadsmodeller på webben : Jämförande studie mellan CityGML och CityJSON
29. Isabelle Andersson (2020) Indoor positioning systems in office environments : a study of standards, techniques and implementation processes for indoor maps
30. Sebastian Roos & August Cnattingius (2021) Covid-19-pandemins konsekvenser på svenskt näringsliv - en Space Syntax analys : Hur har konkursutsattheten och arbetslösheten förändrats för detaljhandel, hotell och restauranger till följd av restriktioner som begränsar besöksnäring?
31. Emelie Ulin (2021) Simuleringar i planeringsprocessen med 3D-stadsmodeller
32. Alfred Hirschfeld & Christoffer Karlsson (2022) Designing and implementing a geospatial mobile application
33. Andreas Ahlström (2022) 3D-stadsmodeller för mindre kommuner - vad är behoven och vilka standarder krävs?
34. Jorun Westman (2022) Looking for shrubs in an alvar - Investigating classification of orthophotos as a way of mapping shrub species *Juniperus communis* and *Dasiphora fruticosa* on Stora Alvaret, Öland