



AALBORG UNIVERSITY
STUDENT REPORT



Remote Sensing of Eelgrass using Object Based Image Analysis and Sentinel-2 Imagery

MASTER THESIS

by

James Ormond Fethers

08 June 2018



AALBORG UNIVERSITY
STUDENT REPORT

**Master of Science (MSc) in Technology, Surveying,
Planning and Land Management, Cand.tech in
Geoinformatics.**

Title:

Remote Sensing of Eelgrass using Object Based
Image Analysis and Sentinel-2 Imagery

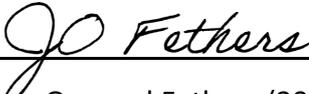
Theme: Master Thesis

Project duration: 6 Months

Submission date: 08 June 2018

University Supervisor: Jamal Jokar Arsanjani

Workplace Supervisor: Mikkel Lydholm Rasmussen


James Ormond Fethers (20161166)

Number of pages: 95

Abstract

Eelgrass (*Zostera marina*) is an ecologically significant and fragile species of seagrass common in Denmark and the Northern Hemisphere. The depth limit of the eelgrass populations is used to evaluate the ecological condition of coastal waters, and coverage is used for evaluation of ecosystem health. Satellite remote sensing has the potential to improve the cost effectiveness of the analysis significantly. Based on a review of existing reported methods, this thesis used Sentinel-2 imagery with object based image analysis and various machine learning algorithms to classify submerged aquatic vegetation at Roskilde Fjord. An ecological model of eelgrass stress parameters was applied to the classification output to produce an empirical classification of eelgrass coverage. The results indicate that Random Forest is the most suitable machine learning algorithm for submerged aquatic vegetation classification, and a scale parameter of 10 produces image objects that obtain the highest classification accuracy. Water column correction and multi-temporal analysis are demonstrated as techniques to improve classification accuracy. The thesis concludes Sentinel-2 imagery may be used for mapping submerged aquatic vegetation but not for the specific identification and analysis of eelgrass.

Preface

This thesis is written by *James Ormond Fethers*, for the Master of Science (MSc) in Technology, Surveying, Planning and Land Management, Cand.tech in Geoinformatics, Aalborg University, Copenhagen.

The thesis was undertaken from 01 February to 08 June 2018 as a collaboration with DHI GRAS, Copenhagen. All chapters are identified by name and corresponding number. All sources are presented in the Bibliography, which is ordered alphabetically. The thesis topic, to evaluate the feasibility of Sentinel-2 Imagery to classify eelgrass, was proposed by my supervisors at DHI GRAS, and has been an educational and rewarding project.

I would like to express my sincere gratitude to several people. Firstly, I would like to thank my supervisor, Associate Professor *Jamal Jokar Arsanjani*, for his guidance and support. Secondly, I would like to thank *Rasmus Eskerod Borgstrøm* for the opportunity to undertake this project with DHI GRAS, and the team at DHI GRAS – particularly *Mikkel Lydholm Rasmussen*, for professional and patient guidance. Lastly, I would like to thank my beautiful wife, *Anne*, for giving me the time and support to write this thesis.

Acronyms

EO	Earth observation
ESA	European Space Agency
EU	European Union
ML	Machine learning
MLA	Machine learning algorithm
MSI	Multispectral instruments
NB	Naïve Bayes
NIR	Near infrared
VNIR	Visible and near-infrared
SWIR	Short-wave infrared
OA	Overall accuracy
OBIA	Object based image analysis
RBF	Radial basis function
RF	Random Forest
SAV	Submerged aquatic vegetation
SVM	Support Vector Machine
SP	Segmentation parameter
WFD	Water Framework Directive

Table of Contents

1	Introduction	1
1.1 Background	2
1.1.1 Remote Sensing in aquatic environments	2
1.1.2 Remote sensing imagery	5
1.1.3 Eelgrass	7
1.1.4 Image preprocessing.....	13
1.1.5 Object based image analysis.....	14
1.1.6 eCognition.....	16
1.1.7 Segmentation.....	16
1.1.8 Machine Learning Algorithms.....	17
1.1.9 Accuracy assessment.....	24
1.1.10	... Water column correction	26
1.1.11	... Multi-temporal analysis.....	27
1.1.12	... Outlook	28
1.2 Problem statement.....	29
1.3 Research questions.....	30
2	Study Area and Data	31
2.1 Study Area.....	31
2.2 Imagery	34
2.2.1 Sentinel-2.....	34
2.2.2 Orthophoto	38
2.3 Training data	38
2.4 Software.....	41
3	Methods	42
3.1 Preprocessing.....	42
3.2 Depth raster	43
3.3 Water Column Correction.....	44
3.4 Object Based Image Analysis	46
3.4.1 Separate land/water/deep	48
3.4.2 Segmentation.....	52
3.4.3 Classification	56

3.5..... Accuracy Assessment.....	59
3.6..... Eelgrass stress model.....	59
4 Results and Discussion	62
4.1..... What is the influence of varying the scale parameter on classification accuracy?...	62
4.2..... Which machine learning algorithm performs best for classification of SAV?.....	66
4.3..... Can Water Column Correction and Multi-Temporal Analysis to improve classification results?	74
4.4..... How can classified SAV be used to guide researchers to find Eelgrass?.....	79
5 Conclusion	82
6 Future direction	84
7 References	88
8 Appendix 1	95

List of Figures

Figure 1: Challenges remote sensing over a water body. (source: http://www.dmu.dk).....	4
Figure 2: Sentinel-2 spectral bands (source: (Noi & Kappas, 2017)	7
Figure 3: Eelgrass (<i>Zostera Marina</i>) is a widespread and ecologically significant species of seagrass.....	8
Figure 4: Relationship between pixel size and object of interest is important when selecting classifier: (a) pixel larger than objects, use sub pixel classifier (b) pixel and objects same size, use pixel classifier (c) pixels significantly smaller than objects, use object classification (source: (Blaschke, 2010)).....	14
Figure 5: Flow diagram illustrating the concept of multiresolution segmentation in eCognition (source: (Trimble Germany GmbH, 2016))	17
Figure 6: Parametric linear model (L) and nonparametric spline model (R). Note the superior fit of the model to the data in the nonparametric approach (source: (James, et al., 2013)) .	19
Figure 7: Linear SVM. The hyperplane is shown as a solid line, the margins are the distance from the solid line to either of the dashed lines, the support vectors are the points within the margins (source: (James, et al., 2013))	22
Figure 8: An SVM with an RBF kernel on non-linear data does a good job separating the two classes. The hyperplane is shown as a solid line, and the margins by the dashed line (source: (James, et al., 2013))	23
Figure 9: Location of Roskilde Fjord.....	31
Figure 10: Sentinel-2 imagery clipped to AOI for (L to R) 12May16, 24Jul16, and 12Sept16 .	36
Figure 11: 12May16 RGB image with percentage clip stretch applied. Clumps of SAV and sand are visible	37
Figure 12: High resolution orthophoto of the southern fjord (L) and zoomed in on visible SAV (R).....	38
Figure 13: Diver transect data from Danish National Monitoring Program	39
Figure 14: Creating training data from orthophoto imagery.....	39
Figure 15: Creating validation samples with Sentinel-2 (L) and Orthophoto (R)	41
Figure 16: Preprocessing workflow.....	42
Figure 17: Depth raster clipped to study area	44
Figure 18: Placing points to extract depths for WCC.....	45
Figure 19: The OBIA workflow	47
Figure 20: eCognition process tree	47
Figure 21: Image bands and depth raster loaded into eCognition	48
Figure 22: Process tree for separating land/water/deep	48
Figure 23: Determining the NIR threshold to separate land and water	49
Figure 24: Misclassified water – clouds (L) and washed up reeds (R)	50

Figure 25: Map with land and water separated	51
Figure 26: Depth raster applied with 4m threshold	52
Figure 27: Multispectral segmentation algorithm was used to segment the water feature class	53
Figure 28: Band 3 (green) showing SAV in shallow and deep areas.....	53
Figure 29: Multiresolution segmentation parameter settings	54
Figure 30: SAV growing along the banks of the fjord. Suspected eelgrass due to growth patterns.....	55
Figure 31: Segmentation of shallow (L) and deep (R) areas	56
Figure 32: Process Tree to undertake classification with MLAs	56
Figure 33: Water feature class classified with 'SAV' and 'Sand' samples visible as green and yellow objects	57
Figure 34: Classifier settings are input during MLA training	57
Figure 35: RF Classification applied	58
Figure 36: Percent values classified as Low, Medium, and High	60
Figure 37: Eelgrass stress model for north fjord, with biomass interpreted into percent coverage.....	60
Figure 38: Overlaying the eelgrass stress model to the SAV classification	61
Figure 39: Eelgrass stress model overlaid on SAV feature class.....	61
Figure 40: Segmentation for SP10 showing variation in image object size between deep and shallow	65
Figure 41: Comparing the overall classification for the 3 MLA used in the study - SVM, RF and Bayes	67
Figure 42: SVM classification map	69
Figure 43: Roskilde Fjord with water column correction applied. Moving clockwise L to R: the entire fjord, the bottleneck, the northern fjord with image stretch applied.....	75
Figure 44:WCC classification with SVM	76
Figure 45: Multi-temporal analysis, with classification overlaid on Sep2016 image	78
Figure 46: Classification with eelgrass modelled stressors overlaid	80
Figure 47: Python based software that detects eelgrass from blade shape (source: (Sengupta , et al., 2018))	85
Figure 48: Batch file calling gdalwarp to resample each band to 10m resolution using nearest neighbour.....	95
Figure 49: gdalbuildvrt command, with -separate, to build a virtual raster placing each input .jp2 image into a separate band	95
Figure 50: gdal_merge.py is a python command that automatically mosaics a set of images	95

Figure 51: gdalwarp command used to clip .vrt with AOI.shp95

List of Tables

Table 1: Confusion matrix	25
Table 2: Aquatic plants in Roskilde Fjord	33
Table 3: Factors that influence water turbidity in Roskilde Fjord	34
Table 4: Overall accuracy for the classifications undertaken with RF at various SP settings. Note: shape and compactness were consistent at 0.5 each	64
Table 5: SVM confusion matrix C=10,000, gamma=0.001	69
Table 6: SVM parameter settings C=1,000,000 and gamma= 0.00001	71
Table 7: SVM parameter settings C=1,000 and gamma= 0.001	71
Table 8: SVM parameter settings C=500 and gamma= 0.001	72
Table 9: RF confusion matrix (nTrees= 500)	73
Table 10: Bayes confusion matrix	74
Table 11: Confusion matrix of SVM classifier with WCC	77
Table 12: Confusion matrix for multi-temporal analysis with SVM classifier	79

1 Introduction

‘Due to the important role that seagrasses play in estuaries, there has been a considerable effort at developing sampling and mapping techniques to quantify the spatial distribution, biomass and health of seagrass communities, and monitor changes over time. Remote sensing approaches have seen increasing application to the mapping of seagrass beds due to their synoptic perspective and cost-effective mapping over large areas’

(Lathrop, et al., 2006).

Eelgrass (*Zostera marina*) is an ecologically significant and fragile species of submerged aquatic vegetation (SAV) that is widespread around Denmark and the Northern Hemisphere. It is of particular interest at regional and state levels because eelgrass depth limit is used as an environmental indicator to evaluate the ecological condition of coastal waters under the EU Water Framework Directive (WFD). The depth limit is surveyed by field sampling methods, which although serve the purposes for EU reporting; are expensive, and ineffective for capturing representative information about the diversity of an eelgrass area (Hossain, et al., 2015). Information about eelgrass coverage can fill this void, and provide guidance for monitoring population dynamics for directing policy and conservation efforts.

The aim of this thesis is to determine whether Sentinel-2 imagery can be used to classify eelgrass coverage, and be used to direct field data collection. The methodology in this thesis is desirable because Sentinel-2 is freely available with good temporal coverage, and remote sensing is a fast and affordable technique for classification over large spatial and temporal scales. Although airborne imagery finds wider application in seagrass monitoring (Lathrop, et al., 2006), satellite imagery such as Sentinel-2 is beneficial due to superior cost and temporal coverage. This thesis also addresses a limitation of Sentinel-2 imagery – the moderate spectral resolution, which is insufficient to distinguish SAV at a

species level, by incorporating an ecological model with remote sensing classification results.

The methodology in this thesis incorporates classification with Object Based Image Analysis (OBIA) and machine learning algorithms (MLAs), which have been the focus of many studies in recent times (Qian, et al., 2015). OBIA combined with MLAs offer superior classification accuracy compared with more traditional methods, such as pixel based classification, however the challenge is optimising the vast array of variables that can influence classification accuracy. This thesis addresses several of these variables, such as MLA selection and segmentation parameter (SP) settings, by building on results from previous studies to optimise these techniques.

This thesis was undertaken in collaboration with DHI GRAS, a remote sensing specialist organisation, to evaluate the feasibility of this technique for method and data evaluation purposes, to enable a validated work history of sentinel-2 classification of marine habitats.

1.1 Background

1.1.1 Remote Sensing in aquatic environments

Remote sensing originated in the 1840's, where cameras attached to balloons were used for topographic mapping applications (Lavendar & Lavendar, 2016). Remote sensing cameras were later mounted to aircraft and were used for reconnaissance and surveillance during World War One, followed by the first images from space with the launch of the Russian V-2 rockets. The first photos acquired from an orbiting satellite were from the US military satellites in the 1960's, which acquired images using film that were dropped back to earth in re-entry capsules and caught mid-air by airplanes (Lavendar & Lavendar, 2016). The 1960's heralded several satellite missions, which culminated in the launch of Apollo 9 in 1968 that captured the first

multispectral image using a four-lens camera. In 1972 Landsat 1 was launched, which represented the first continuous collection of earth observation (EO) data to support research. Since this mission, numerous other satellite missions have been launched, equipped with a variety of sensors that provide continuously improving spatial, temporal and spectral resolution imagery. In addition to satellites, other forms of remote sensing exist, including high resolution cameras mounted to fixed wing aircraft, helicopters and drones.

The principle of remote sensing involves the detection of energy reflected or emitted from the earth as electromagnetic radiation (EMR) (Lavendar & Lavendar, 2016). Light is a form of EMR that exists in a wide range of wavelengths, from high energy radiation such as gamma rays and X-rays, to lower energy forms ranging from UV, visible light, and radio waves. The wavebands that can penetrate water are the visible, UV, and short infrared bands (Hossain, et al., 2015), which makes these bands particularly useful for remote sensing in marine environments. Visible bands penetrate further into the water column than UV and short infrared, which makes these bands useful for detecting bathymetric features. For detecting SAV, the green waveband is considered the best (Klemas, 2016), due to the combined low attenuation in the water column and high reflectivity from chlorophyll pigments in vegetation.

Remote sensing is a desirable method for spatial monitoring, because it can cost and time effectively cover land and inaccessible areas with high frequency (O'Neill, et al., 2011). Marine environments are often characterised by large areas and limited accessibility, therefore remote sensing is particularly desirable as it can monitor and assess short and long-term changes and trends faster, more completely and at lower cost per unit area than field or ship surveys (Klemas, 2016). A rule of thumb for determining whether remote sensing of SAV within a water body may be feasible is whether bottom features are visually distinguishable from a boat, or visible in remote sensing imagery.

In the past four decades, rapid technological and methodological advancements have occurred in the field of remote sensing (Hossain, et al., 2015), with the emergence of new satellites, sensors and data analysis techniques that are highly effective for the monitoring of coastal processes and features (Klemas, 2016), including seagrass habitats (Hossain, et al., 2015). However, despite these advancements in technology, no single method is suitable to measure all seagrass parameters (Hossain, et al., 2015), and therefore it is important to choose remote sensors and data analysis methods that are most appropriate for the specific seagrass study (Klemas, 2016).

Remote sensing in aquatic environments is complicated by several factors, such as atmospheric interference, variability in water depth and bottom albedo, and water column attenuation by scattering and absorption (Figure 1) (Cho, et al., 2012). Accounting for water column attenuation is considered one of the greatest challenges (Yang, et al., 2010), due to the large number of variables associated with absorption and scattering; which increase with depth, concentrations of suspended particles, chlorophyll and dissolved organic matter (Visser, et al., 2013) (Yang, et al., 2010).

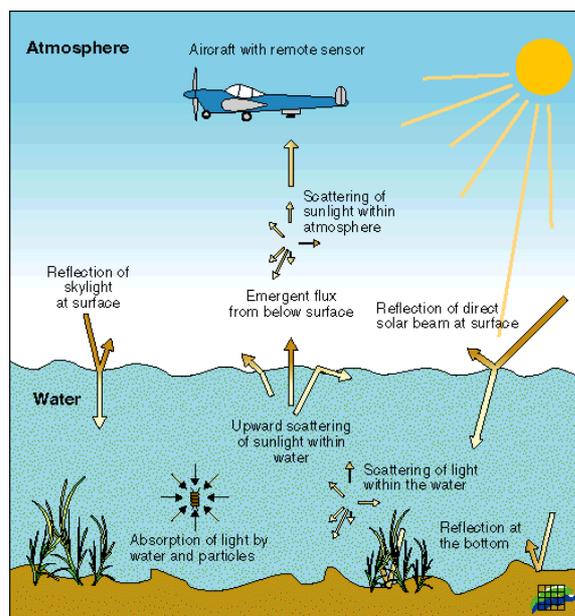


Figure 1: Challenges remote sensing over a water body.

(source: <http://www.dmu.dk>)

The water column effect can be mitigated by undertaking remote sensing in clear, shallow water where seagrasses grow in dense meadows and constitute the only dark features on a sandy bottom, and/or select clear, sunny days to obtain the imagery (Krause-Jensen, et al., 2004). These mitigation factors may not always be feasible however, in which case a water column correction (WCC) may be necessary.

Another challenge with remote sensing in aquatic environments is the separation of the spectral signature for the features of interest, which can be confused particularly by dark features, such as mussel beds, stones or macroalgae (Krause-Jensen, et al., 2004). This was demonstrated in a study by (Kuusemäe, et al., 2016), who reported that macroalgae beds that are also green can be interpreted as eelgrass beds due to the small differences in spectral colour values (Kuusemäe, et al., 2016). Other challenges include movement of the water surface due to wind or currents, and tidal influence which can significantly vary the depth of the water body.

SAV beds, including eelgrass, have high spatial complexity and temporal variability, which requires high spatial, spectral and temporal resolutions to be observed effectively (Klemas, 2016). (Visser, et al., 2013) demonstrated a solution to this problem, by using OBIA to distinguish features of interest based on spatial and textural information, however these trials were conducted in test pools with high resolution imagery at an individual plant level, therefore unlikely to have application in imagery with lower spatial resolution.

1.1.2 Remote sensing imagery

The recent advancements in multispectral and hyperspectral sensors have resulted in an increase in imagery with fine spatial (0.4-4m) and/or spectral (200 narrow bands) resolution (Klemas, 2016). The choice of multispectral and hyperspectral imagery is important, particularly for remote

sensing in marine environments, because hyperspectral imagery enables mapping SAV at a species level, as demonstrated by Phinn, et al. (2008), who used airborne hyper-spectral (CASI-2 sensor using a pixel size of 4.0 m) imagery for mapping seagrass species composition; and Roelfsema, et al., (2014), who combined field data with high spatial resolution imagery (Worldview-2 – 0.5m resolution) in a time series analysis to produce seagrass species and percentage cover maps on a landscape scale (>100km²). Furthermore, the use of hyperspectral imagery can improve classification accuracy, as demonstrated by O’Neill, et al., (2011), who conducted classification of eelgrass using hyperspectral imagery and achieved an overall accuracy of over 85%.

One of the main constraints with hyperspectral imagery is the high cost (Roelfsema, et al., 2014). Therefore, the selection of appropriate imagery depends on funding and the desired outcome of the classification.

Sentinel-2 is an EO mission by the European Space Agency (ESA) that consists of two multispectral satellites – Sentinel 2A and 2B. Launched on 23 June 2015 and 07 March 2017 respectively, the satellites are equipped with Multispectral Instruments (MSI) that sample 13 spectral bands (Figure 2). Sentinel-2 has moderate spatial resolution, which makes it suitable for a variety of purposes, such as vegetation monitoring, soil and water cover, and observation of waterways.

Data are acquired on 13 spectral bands in the VNIR and SWIR (European Space Agency, 2018), at the following spatial resolutions (Figure 2):

- four bands at 10 m: Band 2, Band 3, Band 4, and Band 8
- six bands at 20 m: Band 5, Band 6, Band 7, Band 8a, Band 11, and Band 12
- three bands at 60 m: Band 1, Band 9, and Band 10

Spectral Band	Center Wavelength (nm)	Band Width (nm)	Spatial Resolution (m)
Band 1	443	20	60
Band 2	490	65	10
Band 3	560	35	10
Band 4	665	30	10
Band 5	705	15	20
Band 6	740	15	20
Band 7	783	20	20
Band 8	842	115	10
Band 8a	865	20	20
Band 9	945	20	60
Band 10	1380	30	60
Band 11	1610	90	20
Band 12	2190	180	20

Figure 2: Sentinel-2 spectral bands (source: (Noi & Kappas, 2017))

Sentinel-2 data is available for download with or without atmospheric correction, as a Level 2A or 1C product respectively. Level 1C products have undergone radiometric and geometric corrections, including ortho-rectification and spatial registration, whereas 2A products have undergone additional atmospheric correction (European Space Agency, 2018).

1.1.3 Eelgrass

Eelgrass (*Zostera marina*) (Figure 3) is a common species of seagrass widespread in Denmark and the Northern Hemisphere, which serves as a keystone component of many marine ecosystems (O'Neill, et al., 2011). It serves as an important link in many food chains because it harnesses energy and nutrients that are consumed by organisms, provides critical structural components in benthic environments, and serves as essential habitat for many organisms such as shellfish, crabs, fish and waterbirds (Lathrop, et al., 2006). These ecological functions also have economic benefits, with the production of goods such as shellfish and finfish (Terrados & Borum, 2004).



Figure 3: Eelgrass (Zostera Marina) is a widespread and ecologically significant species of seagrass

Eelgrass meadows can decrease the damaging effects of waves and reduce shoreline erosion (Klemas, 2016). Due to the wide variety of functions of eelgrass, and high sensitivity to degraded water quality (Lathrop, et al., 2006), it is intimately linked to the health of the wider marine environment (Duarte, et al., 2004) and serves as an indicator of the status of the coastal zone (Terrados & Borum, 2004).

Eelgrass also has historical significance, as it has previously been used for stuffing mattresses and cushions, for the construction of dikes for land and sea defences (Haynes, 2000), and is the only known example of a grain that has been harvested from the sea for human consumption (Felger & Moser, 1973).

Eelgrass populations declined worldwide at a rate of 2% per year between 1990 – 2000 (Duarte, et al., 2004), and populations continue an overall trend of decline today (The IUCN Red List of Threatened Species, 2017). This is due to factors such as anthropogenic pressure (O'Neill, et al., 2011), climate change (Klemas, 2016), declining water quality, dredging, macroalgal infestation and disease (Lathrop, et al., 2006). Furthermore, eelgrass depth limit is used as an environmental indicator of waterway health under the WFD

(Kuusemäe, et al., 2016), which is an EU initiative for all member states to determine the status of all waterbodies. To understand these pressures and direct coastal ecosystem management to protect these vital ecosystems, it is important to delineate and monitor eelgrass distribution (Klemas, 2016) (O'Neill, et al., 2011), as well as obtain an understanding of species composition, richness, abundance and spatial patterns (Phinn, et al., 2008). To achieve these objectives, spatial information is required (Phinn, et al., 2008).

Eelgrass is widespread across a variety of habitats in the northern hemisphere. It is found in arctic waters along the northern Norwegian coast, where it can endure several months of ice cover, as well as in Danish fjords, to the warm waters of the Mediterranean (Kuusemäe, et al., 2016). In Danish fjords, eelgrass has undergone widespread reduction in the past century due to wasting disease and eutrophication from nearby agriculture (Kuusemäe, et al., 2016). The depth limit of eelgrass growth in Denmark is 3-5m in fjords and a few metres deeper along the open coastline (The Danish Council for Strategic Research, 2017). Light availability is the limiting factor for depth.

Knowledge of eelgrass growth and reproduction is crucial for eelgrass ecosystem management, as it provides information about recovery times (Marba, et al., 2004). Eelgrass has two mechanisms of reproduction – sexual through seed, and clonal through rhizome growth (Marba, et al., 2004). Sexual reproduction through seed occurs during plant flowering, which releases pollen into the water. Flowering is controlled by water temperature, which in Europe occurs in late spring and summer, where irradiance and water temperature is high (Marba, et al., 2004). Eelgrass plants release large quantities of seeds (thousands per m²), however as they are negatively buoyant, the seeds generally settle within a few metres of the plant and the majority are consumed by crabs and other grazers (Marba, et al., 2004). Seed may also be consumed and transported by waterbirds (Marba, et al., 2004). Clonal growth is the

primary method of seagrass bed expansion (Robbins & Bell, 1994), and the most important process for the maintenance of seagrass meadows (Cunha, et al., 2004). The speed of this process varies greatly amongst various seagrass species. For eelgrass, clonal growth is very slow (on average 16cm/yr) (Cunha, et al., 2004), which results in population recovery occurring over large timescales. During a widespread reduction from wasting disease in the 1930s (Kuusemäe, et al., 2016), recolonization was observed on a timeframe of 2-3 decades (Cunha, et al., 2004). The implications of these findings are that monitoring programmes need to be sustained over long time periods to determine whether populations are regenerating, which makes remote sensing an optimal technique.

The mechanisms hindering eelgrass recovery were identified and modelled by Kuusemäe, et al., (2016). The most significant factors were 1) lack of sediment anchoring capacity, 2) resuspension created by drifting ephemeral macroalgae, 3) seedling uproot by current and wave forces, 4) ballistic stress from attached macroalgae and 5) burial of seeds and seedlings by lugworms. These processes were quantified and modelled using an ecological DHI MIKE 3D model, which is used as a national tool to predict areas where eelgrass restoration effort may be initiated, and has been incorporated in the results of this study. Krause-Jensen, et al., (2004), also identified salinity as a significant factor, which was not included in the model.

Seagrass habitats can be visualized on multiple scales, ranging from millimetres to kilometres (Robbins & Bell, 1994). At the patch level, millimetres to centimetres, seagrass habitats are observed as individual plants to small patches on fine spatial scales. Such observations are only possible with transects or high-resolution imagery (Roelfsema, et al., 2014). At moderate scales (metres to tens of metres), eelgrass can be observed as a bed, which is a spatially contiguous area of similar percent cover composition. At the largest

resolution (hundreds of metres), eelgrass is measured as meadows, which consist of spatially contiguous areas of seagrass beds (Lathrop, et al., 2006). The limitation to visualising these scales lies with the spatial resolution of the imagery. Using 10m resolution imagery, such as with Sentinel-2, eelgrass can only be observed as beds or meadows, as the resolution is too low to observe patches.

Eelgrass coverage can be used to determine distribution and/or abundance (Krause-Jensen, et al., 2004), and detect changes (Duarte, et al., 2004). Distribution refers to where it is growing, abundance refers to cover at a specific depth. Abundance is a useful measure, because it can be directly linked with water quality, since eelgrass demonstrates depth dependence with the highest abundances typically at intermediate water depths where levels of exposure and light are moderate. Any decline in seagrass at this depth can be attributed, at least partly, to increased light attenuation in the water column and therefore changes in water quality (Duarte, et al., 2004) (Krause-Jensen, et al., 2004) – which is why this metric is used to report under the WFD. Abundance can also be used as a measure of eelgrass biomass, as demonstrated by (Carstensen, et al., 2015), who used eelgrass cover and light attenuation to determine eelgrass biomass.

A diverse range of eelgrass monitoring programs are available, ranging from volunteer collected data to more complex scientific techniques, such as remote sensing (Duarte, et al., 2004). The method of measuring eelgrass coverage depends on the monitoring objectives (Krause-Jensen, et al., 2004) and available resources (Duarte, et al., 2004). For cataloguing presence/absence or coarse area distribution, macro-scale maps with low to moderate resolution imagery is suitable. However, for detailed information about distribution, change in seagrass areas or to estimate biomass, high resolution imagery is required. Monitoring on very fine scales (cm's) requires

manual observations in conjunction with DGPS positions, which is very time and cost intensive (Krause-Jensen, et al., 2004)..

Direct observations are collected annually to determine eelgrass abundance for the Danish national monitoring programme to report under the WFD (Krause-Jensen, et al., 2004), and have been used as observation data for several remote sensing studies (e.g. (Carstensen, et al., 2015)). This method is effective in detecting declines, however is time consuming if large areas are to be covered (Duarte, et al., 2004). Observations for the Danish national monitoring program are collected by a diver who swims along depth gradients and estimates percent cover at intervals of 5-10 m. Regular observations of cover, GPS position and water depth are recorded, and the average cover within depth intervals of 1 m is calculated (Krause-Jensen, et al., 2004). Observations are recorded at the same dive sites each year, to determine yearly trends. For Roskilde Fjord, these dive locations were established in 1979 (Bruhn, et al., 2013). This method is reported as being the most repeatable, precise and cost-efficient of several methods tested (Krause-Jensen, et al., 2004), however this study aims to demonstrate that this methodology is limited, since dive locations are determined by nothing more than historical locations, and remote sensing could be used to target eelgrass meadows to ensure dive sites are representative of the study area. This hypothesis is supported by Duarte, et al., (2004), who state that programs that assess changes across entire meadows are far more effective in detecting trends than quadrat-based programs, which can only provide inferences on 'very local scales'.

The use of conventional vegetation indices, such as strong NIR reflectance, and blue and red absorbance, cannot be used for SAV since the upwelling signal from waterbodies contain several sources of interference from the water column (Cho, et al., 2012), and the NIR wavelength is absorbed

strongly by water (Klemas, 2016). As such, the green region of the spectrum is considered best for sensing submerged aquatic vegetation, followed by the red, and red edge regions (Klemas, 2016), as these bands penetrate the water column and are reflected and absorbed respectively, by the SAV. However, using these bands to discriminate SAV is only feasible in shallow water up to 0.5m deep, as the spectral signature attenuates strongly with depth (Klemas, 2016) (Cho, et al., 2012). This value reduces further with increased turbidity or surface disturbance (Cho, et al., 2012). These results are replicated by (Yang, et al., 2010) , who demonstrated the appropriate wavebands for seagrass mapping generally lay between 500 and 630 nm (blue and green) and 680 and 710 nm (red and red edge), and found a strong relationship between the reflectance value at 715 nm and Leaf Area Index. Based on the recommendations from these studies, the green, red, and red edge bands were used in this thesis.

1.1.4 Image preprocessing

Image preprocessing is a form of image rectification and restoration, that is used to correct distorted or degraded image data arising from image acquisition, to create a truer representation of the actual scene (Lillesand & Kiefer, 1999). It involves processing data using an equation, or series of equations, to form a new digital image that may be displayed or further processed. The type and degree of preprocessing is highly dependent upon the characteristics of the sensor used to capture the image, because of the variety of geometric distortions and noise (Lillesand & Kiefer, 1999).

Resampling is the process of geometrically transforming a digital image to appropriately assign the appropriate digital number to an output cell or pixel (Lillesand & Kiefer, 1999). It is necessary when working with raster data sets that have different spatial resolutions, or that use different geographic coordinate systems. Some common methods are Nearest Neighbour, Majority

Resampling, Bilinear Interpolation and Cubic Convolution; which differ broadly by whether the data type is continuous or discrete.

1.1.5 Object based image analysis

OBIA is a technique used to classify digital imagery by grouping together clusters of pixels into image objects based on homogeneity through a process of segmentation, and then classifying these objects based on spectral, textural, neighbourhood and object specific shape parameters (Ouyang, 2015) (Yoon, et al., 2003). 'A considerable amount of research has shown that object-based approaches are superior to traditional pixel-based methods in the classification of high spatial resolution data' (Qian, et al., 2015). It is useful to extract tangible information from imagery for use with GIS, and has great potential for very high resolution imagery (Zhang & Maxwell, 2006) (Drăguț, et al., 2010), where pixel size is significantly smaller than objects of interest (Figure 4), which is increasing in prevalence with the improvements in satellite sensor resolution (Blaschke, 2010).

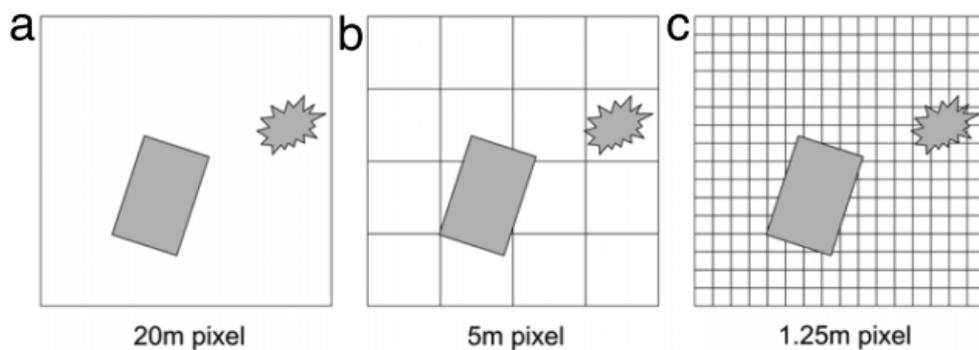


Figure 4: Relationship between pixel size and object of interest is important when selecting classifier: (a) pixel larger than objects, use sub pixel classifier (b) pixel and objects same size, use pixel classifier (c) pixels significantly smaller than objects, use object classification (source: (Blaschke, 2010)).

An advantage of OBIA compared with pixel based methodologies, is it can overcome the salt-and-pepper effect, which is caused by high local spatial heterogeneity between neighbouring pixels (Lillesand & Kiefer, 1999). OBIA

clusters pixels together to remove small scale variation, and results in faster processes because objects, not individual pixels, are classified (Ouyang, 2015).

OBIA hinges on image segmentation, which originated in the 1970's where it was used for industrial image processing. It has gained prominence since the mid-2000s, as evidenced by the sharp increase in peer reviewed journal articles since this time (Blaschke, 2010). Image objects, which enables a user to define the scale at which image features are analysed, are built in a hierarchical structure that enables the user to display image object information at different scales simultaneously (Baatz & Schape, 1999).

OBIA has great potential for classification in marine environments, where spectral information may be distorted, and over large areas. This has been demonstrated in several studies, such as Roelfsema, et al., (2014), who used OBIA and high spatial resolution imagery in a timeseries approach to map seagrass over a 142km² area; and Lathrop, et al., (2006), who used OBIA to map SAV – eelgrass and *Ruppia maritima* (species which are also present in Roskilde Fjord), over a 36,000ha study site.

OBIA was used for this study to measure eelgrass coverage on a landscape scale. Conventional classification techniques are difficult to apply for this purpose because of the inconsistency in spectral response for SAV beds, that change with water depth and turbidity (Lathrop, et al., 2006). Several studies have demonstrated successful use of OBIA techniques with medium and coarse resolution imagery, such as Myint, et al., (2008), who used Landsat TM (28.5m resolution) to conduct change detection following cyclone damage, and demonstrated a 15-20% higher accuracy with the object-oriented approach compared with pixel-based approaches. OBIA also provided the opportunity for knowledge about eelgrass growth to be incorporated into the classification by

specifying the size and shape of the image objects, as demonstrated by Visser, et al., (2013).

1.1.6 eCognition

The emergence of OBIA coincided with the launch of the eCognition software by Trimble in 2000, which was the first commercially available software that enabled image processing and GIS functionality in an object based environment (Blaschke, 2010). This software was later renamed 'Definiens', and then reverted back to 'eCognition', which it is known as today. It is based on the approach originally known as Fractal Net Evolution (Baatz & Schape, 1999) (Blaschke, 2010), which encompasses the formation of image objects using contextual information such as form and texture. This was built into a programmable, user friendly workflow. This software has been used in a significant number of studies, and has led to the development of numerous other similar software applications (Blaschke, 2010).

1.1.7 Segmentation

Image object segmentation is a crucial process as it significantly influences classification results (Zhang & Maxwell, 2006). Shape and texture can be incorporated into the image object SP settings, and may produce better classification accuracy than differentiation of spectral signatures with very high-resolution imagery (Visser, et al., 2013). The basic strategy is to build up a hierarchical network of image objects which allows image information to be represented at different resolutions (scales) simultaneously.

Various segmentation algorithms are available in eCognition, which are broadly divided by whether they operate on a top-down or bottom-up approach (Trimble Germany GmbH, 2016). Multi-resolution segmentation is a commonly used bottom-up approach, which consecutively merges pixels or existing image objects into larger ones, based on the criteria of relative

homogeneity (Qian et al., 2015), and was used in this thesis. Segmentation can be influenced with various scale parameter criteria, such as shape, compactness and scale parameter (Figure 5).

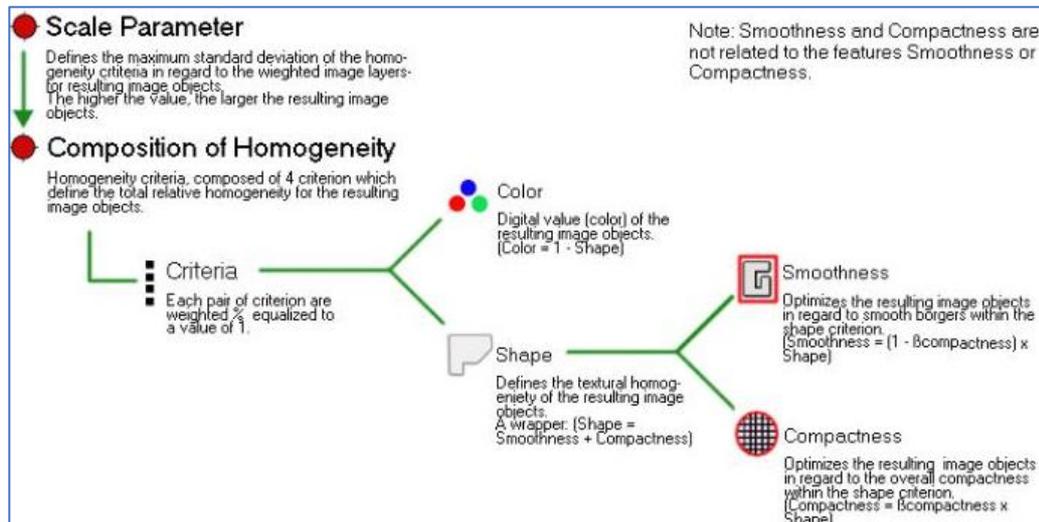


Figure 5: Flow diagram illustrating the concept of multiresolution segmentation in eCognition (source: (Trimble Germany GmbH, 2016))

These parameters are usually set using a process of trial and error, with the objective to produce image objects that only contain a feature of interest. Several studies have developed more objective techniques to set these parameters, such as Zhang & Maxwell, (2006), who used fuzzy logic to determine suitable object SPs; and Drăguț, et al., (2010), who developed a tool for estimating the scale parameter setting that minimised the rate of variance change within image objects. This thesis did not use these techniques; however, it builds on the research by investigating the impact of adjusting the scale parameter on overall classification accuracy.

1.1.8 Machine Learning Algorithms

Machine learning (ML) is a branch of computer science that enables computer systems to learn (i.e. progressively improve performance on a specific task) the underlying behaviour of a system from a set of training data

without being explicitly programmed. It consists of a variety of algorithms that can provide various forms of regression (multivariate, nonlinear and nonparametric) and classification (supervised or unsupervised) (Lary, et al., 2016). It has a wide variety of uses in areas where implementing algorithms is difficult or infeasible, such as email filtering, optical character recognition and computer vision.

ML provides an effective empirical approach for use in nonlinear systems, such as remote sensing, because of its ability to learn underlying trends in a system through the input of training data (Lary, et al., 2016). It has increasing interest in conjunction with OBIA to determine the best performing algorithm (Qian, et al., 2015). ML with training samples enables classification of image objects based on underlying trends in the system which are unknown by the operator (Qian, et al., 2015), which was the approach used in this thesis.

A principle underpinning all supervised ML is the equation:

$$Y = f(X)$$

Where Y is the output variable, f is the functional form of the model, and X is the input variables (James, et al., 2013).

The shape of the model, f , determines how data points are separated and therefore how they are classified (James, et al., 2013). Defining the functional form of f is the purpose of the ML algorithm and the reason why selecting various ML algorithms result in different classification results. If f was already known, it could be used directly in the classification and would not need to be learned using ML algorithms. Central issues in ML are avoidance of overfitting, and balance between simplicity and fit to data.

MLA can be broadly divided as parametric and non-parametric, which describes whether the function used to separate the data for classification makes assumptions about the form of the function f (parametric) or not (non-parametric). A comparison between parametric and non-parametric methods is illustrated at Figure 6, where linear and spline models have been fitted to the data. It can be observed the linear fit is not quite accurate, with the true f demonstrating some curvature which is well captured by the spline approach. This principle applied to spatial data would result in higher classification accuracy for the non-parametric approach.

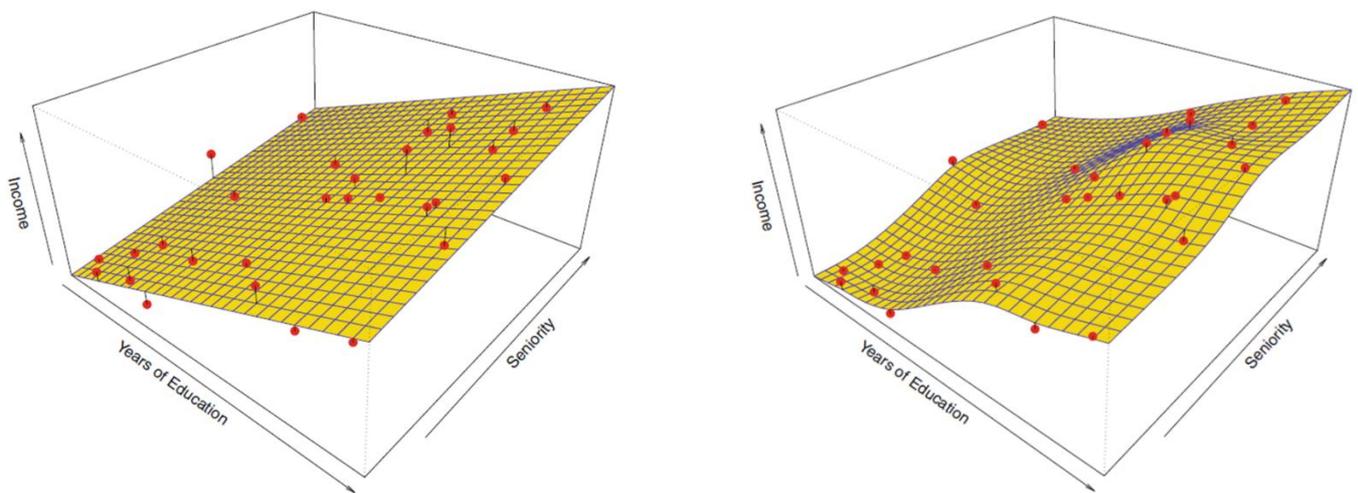


Figure 6: Parametric linear model (L) and nonparametric spline model (R). Note the superior fit of the model to the data in the nonparametric approach (source: (James, et al., 2013))

The MLAs used in this thesis were Naïve Bayes (NB) (parametric), Random Trees (Random Forest – RF) and Support Vector Machine (SVM) (both nonparametric), which are among the most widely used in remote sensing (Qian et al., 2015). It is valuable to trial different MLAs with a dataset to observe how the algorithm interacts with different types of data, and which can provide the more accurate result. A challenge with using various MLAs is determining the optimal tuning of indices, as demonstrated by (Qian et al., 2015).

Parametric models involve a two-step approach, where firstly an assumption is made about the functional form of f , and secondly the training data is used to fit or train the model (James, et al., 2013). Parametric algorithms are relatively simple to implement as they do not require many input parameters, and have fast processing speeds. The simplest form of a parametric model is linear, which assumes a linear relationship between the variables. The training data determines the intercept and slope of f to produce the predictive model. Although this method is simple to apply, it is susceptible to inaccurately representing underlying trends if it does not accurately reflect the relationship amongst the variables, and is best suited to more simplistic problems. Inaccuracy can be addressed by choosing flexible models that fit multiple forms of f to the data, however this requires estimating a greater number of parameters and can lead to overfitting of data, where errors or noise are followed too closely (James, et al., 2013). In relation to remote sensing, parametric models are likely only suitable in images with well defined spectral classes and with a suitable amount of training data (Qian, et al., 2015). In the case of complex imagery where classes are a mix of spectral values, or where training data is insufficient, these methods are likely too simplistic to accurately classify an image.

The NB classifier is a simple probabilistic classifier based on Bayesian statistics, which assumes the strong independence of variables by assuming that the presence or absence of a particular feature, is unrelated to the presence or absence of another (Trimble Germany GmbH, 2016). Therefore, only a small amount of training data is required for classification, because only the variances of the variables need to be determined, rather than the entire covariance matrix (Trimble Germany GmbH, 2016). NB can be a practical choice, as it can achieve similarly high accuracy to SVM without the need to set any tuning parameter, however is sensitive to training sample size (Qian, et al., 2015).

Nonparametric algorithms do not assume the form of f , and instead use the data points to make an estimate of f that best fits the data values and produces a smooth fit. This enables an accurate fit for a wide range of possible f shapes and more accurately reflect complex relationships than parametric approaches, which is the major advantage of this approach. A disadvantage however is that since estimating f is not reduced to a small number of parameters, as with parametric approaches, a larger number of observations is usually required to obtain an accurate estimate of f (Qian, et al., 2015) (James, et al., 2013). The non-parametric algorithms used in this thesis, RF and SVM, are the foremost at producing high classification accuracy (Noi & Kappas, 2017). SVM is less sensitive to training sample size than RF (Noi & Kappas, 2017), however SVM is particularly sensitive to tuning parameter settings (Qian, et al., 2015).

RF is based on the Decision Trees algorithm, which builds an ensemble of decision trees and merges them together using the bagging method. In this way, RF provides an accurate representation of the data during classification because instead of examining a single node from a single tree, it examines a random subset of nodes from multiple trees to decide on the most appropriate way to split the node, and from here, builds a new tree based on these decisions. In this way, it incorporates a broad subset of training data into the classification, and provides a result representative of the input data.

RF default parameters often produce accurate prediction results (Noi & Kappas, 2017) which makes them relatively fast and simple to implement. Two key parameters are the number of trees (n_{Tree}) and the number of features incorporated in each node (m_{try}). Although several studies have stated that the default parameters provide satisfactory results (Duro, et al., 2012) (Liaw & Wiener, 2002), Noi & Kappas, (2017), demonstrated that n_{Tree} 500 was a

threshold at which classification accuracy stopped improving. This setting was used in this thesis.

Support vector machine (SVM) are widely considered one of the best ‘out of the box’ classifiers (James, et al., 2013). A SVM works by constructing a hyperplane, or a set of hyperplanes, to create maximum separation between data classes (Figure 7).

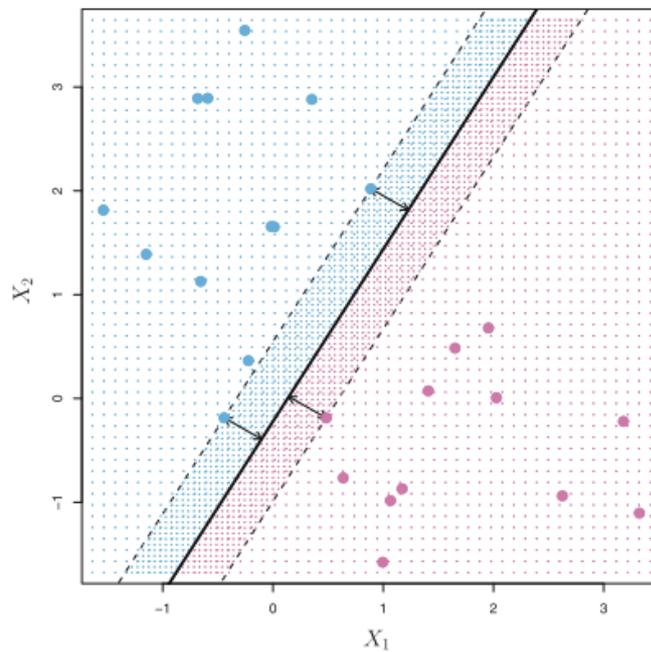


Figure 7: Linear SVM. The hyperplane is shown as a solid line, the margins are the distance from the solid line to either of the dashed lines, the support vectors are the points within the margins (source: (James, et al., 2013))

SVM is good for classifying high dimensional data sets, and it works well on small data sets. The disadvantages are that choosing the correct kernel and parameter settings can be time consuming and computationally intensive. SVM is the least sensitive algorithm to sample sizes because it uses only the support vectors, which are the data that violate the margins (Figure 7), instead of all training samples, to build the separating hyperplane (Qian, et al., 2015). The width of the margins are effectively determined by cost (C), as it determines how tolerant the margins are to violations. A smaller C value is less tolerant to

violations; therefore, the margins are narrower and have a smooth decision boundary, and vice versa with a larger C value. A problem with SVM is overfitting, which is increased with high C values (Qian, et al., 2015). The important aspect of the training samples is to ensure they cover the full spectrum of pixel values in an image, as the extreme class values form the support vectors that determine f . Therefore, adding more training samples may not significantly affect the classification accuracy (Qian, et al., 2015). Setting C is thus a trade-off between a smooth decision boundary and one that classifies all points correctly.

To map non-linear decision boundaries (i.e. data points that cannot be divided linearly), several SVM kernel functions are available, including linear, polynomial, radial basis function (RBF) (Figure 8) and sigmoid kernels. The RBF kernel has been demonstrated to be superior in a variety of studies (Qian, et al., 2015), and was used in this thesis.

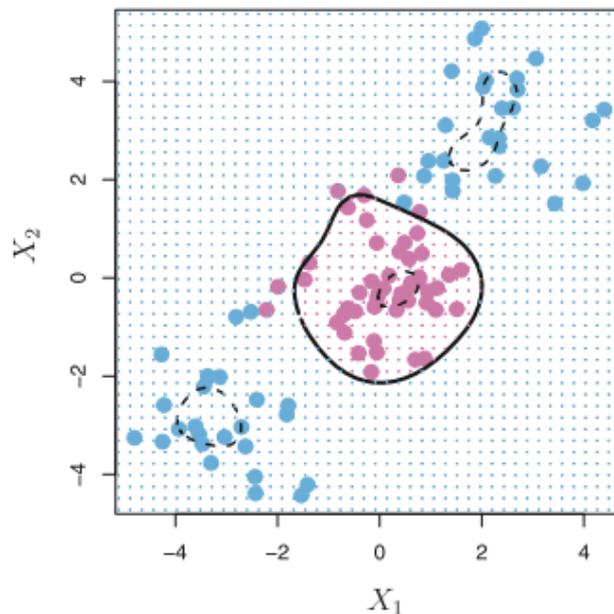


Figure 8: An SVM with an RBF kernel on non-linear data does a good job separating the two classes. The hyperplane is shown as a solid line, and the margins by the dashed line (source: (James, et al., 2013))

The main parameter settings for the RBF kernel is C and gamma. Gamma influences the shape of the separating hyperplane (f) (Qian, et al., 2015), by defining how far the influence of a single training example reaches. Low values have greater influence from far data points, whereas high values have greater influence from close data. Consequently, high gamma values ignore further away points, and can result in a wiggly decision boundary because the points nearby have a large input on the shape, and consequently this setting is susceptible to overfitting data. Low gamma values have the opposite effect, and account for more of the data, and result in a straighter line. Overall, the RBF kernel has very local behaviour, in the sense that only nearby training observations influence the hyperplane (James, et al., 2013).

1.1.9 Accuracy assessment

Accuracy of a classification is expressed using a confusion matrix (Table 1). It indicates how well a classification has categorized a representative subset of pixels used in the training process of a classification, by comparing, on a category-by-category basis, the relationship between known reference data (ground truth) and corresponding results of an automated classification (Lillesand & Kiefer, 1999). The number of rows and columns are equal to the number of categories for the classification being assessed, therefore a confusion matrix is always square. Accuracy assessment is an essential component of remote sensing, which is re-enforced by a quote by (Lillesand & Kiefer, 1999) – ‘A classification is not complete until its accuracy is assessed’

The training cover types are listed in columns, and the pixels classified by the classifier are listed horizontally (Table 1), enabling interpretation of the performance of the classification.

Table 1: Confusion matrix

		Truth			
Predicted	ClassValue	Sand	SAV	Total	U_Accuracy
	Sand	46	28	74	0,622
	SAV	6	39	45	0,867
	Total	52	67	119	0
	P_Accuracy	0,885	0,582	0	0,714

Running along the major diagonal of the error matrix, from upper left to lower right (shaded cells), are the training pixels that were correctly classified (Table 1). Either side of this major diagonal are the classification errors of omission (exclusion) and commission (inclusion), with omission elements displayed in columns and inclusion errors horizontally in rows.

The producer and user accuracies are displayed around the periphery of the table (Table 1). The producer accuracy is derived from dividing the total number of correctly classified pixels by the total number of reference pixels, and provides an indication of how well a particular cover type was classified. The user accuracy is calculated by dividing the number of correctly classified pixels by the total number of pixels within that category. This provides an indication of the probability that a pixel classified in a particular category actually represents that category on the ground (Lillesand & Kiefer, 1999).

The overall accuracy is a measure of the total number of correctly classified pixels, divided by the total number of reference pixels. Caution should be taken when interpreting overall accuracy, because it is an average all classes and can be skewed by particular classes, therefore it does not reveal the distribution between classes (Lillesand & Kiefer, 1999). Overall accuracy should be interpreted in conjunction with user and producer accuracies.

Kappa is an accuracy measure that reflects the difference between actual agreement and agreement expected by chance. This measure was excluded from this thesis as recent peer reviewed articles suggest that the basis for kappa accuracy is fundamentally flawed and should not be used (Foody, 2011) (Lyons, et al., 2012).

Accuracy assessment with validation samples should be undertaken using random sampling, to ensure sampling biases are not present, and the samples are representative of the data set under analysis (Lillesand & Kiefer, 1999). Outlined below are techniques for creating random sample points:

- Simple random sampling: sites are generated independently, tends to under sample small but potentially important areas;
- Stratified random sampling: the study area is split into strata and random samples are generated in each stratum. Strata are allocated samples based on proportionate size of the strata, which ensures samples are geographically representative of the data set; and,
- Equalised stratified random sampling: creates points that are randomly distributed within each strata, where each class has the same number of points. Does not weight strata size, therefore small strata can have a disproportionate weighting on the classification.

1.1.10 Water column correction

The effect of the water column inhibits quantifying seagrass cover, especially in optically shallow water where radiance can be modified by phytoplankton, suspended organic and inorganic matter and dissolved organic substances (Yang, et al., 2010). Water column correction removes the effect of the water column by normalizing an image around a relative bathymetry to remove the effect of depth related light attenuation, for applications such as seagrass habitat maps (Hossain, et al., 2015).

The most commonly cited approach is that of Lyzenga (1978), which uses a correlation between two imagery bands to generate a pseudo-colour band, which is used to normalise the image bands (Yang, et al., 2010) (Lyzenga, 1978). This approach was used in this thesis. Other WCC techniques have been developed, for example Misbari & Hashim, (2016), demonstrated the use of the bottom reflectance index for quantifying seagrass total aboveground biomass. Future studies should focus on optimising WCC for the water body of interest, as this can yield significant classification improvements.

1.1.11 Multi-temporal analysis

Multi-temporal image analysis is a classification technique that merges band combinations from separate dates, to produce a time series of atmospheric and geometrically corrected images to detect changes, such as vegetation growth (Lillesand & Kiefer, 1999). The extent to which this technique improves classification accuracy is clearly a function of the growth of the crop or vegetation type, and the number and timing of the dates used in the classification (Lillesand & Kiefer, 1999). This technique is particularly useful in aquatic environments where detecting SAV can be challenging, and to distinguish growing vegetation from other dark features, such as mussel beds, stones or macroalgae, which can confuse interpretation (Krause-Jensen, et al., 2004). Furthermore, low density meadows may be challenging to detect and sensitivity of mapping is higher for dense meadows (Krause-Jensen, et al., 2004), therefore multi-temporal analysis can improve classification as density increases throughout the growing season.

Various types of multi-temporal analysis can be used to combine multitemporal data for classification. One approach, and the approach trialled in this study, is to combine all spectral bands from all dates into a master dataset for classification. Alternatively, principle component analysis can be used to reduce the dimensionality of the combined dataset, by computing the

first three principle components of each image and merging them to create a final 6 band image for classification. This 6-band image can be stored, manipulated and classified with much greater efficiency than the original 12 band image (Lillesand & Kiefer, 1999). A third strategy is the multitemporal profile approach, where modelling is based on the behaviour of each crop's spectral response pattern over a period of time.

Multi-temporal analysis has yielded high classification accuracies for seagrass mapping on a landscape scale in multiple studies (Roelfsema, et al., 2014). Lyons, et al., (2013), used a trend and time series approach to describe the processes that drive seagrass growth and decline, with a time series of 23 annual maps and 16 monthly maps to demonstrate the inter and intra annual dynamics of seagrass populations. Both studies propose the integration of environmental data, such as water quality, with mapping outputs, to enable assessment of the impacts of management actions on seagrass meadows (Roelfsema, et al., 2014) (Lyons, et al., 2013). This thesis follows these recommendations by applying a model of eelgrass stress factors to a classification output, and proposes additional benefits of using this technique, namely to guide field researchers. This is important because time-series mapping should be used as a compliment, rather than a replacement, for on-ground monitoring efforts (Roelfsema, et al., 2014). This is reinforced by Hossain, et al., (2015), who states 'no single approach is suitable for and capable of measuring all seagrass parameters and assessing change. Integration of field data, imagery, and mapping approaches is therefore required'.

1.1.12 Outlook

There are several emerging technologies within remote sensing in aquatic environments. Significant advances in sensors and software have recently occurred, and continue to occur with the emergence of drone imagery

(Ørberg, et al., 2018). This thesis proposes more significant advancements lie within the automation of field data collection and processing, which have been demonstrated to add significant cost to monitoring programmes (Roelfsema, et al., 2014).

Field data collection in aquatic environments is usually undertaken by snorkelers, with manual photo analysis used to estimate seagrass composition and abundance (Roelfsema, et al., 2014). Sometimes observations of percent cover are undertaken in-situ, however these methods have been demonstrated to be sensitive to observer bias (Krause-Jensen, et al., 2004). Krause-Jensen, et al., (2004), proposes this bias be addressed with adequate training, however this thesis proposes this method is subjective, expensive and time consuming; and computer recognition algorithms of video transects, which are currently being developed, could be a better solution. This sentiment is reflected by Roelfsema, et al., (2014), who states 'automation of this process (collection of field data by snorkelers and manual photo analysis) would reduce processing effort'.

1.2 Problem statement

Eelgrass is an ecologically significant marine plant and highly sensitive to many important water quality parameters, such as nitrogen and phosphorus levels, and turbidity. As such, the depth limit at which it grows is used to assess waterway health under EU regulations. Collection of eelgrass field data for this purpose is undertaken using diver transects in locations established in past years - for Roskilde Fjord these locations were established in 1979. The same locations are dived each year, and eelgrass percentage coverage and depth is recorded to determine any trends. This methodology is flawed because eelgrass could be growing in another location in deeper water which is not captured by the sampling. These data are used to report to the EU, however can also be used for training and validation samples with remote sensing

applications. As such, the methodology by which these data are collected could be improved to more accurately represent current eelgrass standing crop. This study proposes such an approach, by identifying areas of SAV with classification using various MLAs with OBIA and Sentinel-2 imagery, and overlaying a model with eelgrass stress parameters to provide a likelihood about whether classified SAV is eelgrass. Factors effecting classification accuracy are also assessed.

Aim: To evaluate the feasibility of remote sensing with Sentinel-2 imagery to determine eelgrass coverage and provide guidance for eelgrass field data collection.

1.3 Research questions

1. What is the influence of varying the scale parameter on classification accuracy?
2. Which machine learning algorithm performs best for classification of SAV?
3. Can Water Column Correction and Multi-Temporal Analysis to improve classification results?
4. How can classified SAV be used to guide researchers to find Eelgrass?

2 Study Area and Data

2.1 Study Area

This thesis focused on Roskilde Fjord, Zealand, Denmark (Figure 9). The area is a 40km long and narrow fjord, consisting of approximately 30 small islands and islets. It has a surface area of 123km², and a mean depth of 3m (Pedersen, et al., 2014) and tidal amplitude of 0.2m (Kuusemäe, et al., 2016).

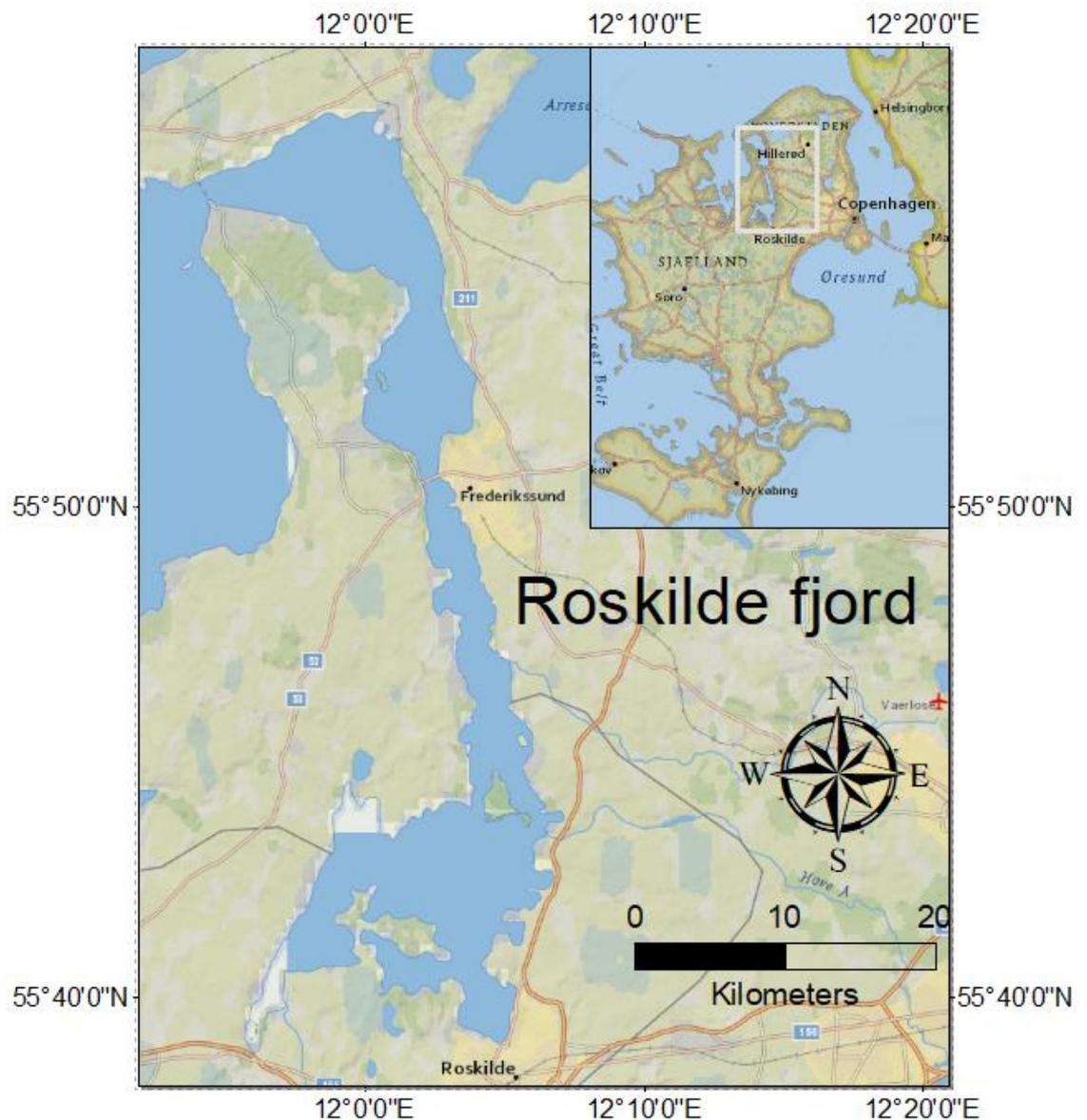


Figure 9: Location of Roskilde Fjord

Eelgrass growth is during spring and summer, where solar radiation is higher. The timeframe for change in coverage is long due to slow epiphytic growth. Eelgrass coverage is approximately 8%, which grows mainly along the shoreline (Kuusemäe, et al., 2016), and can grow up to 1m high. Strong currents are often present at the southern region of the mid-fjord bottle neck (Hansen, et al., 2015), which could result in eelgrass and other SAV lying flat and covering possible sand areas. The presence of currents can also influence eelgrass detection through blade orientation and minimising algae growth.

Roskilde Fjord contains several species of aquatic plants (Table 2) (Hansen, et al., 2015) and the substrate is sand and mud. This is useful to know due to the overlapping spectral signatures of various plants, and the substrate, which can confuse classification.

Table 2: Aquatic plants in Roskilde Fjord

Plant	Description	Image
Eelgrass (<i>Zostera marina</i>)	Most ecologically significant plant in fjord. Scale: <ul style="list-style-type: none"> • Patch: 1-10cm² • Bed: 10-100m² • Meadow: >100m² 	
Gutweed	Green algae that grows in shallow water	
Sea felt and brown algae	Form continuous thick carpets on fjord bed	
Sea lettuce	Large green alga, forms lettuce like sheets, thrives in still, shallow water	
Spiral tasselweed (<i>Ruppia cirrhosa</i>)	Long, thin stems that grow in very shallow water	

Several factors influence water turbidity (Table 3) (Hansen, et al., 2015), which are useful to understand to determine the optimal time of year for remote sensing.

Table 3: Factors that influence water turbidity in Roskilde Fjord

Parameter (source)	Influence	When
Nutrients: nitrate (agriculture) and phosphate (sewage, industry, agriculture)	Eutrophication that can result in algal blooms	After large rain events
Oxygen (increase from plants, atmosphere; decrease from ice coverage)	Adequate oxygen required for plant growth and decomposition	Anoxic periods are rare in Roskilde Fjord
Heat (warm weather)	Promotes algal growth	Warmer months: May – Aug
Salinity (exchange with ocean, evaporation)	Determines which plants and species thrive. Approx. 1.8% in north, 0.8% in south ¹ .	During windy weather

2.2 Imagery

2.2.1 Sentinel-2

Sentinel-2 image tiles were identified using Remotepixel Viewer (<https://viewer.remotepixel.ca>), which enables users to view map tiles from several sensors², with options to change band combinations and histogram cut. Imagery from spring to late summer 2016 was viewed to assess cloud coverage and water visibility. This period was selected to coincide with optimal eelgrass

¹ These salinity ranges are considered low concentrations. Eelgrass populations from low salinity environments have been demonstrated to have remarkable plasticity to thrive in low - high salinity conditions (Salo & Bostrom, 2014), which indicates salinity is not a limiting factor in the fjord.

² Landsat-8, Sentinel-2 and CBERS-4

growth periods (Krause-Jensen, et al., 2004), minimal cloud coverage, and good water visibility³. Ensuring these parameters were optimal gave the best chance of observing eelgrass.

Selected tiles were then downloaded using the python Sentinelhub.aws command, that uses a Sentinel hub API to accesses imagery from the Amazon Web Services (AWS) library. This methodology was preferred to download directly from the Remotepixel website, because it enabled a fast and seamless method to download a complete tile to a specified location (refer Appendix 1 for screen captures of commands).

Tiles for 12May16, 24Jul16, and 12Sep16 were downloaded, pre-processed and clipped to the study area (Figure 10). Three images were selected to examine which was best for classification, and for multi-temporal analysis later in the study.

³ Water visibility in Roskilde Fjord is optimal in May, and several months preceding.

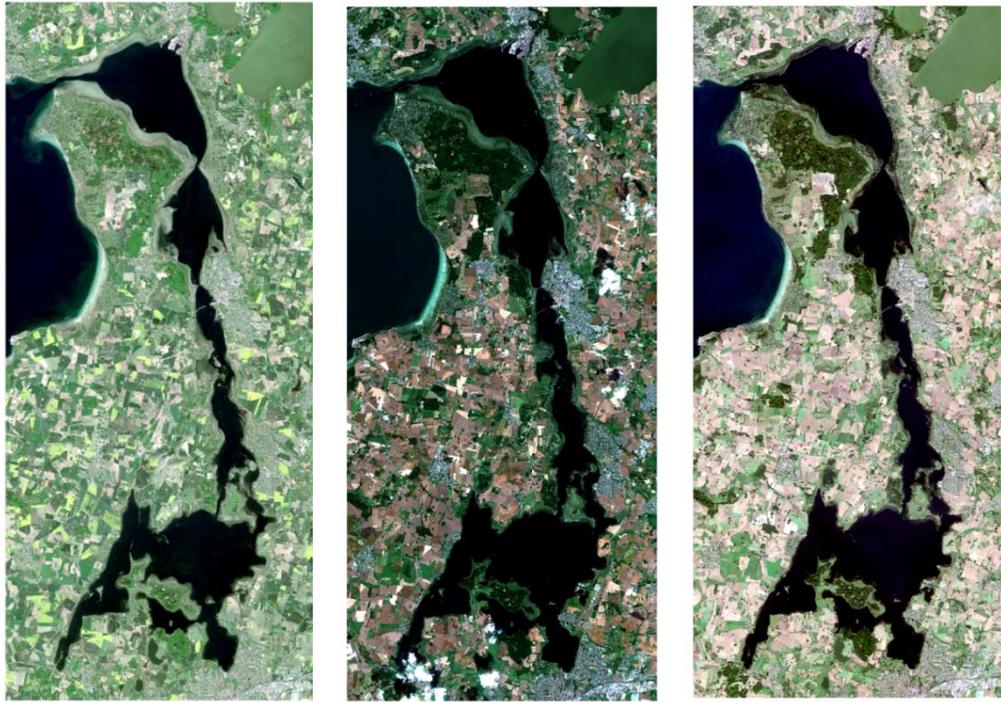


Figure 10: Sentinel-2 imagery clipped to AOI for (L to R) 12May16, 24Jul16, and 12Sept16

It was determined that the 12May16 image would be select to undertake the primary analysis for the study, primarily due to better water clarity. This selection process was a trade-off because greater eelgrass coverage is visible later in the year, however superior water quality was deemed more important to maximise classification accuracy results.

The imagery was inspected as RGB composites to observe water clarity and whether SAV was visible. SAV and suspected eelgrass was visible along the banks of the fjord in all 3 sets of images, with cover increasing later into the year as expected (refer Figure 45). Moving inwards from the banks, the water rapidly became darker, which could have been due to turbidity or depth. As the mean fjord depth is 3m and secchi depth is about 5m (Hansen, et al., 2015), it was expected these dark areas would be visible in the imagery. Therefore, the images were stretched in the darker water areas, which revealed previously unseen features (Figure 11)

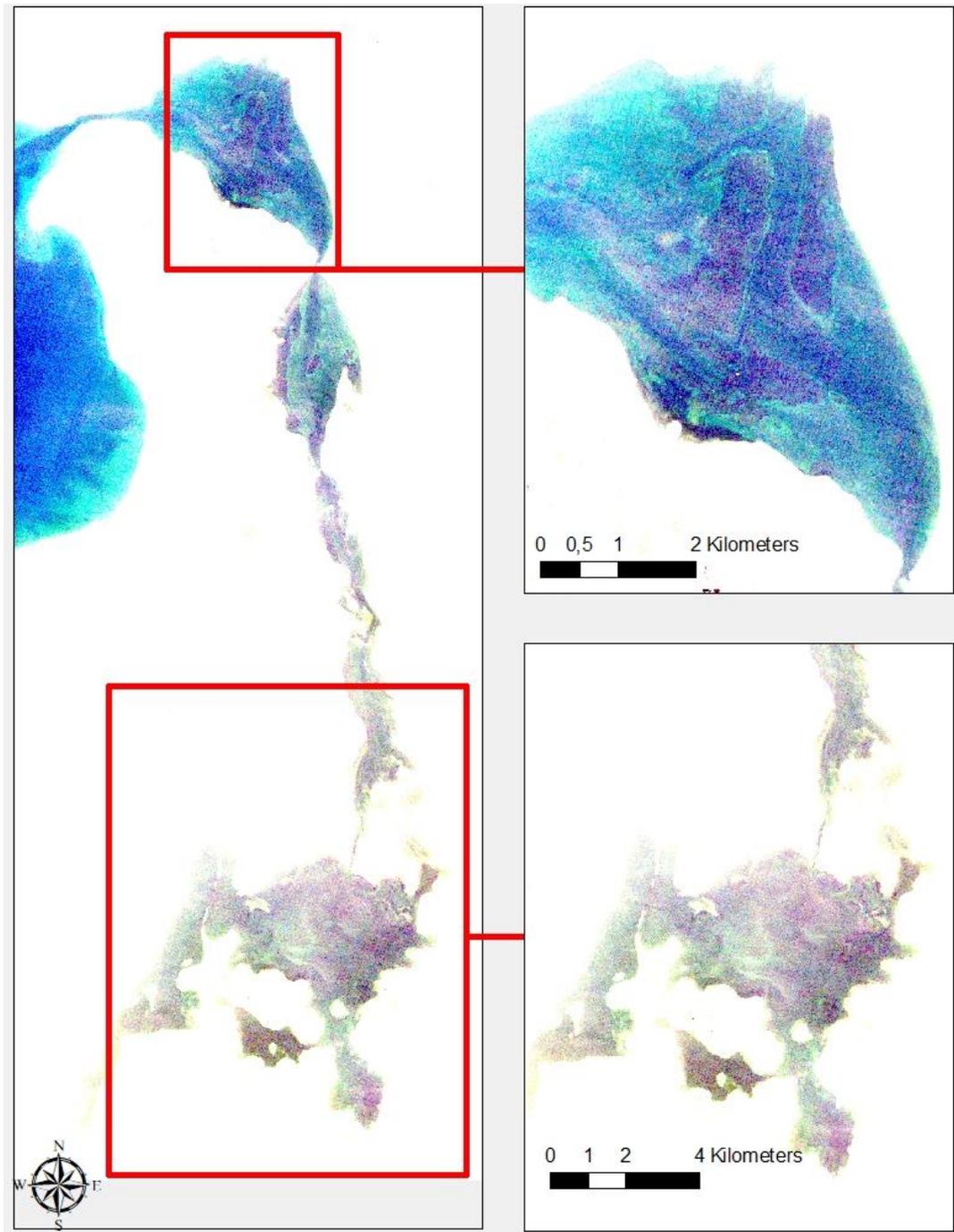


Figure 11: 12May16 RGB image with percentage clip stretch applied. Clumps of SAV and sand are visible

Large meadows of suspected SAV were visible in the deeper sections of the fjord. The depth raster (incorporated later into the study), indicated these

meadows were in water less than 4m deep and were outlined by bands of sand, which increased the likelihood the patches were meadows of SAV.

2.2.2 Orthophoto

A high resolution orthophoto from spring 2016 was downloaded from Kortforsyning (the Danish Geodata Agency), primarily to create training and validation samples, and to assist with interpretation of features in the Sentinel-2 imagery (Figure 12). The higher resolution (2m) also enabled better feature detection, to distinguish and verify image objects.



Figure 12: High resolution orthophoto of the southern fjord (L) and zoomed in on visible SAV (R)

2.3 Training data

Initially, it was intended to use diver transect data of SAV observations, including eelgrass, from 2010 to 2016 which was collected for the Danish National Monitoring Program, and provided by Aarhus University. However, these data proved inadequate for this classification for several reasons.

Firstly, the data did not correlate with the Sentinel-2 and orthophoto imagery, with low observations recorded in areas that appeared have high cover of eelgrass, and vice versa. Secondly, the data was collected on a much finer spatial scale than the Sentinel-2 imagery. This resulted in a small pixel

coverage of the diver data compared with the total area of the fjord, which was insufficient to train the classifier and use as validation samples. Also, adjacent high and low observations were observed on a fine spatial scale (Figure 13), which was undesirable as it could confuse the classification, particularly given the scale disparity.

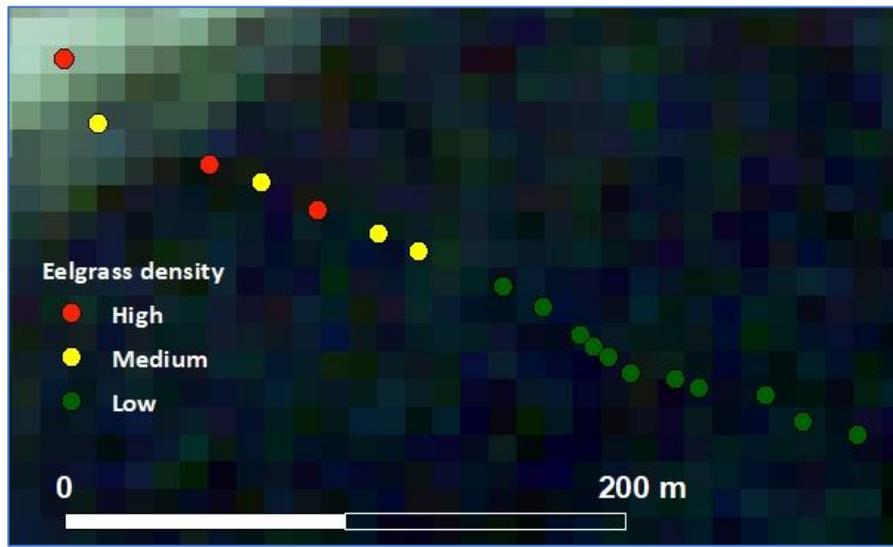


Figure 13: Diver transect data from Danish National Monitoring Program

Therefore, training and validation sample data were instead created with manual observations of Sentinel-2 and high resolution orthophoto imagery (Figure 14).

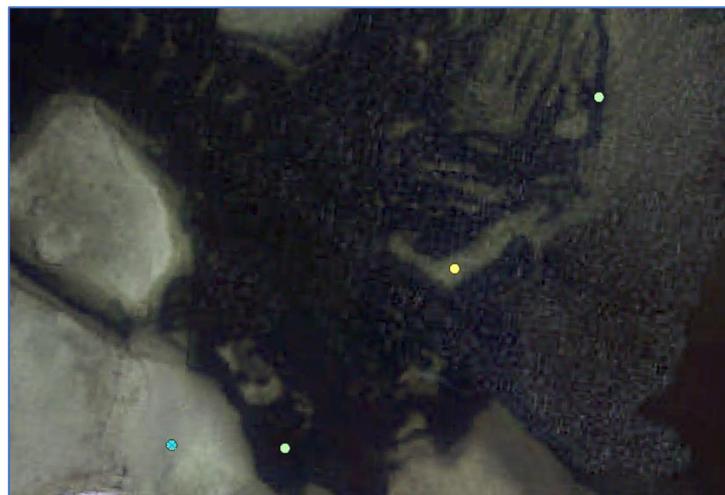


Figure 14: Creating training data from orthophoto imagery

These observations were conducted on pixel by pixel basis over the water area, and labelled either 'Sand' or 'SAV'. Using this method to create manual training data is not ideal, as it introduces bias through operator judgement, on the other hand however, it provided an opportunity to demonstrate the feasibility of this methodology, as field data will not always be available.

A presence/absence classification was deemed more appropriate than percent SAV coverage due to the lack of an accurate measurement method, and the image resolution was assessed as too low to produce a reasonable accuracy. Data were collected as single pixels, however since data points have no geometry, a fishnet raster of 10m x 10m was created to coincide with the pixels from the imagery. An observation was made on the majority coverage of the pixel containing the data point, to ensure the data point represented the pixel.

Training data was then created by manually selecting 125 samples each of SAV and sand, which is the minimum number of training samples per class demonstrated to produce optimal classification accuracy for all the MLAs used in this study (Qian, et al., 2015).

A spectrally diverse array of samples was collected to avoid overfitting the model. These observations were made on a pixel by pixel basis, however this was not essential as OBIA is usually tolerable to training data spanning multiple pixels, so long as it is within an image object. The training samples were updated during the classification to improve the classification if it was observed performing poorly.

Validation samples totalling 60 per class were created using *Equalized stratified random* points, which randomly distributed an equal number of

points within the SAV and sand feature classes, and ensured each class was represented by an unbiased, spatially and spectrally diverse selection of training pixels. Values were then assigned to the points based on the majority content of the pixel containing the point (Figure 15).



Figure 15: Creating validation samples with Sentinel-2 (L) and Orthophoto (R)

Sixty validation samples constituted approximately 50% of the number of training samples for each class (125), which is sufficient for accuracy assessment (Qian, et al., 2015).

2.4 Software

The classification was undertaken with Trimble eCognition® Developer, which provides a comprehensive array of image analysis algorithms to undertake OBIA for a variety of remote sensing applications. Users can develop algorithms or rule sets for the automatic analysis and extraction of remote sensing data, primarily by tinkering with SP settings to influence image objects, and during the classification process, which are selected based on either image-driven or knowledge-driven process.

ArcGIS was used for functions such as image and data visualisation, creating data samples, and map algebra.

3 Methods

3.1 Preprocessing

The image preprocessing workflow is outlined below (Figure 16).

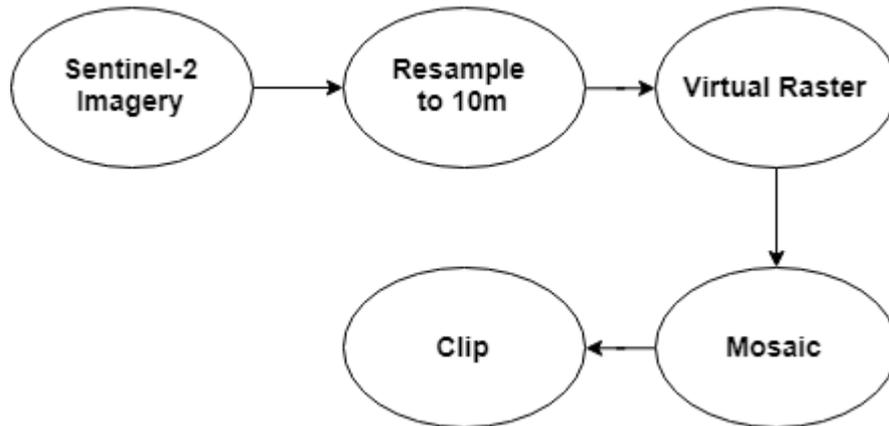


Figure 16: Preprocessing workflow

The first step of preprocessing was to resample the Sentinel-2 image bands to 10m, due to the differing spatial resolutions (refer Figure 2). This was undertaken using *gdalwarp* within a batchfile (.bat) (Figure 48). The .bat file is saved within the folder containing the image bands, and the function is called using a python command. Via this command, each band was resampled to 10m spatial resolution using Nearest Neighbour.

A virtual raster (.vrt) was then created, which is essentially an XML file that enables a lightweight way to manage large datasets with multiple files. This was undertaken using the *gdalbuildvrt* command. The '-separate' command was included to place each file into a separate band, to ensure the resulting image had multiple bands and could be accessed accordingly (Figure 49). The .vrt files were then mosaiced with *gdal_merge.py* to combine the imagery into a single .vrt (Figure 50).

The .vrt was then clipped with a shapefile created for the AOI of Roskilde Fjord, with the *gdalwarp* command (Figure 51). The 'crop_to_cutline'

command was included, which crops the extent of the dataset to the extent of the cutline, which ensures no areas of 'no data' remain.

The AOI shapefile was created in QGIS using the plugin 'rectangles ovals digitizing' to create a perfect rectangular bounding box. A perfect rectangle boundary is preferable because the *gdalwarp -cutline* function creates a bounding box using the upper and lower limits of the clip object, and any overlap with the shapefile results in an area of 'No Data', which is problematic for further analysis. Using a perfectly symmetrical object ensures no overlap, and consequently no 'No Data' areas.

Screen captures of these commands are provided at Appendix 1.

3.2 *Depth raster*

The use of a depth enabled areas of dense SAV or turbid water to be distinguished from deep water. This was useful to visualise the imagery to get an indication of the size and shape of SAV in the dark water areas. It also enabled depth thresholds to be used during the classification based on knowledge of the depth of eelgrass growth and anticipated light penetration in the fjord.

The depth raster was created using the Danish Maritime Safety Administration 50m resolution data from DHI. Land areas had values of infinity, which made it unsuitable to use in eCognition because an integer is required for it to be valid data during classification.

To remove the infinity values, the *raster calculator* in ArcGIS was used to convert these values to 1, which was land for the classification, using the following expression:

```
("DK_merge.tif" < -100) * (1)+("DK_merge.tif" >= -100)*("DK_merge.tif")
```

With 'DK_merge.tif' the identity of the depth raster. Following this process, it was then resampled to 10m, to provide a smoother appearance, and clipped to the study area (Figure 17).

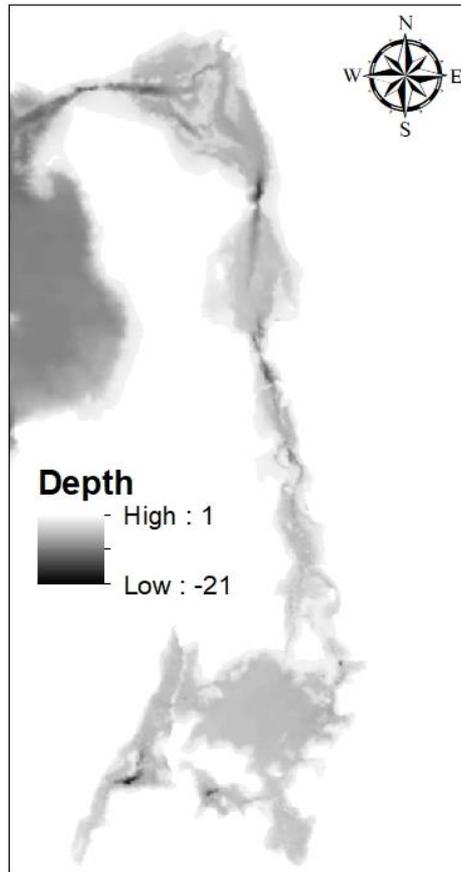


Figure 17: Depth raster clipped to study area

3.3 Water Column Correction

Water column correction (WCC) was considered for this classification to remove the effect of the water column, which varies spatially within the fjord, and temporally during the year. Removing variability from the water column was desirable to achieve consistent classification results and facilitate multi temporal analysis.

The methodology used was based on the method developed by Lyzenga (Lyzenga, 1978). The first step was to explore the band combinations, to determine which resulted in the smoothest depth gradient. Each band

combination was calculated and visualized in ArcGIS. The *raster calculator* was used to perform calculations on the bands, and they were then visualized to determine the most suitable combination for calculating the relative bathymetry. The equation used was (Lyzenga, 1978):

$$D = \frac{\ln(b2)}{\ln(b1)}$$

Where b2 is the band with the longer wavelength, and b1 is the preceding band with the shorter wavelength. The Ln of each band is used because light attenuates in a water column as a logarithmic function of depth. By applying this ratio between two bands, the method becomes closer to becoming bottom type independent. Using this methodology, it was determined that band combination 3/2 (green/blue) was the most suitable to normalise the image bands.

The next step was to create several data points to capture a single bottom type across as wide a range of depths as possible (Figure 18).



Figure 18: Placing points to extract depths for WCC

Values for the relative bathymetry, as well as each band within the raster, were then extracted to the points using the *Extract multi-values to points* tool in ArcGIS. This data was then exported to Excel, and plotted in a scatterplot with the relative bathymetry on the X axis, and the various band values on the Y axis. An ordinary least squares regression was performed for each band combination, with the A value from the resulting equation $A * X + B$ noted for use later in the process.

Next a 'delta relative depth' value was calculated, by using the highest extracted relative depth and subtracting the relative bathymetry pixel value from it. This value served as a normalisation depth for the remainder of the calculation.

The final step was to normalize each of the bands using the A value and 'delta relative depth' with the following formula:

$$'delta\ relative\ depth' * A + pixel\ value$$

This function was performed in ArcGIS using the *raster calculator*, with each band used as 'pixel value'. The resulting normalised bands were then mosaiced using the *composite bands* tool.

3.4 Object Based Image Analysis

OBIA was undertaken using eCognition software, and followed the workflow depicted at Figure 19. The methodology thus far was preparing the imagery for use in this workflow.

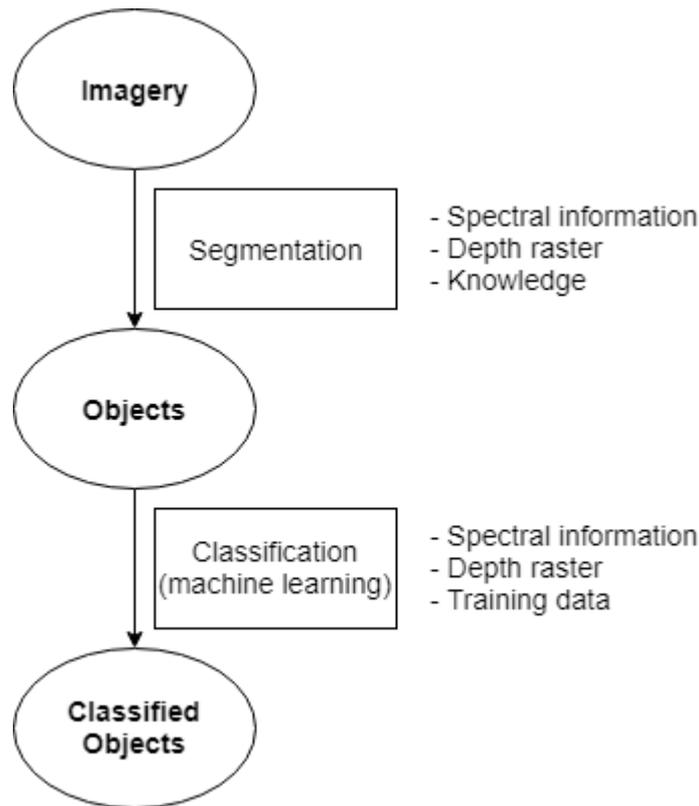


Figure 19: The OBIA workflow

This process consists of a series of segmentation and classification processes, that result in classified image objects. The eCognition process tree developed for this workflow is presented at Figure 20.

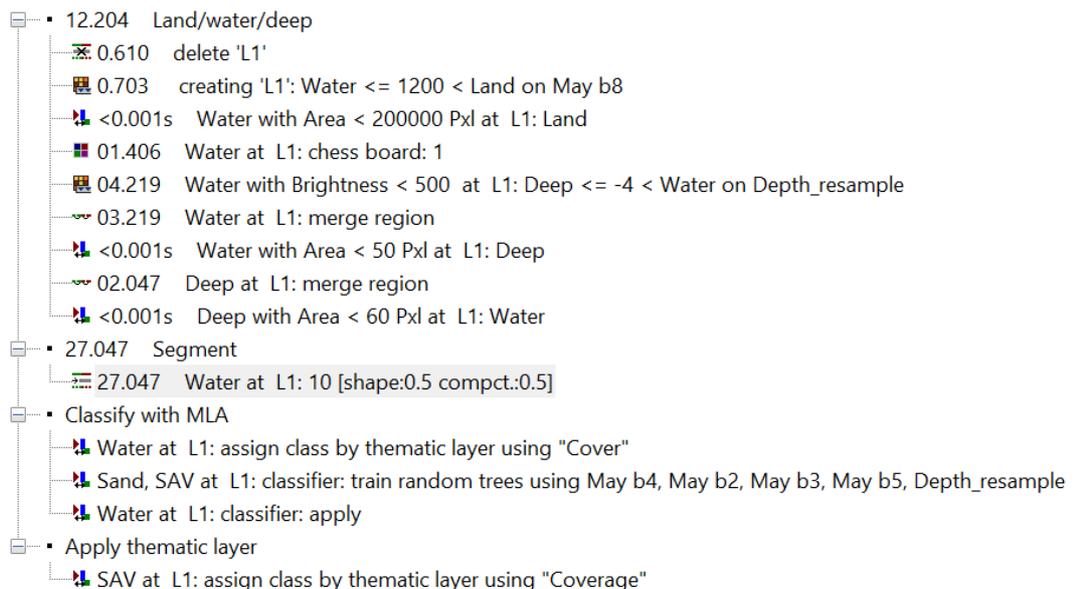


Figure 20: eCognition process tree

The first step was to separately load the 12May16 image bands into eCognition, which enabled individual bands to be accessed during the analysis. The depth raster was also loaded as a layer (Figure 21). The training data was added as a thematic layer to train the various MLAs.

Image Layer	R	G	B
May b1			
May b2			○
May b3		○	
May b4	○		
May b5			
May b6			
May b7			
May b8			
Depth_resample			

Figure 21: Image bands and depth raster loaded into eCognition

3.4.1 Separate land/water/deep

The first stage of the process was to separate land, water and deep areas to focus the classification accordingly. This was undertaken using the Process Tree at Figure 22.

12.204 Land/water/deep		
1	0.610	delete 'L1'
2	0.703	creating 'L1': Water <= 1200 < Land on May b8
3	<0.001s	Water with Area < 200000 Pxl at L1: Land
4	01.406	Water at L1: chess board: 1
5	04.219	Water with Brightness < 500 at L1: Deep <= -4 < Water on Depth_resample
6	03.219	Water at L1: merge region
7	<0.001s	Water with Area < 50 Pxl at L1: Deep
8	02.047	Deep at L1: merge region
9	<0.001s	Deep with Area < 60 Pxl at L1: Water

Figure 22: Process tree for separating land/water/deep

The *delete level* algorithm was inserted to reset the process tree after trialling various settings and processes (step 1 Figure 22).

The *multi threshold segmentation* algorithm enabled water and land to be separated based on a threshold NIR value (step 2 Figure 22). This value was determined by displaying only band 8 (NIR), and hovering the mouse over land and water areas to determine a NIR threshold value (Figure 23). NDWI, a popular band ratio algorithm for detecting shallow water bodies, was not used in this study to separate land and water because the fjord is generally too deep for it to work effectively, and NIR was sufficient for this purpose.

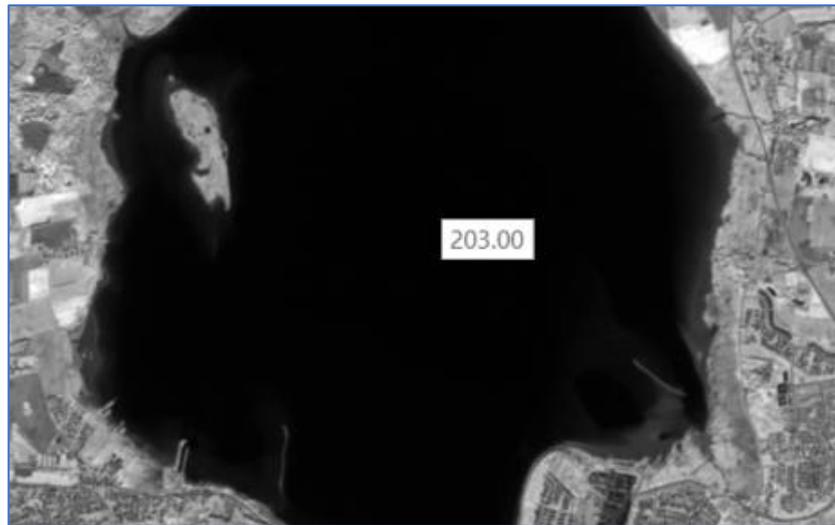


Figure 23: Determining the NIR threshold to separate land and water

Large NIR differences were observed between land and water, with the majority of values <1000 over water, and several thousand over land. The value of 1200 was selected (step 2, Figure 22), as this value generally separated land and the main water body well, with a few exceptions.

Firstly, areas with shadows from clouds or trees have low NIR values, so these areas were often classified as water, and resulted in small patches of incorrectly classified water (Figure 24). Small patches of water were also produced by dams and lakes, which were not wanted in the classification.



Figure 24: Misclassified water – clouds (L) and washed up reeds (R)

These small water bodies were removed by changing their classification to land by specifying a minimum threshold size with the *assign classification* algorithm (step 3, Figure 22). As Roskilde Fjord was the only water body of interest, and significantly larger than small dams, lakes and shadows in the scene; this threshold value was not conservative.

Secondly, Vegetation along the fjord banks, such as washed up reeds or algae, have a high NIR value and therefore are classified as land (Figure 24). These instances were fairly rare and inconsequential for the classification, and were therefore ignored. The map with land and water separated is shown at Figure 25.



Figure 25: Map with land and water separated

Deep and shallow areas were separated to account for dark objects in the water which were not patches of SAV, and account for areas too deep for eelgrass to grow. This was done with the depth raster (Figure 17), which was loaded into the *assign class* algorithm and enabled 'Deep' and 'Water' areas to be thresholded (step 5, Figure 22).

Eelgrass depth limit in Roskilde Fjord is reported as 2-3m and the densest beds are in waters 1-2m deep (Hansen, et al., 2015), however as light is the main limiting factor to eelgrass growth (Kuusemäe, et al., 2016), and secchi depth in 2016 was 5.5m (Hansen, et al., 2015), it was anticipated eelgrass may be detected at 4m.

Depth was selected based on examining the image with stretch applied (Figure 11) to determine at what depth SAV and bathymetric features were visible. Depths were reduced at 0.5m intervals to determine the depth at which

features were no longer visible. Using this methodology, 4m depth was selected (Figure 26).



Figure 26: Depth raster applied with 4m threshold

Small patches of deep water, and small patches of water within the deep feature class were removed using *merge* and *assign class* algorithms for each (steps 6 – 9, Figure 22).

3.4.2 Segmentation

Segmentation was undertaken on the water feature class to create image objects that distinguished SAV from sand. As illustrated at Figure 19, input sources for the segmentation were spectral information, the depth raster, and expert knowledge about eelgrass growth patterns, which were input into the segmentation using the *multispectral segmentation* algorithm (Figure 27)

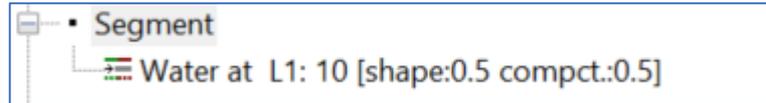


Figure 27: Multispectral segmentation algorithm was used to segment the water feature class

Spectral information was incorporated into the *multispectral segmentation* algorithm by assigning a weight to each to each band (Figure 29). As clusters of SAV were the objects of interest, band 3 (green) was assigned a weight of 2, due to SAV having a greater absorbance at the green wavelength (Hossain, et al., 2015), which was also observed in the imagery (Figure 28).

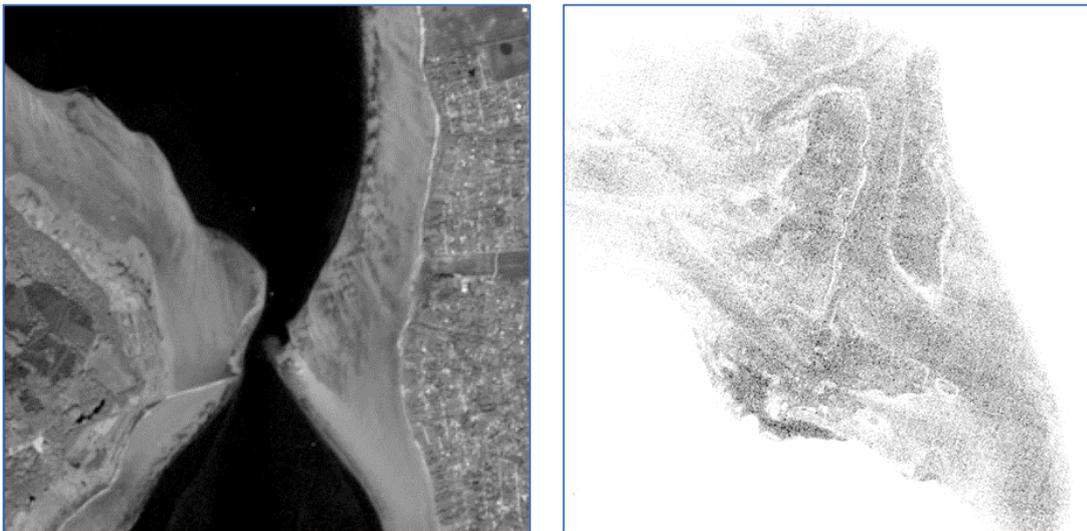


Figure 28: Band 3 (green) showing SAV in shallow and deep areas
Other bands were given a weighting of 1 (Figure 29).

Parameter	Value
▲ Image Layer weights	1, 1, 1, 2, 1, 1, 1, 1, 1
Depth_resample	1
May b1	1
May b2	1
May b3	2
May b4	1
May b5	1
May b6	1
May b7	1
May b8	1
▷ Thematic Layer usage	No, No, No
Scale parameter	10
▲ Composition of homogeneity criterion	
Shape	0.5
Compactness	0.5

Figure 29: Multiresolution segmentation parameter settings

The scale parameter was adjusted to create image objects based on larger or smaller amounts of pixel homogeneity, which was examined in detail.

Knowledge and observations of eelgrass growth patterns were input into the segmentation with the shape and compactness settings within the *multiresolution segmentation* algorithm (Figure 27). The weighting of the shape parameter determines the influence of shape compared with colour. Shape homogeneity is based on the deviation of a compact (or smooth) shape, and colour homogeneity is based on the standard deviation of the spectral colours (Trimble Germany GmbH, 2016). Compactness weighting determines the influence of compactness compared with smoothness. Compactness is a measure of how compressed an image object is, and smoothness is a measure of the evenness of the object boundary.

Eelgrass is characterised by bands of growth which have a low compactness, as opposed to other SAV in the fjord such as sea lettuce or gutweed which are more compact (Table 2), which is also observed in the imagery (Figure 30).

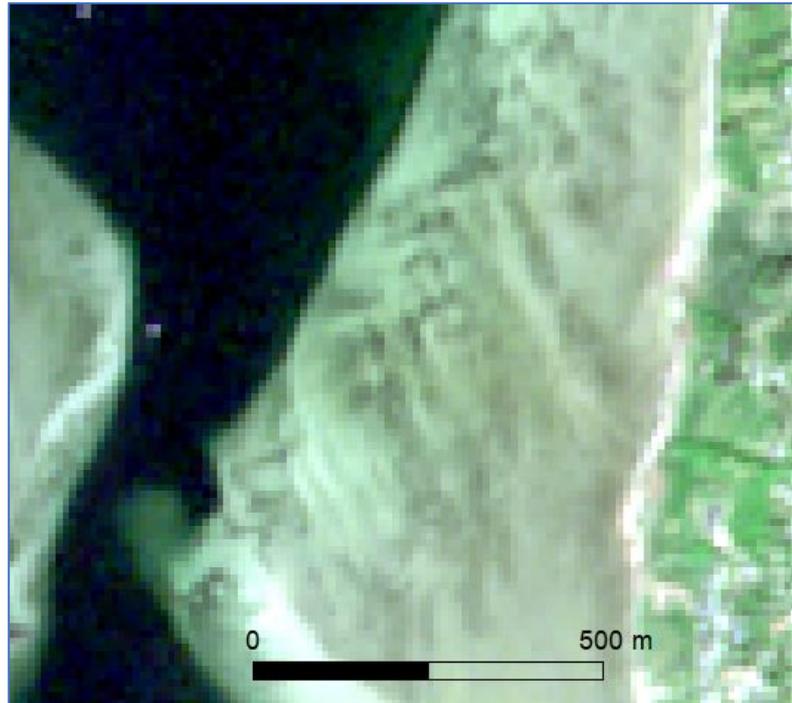


Figure 30: SAV growing along the banks of the fjord. Suspected eelgrass due to growth patterns

Determining size, shape and spectral parameters can be done based on knowledge, however the process is essentially trial and error to find the settings that segment the image appropriately. The rule of thumb is to ensure that the feature of interest does not contain other features in image objects. The combination used for this thesis was shape 0.5 and compactness 0.5, which gave even weight to all settings to attempt to capture the features along the banks and deep areas evenly.

The segmentation results are presented at Figure 31. It can be observed that image objects cover areas of SAV, and have captured features in shallow and deep areas reasonably well.

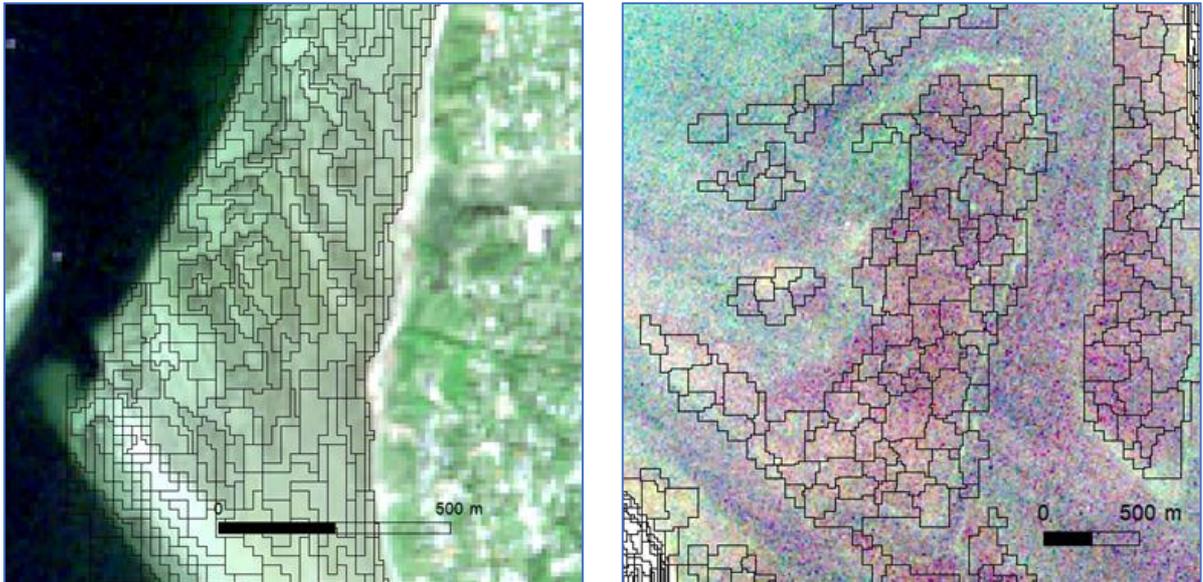


Figure 31: Segmentation of shallow (L) and deep (R) areas

3.4.3 Classification

Classification was the second stage of the OBIA workflow (Figure 19), where image objects are classified based on their attributes, such as spectral characteristics, shape, colour and context (Trimble Germany GmbH, 2016). The MLAs used were Bayes (parametric), RF and SVM (both nonparametric). It was anticipated SVM would produce the highest accuracy based on findings from previous studies (e.g. (Noi & Kappas, 2017)). The steps to apply these algorithms in the eCognition OBIA workflow is presented at Figure 32.



Figure 32: Process Tree to undertake classification with MLAs

Training data was loaded using the *assign class by thematic layer* algorithm, and applied to the water feature class, as illustrated at Step 1. This classified the 'Water' feature class with the sample value 'SAV' or 'Sand' (Figure 33).

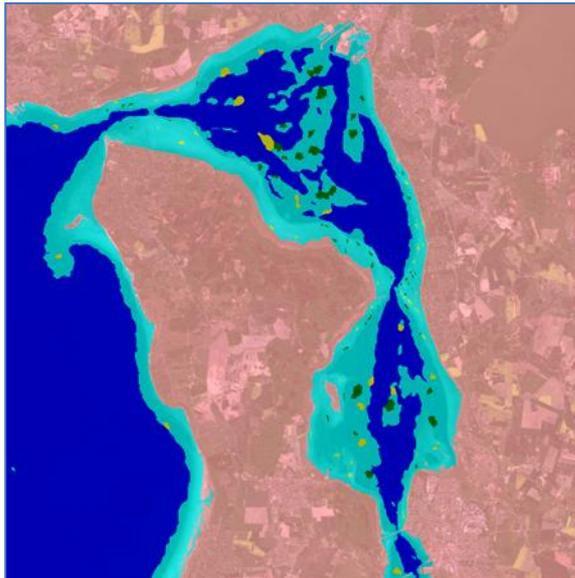


Figure 33: Water feature class classified with 'SAV' and 'Sand' samples visible as green and yellow objects

Step 2 was to train the various MLAs with the newly created 'SAV' and 'Sand' feature classes. The example at Figure 32 used the RF classifier, and the features B4 (red), B3 (green), B2 (blue), B5 (red-edge 1), and Depth raster as the input values. During this step, classifier tuning parameter are set (Figure 34).

Classifier	
Type	Random Trees
Depth	20
Min sample count	0
Use surrogates	No
Max categories	16
Active variables	0
Max tree number	500
Forest accuracy	0.01
Termination criteria type	Both

Figure 34: Classifier settings are input during MLA training

Tuning parameter setting is significant to optimise classification accuracy, and can have a more significant impact on overall accuracy than the number of training samples (Qian, et al., 2015). Tuning parameter settings must

be optimised based on the training parameter sample size (Qian, et al., 2015) (Noi & Kappas, 2017).

The key parameter settings for the SVM RGB Kernel are C and gamma. Qian, et al., (2015), demonstrated that the optimal settings varied with sample size, however for training sample sizes of 125 per class (which matched the training sample size for this study), the optimal values of C and gamma were 10,000 and 0.001, respectively. These results were backed up by Noi & Kappas, (2017), who concluded ‘in general, a high value of C and a low value of gamma produced the lowest error’. Various combinations of these values were tested for this thesis.

For DT parameter settings, the key parameter settings are depth and nTree (Noi & Kappas, 2017). Depth and nTree were set at 20 and 500 respectively, which were demonstrated to produce optimal results (Noi & Kappas, 2017).

Step 3 (Figure 32) applied the classification to the water feature class (Figure 35).



Figure 35: RF Classification applied

3.5 Accuracy Assessment

An accuracy assessment was undertaken in ArcGIS with the *Compute Confusion Matrix* tool, using the validation samples and classification results as input. Previous studies have demonstrated the validity of using of manually created sample data undertake accuracy assessment (e.g. (Roelfsema, et al., 2009)).

3.6 Eelgrass stress model

An eelgrass stress model was applied to the SAV feature class to indicate the likelihood of the classified SAV being eelgrass. This model was provided by (Kuusemäe, et al., 2016), following their study into modelling stressors on eelgrass recovery. The inputs to the model were:

- Concentrations of nutrients N and P,
- Turbidity resulting from macroalgae,
- Mechanical damage from stones attached to drifting macroalgae, and,
- Turnover of sediment by lugworms.

Values were provided in the model as g C m^2 (grams of Carbon per metre squared), which were converted to percentage coverage using a conversion factor of 1% coverage equals to 2 g C m^2 (Carstensen, et al., 2015). The resulting percentage values ranged from 0 to 47 %, with a large proportion of 0 values. These values were classified as Low (>10), Medium (10 to ≥ 30) and High (>30), to provide an optimal representation, as demonstrated at Figure 36, and visualised at Figure 37.

Input Table	EC_May_ros_C10_el305-84_elements_13686
Field Name	Coverage
Expression	classify(!Percent_co!)
Expression Type (optional)	PYTHON
Code Block (optional)	<pre>def classify(percent_co): if (percent_co <= 10): return 'Low' elif (percent_co > 10) and (percent_co <= 30): return 'Medium' elif (percent_co > 30): return 'High'</pre>

Figure 36: Percent values classified as Low, Medium, and High

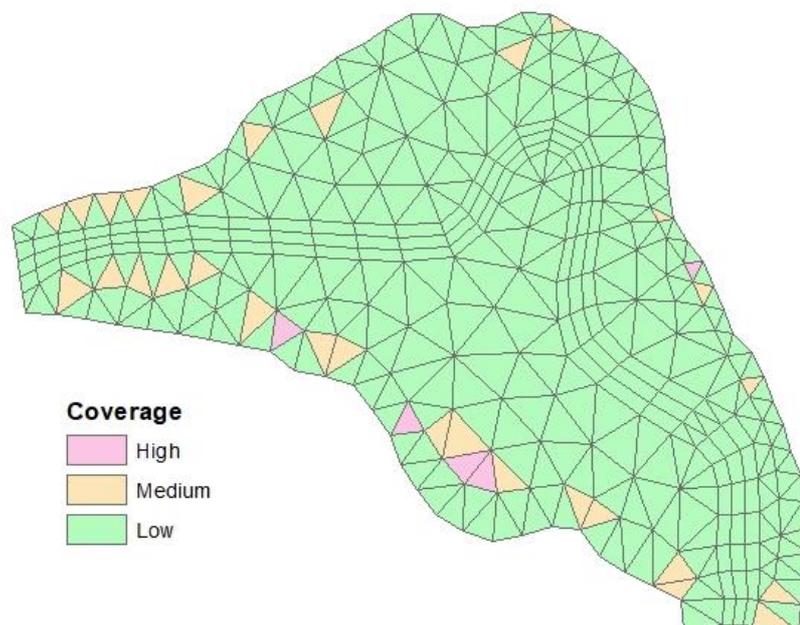


Figure 37: Eelgrass stress model for north fjord, with biomass interpreted into percent coverage

The output of this model is eelgrass biomass, however this has been interpreted as eelgrass probability, as areas of higher biomass are more likely to contain eelgrass than lower biomass areas.

The above model was loaded into eCognition as a thematic layer labelled 'Coverage', and overlaid onto the SAV classification using the *assign class* algorithm (Figure 38).

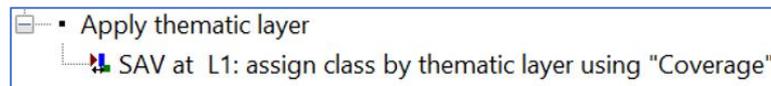


Figure 38: Overlaying the eelgrass stress model to the SAV classification

The result is presented at Figure 39. Areas that were classified SAV have been reclassified as a high, medium or low probability of being eelgrass based on the eelgrass stress parameter model.

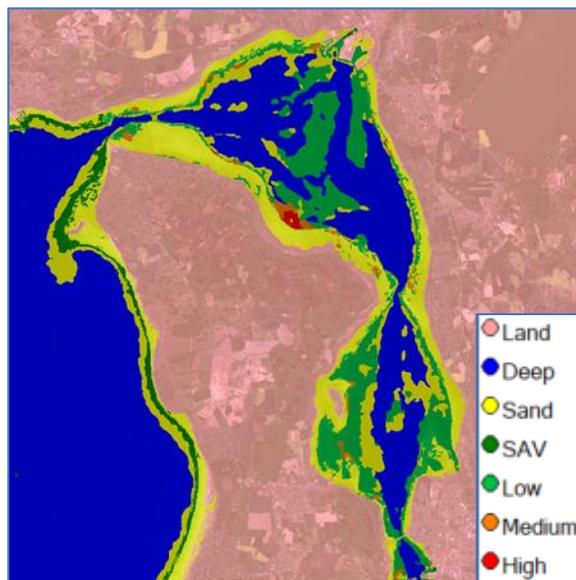


Figure 39: Eelgrass stress model overlaid on SAV feature class.

4 Results and Discussion

The aim of this project was to evaluate the feasibility of remote sensing with Sentinel-2 imagery to determine eelgrass coverage and provide guidance for eelgrass field data collection. To achieve this aim, four research questions were proposed:

1. What is the influence of varying the scale parameter on classification accuracy?
2. Which machine learning algorithm performs best for classification of SAV?
3. How can one improve the accuracy of eelgrass detection for future studies?
4. How can classified SAV be used to guide researchers to find Eelgrass?

Presented below are the results and discussion in the context of the research questions, and the strengths and weakness of the project.

4.1 *What is the influence of varying the scale parameter on classification accuracy?*

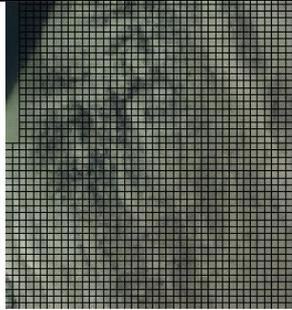
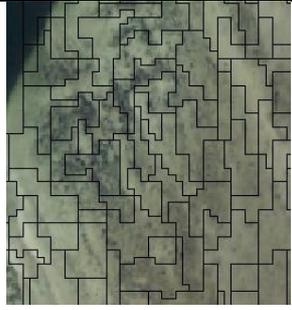
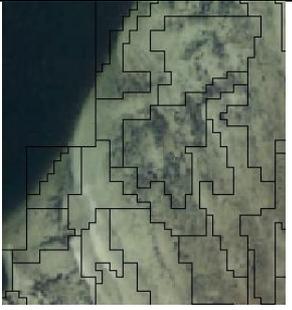
The scale parameter (SP) is one of several settings, including shape and compactness, within the *multiresolution segmentation* algorithm in eCognition; which have been demonstrated in several studies to impact OA (Smith, 2010) (Drăguț, et al., 2010). The SP effectively determines the size of image objects by setting a threshold value for the degree of spectral heterogeneity permitted within each image object. Larger SP values produce larger image objects, and vice versa for small values (Trimble Germany GmbH, 2016). The size and shape of image objects underpin classification accuracy, because training data and classification algorithms are applied to these objects. Varying the SP and observing the influence on classification, was selected as a research question for this study to demonstrate the effect on OA with classifying SAV with

moderate resolution imagery, and provide recommendations for the optimal SP setting.

RF was selected to test the various SP settings from amongst the 3 MLA's used in this study. RF was demonstrated by (Pal, 2005) to perform equally well to SVM in terms of OA, and required fewer and easier to define parameters to obtain similar classification accuracy results. RF was selected over Bayes due to the anticipated superior performance of a nonparametric algorithm versus parametric.

Varying the SP settings influenced OA, as demonstrated at Table 4. Image objects generally grew larger with increasing the scale parameter, as spectral heterogeneity increased.

Table 4: Overall accuracy for the classifications undertaken with RF at various SP settings. Note: shape and compactness were consistent at 0.5 each

SP	1	10	20
OA (%)	68.1%	72.2%	66.4%
			
SP	30	40	50
OA (%)	66.4%	68.1%	67.2%
			

The highest OA was achieved with a scale parameter of 10, with OA of 72.2%. The second highest OA of 68.1% was obtained with both SP 1 and 40, followed by SP 50 which obtained OA of 67.2%. The lowest OA of 66.4% was obtained with both SP20 and 30.

SP1 used a *chessboard* segmentation, and represents a pixel based classification. It demonstrates the improvement in OA possible with OBIA. SP10 effectively captured areas of SAV and excluded sand. As SP increased, more sand is included in the image objects, as the minimum spectral homogeneity

increases. SP30 and 40 appear to have identical image objects, however SP40 has a slightly higher OA, and second highest OA overall. This indicates the increased spectral homogeneity is beneficial in some areas of the fjord, and beneficial overall for the classification, compared with SP30.

Image objects for SP10 in the deep and along the shore are illustrated at Figure 40. It can be observed that the smallest image objects, 1 to 2 pixels in size, are clustered around the shoreline, and increase in size moving away from the shore. Larger image objects, up to 2000m², are in the deeper areas. These large segments could be a source of bias during classification, as the entire object is classified by a single training point, which is only a single pixel. This demonstrates how classification accuracy can be impacted through segmentation.



Figure 40: Segmentation for SP10 showing variation in image object size between deep and shallow

It is not possible to attribute the SP directly with image object size, as demonstrated by the degree of image object size variation in shallow and deep areas (Figure 40), and the minimal influence on image object size by varying SP from 20 to 40 (Table 4); therefore it is not possible to make conclusions about the influence of image object size on classification accuracy.

Varying the SP modifies spectral heterogeneity within an image object, which influences classification accuracy based on how well this setting reflects the features in the scene. The challenge for the user is to determine this setting, in conjunction with shape and compactness.

It is recommended segmentation be undertaken on multiple levels when features vary in size, shape and/or spectrally. This is particularly applicable in marine environments, where features vary in the shallow and deep areas, and water quality can influence the spectral signature, which can vary over small spatial scales. This is demonstrated with feature size and shape varying in the shallow and deep regions of the fjord, and the southern region being more spectrally homogenous than the north (Figure 11 and Figure 12). It is expected that dividing the north and south regions, and undertaking a separate segmentation for shallow and deep areas would improve classification accuracy.

4.2 Which machine learning algorithm performs best for classification of SAV?

The classification outputs for the highest overall accuracy (OA) achieved with each MLA are presented at Figure 41. The highest OA was achieved with SVM with 73.1%, followed by RF and Bayes with 72.3% each. The OA for SVM and RF was obtained using identical SPs, which used a scale parameter of 10 to create image objects, whereas the OA for Bayes was obtained with a scale

parameter of 20. All other SP settings were identical (refer to Section 3.4.2 for the specific SP settings).

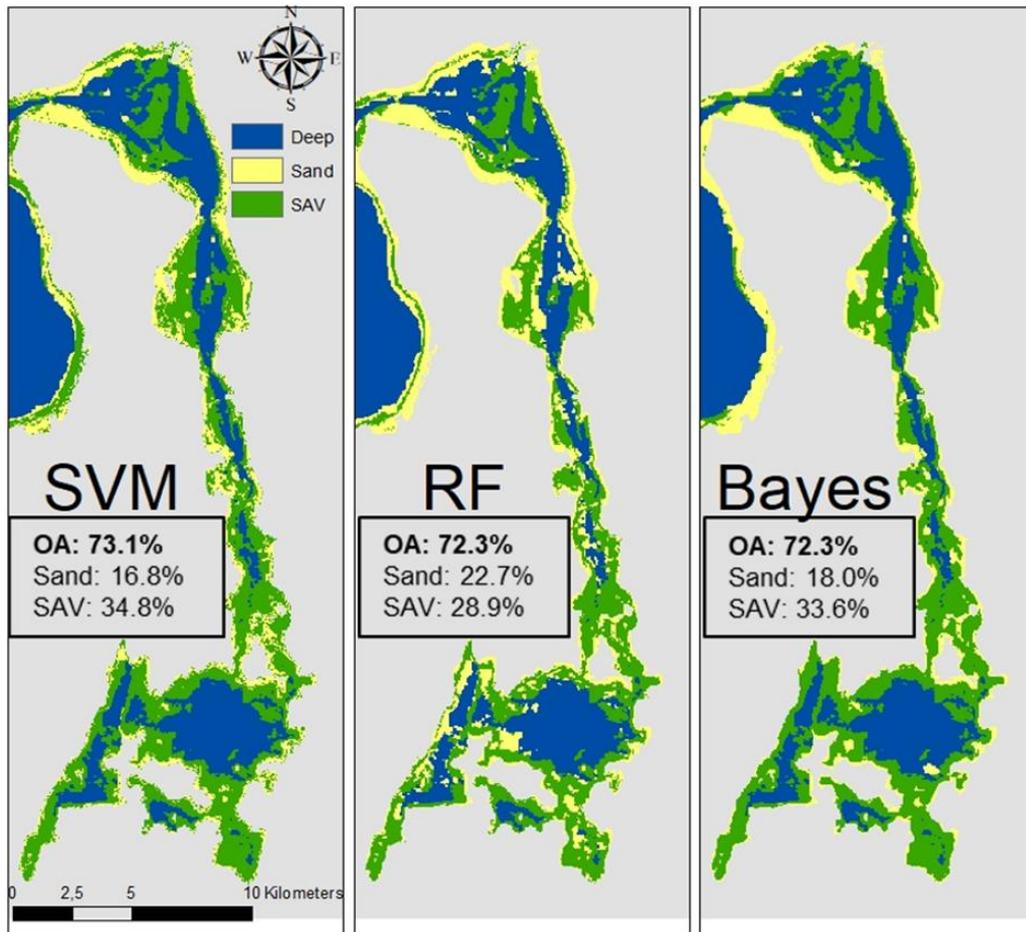


Figure 41: Comparing the overall classification for the 3 MLA used in the study - SVM, RF and Bayes

SVM and Bayes classified a similar overall SAV coverage of 34.8% and 33.6% respectively, whereas RF classified SAV lower at 28.9% (Figure 41). It can be observed that RF classified the most sand (22.7%) and least SAV, due to classifying sand in the southern and mid-northern regions of the fjord. All MLAs produced visually similar results for total SAV cover in the far northern region of the fjord.

The official percentage eelgrass coverage in Roskilde Fjord in 2016 was 9.4% (Kuusemäe, et al., 2016), which indicates the results obtained in this study captured additional SAV growing in Roskilde Fjord, which supports the hypothesis that that it would not be possible to distinguish eelgrass from other SAV.

The performance of each MLA is discussed below. Manual creation of verification and training data was an error source due to operator inconsistencies in selecting Sand and SAV samples. For analysing the performance of each MLA however, it is assumed the training and verification data is accurate.

SVM was examined in detail due to the expectation of higher OA, based on results from other studies (e.g. (Qian, et al., 2015)). It is also the most sensitive to tuning parameter settings ((Noi & Kappas, 2017) (Qian, et al., 2015)), therefore is most suitable to examine the effects of varying parameter settings.

The highest overall classification accuracy obtained with SVM used parameter settings of $C = 10,000$ and $\gamma = 0.001$, which is presented in detail at Figure 42 and Table 5.

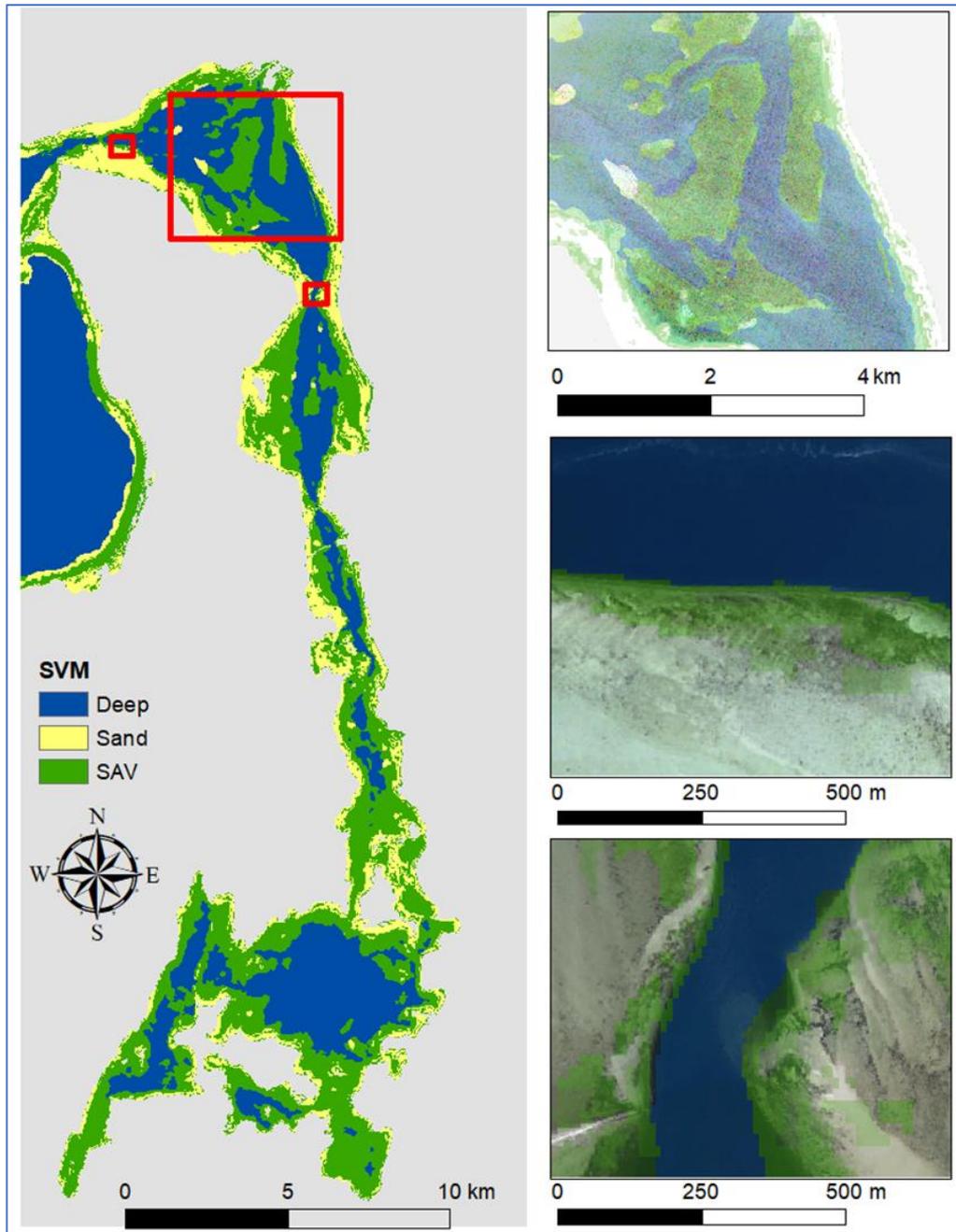


Figure 42: SVM classification map

Table 5: SVM confusion matrix $C=10,000$, $\gamma=0.001$

SVM: $C= 10,000$, $\gamma= 0.0001$					
ClassValue	Sand	SAV	Total	U_Accuracy	
Sand		29	9	38	0,763
SAV		23	58	81	0,716
Total		52	67	119	0
P_Accuracy	0,558	0,866	0	0,731	

The P_Accuracy (producer accuracy) indicates OA was biased by under classification of sand and over classification of SAV, due to the low sand producer accuracy of 55.8%. This indicates that 55.8% of sand validation samples were mapped as sand, and since there were only 2 categories used in the classification, these areas were instead classified SAV. This is also evident at Figure 42, with large areas of the southern region of the fjord classified as SAV. An explanation for this result is that SVM is overfitting the training data, and the training data is not representative of the entire fjord. This is demonstrated in the southern region of the fjord, which is relatively more spectrally homogenous, and therefore more susceptible to overfitting if the training data is not accurate. Another reason could be due to SP settings, as demonstrated at Section 4.1.

To examine the effect of adjusting tuning parameters, C and gamma were adjusted, which alters the margins and shape of the classification kernel respectively, to determine the effect on the classification. Results from the study by (Qian, et al., 2015) were used as a basis to set the parameters. This study demonstrated high OA for parameter settings $C=1,000,000$ and $\gamma=0.0001$, with sample sizes higher and lower than 125. For very low sample sizes (25 per class) parameter settings $C=10,000$ and $\gamma=0.001$ obtained the highest OA. Three different settings were tested and are presented below (Table 6 – Table 8).

Increasing C and reducing gamma had a negative effect on the classification in every accuracy measure (Table 6).

Table 6: SVM parameter settings $C=1,000,000$ and $\gamma=0.00001$

SVM: $C=1,000,000$, $\gamma=0.00001$				
ClassValue	Sand	SAV	Total	U_Accuracy
Sand	27	54	81	0,333
SAV	25	13	38	0,342
Total	52	67	119	0
P_Accuracy	0,519	0,194	0	0,336

Increasing C widens the margins and increases the fit of the model to the data, whereas reducing γ counteracts this by increasing the distance of the training data that influence the model (refer Figure 7 and Figure 8). The results presented at Table 6 indicate the model was neither under or overfit to the data, but simply a bad fit, because all accuracy results were low.

The next alteration of the SVM parameters was to lower C to 1000 and increase γ to 0.001, which positively influenced classification accuracy (Table 7).

Table 7: SVM parameter settings $C=1,000$ and $\gamma=0.001$

SVM: $C=1000$, $\gamma=0.001$				
ClassValue	Sand	SAV	Total	U_Accuracy
Sand	46	28	74	0,622
SAV	6	39	45	0,867
Total	52	67	119	0
P_Accuracy	0,885	0,582	0	0,714

OA was 71.4%, which was marginally lower than highest OA (73.1%), however this result was characterised differently. This combination of C and γ narrowed the support vectors, but increased their sensitivity to data points further away, which means the function uses less data to set the hyperplane, but more data to determine the shape.

A more precise classification of SAV was observed with a U_Accuracy (user accuracy) of 86.7% (Table 7), however the producer accuracy for SAV was low at 58.2%, which indicates a poor correlation with the validation sample. For use in the field, the classification at Table 7 would probably be more useful than

the classification at Table 5, especially given the large size of Roskilde Fjord, because a researcher could be more certain to find SAV at the locations defined at Table 7.

To test the influence of reducing the C parameter, C was reduced to 500 and gamma was maintained at 0.001 (Table 8). Overall accuracy improved to 72.2%, and small improvements were generally observed for the user and producer accuracies.

Table 8: SVM parameter settings C=500 and gamma= 0.001

SVM: C= 500, gamma= 0.001				
ClassValue	Sand	SAV	Total	U_Accuracy
Sand	46	27	73	0,630
SAV	6	40	46	0,870
Total	52	67	119	0
P_Accuracy	0,885	0,597	0	0,723

Reducing C reduces the hyperplane margins, which appears to have benefited the classification in this instance. It is hypothesised this was beneficial because it increased the robustness of the model to account for the spectral heterogeneity within classes across the fjord, particularly in the northern and southern regions (refer Figure 11). A more robust model has the versatility to classify these regions more accurately, because it treats spectral variability as noise rather than training data.

Comparing Table 7 and Table 8, sand was classified similarly for producer and user accuracy, whereas values improved for SAV. The producer accuracy for SAV increased slightly from 58.2% to 59.7%, which indicates these parameter settings increase the agreement with the training and validation sample, however this value is still low. SAV areas are classified as sand, more often than vice versa, which indicates sand is being mistaken for SAV approximately 40% of the time. This highlights one of the main challenges with this study, because many sand areas have a green tinge from the water body, and are therefore susceptible to being classified as SAV.

Increasing or decreasing C is a trade-off between bias and variance. Reducing C reduces the margins of the classifier and thereby the number of training samples that can influence the classification. This increases bias (the number of assumptions about the form of the target function) and reduces variance (the amount the target function will change given the introduction of new training data) (James, et al., 2013), with the opposite effect when C is increased. This highlights the challenge with small training data sets, as a higher proportion of the data set needs to be incorporated into the model to represent the ground truth, however this can make the model susceptible to overfitting.

RF is resilient to overfitting data due to each decision tree using only a subset of the total data for training (Noi & Kappas, 2017). In these results it demonstrated the most versatility across the entire fjord, highlighted by the evenness of producer accuracy for sand and SAV, the relatively high accuracies obtained for all categories, and classification of sand and SAV in the relatively spectrally homogenous southern fjord region (Figure 41 and Table 9.

Table 9: RF confusion matrix (nTrees= 500)

RF: nTree= 500					
ClassValue	Sand	SAV	Total	U_Accuracy	
Sand		37	18	55	0,673
SAV		15	49	64	0,766
Total		52	67	119	0
P_Accuracy	0,712	0,731	0	0,723	

RF produced an OA of 72.3%, which was slightly lower than SVM (73.1%). The relatively even results achieved for user and producer accuracies for both classes highlight the ability of RF to account for all variables equally during parameter selection.

The OA classification result achieved with Bayes (72.3%) (Table 10) was comparable to SVM and RF, however closer inspection of the confusion matrix indicates the OA was driven by biased classification of SAV over sand, which

highlights why interpreting solely OA as an indication of the classification accuracy can be misleading.

Table 10: Bayes confusion matrix

ClassValue	Bayes			U_Accuracy
	Sand	SAV	Total	
Sand	29	10	39	0,744
SAV	23	57	80	0,712
Total	52	67	119	0
P_Accuracy	0,558	0,851	0	0,723

Bayes achieved the highest OA with a SP of 20, which resulted in image objects with greater spectral heterogeneity than used for the other MLAs. These SP settings however biased the classification to over classify SAV, as demonstrated in the producer accuracy (Table 10), which indicates that 85.1% of the validation samples were classified correctly for SAV, however only 55.8% for sand.

Overall, it is assessed that RF performs best for classification of SAV, due to the following reasons:

- High overall accuracy,
- Even producer and user accuracy for both feature classes, and,
- Versatility to classify accurately across a large study area.

However, selection of MLA for future studies should be undertaken with consideration to the size of the training sample.

4.3 Can Water Column Correction and Multi-Temporal Analysis to improve classification results?

Many techniques are available to improve classification accuracy throughout the OBIA workflow (refer Figure 19). Several of these techniques have already been demonstrated, such as the selection of appropriate MLA, segmentation algorithm selection and parameter settings, and training

samples. In addition to these, two other techniques were examined – Water Column Correction (WCC) and Multi-Temporal Analysis (MTA).

Water column correction (WCC) removes the effect of the water column on the underlying bathymetry, thereby increasing homogeneity and improving suitability for segmentation and classification (Figure 43). It is necessary for MTA because it normalises the water body within each image.

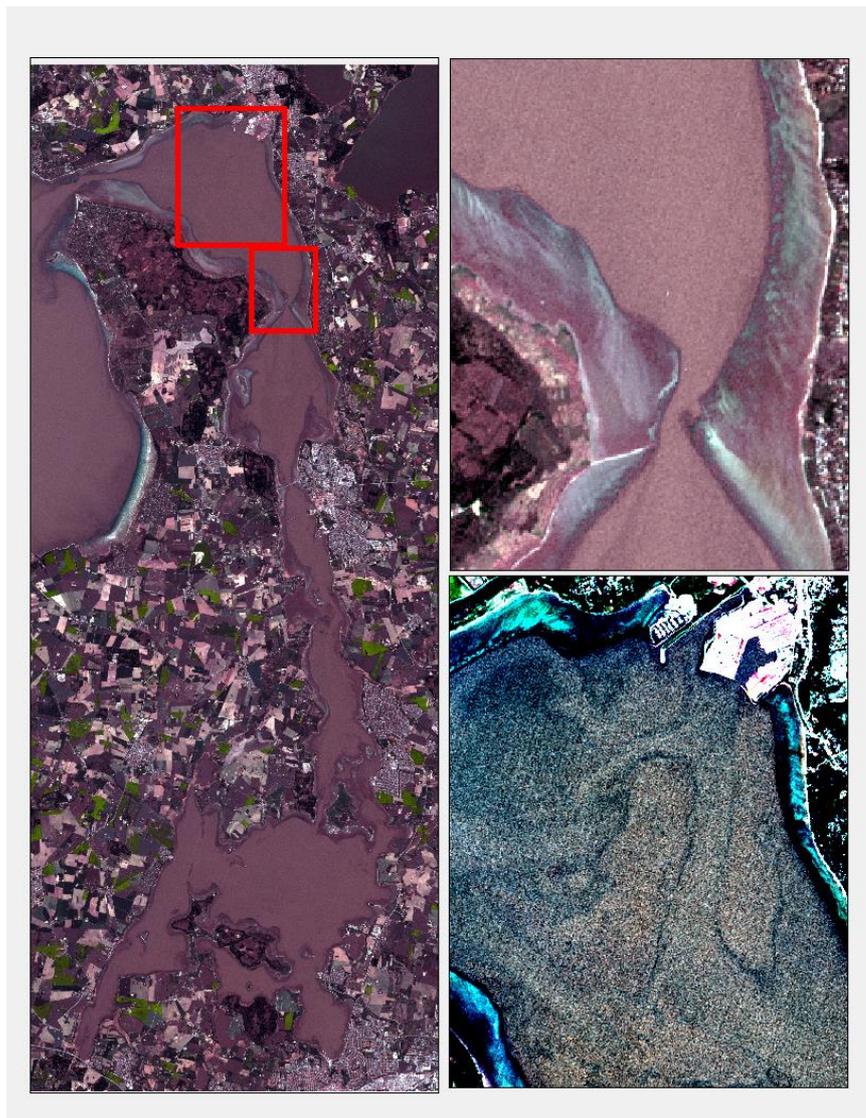


Figure 43: Roskilde Fjord with water column correction applied. Moving clockwise L to R: the entire fjord, the bottleneck, the northern fjord with image stretch applied.

It can be observed that the bathymetry is spectrally more homogenous, and SAV beds and meadows remain visible along the banks and deep areas of the fjord (Figure 43). This imagery was classified using the SVM classifier, with the same parameter and segmentation settings that achieved the highest OA (SP10) (refer Section 4.2) (Figure 44).

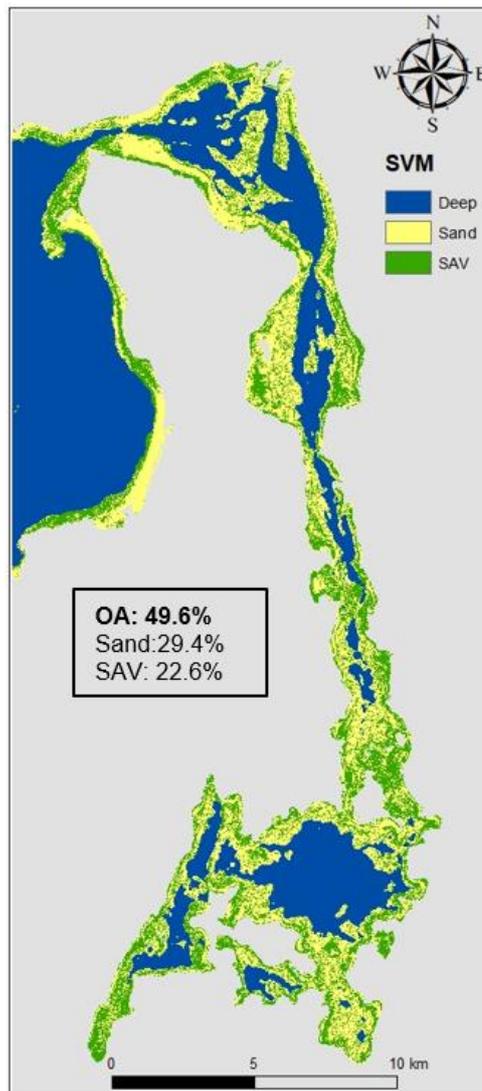


Figure 44: WCC classification with SVM

It can be observed that the improved spectral homogeneity from the WCC resulted in smaller image objects (Figure 44). More sand (29.4%) and less SAV (22.6%) was detected.

Table 11: Confusion matrix of SVM classifier with WCC

WCC - SVM: C= 10,000, gamma= 0.0001				
ClassValue	Sand	SAV	Total	U_Accuracy
Sand	29	36	65	0,446
SAV	22	30	52	0,577
Total	52	67	119	0
P_Accuracy	0,558	0,448	0	0,496

OA was 49.6 % (Table 11), which is significantly lower than the accuracy obtained by the SVM classifier (73.3%) with the non-WCC image. From these results, it is evident the WCC did not improve classification accuracy.

There are several possibilities for these observations. Firstly, the large size of the fjord could have resulted in a single delta relative depth being inadequate to normalise the entire scene using this value alone. The southern region exhibits greater water turbidity than the north, which could influence the overall WCC result because an area in the northern section of the fjord was used to extract the relative depth values. This methodology could be improved by extracting separate relative bathymetries at multiple areas of the fjord to better account for this variability. Secondly, the creation of points to extract the relative depths is a subjective process and therefore prone to operator error. The objective when creating depth extraction points is to maximise coverage of depth values across a single bottom type. This is based on the operator's judgement to identify the bottom type based on appearance in the imagery. This could be improved by trial and error.

Another technique that can yield classification improvements is MT analysis. This method was hypothesised to be particularly useful for this study to visualise seasonal changes, and distinguish eelgrass from features that could confuse the classification, such as black mussels and rocks. The imagery displayed visible changes across the timespan, with increased SAV coverage particularly evident along the banks (Figure 45), which is expected with

increased sunlight. These areas are in relatively shallow water, and the effect of the water column is not particularly visible, therefore it was anticipated this technique could yield improved results without undertaking WCC for each scene. The image was classified with SVM, using the same parameter settings that achieved highest OA (refer Section 4.2). The result of this classification is overlaid on the Sep2016 image (Figure 45).

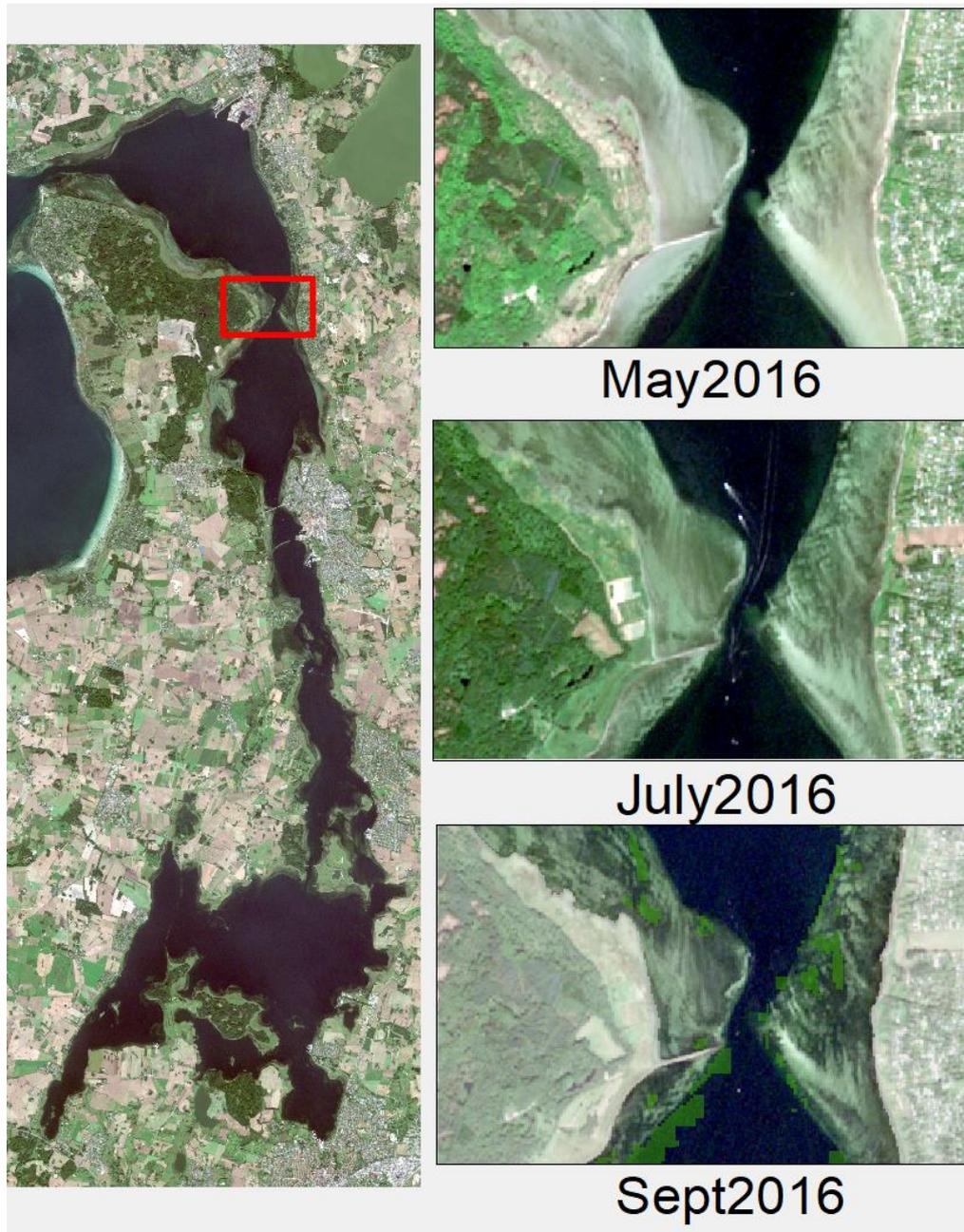


Figure 45: Multi-temporal analysis, with classification overlaid on Sep2016 image

It can be observed at Figure 45 the classifier failed to capture SAV growth moving later into the season, and given OA reduced (66.4%) (Table 12), it seems to have added redundant and confusing information.

Table 12: Confusion matrix for multi-temporal analysis with SVM classifier

Multi-temporal - SVM: C=10,000, gamma=0.0001				
ClassValue	Sand	SAV	Total	U_Accuracy
Sand	40	28	68	0,588
SAV	12	39	51	0,765
Total	52	67	119	0
P_Accuracy	0,769	0,582	0	0,664

This result is hypothesised to have occurred for the following reasons. Firstly, the water column could have influenced the result, despite not displaying visual differences as described earlier, which highlights the subtleties within remote sensing. WCC was not undertaken on each image, due to the lack of improvement in the OA (as earlier described), however if WCC was optimised it is hypothesised it would be beneficial. Secondly, this result could be improved with more accurate training data, that includes each species of SAV and other aquatic features, and reflects eelgrass phenology throughout the year.

Although WCC and MTA did not improve classification accuracy in this thesis, it is hypothesised both would be beneficial if the techniques are optimised.

4.4 *How can classified SAV be used to guide researchers to find Eelgrass?*

It is assessed SAV was classified with reasonable accuracy across the fjord, based on visual observations of classification performance (Figure 42). To provide an empirical approach to indicate the likelihood of whether the classified SAV is eelgrass, output from a 2011 DHI MIKE 3D model of eelgrass stress factors was overlaid on the classification output (Figure 46).

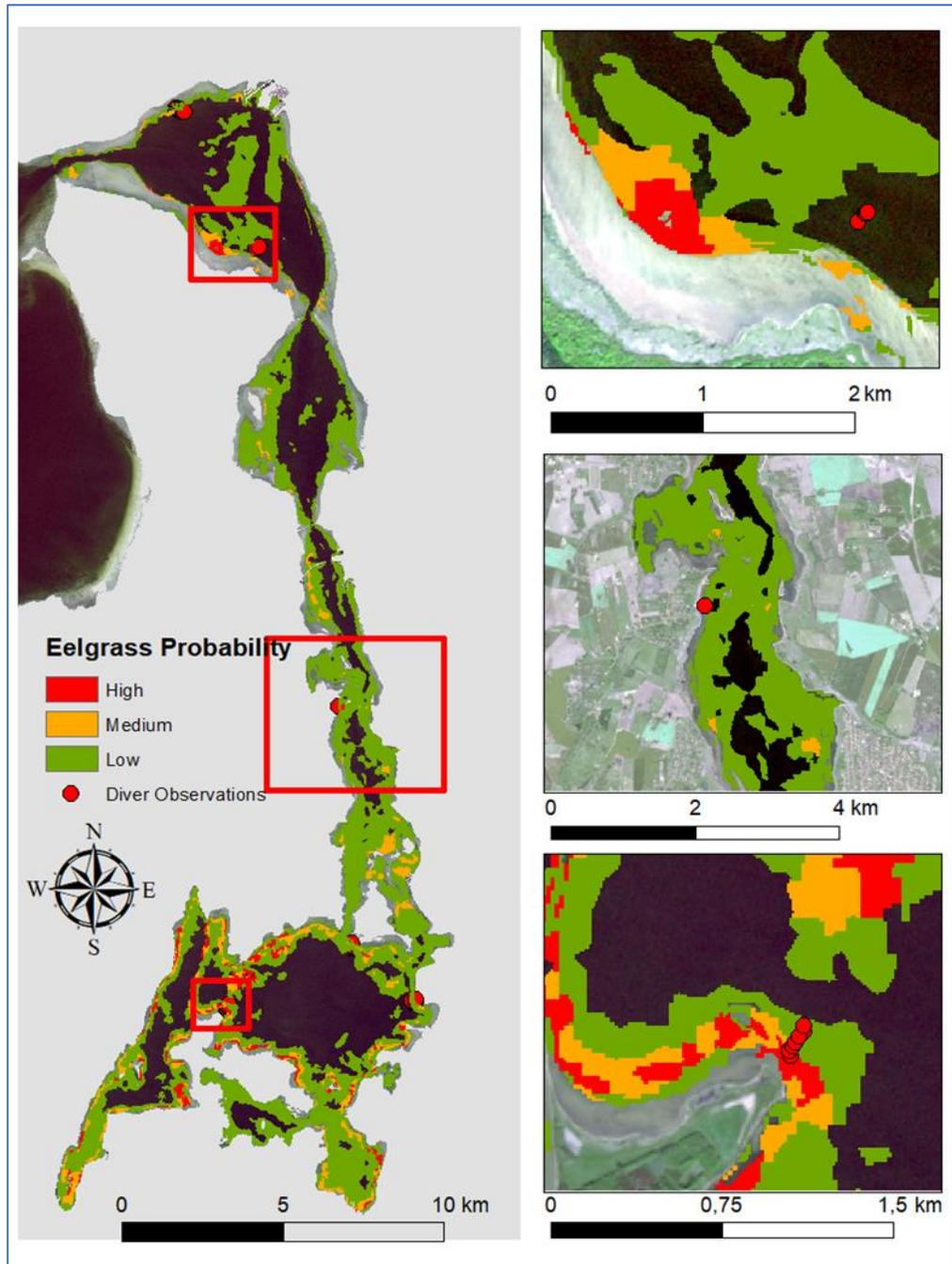


Figure 46: Classification with eelgrass modelled stressors overlaid

The areas in Roskilde Fjord where eelgrass is likely to grow is illustrated at Figure 46. It can be observed there are generally more areas in the southern region where high probabilities of eelgrass are predicted. Nutrient influx has been demonstrated to be a major eelgrass stress factor (Kuusemäe, et al., 2016), which is heavily influenced by surrounding land use. It is unknown what the surrounding land use is, however crops are visible in the satellite imagery

in many areas of the southern fjord. This could be an indication that fertilizer application is not adversely affecting water quality, and is being properly managed.

Locations where diver data was collected as part of the Danish National Monitoring Program are indicated Figure 46. It can be observed that the dive sites generally correlate with areas of high probability of eelgrass growth. However, there are also areas where the dive sites do not correlate with these areas, such as in the northern region of the fjord (top right box, Figure 46). Eelgrass is more likely to be found along the shore than deeper areas, which supported by the literature (Hansen, et al., 2015).

The DHI MIKE 3D model of eelgrass stress factors was generated in 2011, so it is unlikely these predictions are accurate today, as water quality has improved due to the reduction of nutrient inflow from surrounding land use (Hansen, et al., 2015). Nevertheless, this technique demonstrates a robust method to illustrate the location of eelgrass crop from moderate resolution imagery.

This technique could be improved by undertaking a regression to determine the correlation between dive sites and probability of areas likely to contain eelgrass. Field data collection to verify the accuracy of the classification and model would be an interesting next step.

5 Conclusion

The aim of this thesis was to evaluate the feasibility of remote sensing with Sentinel-2 imagery to provide guidance for eelgrass field data collection, using Roskilde Fjord as a case study.

It demonstrated that classifying eelgrass using spectral signature alone with moderate spatial resolution imagery it not feasible due to spectral mixing from varies SAV types, scattering and attenuation in the water body, spectral signature from the seabed, and atmospheric and water column effects; however, detecting SAV is feasible.

Scale parameter was demonstrated as important for determining the spectral heterogeneity of image objects, which influences classification accuracy depending on how accurately the image objects represent the spectral signature of the features of interest. It demonstrated that image segmentation is important for determining OA.

Although the highest classification accuracy was achieved with SVM, RF is recommended for future classification of SAV due to high OA and tendency to not overfit the training data. An interesting next step would be to collect field data to validate the MLA results.

Several techniques to improve the classification accuracy of eelgrass detection were presented, however it is hypothesised accurate WCC could yield the highest improvements, particularly over large study areas where the water column turbidity is highly variable. It is anticipated MT analysis could also yield improved results, provided accurate atmospheric correction and WCC is undertaken, and thorough training data available.

A technique to guide field studies to eelgrass locations was demonstrated by overlaying a model simulating eelgrass stress parameters with classified SAV. Combining these techniques provided an empirical approach to locating eelgrass, rather than speculating about the identity of SAV which can vary significantly throughout a study area. This technique could be improved by overlaying a model of eelgrass stress parameters with up-to-date data. An interesting next step would be to validate the research by undertaking field samples on the high, medium and low probability areas to determine if they correlate with the predicted values.

The results from this study are significant because they build on existing research into remote sensing of SAV, and demonstrate the benefits of fusing spectral characteristics with modelling results to provide an empirical result. A limitation of the study was a lack of accurate field data, however this also provided an opportunity to demonstrate a technique to classify SAV from manual observations. This is valuable because it is likely accurate field data will not usually be available, particularly in remote locations or very large study areas.

6 Future direction

This study provides a basis to implement several current and emerging technologies to build on the results.

1. Use imagery with higher spectral resolution

One of the key challenges with this study was distinguishing features from the moderate spectral resolution of the imagery. Imagery with higher spectral resolution can distinguish features more accurately, such as hyperspectral imagery, which obtains spectral signatures on much narrower bandwidths than multi-spectral imagery. This is accomplished by hyperspectral sensors that acquire images in many, very narrow, contiguous spectral bands across a broad range of spectra, which constructs an effectively continuous reflectance for every pixel in a scene (Lillesand & Kiefer, 1999). Such imagery is useful for distinguishing features that have very similar spectral signatures. Several studies have collected in-situ reflectance measurements for eelgrass (e.g. (O'Neill, et al., 2011)), and the species composition and bottom composition of Roskilde Fjord has been well studied and documented (e.g. (Krause-Jensen, et al., 2004)). If the spectral signatures (endmembers) of all SAV and bottom composition could be established, an accurate classification using hyperspectral imagery could be undertaken.

Hyperspectral imagery can be collected from sensors mounted to a variety of vehicles, such as drones, aircraft and satellites (Lillesand & Kiefer, 1999). Aircraft imagery, collected at relatively low altitudes, is high resolution however is expensive due to the cost of aircraft operation and can also be hazardous for the aircraft to fly in certain locations, altitudes and weather conditions. Drones provide a solution to these challenges because they are unmanned and can be programmed, and as a result are low cost and low risk. Drones can be operated at very low altitudes, which enables collection of very

high spatial resolution imagery that can be used for classification, and is an emerging technology in this field.

2. Collect more accurate training data

As documented, a major limitation of this study was the absence of accurate training data, collected from in-situ observations. While collection of such data is unlikely to improve OA, it does enable more confident interpretation of results because the values do not incorporate any speculation about identity of features.

One of the major challenges with collecting submerged training data is the time and effort required to collect it, as it usually involves diver transects which are slow and laborious (Roelfsema, et al., 2014) (Armstrong, 2016). However, advances in technology are providing solutions. Software algorithms that use a python code to detect eelgrass from the shape of the blades enable eelgrass leaf area to be extracted automatically from video transects (Figure 47) (Sengupta, et al., 2018). This greatly improves the efficiency of data collection because traditional methods rely on estimation of plant coverage over a particular area.



*Figure 47: Python based software that detects eelgrass from blade shape
(source: (Sengupta, et al., 2018))*

Methods for collection of video transects are also being developed. Diver transects are traditionally recorded by a diver with a camera, or a suspended camera under a vessel. These methods are relatively slow and laborious, particularly the diver collected data, as the diver is required to surface regularly to record GPS location. In situ platforms, such as autonomous underwater vehicles (AUVs), remotely operated vehicles (ROVs), towed platforms and drop cameras are being developed primarily for data collection in areas where remote sensing is not feasible due to water depth or turbidity (Armstrong, 2016). However, these technologies could also be used to collect field data for remote sensing to reduce time and cost, particularly in large and/or hazardous areas.

3. Reduce scene area

This thesis demonstrated the challenges with classifying a large scene, such as accounting for the spectral variation in the northern and southern sections of the fjord. Dividing the scene to isolate these areas, and performing separate classifications on each, is expected to improve OA, assuming accurate WCC cannot be performed for the entire fjord.

4. Segment SAV on multiple image object levels

A visual inspection of the scene indicates that SAV grows in Roskilde Fjord on multiple spatial scales, with smaller patches usually growing along the banks and larger beds and meadows in the deeper areas. Such areas would benefit from segmentation on multiple scales, which could be achieved by segmenting according to an image object hierarchy, as demonstrated by (Lyons, et al., 2012).

5. Account for within scene variation of atmosphere and water column

As demonstrated with the WCC (refer Section 4.3), accurately accounting for light attenuation at varying depths and water clarity is a

challenge. More stable and effective WCC techniques can provide a purer spectral response from the underlying substrate and SAV. To determine the effect of the water column, further research could focus on conducting this classification of the same scene several times throughout the year, and observe the OA as water quality varies.

The same can be said for atmospheric correction, although this is less important than accurate WCC, because it is likely to vary on a smaller spatial scale than water column variation, and therefore less likely to vary significantly within a scene. As a result, atmospheric correction is only important if a multi-temporal analysis is to be conducted, to account for atmospheric variation within each scene.

An atmospheric correction technique that is specialised for use over water bodies is Case-2 Regional/Coast Colour (C2RCC), which is implemented through the Sentinel Application Platform (SNAP) software from ESA. This application is specialised for atmospheric correction over water bodies because it provides water leaving reflectance.

6. Bathymetric LiDAR

Bathymetric lidar uses a laser in the green spectrum to measure bathymetry. Since vegetation is highly reflective in this region, bathymetric lidar could be used to determine the presence/absence of SAV. Given the very high resolution of lidar point cloud data, this could be used in addition to remote sensing classification to distinguish SAV from other features, such as rocks and mussels. A setback of this technique is the high cost.

7 References

- Armstrong, R. A., 2016. Landscape-Level Imaging of Benthic Environments in Optically-Deep Waters. In: C. W. Finkl & C. Makowski, eds. *Seafloor Mapping Along Continental Shelves*. Switzerland: Springer International Publishing, pp. 261-273.
- Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 65, pp. 2-16.
- Breiman, L., 2001. Random Forests. *Machine Learning*, 45(1), pp. 5-32.
- Bruhn, A. et al., 2013. *Ålegræs og anden vegetation på kystnær blød bund*, Aarhus: Aarhus Universitet.
- Baatz, M. & Schape, A., 1999. *Object-Oriented and Multi-Scale Image Analysis in Semantic Networks*. Enschede, The Netherlands, 2nd International Symposium: Operationalization of Remote Sensing.
- Carstensen, J., Krause-Jensen, D. & Blasby, T. J., 2015. Biomass-Cover Relationship for Eelgrass Meadows. *Estuaries and Coasts*, 39(1), pp. 440-450.
- Cho, H. J., Mishra, D. & Wood, J., 2012. Remote sensing of submerged aquatic vegetation. *Remote Sensing – Applications*, pp. 297-308.
- Cunha, A. H., Duarte, C. M. & Krause-Jensen, D., 2004. How long time does it take to recolonize seagrass beds?. In: J. Borum, C. M. Duarte, D. Krause-Jensen & T. M. Greve, eds. *European seagrasses: an introduction to monitoring and management*. Munich: The M&MS project, pp. 72-76.
- Dogan, O. K., Akyurek, Z. & Beklioglu, M., 2009. Identification and mapping of submerged plants in a shallow lake using quickbird satellite data. *Journal of Environmental Management*, Volume 90, pp. 2138-2143.
- Drăguț, L., Tiede, D. & Levick, S. R., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science*, pp. 859-871.

- Duarte , C. M., Marba, N. & Santos, R., 2004. What may cause loss of seagrasses?. In: *European seagrasses: an introduction to monitoring and management*. Munich: The M&MS project Date, pp. 24-33.
- Duarte, C. M., Alvarez, E., Grau, A. & Krause-Jensen, D., 2004. Which monitoring strategy should be chosen?. In: *European seagrasses: an introduction to monitoring and management*. Munich: The M&MS project Date, pp. 41-45.
- Duro, D. C., Franklin, S. E. & Dube, M. G., 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment*, 118(1), pp. 259-272.
- European Space Agency, 2018. *Sentinel-2 - Radiometric Resolutions*. [Online] Available at: <https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/radiometric>
- Felger, R. & Moser, M. B., 1973. Eelgrass (*Zostera marina* L.) in the Gulf of California: Discovery of its Nutritional Value by the Seri Indians. *Science*, 181(12), pp. 355-356.
- Foody, G. M., 2011. Impacts of imperfect reference data on the apparent accuracy of species presence–absence models and their predictions. *Global Ecology and Biogeography*, 20(1), pp. 498-508.
- Gascon, F. et al., 2017. Copernicus Sentinel-2A Calibration and Products Validation Status. *Remote Sensing*, 9(6), p. 584.
- Hansen, E., Roark, L. & Parfitt, K., 2015. *Roskilde Fjord Portal*. [Online] Available at: www.roskilde-fjord.dk [Accessed 30 May 2018].
- Haynes, R. R., 2000. *Zosteraceae: Eel-grass Family*. 1st ed. New York: Oxford University Press.

Hossain, M. S., Bujang, J. S., Zakaria, M. H. & Hashim, M., 2015. The application of remote sensing to seagrass ecosystems: an overview and future research prospects. *International Journal of Remote Sensing*, 36(1), pp. 61-114.

James, G., Witten, D., Hastie, T. & Tibshirani, R., 2013. *An Introduction to Statistical Learning with Applications in R*. New York: Springer.

Klemas, V. V., 2016. Remote Sensing of Submerged Aquatic Vegetation. In: C. W. Finkl & C. Makowski, eds. *Seafloor Mapping along Continental Shelves - Research and Techniques for Visualizing Benthic Environments*. Switzerland: Springer International Publishing, pp. 125-141.

Krause-Jensen, D., Marba, N., Alvarez, E. & Grau, A., 2004. How is seagrass habitat quality monitored?. In: J. Borum, C. M. Duarte, D. Krause-Jensen & T. M. Greve, eds. *European seagrasses: an introduction to monitoring and management*. Munich: The M&MS project, pp. 63-67.

Krause-Jensen, D., Quaresma, A. L., Cunha, A. H. & Greve, T. M., 2004. How are seagrass distribution and abundance monitored?. In: J. Borum, C. M. Duarte, D. Krause-Jensen & T. M. Greve, eds. *European seagrasses: an introduction to monitoring and management*. Munich: The M&MS project Date, pp. 45-54.

Kuusemäe, K., Rasmussen, E. K., Canal-Vergés, P. & Flindt, M. R., 2016. Modelling stressors on the eelgrass recovery process in two Danish estuaries. *Ecological Modelling*, pp. 11-42.

Lary, D. J., Alavi, A. H., Gandomi, A. H. & Walker, A. L., 2016. Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, pp. 3-10.

Lathrop, R. G., Montesano, P. & Haag, S., 2006. A Multi-scale Segmentation Approach to Mapping Seagrass Habitats Using Airborne Digital Camera Imagery. *Photogrammetric Engineering & Remote Sensing*, 72(6), pp. 665-675.

- Lavendar, S. & Lavendar, A., 2016. *Practical Handbook of Remote Sensing*. 1st ed. Florida: CRC Press.
- Liaw, A. & Wiener, M., 2002. Classification and Regression by randomForest. *R News*, 2(3), pp. 18-22 .
- Lillesand, T. M. & Kiefer, R. W., 1999. *Remote Sensing and Image Interpretation*. 4th ed. New York: John Wiley & Sons.
- Lyons, M. B., Phinn, S. R. & Roelfsema, C. M., 2012. Long term land cover and seagrass mapping using Landsat and object-based image analysis from 1972 to 2010 in the coastal environment of South East Queensland, Australia. *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 71, pp. 34-46.
- Lyons, M. B., Roelfsema, C. M. & Phinn, S. R., 2013. Towards understanding temporal and spatial dynamics of seagrass landscapes using time-series remote sensing. *Estuarine, Coastal and Shelf Science*, Volume 120, pp. 43-53.
- Lyzenga, D., 1978. Passive remote sensing techniques for mapping water depth and bottom features. *Applied Optics*, 17(1), pp. 379-383.
- Marba, N., Duarte, C. M., Alexandra, A. & Cabaco, S., 2004. How do seagrasses grow and spread?. In: *European seagrasses: an introduction to monitoring and management*. Munich: The M&MS project Date, pp. 11-19.
- Maxwell, T., 2005. *Object-oriented classification: classification of pan-sharpened Quickbird imagery and a fuzzy approach to improving image segmentation efficiency*, Fredericton, N.B: Department of Geodesy and Geomatics Engineering, University of New Brunswick.
- Misbari, S. & Hashim, M., 2016. Change Detection of Submerged Seagrass Biomass in Shallow Coastal Water. *Remote Sensing*, 8(12), pp. 1-29.
- Mumby, P. J., Green, E. P., Edwards, A. J. & Clark, C. D., 1997. Measurement of seagrass standing crop using satellite and digital airborne remote sensing. *Marine Ecology Progress Series*, Volume 159, pp. 51-60.

Myint, S. W., Yuan, M., Cervený, R. S. & Chandra, G. P., 2008. Comparison of Remote Sensing Image Processing Techniques to Identify Tornado Damage Areas from Landsat TM Data. *Sensors*, 8(2), pp. 1128-1156.

Noi, P. T. & Kappas, M., 2017. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors*, 18(1), pp. 1-20.

O'Neill, J. D., Costa, M. & Sharma, T., 2011. Remote Sensing of Shallow Coastal Benthic Substrates: In situ Spectra and Mapping of Eelgrass (*Zostera marina*) in the Gulf Islands National Park Reserve of Canada. *Remote Sensing*, 3(5), pp. 975-1005.

Ouyang, Z., 2015. *Object-Based Classification & eCognition*. [Online] Available at: <http://lees.geo.msu.edu>

Pal, M., 2005. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing ISSN;* 26(1), pp. 217-222.

Pedersen, T. M., Sand-Jensen, K., Markager, S. & Nielsen, S. L., 2014. Optical Changes in a Eutrophic Estuary During Reduced Nutrient Loadings. *Estuaries and Coasts*, 37(4), pp. 880-892.

Phinn, S. et al., 2008. Mapping seagrass species, cover and biomass in shallow waters: An assessment of satellite multi-spectral and airborne hyper-spectral imaging systems in Moreton Bay (Australia). *Remote Sensing of Environment*, Volume 112, pp. 3413-3425.

Qian, Y. et al., 2015. Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery. *Remote Sensing*, Volume 7, pp. 153-168.

Remotepixel, 2018. *Remotepixel Viewer*. [Online] Available at: <http://viewer.remotepixel.ca>

Robbins, B. D. & Bell, S. S., 1994. Seagrass landscapes: a terrestrial approach to the marine subtidal environment. *Trends in Ecology and Evolution*, 8(9), pp. 301-304.

Roelfsema, C. M. et al., 2014. Multi-temporal mapping of seagrass cover, species and biomass: A semi-automated object based image analysis approach. *Remote Sensing of Environment*, Volume 150, pp. 172-187.

Roelfsema, C. M., Phinn, S. R., Udy, N. & Maxwell, P., 2009. An integrated field and remote sensing approach for mapping Seagrass Cover, Moreton Bay, Australia. *Journal of Spatial Science*, 54(1), pp. 45-62.

Salo, T. & Bostrom, C., 2014. Population specific salinity tolerance in eelgrass (*Zostera marina*). *Journal of Experimental Marine Biology and Ecology*, 461(1), pp. 425-429.

Sengupta, S., Stockmarr, A. & Ersbøll, B., 2018. *Eel grass detection from underwater videos using a robust line based algorithm*. s.l.:Ongoing research project at DTU.

Smith, A., 2010. Image segmentation scale parameter optimization and land cover classification using the Random Forest algorithm. *Journal of Spatial Science*, 55(1), pp. 69-79.

Terrados, J. & Borum, J., 2004. Why are seagrasses important? – Goods and services provided by seagrass meadows. In: *European seagrasses: an introduction to monitoring and management*. Munich: The M&MS project, pp. 8-11.

The Danish Council for Strategic Research, 2017. *novagrass - innovative eelgrass restoration techniques*. [Online]
Available at: <http://www.novagrass.dk/en/facts-about-eelgrass/eelgrass/>
[Accessed 30 May 2018].

The IUCN Red List of Threatened Species, 2017. *Zostera marina (Eelgrass)*. [Online]

Available at: <http://www.iucnredlist.org/details/153538/0>

[Accessed 27 May 2018].

Trimble Germany GmbH, 2016. *eCognition Developer 9.2 User Guide*. 9.2.0 ed. Munich: Trimble Germany GmbH.

Visser, F., Wallis, C. & Sinnott, A., 2013. Optical remote sensing of submerged aquatic vegetation: opportunities for shallow clear water. *Limnologica - Ecology and Management of Inland Waters*, 43(5), pp. 388-398.

Yang, C. et al., 2010. Analysis of seagrass reflectivity by using a water column correction algorithm. *International Journal of Remote Sensing*, 31(17-18), pp. 4595-4608.

Yoon, G.-w., Cho, S. I., Jeong, S. & Park, J.-h., 2003. *Object oriented classification using Landsat images*. Busan, South Korea, 24th Asian conference on remote sensing.

Zhang, Y. & Maxwell, T., 2006. *A fuzzy logic approach to supervised segmentation for object-oriented classification*. Reno, Nevada, ASPRS 2006 Annual Conference.

Ørberg, S. B. et al., 2018. *Kortlægning af ålegræsenge med ortofotos*, Aarhus: Aarhus Universitet, Institut for Bioscience.

8 Appendix 1

```
1 for %%f in (*) do gdalwarp -tr 10 10 -r near %%f E:\James\RoskildeFjord\Imagery\S2Raw\S2A_OPER_MSI_L1C_TL_SGS_20160912T160036_A006397_T32VPH_N02.04\10m\%%f
```

Figure 48: Batch file calling gdalwarp to resample each band to 10m resolution using nearest neighbour

```
C:\Users\jafe\S2A_OPER_MSI_L1C_TL_SGS_20160912T160036_A006397_T33UUB_N02.04\IMG_DATA>gdalbuildvrt -separate output.vrt C:\Users\jafe\RoskildeFjord\S2Raw\S2A_tile_20160512_33UUB_\10m/*.jp2  
0...10...20...30...40...50...60...70...80...90...100 - done.
```

Figure 49: gdalbuildvrt command, with -separate, to build a virtual raster placing each input .jp2 image into a separate band

```
(sentinel2) C:\Users\jafe\RoskildeFjord\S2Raw>python C:\OSGeo4W64\bin\gdal_merge.py -o gdalmerge.vrt output2.vrt output1.vrt
```

Figure 50: gdal_merge.py is a python command that automatically mosaics a set of images

```
C:\Users\jafe\RoskildeFjord\S2Raw>gdalwarp -cutline AOI.shp -crop_to_cutline gdalMerge.vrt 20160912.tif
```

Figure 51: gdalwarp command used to clip .vrt with AOI.shp

i