AALBORG UNIVERSITY
STUDENT REPORT

A TILE BASED PROTOCOL FOR THE CREATION OF SENTINEL 2 MONTHLY IMAGE COMPOSITES FOR CROP TYPE CLASSIFICATION

VLAD MIHAI ROSCA

MASTER THESIS – JUNE 2018

COPENHAGEN 2018

Title:
A Tile Based Protocol for the Creation of Sentinel 2 Monthly Image Composites for Crop Type Classification

Keywords:
Remote Sensing, Sentinel, Satellite Image Compositing, Geometric Median, Medoid, Crop Type Classification

Author:
Vlad Mihai Rosca

Principal Supervisor:
Jamal Jokar Arsanjani, PhD. AAU Geoinformatics

Secondary Supervisor:
Kenneth Grogan, PhD. DHI GRAS

Project period:	1st September 2017 – 8th June 2018
ECTS:	60 ECTS
Education:	MSc in Geoinformatics
University:	Aalborg University Copenhagen

AAU Master of Science Programme in Geoinformatics
A. C. Meyers Vænge 15
2450 København SV
Denmark

Preface

This thesis has been submitted as a capstone project for the Master of Science in Geoinformatics at Aalborg University Copenhagen.

The thesis was initiated and completed in close collaboration with DHI GRAS, where I received a great deal of assistance from all my colleagues during the data collection and processing stages of this project.

Many thanks to my two supervisors, Jamal and Kenneth. Jamal for his help and support throughout university, and Kenneth for his assistance in initiating, developing and implementing this project at DHI GRAS.

Vlad Rosca

Copenhagen, June 2018.

Abstract

Cloud cover presents many challenges to the analysis of satellite imagery. While high temporal resolution can alleviate some of these challenges, for time sensitive objectives, such as crop type classification and crop yield estimation, the implementation of image composites may provide a big advantage.

In this report, three tile based, Sentinel 2 monthly image compositing methodologies are presented. The resulting composites are then compared against each other and against a clearest, least cloudy image for the purposes of crop type classification.

The study area is Sentinel 2 tile 32UNG, an area which saw very high cloud cover throughout the growing season in 2017. The initial acquired images are atmospherically corrected and cloud masked before composites are created for the months of March and July 2017 utilizing the maximum NDVI, medoid, and geometric median compositing methodologies. Indexes are calculated on the resulting composites to assist with crop type classification.

The early (March) and late (July) growing season images are stacked and prepared for classification. The crop type classification in this study is performed using random forests, with the number of trees set at 400. The overall and per crop classification accuracies of the composites are compared against a clearest, most cloud free image from the same period.

The composite images provided an increase of 23.5% in the total number of fields available for classification. The overall accuracies for this study were 84.79%, 84.21%, 83.57%, and 81.35% for the clearest image, medoid, geometric median, and maximum NDVI classifications, respectively. Our conclusion is that these results indicate that monthly image composites can be beneficial for the purpose of cloud-free crop type monitoring and classification. The advantages of high-dimensional image composites for other classification purposes should be further researched.

Table of Contents

List of Abbreviations

AVHRR – Advanced Very High Resolution Radiometer

BRDF – Bidirectional Reflectance Distribution Function

CAP – Common Agricultural Policy

ESA – European Space Agency

GDAL – Geospatial Data Abstraction Library

IPCC – International Panel on Climate Change

NASA – National Aeronautics and Space Agency

NDI – Normalized Difference Index

NDII – Normalized Difference Infrared Index

NDVI – Normalized Difference Vegetation Index

NDWI – Normalized Difference Water Index

NRCAN – Natural Resources Canada

SAR – Synthetic Aperture Radar

SWIR – Short Wave Infrared

USGS – United States Geological Survey

List of Equations

List of Figures

List of Tables

1 Introduction

In many areas of the world, high cloud cover presents many challenges to the analysis of satellite imagery. While high temporal resolution can alleviate some of these challenges, for time sensitive objectives, such as crop type classification and crop yield estimation, the implementation of image composites may provide a big advantage.

In the following sections of this chapter, the report will introduce the background theory and research in the areas of remote sensing, image compositing and image classification and outline the scope and research objectives of his study.

1.1 Background

A general definition of remote sensing is that it is the process of "detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation at a distance" (USGS). For the purposes of this project, remote sensing will refer to the monitoring of the physical characteristics of the earth's surface and the reflectance emitted by it when measured via passive remote sensing. Passive sensors measure the energy which is reflected from the earth's surface (NRCAN). Active sensors on the other hand, such as radar, emit their own source of energy towards the target and reflectance radiation is then measured (Richards & Jia, 2006).

Sensors on board different satellites measure different radiation at specific ranges, at specific spatial, temporal and radiometric resolutions. The resulting images thus have different characteristics depending on the sensor.

1.1.1 Spatial Resolution

The spatial resolution of a satellite image is concerned with the size of the area covered within one pixel. This is generally expressed as the size of the area covered by each pixel in meters. The higher the quality of the resolution is, the greater the level of detail of an object observed on the ground.

Introduction

Figure 1 below shows an example of how different spatial resolutions can affect how objects are scene in a satellite image.

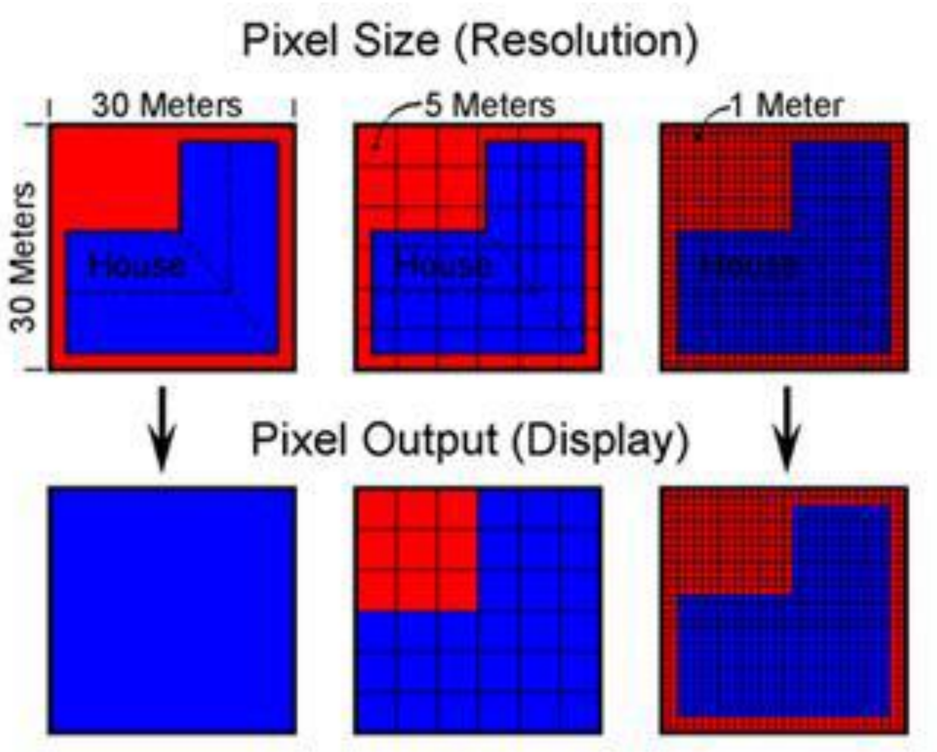

Figure 1 Visual demonstration of pixel sizes. From Satellite Imaging Corporation. (2018). Spatial Resolution [Image]. Retrieved from https://www.satimagingcorp.com/services/resources/characterization-of-satellite-remote-sensing-systems/

Satellites are usually put into four categories based on their spatial resolution:

Table 1 Spatial resolution categories

Category	Spatial Resolution
Low Spatial Resolution	> 1000m
Medium Spatial Resolution	100m – 1000m
High Spatial Resolution	5m – 100m
Very High Spatial Resolution	< 5m

Different factors can negatively or positively influence the spatial resolution of an image. These factors can include:

- The image scale factor - spatial resolution decreases as the scale factor increases.
- The quality of the optical system
- The grain structure of the photographic film
- The contrast of the original objects
- Atmospheric scattering effects – can lead to reduced contrast and resolution
- Image motion – the relative motion between the ground and sensor can cause distortion.

1.1.2 Temporal Resolution

The temporal resolution specifies the revisiting frequency of a satellite sensor for a specific location. Satellite Imaging Corporation (2018) utilizes the following three categories:

Table 2 Temporal resolution categories

Category	Temporal Resolution
Low Temporal Resolution	< 24 hours – 3 days
Medium Temporal Resolution	4 – 16 days
High Temporal Resolution	> 16 days

The revisit period refers to how long a satellite takes to complete an orbit of the earth and return to capture a new image of the same area. The temporal resolution of a satellite is generally inversely correlated with spatial resolution, where very high spatial resolution satellites tend to have a lower temporal resolution and vice versa. The temporal resolution thus, depends on several factors, including swath overlap, satellite capabilities and latitude. (Karimi, 2004)

An overlap of the swath tends to increase the temporal resolution of a satellite, however this is limited to narrow strips as in an ideal orbit the overlap of swaths would be minimised. Due to the curvature of the earth, areas at high latitudes tend to have significantly higher revisit frequencies than those located at the equator.

The time of day that images are captured tends to have an impact on the resulting image and its usage, due primarily to the position of the sun. Sun-synchronous orbit satellites, such as Landsat 8 and Sentinel 2, tend to capture each location at between 10:00 and 12:00 local sun time. This is in order to minimize the effects of shadows and other atmospheric impacts.

1.1.3 Spectral Resolution

A sensor's spectral resolution specifies the number of spectral bands in which the sensor can collect reflected radiation. The number of bands however, is not the only important aspect of spectral resolution. The position and the widths of bands in the electromagnetic spectrum is important as well.

Multi Spectral Instruments record reflectance values in separate bands at various wavelengths. Hyperspectral Instruments can record values in hundreds of separate bands on very narrow wavelengths, showing subtle different between various objects at very high spectral resolutions.

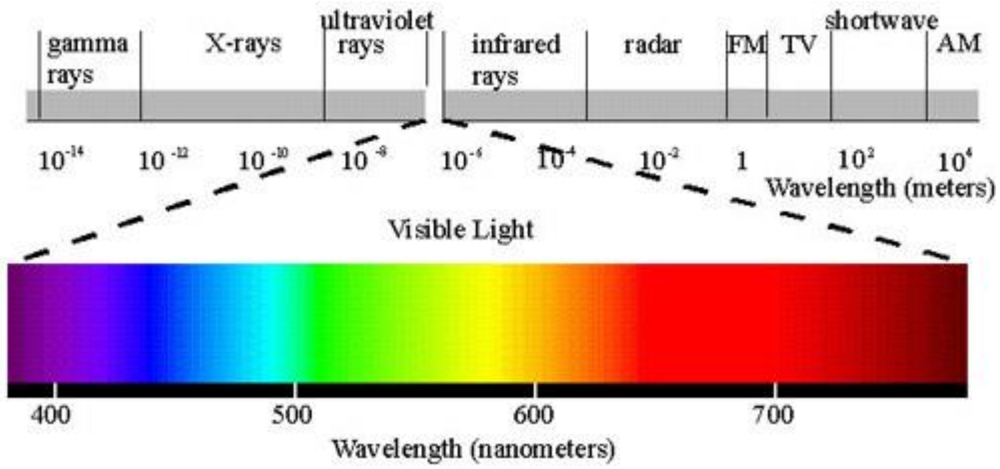

Figure 2 Spectral wavelengths. From Satellite Imaging Corporation. (2018). Spectral Resolution [Image]. Retrieved from https://www.satimagingcorp.com/services/resources/characterization-of-satellite-remote-sensing-systems/

Satellite Imaging Corporation (2018) utilizes the following three categories:

Table 3 Spectral resolution categories

Category	Number of Spectral Bands
Low Spectral Resolution	3 Bands
Medium Spectral Resolution	3 – 15 Bands
High Spectral Resolution	15 – 220 Bands

1.1.4 Radiometric Resolution

Radiometric resolution represents the smallest difference in energy that can be detected by a sensor, also known as the sensitivity of the sensor. The finer the radiometric resolution, the more sensitive it is to detecting small differences in reflectance values.

Radiometric resolution refers to the quantity of information in a pixel and expressed in units of bits. This is represented as a value between 0 and a selected power of 2 minus 1. Thus a radiometric resolution of 8 bits means the pixel has 256 possible intensity values (0 – 255) and 16 bits can represent 65,536 different intensity values. Advances in technology have led to an increase in the radiometric resolution of most satellite systems, now making it easier to detect subtle changes in reflectance.

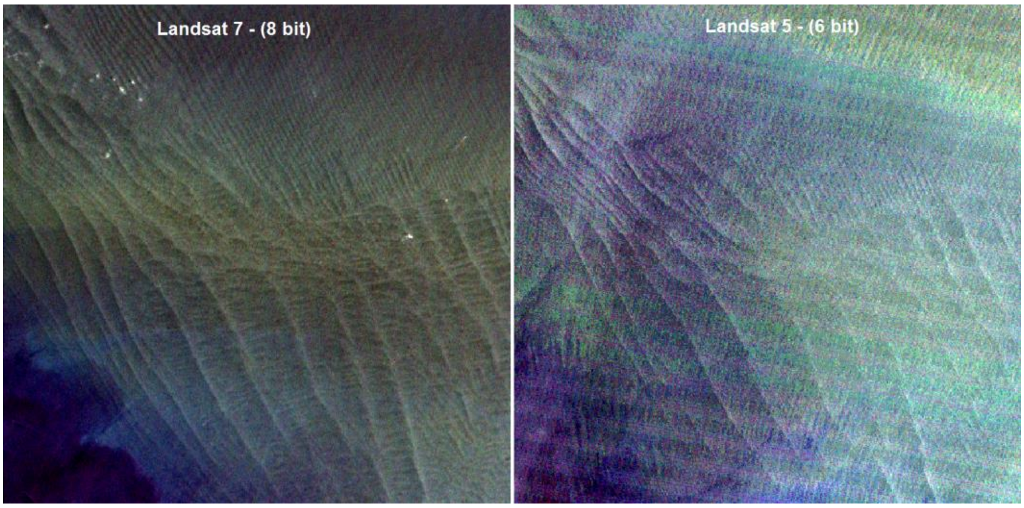

Figure 3 Comparison of different radiometric resolutions

1.2 Image Preprocessing

In order to remove outlying data for the compositing process, images first need to be cloud masked. While most cloud masks are not perfect, this process increases the likelihood that only clear, relevant pixels will be used in the compositing process.

As images in a compositing process are acquired on different dates, different atmospheric factors affect each individual satellite image. In order to obtain bottom of atmosphere reflectance values, atmospheric correction must be implemented in this preprocessing step. In recent times, these atmospheric correction algorithms have evolved and are now based on “rigorous radiative transfer modelling approaches” (Gao, Montes, Davis & Goetz, 2009).

Finally, as Sentinel 2 bands are collected at different spatial resolutions, a resampling algorithms should be applied in order to ensure that all bands utilized in a study are at the same spectral resolution.

1.2.1 Cloud Masking

Satellite imagery is available at a higher temporal resolution than ever before, however, many of these images inevitably capture clouds over many areas of interest. Clouds and cloud shadows are significant sources of noise in satellite images (Zhu & Woodcock, 2012) and create many problems when attempting to identify and classify what is on the ground.

Cloud and cloud shadow identification and removal is an important initial step in most satellite image analysis tasks as their presence can cause issues with atmospheric correction, index calculations and land cover classification (Zhu & Woodcock, 2012).

Generally, thick, opaque clouds can be more easily identified and removed from the image as they have very bright reflectance values when compared to the rest of the image. Their shadows however can still cause issues. In general cloud shadows are darker when compared to the average pixel in the image, however this is not always the case, as shadows have a lesser impact on bright objects in the image such as urban areas or coastal sand.

Secondly, thin, semitransparent, cirrus type clouds can distort the reflectance values observed and are much more difficult to identify and remove due to the fact that their signal includes the cloud, as well as the land beneath it (Gao & Kaufmann, 1995). Sentinel 2 however provides a separate 60m resolution SWIR – Cirrus band (Band 10, 1,375 nm) to assist with thin cloud identification and removal

Many object and pixel based classification algorithms have been developed for the automatic identification and removal of clouds, as manual processes are extremely inefficient and time consuming.

1.2.2 Atmospheric Correction

Reflectance values transferred from the earth's surface to the satellite are highly influenced and modified by interaction with the earth's atmosphere (Hadjimtsis, Papadavid, Agapiou, Themistocelous & Hadjimitsis, 2010). In order correct for these influences of the atmosphere, atmospheric correction algorithms must be implemented in the preprocessing stages of any remote sensing analysis.

Atmospheric correction is one of the most important preprocessing steps when working with remote sensing data (Hadjimtsis et al, 2010). In order to properly study the images acquired by satellites, an accurate removal of these atmospheric effects is required (Gao et al, 2009).

Martins, Barbosa, de Carvalho, Jorge, Lobo & Novo (2017) note that the quality of the final product is highly dependent on the atmospheric correction method and as such proper care must be taken when processing images, especially when working with multi temporal data, as is the case of this project.

Atmospheric correction methods can be largely grouped into three main categories:

1) Empirical methods
2) Radiative Transfer Modelling Approaches
3) Hybrid Approaches

Empirical methods grew in popularity and were mainly developed during the 1980s (Gao, Montes, Li, Dierssen & Davis, 2007). These approaches do not depend on knowledge of outside climatological effects, and the correction is performed on properties derived from the image observation (Martins et al, 2017). Conel, Green, Vane, Bruegge, Alley & Curtiss (1987) have also introduced an empirical line approach which requires field measurements of darkest and brightest objects in the scene, and surface values for all other pixels are then linearly regressed from these observed dark and bright values.

The radiative transfer modelling approaches were first realized in 1993, with development of ARTEM (Gao, Heidebrecht & Goetz, 1993). These approaches use theoretical models which simulate the absorption and scattering effects of atmospheric gases on radiance values (Gao et al, 2007).

The hybrid approaches use a combination of image observed properties and simulated values in order to try and provide a mixed approach, based on empirical measurements and theoretical simulations. Sen2Cor, the atmospheric correction method used in this project is one such approach.

1.3 Image Compositing

Holben (1986) first introduced the idea of a maximum value composite using data from the Advanced Very High Resolution Radiometer (AVHRR) sensor. The goal of the composite was to create cloud-free and contaminant-free spatially continuous images to study vegetation dynamics. The outcome of the compositing algorithm introduced by Holden was that the composite images could be used for more concrete vegetation studies, as the compositing technique minimized cloud interference, atmospheric contamination, sun-angle and shadow effects, and increased the availability of data in the study area.

The idea behind Holben's algorithm still stands today, primarily due to cloud contamination in satellite imagery. It is often highly difficult to acquire fully cloud free images for an area of study, acquired at an agreeable time. Since Holben's introduction of the maximum value composite, a number of different methodologies have been introduced, each aiming to improve classification or monitoring techniques for certain remote sensing projects.

Best available pixel composites (Griffiths, van der Linden, Kuemmerle & Hostert, 2013; White, Wulder, Hobart, Luther, Hermosilla, Griffiths, …, & Guindon, 2014) is a compositing methodology which relies on user generated weights to create a final composite for analysis. The compositing algorithms requires significant user input to decide on the final pixel of the composite. The user can assign different weights to variables such as distance to clouds, sun angle, image acquisition date or atmospheric values within each image. If multiple clear pixels are available to choose from, a scoring mechanism is implemented based on the inputted weights, and the pixel with the highest score will be written into the composite. This methodology has shown to be useful in a number of applications, primarily in forestry and land use and land cover change detection. One downside of this method however is its reliance on user input for weighing the variables. While this may not cause issues for a small area of interest, applications to larger scales, such as country, or continent wide studies may be more difficult due to variability in cloud cover, atmospheric conditions or land cover within the area.

Maximum NDVI is another popular compositing approach for vegetation and agriculture studies (Flood, 2013). NDVI is arguably the most widely implemented vegetation index in remote sensing (Robinson, Allred, Jones, Moreno, Kimball, Naugle … & Richardson, 2017). It has wide applications in agriculture, forestry, and crop type classification to name a few. The compositing methodology behind the Maximum NDVI algorithm is rather straightforward. Once a suitable compositing period has been decided on and the images have been preprocessed for cloud removal and atmospheric correction, NDVI is calculated for each input image. For each pixel in the final composite, the NDVI value in each input image is compared, and the values from the date which observed the highest NDVI value are written to the final composite. This compositing methodology is rather useful in studies which aim to classify or monitor vegetation, however it is not as useful for areas such as change detection in urban areas due to its reliance on NDVI. Some other downsides of this compositing methodology are that the final composite image will likely only include pixels from the spring or summer season, when NDVI values tend to be the highest. Due to the reliance on high NDVI values, this compositing methodology may also cause issues in the differentiation of green areas, such as fallows and forests, as the pixels for both areas will be recorded at their greenest. One highly positive trait of this algorithm however is cloud removal. Clouds tend to have lower NDVI values than the surface beneath them, and as such using a maximum NDVI value for compositing will have a secondary cloud removal effect in cases where the cloud mask failed in preprocessing.

Flood (2013) introduced the idea of creating seasonal composites using the medoid, a multi-dimensional median. In this method, for each pixel, the medoid over the period is selected and written in the final composite. The medoid is an actual observed value within the time series, and aims to maintain the spectral relationship between the spectral bands. This methodology appears to be robust against extreme values, such as residual clouds, or fields which have been harvested early. The outcomes of Flood's study were that this methodology appears to be more representative of the time series than the maximum NDVI composites, and create a smoother looking composite.

A high-dimensional median compositing methodology is introduced by Roberts, Mueller & McIntyre (2017). This relies on calculating the geometric median of each pixel in the time series and "effectively trades a temporal stack of poor quality observations for a single high-quality pixel composite with reduced spatial noise". One particular strength of this methodology is that outliers, such as residue clouds or early harvested fields do not have a large effect on the final composite value. As such, even though cloud masking algorithms tend to miss clouds and cloud edges fairly often, this methodology does not require a perfect input stack, and can deal with some anomalies in the data. Other compositing methodologies exist, relying on median values for each band, or on acquisition times of the images utilized in the study, such as a ‘peak summer’, ‘earliest image’, or ‘latest image’ composite.

1.4 Image Classification

For the purposes of remote sensing, image classification is an automated process of extracting information and classifying pixels from a remotely sensed raster image. The resulting raster from the image classification process can then be used to create thematic maps for use in land cover and land use classifications. There are two types of classification depending on the interaction between the operator and the computer during classification: supervised and unsupervised (Aggarwal 2004).

A supervised classification method is employed in this project. Supervised classification is a method which involves the user selecting sample pixels or objects in an image that represent specific classes and then instructs the image processing software to implement these training points as reference for the classification of all other pixels or objects in the image. These training sites (also known as reference sets or input classes) are positioned based on the confirmed knowledge of the operator, gained either via on site surveys or acquired information from third parties.

A second decision in classification is whether to employ a pixel based or object based classification methodology. In a pixel based classification, each pixel is seen as and assessed as individual from the surrounding pixels. For the purposes of crop type classification this might mean that two pixels may be classified to different classes, even though they are within the same field.

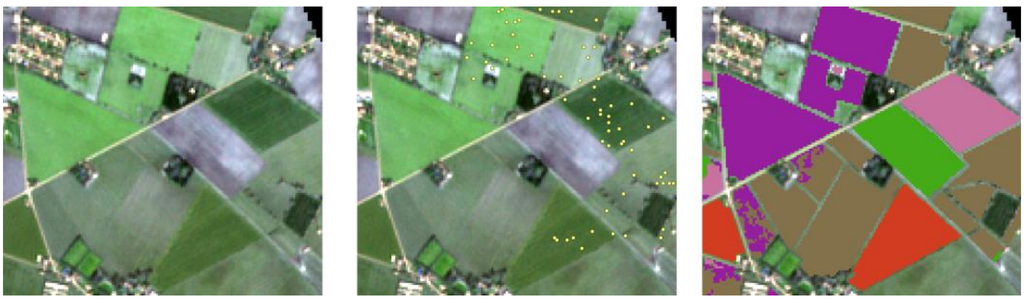

Figure 4 Pixel based classification process – Satellite Image – Training Points – Classified Fields

1.4.1 Crop Type Classification

Land cover and land use classification is a tremendously popular concept within remote sensing. The development of land classification systems such as CORINE, USGS's classification system and other more specific systems such as IPCC's Land-Use Change and Forestry guide have led to an increase in the study of image classification from a remote sensing perspective. However, little emphasis has been placed on developing detailed crop maps (Wardlow, Egbert & Kastens (2007), despite the introduction of various economic processes, sustainable development and agriculture goals, and food security concerns. Only recently, the European Space Agency's Common Agricultural Policy (CAP) has started relying on satellite imagery from crop monitoring and classification.

Most crop type classification research in recent times has focused on the use of classification algorithm, intending to develop better and more accurate deep learning models for crop health monitoring, classification and yield estimation.

Studies in the Ukraine (Kussul, Lavreniuk, Skakun & Shelestov, 2017) and Japan (Sonobe, Tani, Wang, Kobayashi & Shimamura, 2014) have also focused on using synthetic aperture radar (SAR) data for use in crop type classification with good results.

For this project, the focus is on using only optical data from a single sensor (Sentinel 2) for the purposes of crop type classification. The focus of the study is on improving the data used as input for the classification, by creating composite images, instead of the current popular approach of utilizing a single image.

1.4.2 Random Forests

Breiman (2001) introduces random forests as a combination of tree predictors, an ensemble learning method which generates multiple classifiers and clusters their results. The paper concludes that random forests are an accurate and effective tool in classification, performing better than previous bagging or boosting methods.

Belgiu & Dragut (2016) describe the methodology behind random forests and their application to remote sensing classification. The random forest classifier draws a random subset of training samples through a replacement (bagging) approach. This means that certain samples may be selected more than once, while others may be ignored by the classifier. Breiman (2001) outlines that about two thirds of the samples are used to train the trees, and the remainder are used to test how the model performs as a cross-validation technique.

The classifier requires two user inputs – the number of trees to be generated, and the number of variables to be selected and tested for the best split when growing the tree. Due to the law of large numbers, the classifier is less sensitive to the number of trees, and a large enough number, such as 500, appears to be large enough for remote sensing applications (Lawrence, Wood & Sheley, 2006). The number of variables to be selected is usually set to the square root of the number of input variables (Gislason, Benediktsson & Sveinsson, 2006).Figure 5 below describes the random forest classifier as discussed by Belgiu & Dragut (2016)

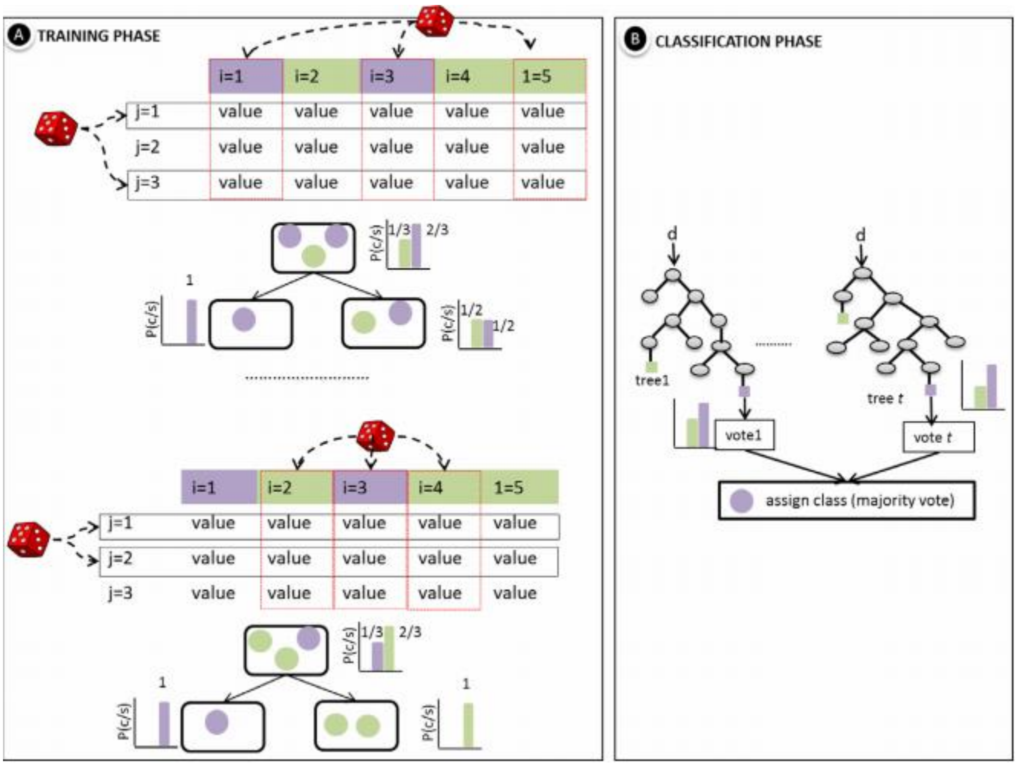

Figure 5 Training and classification phases of random forest classifier: i = samples, j = variables, p = probability, c = class, s = data, t = number of trees, d = new data to be classified, value = the different values that the variable j can have. From "Random forest in remote sensing: A review of applications and future directions" by M. Belgiu & L. Dragut, 2016, ISPRS Journal of Photogammetry and Remote Sensing, 114, p. 26.

Random forests have been shown to perform very well when compared to other classification algorithms including support vector machines, maximum likelihood, and neural networks. Some advantages of using random forests for classification are that overfitting to the training data is not a problem, and that inaccuracies in the training data can usually be ignored when using a large enough number of trees (Breiman, 2001). This makes random forests an excellent choice as a classification algorithm for this project.

1.4.3 Accuracy Assessment

Crop type classification maps are a type of thematic map, similar to land cover and land use maps, and other thematic maps derived from remote sensing data. As these maps are created based on a classification algorithm which relies on only a small subset of training data, an accuracy assessment of the classification results must be performed in order to ensure that the predicted results are accurate and can be relied upon.

This process of results analysis and evaluation is critical to the quality of the product. A comparison of the ground truth data and the classified data has to be carried out in order to find the accuracy of the results. While many methods have been described and discussed for the purposes of remote sensing (Aronoff, 1982, 1985; Koukoulas & Blackburn, 2001), Foody (2002) mentions that the mostly widely used and adopted methodology is the confusion matrix, which is currently at the core of accuracy assessment literature.

The confusion matrix makes it easily evident to the researcher about discrepancies between the training and testing data, as well as easy to spot and interpret differences within classes. One of the more popular measures of classification accuracy derived from the confusion matrix is overall accuracy, or the percentage of cases correctly classified. Two other popular measures are the producer’s and the user’s accuracies.

Overall accuracy is the basic measure revealing how many classification attempts were correct in total, a simple calculation showing the proportion between the correctly classified points of all classes and the total number of test points.

Producer’s accuracy defines the proportion between the number of points belonging to a specific class X which were correctly classified as points of class X, excluding the points which the classifier agent omitted - false negatives, and the total number of ground true points under class X. This shows the overall accuracy of the classifier agent in context of each class separately.

User’s accuracy defines the proportion between the number of points which were correctly classified as belonging to class X and the total number of points classified as class X – including false positives – points of classes Y and Z incorrectly classified as class X points.

1.5 Problem Statement

Satellite imagery and satellite image compositing is widely used in a variety of fields including forestry (Potapov, Turubanova & Hansen, 2011), oceanography (Breaker, Armstrong & Endris, 2010), and coastal monitoring (Sagar, Phillips, Bala, Roberts & Lymburner, 2018). The vast archive of available data is increasing at a more rapid than ever pace, with both publicly, and privately funded satellites now providing near daily coverage of all areas around the globe. With the exponentially growing data availability, there is now a strong need for effective and efficient processes to interpret, analyse and derive valuable information from this data.

The opening up of the Landsat archive in 2008 (Woodcock, Allen, Anderson, Belward, Bindshadler, Cohen … & Nemani, 2008) provided scientists with free access to millions of images acquired by the Landsat mission (Wulder, White, Loveland, Woodcock, Belward, Cohen … & Roy, 2016). Since then, a large number of tools for processing remote sensing images from various satellites have been further developed, including tools for cloud detection (Hagolle, Huc, Pascual & Dedieu, 2010; Zhu & Woodcock, 2012; Brockmann, Paperin, Danne & Ruescas, 2013), atmospheric correction (Gao, et al 2009), and image compositing (Griffiths et al., 2013).

This study explores how these methods, including cloud masking, atmospheric correction, and image compositing, can increase the data availability for crop monitoring from a single image, to an entire month or growing season. This can be particularly beneficial for areas with high cloud coverage where a single image classification approach will not work for an entire area of interest due to interference from clouds.

1.6 Research Objectives

Several studies have proposed different compositing methodologies for satellite imagery. The Landsat archive, due to its long temporal coverage has been the primary interest, with focus on developing composites for various areas, including the contiguous United States (Roy, Ju, Kline, Scaramuzza, Kovalskyy, Hansen … & Zhang, 2010), the Carpathian region (Griffiths, Kuemmerle, Baumann, Radeloff, Abrudan, Lieskovsky ... & Hostert, 2014), and Canadian forests (White, Wulder, Hobart, Luther, Hermosilla, Griffiths … & Guindon, 2014). This study however investigates the possibility of using data from Sentinel 2, a much newer sensor, which provides better temporal and spatial resolution than Landsat 8.

The focus of the study is optical data, provided by the Sentinel 2 pair of satellites, Sentinel 2A and Sentinel 2B, launched by the European Space Agency under the Copernicus program on 23rd June 2015, and 7th March 2017, respectively (European Space Agency, 2018).

The first part of this project aims to provide a compositing workflow focused on creating high quality, analysis ready image composites, applicable on a tile level for Sentinel 2 imagery.

The second part of the study focuses on investigating the suitability of monthly satellite image composites for the purpose of crop type classification.

In order to achieve these two aims, the following research questions are defined:

- Which preprocessing techniques are the most important in the compositing process?
- What different compositing methodologies can be applied to Sentinel 2 data on a tile level?
- Which remote sensing indexes are most appropriate for crop type classification?
- How does the accuracy of crop type classification performed on image composites compare to that of a classic clearest image approach?

2 Data and Materials

This section of the report introduce the study area where the case study has been prepared, as well as the satellite images and agricultural data acquired. The hardware and software utilized during this project are also presented and described.

2.1 Study Area

The study area for this project lies in south western Denmark, with its center point located at 9° 52’ 3.57” E, 55° 26’ 57.70” N. The area covers parts of Jutland and the island of Fyn and it covers the entirety of Sentinel 2 acquisition tile 32UNG.

Figure 6 Study area outline shown within Denmark

This was deemed to be a suitable study area for this project as a large portion is used for growing varied crops, primarily Spring Barley, Winter Wheat and Corn. According to the CORINE Land Cover Classification System, approximately 57.5% of the study area is classified as ‘Agricultural Areas’. The further breakdown is as follows:

Table 4 CORINE level 3 classification for study area

CORINE Land Code	CORINE Level 3 Classification	Area (ha)	Area (%)
211	Non-irrigated Arable Land	598,771	49.67%
222	Fruit Trees and Berry Plantations	1,968	0.16%
231	Pastures	5,928	0.49%
242	Complex Cultivation Patterns	11,674	0.97%
243	Land principally occupied by agriculture, with significant areas of natural vegetation	74,824	6.21%
	Total Agricultural Area	693,165	57.5%

Additionally, the area had no days with 0% cloud cover over the study period of spring and summer 2017, further showing the need for image compositing algorithms in order to make full use of remotely sensed data.

This study provides three different compositing algorithms to achieve the image composites, namely maximum NDVI, geometric median and medoid. The images are preprocessed using open source tools for cloud masking and atmospheric correction, however the accuracies and performance of these processes are not evaluated, as this was outside the scope of this study. The compositing methodologies are implemented in Python and make use of various open source libraries.

The resulting image composites are then used for classifying various crop types in the study area. Random Forests are used as the classifier of choice and the classification results from each composite image, as well as a ‘clearest, most cloud free image’ are compared based on their overall accuracy.

2.2 Satellite Image Data

For this project, only Sentinel 2 optical data has been used in order to perform the compositing and the classification. The data for this part of the project has been downloaded from the European Space Agency's SciHub download center, available online at https://scihub.copernicus.eu/dhus/#/home

The data for this study area covers the entire Sentinel scene 32UNG and the time period for this research includes the spring (March, April, May) and summer (June, July, August) of 2017. The data is easily accessible via SciHub and can be downloaded in ESA's .SAFE format on a tile basis. The .SAFE folder structure includes all 10m, 20m, and 60m bands captured by Sentinel 2 as well as any required metadata. The scenes are downloaded as a Level-1C product. This product consists of 100 x 100km tiles in UTM Zone 32N projection. The level 1C products are resampled and orthorectified and the per pixel radiometric measurements are provided in top of atmosphere reflectances. (ESA)

The data downloaded from SciHub includes 47 different scenes acquire between March 1st 2017 and August 31st 2017. Due to the launch of Sentinel 2B, a twin satellite to Sentinel 2A, we can see a large increase in the number of scenes available in July and August. The two satellites are in polar opposite sun synchronous orbits and aim to provide a new image of each every 2.5 days on average.

The temporal distribution of the acquisitions is as follows:

Table 5 Temporal distribution of images

Month	Number of Scenes	Acquisition Dates
March	7	1st, 4th, 11th, 14th, 21st, 24th, 31st
April	6	3rd, 10th, 13th, 20th, 23rd, 30th
May	5	3rd, 10th, 13th, 23rd, 30th
June	6	2n, 9th, 12th, 19th, 22nd, 29th
July	12	2nd, 4th, 7th, 9th, 12th, 14th, 17th, 19th, 22nd, 24th, 27th
August	11	1st, 3rd, 6th, 8th, 11th, 13th, 18th, 21st, 23rd, 26th, 31st

Sentinel 2 captures data in the following spectral bands, with spatial resolutions ranging from 10m to 60m:

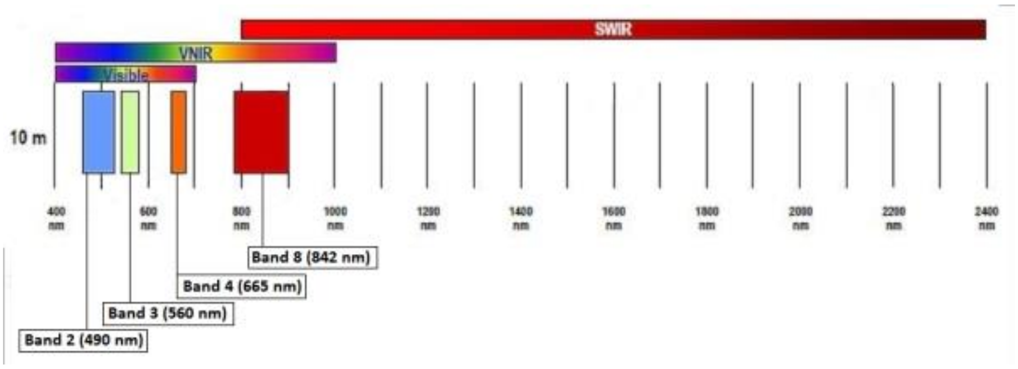

Figure 7 Sentinel 2 10m bands. From European Space Agency. (2018). Sentinel-2 10m spatial resolution bands [Image]. Retrieved from https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial

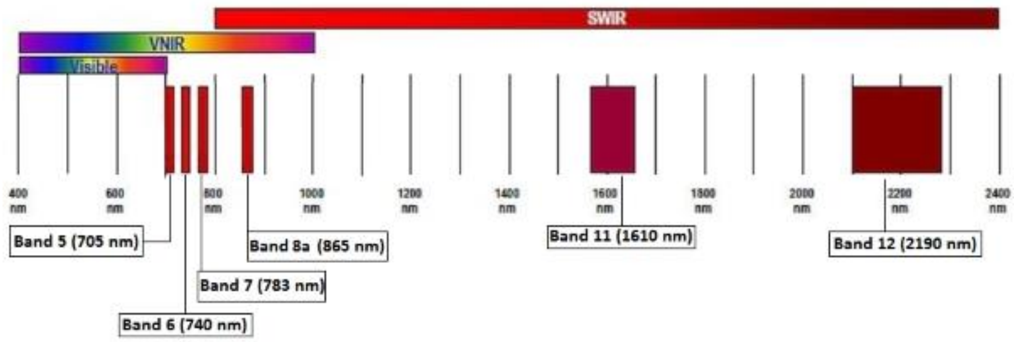

Figure 8 Sentinel 2 20m bands. From European Space Agency. (2018). Sentinel-2 20m spatial resolution bands [Image]. Retrieved from https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial

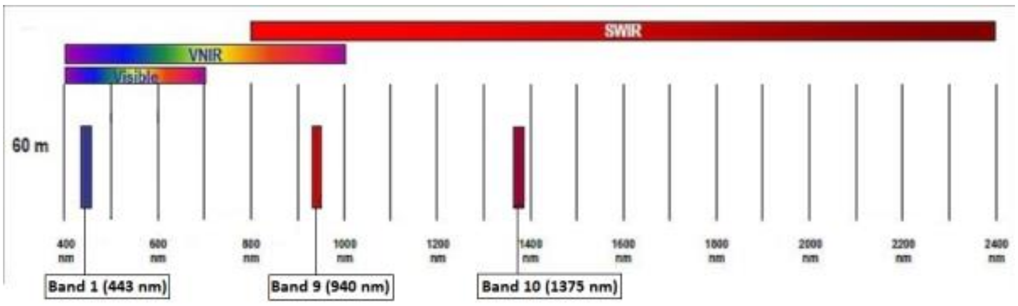

Figure 9 Sentinel 2 60m bands. From European Space Agency. (2018). Sentinel-2 60m spatial resolution bands [Image]. Retrieved from https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial

2.3 Agricultural Data

The Danish AgriFish Agency provides a nationwide, annually updated shapefile consisting of digitized agricultural fields. This shapefile contains the annual reporting of farmers making use of subsidies provided by the European Union. The shapefile contains the fields, field sizes and the crops being grown on the field.

The data has been clipped to include only the study area and contains a total of 87,863 fields.

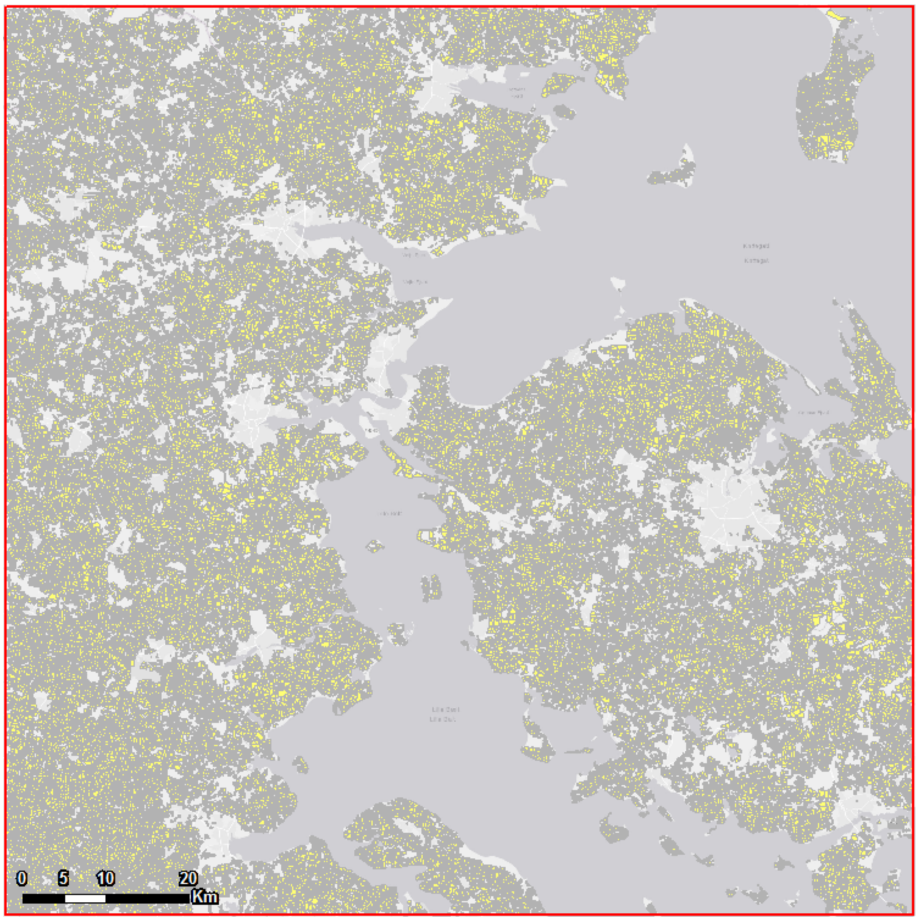

Figure 10 Agricultural fields shown within the study area

The latest shapefile, generated in 2017, has been used throughout this study. While the accuracy of the shapefile cannot be verified, it is assumed to be a fairly comprehensive and accurate representation of crops grown in the study area.

2.4 Hardware and Software

All processing for this study has been performed on a Lenovo Thinkpad laptop with the following specifications:

Table 6 Hardware specifications

Operating System	Windows 10 Enterprise 64-bit
Processor	Intel Core i7-3610QM @ 2.30GHz
Number of Logical Processors	8
Memory	16384MB RAM

The GIS software utilized for this study includes QGIS 2.18.9 and ArcMap 10.5. These software packages were utilized for some data filtering and creation of graphics and data visualizations.

The majority of data filtering, image compositing and image preprocessing was done in Python 3.6. Separate Anaconda environments were created for the preprocessing and image compositing steps.

The random forest classification was performed in R utilizing the RandomForest package.

The accuracy assessment and error matrix creation was performed within QGIS, utilizing the GRASS Plugin r.kappa.

3 Methodology

This study provides three different compositing algorithms to achieve the image composites, namely maximum NDVI, geometric median and medoid. The images are preprocessed using open source tools for cloud masking and atmospheric correction, however the accuracies and performance of these processes are not evaluated, as this was outside the scope of this study. The compositing methodologies are implemented in Python and make use of various open source libraries.

While images for the entire spring and summer seasons of 2017 have been downloaded, for this project, composites have been created only for the months of March and July, as they were deemed to provide the most detailed information in regards to crop growth. Additionally, some images were discarded due to having too many clouds or for not capturing 100% of the study area. The final composites were created using four acquisitions from March and five acquisitions from July.

After the composites for each month have been created, four indexes were calculated on each resulting composite and added to stack of ten Sentinel 2 bands creating a 14 band Geotiff for each month.

In order to prepare the images for classification, a 28 band stack was then created by combining the March and July images.

The resulting image composites are then used for classifying various crop types in the study area. Random forests are used as the classifier of choice and the classification results from each composite image, as well as a ‘clearest, most cloud free image’ are compared based on their overall accuracy.

3.1 Compositing

This section of the report provides details of the compositing methodologies utilized in this study. The data, materials, implementation, and results of this part of the study are then presented. The following diagram explains the process involved in this part of the report, focused on creating the composite images.

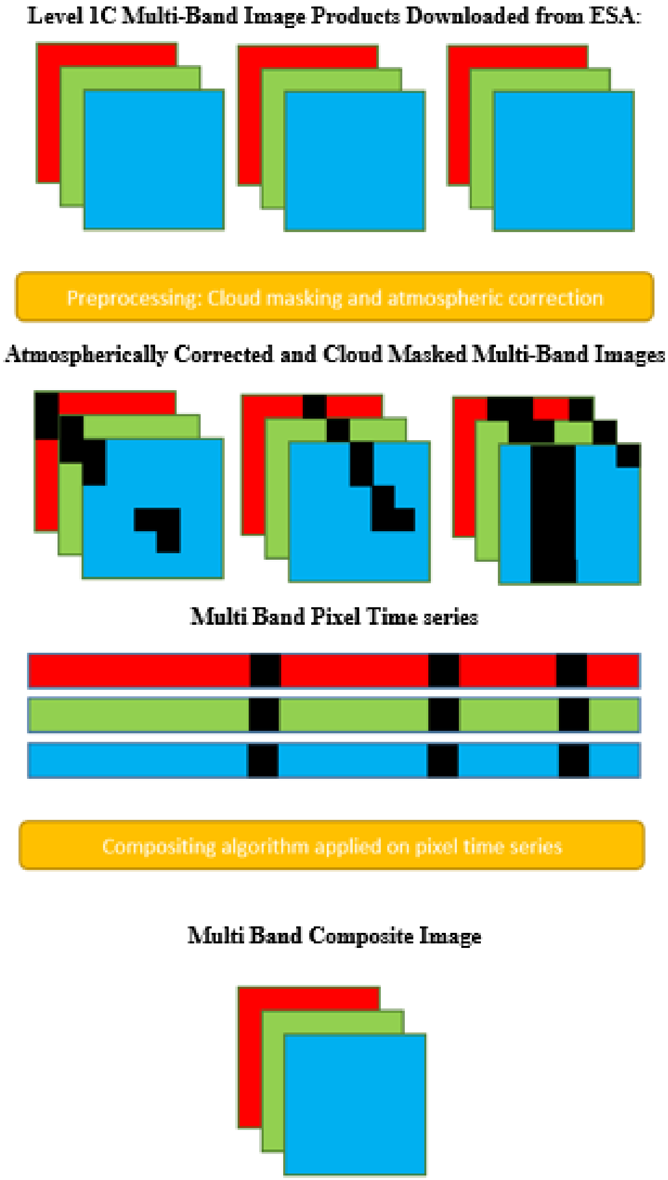

Figure 11 Compositıng process

3.1.1 Preprocessing

In order to obtain the most relevant and accurate information from the satellite images, a number of preprocessing techniques have been implemented in this study.

Cloud Masking

The cloud removal process in the preprocessing stages of this project uses a combination of three state of the art cloud masks – fmask (Zhu, Wang & Woodcock, 2015), Sen2Cor (Muller-Wilm, Louis, Richter, Gascon & Niezette, 2013), and IDEPIX (Brockmann et al, 2013). The combination of these cloud masks creates a more aggressive cloud mask for this project ensuring near full removal of cloud objects.

All scenes acquired in the data acquisition step have been cloud masked using the combined mask. The combined mask also includes a 20pixel buffer around cloud objects in order to ensure full removal of clouds. The cloud mask is written in Python and developed as a command line utility making it easy for users to process multiple scenes. The following command line can be used to generate a cloud mask for a single Sentinel 2 scene:

```
cloud_mask s2 –i C:/scene.SAFE –o cloud_mask.tif –s 10 --cm_buffer 20
```

In the above command the creation options are as follows:

S2 – Indicates this is a Sentinel 2 scene

-i – Path input to the ESA downloaded Level 1C .SAFE folder

-o – Path output for the created cloud mask

-s – Processing pixel size (10m in the case of this study)

-cm_buffer – Buffer to be added to the resulting cloud masks (in meters)

The resulting cloud masks for March are as follows:

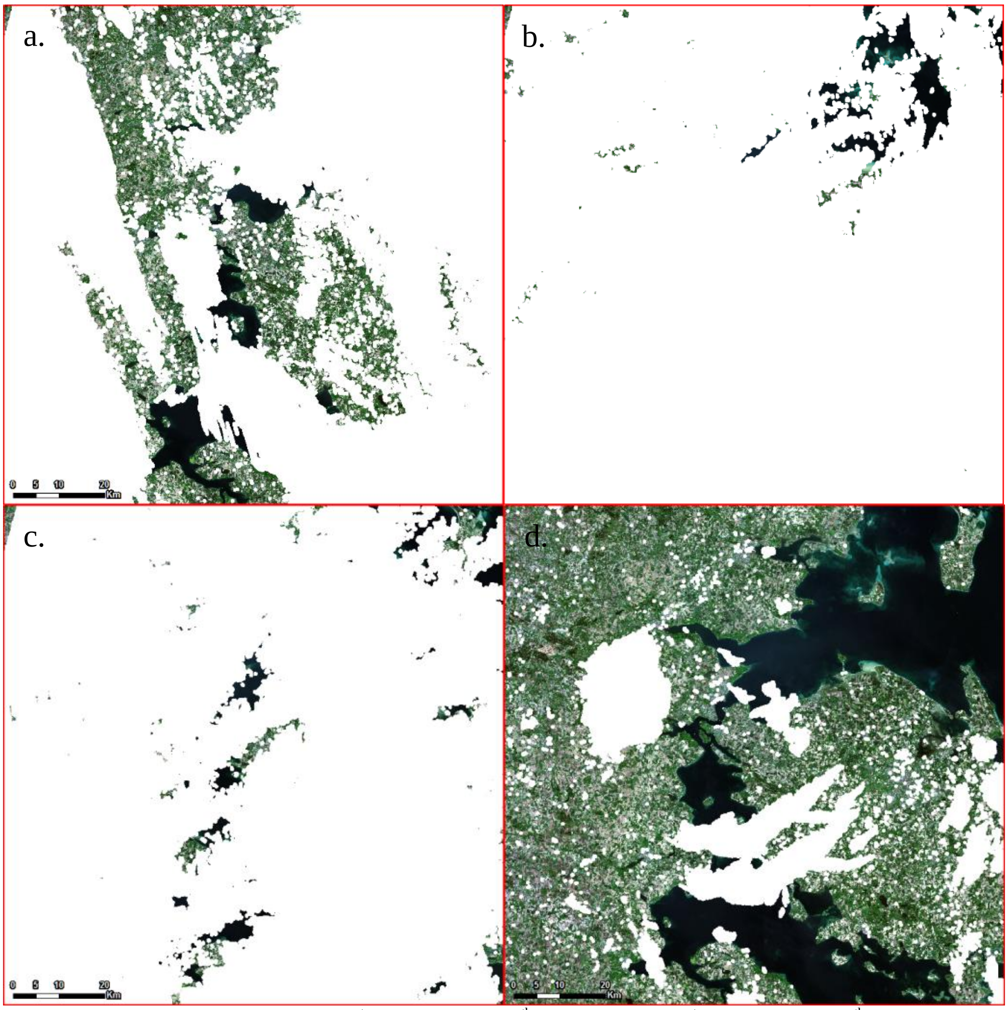

Figure 12 March cloud masks a. March 1st 2017 b. March 11th 2017 c. March 21st 2017 d. March 24th 2017

The resulting cloud masks for July are as follows:

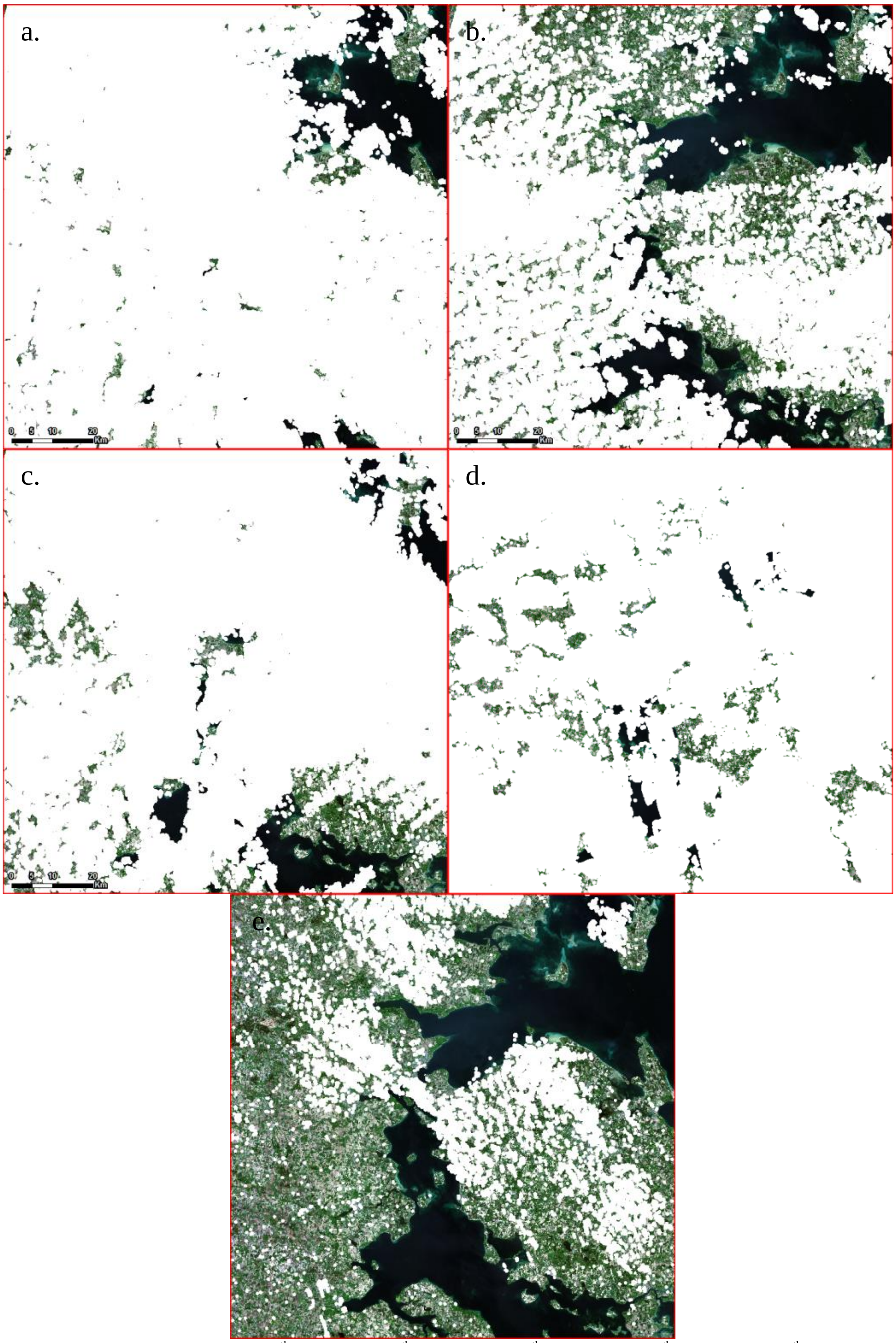

Figure 13 July cloud masks a. July 4th 2017 b. July 9th 2017 c. July 14th 2017 d. July 17th 2017 e. July 19th 2017

The following figure shows the number of clear observations for each pixel in March:

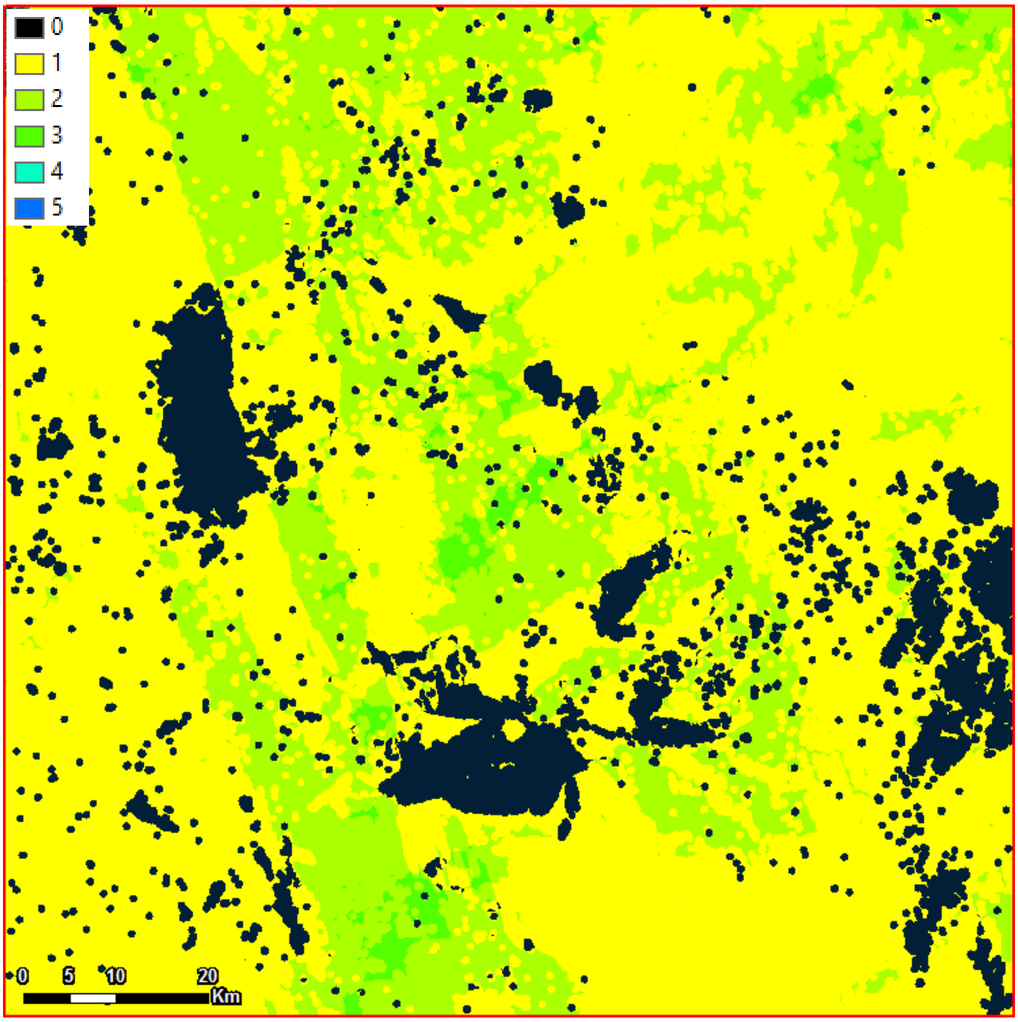

Figure 14 March data availability

The following figure shows the number of clear observations for each pixel in July:

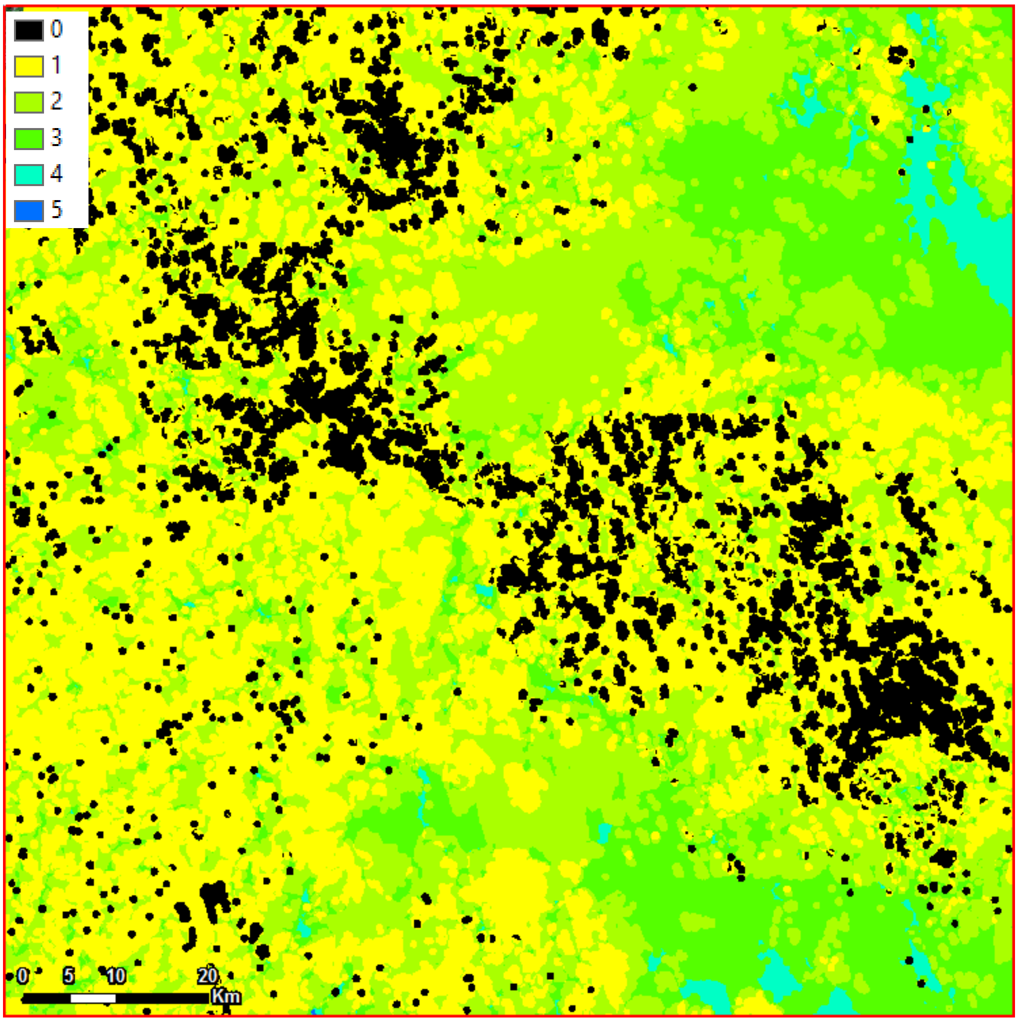

Figure 15 July data availability

Due to pixel removal from the cloud masks, data availability decreased drastically. The data available will need to be further filtered at a later stage to be prepared for input into the random forest classification algorithm in order to create a mask consisting of pixels which have clear observations available in both March and July.

Atmospheric Correction

Sen2Cor is a Level 2A processing toolbox for Sentinel 2 imagery over land and was developed by the European Space Agency (Muller-Wilm et al, 2013). For atmospheric correction, this processor integrates image observations with look-up tables from the LibRadtran model to remove atmospheric effects (Martins et al, 2017). Version 2.5.5 of the processor, released March 23, 2018, was utilized throughout this project.

The scenes were processed using the command line utility provided by the Sen2Cor package. The package provides a simple interface for users to process multiple scenes rapidly. Atmospheric correction on a single scene took approximately 30 minutes. The level 2A atmospheric correction processor can be run in the command line:

```
L2A_Process path/to/SAFE/folder
```

The following configuration parameters for atmospheric correction have been set using the GIPP file for user inputs.

Table 7 Atmospheric correction user inputs

Option	Value
Aerosol Type	Rural
Mid-Latitude	Auto
Ozone Content	Best Guess from Metadata
Water Vapour Correction	True
Water Vapour WaterMask	True
Cirrus Correction	True
Water Vapour Threshold Cirrus	0.25
BRDF Correction	False

No BRDF correction has been performed as the terrain in the study area is mostly flat.

3.1.2 Compositing Algorithms

Throughout this project, three different image compositing methodologies were implemented and compared, using three different algorithms, maximum NDVI, medoid, and geometric median. The algorithms were written, implemented and tested in Python 3.6.

Prior to compositing, a single image stack consisting of ten bands is created for each input date. The 20m resolution bands are first resampled to 10m resolution.

The stack is a ten band GeoTiff consisting of all 10m and resampled 20m resolution bands acquired by Sentinel 2 for each date.

Maximum NDVI

Maximum NDVI is a fairly common algorithm for compositing satellite images. In this approach, NDVI is calculated for each individual image. In cases where more than one clear pixel is available, the values corresponding to the date which had the highest NDVI value are written into the final composite. This process is executed in Python using the rasterio and numpy libraries. The resulting composite consist of actually observed values, however neighbouring pixels may be from different acquisition dates, thus creating some artefacts in the final composites.

This compositing process is the least computationally intensive of all three methodologies evaluated in this study. The compositing process took approximately five minutes for each monthly composite created.

Medoid

The medoid method is a multi-dimensional median method for compositing. The medoid is a “measure of center” in a multi-dimensional set of points, similar to the median in a unidimensional space (Flood, 2013).

In the context of remote sensing and image compositing from a time series of observations, the multiple dimensions working in this process are the different bands of the image and the different acquisition dates for every captured pixel.

A calculation formula for the medoid is:

$$x_{medoid} = argmin_{y \in \{x_1, x_2, \cdots, x_n\}} \sum_{i=1}^{n} d(y, x_i)$$

Equation 1 Medoid

This measure of center is defined as the point which minimizes the sum of the distances to all the points in the dataset, with the additional constraint that the selected point has to be an observed value (Flood, 2013)

The medoid compositing algorithm was implemented in this project in Python using the rasterio, numpy and hdmedians (Roberts, 2017) packages. Additionally, the script was modified to be run in parallel using concurrent.futures in order to speed up the process. This compositing process was the 2nd most computationally intensive compositing algorithm, with composites for each month taking approximately 20 minutes to create.

Geometric Median

The geometric median is a similar concept to that of the medoid, a multi-dimensional median method for compositing. The main difference between the two methods is that the geometric median is not necessarily an observed value. In order to calculate the geometric median, an artificial value, which minimizes the distances between points in all dimensions is created. This artificial value is then written into the final composite.

The calculation formula for the geometric median is:

$$\underset{y \in \mathbb{R}^n}{argmin} \sum_{i=1}^{m} ||x_i - y||_2$$

Equation 2 Geometric median

The strength of both the medoid and geometric median compositing methodologies is that they maintain the radiometric relationship between the bands in the final composite. As such, we can easily calculate accurate and relevant indexes on the composite, despite the values being artificial, in the case of the geometric median.

This compositing algorithm was implemented in Python and is based on the hdmedians package (Roberts, 2017). The same parallel processing methods were used as in the medoid implementation in order to reduce processing time. This was the most computationally intensive process, with composites taking approximately 90 minutes to create for each month on the available hardware.

3.2 Crop Type Classification

The following part of the report describes the how the agricultural data and materials used in this study were implemented, with a focus on the classification process and a comparison of accuracy of the three resulting composite against each other, and against a 'clearest, most cloud free' image classification.

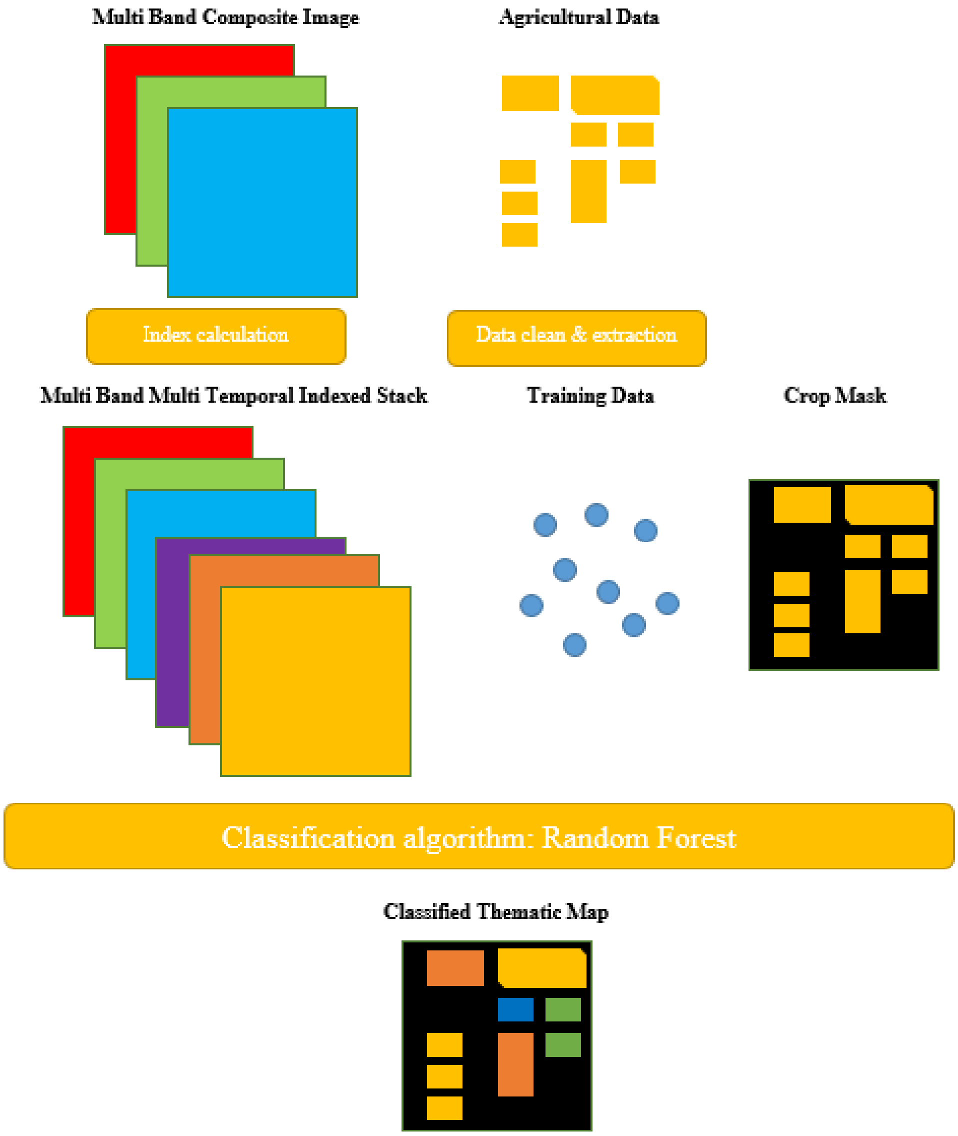

Figure 16 Classification process

3.2.1 Image Preprocessing

The random forest classification has been performed using the R package 'randomForest'. The package takes multiple band Geotiffs as input for the classification. In order to prepare the composite images for input into the classification package, a number of preprocessing steps, such as index calculations and band stacking have been performed in this stage.

In order to increase the accuracy of the classification a number of spectral indexes have been calculated. These indexes have been shown in previous research to improve the classification accuracy. The following indexes have been calculated for use in this project:

Normalized Difference Vegetation Index (NDVI)

NDVI is probably one of the most important indexes in the monitoring and classification of land and agricultural areas (Robinson et al., 2017). The index is calculated using the following formula:

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)}$$

Equation 3 NDVI

For Sentinel 2 NDVI is calculated using bands 8 and 4, with the central wavelengths recorded at 842nm and 665nm, respectively.

Red Edge NDI

This index is particularly important for the calculation of green leaf area index, and chlorophyll content (Delegido, Verrelst, Alonso & Moreno, 2011)

The index is calculated using the following formula:

$$Red\ Edge\ NDI = \frac{(NIR - Red\ Edge)}{(NIR + Red\ Edge)}$$

Equation 4 Red Edge NDI

For Sentinel 2 Red Edge NDI is calculated using bands 8a and 6, with the central wavelengths recorded at 865nm and 740nm, respectively.

Normalized Difference Infrared Index (NDII)

This index uses a normalized difference formulation. It is a reflectance measurement sensitive to changes in water content of plant canopies.

The index is calculated using the following formula:

$$NDII = \frac{(NIR - SWIR)}{(NIR + SWIR)}$$

Equation 5 NDII

For Sentinel 2 NDII is calculated using bands 8a and 11, with the central wavelengths recorded at 865nm and 1610nm, respectively.

Normalized Difference Water Index (NDWI)

Introduced by Gao (1996) this index is used in remote sensing of vegetation liquid water.

The index is calculated using the following formula:

$$NDWI = \frac{(NIR - SWIR)}{(NIR + SWIR)}$$

Equation 6 NDWI

For Sentinel 2 NDWI is calculated using bands 8a and 12, with the central wavelengths recorded at 865nm and 2190nm, respectively.

After calculation, the indexes are added to the image stack, further increasing the number of bands in the GeoTiff. The resulting monthly composites then consist of the ten 10m and resampled 20m Sentinel 2 bands, plus the four calculated indexes, resulting in a 14 band stack for each monthly composite.

The two monthly composites are then stacked again, but this time only pixels which have clear observations in both March and July are written into the final stack. The final stack thus consists of a total of 28 bands. Pixels which had a clear observation in only one of the two months are removed from the analysis, in order to improve classification accuracy.

3.2.2 Agricultural Data Preprocessing

The raw agricultural data downloaded from the Danish AgriFish Agency contains multiple similar crop types (e.g. different types of grass) which will need to be merged into one larger class in order to facilitate classification. After clipping the data to include only the fields within the project's study area, a filtering process was put in place in order to remove fields that were too narrow or too small to rasterize for classification. This process involved applying a negative buffer to the fields, and then removing all fields with an area of less than 100m^2 (one Sentinel 2 pixel size) and all fields narrower than 5m.

Once the fields have been reduced to include only fields which can be reliably classified using the 10m resolution imagery, the fields were then reclassified into ten classes for easier classification:

Table 8 Reclassified agricultural data

Class	Number of fields	Area (based on number of pixels)
Grass	18,295	365 ha
Spring Barley	17,220	782 ha
Winter Wheat	16,816	1104 ha
Corn	5,440	295 ha
Winter Rapeseed	4,399	327 ha
Winter Barley	4,371	262 ha
Winter Rye	3,647	166 ha
Spring Oats	2,228	93 ha
Potatoes	1,485	66 ha
Other	5,205	292 ha

The 'Other' class consists of all other crop types with less than 1,000 individual fields.

3.2.3 Training and Testing Data Selection

Selection of accurate, representative and spatially distributed training data is paramount to the performance of the classification algorithm. For the purposes of this project, 25,000 training pixels have been selected to use as training data for the classification. These pixels were created using a stratified random sampling technique, ensuring that the training number of pixels is representative of the total number of fields for each class.

In order to ensure a fair comparison of composites to the clearest available image however, the training has only been selected from pixels which have had a minimum of two clear observations in both March and July, as show in the map below.

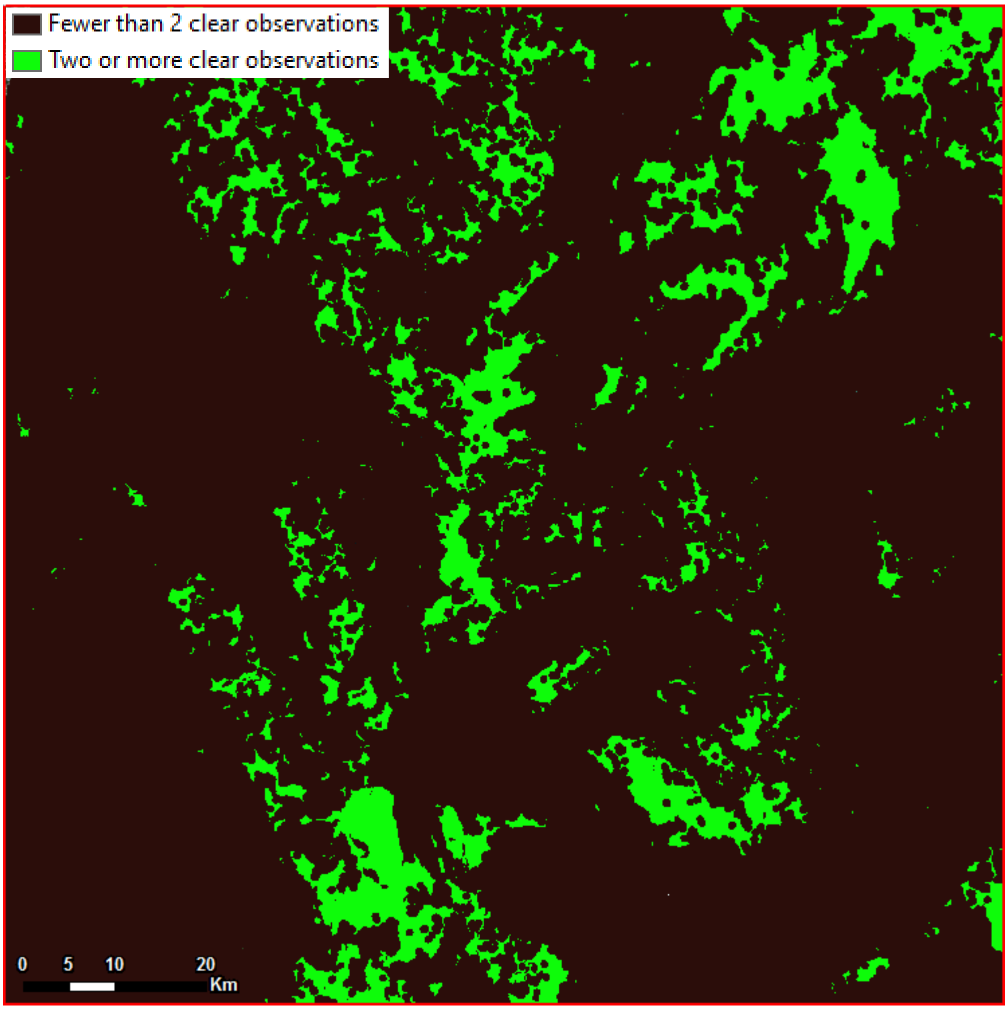

Figure 17 Pixels with two or more clear observations in March and July

The map below shows the point distribution of the training pixels:

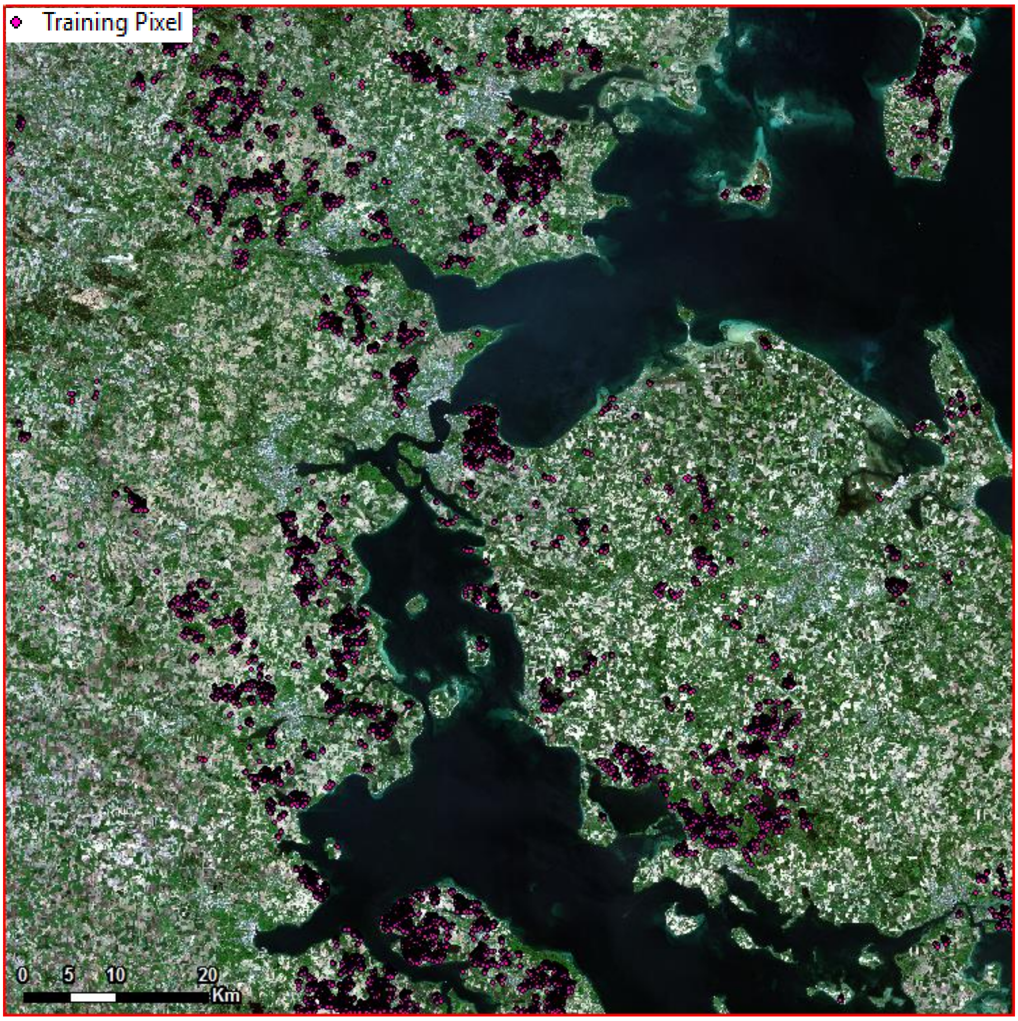

Figure 18 Spatial distribution of training pixels

In order to ensure that both composite and single values were classified correctly, the selection of the testing pixels within the image was less stringent. The locational constraint on testing pixels location is that at least one clear observation must be available in both March and July for the clearest image in each of those months. The second constraint was that testing points must not be selected from pixels which were already used for training the model.

This testing point selection process ensures that all composites and the clearest image approach are treated evenly and ensures a fair comparison between all classification results. A total of 24,966 testing pixels were generated. As was the case with the training pixels, the testing pixels were generated using a stratified random sampling method.

The map below shows the spatial distribution of the testing pixels:

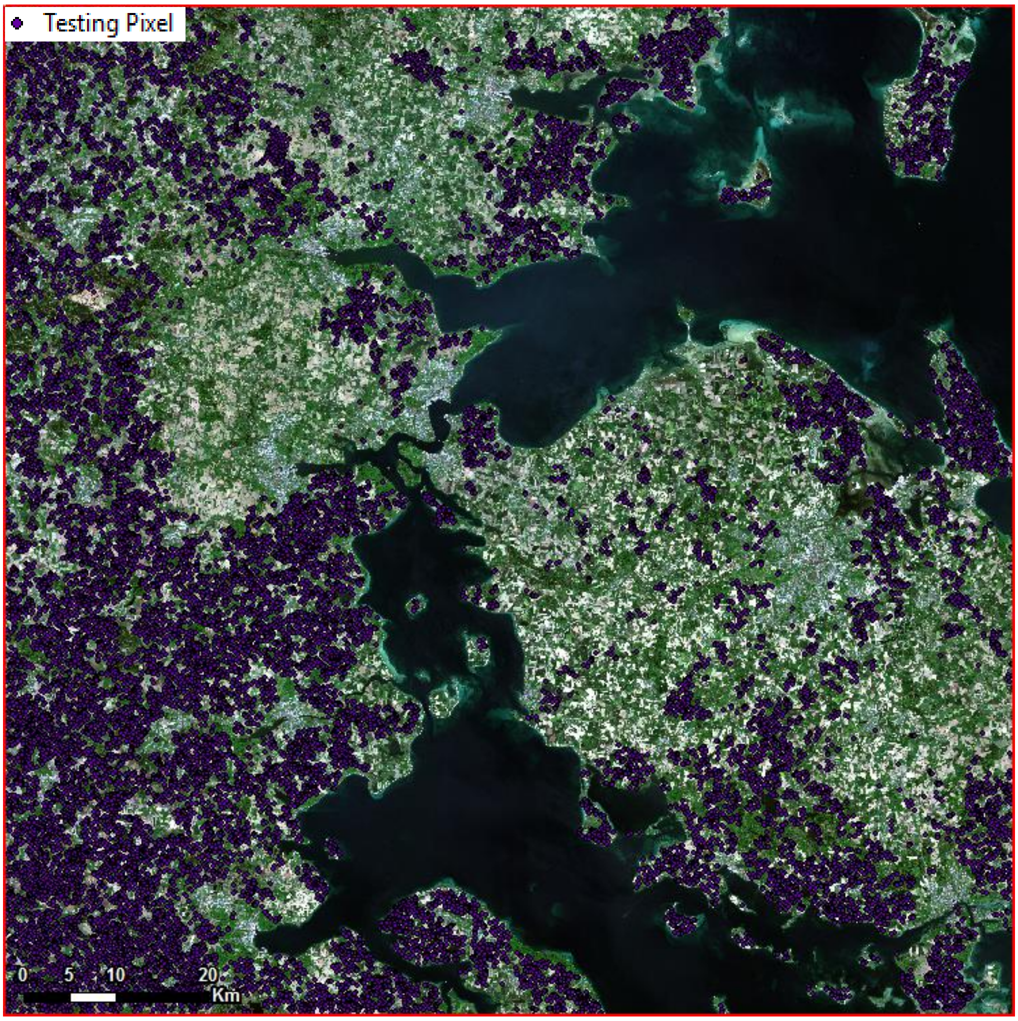

Figure 19 Spatial distribution of testing pixels

3.2.4 Classification Algorithm

The classification has been performed in R using the RandomForest algorithm. This algorithm takes the image stack, training data, and a crop mask as input for the classification. The crop mask indicates which pixels in the image will be classified, and which will be ignored by the classification algorithm. The crop mask was generated to include all fields, where all the pixels within a given field contain at least one clear observation in both the March clearest image and the July clearest image. This filtering of fields was done in order to allow the algorithm to classify only data which has reliable values for the field in every composite and clear image compared in this project.

The map below shows the crop mask utilized in this classification:

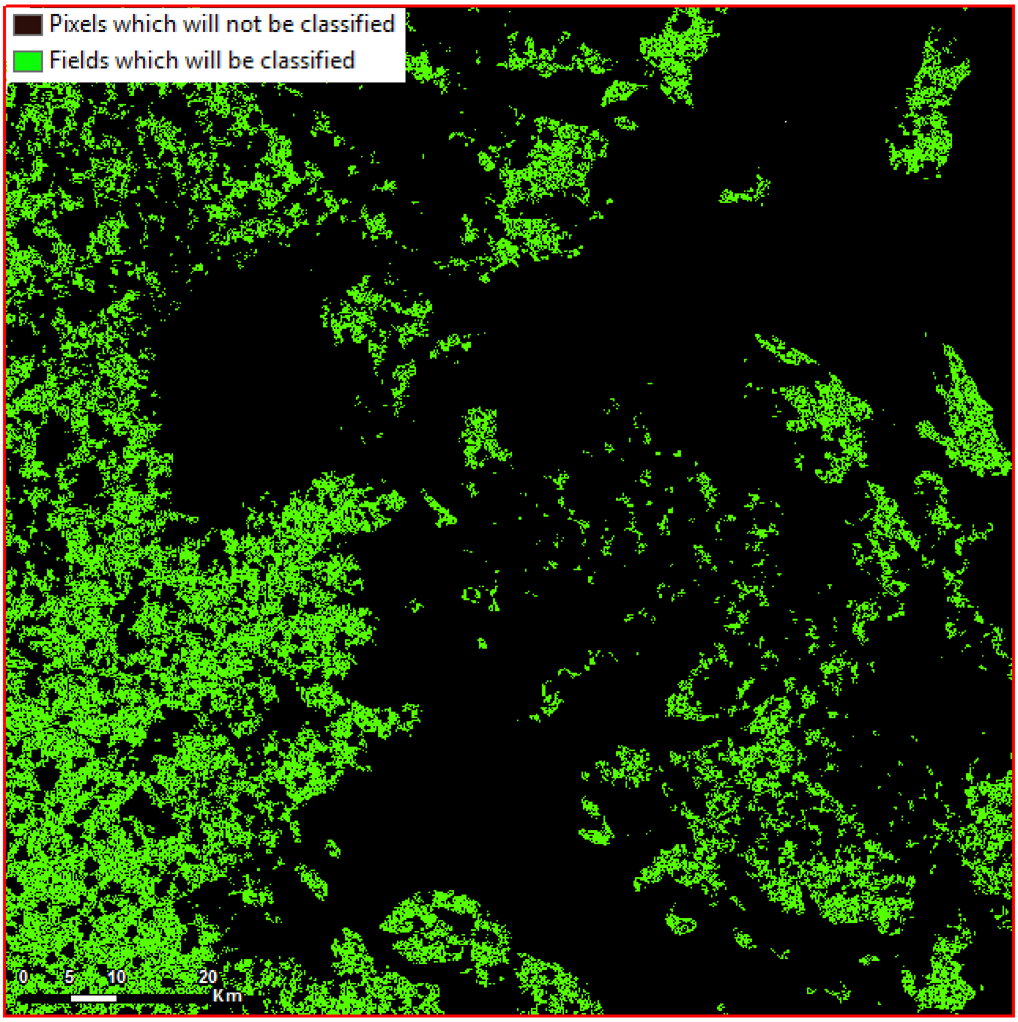

Figure 20 Crop mask

3.2.5 Classification Postprocessing

The outputted thematic map from the classification algorithm has been further postprocessed in order to eliminate any stray pixels and create a smoother classification map. This process was implemented in QGIS using the GDAL sieving tool. The tool creates a user defined minimum mapping unit of eight pixels. This ensures that the minimum number of adjoining pixels must be at least eight, otherwise the pixels are reclassified based on surrounding pixels. This tool eliminates minor errors of stray 'crops' within fields.

4 Results

The following sections will describe and display the results of all the compositing algorithms and the final crop classification maps generated from this project. The final subsection will focus on providing an accuracy assessment for each classification performed during this study.

4.1 Compositing Results

The main advantage of creating image composites is increased data availability. In the case of this study, the composite images provided an additional 10,675 fields for classification. This is an increase of 23.5% over the clearest image classification approach.

This percentage increase in data availability includes only data relevant to this study, i.e. agricultural fields. The overall increase in data availability is even higher when urban areas, forests and water are included, further showing the advantages of using image composites.

The figures beginning overleaf show the final, cloud masked, clearest images and composite images used in this study. Clouded areas have been removed from the image and are displayed as blacked out areas.

4.1.1 Clearest Image Approach

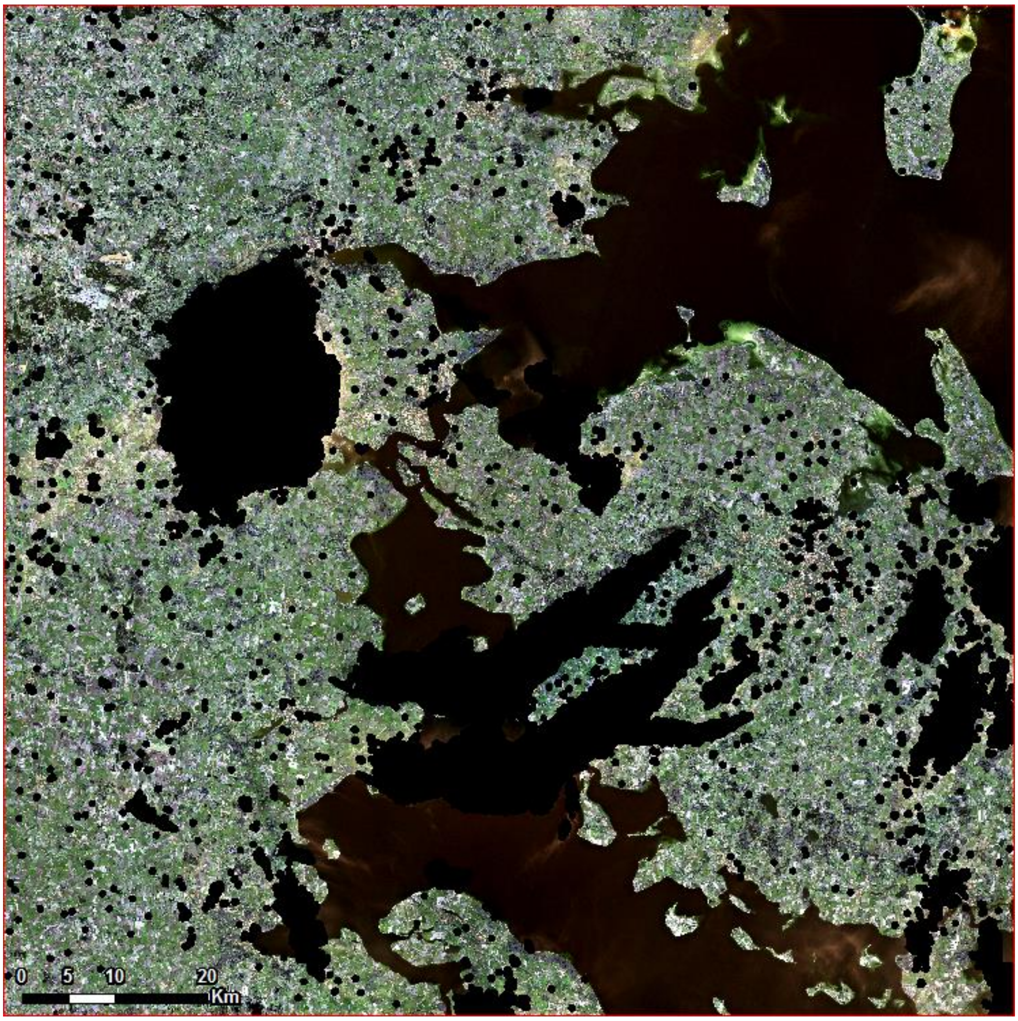

Figure 21 Clearest, least cloudy image in March (acquired March 24 2017)

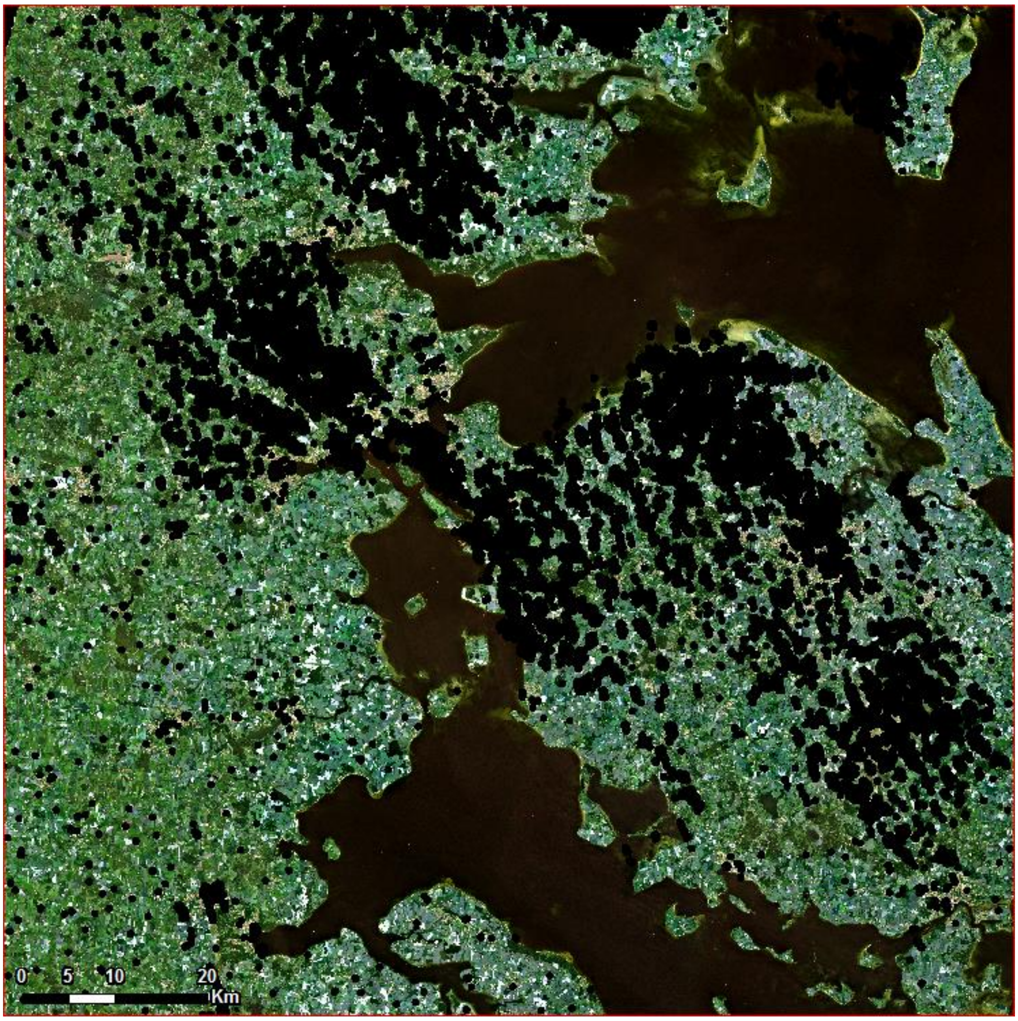

Figure 22 Clearest, least cloudy image in July (acquired July 19 2017)

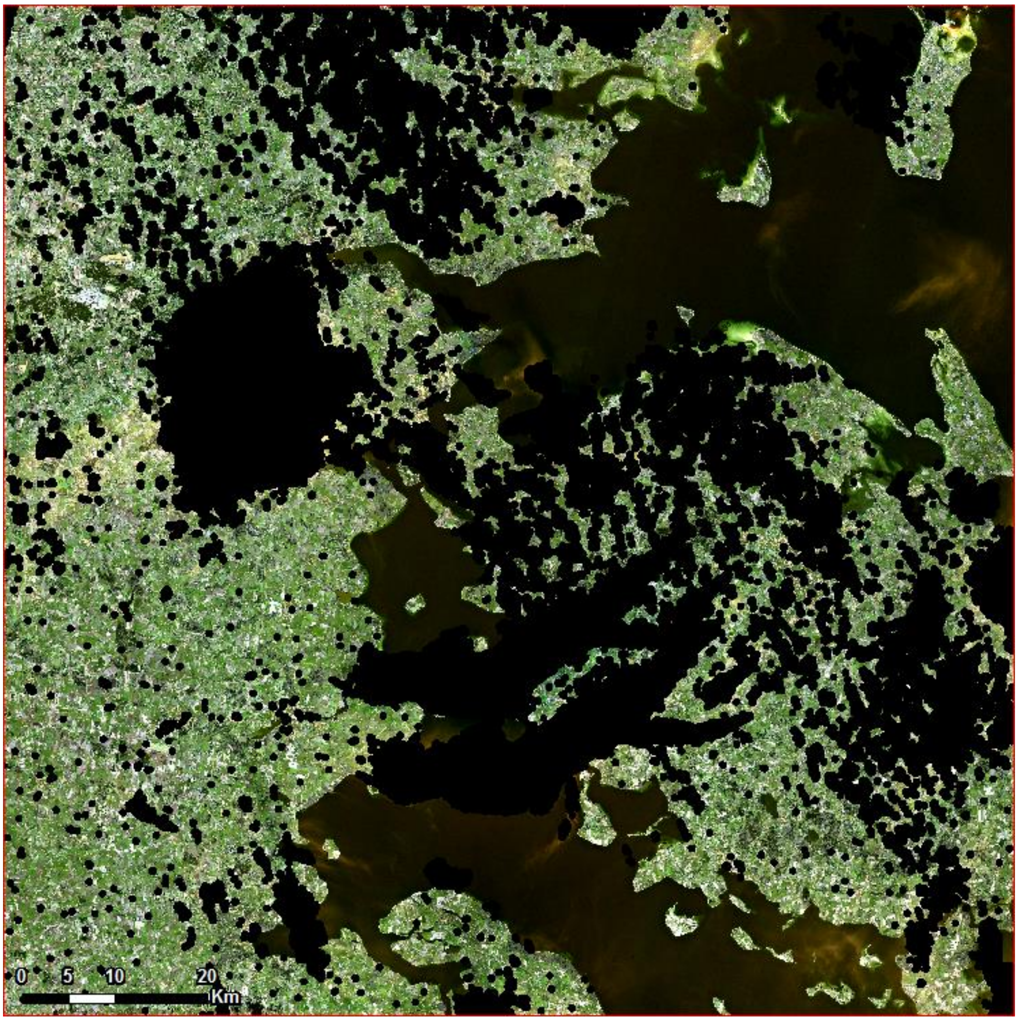

Figure 23 Clearest, least cloudy image classification input stack

Figure 23 above is the final input stack utilized for the classification using the clearest image approach. Due to constraints in the classification algorithm, ‘NoData’ values are not accepted for training values. The data has thus been further masked, and the masked out areas did not have a clear observation in neither March, nor July, thus further decreasing the available data. The data includes only the commonly clear area for both March and July.

4.1.2 Maximum NDVI Composite

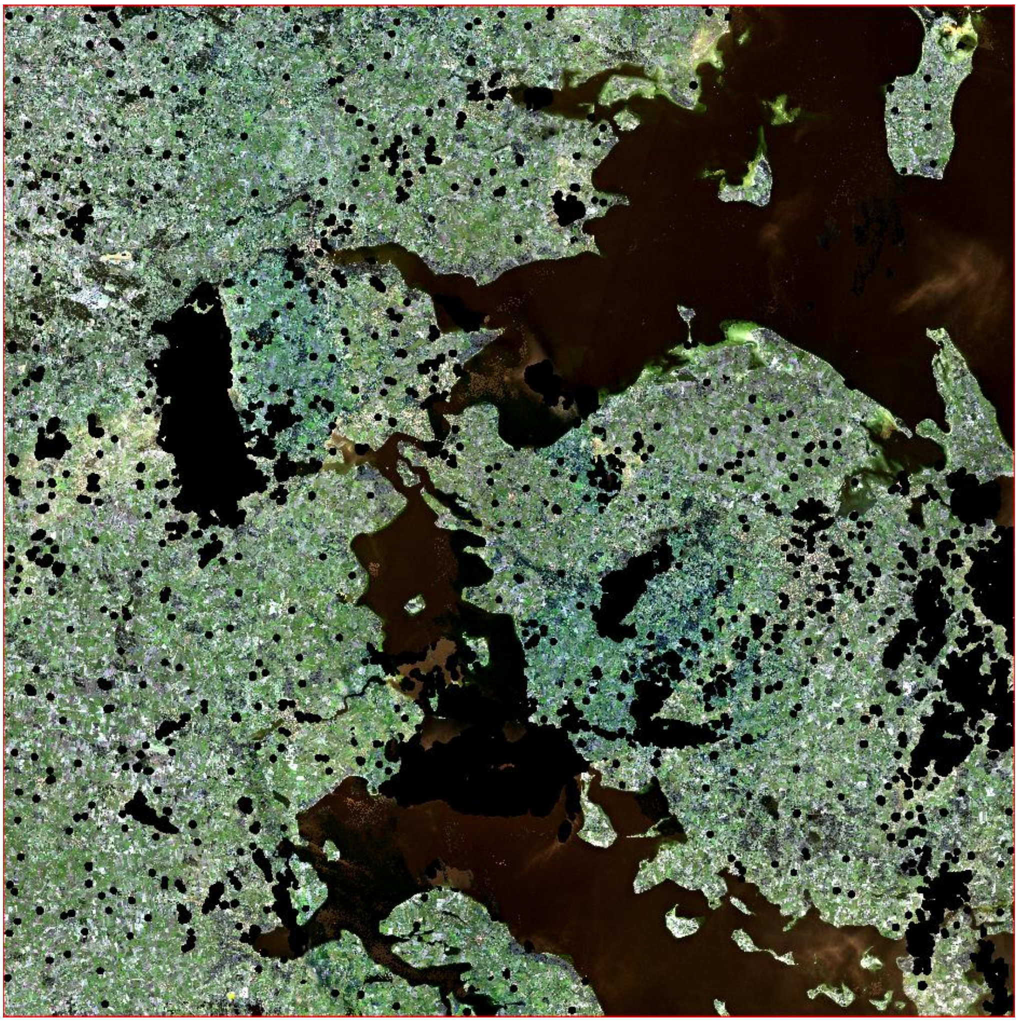

Figure 24 Maximum NDVI composite for March

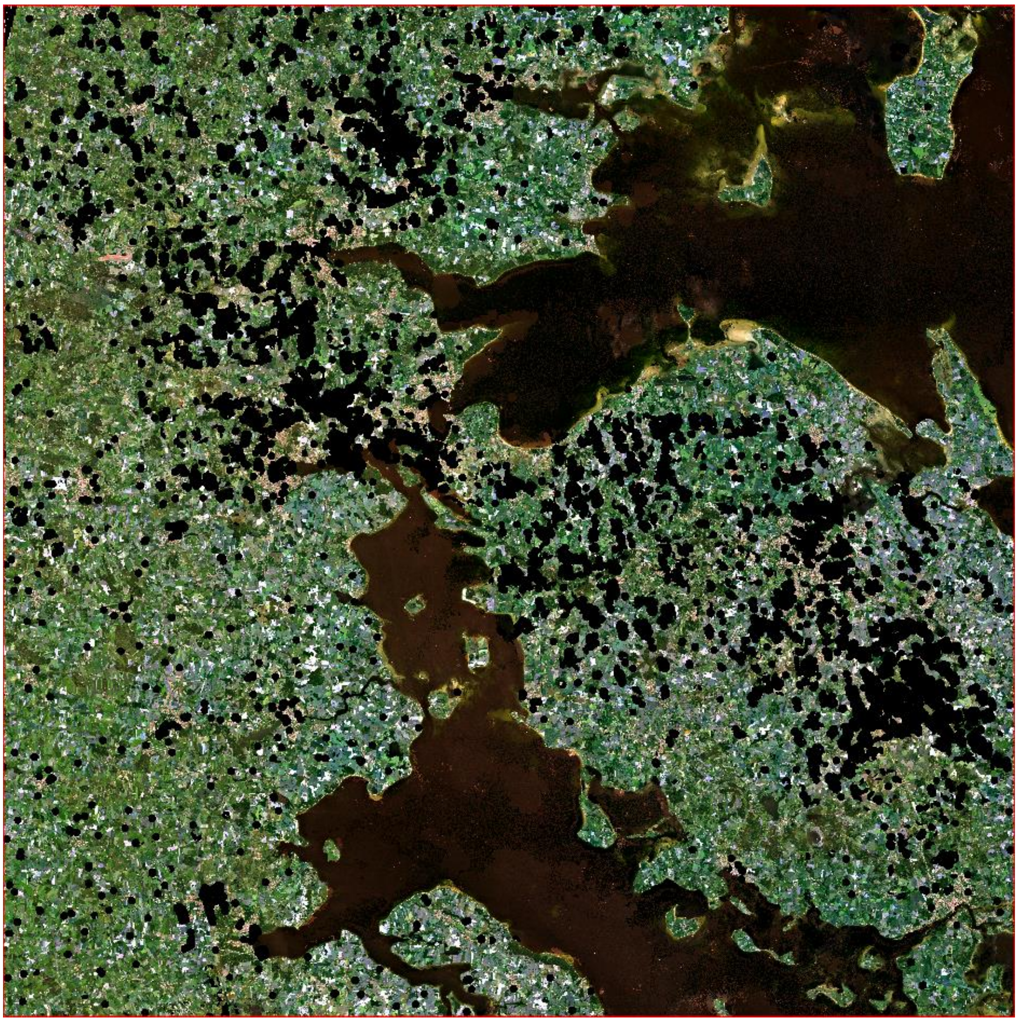

Figure 25 Maximum NDVI composite for July

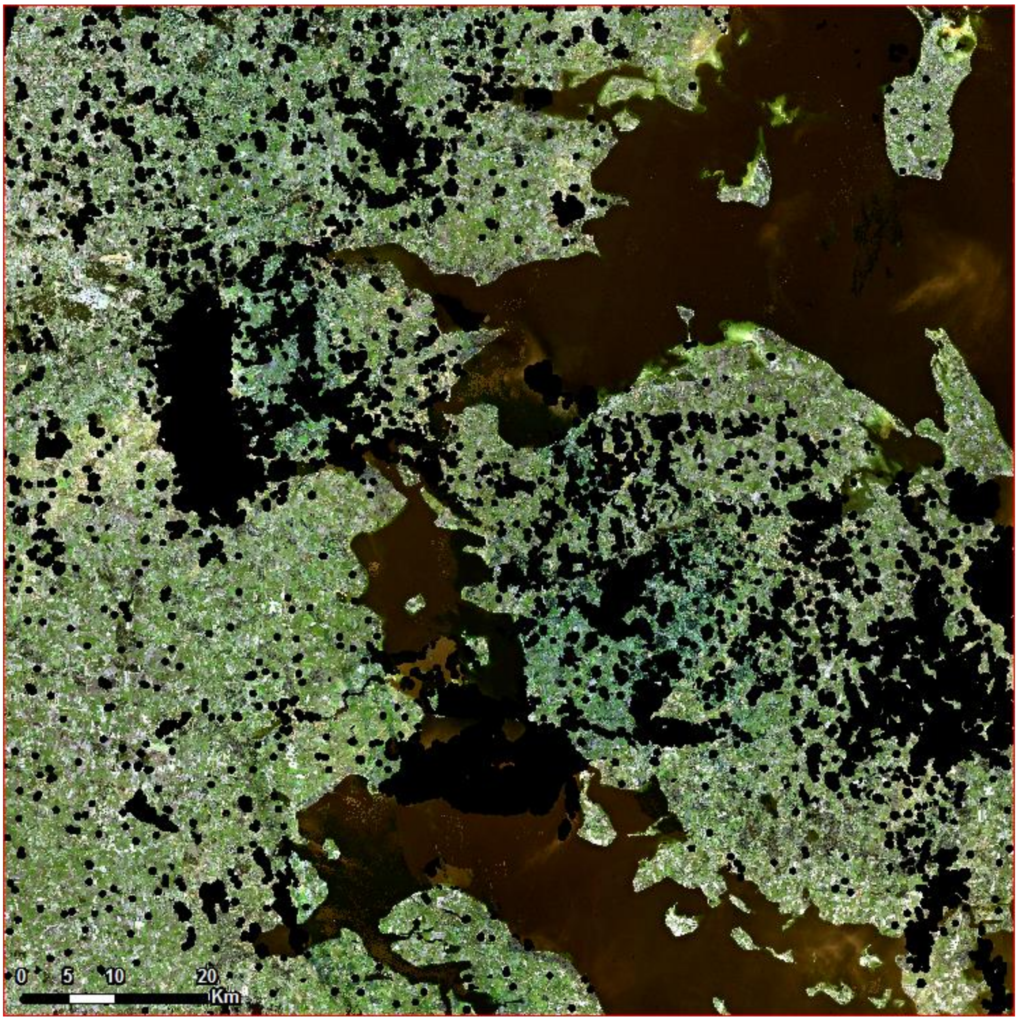

Figure 26 Maximum NDVI composite classification input stack

Figure 26 above is the final input stack utilized for the classification using the maximum NDVI compositing approach. Due to constraints in the classification algorithm, 'NoData' values are not accepted for training values. The data has thus been further masked, and the masked out areas did not have a clear observation in neither March, nor July, thus further decreasing the available data. The data includes only the commonly clear area for both March and July.

4.1.3 Medoid Composite

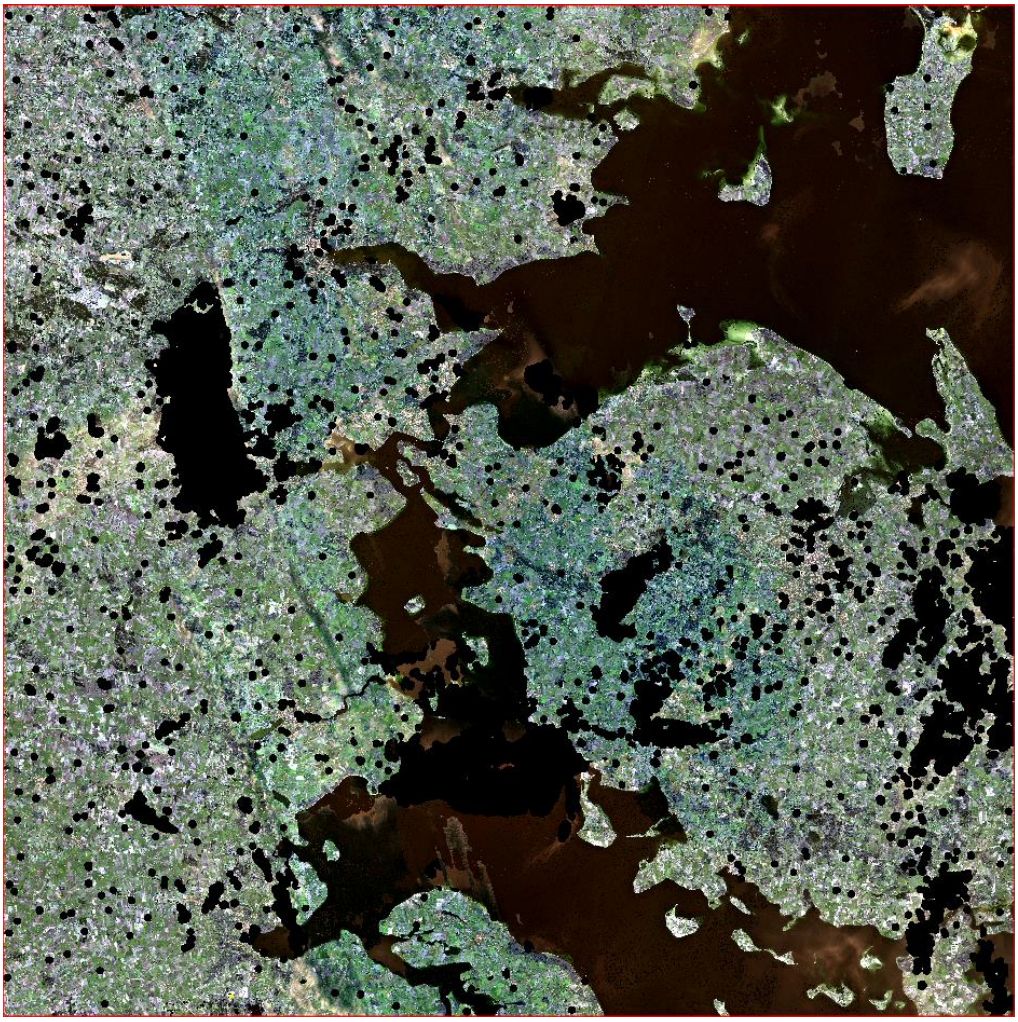

Figure 27 Medoid composite for March

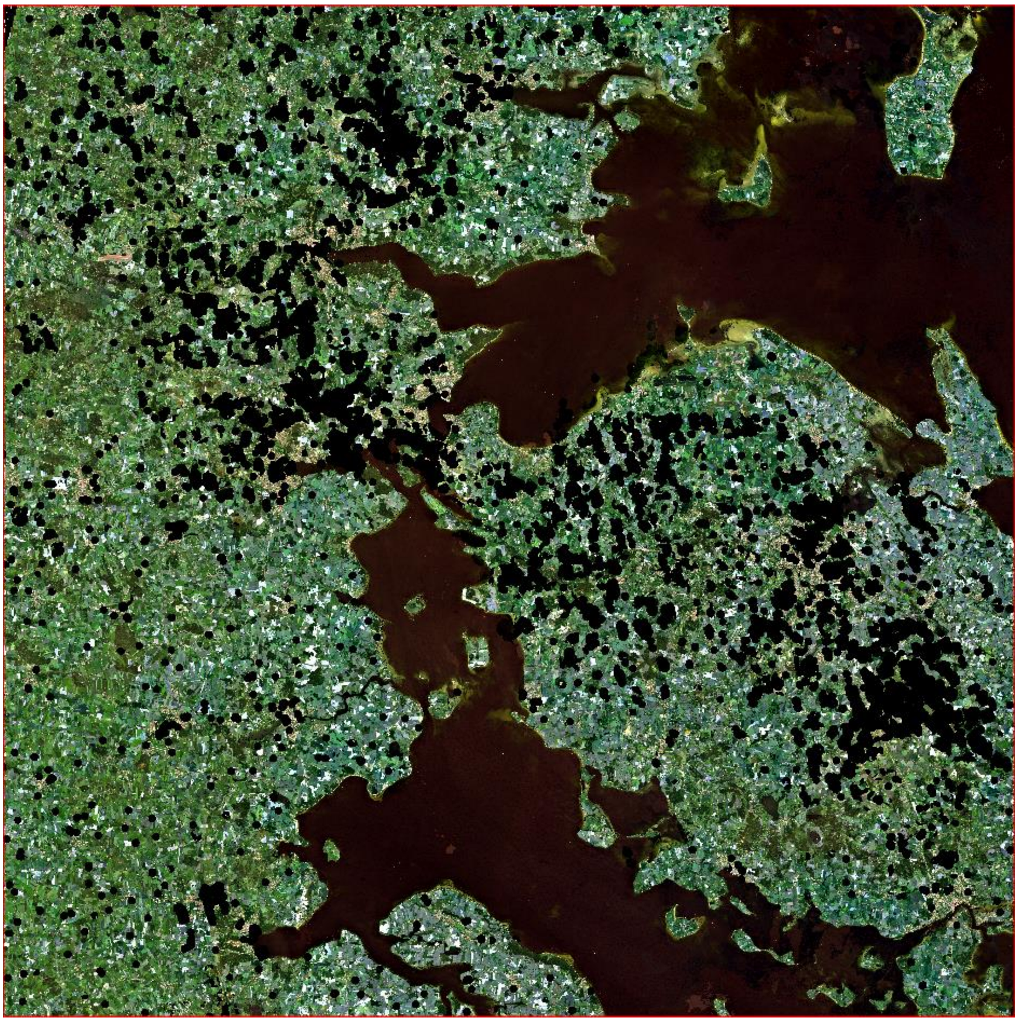

Figure 28 Medoid composite for July

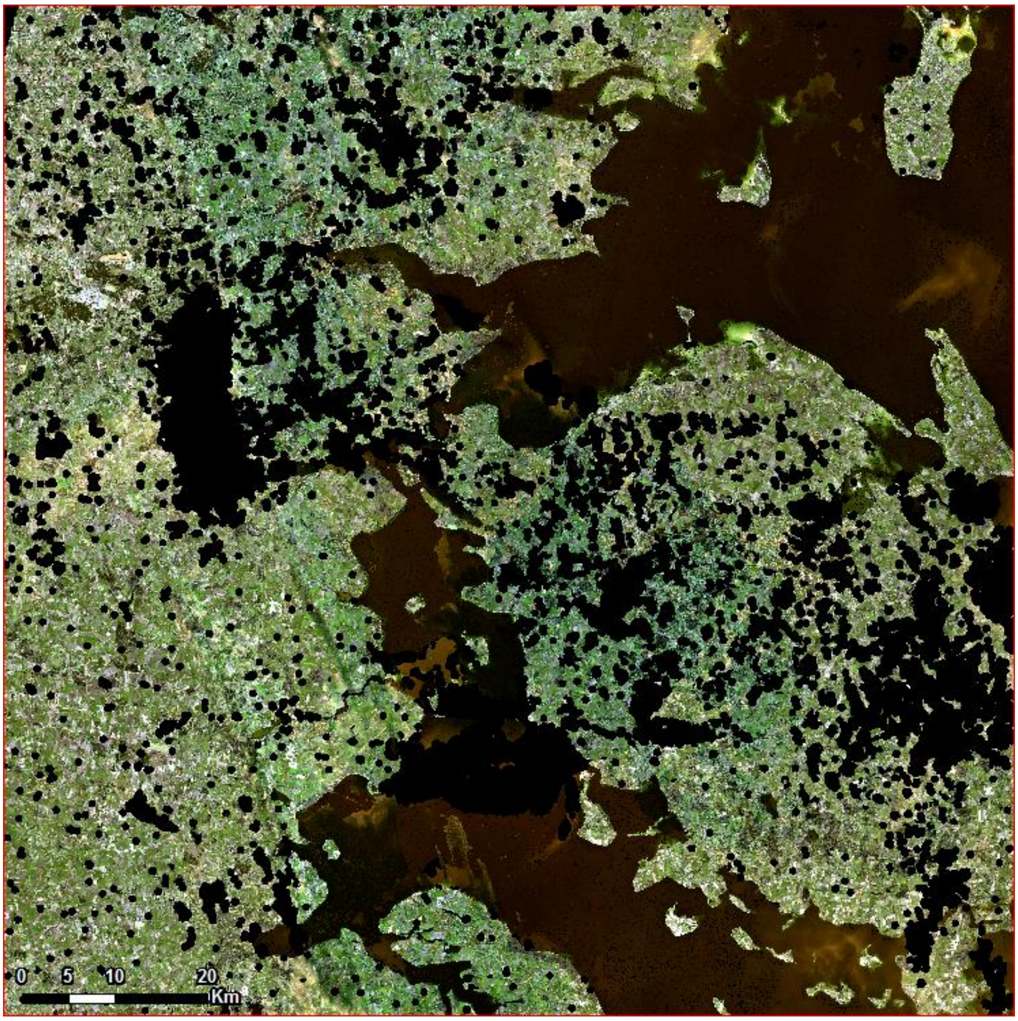

Figure 29 Medoid composite classification input stack

Figure 29 above is the final input stack utilized for the classification using the medoid compositing approach. Due to constraints in the classification algorithm, 'NoData' values are not accepted for training values. The data has thus been further masked, and the masked out areas did not have a clear observation in neither March, nor July, thus further decreasing the available data. The data includes only the commonly clear area for both March and July.

4.1.4 Geometric Median Composite

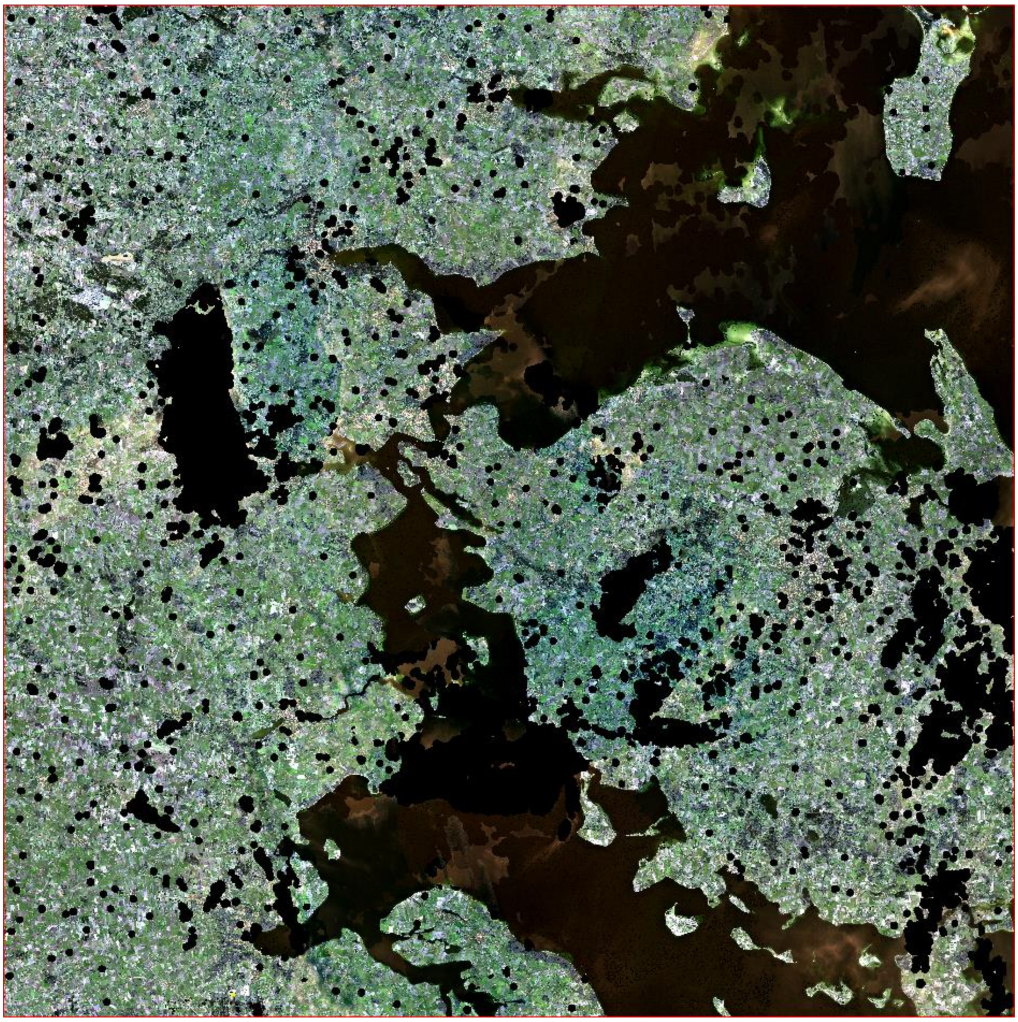

Figure 30 Geometric median composite for March

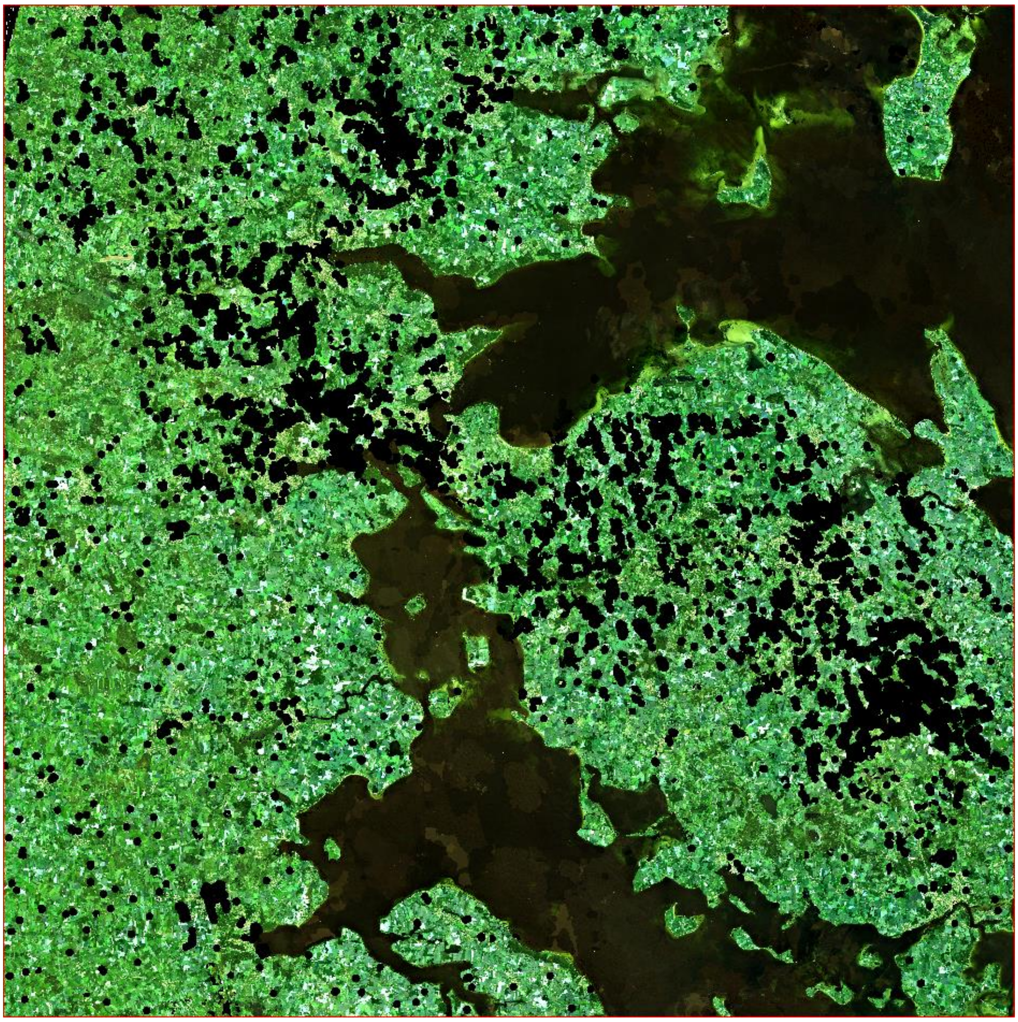

Figure 31 Geometric median composite for July

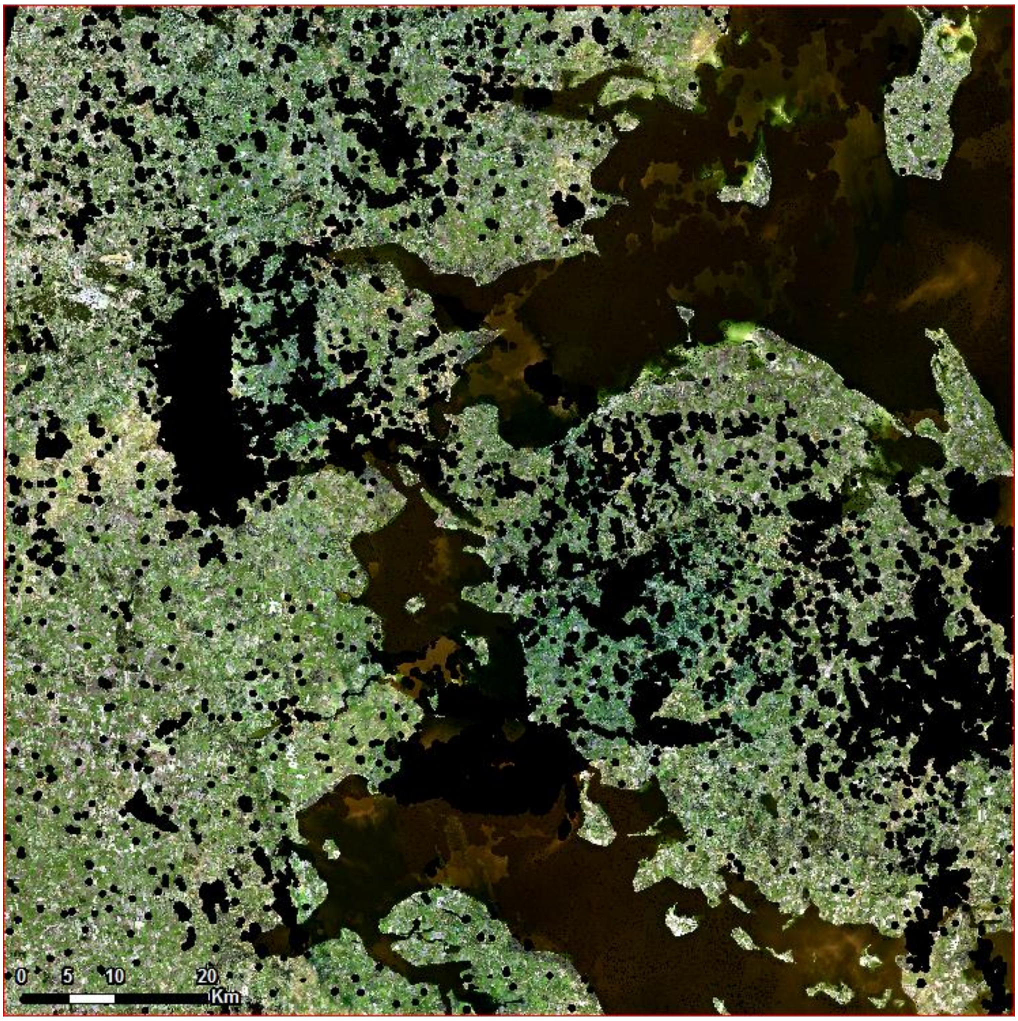

Figure 32 Geometric median composite classification input stack

Figure 32 above is the final input stack utilized for the classification using the geometric median compositing approach. Due to constraints in the classification algorithm, ‘NoData’ values are not accepted for training values. The data has thus been further masked, and the masked out areas did not have a clear observation in neither March, nor July, thus further decreasing the available data. The data includes only the commonly clear area for both March and July.

4.2 Classification Results

The following maps display the results of the random forest classification when applied to each composite and clearest image approach. The accuracy results will be discussed in the accuracy assessment section to follow.

4.2.1 Clearest Image Approach

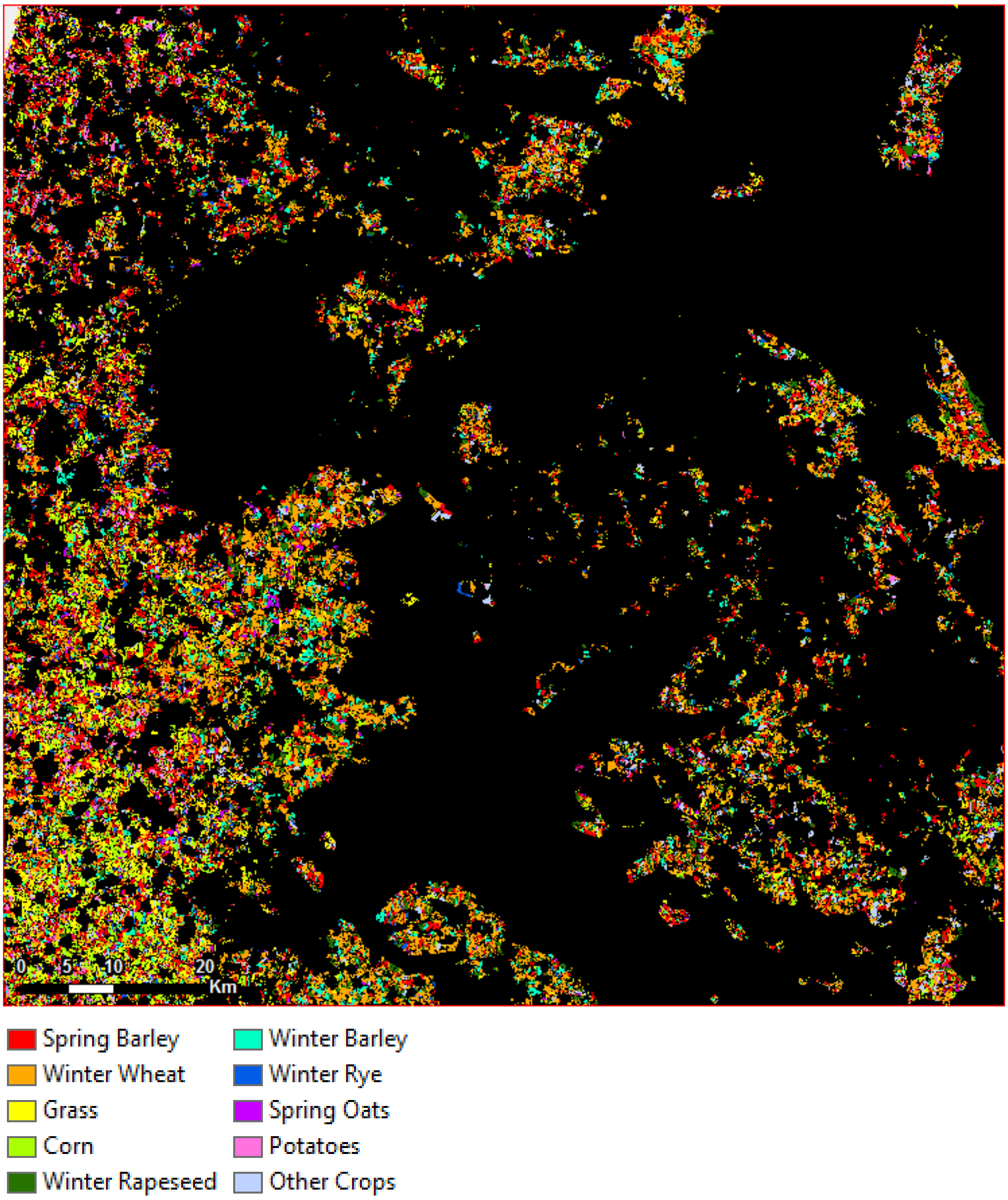

Figure 33 Clearest image classification results

4.2.2 Maximum NDVI Composite

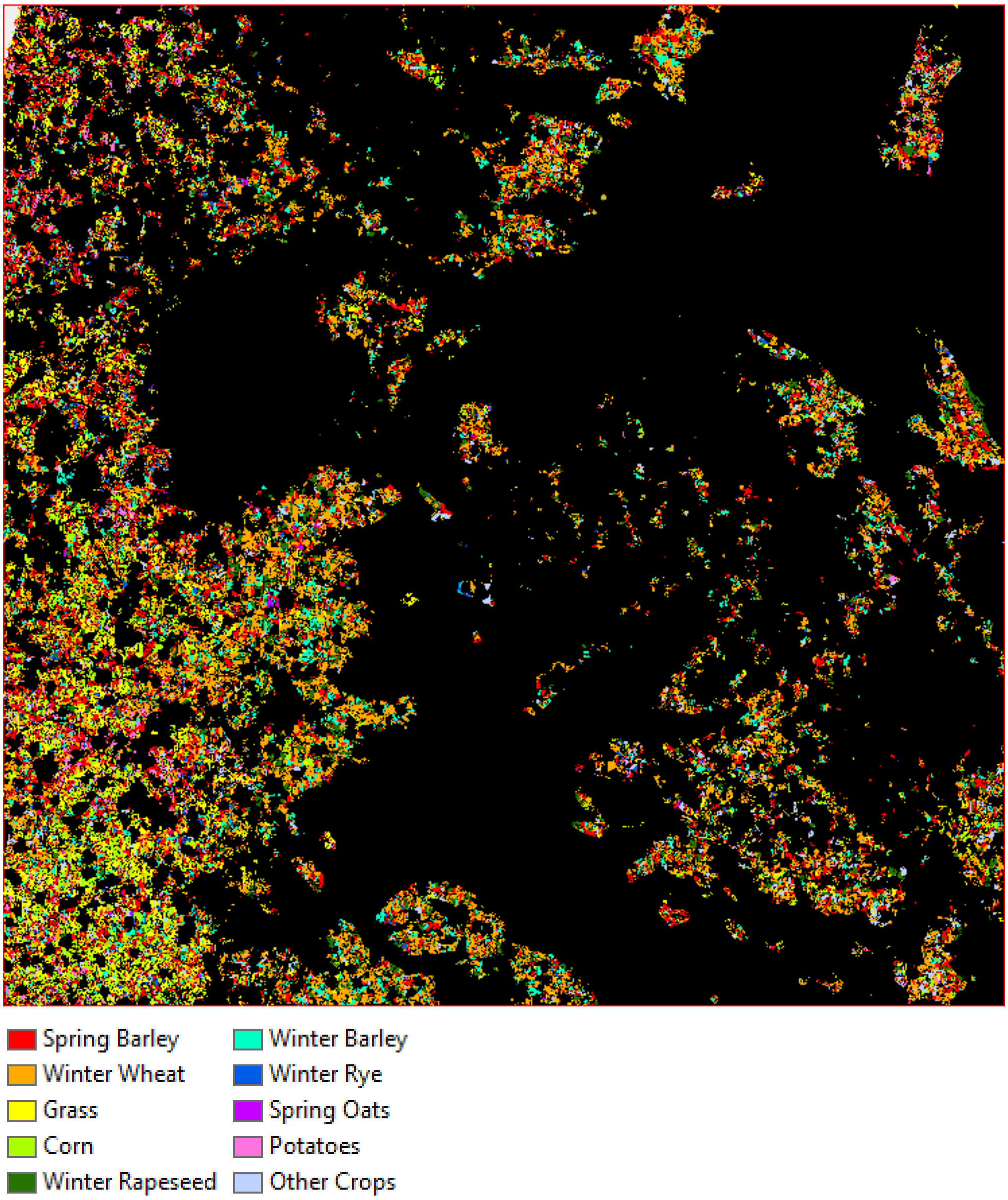

Figure 34 Maximum NDVI composite classification results

4.2.3 Medoid Composite

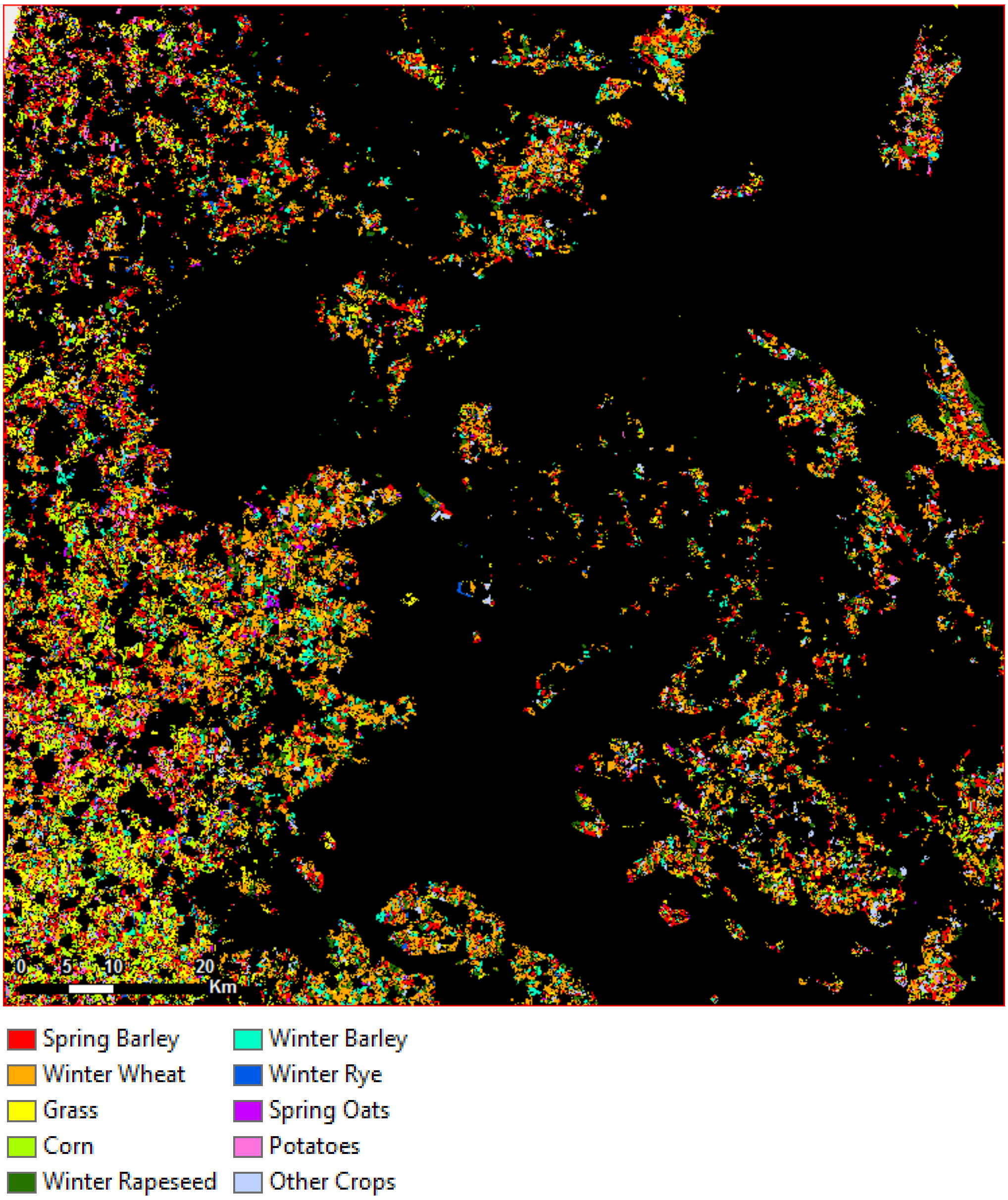

Figure 35 Medoid composite classification results

4.2.4 Geometric Median Composite

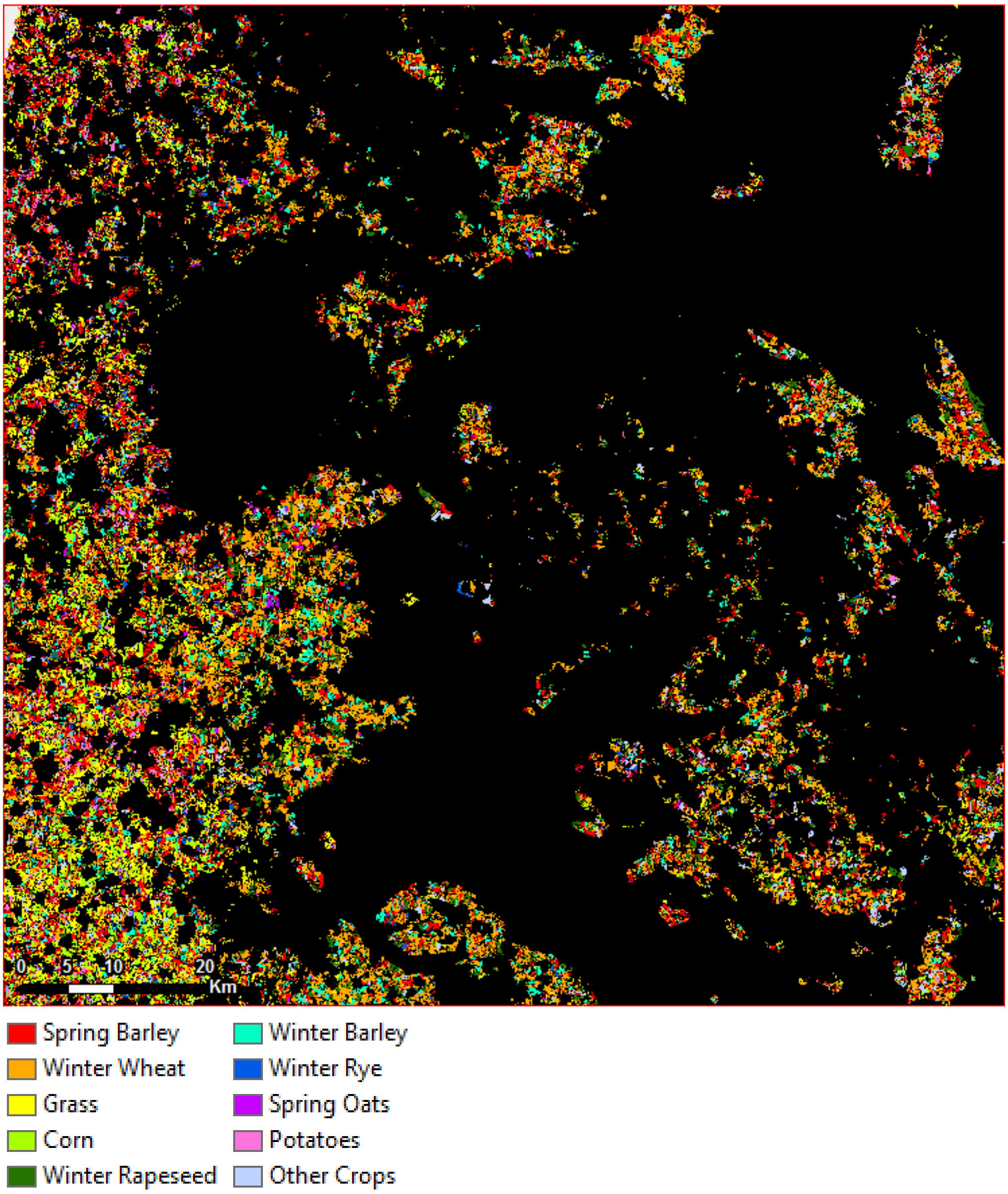

Figure 36 Geometric median composite classification results

4.3 Accuracy Assessment

An accuracy assessment has been performed on the resulting classifications using the GRASS GIS tool r.kappa. For testing purposes, 24,966 pixels were selected using a stratified random sample.

The clearest image approach achieved the highest overall accuracy of **84.79%**, the two high-dimensional median composite approaches, medoid and geometric median performed very similarly with overall accuracies of **84.21%** and **83.57%,** respectively. The maximum NDVI compositing approach performed the worst, achieving an overall accuracy of **81.35%**.

The consumer's accuracy for each class was fairly consistent across classes and across compositing methodologies, with no extreme outliers. The 'Other' class performed the worst in this measure with consumer's accuracies ranging from 67.12% to 72.60%. This outcome was somewhat expected, as the 'Other' class contains multiple crop types. On the producer's accuracy side, two crops achieved low accuracy results, Winter Rye and Spring Oats. Neither the composite images, nor the single image approach achieved high results in these classes, with the Maximum NDVI composite performing the worst and the single image approach performing the best.

These two classes cover a small portion of the total agricultural zone within the study area, and it is likely that the training data mask, which excluded pixels which did not have two or more clear observations in both March and July, masked out a lot of the fields containing these two crops, potentially reducing the amount of training data within these fields.

The composite images produced comparable per class and overall accuracies to the single image approach, showing that composite images are a suitable alternative for crop type classification. This is especially important since compositing largely increased the total number of useable pixels, with very little compromise in classification accuracy. The error matrices for each of the classification approaches, clearest image, and the three composite images can be seen starting overleaf.

4.3.1 Clearest Image Approach

Table 9 Clearest image error matrix

	Spring Barley	Winter Wheat	Grass	Corn	Winter Rapeseed	Winter Barley	Winter Rye	Spring Oats	Potatoes	Other	Row Sum	Consumer's Accuracy
Spring Barley	4608	114	53	82	10	9	31	352	8	244	5511	**83.61%**
Winter Wheat	338	6121	47	23	153	38	565	21	4	150	7460	**82.05%**
Grass	204	18	2758	160	3	3	17	30	41	218	3452	**79.90%**
Corn	27	3	19	2031	0	0	1	8	10	20	2119	**95.85%**
Winter Rapeseed	2	16	0	0	1734	0	6	0	0	3	1761	**98.47%**
Winter Barley	7	31	5	3	11	1600	20	0	1	29	1707	**93.73%**
Winter Rye	7	46	0	0	7	3	515	0	0	7	585	**88.03%**
Spring Oats	61	0	3	1	0	0	1	311	1	33	411	**75.67%**
Potatoes	28	0	24	10	0	2	2	1	480	19	566	**84.81%**
Other	119	11	96	80	2	8	4	21	41	1012	1394	**72.60%**
Col Sum	5401	6360	3005	2390	1920	1663	1162	744	586	1735	24966	
Producer's Accuracy	**85.32%**	**96.24%**	**91.78%**	**84.98%**	**90.31%**	**96.21%**	**44.32%**	**41.80%**	**81.91%**	**58.33%**		**OA 84.79%**

4.3.2 Maximum NDVI Composite

Table 10 Maximum NDVI composite error matrix

	Spring Barley	Winter Wheat	Grass	Corn	Winter Rapeseed	Winter Barley	Winter Rye	Spring Oats	Potatoes	Other	Row Sum	**Consumer's Accuracy**
Spring Barley	4280	132	67	115	9	8	16	542	21	289	5479	**78.12%**
Winter Wheat	633	5926	42	42	113	40	529	46	2	145	7518	**78.82%**
Grass	230	20	2689	146	1	3	15	38	58	186	3386	**79.42%**
Corn	21	3	21	2017	0	0	1	3	9	11	2086	**96.69%**
Winter Rapeseed	4	34	2	0	1768	1	12	0	0	8	1829	**96.66%**
Winter Barley	30	175	17	1	24	1601	89	0	2	69	2008	**79.73%**
Winter Rye	4	39	1	0	3	2	479	0	0	12	540	**88.70%**
Spring Oats	12	0	0	1	0	0	0	93	0	5	111	**83.78%**
Potatoes	26	0	13	8	0	0	2	1	461	14	525	**87.81%**
Other	161	31	153	60	2	8	19	21	33	996	1484	**67.12%**
Col Sum	5401	6360	3005	2390	1920	1663	1162	744	586	1735	24966	
Producer's Accuracy	**79.24%**	**93.18%**	**89.48%**	**84.39%**	**92.08%**	**96.27%**	**41.22%**	**12.50%**	**78.67%**	**57.41%**		**OA 81.35%**

4.3.3 Medoid Composite

Table 11 Medoid composite error matrix

	Spring Barley	Winter Wheat	Grass	Corn	Winter Rapeseed	Winter Barley	Winter Rye	Spring Oats	Potatoes	Other	Row Sum	Consumer's Accuracy
Spring Barley	4579	128	60	81	10	9	22	394	10	256	5549	**82.52%**
Winter Wheat	368	6076	44	29	146	43	599	26	4	137	7472	**81.32%**
Grass	205	20	2752	182	0	4	15	35	47	193	3453	**79.70%**
Corn	24	4	22	2014	0	0	1	4	8	15	2092	**96.27%**
Winter Rapeseed	5	45	2	0	1745	0	10	0	0	11	1818	**95.98%**
Winter Barley	7	33	8	2	13	1590	25	0	1	28	1707	**93.15%**
Winter Rye	8	39	0	0	1	2	478	0	0	3	531	**90.02%**
Spring Oats	57	2	1	4	0	0	0	258	1	20	343	**75.22%**
Potatoes	29	0	15	9	0	0	2	1	472	13	541	**87.25%**
Other	119	13	101	69	5	15	10	26	43	1059	1460	**72.53%**
Col Sum	5401	6360	3005	2390	1920	1663	1162	744	586	1735	24966	
Producer's Accuracy	**84.78%**	**95.53%**	**91.58%**	**84.27%**	**90.89%**	**95.61%**	**41.14%**	**34.68%**	**80.55%**	**61.04%**		**OA 84.21%**

4.3.4 Geometric Median Composite

Table 12 Geometric median composite error matrix

	Spring Barley	Winter Wheat	Grass	Corn	Winter Rapeseed	Winter Barley	Winter Rye	Spring Oats	Potatoes	Other	Row Sum	**Consumer's Accuracy**
Spring Barley	4537	127	54	77	10	9	24	464	11	287	5600	**81.02%**
Winter Wheat	410	6021	32	24	114	28	497	26	2	122	7276	**82.75%**
Grass	229	19	2736	228	1	5	15	40	47	196	3516	**77.82%**
Corn	28	1	27	1975	0	0	1	2	19	22	2075	**95.18%**
Winter Rapeseed	11	42	3	0	1777	3	20	0	0	12	1868	**95.13%**
Winter Barley	11	76	11	1	11	1601	38	0	2	34	1785	**89.69%**
Winter Rye	5	46	0	1	5	2	545	0	0	8	612	**89.05%**
Spring Oats	28	2	1	5	0	0	0	185	0	10	231	**80.09%**
Potatoes	27	0	12	9	0	0	2	1	451	9	511	**88.26%**
Other	115	26	129	70	2	15	20	26	54	1035	1492	**69.37%**
Col Sum	5401	6360	3005	2390	1920	1663	1162	744	586	1735	24966	
Producer's Accuracy	**84.00%**	**94.67%**	**91.05%**	**82.64%**	**92.55%**	**96.27%**	**46.90%**	**24.87%**	**76.96%**	**59.65%**		**OA 83.57%**

5 Discussion

This chapter provides insights to the results of this study and evaluates potential strengths and weakness of the methodologies implemented in this report. Applications of the results and limitations of this research are also provided.

Cloud masking and atmospheric correction were identified as the two most important techniques during the preprocessing stages. Research in the areas of cloud masking and atmospheric correction are ongoing, and the two processes implemented in this study are a state of the art combined cloud mask, and an open source Sentinel 2 specific atmospheric correction tool. These tools provided adequate performance when preparing the images for compositing in this study.

The three different methodologies compared in this study, maximum NDVI, medoid, and geometric median provided adequate results in terms of increasing the data availability. The maximum NDVI method has been previously used in research and has shown good results in this study as well. However, the newer, high-dimensional median compositing approaches provided a much higher increase in classification accuracy. These two compositing approaches should be researched further for the purposes of classification as the better accuracy results far outweigh the slightly longer processing times.

Four main indexes were included in this classification. These indexes are NDVI, Red Edge NDI, NDII, and NDWI. All the indexes utilized in this study have shown to increase accuracy results in past studies

The outcome of the study is that image composites can provide a large increase in the availability of data with very little compromise in classification accuracy when utilized for crop type classification. The differences in accuracies even between the best performing composite (Medoid – 84.21% OA) and the single image approach (84.79%) are statistically significant when testing using McNemar's test (McNemar, 1947). This difference however is rather insignificant in a real world classification scenario.

In addition to the similarly in overall accuracies, the composites performed on par with the single image approach when compared on a per class basis.

The primary benefit provided by the composites, far outweighs the small decrease in overall classification accuracy. For the purposes of this study, over 23.5% more fields were available to classify in the monthly composites, than in the clearest image approach.

From a classification perspective, the only clearly visible issues arose with the producer's accuracy of two of the smaller classes in this study: Winter Rye and Spring Oats. These two classes had the lowest accuracies, and were primarily confused with two of the largest classes, Winter Wheat and Spring Barley respectively. While this is not an issue with the composites, it appears to be an issue with the training data used within this study and it should be investigated further.

5.1 Limitations of the Study

The compositing methodologies, while beneficial for this study are however not perfect for visualization purposes. Preprocessing steps, especially cloud masking, can be drastically improved in order to ensure better data preparation for compositing. Cloud masking and atmospheric correction are paramount to achieving high accuracies in any type of remote sensing classification. The combined cloud mask utilized in the study, while it seemed to remove the majority of clouds, was too aggressive in certain areas, fully removing non cloudy pixels which contained bright buildings with a similar spectral signature to that of clouds, as can be seen in figure 37 overleaf.

Figure 37 Cloud mask erroneously removing bright buildings

The improvement of cloud masks and atmospheric correction methodologies is an ongoing topic in remote sensing, with many researchers attempting to utilize machine learning and deep learning approaches to solve these problems.

A second issues with composites is the pixel homogeneity within a field. In a single image approach, all the pixels within a field were captured on the same date, and assuming the field contains a single crop, with the entire field planted on the same date, only minor spectral differences should be visible. In a composite however, pixels within the same fields may have been acquired on different dates, creating minor artefacts within the image. While this study has shown this to not be a major issue when using the composites for classification, the artefacts within the image make for a less visually appealing image, as seen in figure 38 overleaf. The geometric median composite displayed the fewest artefacts when compared to other compositing methodologies, perhaps indicating that a geometric median composite could be the most beneficial when utilized for classification over a larger scale.

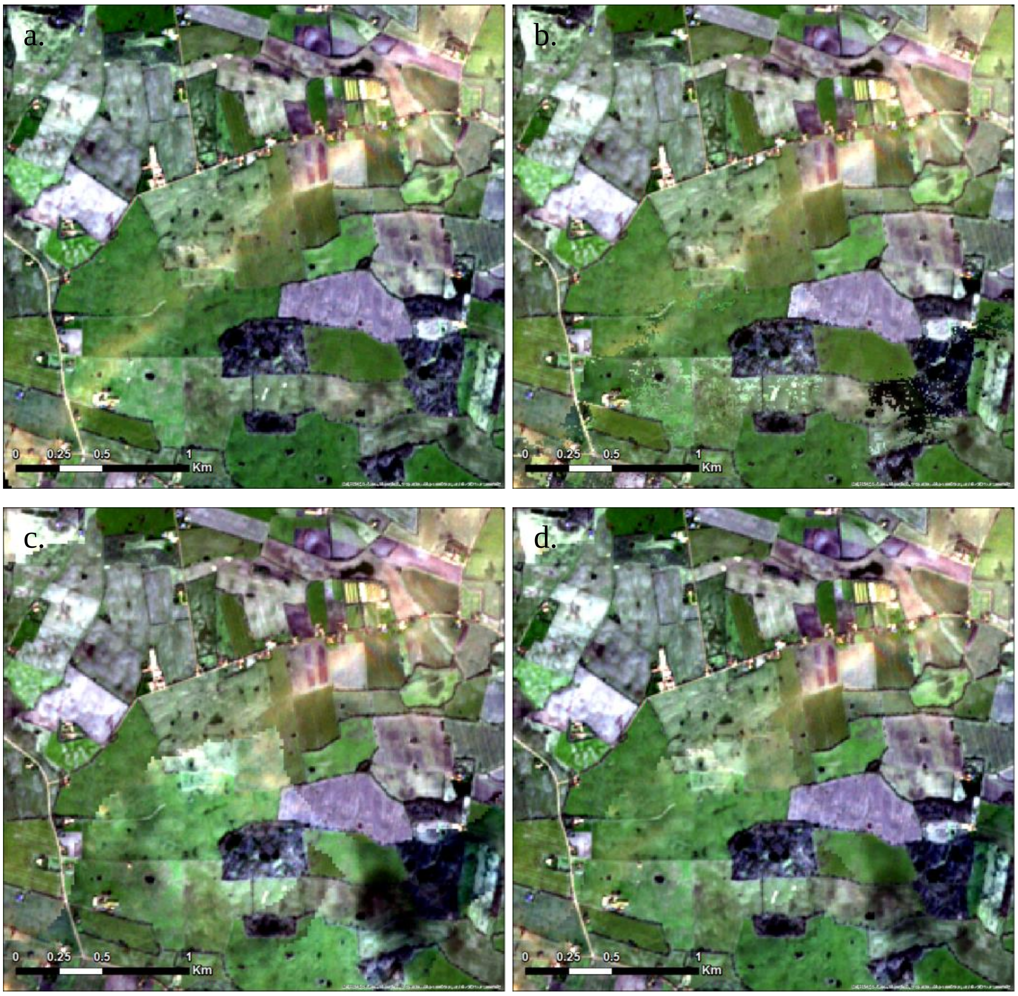

Figure 38 Artefacts within composite images a. Single image b. Maximum NDVI composite c. Medoid composite d. Geometric median composite

A second limitation of this study is the training data. The training and testing data for this project was based on open government data, which might not be available on a worldwide scale, thus limiting the replicability of this study. Within the European Union however, due to the Common Agricultural Policy reporting rules, this data should be available, however it may not always be accurate. The data is self-reported by farmers in order to apply for agricultural subsidies, and while The Danish AgriFish Agency expects approximately 95% accuracy in this self-reported data, this high level of accuracy may not be applicable everywhere.

5.2 Applications of the Results

The results of this study showed fairly high accuracy results, considering only optical data has been used for classification. The composites performed fairly well when compared to the single image approach, showing only a slight decrease in classification accuracy, however the composites generated clear pixels for over **23.5%** more fields than the clearest image approach. This is the main advantage of using composites over a single image approach, as more data could potentially mean that less field visits would be required for policy makers, potentially decreasing costs associated with controlling farmers.

On a wider scale, the composites provide a big opportunity for monitoring and classification of land cover and land use in cloudy and rainy areas in certain periods, such as a wet season in Africa or South East Asia. The composites could make crop yield estimates more viable in such areas.

6 Conclusion

This thesis explored how monthly satellite image composites can be used for the purposes of crop type classification, and how the accuracy results of these composites compare to a more classic, clearest, least cloudy image classification approach. In order to assist with the research objectives for this study, a tile based compositing methodology for Sentinel 2 tiles was developed, using tile 32UNG in Denmark as a case study, where monthly composites for March and July 2017 were created using three different methodologies, maximum NDVI, medoid, and geometric median.

The significance of utilizing composite images for the purpose of crop type classification has been shown. Results indicated that monthly image composites can be beneficial for the purposes of crop type classification. The composite images provided an increase of 23.5% in the total number of fields available for classification. The overall accuracies for this study were 84.79%, 84.21%, 83.57%, and 81.35% for the clearest image, medoid, geometric median, and maximum NDVI classifications, respectively.

Progress in remote sensing technologies, such as improvements in cloud masking and atmospheric correction can potentially show further increase in the accuracy results for crop type classification, and improvements in these sectors should be explored in the future.

This study has shown the applicability of monthly image composites for classification at a tile based scale, however future research in tasks involving larger scale areas, such as land cover and land use classification tasks for entire countries could be beneficial to increase the popularity of the high-dimensional median compositing approaches which performed well in this study.

6.1 Future Directions

The study established that monthly composites can be utilized for the purpose of crop type classification. The study was however performed on only one Sentinel 2 tile, limiting the size of the area. Applications over larger areas, containing multiple acquisition paths, where same day acquisitions for a single image approach are not possible would be an interesting future direction. The composite will smooth the values over the month, likely creating a more uniform data set for large scale, such as country wise classifications.

Due to the limitations in the data availability, the final composites in this study had a number of pixels where no clear observations could be computed into a final composite, a second step into future directions could focus on gapfilling approaches for high-dimensional median composites.

Lastly, high-dimensional median composites have only been recently introduced in remote sensing applications. Further testing for different classification processes, such as land cover/land use classifications or change detection using the medoid and geometric median approaches could prove valuable.

References

Aggarwal, S. (2004). Principles of remote sensing. *Satellite remote sensing and GIS applications in agricultural meteorology*, 23-38.

Aronoff, S. (1982). Classification accuracy: a user approach. *Photogrammetric Engineering and Remote Sensing*, *48*(8), 1299-1307.

Aronoff, S. (1985). The minimum accuracy value as an index of classification accuracy. *Photogrammetric Engineering and Remote Sensing*, *51*(1), 99-111.

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24-31.

Breaker, L. C., Armstrong, E. M., & Endris, C. A. (2010). Establishing an objective basis for image compositing in satellite oceanography. *Remote Sensing of Environment*, *114*(2), 345-362.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Brockmann, C., Paperin, M., Danne, O., & Ruescas, A. (2013, September). Multi-Sensor Cloud Screening and Validation: IdePix and PixBox. In *Proceedings of the 2013 European Space Agency Living Planet Symposium.*

Conel, J. E., Green, R. O., Vane, G., Bruegge, C. J., Alley, R. E., & Curtiss, B. J. (1987). AIS-2 radiometry and a comparison of methods for the recovery of ground reflectance.

Delegido, J., Verrelst, J., Alonso, L., & Moreno, J. (2011). Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. *Sensors*, *11*(7), 7063-7081.

European Space Agency (ESA). (2018). Level-1C Product Types. Retrieved from https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/product-types/level-1c

Flood, N. (2013). Seasonal composite Landsat TM/ETM+ images using the medoid (a multi-dimensional median). *Remote Sensing*, *5*(12), 6481-6500.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote sensing of environment*, *80*(1), 185-201.

Gao, B. C., Heidebrecht, K. B., & Goetz, A. F. (1993). Derivation of scaled surface reflectances from AVIRIS data. *Remote sensing of Environment*, *44*(2-3), 165-178.

Gao, B. C., & Kaufman, Y. J. (1995). Selection of the 1.375-μm MODIS channel for remote sensing of cirrus clouds and stratospheric aerosols from space. *Journal of the atmospheric sciences*, *52*(23), 4231-4237.

Gao, B. C., Montes, M. J., Davis, C. O., & Goetz, A. F. (2009). Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. *Remote Sensing of Environment*, *113*, S17-S24.

Gao, B. C., Montes, M. J., Li, R. R., Dierssen, H. M., & Davis, C. O. (2007). An atmospheric correction algorithm for remote sensing of bright coastal waters using MODIS land and ocean channels in the solar spectral region. IEEE Transactions on Geoscience and Remote Sensing, 45(6), 1835-1843.

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

Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V. C., Abrudan, I. V., Lieskovsky, J., ... & Hostert, P. (2014). Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. *Remote Sensing of Environment*, *151*, 72-88.

Griffiths, P., van der Linden, S., Kuemmerle, T., & Hostert, P. (2013). A pixel-based Landsat compositing algorithm for large area land cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*(5), 2088-2101.

Hadjimitsis, D. G., Papadavid, G., Agapiou, A., Themistocleous, K., Hadjimitsis, M. G., Retalis, A., ... & Clayton, C. R. I. (2010). Atmospheric correction for satellite remotely sensed data intended for agricultural applications: impact on vegetation indices. *Natural Hazards and Earth System Sciences*, *10*(1), 89-95.

Hagolle, O., Huc, M., Pascual, D. V., & Dedieu, G. (2010). A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENμS, LANDSAT and SENTINEL-2 images. *Remote Sensing of Environment*, *114*(8), 1747-1755.

Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International journal of remote sensing*, *7*(11), 1417-1434.

Karimi, H. A. (Ed.). (2004). *Telegeoinformatics: Location-based computing and services*. CRC press.

Koukoulas, S., & Blackburn, G. A. (2001). Introducing new indices for accuracy evaluation of classified images representing semi-natural woodland environments. *Photogrammetric Engineering and Remote Sensing*, *67*(4), 499-510.

Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, *14*(5), 778-782.

Lawrence, R. L., Wood, S. D., & Sheley, R. L. (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). *Remote Sensing of Environment*, *100*(3), 356-362.

Martins, V. S., Barbosa, C. C. F., de Carvalho, L. A. S., Jorge, D. S. F., Lobo, F. D. L., & Novo, E. M. L. D. M. (2017). Assessment of atmospheric correction methods for Sentinel-2 MSI images applied to Amazon floodplain lakes. *Remote Sensing*, *9*(4), 322.

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, *12*(2), 153-157.

Muller-Wilm, U., Louis, J., Richter, R., Gascon, F., & Niezette, M. (2013, September). Sentinel-2 level 2A prototype processor: Architecture, algorithms and first results. In *Proceedings of the ESA Living Planet Symposium, Edinburgh, UK* (pp. 9-13).

Natural Resources Canada (NRCAN). (2018). Passive vs. Active Sensing. Retrieved from http://www.nrcan.gc.ca/node/14639

Potapov, P., Turubanova, S., & Hansen, M. C. (2011). Regional-scale boreal forest cover and change mapping using Landsat data composites for European Russia. *Remote Sensing of Environment*, *115*(2), 548-561.

Richards, J. A., & Jia, X. (2006). Remote sensing digital image analysis. *New York: Springer-Verlag Berlin Heidelberg*, 55.

Roberts, D., Mueller, N., & Mcintyre, A. (2017). High-dimensional pixel composites from earth observation time series. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(11), 6254-6264.

Robinson, N. P., Allred, B. W., Jones, M. O., Moreno, A., Kimball, J. S., Naugle, D. E., ... & Richardson, A. D. (2017). A Dynamic Landsat Derived Normalized Difference Vegetation Index (NDVI) Product for the Conterminous United States. *Remote Sensing*, *9*(8), 863.

Roy, D. P., Ju, J., Kline, K., Scaramuzza, P. L., Kovalskyy, V., Hansen, M., ... & Zhang, C. (2010). Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, *114*(1), 35-49.

Satellite Imaging Corporation. (2018). Characterization of Satellite Remote Sensing Systems. Retrieved from https://www.satimagingcorp.com/services/resources/characterization-of-satellite-remote-sensing-systems/

Sagar, S., Phillips, C., Bala, B., Roberts, D., & Lymburner, L. (2018). Generating Continental Scale Pixel-Based Surface Reflectance Composites in Coastal Regions with the Use of a Multi-Resolution Tidal Model. *Remote Sensing*, *10*(3), 480.

Sonobe, R., Tani, H., Wang, X., Kobayashi, N., & Shimamura, H. (2014). Random forest classification of crop type using multi-temporal TerraSAR-X dual-polarimetric data. *Remote Sensing Letters*, *5*(2), 157-164.

United States Geological Survey (USGS). (2018). What is Remote Sensing and What is it Used for?. Retrieved from https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used?qt-news_science_products=7#qt-news_science_products

Wardlow, B. D., Egbert, S. L., & Kastens, J. H. (2007). Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains. *Remote Sensing of Environment*, *108*(3), 290-310.

White, J. C., Wulder, M. A., Hobart, G. W., Luther, J. E., Hermosilla, T., Griffiths, P., ... & Guindon, L. (2014). Pixel-based image compositing for large-area dense time series applications and science. *Canadian Journal of Remote Sensing*, *40*(3), 192-212.

Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., ... & Nemani, R. (2008). Free access to Landsat imagery. *Science*, *320*(5879), 1011-1011.

Wulder, M. A., White, J. C., Loveland, T. R., Woodcock, C. E., Belward, A. S., Cohen, W. B., ... & Roy, D. P. (2016). The global Landsat archive: Status, consolidation, and direction. *Remote Sensing of Environment*, *185*, 271-283.

Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, *159*, 269-277.

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote sensing of environment*, *118*, 83-94.