Student thesis series INES nr 573

Early-stage detection of bark beetle infested spruce forest stands using Sentinel-2 data and vegetation indices

Karl Piltz

2022

Department of

Physical Geography and Ecosystem Science

Lund University

Sölvegatan 12

S-223 62 Lund

Sweden

Karl Piltz (2022).
Early-stage detection of bark beetle infested spruce forest stands using Sentinel-2 data and vegetation indices

Identifiering av Barkborre angripna granskogsbestånd i ett tidigt stadie med hjälp av Sentinel-2 data och vegetationsindex
Master degree thesis, 30 credits in *Geomatics*
Department of Physical Geography and Ecosystem Science, Lund University

Level: Master of Science (MSc)

Course duration: *January* 2022 until *June* 2022

Disclaimer

Early-stage detection of bark beetle infested spruce forest stands using Sentinel-2 data and vegetation indices

---

Karl Piltz

Master thesis, 30 credits, in *Geomatics*

Per-Ola Olsson
Dep. Of Physical Geography and Ecosystem Science, Lund University

Exam committee:
Torbern Tagesson
Dep. Of Physical Geography and Ecosystem Science, Lund University

Acknowledgement

First, I would like to thank my supervisor Per-Ola Olsson for the support throughout this thesis! You have made this project into a very positive experience in which I have learned a lot and came through with a newfound confidence thanks to our many discussions and your excellent supervision. I would also like to thank my friends, my girlfriend and my mother who has been behind me every step of the way, supporting me with many laughter's, coffees, some sauna and helpful discussions even though they were tired of my never-ending thesis talk.

Abstract

The European spruce bark beetle is an insect that is often referred to as a pest. Responsible for the destruction of over 150 million m^3 of Norwegian spruce forest in Europe over the last 50 years makes this insect one of the major disturbances to the forest industry. With global warming at large, changes of the distribution of the bark beetle have emerged with large outbreaks now regularly occurring in northern Europe due to warmer and prolonged summer seasons. Sweden has since 2018 been affected by mass outbreaks that have destroyed over 27 million m^3 of spruce forest. In order to mitigate the disturbance of forest and limit the spread of further attacks, it is important to detect and cut down infested trees at an early stage. In recent time, many studies have focused on early-stage bark beetle detection using remote sensing methods. Efforts to detect early-stage infestation, i.e., "green attacks" using vegetation indices (VI) on a pixel level, have found varying levels of success but have shown the potential of using VIs sensitivity of changes to biochemical leaf properties to detect early-stage infestation. Hence, the aim of this thesis was to study if the variability between pixels for a vegetation index on a forest stand scale change during a spruce bark beetle outbreak and to test if the variability can be used as an early indicator for bark beetle infestation. This was done by calculating the coefficient of variation between 2017 and the outbreak year 2018 from Sentinel-2 data for four different VIs (NDVI, NDWI, DRS, RDI) in 17 spruce forest stands where 10 stands were infested and 7 were healthy. The coefficient of variation was used to classify the stands into healthy and infested by computing the cumulative sum of each VI in each stand. The classification performance for each VI was evaluated using receiver operating characteristics graphs which were used to find the optimal threshold for each classification. The classification was done using the cumulative sum for two different timeframes, early stage (1^{st} of May-1^{st} of July 2018) and the whole bark beetle season (30^{th} of April-30^{th} of September) The results of the thesis, indicated that the variability within a forest stand does change during a bark beetle outbreak with increased variability over time within stands that have been attacked. It also showed that changes in variability have the potential to be used as an early indicator for bark beetle infestation, and the variability can be used to detect and classify individual forest stands that were infested at an early stage, i.e., green attack stage. It was found that NDWI was the most suitable index during the period May-July to detect infested forest stands. However, for the whole season NDVI and RDI also displayed potential as both were able to detect high rates of infested forest stands while limiting the misclassification of healthy ones. An infested forest stand could be detected as early as ca. the 29^{th} of May. Meaning that the infestation is still in the green attack stage, in which mitigation is still possible to eliminate the spread of further bark beetle attacks.

Table of contents

List of Figures

List of Tables

List of abbreviations

AOI	Area of interest
VI	Vegetation index
NIR	Near infrared
SWIR	Shortwave infrared
NDVI	Normalized difference vegetation index
NDWI	Normalized difference water index
DRS	Distance Red SWIR
RDI	Ratio drought index
ROC	Receiver operating characteristics
Tpr	True positive rate
Fpr	False positive rate

1. Introduction

The European Spruce bark beetle is an insect that is often referred to as a pest for the forestry industry (Huo et al., 2021; Wulff & Roberge, 2021). Throughout Europe the bark beetle has been the culprit for large economic as well as ecological loss due to infestation and destruction of over 150 million m^3 of Norway spruce forest stands over the last 50 years (Huo et al., 2021). The main region of infestation has historically been located around central Europe however with earlier seasonal warming and longer warm periods, large outbreaks have started to occur regularly in northern Europe as well (Abdullah, Skidmore, et al., 2019a; Huo et al., 2021). Sweden has become increasingly affected by mass infestation of Norwegian spruce forest due to Bark Beetles attacks with over 27 million m^3 forest destroyed since 2018 (Huo et al., 2021; Martin Schroeder, 2020; Wulff & Roberge, 2021). The year 2018 was a record year for spruce bark beetle infection in Sweden and was associated with the heatwaves that occurred during the summer (Martin Schroeder, 2020; Wilcke et al., 2020; Wulff & Roberge, 2021). The increased temperature led to heightened vulnerability within the forest stands mainly through drought (Bárta et al., 2021; Martin Schroeder, 2020; Netherer & Hammerbacher, 2022; Wulff & Roberge, 2021). The European spruce bark beetle feed and spawn beneath the bark of the spruce tree and are especially drawn to trees stressed by drought, storms or disease (Blennow, 2004; Martin Schroeder, 2020). With the prolonged warm summer season increased reproduction opportunity the bark beetle population was able to spawn in two waves leading to at the time a record outbreak (Huo et al., 2021; Wulff & Roberge, 2021). Under normal circumstances the bark beetle has a positive effect on the ecosystem by disforesting weak trees (Bárta et al., 2021; Huo et al., 2021). However, during mass outbreaks as those seen since 2018 in Sweden, the population of Bark beetles can grow to numbers high enough to attack healthy trees too, leading to damages to ecosystem values and large economic loss for the forestry industry (Bárta et al., 2021; Martin Schroeder, 2020). As climate change is projected to increase causing higher temperatures and extreme weather events, this too is projected to have an enhancing effect on Bark beetle swarming in Sweden (Abdullah, Skidmore, et al., 2019a; Huo et al., 2021). With the already alarming pattern over the last 4 years, it is highly important to detect and combat the issue at an early stage to limit economic and ecological loss.

To counteract the swarming of Bark beetles it is important to detect infested trees early. By identifying affected trees at an early stage when the offspring beneath the bark are still larvae (Huo et al., 2021), it is possible to remove the attacked tree(s) from the forest and limit the spread to other trees (Huo et al., 2021). Detection of infested trees has previously mainly been done manually through forest inventory which is often ineffective, as it is cumbersome, and difficult to do on a larger scale (Abdullah, Darvishzadeh, et al., 2019). Instead, technical solutions have emerged with the use of remote sensing as a tool for detecting bark beetle infested trees. The detection is based on the changing spectral properties of an infested tree which typically goes through three stages (Green, Red & Gray attacks) that are spectrally different to each other (Abdullah, Darvishzadeh, et al., 2019). Green attack is the first phase where the needles are green, and the insects are still larvae (Abdullah, Darvishzadeh, et al., 2019). Red attack is the second phase in which the tree´s needles changes colour from green towards red as an indication of stress (Abdullah, Darvishzadeh, et al., 2019). Gray-attack is the final phase in which the tree no longer can support its foliage and sheds all needle leaving only the bare grey trunk and branches (Abdullah, Darvishzadeh, et al., 2019).

Previous studies have investigated how the spectral properties change when a tree is infested and various vegetation index (VI) have been compared to identify which indices that are more suitable to detect different phases of infestation (Abdullah, Darvishzadeh, et al., 2019; Abdullah, Skidmore, et al., 2019a; Huo et al., 2021). Especially successful was the identification of red attacks and grey attacks due to the larger difference in reflectance compared to healthy trees (Huo et al., 2021). Green attacks on the other hand have been found to be more difficult to detect due to the smaller differences in reflectance between attacked and healthy trees, especially in the visual spectrum (Huo et al., 2021). As early detection is key to mitigate the bark beetle problem, several recent studies have focused on detection of green attacks. This has been done by linking leaf parameters (such as water content and chlorophyll) to foliage reflectance changes due to bark beetle infestation (Abdullah et al., 2018). Abdullah et al. (2018) and Abdullah, Darvishzadeh, et al. (2019) explored this link and showed a high potential to detect early-stage green attacks using VIs based on Red, Green, Blue (RGB), Near infrared (NIR) and shortwave infrared (SWIR) wavelength bands derived from the Landsat-8 satellite. In a later study, Abdullah, Skidmore, et al. (2019) found Sentinel-2 satellite data to be more sensitive to changes in reflectance than Landsat resulting in higher accuracy for green attack detection. Further studies focusing on premeditated stressors as an indicator instead of direct green attack identification using sentinel-2 satellite data combined with VIs displayed a considerably higher accuracy (80-88%) when detecting early stage bark beetle attacks (Huo et al., 2021).

Previous studies have presented methods for early detection by studying spectral alterations on a pixel level as an indicator. To enable detection, the change in reflectance must be sufficiently large within a pixel. What has not been investigated is if spectral variability between pixels within a homogenous forest stand could be used as an indicator for early detection. During an outbreak, bark beetles attack individual or smaller groups of trees within a stand. These attacked trees might result in an increased variability between pixels that is detectable even if the change is not sufficiently large for detection in single pixels.

1.1 Aim of study

The aim of the thesis is to study if the variability between pixels for a vegetation index on a forest stand scale change during a spruce bark beetle outbreak and to test if the variability can be used as an early indicator for bark beetle infestation. Which vegetation indices to study is decided by the literature review. The specific research questions are:

Can variability in vegetation index values between pixels within a forest stand be used as an indicator for early-stage bark beetle infestation?

Which vegetation index(es) is most suitable to detect early-stage bark beetle infestations?

How early can infested trees be detected?

2. Background

2.1 European Spruce bark beetle

The European spruce bark beetle (*Ips typographus* L.) is one of the main forest disturbances in Sweden (Jönsson et al., 2007; Netherer et al., 2021). The insect is from the Curculionidae family which contain over 6000 different species of bark beetles (Jönsson et al., 2011). The specie is widespread around central Europe and Scandinavia, and its distribution is often connected with the occurrence of Norway spruce trees (Jakoby et al., 2019). Within the diverse family of bark beetles, not all pose a threat like the spruce bark beetle as many are specialized in targeting other tree species and less critical parts of the tree (Bentz & Jönsson, 2015; Krokene, 2015). The spruce bark beetle however targets the stem of the tree by feeding on and reproducing under the tree bark (Bentz & Jönsson, 2015; Hlásny et al., 2021). The infestation of a tree begins by a male locating a pheromone signal emitted from the tree indicating distress (Fettig & Hilszczański, 2015; Wermelinger, 2004). As the male works its way through the tree bark the tree tries to defend itself by drowning the bark beetle(s) with resin (Fettig & Hilszczański, 2015; Krokene, 2015; Wermelinger, 2004). The beetle is able use the resin to send out a chemically altered pheromone signal that other individuals pick up and gather to overpower the defence mechanisms of the tree through mass attacks (Fettig & Hilszczański, 2015; Krokene, 2015; Vega & Hofstetter, 2014). Tunnels are built beneath the bark and female beetles arrive to reproduce and lay their eggs within the tunnels (Cognato, 2015). As the eggs hatch larvae expand the tunnels to then incubate themselves in cocoons (Raffa et al., 2008). It takes around 7 weeks for the offspring to hatch and reach maturity and after 10 weeks the new generation is ready to leave the infested tree in search for a new host (SKOGSSTYRELSEN, 2019). During this time the tree dies as it is consumed by the beetles, hindering nutrient and water transportation within the tree and weakening its defence mechanisms (Abdullah et al., 2018). The preferred target is forest around 60 years old which have more available substrate, maximising reproduction potential (Hlásny et al., 2021). The period of bark beetle swarming is thermally dependent, meaning that for the specie to swarm the temperature must be at least 16 °C with + degrees for a longer time period (Abdullah, Skidmore, et al., 2019b; Jönsson et al., 2007; Jönsson et al., 2011). Usually, the bark beetle season occur around April-May and May-June and hibernate during winter periods in the ground or beneath the bark (Netherer & Hammerbacher, 2022).

The European spruce bark beetle is under normal circumstances ecologically valuable to the forest ecosystem (Hlásny et al., 2021). The infestation of weak trees contributes to enhanced biodiversity by breaking up homogenous tree stands as well as nutrient cycling (Hlásny et al., 2021). However, under certain conditions, like the warm summer in Sweden 2018 the population can grow into numbers high enough to be regarded an epidemic and hence become destructive to the forest system (Hlásny et al., 2021; SKOGSSTYRELSEN, 2019). The conditions that favour epidemic growth are warm, dry and long summer seasons with large amount of substrate to take advantage off (water deficient trees and/or wind fell trees) (Blennow, 2004; Martin Schroeder, 2020). Normally, the bark beetle is dependent on seasonality and can be somewhat predictive, however during years with favourable conditions can lead to a second generation of beetles within the same season (Blennow, 2004). In which case the population density allows for powerful aggregated attacks not only on weak old trees but also younger and healthy ones (Bárta et al., 2021; Hlásny et al., 2021; Kirkendall et al., 2015). When the population breach the epidemic threshold in this fashion as it has done on

multiple occasions in Europe and Sweden it leads to mass tree mortality events and heavily damage economic and ecological values of the forest (Grégoire et al., 2015; Hlásny et al., 2021). Sweden has over the years 2010-2021 experienced a pattern of increasingly larger and more frequent mass infestation events of Norwegian spruce forest triggered by the hot drought in 2018 (SKOGSSTYRELSEN, 2019; Wulff & Roberge, 2021). Recorded since the outbreak was triggered in 2018 is an alarming trend with 3-4 million m^3 of forest killed in 2018, 7 million m^3 in 2019, 7.7 million m3 in 2020 and 8.2 million m^3 in 2021 (Huo et al., 2021; Martin Schroeder, 2020; Wulff & Roberge, 2021).

2.1.2 Effect on the Spruce trees from a bark beetle infestation

A successful bark beetle attack will eventually lead to the death of the attacked spruce tree. A single tree can host on average 2 generations of beetles until it is depleted of sustenance that nourish the population. During this time the tree slowly dies as an effect of hosting the insects due to the cut-off of water and nutrient circulations from the roots as well as chloroplast degradation (Abdullah, Skidmore, et al., 2019a). The beetles effectively act as a disease from the point of entry until leaving for the search of a new host. Like any deadly disease, the tree slowly gets worse over time emphasised by different stages of sickness. The stress that spruce trees experience from bark beetle infestation is typically broken down into 3 different stages emphasised by biochemical alteration and spectral changes (Abdullah et al., 2018; Huo et al., 2021). These stages may vary depending on the specie of bark beetle; however, the European Spruce bark beetle has 3 pronounced phases (Abdullah et al., 2018). These phases are correlated with the phenology and life cycle of the beetles and is hence as indicative to the bark beetle life cycle as it is forest stress. The first phase of stress is denoted green-attack and emphasises that the foliage on the affected tree is still green on this stage (Abdullah, Skidmore, et al., 2019a; Bárta et al., 2021). This stage is apparent early in the infestation process, the bark beetle has been established within the host, but the newly hatched generations are still larvae and is yet to grow an exoskeleton and leave the tree (Abdullah, Darvishzadeh, et al., 2019; Abdullah, Skidmore, et al., 2019a). During this time the tree is overwhelmed and stressed but show no sign of stress in its foliage (Abdullah, Darvishzadeh, et al., 2019; Abdullah, Skidmore, et al., 2019a). The second phase is known as the red attack; here the stress is more pronounced as the needles have lost pigment turning reddish from lack of water and nutrient availability (Abdullah, Darvishzadeh, et al., 2019). During this phase the new generation of bark beetles have already gained maturity and left the tree (Abdullah, Darvishzadeh, et al., 2019). The third and final stage is the grey phase, in which the stressed spruce tree is diseased and can no longer support its foliage and in turn is left with only its bare grey stem (Abdullah, Darvishzadeh, et al., 2019; Bárta et al., 2021). In order to effectively reduce the impact of bark beetle disturbance and preventing mass outbreak, it is crucial to identify and remove a tree whilst in green stage when the larvae are still contained. Hence, methods for large scale early detection are required to counteract present and future problems.

2.2 Remote sensing

2.2.1 Remote sensing of terrestrial vegetation

Remote sensing is a technology to observe objects from a distance that has provided an efficient tool for monitoring of the earth´s surface on a scale not previously obtainable through the eyes of man. The utilization of remotely sensed information such as imagery from satellite, airplane, UAV, and ground mounted cameras provide a more efficient, detailed and cheaper alternative to traditional in situ monitoring.

Satellite based sensors like Landsat-8 and Sentinel-2 have contributed immensely to how vegetation can be studied and what can be learned from observing the terrestrial landscape (Liang & Wang, 2019; Roberts, 2014). Vegetation is perceived through the human eye by the absorption and reflection of solar radiation (Jones & Vaughan, 2010). Different types of vegetation reflect more or less of different wavelength of the electromagnetic spectrum which make it possible to distinguish between different species (Jones & Vaughan, 2010). The human eye is sensible to light within the visual spectrum (Red, Green & Blue) and hence can display colours (Jones & Vaughan, 2010). For example, forests are perceived as green since it reflects strongly within the green wavelength band of the electromagnetic spectrum while absorbing red and blue light (Jones & Vaughan, 2010). While the human eye is limited to the visual spectrum, multispectral and hyperspectral sensors such as those used on satellites can pick up reflectance from wavelengths outside of the visual (Jones & Vaughan, 2010). Vegetation types differ from each other as they all have unique spectral characteristics i.e., a spectral signature which can be used to identify and monitor specific species (Jones & Vaughan, 2010). Furthermore, spectral signatures create the prospect of observing changes over time ex. a vegetative season to study changes and response patterns (Jones & Vaughan, 2010). Remote sensing technology has hence been applied to identify stress and diseases which is made possible by the changes of spectral properties resulting in deviation from the spectral signatures of healthy vegetation (Roberts, 2014). The use of multispectral and hyperspectral sensors takes advantage of wavelength bands that are sensitive to biophysical/chemical properties of vegetation which are strongly affected by change and stress (Hanes, 2013; Jones & Vaughan, 2010). Different bands are sensitive to different biochemical properties, for example near infrared (NIR) is sensitive to leaf cell structure whilst shortwave infrared (SWIR) is sensitive to leaf water content (Hanes, 2013; Jones & Vaughan, 2010). It is therefore possible to extract and study internal leaf properties and alterations which then can be used as proxy for early-stage disease detection (Roberts, 2014). Furthermore, when working with vegetation biophysical properties the reflectance of individual bands is not always sufficient to detect changes due to background noise from due to e.g., the ground, atmosphere, or topography. Instead, vegetation indices are often computed from a combination of multiple bands (Hanes, 2013; Huete, 2014; Jones & Vaughan, 2010). A vegetation index is normally calculated using two contrasting bands of which one is highly absorbed (often red or blue band) by the target and the other highly reflected (NIR or SWIR) to be able to clearly extract the desired property (Hanes, 2013). Like individual bands different indexes are more sensitive to different leaf traits for example a combination of NIR and red is sensitive to greenness related properties while a combination with SWIR is sensitive to water related traits (Hanes, 2013).

2.2.2 Remote sensing of bark beetle attacks

The implementation of remote sensing to monitor bark beetle attacks has been a growing topic over the last decade. As global temperature has risen, and more pronounced and regular mass outbreaks of bark beetles has occurred, researchers has worked more intensely to find methods to detect early-stage infestation.

Abdullah et al. (2018), highlighted the impact and spectral alteration that occur on spruce needles during a successful attack. In a lab experiment, needles were made to experience the same biochemical changes that would occur naturally from bark beetle attacks. Abdullah et al. (2018) measured the spectral responses of different biochemical parameters at multiple wavelengths of the electromagnetic spectrum to differentiate between needles from healthy and infested "trees". Furthermore, they studied which wavelength band(s) that could be deemed significant and most clearly displayed reflection, based on the biochemical parameters and conducted a band separation analysis based on the results presenting suggestions for what bands most efficiently highlight bark beetle activity. Abdullah et al. (2018) found that there was a significant difference in reflectance between infested and healthy samples especially in the NIR and SWIR. The findings showcased the potential use of multi/hyperspectral remote sensing techniques for early detection of bark beetle infestation.

In a later study, Abdullah et al. (2019) attempted to identify early stages green attacks of bark beetles in the Bavarian Forest National Park (Germany) using spectral vegetation indices and canopy surface temperature obtained from Landsat-8 satellite images. The study included multiple vegetation indices using the visual spectrum as well as NIR and SWIR bands. The bands where chosen based on their sensitivity to biochemical alteration such as chlorophyll and leaf water content. The study produced stress maps of healthy and infested trees and concluded that the leaf traits measured from the VIs where all significantly different when comparing healthy and infested trees, suggesting that early detection can be facilitated using NIR and SWIR wavelength bands and vegetation indexes sensitive to leaf water content (ex. NDWI & RDI). Moreover, the study highlighted the role of temperature as in indicator for biochemical alteration from changes in leaf water content, chlorophyll and stomatal conductance due to bark beetle activity.

Furthermore, (Abdullah, Skidmore, et al., 2019a) compared Landsat-8 and Sentinel-2 satellite data for early detection and found that while Landsat-8 was able to distinguish infested pixels using vegetation indexes sensitive to leaf water content (such as NDWI, RDI and DSWI) Sentinel-2 was able to classify stress using vegetation indices sensitive to both water content as well as with indices sensitive to chlorophyll. Moreover, the study showed that Sentinel-2 imagery proved to be more accurate than Landsat-8 where Sentinel-2 reached 67% accuracy and Landsat-8 36% when classifying data into healthy and attacked pixels. The study showcased the prospective use of Sentinel-2 data for early bark beetle detection as well the potential of the red-edge wavelength bands.

Huo et al. (2021) conducted a study in Sweden trying to detect bark beetle green-attack stage using Sentinel-1 and Sentinel-2 data. The study analysed how the spectral properties of stressed spruce trees changed over the vegetation season. It was found that characteristic anomalies could be identified already prior to the actual infestation of bark beetles while spectral differences during a green-attack stage was more difficult to separate. Hence, the study used spectral alteration from stress prior to infestation as an indicator for bark beetle activity later in

the vegetative season. (Huo et al., 2021) also suggested a new vegetation index the Normalized Distance Red & SWIR (NDRS) index which instead of utilizing the NIR has a combination of the red and SWIR bands. The use of NIR has since (Abdullah et al., 2018) found contradicting for detection on a larger scale which is highlighted by (Huo et al., 2021; Ortiz et al., 2013) who found the reflectance in the Red and Red-edge to be better suited for early detection compared to NIR. The study resulted in higher accuracies than recorded in previous studies with 80-82% in early stage (May-July) and 81-91% in mid to late stage (August – October). The study emphasized the use of SWIR which was found to display considerably high separability compared to other bands when measuring classification accuracy (Figure 1). Furthermore, the study suggest that the Red band should be used instead of NIR when working with VIs (Figure 2) and (Huo et al., 2021) highlight the potential of utilizing stress indication and vulnerability to trace potential infestation later in the season.

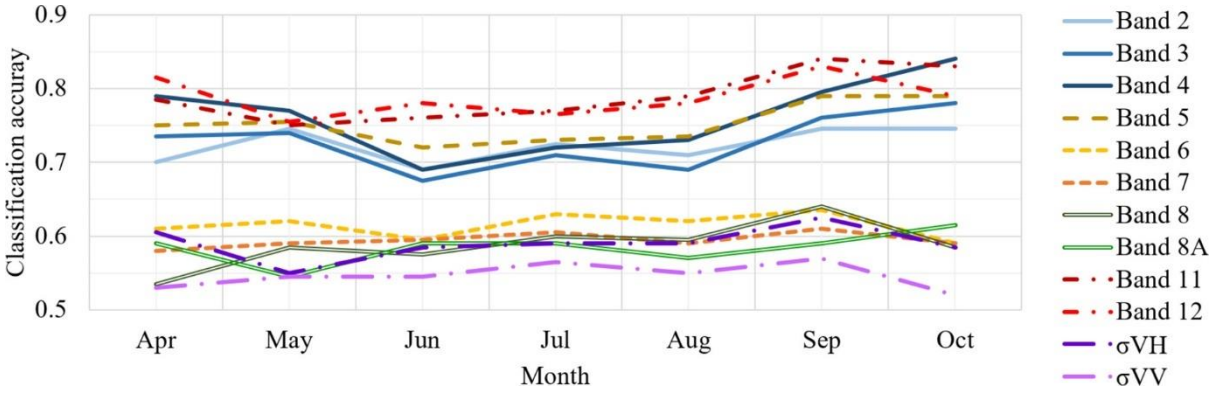


Figure 1: Graph from (Huo et al., 2021)(open access license CC BY) displaying separability of individual bands based on classification accuracy. The graph show that the SWIR bands (11 & 12) have the highest classification accuracy though the red band (4) and Red edge (5) separates themselves with high accuracy.

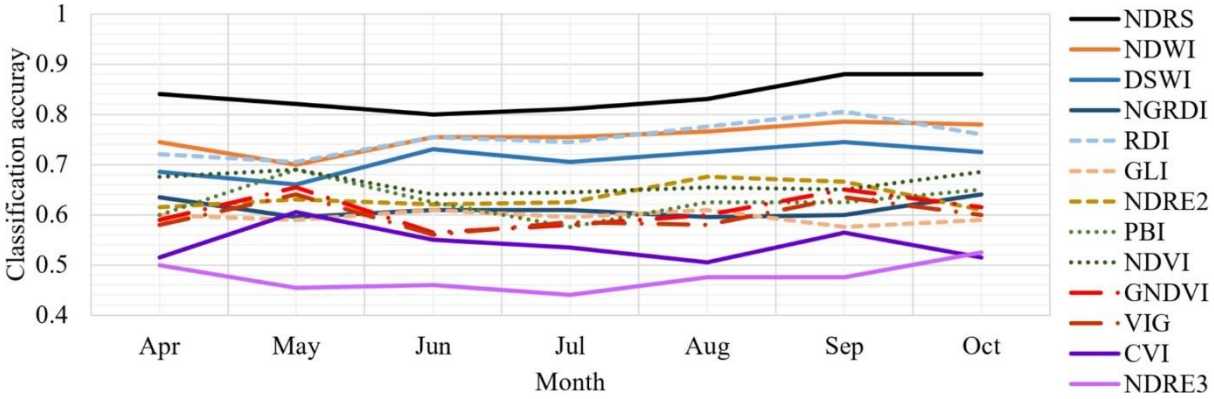


Figure 2: Graph from (Huo et al., 2021)(open access license CC BY) displaying separability of individual vegetation indices based on classification accuracy. Highest accuracy is seen from NDRS (combination of Red and SWIR) followed by indices with some utilization of the SWIR bands (NDWI, DSWI & RDI).

3. Materials and Methods

The method consisted of four main parts: Detection of areas of interest i.e., spruce forest stands (AOI), extraction of VI coefficient of variation from all AOIs, classification of AOIs into healthy and infested based on the variability from VI values and finding the optimal classification threshold for each VI (Figure 3). Additionally, Bark beetle swarming data was used to get an indication of when the bark beetles swarmed (Figure 3).

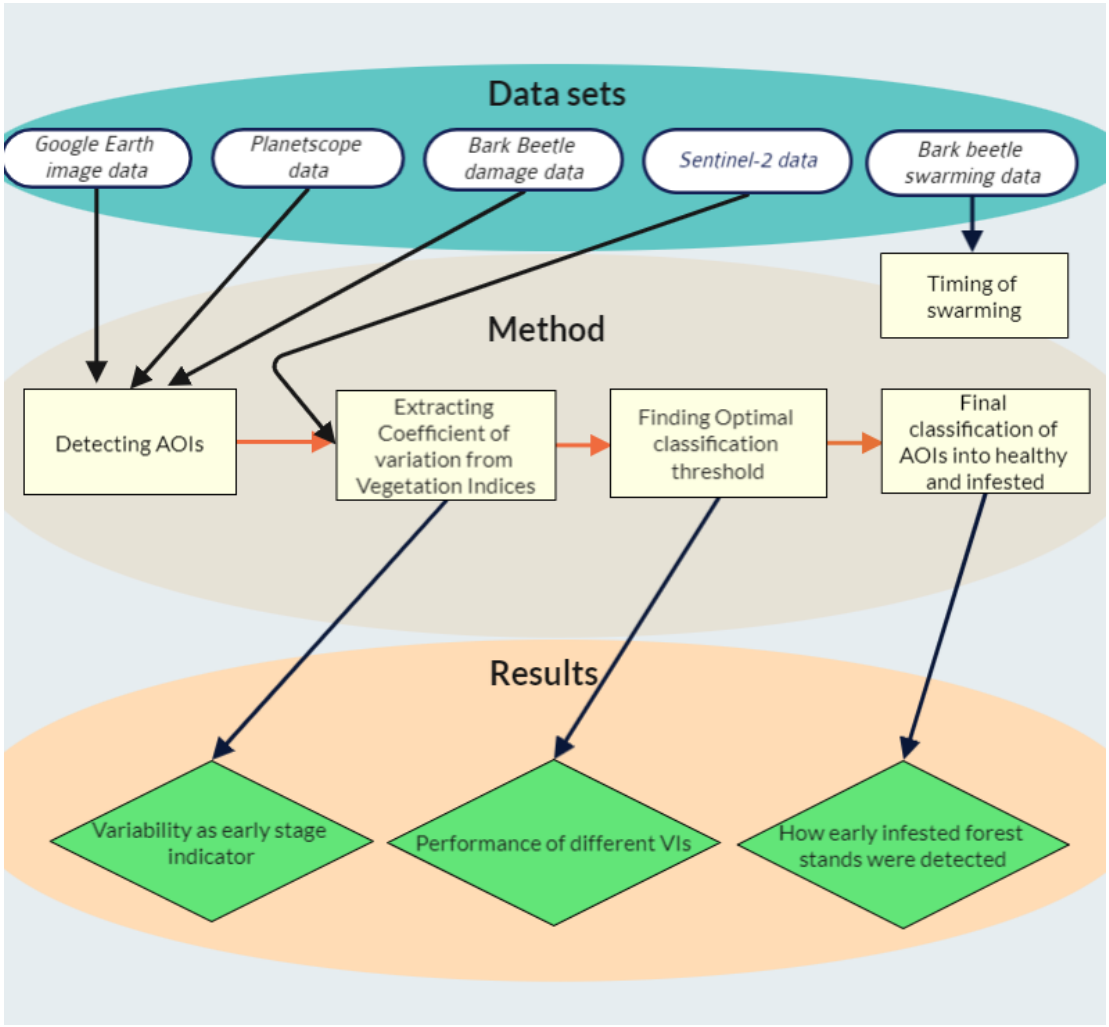


Figure 3: Flowchart of the workflow of the thesis with Data sets used, methodology and what the resulting output of the method.

3.1 Study area

The study area is in Östergötland county and located between Jönköping and Norrköping east of the lake Vättern (Figure 4). Östergötland county is located on the east coast of Sweden and is bordered by Baltic Sea in the east and Vättern in the west. The area has a temperate climate with an annual mean temperature of ca. 6 °C and regular snow cover in the winter, though along the coast the climate is more maritime with milder autumn/winter season (SMHI, 2012). Östergötland has a large forest industry with approximately 10 000 private owners and 60% of the land is covered by forest (Thomas Heldmark, 2020). The region has since 2018 experienced mass outbreaks of Spruce bark beetles.

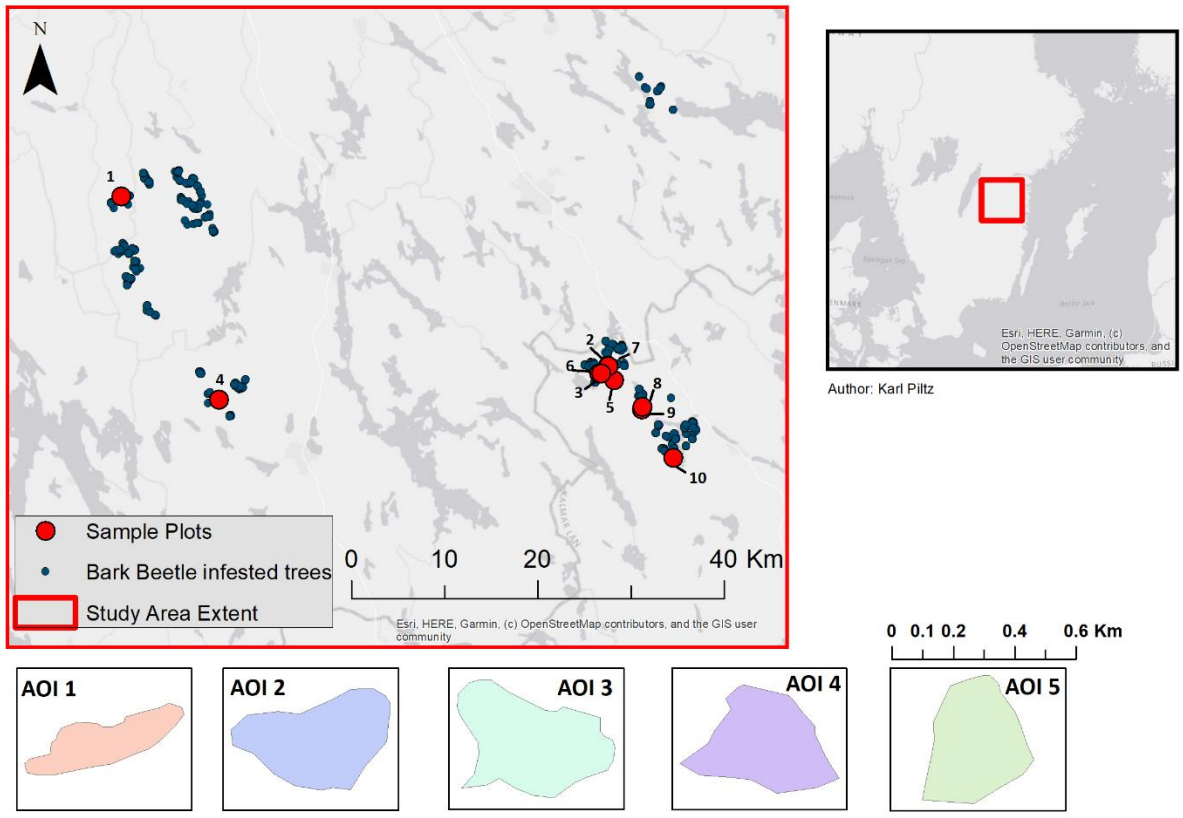


Figure 4: Location of the study area (upper-right) and the AOIs location within the study areas (upper-left) together with examples of different AOI extents and are represented by the sample plots in the upper left map (Shown are AOI 1 – 5) (bottom). (Source Sveaskog; Bark beetle damage data)

3.2 Data

The main datasets used in this study were Sentinel-2 data and bark beetle damage data (Table 1). Additional data sets used were imagery from Google (RGB) and satellite images from Planet (RGB, NIR) to identify areas of interest (AOI) and bark beetle swarming data to get an indication of when the bark beetles were swarming (Table 1).

Sentinel-2 data

Sentinel-2 data was used to compute the VIs that were applied in the study and to calculate the variability between pixels used in the time-series analysis for detecting bark beetle infestation. The Sentinel-2 data was in level 2 and was accessed through the Google Earth Engine Copernicus cloud storage collection which contained data from 31st of March 2017 and onwards. The study utilized four wavelength bands, Red (665 nm), NIR (865 nm), SWIR (1610 nm) and SWIR (2190 nm). Spatial resolution ranges from 10 m (Red) to 20 m (NIR and SWIR) (Table 1). Data was extracted for the period 31st of March 2017 – 31st of December 2018.

Bark beetle damage data

The bark beetle damage data was used to locate areas with large occurrence of bark beetle disturbance. The data set is point data distributed by Sveaskog, where each point represents trees that have been cut down due to bark beetle infestation as well as when it was felled (Table 1). The data was collected by Sveaskog since 2019 by tagging the location and time at which trees within their forests are cut down due to bark beetle infestation. The data used from the dataset was limited to trees cut down between April and May in 2019 as these would be the trees attacked at an early stage in 2018.

Google imagery/Planet to identify areas of interest (AOI)

Ten areas of interest (AOIs) to study the variability between pixels were selected within the study area. The areas were chosen based on certain preferences in order to be of use for the study. The area had to be homogenous spruce forest and must have had a bark beetle outbreak. Bark beetle infestation data from Sveaskog was used to identify forest stands with large occurrence of infestation. The areas of high occurrence were then observed using satellite imagery (Google earth and Planetscope) in ArcMap to determine whether the stand was homogenous spruce. The imagery from google earth was the latest updated web map service from Google satellite that was connected to ArcMap with tiles containing temporal coverage varying between 2017-2022. The Planetscope data cover was from April-May 2018 prior to when infested stands were cut down. When identifying areas with bark beetle infestation the search was limited to the year of 2018 because it was the first year with mass outbreak in recent time in Sweden and hence it can be certain that the forest was still intact i.e., not cut down at this point (due to bark beetle infestation). Additionally, seven areas without any recorded bark beetle activity were chosen to compare the variance response pattern between infested and non-infested forest stands and to be able to create a classification model (Table 1). The areas without recorded outbreaks were chosen the same way as stands that had been infested, by utilizing the bark beetle damage data from Sveaskog locating homogenous spruce stands without any record. These were chosen within the same study area with varying proximity to infested stands.

Bark beetle swarming data

Skogsstyrelsen (Swedish forest agency) use pheromone traps to perform weekly counts of the bark beetle population. In 2018, two stations were in active use in southern Sweden, one in Nässjö which was in the south-west of the study area, and another located in Lagan approximately 100 kilometres south of the study area. At both stations there were recorded high swarming activity predominantly in the end of April to mid-May as well as a smaller peak between mid-June to mid-July (Figure 5; Table 1).

Table 1: Data used in the study including source, type and spatial/spectral resolution.

DATA	TYPE	RESOLUTION	SOURCE
Bark beetle damage data	Point	Spatial: No Value Spectral: No Value	Sveaskog
Swarming data	Tabular	Spatial: No Value Spectral: No Value	Skogsstyrelsen (Swedish forest agency)
Sentinel-2	Raster Satellite imagery	Spatial: 10 - 20 m Spectral: 4 Bands	Copernicus
Google Earth	Raster Satellite imagery	Spatial: Ranging form 15m - 15cm Spectral 3 Bands	Google
Planetscope	Raster Satellite imagery	Spatial: 3 m Spectral 4 Bands	Planet

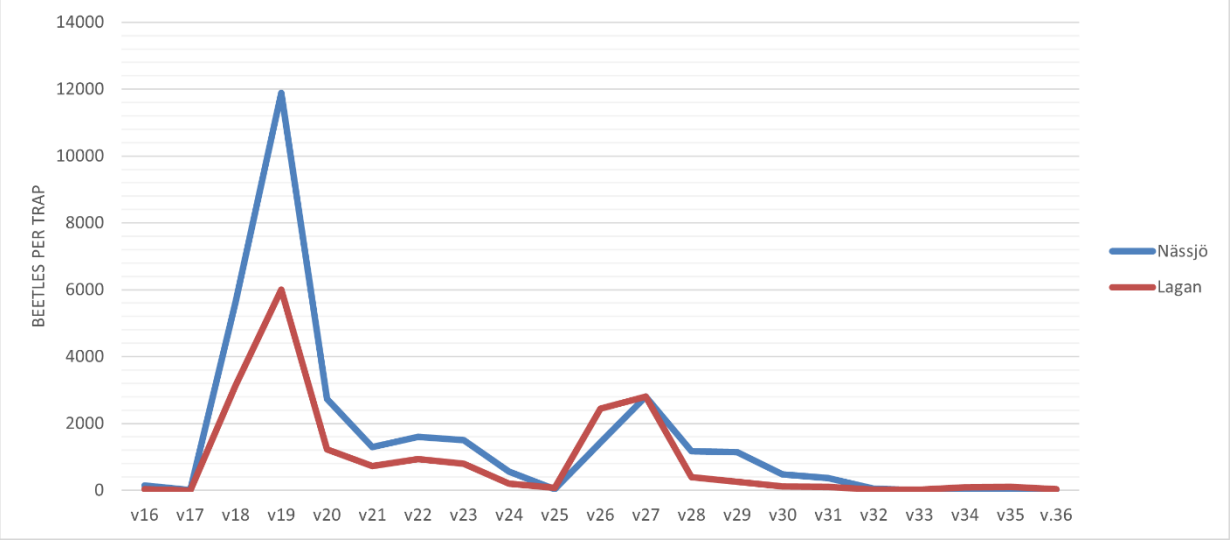


Figure 5: The number of beetles per trap for the two stations Nässjö and Lagan on a weekly basis 2018. (Skogsstyrelsen)

3.3 Vegetation indices

The VIs used in the time-series analysis were selected based on the literature review and resulted in the use of four VIs (NDVI, NDWI, DRS & RDI) with two of the indices sensitive to chlorophyll and two sensitive to water.

The Normalized Difference Vegetation Index (NDVI) was chosen as it has been one of the most used indices for vegetation monitoring to date (Gao, 1996). Though in recent empiric studies the classification of early-stage bark beetle infestation using NDVI has been found to be of moderate accuracy (60-70%) (Huo et al., 2021) it was chosen as it is a good baseline VI to utilize for the study as well as having a high spatial resolution of 10 m. NDVI was computed from the NIR (865 nm) and Red (665 nm) wavelength bands (Eq. 1). NDVI being a combination of NIR, and Red is sensitive to chlorophyll and cell structure related leaf alteration which makes it susceptible to detection of forest disturbances.

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad \text{(Eq. 1)}$$

The Normalized Difference Water Index (NDWI) was chosen as it has been found to reach high classification accuracy when mapping early-stage bark beetle infestation (Huo et al., 2021) as well as being a widely used index for monitoring of vegetation health (Gao, 1996). NDWI is computed using the NIR and SWIR bands and is sensitive to water related leaf alteration instead of chlorophyll and cell structure sensitivity (Gao, 1996). It was incorporated to the study partly based on previous results but also due to the combination of NIR and SWIR which have been heavily suggested in recent studies (Abdullah et al., 2018; Huo et al., 2021). Additionally, the sensitivity to water related changes makes the VI suitable for this study as both cause and effect of a bark beetle attack can be related to water content in leaf/tree as well as an instigating factor through periods of drought. NDWI was computed using SWIR (1610 nm) and NIR (865 nm) with a spatial resolution of 20 m, a combination which have been found highly effective when studying effects of bark beetle attacks (Eq. 2).

$$NDWI = \frac{NIR - SWIR}{NIR + SWIR} \quad \text{(Eq. 2)}$$

The Distance Red SWIR (DRS) index was chosen based on the results from (Huo et al., 2021) in which a new VI, NDRS, was computed with the intention of detecting early stage bark beetle attacks. NDRS is a normalized version of DRS and was in that study normalized using the minimum and maximum pixel value from pixels covering only homogenous spruce forest to make the VI sensitive to changes only related to spruce cover. NDRS was found to have the highest classification accuracy for both early and late-stage bark beetle attacks. Moreover, DRS was included in the study due to the combination of Red and SWIR which has been suggested due to its sensitivity of water content. As this study was done by studying individual homogenous spruce forest stands not including other surrounding land cover types, a normalization of DRS was not deemed beneficial. Hence, the study utilized solely DRS which was computed by a combination of the Red (665 nm) and SWIR (2190 nm) bands (Eq. 3).

$$DRS = \sqrt{(Red)^2 + (SWIR)^2} \quad \text{(Eq. 3)}$$

The Ratio Drought Index (RDI) was included in the study based on similar grounds as NDWI. RDI displayed high and stable classification accuracy for both early and late-stage bark beetle attacked trees and utilize a combination of NIR and SWIR which has been fruitful from previous recorded studies (Huo et al., 2021). Also, RDI is water content sensitive which has been found to generally be more sensitive to detecting bark beetle induced disturbance (Huo et al., 2021). RDI is computed by the combination of NIR (865 nm) and SWIR (2190 nm) with a spatial resolution of 20 m (Eq. 4).

$$RDI = \frac{SWIR}{NIR} \quad \text{(Eq. 4)}$$

3.3.1 Vegetation indices variability within areas of interest

The analysis of VI variability within the AOIs was done using Google Earth Engine (GEE). GEE is a cloud based geospatial computational platform which allows for interactive user activity connecting GIS and remote sensing processing to public access databases from a vast amount of sources (Gorelick et al., 2017). GEE was chosen for the analysis of this study because it allows for easy access and extraction of Sentinel-2 satellite data from the Copernicus database with coverage since 2017. Furthermore, the application is widely used for remote sensing processing and analysis with effective computation of indices, cloud filtering and pre-processing of imagery.

GEE was used to calculate the coefficient of variation of pixel values from different VIs computed from Sentinel-2 images in a time-series. The coefficient of variation was the chosen metric ahead of variance to minimize the effect of seasonal variation and was calculated according to Eq. 5:

$$Coefficient\ of\ Variation = \frac{\sigma}{\mu} \quad \text{(Eq. 5)}$$

Where σ is the spatial standard deviation and μ is the spatial mean.

The script used, computes VI time-series from Sentinel-2 imagery using polygon inputs. The script was created using a python package called eemont which extends the application of python in GEE for pre-processing and VI computation of Sentinel-2 image collections (Montero, 2021). The script was divided into 3 parts: the first part did cloud masking on each image of the time-series, the second part calculated VIs from the Sentinel-2 image collection, and the third part extracted the time-series with zonal statistical computations from the VIs in the image collection.

The cloud masking was done using the s2Cloudless dataset which calculated cloud probability and using a cloud filter. Cloud shadows masking was done by identifying dark pixels in the NIR surface reflectance. Firstly, the mask parameters were set, allowance of 50% cloud cover per image was set, cloud probability of 30% or higher was set to be defined as clouds, a dark pixel threshold where a digital number of 0.15 or less was counted as potential cloud shadows and then a maximum search distance of 2 km for shadows from the edge of an identified cloud

feature. The mask parameters were then transformed into band components and used as a filter which was applied on each Sentinel-2 image in the time-series eliminating images with more than 50% cloud cover, masking pixels with larger than 30% cloud probability and masking dark pixels identified as shadows.

VIs were computed from the eemont-package library (github.com/davemlz/awesome-ee-spectral-indices) as well as using a function to compute user defined VIs. NDVI, NDWI, RDI and DRS were computed and transformed into band components which information could be extracted from in the next stage.

Extraction of time-series was done with zonal statistics using the polygon of the AOIs as input features to get results from the different bands. Here the Coefficient of variation was calculated and extracted from the different VIs for the time-series (31st March 2017 – 31st December 2018). The results were then plotted in resulting in graphs displaying the coefficient of variation (%) for NDVI, NDWI, DRS and RDI over the time-series for each AOI.

The time-series were cleaned up by eliminating values during the winter and snow season from October to April when anomalies where present due to snow cover and low solar angles.

3.4 Classifying AOIs into attacked and healthy

To answer whether it is possible to identify bark beetle outbreaks using variability (Coefficient of Variation) from the extracted VIs, the results had to be classified into attacked and healthy and then validated in order to evaluate the performance of each VI. The classification was based on cumulative sums of the spatial variability of the time-series and Receiver Operating Characteristics (ROC) graphs were used to find the optimal threshold for the classification.

3.4.1 Detecting AOIs with bark beetle infestation

To use the coefficient of variation to detect AOIs with bark beetle outbreak, the values in the time-series was standardized for each VI at every AOI. This was done to eliminate local variation of the different AOIs and to be able to compare the results consistently. The standardization was done using the mean and standard deviation from a stable period i.e., when it is assumed to not display spectral alteration due to bark beetles (Eq. 6). The stable period chosen was between 31st of March 2017 (prior to the mass outbreak in 2018) and 30th April 2018 (prior to first recorded swarming activity in 2018 (Figure 6)). The choice of using a single year (2017 until the first recorded swarming activity in 2018) was done as the earliest Sentinel-2 data available through Google earth engines archive was in March 2017.

$$Standardized\ coeff\ of\ Variation = \frac{Coefficient\ of\ variation - mean(stable\ period)}{Standard\ deviation(stable\ period)} \quad \text{(Eq. 6)}$$

The standardized value at each point in the time-series had either a negative or positive value where the positive values were assumed to be increase of spatial variation within the AOI due to bark beetle infestation. To classify the AOIs into attacked or healthy the cumulative sum was calculated according to Eq. 7:

$$Cusum(t) = Cusum(t-1) + \max(CV_{stand}(t), 0) \quad \text{(Eq. 7)}$$

$$C(0) = 0$$

Where *Cusum* is the cumulative sum, CV_{stand} is the standardized coefficient of variation and *t* is the time step i.e., all negative values were turned to a value of 0 and all positive values from the point of the first recorded bark beetle activity was summed up. This was done for each VI at every AOI and the AOIs were then classified into attacked or healthy based on the threshold of the cumulative sum.

This method was applied for two time periods, first using the cumulative sum and threshold values representing the whole time period (30th April – September) and second for the early stage of the time-series (30th April – 1st July) in order to see if the use of threshold from early or whole period result in better output.

3.4.2 Finding the optimal threshold for classification

The evaluation to find the optimal threshold for the classification was done using receiver operating characteristics graphs (ROC) which is a method that evaluates the performance of a binary classification and visualises it with a curve (Fawcett, 2006). ROC was the chosen method as it is useful when classifying and evaluation of binary classifiers (Fawcett, 2006). Additionally, it has shown promise in a previous study where the cumulative sum and ROC was used to detect birch forest defoliation from insect outbreaks in northern Sweden with the use of MODIS derived NDVI (Olsson et al., 2016).

The ROC graph and curve is computed from a classification model with a set of training data that are to be given a class membership and a set of evaluation data that are used to validate the classification. The classification model is then evaluated on the same basis as a confusion matrix where each value can either be evaluated as a True positive (positive classified as positive), False negative (positive classified as negative), True negative (negative classified as negative) and False positive (negative classified as positive). The ROC graph then plots the True positive rate (Tpr; Eq. 8) on the y-axis against the False positive rate (Fpr; Eq. 9) on the x-axis which display the trade-off between positives classified correctly and negatives classified as positive between 0 - 1.0 for each threshold used in the classification (Fawcett, 2006). The ROC-curve hence displays the classification model performance at each threshold which allow the user to find the optimal threshold to use as well as comparing different models against each other where a perfect classification is (Fpr=0, Tpr=1) i.e. finding all true positives without misclassifying any negatives as positive (Figure 6) (Fawcett, 2006). The thresholds which were tested were derived from the predicted dataset using a decision function which tests all possible thresholds against FPR and TPR in which it drops suboptimal thresholds i.e. thresholds that result no instances being classified (Pedregosa et al., 2011).

$$True\ positive\ rate = \frac{True\ Positives}{Total\ Positives} \quad \text{(Eq. 8)}$$

$$False\ postitive\ rate = \frac{False\ Positives}{Total\ Negatives} \quad \text{(Eq. 9)}$$

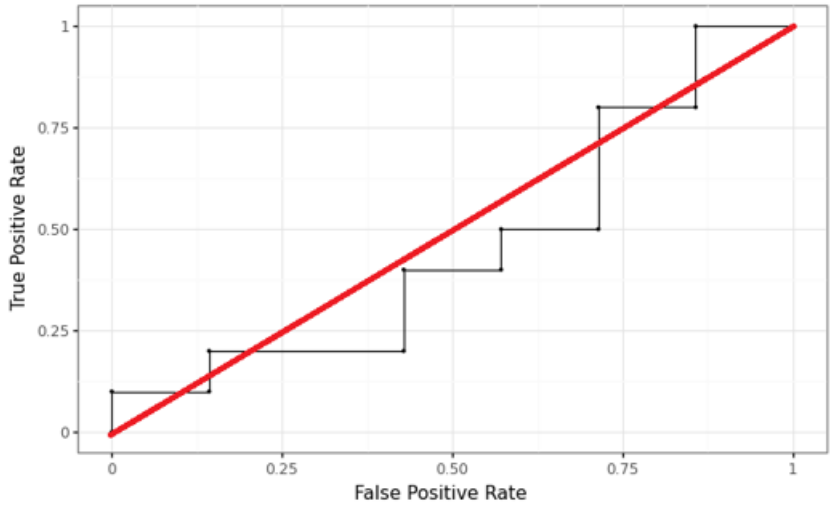


Figure 6: Illustration of ROC graph with true positive rate on the y-axis, false positive rat on the x-axis, ROC-curve displaying the TPR & FPR for each classification threshold and a red line displaying a random classification. ROC-curve above red line is better than random and below is worse than random.

ROC-graphs were created to evaluate the performance of each VI using the cumulative sum of the standardized coefficient of variation as training data and Sveaskogs infestation data as evaluation data. Both training and evaluation data were derived from the 17 AOIs where 10 had recorded bark beetle outbreaks and 7 had no recorded outbreak. Finally, the AOIs were classified into infested and healthy based on the threshold. The evaluation of performance can be done through multiple methods i.e., ROC-curve distance from perfect classification (0,1) or distance to random classification (y=x) however in this study the performance was evaluated by the trade-off between Tpr and Fpr e.g., largest difference between Tpr and Fpr (Figure 6) (Fawcett, 2006).

In order to see how early in the season each VI can identify bark beetle attacks, the optimal threshold was chosen based on the trade-off between true positive rate and false positive rate. The optimal threshold was then used to compare to when the cumulative sum of the standardized positive coefficient of variation for each index reach this threshold. Where the cumulative sum reached the optimal threshold is the earliest detection.

4. Results

4.1 Coefficient of variation

The result from the GEE computations of the coefficient of variation (%) for each of the 10 AOIs with recorded bark beetle infestation (AOI 1-10; Section 4.1.1) and the 7 AOIs without bark beetle infestation (Section 4.1.2) are presented separately.

4.1.1 Coefficient of Variation for AOIs with Bark Beetle Infestation

The coefficient of variation for the AOIs are in the results section presented by four graphs that are representable for the response in variation for all AOIs to make the results more easily readable. All graphs of coefficients of variation are found in the appendix (A1, A2 & A3).

The coefficient of variation for AOIs with bark beetle infestation in general display a well-defined peak in April-May 2018 for all VIs around the same time as the first recorded swarming activity the same year (Figure 7, AOI 2 & AOI 8). The coefficient of variation seemingly increases during 2018 and have a defined peak/increase in July around the timing of the second swarming event in the same period in 2018 (Figure 7, AOI 2, AOI 8 & AOI 7). DRS and RDI display larger coefficient values between 10-15% as in AOI 2 but instances of higher values are found in some AOIs also (Figure 7, AOI 8) (A1, A2, A3). NDVI and NDWI display a less fluctuating coefficient with a clearer coefficient increase and defined peak, especially in the later part of the season with a coefficient ranging between 4-6% (Figure 7, AOI 2, AOI 7). Additionally, it can be seen for many AOIs that the coefficient values are generally higher in 2018 than in 2018 as seen in AOI 7 (Figure 7). In some cases, however, the peak found early 2018 (April-May) is not present or not as well defined like in AOI 7 (Figure 7). Also, instances of highly fluctuating DRS and RDI is found where they seemingly display no pattern and are highly variable over the whole period (AOI 1) (Figure 7).

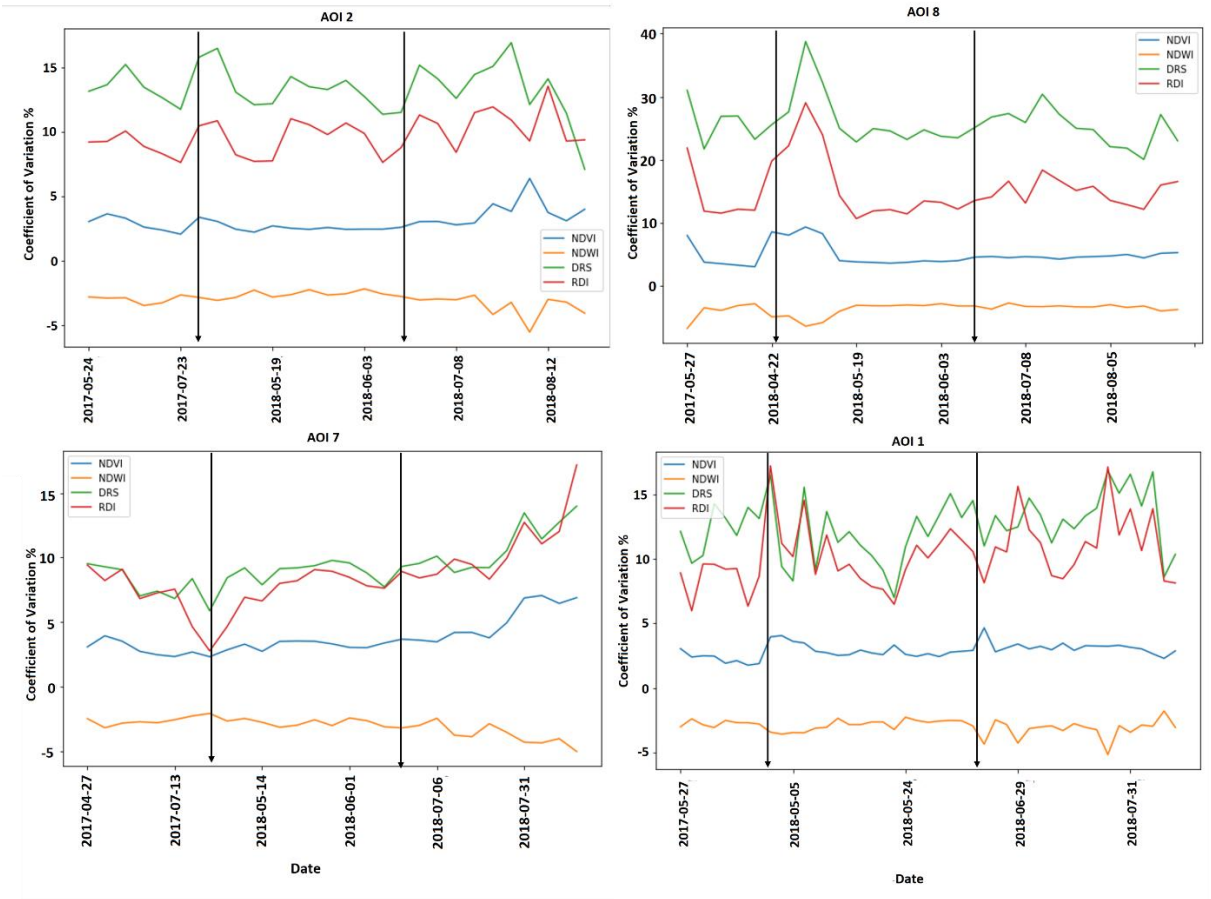


Figure 7: Coefficient of variation (%) for AOIs 2, 8, 7 & 1 between 2017-2018, where green curve is DRS, red is RDI, blue is NDVI and yellow is NDWI. Black arrows are the timing of bark beetle swarming activity in 2018 where the first arrow is on 23rd of April and the second on the 18th of June. (Skogsstyrelsen 2021)

4.1.2 Coefficient of Variation for areas without Bark Beetle Infestation

The coefficient of variation for AOIs without recorded bark beetle outbreaks (1-7) are presented in the results section by four graphs that are representable for the response in variation for all AOIs to make the results more easily readable. All graphs of coefficients of variation are found in the appendix (A4, A5).

The coefficient of variation for AOIs without any recorded bark beetle outbreak display low coefficient values for al VIs (Figure 8). NDVI and NDWI are generally more stable over the whole period with a coefficient around (+-) 2% while DRS and RDI are more variable ranging between 4-10%. For most AOIs there are not found any peak over the period or in the timing of the swarming activity in 2018 (AOI 6 & AOI 7) (Figure 8). The coefficient is seemingly not higher in 2018 than in 2017 with lower values in AOI 6 and no visible difference in AOI 7 (Figure 8). Slight increase in coefficient can be seen in the AOIs during 2018 which can be seen for NDVI and NDWI in AOI 7, AOI 2 and AOI 5 (Figure 8). Not applicable for all AOIs without bark beetle infestation as can be seen in AOI 2 which display a clear increase around the first swarming event and with generally higher values in 2018 than in 2017 (Figure 8). Similar can be found in AOI 5 which display a somewhat defined early peak for NDVI and NDWI while having a clear increase after the second swarming event in 2018 for DRS and RDI (Figure 8).

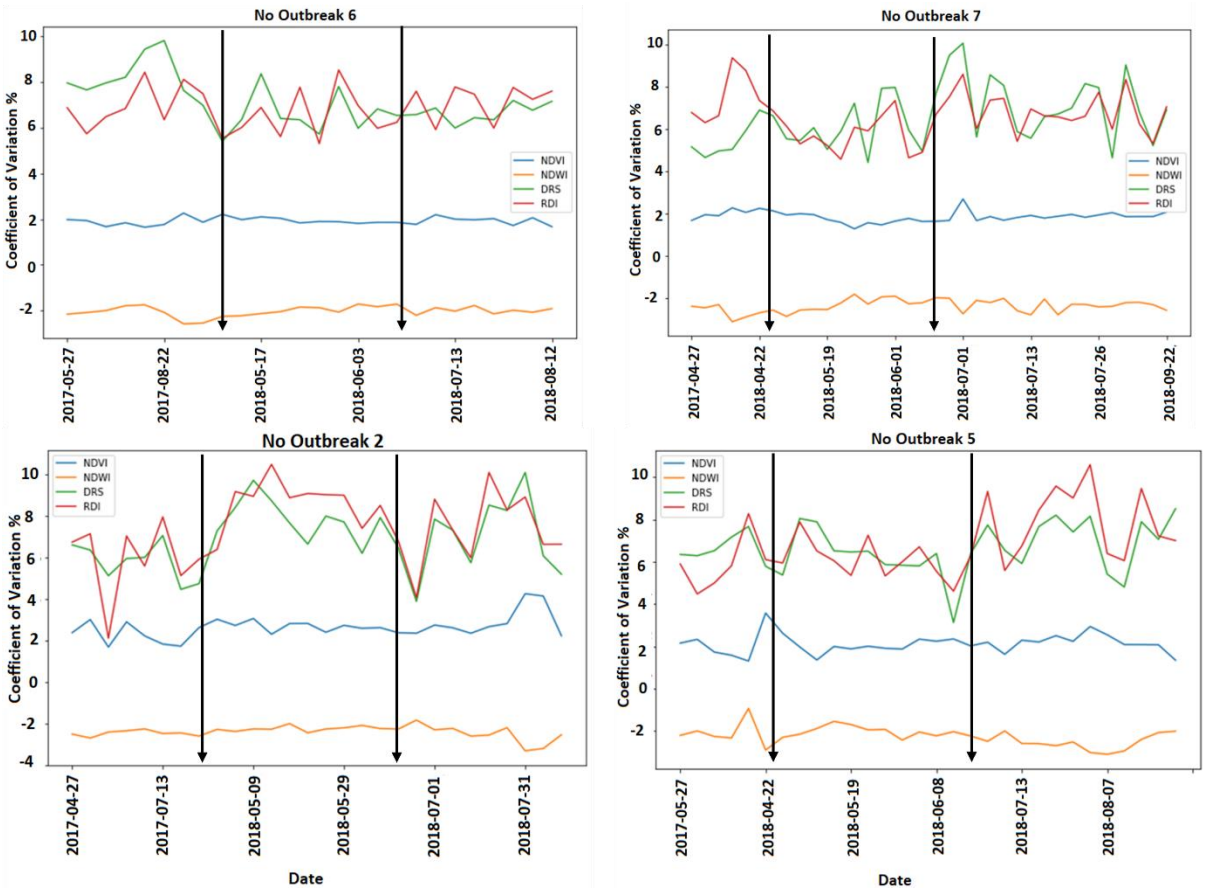


Figure 8: Coefficient of variation (%) for AOIs 6, 7, 2 & 5 without bark beetle infestation between 2017-2018, where green curve is DRS, red is RDI, blue is NDVI and yellow is NDWI. Black arrows are the timing of bark beetle swarming activity in 2018 where the first arrow is on 23rd of April and the second on the 18th of June. (Skogsstyrelsen 2021)

4.2 Finding the optimal thresholds for classification

The analysis to find the optimal threshold for the classification into attacked and healthy areas with ROC-curves resulted in one ROC-graph for each VI for both the whole Bark beetle season (from April 30th to September 30th, 2018) and for the early stage of the bark beetle season (1st May-1st July 2018)

4.2.1 ROC-curve for the whole period

The ROC-curves for the whole season show that the classification model of DRS had a poor performance with a low rate of true positives and high rate of false positives. The best performance was found for a threshold value of 6.68 where the classification had a Tpr of 0.9 and an Fpr of 0.71 (Figure 9). The ROC-curve for NDWI obtained a similar performance as DRS with an optimal threshold of 5.75 with a Tpr of 0.9 and an Fpr of 0.71 (Figure 9). The performance of NDVI was better with a Tpr of 0.8 and an Fpr of 0.14 with the optimal threshold 14.91 (Figure 9). RDI also performed better than DRS and NDWI with a Tpr of 0.8 and an Fpr 0.28 with a threshold of 10.38 (Figure 9)

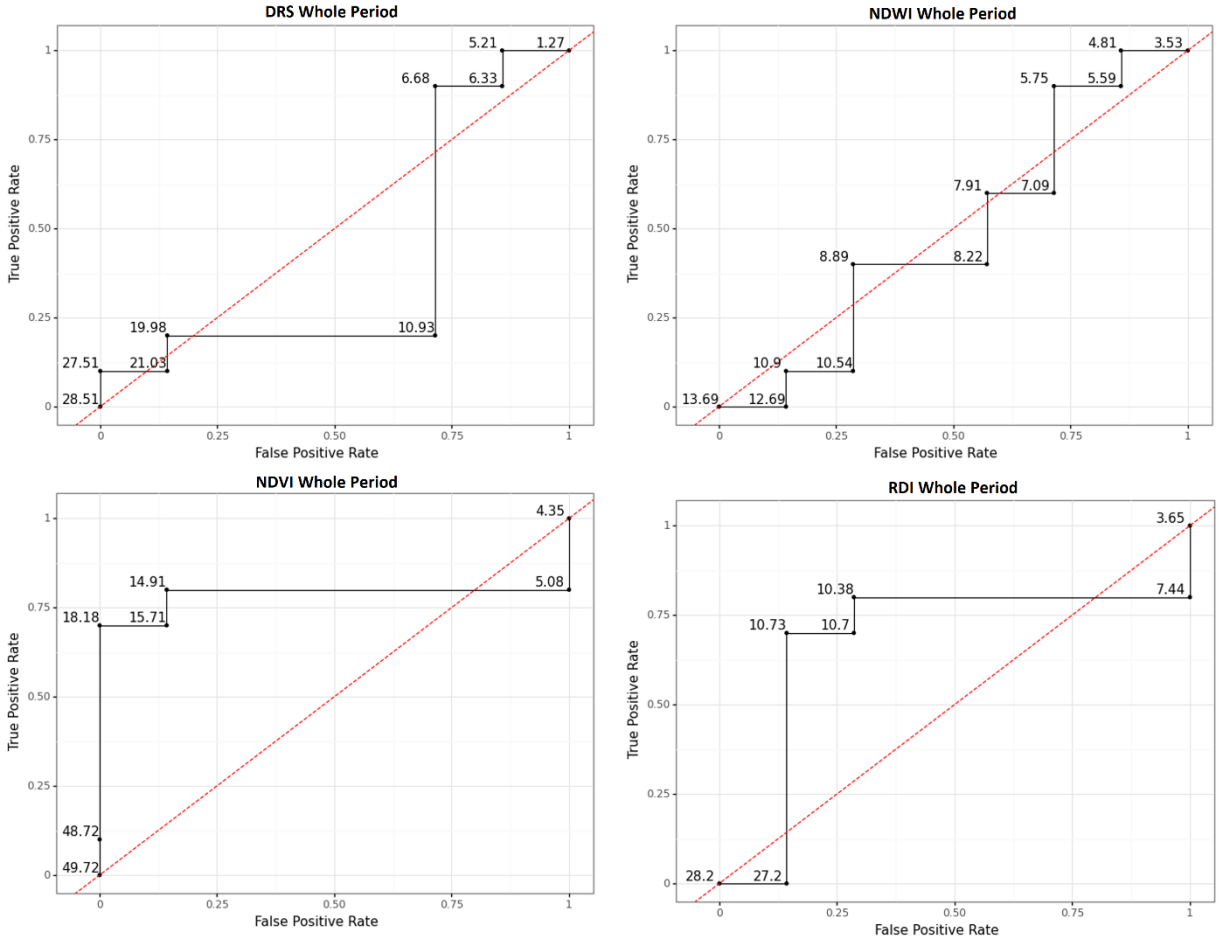


Figure 9: ROC-Curves for DRS, NDWI, NDVI and RDI for the whole bark beetle season (April 30th – September 30th). The curve displays the performance of each threshold measured in the trade-off between Tpr and Fpr. Red dashed line represents a random classification.

When applying the optimal threshold for each VI to perform the classification the earliest detection of bark beetle infestation for NDVI was 13th of July (AOI 4), the latest detected area was AOI 1 at 8th of August with an average detection date around 30th July (Table 2). NDWI detected the earliest outbreak on the 19th of May (AOI 6), latest 10th August (AOI 3) with and average detection date around the 19th of June (Table 2). The earliest detection date for DRS was on the 23rd of June (AOI 1), the latest on 16th of September (AOI 6) and average was 23rd of July (Table 2). RDI detected the first outbreak on the 16th of July (AOI 2), latest on the 19th of September (AOI 10) with an average around 2nd of August (Table 2).

Table 2: The dates on which the optimal threshold detected an outbreak in every AOI for each VI over the whole period (April 30th – September 30th). Additional information: Tpr, Fpr, earliest detection, latest detection and average date of detection.

Whole Period	NDVI	NDWI	DRS	RDI
Optimal Threshold	14.91	5.75	6.68	10.37
TPR	0.8	0.9	0.9	0.8
FPR	0.14	0.71	0.71	0.28
AOI 1	08-Aug	30-May	23-Jun	21-Jul
AOI 2	02-Aug	08-Jun	26-Jul	16-Jul
AOI 3	31-Jul	10-Aug	31-Jul	02.Aug
AOI 4	13-Jul	01-Jun	21-Jul	26-Jul
AOI 5	No Detection	08-Jun	No detection	No Detection
AOI 6	05-Aug	19-May	16-Sep	No Detection
AOI 7	26-Jul	06-Jul	06-Jul	31-Jul
AOI 8	No Detection	No Detection	13-Jul	No Detection
AOI 9	07-Aug	06-Jul	06-Jul	26-Jul
AOI 10	31-Jul	01-Jul	02-Aug	19-Sep
Earliest	13-Jul	19-May	23-Jun	16-Jul
Latest	08-Aug	10-Aug	16-Sep	19-Sep
Average	30-Jul	19-Jun	23-Jul	02-Aug

4.2.2 ROC-Curve for the period May-July

The ROC-curve of the for early season (1st May-1st July) show that the classification model of DRS performed poorly with a trade-off of 0.6 Tpr and 0.42 Fpr at an optimal threshold of 3.66 (Figure 10). NDWI had a better performance with an optimal threshold of 4.8 where the model was able to correctly identify all true positives while having an Fpr of 0.71 (Figure 10). The performance of NDVI was similar to DRS, obtaining a Tpr of 0.6 and an Fpr of 0.71 with an optimal threshold of 3.28 (Figure 10). The ROC-Curve of RDI obtained slightly better performance than NDVI and DRS with a Tpr of 0.7 and an Fpr 0.42 with optimal threshold being 4.36 (Figure 10).

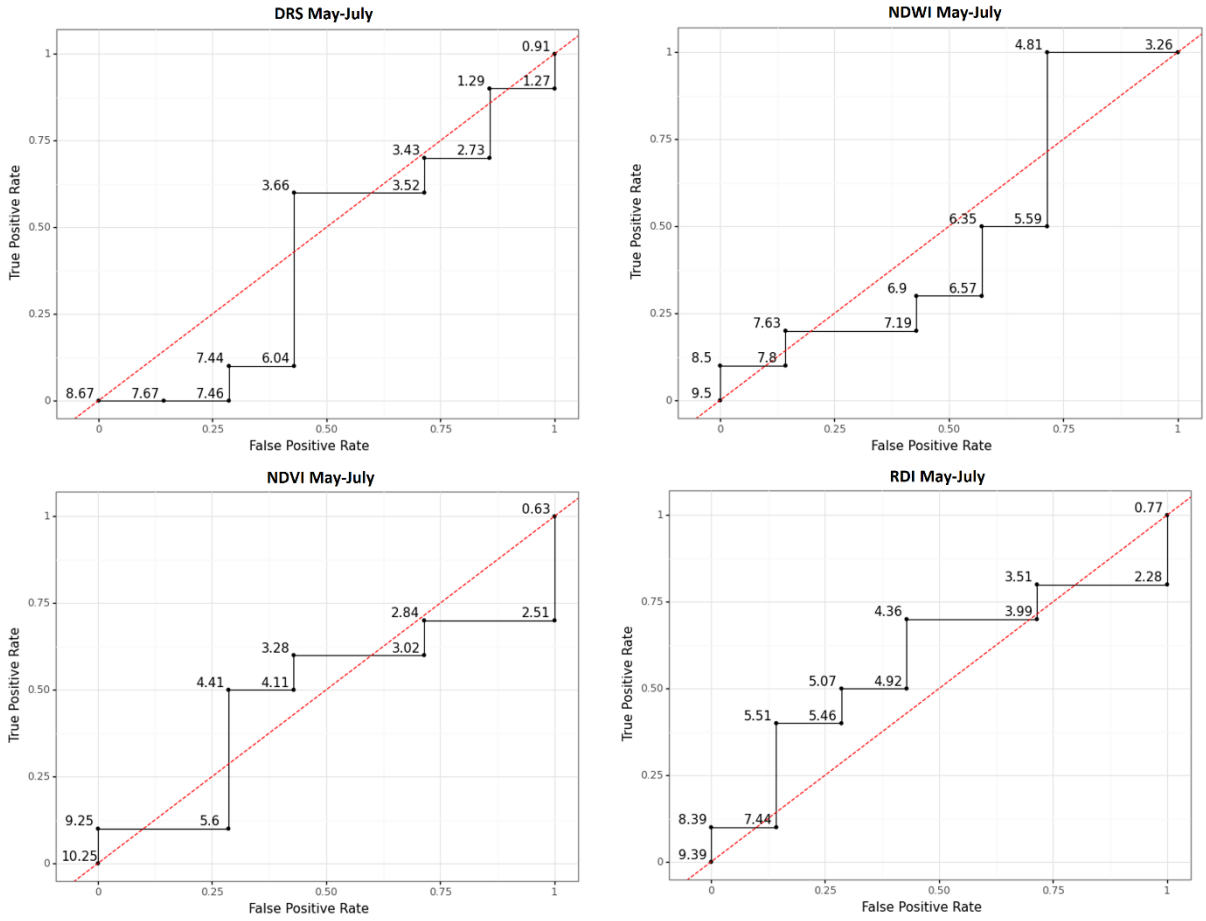


Figure 10: ROC-Curves for DRS, NDWI, NDVI and RDI for the early-stage bark beetle season (1st May-1st July). The curve displays the performance of each threshold measured in the trade-off between Tpr and Fpr. Red dashed line represents a random classification.

When performing the classification for each VI by applying the optimal threshold it was found that for NDVI the earliest detected bark beetle infestation was on the 7th of May (AOI 1), the latest on the 28th of June (AOI 6) and with an average detection date around the 1st of June (Table 3). NDWI detected the infestation earliest on the 14th of May (AOI 6), the latest on the 28th of June (AOI 8) with an average around the 1st of June (Table 3). DRS detected the first outbreak on the 9th of May (AOI 8), latest on the 23rd of June (AOI 4) with an average around the 1st of June (Table 3). Earliest RDI was detected on the 9th of May (AOI 8), latest on the 28th of June (AOI 7) with an average around the 3rd of June (Table 3).

Table 3: The dates on which the optimal threshold detected an outbreak in every AOI for each VI (1st May-1st July). Additional information: Tpr, Fpr, earliest detection, latest detection and average date of detection.

May - July	NDVI	NDWI	DRS	RDI
Optimal Threshold	3.28	4.8	3.65	4.36
TPR	0.6	1	0.6	0.7
FPR	0.42	0.71	0.42	0.42
AOI 1	07-May	29-May	01-Jun	01-Jun
AOI 2	No Detection	03-Jun	No Detection	01-Jun
AOI 3	No Detection	29-May	No Detection	No Detection
AOI 4	23-Jun	29-May	23-Jun	23-Jun
AOI 5	No Detection	01-Jun	No Detection	No Detection
AOI 6	28-Jun	14-May	No Detection	No Detection
AOI 7	29-May	03-Jun	29-May	28-Jun
AOI 8	09-May	28-Jun	09-May	09-May
AOI 9	03-Jun	01-Jun	03-Jun	03-Jun
AOI 10	No Detection	03-Jun	03-Jun	24-May
Earliest	07-May	14-May	09-May	09-May
Latest	28-Jun	28-Jun	23-Jun	28-Jun
Average	01-Jun	01-Jun	01-Jun	03-Jun

5. Discussion

5.1 Coefficient of variation

The results indicate that there is a change in variability between pixels within a stand (AOI) that was attacked by bark beetles during the outbreak year 2018 (Section 4.1.1). For most of the AOIs there seems to be a difference between 2017, the year prior to the outbreak and 2018 in the time-series of coefficient of variation, with an increase in variability, with two peaks in April-May and around late-June-August in 2018, for most AOIs, where the first peak generally is larger. Important to note is that there is a gap in the time-series between October 2017 and April 2018 due to noise (clouds, snow, short days) which could influence the observed increase between 2017 and 2018. However, despite this, the peaks found in 2018 are not present to the same extent in 2017 and the coefficient of variation is generally larger from June-September in 2018 than in 2017. The first peak found around April-May in 2018 is a positive indicator that the variation displayed by the coefficient of variation could be bark beetle induced as the swarming data tells us that the first large swarming outbreaks predominantly occurred between the start of April to mid-May (Section 3.2, Figure 5). Similarly, the peak/increase later in the season from late-June-August could be connected to the peak found in the Swarming data during the same time (Section 3.2, Figure 5). However, it is difficult to conclude that the change in variability is due to bark beetle infestation. The increasing variability might also be caused by the drought in 2018, but since the results indicate that there is a larger increase in variability for bark beetle infested AOIs it is likely that the actual infestation contributes to the higher coefficient of variation. Additionally, if the change in variability was due to other factors like the drought of 2018, the results would still valuable and interesting as when comparing with another study by (Huo et al., 2021) where the intention was to detect stress symptoms prior to an outbreak. Here it was discovered that using stress symptoms as early warning systems to classify potential outbreak areas reached rather high accuracy for detecting later bark beetle outbreaks.

The results indicate that the coefficient of variation is generally lower for the areas without recorded bark beetle outbreak compared to AOIs with bark beetle activity (Section 4.1.2). The difference is more pronounced for NDVI and NDWI which have less fluctuation in the coefficient of variation. Also notable, is that for the areas without recorded bark beetle outbreak the coefficient of variation is more stable in the time-series and does not show a large a difference between 2017 and 2018 as was generally found for the AOIs with bark beetle activity (with some exceptions; area 2, 3 & 5) (Figure 8, A4). Though the general trend implies that the coefficient of variation increased during 2018, there are no well-defined peaks at the early and late stage of 2018 (for most areas) as was found in AOIs with bark beetle activity. Hence, this could instead be an indication that the increase is not bark beetle induced as when comparing infested AOIs with the timing of swarming, a large signal for both variability and swarming was found within the same timeframe. The same response was not found for healthy AOIs. More plausible is that the more stable increase in the coefficient of variation for areas with no infestation record could be due to the major drought and temperature that the area experienced during 2018 (Wilcke et al., 2020). Comparing the coefficient of variation response in areas with and without recorded bark beetle infestations, it could be said that though there are some examples of inconclusive results where areas with supposedly no infestation display some similar trends as areas with bark beetle occurrence, the consensus is that the response is different and that the variation between pixel VI values change and have the potential to

indicate bark beetle infestations. Due to the short length of the time-series no statistical tests were done to test if the change in variability were significant or not. The comparison of variability was made purely through visual analysis. Instead, the indication that the variability did change during the bark beetle infestation in 2018 comes from the classification, since it was possible classify the AOIs into attacked and healthy forest stands.

5.2 Finding optimal thresholds with ROC-Curves

The results from the different VI classification models displayed varying performance (Section 4.2.1; Figure 12). Looking at the performance for the whole period, NDVI and RDI performed better than DRS and NDWI, with NDVI as the best performing classifier. DRS and NDWI were able to correctly classify AOIs infested by bark beetles at a rate of 0.9 while misclassifying areas with no recorded outbreak by 0.71. This is a trade-off which is just slightly better than random and does not sufficiently detect bark beetle affected forest stands without misclassifying a large number of healthy areas. NDVI and RDI were able to correctly classify infested forest stands with a rate of 0.8 while misclassifying areas without outbreaks as infested by 0.14 (NDVI) and 0.28 (RDI). The trade-off for both NDVI and RDI could be said to be acceptable as the goal is to detect forest stands that have been infested, hence the misclassification of some stands that have not experienced an outbreak is less of a problem if the result of the method is to mitigate bark beetle disturbance. The performance of NDVI compared to the other Vis could potentially be due to that the variation would increase over a longer period as the bark beetle attack persists, in which case the higher spatial resolution of 10 m might be beneficial compared to the 20 m of DRS, RDI and NDWI.

For the early-stage period (1st May – 1st July), NDWI performed best i.e., the best trade-off between Tpr and Fpr, correctly detecting and classifying all forest stands affected by bark beetles while misclassifying healthy stands at a rate of 0.71 (Section 4.2.2). The classification models for NDVI, DRS and RDI performed worse than NDWI for the early period with a generally poor classification accuracy and are hence not reliable classifiers in this case. The performance of NDWI show great potential as it correctly classified all infested stands and could be deemed acceptable even if having an Fpr of 0.71 as the importance of detecting infested stands overweight the risk of misclassifying healthy stands. By detecting all areas with bark beetle infestation, it heavily insinuated that the response in NDWI is due to bark beetle outbreaks. Though not possible to say for certain as the signal could also be due to the drought, however, more likely by comparing to swarming data is that the signal is bark beetle induced. Why NDWI performed the best for the early period could potentially be due its high sensitivity to water content, at the early stage of an infestation water content plays a major factor. Bark beetles are both drawn to water stressed hosts as well as cutting off water supplies within the tree when attacking it (Abdullah, Skidmore, et al., 2019a; Huo et al., 2021). Especially when in the early stage where the foliage is still spectrally unaffected, the sensitivity to water of NDWI could play a key part. RDI and DRS performed poorly potentially due to the very large fluctuation in the variation coefficient (Section 4.1.1, Section 4.1.2) which in a shorter timeframe negatively affect the separability of areas with and without bark beetle outbreaks.

5.3 Timing of detection

Using the optimal threshold from NDVI and RDI for the model over the whole season show that the earliest detection date for an outbreak was 13th July for NDVI (AOI 4) and 16th July for RDI (AOI 2) (Section 4.2.1). The average detection date for NDVI was the 30th of July which is a 2-week difference from the earliest date. The earliest date of detection with NDVI

(AOI 4) stands out compared to the detection dates for the other AOIs (NDVI) which were detected at dates closer to the average (30th of July), hence, it is possible that the detection date of AOI 4 is a bit overestimated due to anomalies. More reasonable is that the earliest detection with NDVI occurred around the 20th of July. This is plausible as the green attack stage which extends to July would have a stronger response signal around this time which would be possible for NDVI to detect (Huo et al., 2021). RDI had an average detection date around the 2nd of August with most AOIs detected closer to the average than the 16th of July. Hence it is in this case also likely that this date have been overestimated and in reality would be closer to the end of July while the bark beetle attack is transitioning into red stage (Huo et al., 2021).

The earliest detection date for NDWI in the early stage of the season was on the 14th of May, which is possibly also somewhat overestimated as the other AOIs (NDWI) were detected closer to the average detection date 1st of June. Hence it is more likely that the earliest detection date for NDWI is closer to the 29th of May. This result is important as the date coincides in the middle of the green attack-stage, which would point to that the bark beetles have yet to leave their hosts, increasing the probability of successful mitigation for the forestry industries stakeholders (Abdullah, Darvishzadeh, et al., 2019).

5.4 Sources of error

During the analysis of this thesis there are sources of error to consider when interpreting the results. Firstly, the lengths of the time-series in this study are rather short due to Sentinel-2 data only being available from the 30th of March 2017 when using the Copernicus database via Google earth engine. Additionally, images during late autumn and winter were dismissed due to noise from snow cover and low solar angle which created gaps in the time-series. This influenced the study as there was not a full season of Sentinel-2 data prior to the outbreaks in 2018 to compare with, which made it more difficult to evaluate how an outbreak affects the vegetation season. With a longer time-series it would be possible to compare how the coefficient of variation act during a bark beetle outbreak in relation to how the coefficient of variation behaved during a growing season without bark beetle outbreak. It would then be possible to analyse how the coefficient of variation differed during an outbreak compared to earlier years for the same forest stand. That would enable a comparison between healthy and infested condition for a stand rather than comparing different stands.

Secondly, as images with cloud cover were eliminated with the cloud filter, the different AOIs and areas had unequal number of images from which VI values were extracted. This led to some difficulties when classifying and searching for optimal thresholds of the forest stands since this was done with the cumulative sum i.e., in some cases stands had a higher cumulative sum due to a larger number of images included in the sum. This could potentially have affected the performance of the VI models as some stands could have over-or-underestimated the response signal for one stand compared to the other.

Thirdly, the bark beetle damage data which was used when identifying forest stands with and without recorded berk beetle infestation tells where the has been an outbreak. However, it is possible that there are areas in the data that have been missed or not recorded perhaps due to different forest owners. As the study area is not that big it could be possible that areas with no recorded infestation could have been affected without being detected/recorded.

6. Conclusion

The aim of this thesis was to study if the variability between pixels for a vegetation index on a forest stand scale change during a spruce bark beetle outbreak and to test if the variability can be used as an early indicator for bark beetle infestation.

The results indicate that the variability within a forest stand does change during a bark beetle outbreak with increased variability over time within stands that have been attacked by bark beetles.

It was found that NDWI was the most suitable index during the period May-July as it was able to detect all infested forest stands. However, for the whole season NDVI and RDI also display potential as both were able to detect high rates of infested forest stands while limiting the misclassification of healthy ones.

An infested forest stand could be detected as early as ca. the 29th of May. Meaning that the infestation is still in the green attack stage, in which mitigation is still possible to eliminate the spread of further bark beetle attacks.

The results show that changes in variability have the potential to be used as an early indicator for bark beetle infestation, and the variability can be used to detect and classify individual forest stands that were infested at an early stage i.e., green attack stage.

7. References

Abdullah, H., Darvishzadeh, R., Skidmore, A. K., Groen, T. A., & Heurich, M. (2018). European spruce bark beetle (Ips typographus, L.) green attack affects foliar reflectance and biochemical properties. *International journal of applied earth observation and geoinformation, 64*, 199-209.

Abdullah, H., Darvishzadeh, R., Skidmore, A. K., & Heurich, M. (2019). Sensitivity of Landsat-8 OLI and TIRS data to foliar properties of early stage bark beetle (Ips typographus, L.) infestation. *Remote sensing, 11*(4), 398.

Abdullah, H., Skidmore, A. K., Darvishzadeh, R., & Heurich, M. (2019a). Sentinel-2 accurately maps green-attack stage of European spruce bark beetle (Ips typographus, L.) compared with Landsat-8. *Remote sensing in ecology and conservation, 5*(1), 87-106.

Abdullah, H., Skidmore, A. K., Darvishzadeh, R., & Heurich, M. (2019b). Timing of red-edge and shortwave infrared reflectance critical for early stress detection induced by bark beetle (Ips typographus, L.) attack. *International Journal of Applied Earth Observation and Geoinformation, 82*, 101900.

Bárta, V., Lukeš, P., & Homolová, L. (2021). Early detection of bark beetle infestation in Norway spruce forests of Central Europe using Sentinel-2. *International journal of applied earth observation and geoinformation, 100*, 102335.

Bentz, B. J., & Jönsson, A. M. (2015). Modeling bark beetle responses to climate change. In *Bark beetles* (pp. 533-553). Elsevier.

Blennow, K. (2004). *Osäkerhet och aktiv riskhantering: aspekter på osäkerhet och risk i sydsvenskt skogsbruk*. Sustainable Forestry in Southern Sweden (SUFOR).

Cognato, A. I. (2015). Biology, systematics, and evolution of Ips. In *Bark Beetles* (pp. 351-370). Elsevier.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters, 27*(8), 861-874.

Fettig, C. J., & Hilszczański, J. (2015). Management strategies for bark beetles in conifer forests. In *Bark beetles* (pp. 555-584). Elsevier.

Gao, B.-C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment, 58*(3), 257-266.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202*, 18-27.

Grégoire, J.-C., Raffa, K. F., & Lindgren, B. S. (2015). Economics and politics of bark beetles. In *Bark beetles* (pp. 585-613). Elsevier.

Hanes, J. (2013). *Biophysical applications of satellite remote sensing*. Springer Science & Business Media.

Hlásny, T., König, L., Krokene, P., Lindner, M., Montagné-Huck, C., Müller, J., Qin, H., Raffa, K. F., Schelhaas, M.-J., & Svoboda, M. (2021). Bark beetle outbreaks in Europe: State of knowledge and ways forward for management. *Current Forestry Reports, 7*(3), 138-165.

Huete, A. (2014). Vegetation Indices. In E. G. Njoku (Ed.), *Encyclopedia of Remote Sensing* (pp. 883-886). Springer New York. https://doi.org/10.1007/978-0-387-36699-9_187

Huo, L., Persson, H. J., & Lindberg, E. (2021). Early detection of forest stress from European spruce bark beetle attack, and a new vegetation index: Normalized distance red & SWIR (NDRS). *Remote Sensing of Environment, 255*, 112240.

Jakoby, O., Lischke, H., & Wermelinger, B. (2019). Climate change alters elevational phenology patterns of the European spruce bark beetle (Ips typographus). *Global change biology, 25*(12), 4048-4063.

Jones, H. G., & Vaughan, R. A. (2010). *Remote sensing of vegetation: principles, techniques, and applications*. Oxford university press.

Jönsson, A. M., Harding, S., Bärring, L., & Ravn, H. P. (2007). Impact of climate change on the population dynamics of Ips typographus in southern Sweden. *Agricultural and Forest Meteorology, 146*(1-2), 70-81.

Jönsson, A. M., Harding, S., Krokene, P., Lange, H., Lindelöw, Å., Økland, B., Ravn, H. P., & Schroeder, L. M. (2011). Modelling the potential impact of global warming on Ips typographus voltinism and reproductive diapause. *Climatic Change, 109*(3), 695-718.

Kirkendall, L. R., Biedermann, P. H., & Jordal, B. H. (2015). Evolution and diversity of bark and ambrosia beetles. In *Bark beetles* (pp. 85-156). Elsevier.

Krokene, P. (2015). Conifer defense and resistance to bark beetles. In *Bark beetles* (pp. 177-207). Elsevier.

Liang, S., & Wang, J. (2019). *Advanced remote sensing: terrestrial information extraction and applications*. Academic Press.

Martin Schroeder, D. F. (2020). *Granbarkborrens förökningsframgång i dödade träd under sommaren 2020 i sydöstra Småland, Värmland och Uppland/Västmanland*.

Montero, D. (2021). eemont: A Python package that extends Google Earth Engine. *Journal of Open Source Software, 6*(62), 3168.

Netherer, S., & Hammerbacher, A. (2022). The Eurasian spruce bark beetle in a warming climate: Phenology, behavior, and biotic interactions. In *Bark Beetle Management, Ecology, and Climate Change* (pp. 89-131). Elsevier.

Netherer, S., Kandasamy, D., Jirosová, A., Kalinová, B., Schebeck, M., & Schlyter, F. (2021). Interactions among Norway spruce, the bark beetle Ips typographus and its fungal symbionts in times of drought. *Journal of Pest Science, 94*(3), 591-614.

Olsson, P.-O., Lindström, J., & Eklundh, L. (2016). Near real-time monitoring of insect induced defoliation in subalpine birch forests with MODIS derived NDVI. *Remote sensing of environment, 181*, 42-53.

Ortiz, S. M., Breidenbach, J., & Kändler, G. (2013). Early detection of bark beetle green attack using TerraSAR-X and RapidEye data. *Remote sensing, 5*(4), 1912-1931.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, *12*, 2825-2830.

Raffa, K. F., Aukema, B. H., Bentz, B. J., Carroll, A. L., Hicke, J. A., Turner, M. G., & Romme, W. H. (2008). Cross-scale drivers of natural disturbances prone to anthropogenic amplification: the dynamics of bark beetle eruptions. *Bioscience*, *58*(6), 501-517.

Roberts, D. (2014). Forestry. In E. G. Njoku (Ed.), *Encyclopedia of Remote Sensing* (pp. 210-218). Springer New York. https://doi.org/10.1007/978-0-387-36699-9_48

SKOGSSTYRELSEN. (2019). En branschgemensam satsning. GRANBARKBORREPROJEKTET-stoppa borrarna!

SMHI. (2012, 01/03/2022). *Östergötlands klimat*. https://www.smhi.se/kunskapsbanken/klimat/klimatet-i-sveriges-landskap/ostergotlands-klimat-1.4910

Thomas Heldmark, V. (2020). *Östergötlands gröna hjärta - samtal om skogen och dess värden*.

Vega, F. E., & Hofstetter, R. W. (2014). *Bark beetles: biology and ecology of native and invasive species*. Academic Press.

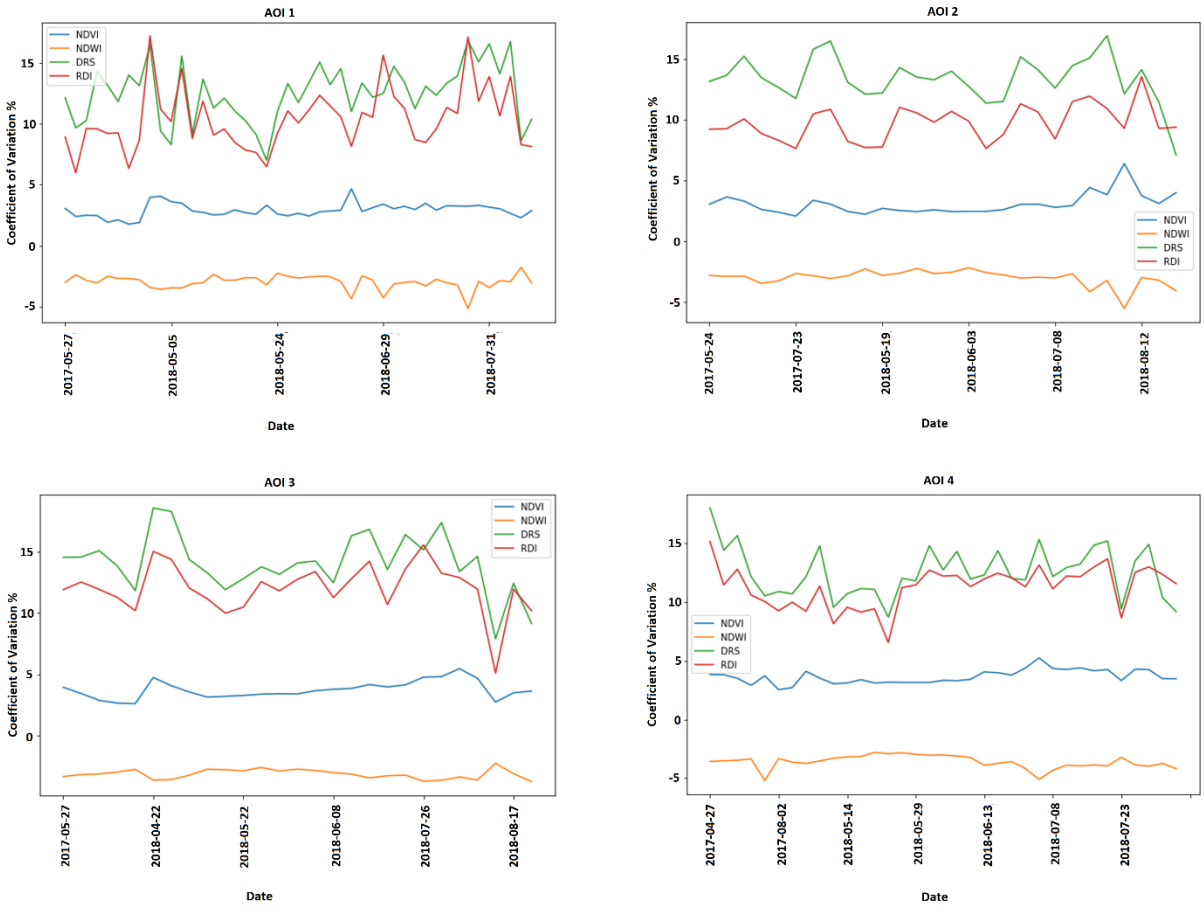
Wermelinger, B. (2004). Ecology and management of the spruce bark beetle Ips typographus—a review of recent research. *Forest Ecology and Management*, *202*(1-3), 67-82.

Wilcke, R. A. I., Kjellström, E., Lin, C., Matei, D., Moberg, A., & Tyrlis, E. (2020). The extremely warm summer of 2018 in Sweden–set in a historical context. *Earth System Dynamics*, *11*(4), 1107-1121.

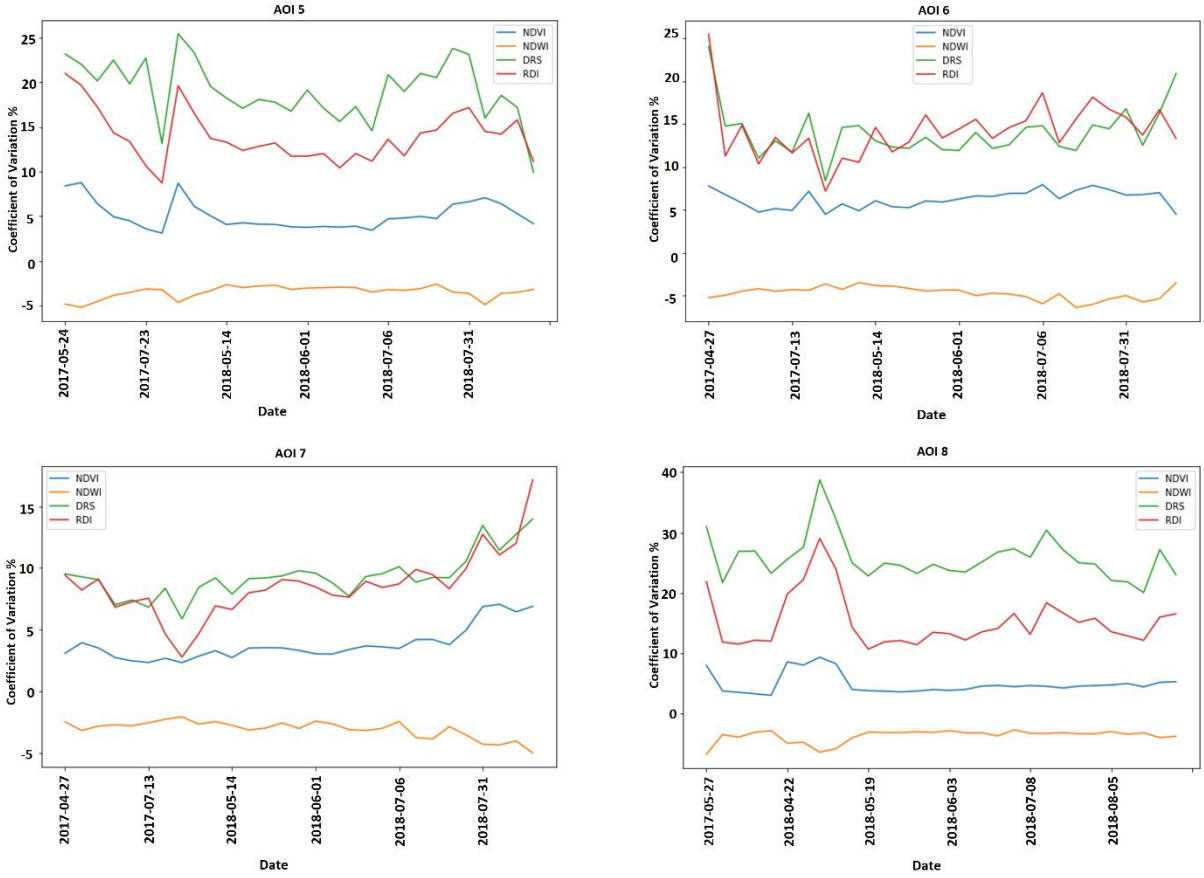
Wulff, S., & Roberge, C. (2021). Nationell Riktad Skogssadeinventering (NRS)—Inventering av Granbarkborreangrepp i Götaland och Svealand 2020. *Institutionen för Skoglig Resurshushållning: Umeå, Sweden*.

Appendix

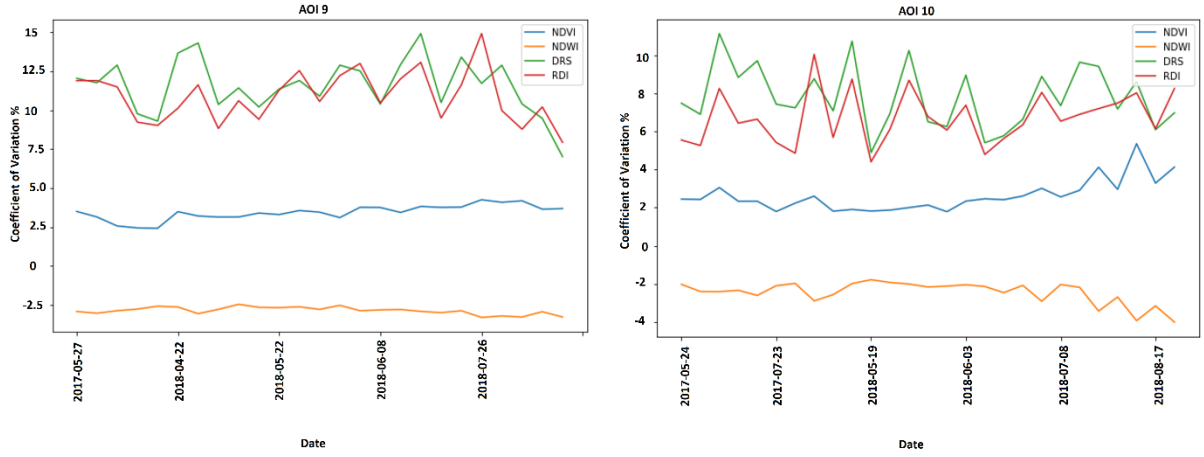
Following appendix section display the coefficient of variation in each AOI where figure A1, A2 and A3 are AOIs with recorded bark beetle infestation. A4 and A5 are AOIs without recorded bark beetle infestation.



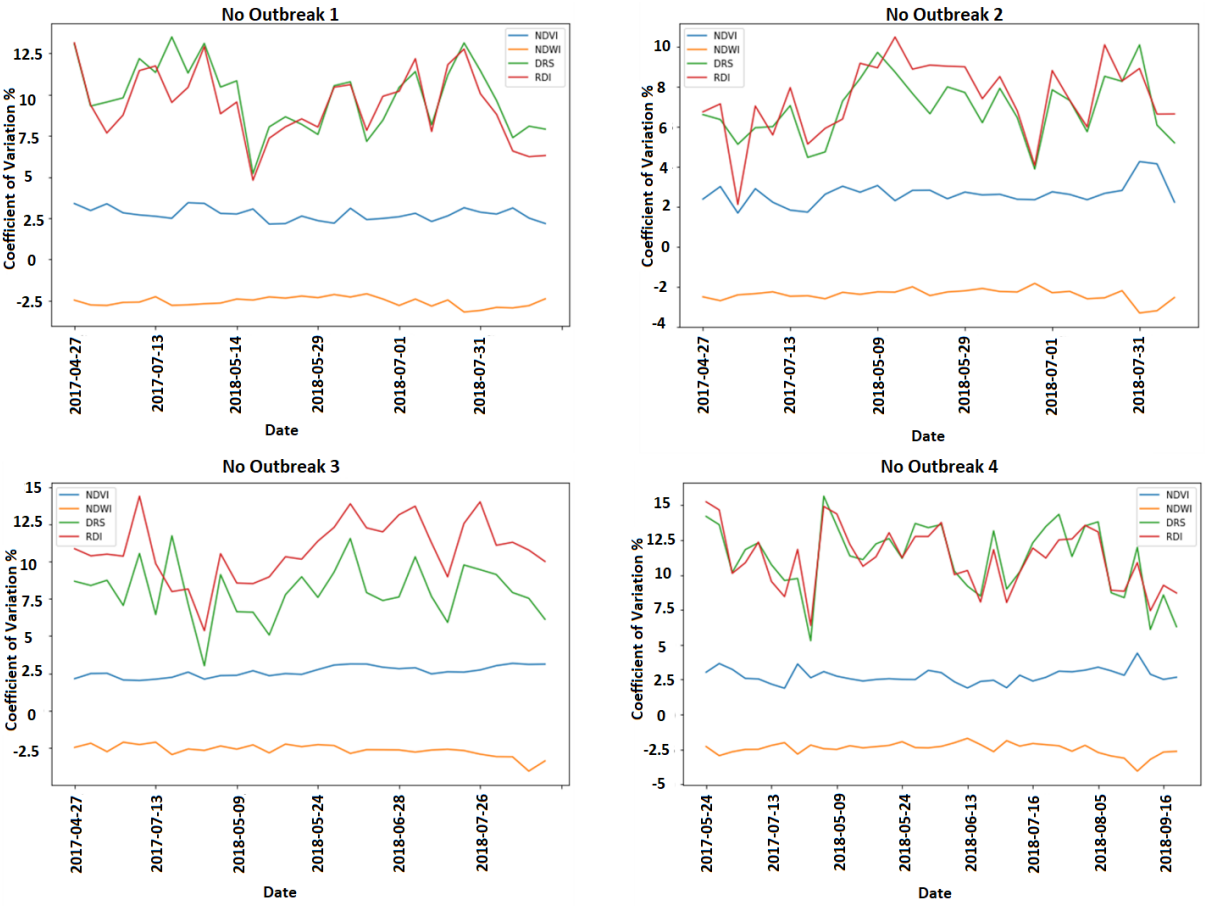
A1: Coefficient of variation (%) for AOI 1-4 between 2017-2018, where green curve is DRS, red is RDI, blue is NDVI and yellow is NDWI.



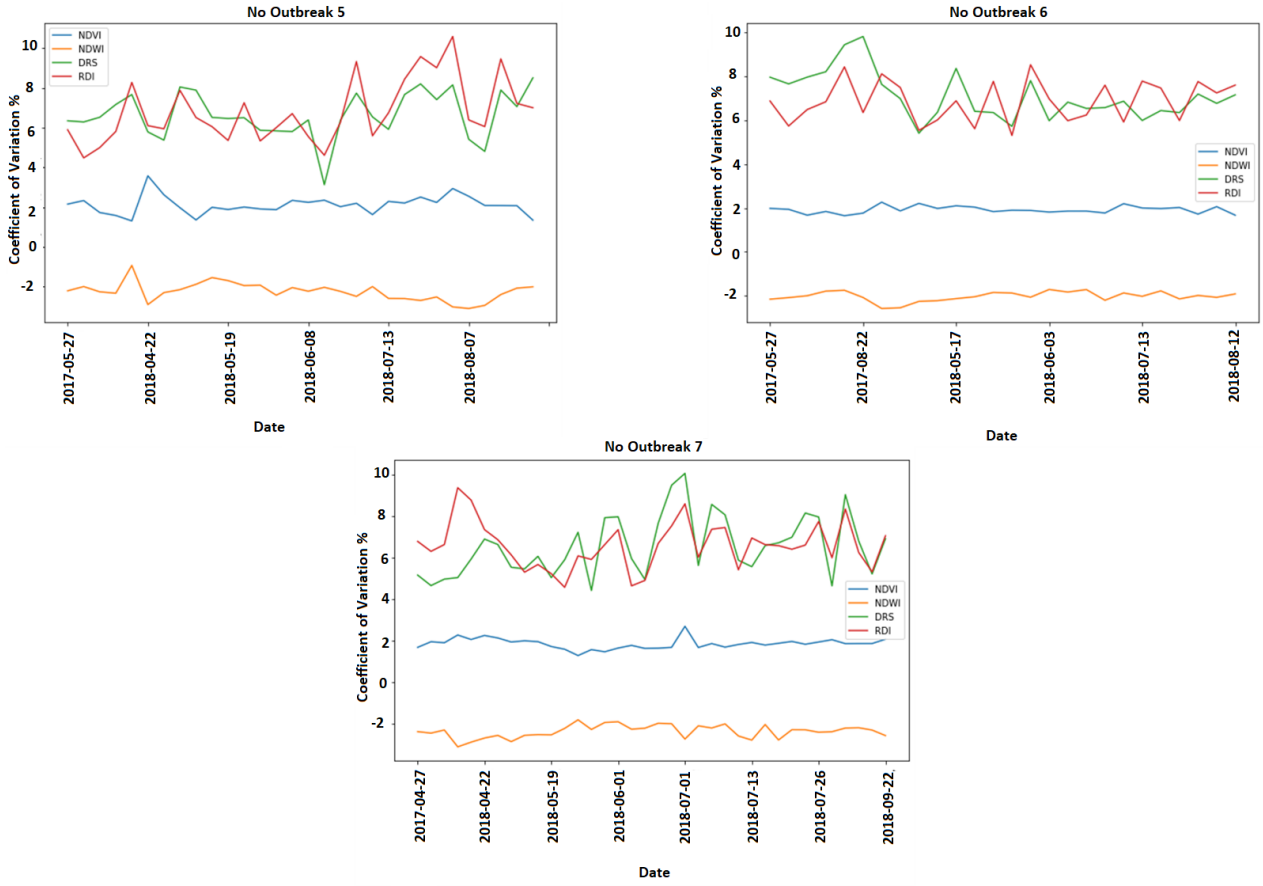
A2: Coefficient of variation (%) for AOI 5-8 between 2017-2018, where green curve is DRS, red is RDI, blue is NDVI and yellow is NDWI.



A3: Coefficient of variation (%) for AOI 9 and AOI 10 between 2017-2018, where green curve is DRS, red is RDI, blue is NDVI and yellow is NDWI.



A4: Coefficient of variation (%) between 2017-2018 for area 1-4 where no bark beetle infestation has been recorded, where green curve is DRS, red is RDI, blue is NDVI and yellow is NDWI.



A5: Coefficient of variation (%) between 2017-2018 for area 5-7 where no bark beetle infestation has been recorded, where green curve is DRS, red is RDI, blue is NDVI and yellow is NDWI.