ASSESSING THE POTENTIAL OF UNMANNED AERIAL VEHICLE DATA IN DETECTING THE BARK BEETLE INFESTATION ON EUROPEAN SPRUCE

EUPHRASIA CHILONGOSHI

July 2021

SUPERVISORS:
Ir. L.M. van Leeuwen-de Leeuw (First Supervisor)
Dr. I. van Duren (Second Supervisor)

ASSESSING THE POTENTIAL OF UNMANNED AERIAL VEHICLE DATA IN DETECTING THE BARK BEETLE INFESTATION ON EUROPEAN SPRUCE

EUPHRASIA CHILONGOSHI
Enschede, The Netherlands, July 2021

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Natural Resource Management

SUPERVISORS:
Ir. L.M. van Leeuwen-de Leeuw
Dr. I. van Duren

ADVISOR:

Dr.Y.A. Hussin

THESIS ASSESSMENT BOARD:
Dr. L.L.J.M. Willemen (Chair)
Dr. T. Kauranne (External Examiner)

ABSTRACT

Insect outbreaks have caused more natural disruptions in coniferous forest ecosystems in recent decades. Early detection of bark beetle green attack, a period when trees start displaying visible signs of infestation stress, is critical to effective and timely forest management to minimize economic loss and prevent a widespread breakout. The bark beetle infestation stages include green, red, and grey attacks. The green attack is the first stage; eggs are laid beneath the tree's bark. The second stage is the red attack; the Europen spruce tree's needles turn from green to yellow or yellow-brown foliage. The last stage is the grey attack, which involves falling off the needles. The infestation stage identification is primarily based on field surveys, which is costly and time-consuming, especially in thick forests, as it is practically impossible to view the tree crown because of overlaps. Hence, the identification of bark beetle infestation, especially the green attack using remote sensing to complement field survey, is critical to mitigating the spread. Most studies have focused on detecting bark beetle infestation stages using remote sensing data obtained from satellites. The use of Unmanned Aerial Vehicle (UAV) offers new tools and methods for better improved early detection of bark beetle infestation by offering datasets with very high spatial resolution. The data collected by the UAV platform can be used in a variety of ways and methods to detect bark beetle infestations. The study's main primary objective was to assess the potential of unmanned aerial vehicle data in detecting the bark beetle infestation on European spruce. To obtain the multispectral UAV images, a parrot Sequoia multispectral camera mounted on a Phantom 4 UAV was flown on the three sampled blocks (one, two, and three). The field-based data collection considered the different bark beetle infestations (green, red, and grey attacks), including healthy European spruce, tree location X, Y coordinates, and block number.

Tree crowns of bark beetle infestations collected in the field were manually delineated to be used as spectral signatures to classify the UAV multispectral images based on the three classifiers Random Forest (R.F), Support Vector Machine (SVM), and Maximum Likelihood (ML). Further, this study also used object-based image analysis (OBIA) to ensure that the European spruce tree crowns were divided into non-overlapping categories for accurate results. In addition, the acquired UAV multispectral images were also used to calculate the three vegetation indices, which include Green Normalized Difference Vegetation Index (GNDVI), Normalized Difference Red Edge (NDRE), and Normalized Difference Vegetation Index (NDVI). The GNDVI, NDRE, and NDVI were further used as input variables in the Species Distribution Model (SDM) to determine the performance of the vegetation indices in predicting the probability of the bark beetle infestation stages, especially the green attack. The classifiers and vegetation indices' performance was assessed to identify which classifier and vegetation index gives the higher accuracy. The findings of this study showed that a UAV multispectral image could be utilized to detect and predict the probability of bark beetle infestations stages on European spruce, especially the green attack, with the possibility of giving higher accuracies. The research found out that among the three classifiers (R.F, SVM, and ML), ML outperformed R.F and SVM. Furthermore, among the three vegetation indices, NDRE outperformed GNDVI and NDVI. Additionally, the results of this study can be extrapolated to other areas because they are promising as they show that the UAV multispectral images have potential in detecting and predicting the probability of the bark beetle infestation stages, especially the green attack.

Keywords: *Unmanned Aerial Vehicle (UAV), Multispectral Images; Bark beetles; Infestation stages; Classifiers; Vegetation indices; Classification; Species Distribution Model (SDM); Probability; Accuracy Assessment.*

ACKNOWLEDGEMENTS

Thank you, Almighty God, for providing me with the knowledge and insight I needed most to pursue my MSc academic achievement.

I would also like to thank Ir. L.M. van Leeuwen-de Leeuw for giving me the opportunity to work on this thesis, for the criticism, remarks, and treasured advice and supervision that helped me write my thesis report. I am also very thankful to Dr. I. van Duren for her remarks, suggestions, and valuable comments on this research.

I am also grateful to Mr. Timothy Roberts, the UAV drone pilot, during the UAV multispectral image capturing.

Special thanks go to Haagse Bos management and community for giving us permission to conduct our UAV flight and carry out a data collection exercise in the forest. Further, many thanks to Takkenkamp company manager Mr. Teun Schoolkate, who took his time to take us around the forest and understand the challenges they face as forest managers in terms of bark beetle outbreak.

Many thanks go to all ITC staff, specifically from the Natural Resources Department (NRM), for providing exceptional facilities and organization of masters in geoinformation science. Likewise, I would like to convey my thanks to the team I worked with during data collection; my work was made easier with their support. Also, many thanks to all NRM students 2019-2021 intake for their support, advice, and encouragement.

I would also like to express my appreciation to the Netherlands fellowship program for giving me an opportunity to participate in the Master of Geo-information Science program at ITC, University of Twente.

Finally, I want to convey my gratitude to my family for their unwavering support and encouragement during my MSc academic journey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

ABBREVIATIONS

AOI: Area of Interest

AUC: Area Under the Curve

CHM: Canopy Height Model

CBD: Central Business District

DGNSS: Differential Global Navigation Satellite System

GCP: Ground Control Point

GIS: Geographic Information Systems

GPS: Geographical Position System

GNDVI: Green Normalized Difference Vegetation Index

GSD: Ground Sampling Distance

MLC: Maximum likelihood classifier

MS: Multispectral

NDRE: Normalized Difference Red Edge

NDVI: Normalized Difference Vegetation Index

NRM: Natural Resouce Management

OBIA: Object-Based Image Analysis

OOB: Out Of Bag

RF: Random Forest

RTK: Real-Time Kinematic

RMSE: Root Mean Square Error

SDM: Species Distribution Model

SVM: Support vector machine

QGIS: Quantum Geographic Information Systems

UAV: Unmanned Aerial Vehicle

1. INTRODUCTION

1.1 Background

Forests play an essential role in social, economic, and environmental functions such as biodiversity conservation by providing habitation to several world's faunae and flora (Mauri, Strona, & San-Miguel-Ayanz, 2017). According to Christopoulou (2011), the global extent of forests and other woodland is approximately 31% of the total land area. Europe's continent has about 182 million hectares of forest, accounting for 5% of the world's total forest. Europe's forest cover takes up approximately 43% of its entire land area (The European Union and Forests, 2019). Furthermore, the Netherlands' forest cover accounts for roughly 10% of the country's total land surface area. Coniferous, broadleaved, and mixed coniferous and broadleaved are the three major types of forest found in the Netherlands. Some of the dominant tree species found in the three forests include spruce, Larch, Pine, Douglas fir, Oak, Beech, Birch, Ash, and other exotic deciduous trees (Fahy, 1999). The Netherlands is a country situated in northwestern Europe (Schuck & van der Maaten, 2013). Further, Chazdon et al. (2016) describe a forest as an ecological unit comprised of trees and numerous biological diversity forms.

According to Barbati et al. (2014) and Portoghesi (2006), the coniferous forest comprises of cone-bearing needle-leaved or evergreen trees found in the temperate western part of Europe. One of the main tree species found in coniferous forests includes spruce. The spruce can grow tall, reaching 50-60 meters in height with a straight and regular trunk, making it vital for many uses, as shown in figure 1 (Caudullo, Commission, & Rigo, 2017).

Figure 1: Straight and regular trunk for European spruce

Source: Field data (2020)

The European spruce is considered one of Europe's essential trees due to its long tradition of cultivation. European spruce plays a significant role in providing environmental and economic values (Caudullo et al., 2017). The European spruce provides many environmental values, such as ecosystem services. The ecosystem services include water and nutrient cycles, soil and wind erosion protection, and a carbon sink. European spruce's economic values include forest resources, such as timber for construction and paper production. Besides, the wood extracted from European spruce is also vital to the local people producing wooden products such as furniture, artwork (European Union, 2019). However, the bark beetle infestation cases on European spruce have been reported in the northern, eastern, and central parts of Europe, Asia, Canada, Romania, and Russia (FAO, 2005; Fettig & Hilszczański, 2015). According to USDA Forest Service (2014), a bark beetle is a tiny insect with a tough, cylindrical body that breeds beneath the tree's bark and eventually kills the tree, as shown in figures 2 and 3.

Figure 2: Bark beetle insect with tough, cylindrical body shape

Source: (Forestry Commission, 2018)

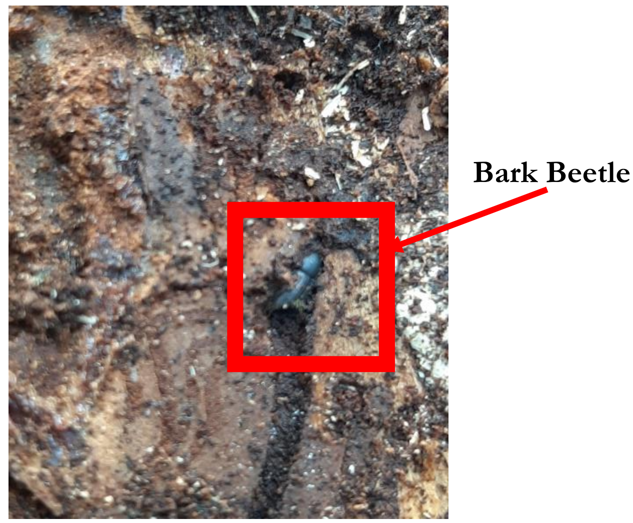


Figure 3: Bark beetle breeding beneath the tree's bark

Source: (Field data, 2020)

The bark beetle infestation has caused very high mortality to European spruce, affecting forest structure and composition causing major economic losses to forest owners (Abdullah et al., 2018; Groen & Heurich, 2018; Jönsson & Bärring, 2011). The bark beetle infestation on European spruce has also been reported in the Netherlands' temperate coniferous forest. According to Wermelinger (2004), the harsh storms that occurred in central Europe in the 1990s have sparked the bark beetle outbreak on the European spruce. In addition, the monoculture way of European spruce planting makes the population vulnerable to pests and diseases (Felton, Lindbladh, Brunet, & Fritz, 2010).

The bark beetle life cycle consists of three main stages: dispersal, attack, and colonization. The clear signs of bark beetle infestation in European spruce can show within four to five weeks once the fully developed bark beetle, carrying fungae, houses the tree. Firstly, during the dispersal period, the adult beetles carrying the fungi mycangia or exoskeleton are dispersed by wind. Secondly, the attack phase involves tree selection by the discoverer beetle. At this stage, both male and female bark beetles enter the tree and release the pheromone. The pheromone attack happens within 2 to 5 days. Lastly, the colonization stage involves laying the eggs and the injection of fungi into the phloem. It is at this stage that the extensive burrowing of the phloem happens. The tree's killing is begun through the pheromone mediate mass attack by the newly developed adults through the spore suckling (Six & Wingfield, 2011). Figure 4 shows the three main stages of the bark beetle life cycle: dispersal, attack, and colonization.

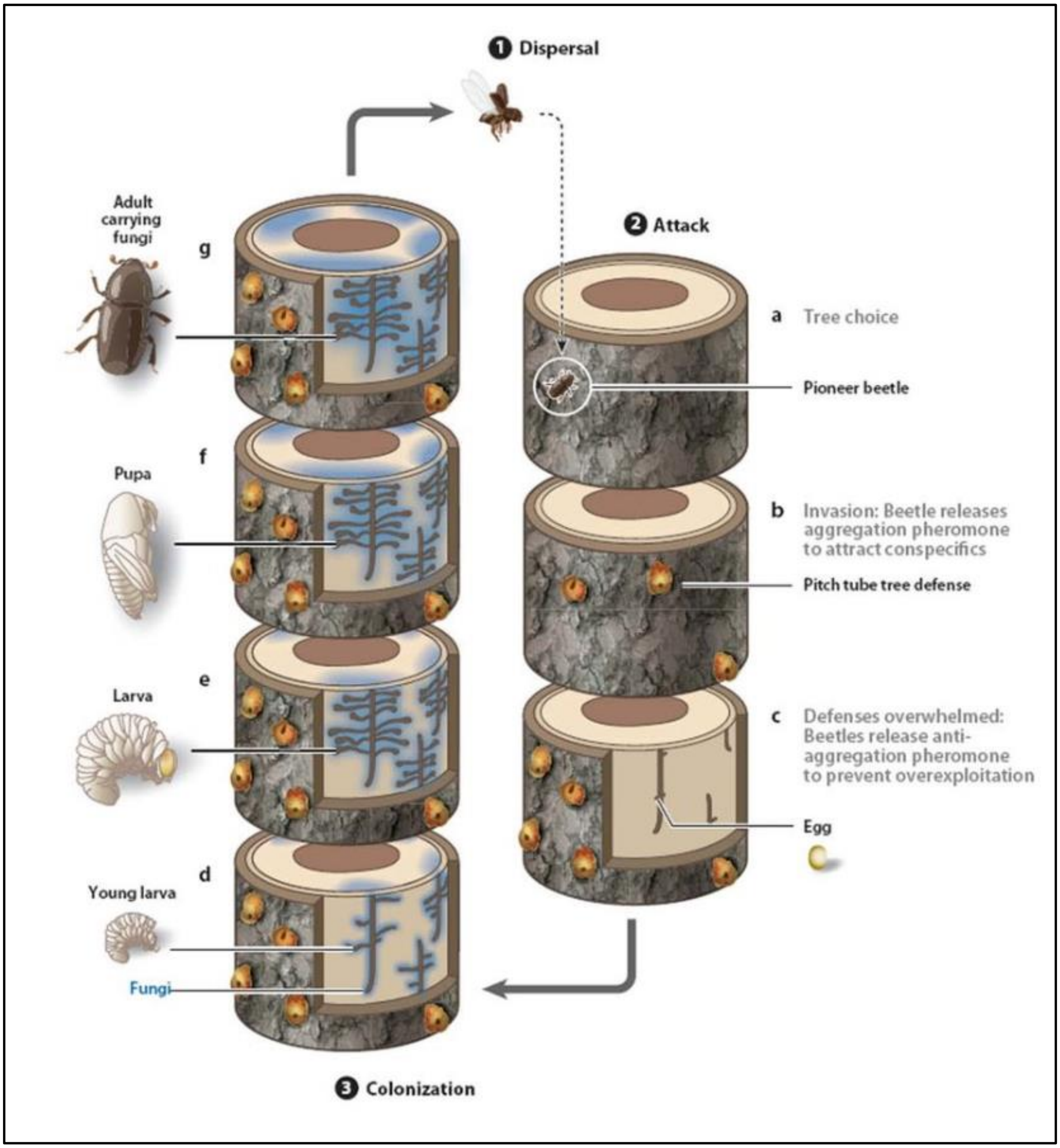


Figure 4: Three main stages of the bark beetle life cycle: dispersal, attack and colonization.

Source: (Six & Wingfield, 2011)

Bark beetles feed and reproduce within the phloem tissue under the infested trees' bark (Food and Agriculture Organization of the United Nations, 2005). The spread of bark beetle infestation is dependant on old trees with a diameter breast height (DBH) greater than 25 cm, weakened by storm or drought (Coggins et al., 2011; Fettig & Hilszczański, 2015). This makes the European spruce stands of seventy years and above more vulnerable to bark beetle infestation. In dry summers, a continuous water deficit decreases water usage and net production, deteriorating old trees in the end. Further, nutrient recycling also slows down, making the European spruce trees more vulnerable to bark beetle infestation (Oliech, 2019). The climatic condition is also a contributing factor to the spread of bark beetles. Dry years enhance the mass reproduction of beetles, such as warm spring. The bark beetles actively swarm when the temperatures are above 16° -20° C (Heurich, 2009).

Consequently, in the early stages of the attack, the European spruce trees produce large amounts of oleoresins following the wounding, which is the primary defense mechanism against the bark beetle attack (Lewinsohn, Gijzen, & Croteau, 1991). Fettig & Hilszczański (2015) state that when the moisture flow is adequate and oleoresin transudation pressures occur, bark beetles that instigate host selection are killed mainly by drowning when the population is low. However, the bark beetle carries fungal pathogen such as satin fungi (Ophiostoma and Cerato-cystis species), affecting the European spruce. The introduction of the diseases by the bark beetles may stress and overwhelm host trees' defense mechanisms and eventually kill the tree (Food and Agriculture Organization of the United Nations, 2005). Further, the beetle larva and the fungi cut the transportation of water, sugar, and other nutrients within the tree's trunk when they penetrate through the living phloem and xylem cells (Abdullah et al., 2018).

There are three different bark beetle infestation stages: green, red, and grey attacks (Wermelinger, 2004; Fettig & Hilszczański, 2015; Six & Wingfield, 2011). The green attack is the first stage; eggs are laid beneath the tree's bark; according to figure 4, this happens under the attack (2) stage (Fettig & Hilszczański, 2015). Coggins, Coops, & Wulder (2011) indicates that, during the first few weeks of the attack, the foliage can stay green and look healthy. However, the beetles' action beneath the bark limits water and nutrients' translocation from roots to the foliage; the tree will show withering signs, a period when the tree show visual signs of infestation stress. The second stage is the red attack; the Europen spruce tree's needles turn from green to yellow or yellow-brown foliage. These signs will show within five and six weeks from the infestation date. The last stage is the grey attack, which involves falling off the needles, limited or no foliage remains; only the bark remains till death (Abdullah et al., 2019; Fettig & Hilszczański, 2015). Both the red and the grey attack occurs at the colonization stage, refer to figure 4.

Some of the management strategies to mitigate the spread of the bark beetle include attractants and clear cuts. The attractant is the process where bark beetles are trapped and collected in masses, as shown in figure 5. This method has proved ineffective because some bark beetles may spill and infest nearby European spruce leading to more infection (Fettig & Hilszczański, 2015). Clear-cutting of infected spruce trees has been another method to mitigate the spread (Heurich, 2009). Figure 6 shows the clear cut in the study area.

Figure 5: Attractant mitigation measure

Adapted from: (Forest Protection, n.d.)

Figure 6: Clear cut mitigation measure

Source: (Field data, 2020)

Nevertheless, clear-cutting contributes to habitat loss for animal and plant species leading to species extinction. Clear cuts also lead to an immense amount of greenhouse gases being released into the atmosphere leading to global warming (Tonks, 2020).

However, the mitigation measures used have not been beneficial to forest managers as the mortality rate keeps increasing. There is a need to detect the bark beetle infestation in its early stages, called green attack. The detection of the green attack at an early stage will enable forest managers to take quick action to prevent the spread and enhance effective forest management to lessen European spruce's economic loss (Abdullah et al., 2019a).

1.2 Detection of bark beetle using remote sensing

Studies have shown that identifying different bark beetle infestation stages using remote sensing data potentially gives useful information. Different sensors have their advantages and disadvantages. New sensors and methods are also continuously developed. However, we are yet to evaluate which sensor and method perform best.

Active and passive sensors have been used to determine different bark beetle infestation stages on European spruce. By definition, active sensors have their source of light on board, emitting energy in specific wavebands. In contrast, passive sensors depend on the sun's energy (Erdle, Mistele, & Schmidhalter, 2011). Passive sensors include moderate-spatial resolution imaging spectroradiometer (MODIS), Landsat, SPOT, and multispectral sensors. Further, (Light Detection and Ranging) LiDAR and RADAR have also been used as active sensors (Chen & Meentemeyer, 2016; Bright, Hudak, McGaughey, Andersen, & Negrón, 2013; Puntsho, 2020; Çöltekin et al., 2020).

1.2.1 Active sensors

The Radar sensor is not sensitive to cloud covers; it penetrates through the canopy and has broad spatial coverage. Therefore, Radar sensors might overcome the cloud cover problem in passive sensors. The only challenge is that it has a low spatial resolution of 30 meters, making it difficult in detecting infestations Stages in individual trees, and it is also commercial (Ortiz, Breidenbach, & Kändler, 2013; Marshall, Rees, & Dowdeswell, 1993; Puntsho, 2020; Çöltekin et al., 2020).

In addition, the Lidar sensor has also been used to detect bark beetle infestation in coniferous forests using the irradiance information that provides some signs of the state of photosynthetic level in trees. The detection of bark beetle infestation stages is done by measuring the intensity of light reflected to the sensor. The Lidar intensity values tend to decrease in infected trees (Bright et al., 2013). The spatial coverage depends on the Field of View (FOV). Further, Lidar's spatial resolution ranging from 1 meter to 150 meters, though the low spatial resolution limits the detection of different bark beetle infestation stages. Lidar sensors are not affected by cloud covers because they do not depend on the sun as an energy source (Blanchard et al., 2014). Most Lidar cameras are commercial, making the exercise costly (Royo & Ballesta-Garcia, 2019).

1.2.2 Passive sensors

Chen & Meentemeyer (2016) indicate that sensors like (Landsat, MODIS, and SPOT) have been used for identifying different infestation stages. However, the mentioned sensors have limited capacity due to low spatial resolution (Chen & Meentemeyer, 2016). Low-spatial resolution images do not give accurate results due to mixed pixels; this means many features per pixel (Wulder, Dymond, White, Leckie, & Carroll, 2006). In addition, the lack of clear visual indicators in the needles in the green attack stage makes it difficult to distinguish between health and the early stage (Abdullah et al., 2019). Many studies have been conducted using spectral vegetation indices from low to high-spatial-resolution satellite data to determine different bark beetle infestation stages, especially the early stage. However, the results have not been accurate enough (Abdullah et al., 2019).

On the one hand, high-spatial-resolution images have new possibilities for detecting and mapping bark beetle infestations. The spectral signatures from different electromagnetic spectrum bands have been linked to various functional and structural plant traits such as 1100 – 2400nm is used to detect spectral reflectance properties, mainly using the red edge band (Abdullah et al., 2019a). According to Chen & Meentemeyer (2016) findings, high-spatial-resolution images are appropriate for fine-scale detection. Because most high-spatial-resolution images cover a vast region, with spatial resolutions ranging from 0.5-2m, they may be stable when most infected trees are in small, discrete patches. RapidEye, Worldview, and HyMap are examples of such images (Abdullah et al., 2019a). High-spatial resolution images, on the other hand, are costly. They're usually commercial, and image collection requires placing an order ahead of time. Furthermore, cloud cover often obscures some information in high-spatial-resolution images (Wulder, White, Coops, & Butson, 2008; Gartner, Veblen, Leyk, & Wessman, 2015).

Unmanned Aerial vehicles (UAVs) have become popular in the recent past due to the low prices and flexibility, limited expertise required to operate, and no cloud cover problem. Therefore, UAVs are economical compared to other satellite platforms in data acquisition and downloading costs (Nowak, Dzińb, & Bogawski, 2019). UAV's are capable of identifying individual trees and even leaves due to their high spatial resolution at a local scale ranging from 1 to 5cm (Klouček et al., 2019). Furthermore, UAV's have a small spatial coverage and must be operated within a visual line of sight from the ground station, making it possible for a collision.UAVs are very sensitive to atmospheric conditions such as strong winds and rain, affecting images' quality. UAV's have complicated flying regulations such as safety and security issues (Paneque-Gálvez, McCall, Napoletano, Wich, & Koh, 2014; Kanellakis & Nikolakopoulos, 2017).

Until now, UAVs' application in detecting the different infestation stages of bark beetle is still limited. It is challenging to detect different stages of bark beetle infestation on European spruce, especially the early stage (green attack) when using a low-cost camera with low spectral, spatial resolution (RGB) because the distinction between the healthy and green attack stage is minimal. (Klouček et al., 2019). However, other UAV cameras such as multispectral, hyperspectral, and thermal have been used to provide information concerning bark beetle infestation stages.

Figure 7 shows the conceptual diagram for assessing the potential of unmanned aerial vehicle data in detecting the bark beetle infestation on European spruce. The conceptual diagram depicts the arrangement and relationships of key attributes within the system in the study area. The key attributes in the conceptual diagram include Haagse Bos, coniferous forest comprised of healthy and unhealthy European spruce, the causes of bark beetle outbreak, the size of European spruce trees prone to bark beetle infestation, remote sensing platform used to capture the study area, and finally how useful the information captured to the forest managers.

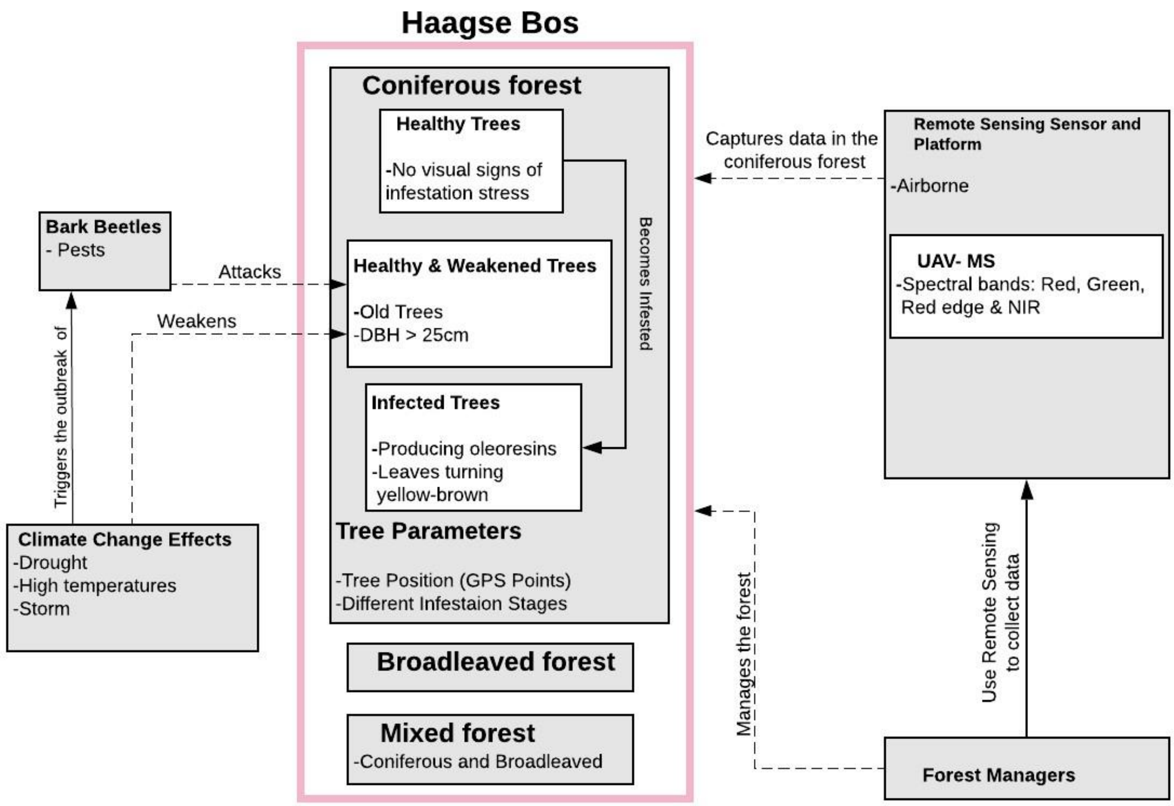


Figure 7: Conceptual diagram explaining the interacting systems within the study area

1.3 Image classification

Furthermore, (Lin, Chin TengLi, Kuo, & Huang, 2012) believe that the UAV multispectral images' spectral and spatial resolution information makes it suitable, efficient, robust, and accurate in determining the different bark beetle infestation stages on European spruce. Therefore, several remote sensing image classification and prediction model techniques are available that have used multispectral images. Maximum likelihood classifier (MLC), machine learning classifiers such as Random Forest (RF), and Support vector machine (SVM) methods have been used before because of the ability and potential to the classification of remotely sensed imagery (Li et al., 2012). In addition, the species distribution model (SDM) has also been used to predict different bark beetle infestation stages (Elith & Leathwick, 2009). However, the best method to determine the bark beetle infestation stages on European spruce, especially the green attack, is yet to be determined. Therefore, the two different approaches (classifier and SDM) were used in this study to determine which approach gives higher accuracy in detecting and predicting the infestation stages based on the UAV multispectral images.

Image classification is the grouping of pixels that are similar in spectral characteristics. There are two types of image classification methods: which include parametric and non-parametric. MLC is a good example of a parametric technique that verifies the distribution of a known class with signatures based on statistics for the class (Al-doski, Mansor, Zulhaidi, & Shafri, 2013). RF and SVM are good examples of non-parametric technique image classification. The signature is not based on statistics but discrete objects such as polygons (Breiman, 2001). RF and SVM are machine learning techniques that use an ensemble learning method for classification. With RF, multiple decision trees are trained on the bootstrapped sample of the original training data. When training samples bootstrapping occurs at each tree in RF, it learns from a random sample of the data set; the samples are drawn with replacement. Since some samples are drawn with replacement, some samples will be used more than others in a single tree (Gislason, Benediktsson, & Sveinsson, 2006). One parameter amongst a randomly chosen subset of input parameters at every decision tree node is selected as the best split and used for node splitting. Every single tree uses only a portion of the input samples for training. Simultaneously, the remaining referred to as Out of Bag (OOB) of the samples are used to validate the output's accuracy. RF has been primarily used to determine forest healthy (Breiman, 2001; Yuan & Hu, 2016).

In addition, SVM has been widely used in image classification because of its accuracy techniques in pattern and image classification. SVM is a supervised learning technique used in image classification because of its computational model (Gove & Faytong, 2012). SVM operates its classification by constructing a hyperplane in higher dimensions. The decision planes are referred to as hyperplanes. The hyperplane differentiates a group of data of one type from the other (Gove & Faytong, 2012). A dataset is classified by creating a hyperplane, which is non-linear mapping for higher-dimensional space. SVM uses vector points which are referred to as support vectors. Support vectors are distinguished by the outcome margin, which can either be a small or large marginal separation between classes (Robert Gove, 2012). Tzotsos & Argialas (2008) indicate that SVM separates the classes with maximum marginal distance in the decision plane. SVM classifier can handle high dimensional multispectral data and have low computational time, and capable of handling multi-class challenges. Furthermore, because of its success in differentiating classes when the number of dimensions is more than the number of samples, an SVM classifier has been employed in remote sensing to detect insect defoliation and map bark beetle infestation (Lin, Chin TengLi, et al., 2012). Figure 8 shows the small and large margins used to separate classes in SVM.

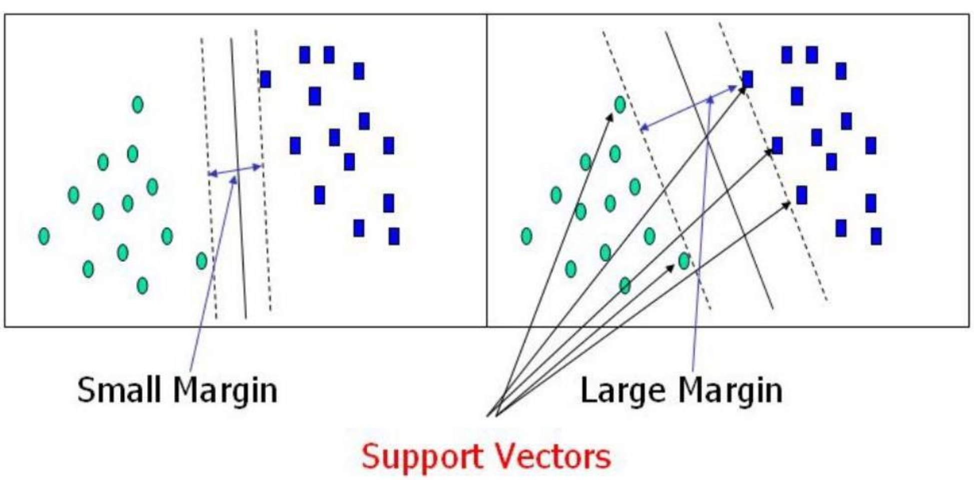


Figure 8: Support vectors with small and large margins for separating classes in SVM

Source: (Rohith Gandhi, 2018).

MCL undertakes the statistics for each category in every participating band and determines if they are normally distributed. It computes a class's probability where a given pixel belongs. The normality theory is dependent on the training samples (Sun, Yang, Zhang, Yun, & Qu, 2013). In addition, the MCL is believed to be one of the most accurate classifications with high accuracy and precision. The method has been used before determining forest healthy and spruce tree infestation (Kranjčić & Rezo, 2018).

Further, the species distribution model (SDM) is a statistical technique that predicts the distribution of species within geographical areas and temporal resolution. The SDM can discriminate the presence and absence of occurrence based on the binary predictor (0, 1). The model's input data are dependant variables that link to the presence and absence of a predicted variable. In addition, the independent variables are used as predictors. The predictor variables are either numeric, binary, or categorical (Elith & Leathwick, 2009). Environmental variables that show the actual locations of each species are also required by the model. Georeferenced raster maps and a list of coordinates are among the environmental data. For example, species data can be represented as geographical points. In contrast, environmental predictors might be represented as a collection of raster maps with the same projection. Thus, according to the concept behind SDM, it might predict distinct stages of bark beetle infestation since it can classify between the presence and absence of different infestations (healthy, green, and red attacks) using binary predictions of 0 or 1 (Naimi & Araújo, 2016). The task can be done by assigning one (1) to the infestation stage to be predicted and zero (0) to the infestation stage that is not to be predicted. Tree position X, Y coordinates, extracted standard deviation values of vegetation indices from digitized polygons, and three vegetation indices maps. The three vegetation indices include Green Normalized Difference Vegetation Index (GNDVI), Normalized Difference Red Edge (NDRE), and Normalized Difference Vegetation Index (NDVI).

The SDM is an example of the logistic regression known as a generalized linear model (GLM) because the outcome is constantly dependent on the sum of the inputs and parameters. The GLM is an extension of the traditional linear regression techniques. The difference between traditional linear regression and the GLM is that the dependent variable in a GLM is changed automatically using a link-function, which defines mathematically how the dependent variable is converted relative to its mean value and linked to the linear predictor. On the contrary, this is impossible with linear regression techniques as it usually gives skewed distribution (Kienast, Bolliger, & Zimmermann, 2012). The advantage of the SDM technique is that the code used in SDM has the benefit of inheritance as it can be shared for related studies (Alfons, Templ, & Filzmoser, 2010).

1.4 Vegetation indices

Vegetation indices have been used before in remote sensing to determine vegetation healthy based on the correlation between forest stress and vegetation indices (Lottering, Mutanga, & Peerbhay, 2018). By definition, vegetation indices are a mathematical combination of two or more spectral bands, highlighting the distinction between vegetation reflectance from two or more wavelengths to show specific vegetation characteristics (Z. Wang et al., 2016).

According to Abdullah et al. (2018), vegetation indices such as Normalized Difference Red Edge (NDRE), Normalized Difference Water Index (NDWI), Decrease Water Stress Index (DSWI), Normalized Difference Vegetation Index (NDVI), Green Normalized Difference Vegetation Index (GNDVI) and Leaf Water Content

Index (LWCI) have been used before in the detection and mapping of the bark beetle green attack infestation stage. Besides, Abdullah et al. (2019a) believe that the outlined vegetation indices highlight any slight changes within the European spruce foliage due to bark beetle infestation. The highlight in any small change is because sensors with particular spectral bands in red (R), green (G), near-infrared (NIR), and red edge (RE) are mostly used to establish the amount of photosynthetically active green foliage in plants such as a UAV multispectral. The detection in any slight change is because the red band absorbs red light, the green band reflects green light, and NIR reflects the entire spectrum making it possible to establish chlorophyll content in the plant (Teillet, Staenz, & Williams, 1997). In addition, since the chlorophyll is green, the UAV multispectral green band is a good absorber of electromagnetic energy when the vegetation is healthy. At the same time, the red and near-infrared bands are good reflectors of electromagnetic energy (Kumar, 2017).

1.5. Problem statement

Insect outbreaks have caused more natural disruptions in coniferous forest ecosystems in recent decades. Early detection of bark beetle green attack, a period when trees start displaying visible signs of infestation stress, is critical to effective and timely forest management to minimize economic loss and prevent a widespread breakout. The bark beetle infestation has caused very high mortality to European spruce in coniferous forests, affecting forest structure and composition. Thus, there is a need for early detection of bark beetle infestation to mitigate the challenges faced by forest managers. The identification of different infestation stages is mostly field-based: which is expensive, time-consuming, especially in large or inaccessible terrain due to trees' height. In this view, there is a need to identify cost-effective remote sensing methods for determining the different bark beetle infestation stages, especially the early stage (green attack): to mitigate the bark beetle's spread through early removal of infested trees. Remote sensed-based methods are more transparent and accurate when high-spatial-resolution images are used as opposed to medium spatial resolution images such as (Landsat, MODIS, and SPOT) which do not clearly show infestation stages due to coarse spatial resolution (Eitel et al., 2011). UAV images provide a very high spatial resolution; this relatively new technology has not been broadly used.

Therefore, this research seeks to explore the potential of UAV multispectral images to detect different bark beetle infestation stages on European spruce, especially the green attack, as little information has been documented on the methods used.

1.6. Main objective

To assess the accuracy of the existing methods in detecting the different bark beetle infestation stages on European spruce using UAV multispectral images, especially the green attack.

1.6.1. Specific objectives and research questions

No.	Specific Objectives	Research Questions
1.	To assess which classifier among RF, SVM, and ML gives the highest accuracy in detecting the green attack of bark beetle infestation stages on European spruce from UAV multispectral images.	What classifier among RF, SVM, and ML gives the highest accuracy in detecting the green attack of the bark beetle infestation stage on European spruce from UAV multispectral images?
2.	To determine which vegetation indices (GNDVI, NDRE, NDVI) from UAV multispectral images give the highest probability in predicting the different bark beetle infestations stages, especially the green attack using UAV multispectral images?	Which vegetation indices (GNDVI, NDRE, NDVI) from UAV multispectral images have the highest probability of detecting and predicting the green attack of bark beetle infestation stage on European spruce using UAV multispectral images?
3.	To compare the accuracies from the classifiers (RF, SVM, ML) and vegetation indices (GNDVI, NDRE, NDVI) in detecting the different bark beetle infestation stages, especially the green attack using UAV multispectral images.	Which of the classifier (RF, SVM, ML) and vegetation indices (GNDVI, NDRE, and NDVI) has the highest accuracy in detecting the different bark beetle infestations stages using, especially the green attack using UAV multispectral images?

2 MATERIAL AND METHODS

2.1 Description of the study area

Since the research aimed to determine how useful data from unmanned aerial vehicles could be in detecting bark beetle infestations on European spruce. As a result, I sought a location with European spruce affected by a bark beetle infestation. Having the bark beetle outbreak in Haagse Bos coniferous forest, particularly among European spruce tree species, the Haagse Bos met the study's objective. The Haagse Bos is northeast of Enschede central business district (CBD) in the Netherlands, Overijssel province. The Haagse Bos lies between 261000 meters and 264000 meters East and 474500 meters to 478200 meters North, as shown in figure 9 (Primasatya, Hussin, & Van Leeuwen, 2016). The area is relatively flat, with elevation varying from 49 to 65 meters above sea level (Wikiloc Haagse Bos Enschede Losser Trail, 2020). The total forest area is 43 hectares and is privately owned. It consists of broadleaf, coniferous, mixed forest of native and exotic tree types. Previously, the forest was managed for timber production, particularly coniferous tree wood (Primasatya et al., 2016). The total sampled area for bark beetle infestation covers 3.7 hectares, with European spruce (Picea abies) being the major coniferous tree species. The other two major coniferous tree species include Douglas fir and Scots pine (Pinus syvestris). In winter, the weather condition is cold and windy, with temperatures dropping below zero on occasion. Summer is hot, with typical monthly temperatures ranging from 12 to 25 degrees celsius and an annual total rainfall of 841 millimeters (Gaden, 2020). The three sampled blocks within the Haagse Bos include one, two, and three, as shown in figure 9.

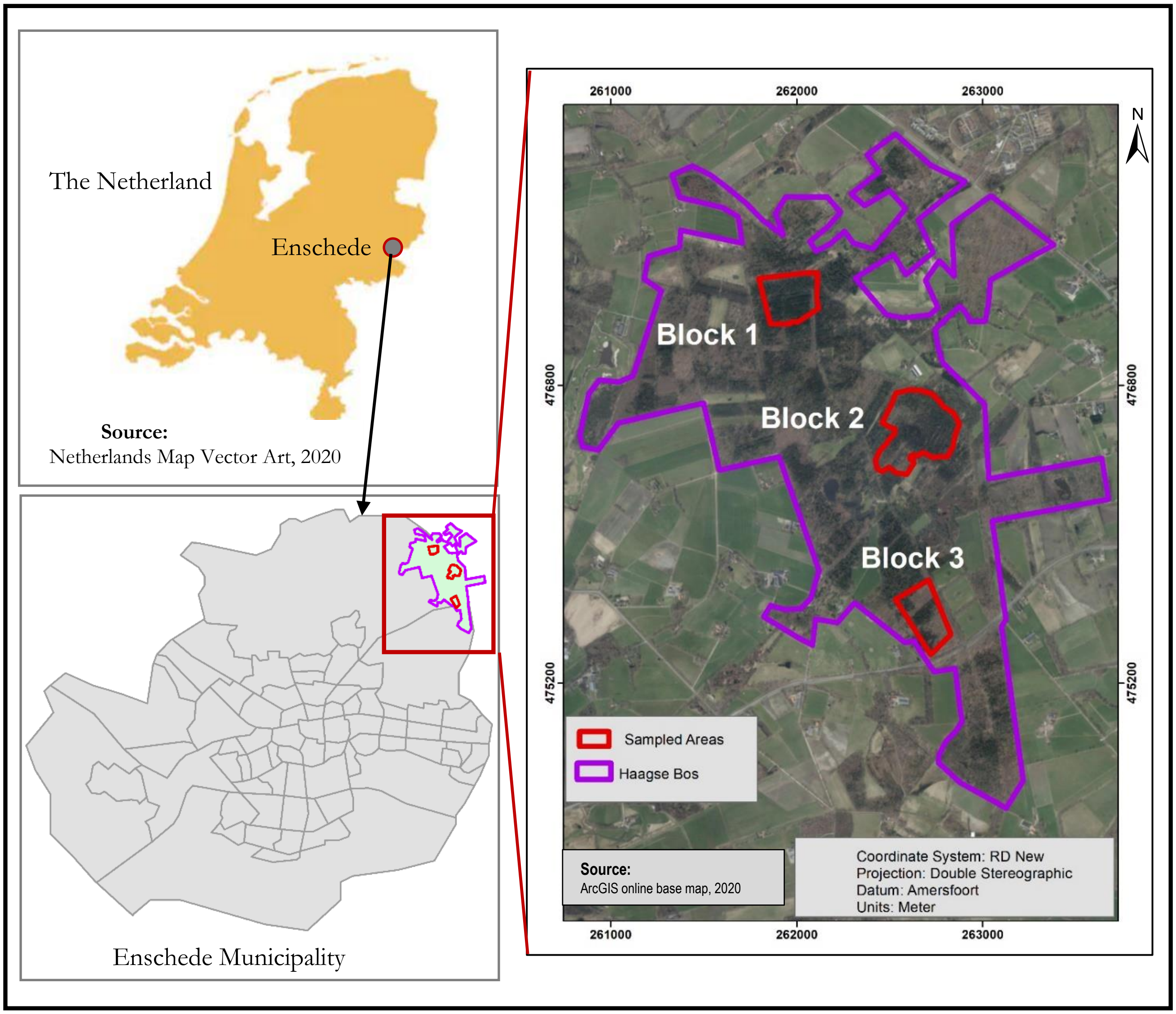


Figure 9: Three sampled blocks (one, two and three) within Haagse Bos infested with bark beetle outbreak

2.2 Research methods

There were nine parts to the approach used in this study: (1) fieldwork planning, (2) UAV multispectral image acquisition and field data collection, (3) image processing, (4) image segmentation; feature extraction of mean and standard deviation brightness values for band red, green, red edge and near-infrared, (5) digitizing of different bark beetle-infested and healthy tree crowns, (6) extraction of individual tree crowns, and image segmentation accuracy evaluation, (7) calculation of vegetation indices and extraction of vegetation indices standard deviation values for infested trees and healthy, (8) image classification and SDM model development; and (9) classification accuracy assessment and model validation. The nine mentioned parts in this section are linked to figure 10 for a detailed explanation of each step.

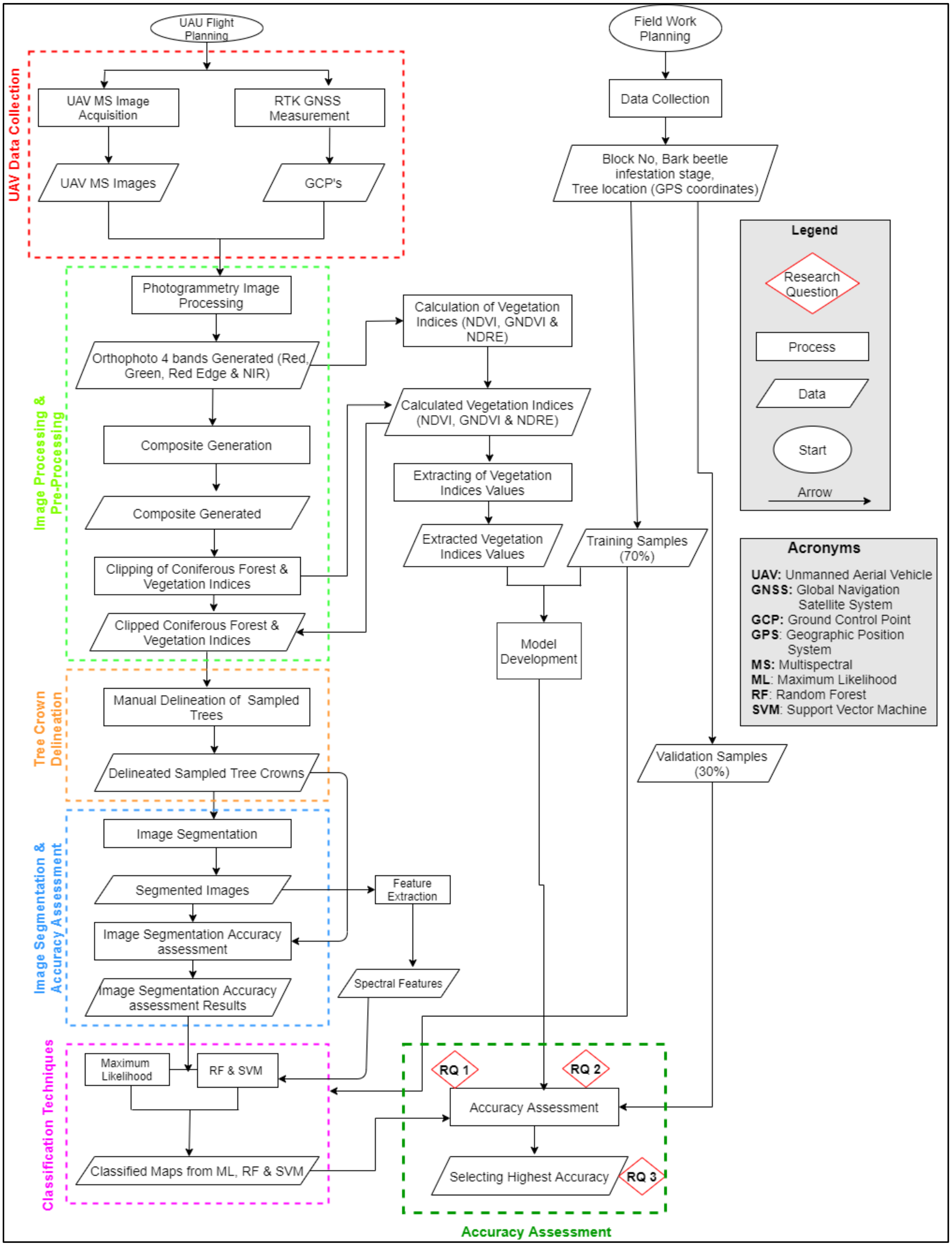


Figure 10: Workflow describing the methods applied in this study

2.3 Sampling design

In remote sensing, samples are collected for two reasons: training data for classification, model development, image accuracy assessment, and model validation (Ramezan, Warner, & Maxwell, 2019). Purposive sampling approaches was used to collect data for this study. Purposive sampling is a method that relies on the researchers' judgment in selecting representative sample sites and samples. On the other hand, purposive sampling uses expert knowledge to choose samples, making the process more efficient and accurate (Hewson, 2007; Ramezan, Warner, & Maxwell, 2019). Consequently, to fulfill the research goal, a purposive sample technique was used in this study. The selection of the three sampled blocks and the bark beetle infestations in Haagse Bos were intentionally dependent on the European spruce presence and bark beetle outbreak. Furthermore, specialists' knowledge of distinguishing one infestation stage from the other was employed to ensure that the sample collected was correct and effective.

2.3.1 Pre-field work

After an initial visit to the study area on October 8, 2020, sample blocks were identified using Google Earth Pro prior to fieldwork. The blocks in this study refer to specific locations where samples of different stages of bark beetle infestation on European spruce were collected from.

2.3.2 Datasets collected

Two datasets were collected for this study: UAV multispectral images and different bark beetle-infested and healthy European spruce. The equipment used for primary data collection is shown in table 1.

Table 1: List of field equipment used for primary data collection

No	Equipment	Purpose
1	Unmanned Aerial Vehicle (UAV) Phantom 4	Mounting on of the sequoia camera
2	Parrot Sequoia Camera	Capturing of Multispectral images
3	differential global navigation satellite system (DGNSS)	To measure Ground Control Points (GCPs)
4	Ground Control Points (GCP) Markers	To mark the GCP's
5	Avenza Maps Software	Data collection of individual trees and navigation

2.3.3 Software

The facilitation of the study was made possible using the different software packages. Table 2 shows the software that enabled the processing and analysis of the datasets in this study.

Table 2: Software packages used for processing and analyzing the collected data

No	Software	Purpose
1	PIX4DMapper 4.6.1	Photogrammetry Processing of UAV images
2	ArcGIS 10.7	Delineation of the tree canopy, tree masking, symbology, map layout, vegetation indices calculation, and standard deviation value extraction
3	Ecognition Developer	Image segmentation and feature extraction
4	QGIS 3.14	RF classification,
5	ENVI 5.6	SVM classification
6	ERDAS Imagine	ML classification
7	RStudio	Model development for the prediction of different bark beetle infestation stages

In reference to table 1; During data collection, the sequoia camera was mounted on a UAV platform. The ground control station, the aerial platform, and the communication link make up the UAV platform. During manual operation mode, the UAV is operated from the ground control station. The flying drone serves as the aerial platform. The communication connection is a wireless data link between the transmitter and receiver for remote communication with a UAV (Rath & Shekarappa, 2007).

The multispectral images from the study area were captured using a sequoia camera. A Sequoia camera is a multispectral camera that captures high spatial resolution data from a target area. It is designed to be compatible with and mountable on nearly all UAV systems. The sequoia camera's four multispectral sensors capture green, red, red-edge, and near-infrared spectral bands. The initial measurement unit (IMU) and reflection data and the GPS location utilized by the processing software to reconstruct the data are recorded by the sequoia camera (Deng et al., 2018). In addition, the spatial resolution of the multispectral images acquired for this study was 10.49 cm x 10.49 cm. The selected spatial resolution was arrived at because high spatial resolution improves branch, stem, and leaf structure (Kellner et al., 2019).

To offer positional adjustments to GPS signals, eliminate pseudorange errors, and improve the quality of position data, the differential global navigation satellite system (DGNSS) was used to capture the ground control points (GCPs). A DGNSS is a satellite-based system used for navigation and monitoring ground control points (GCPs). The user, control, and space segment are the three essential components of a DGNSS. The user segment consists of the people who utilize the system. The control section guides the spacecraft through its orbits and reports any problems detected during navigation. In addition, constellation satellites make up the space portion. The binary codes are sent through high electromagnetic fields to communicate between the segments through high electromagnetic radiation known as carrier waves (wave modulation) (Hofmann-Wellenhof et al., 2008).

The GCP makers were used to register the points during UAV multispectral image acquisition. The GCP markers were essential for orthorectification during photogrammetry pre-processing. At least the minimum number of three GCPs is required, but the higher the number, the higher accuracy of the final results during orthorectification (Oniga, Breaban, & Statescu, 2018). The three blocks had a different number of GCPs. Block one had six GCPs, block two had five, and block three had three, respectively. Figure 11 shows the GCP.

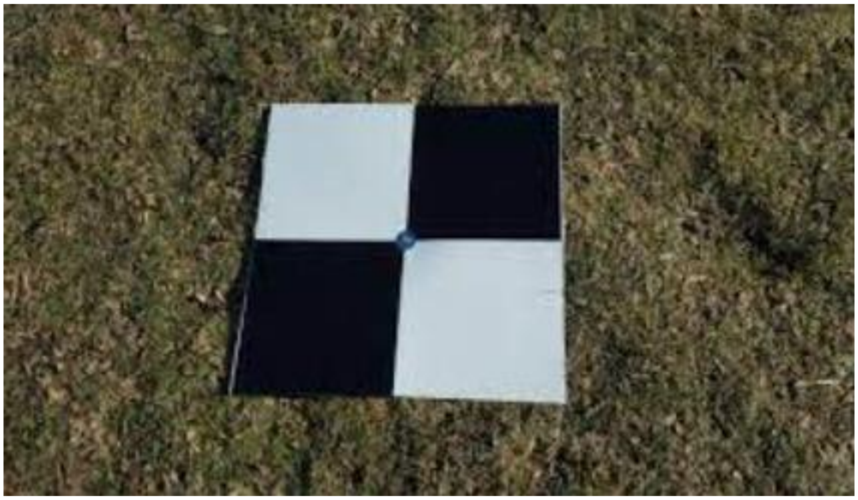


Figure 11: Ground Control Point maker (CGP) used in the field

2.3.4 Multispectral UAV image acquisition

Between September 2020 and November 2020, the drone was flown for data acquisition; the weather condition favored the drone flight because the sky was fully cloudy with no wind and rain. The Parrot Sequoia camera was mounted on the Phantom-4 UAV platform to acquire multispectral images of the study area. Table 3 shows the Parrot Sequoia camera's sensor bands spectral ranges for UAV multispectral images. The four spectral bands include two at a visible wavelength (Green: 530nm - 570nm and Red: 640nm - 680nm), red edge (RE) band (730nm - 740nm), and near-infrared (NIR) band. (770nm - 810nm).

Table 3: Parrot Sequoia camera's sensor bands and the spectral ranges

Sensor Bands	Spectral Range
Green	530nm - 570nm
Red	640nm - 680nm
Red edge	730nm - 740nm
Near-Infrared	770nm - 810nm

The flight pattern, altitude, speed, camera angle, side, forward overlap, and spatial resolution were some of the essential characteristics of the flight. Table 4 shows the flight settings chosen for this study enhance dense point cloud formation while also allowing for the acquisition of very high-spatial-resolution images. The branch, stem, and leaf structure may be easily resolved using the 3D model built using the specified parameters (Kellner et al., 2019). The camera was calibrated prior to enhance higher accuracy in the captured images (Cramer, Przybilla, & Zurhorst, 2017).

Table 4: UAV flight parameters and corresponding variables

No	Parameter	Value (Multispectral)
1	UAV Images	Sequoia multispectral sensor
2	Camera Angle	90o
3	Forward Overlap	90%
4	Side Overlap	80%
5	Speed	Slow
6	Altitude	100-110meters
7	Spatial resolution	0.10493 cm x 0.10493 cm

In addition, the flight height was determined based on the requested UVA multispectral image properties. A flight height of between 100 and 110 meters was chosen to decrease the effect and eliminate ambiguities that arise as the distance between sensors and foliage increases (Seifert et al., 2019). The flight height and the pixel size were determined based on the requested image properties. The spatial resolution of the UAV multispectral images was at 10.49 cm x 10.49 cm. The WGS 1984 UTM Zone 32N projected coordinate system was employed as a datum for the Netherlands. A camera angle scan of 90° was also considered to guarantee a high spatial resolution of 10.49 cm x 10.49 cm. Furthermore, a camera view angle of 90 degrees provides the best point cloud density, reducing mistakes that could impair tree canopy estimation. Moreover, the UAV's slow speed allows for creating dense point clouds for a 3-dimensional model (Dandois et al., 2015; Hirschmugl et al., 2007; Kellner et al., 2019).

Figure 12 shows a double grid flight plan with 80 percent side overlap and 90 percent forward overlap utilized to optimize point cloud density.

Figure 12: Double grid flight plan with forward and side overlap used in this study

2.3.5 Field data collection

The Avenza Maps App was used for navigation and capturing tree locations (X Y) coordinates for different bark beetle infestation and healthy European spruce. Avenza Maps software is an open-source App that can be downloaded on a mobile phone and can load a georeferenced base map for use during the navigation process. The App reduces setup time at the site and is highly accurate (Nyan, Kamarudin, Wahab, & Ismail, 2018). Further, the field data collection was conducted between October 17 and December 14, 2020. The total samples of 263 were collected from the three selected blocks, including 80 healthy trees, 88 green, and 95 red attacks. Block number, infestation stage, tree location using X, Y coordinates, and photos of different infestation stages are among the field parameters collected. The total samples captured from each of the three sampled blocks are shown in Table 5. Figure 13 depicts photos of different stages of bark beetle infestation on European spruce trees in the study area, including healthy, green, red, and grey attacks, respectively.

Table 5: Total number of samples collected at each infestation stage from the three sampled blocks

Block No	Infestation Stage	Total Number of Trees Sampled
1		
	Healthy	26
	Green attack (Stage 1)	33
	Red attack (Stage2)	45
Total		104
2		
	Healthy	15
	Green attack (Stage 1)	23
	Red attack (Stage2)	29
Total		67
3		
	Healthy	39
	Green attack (Stage 1)	32
	Red attack (Stage2)	21
Total		92
Total Number of Trees Sampled in the Selected Three Blocks		263

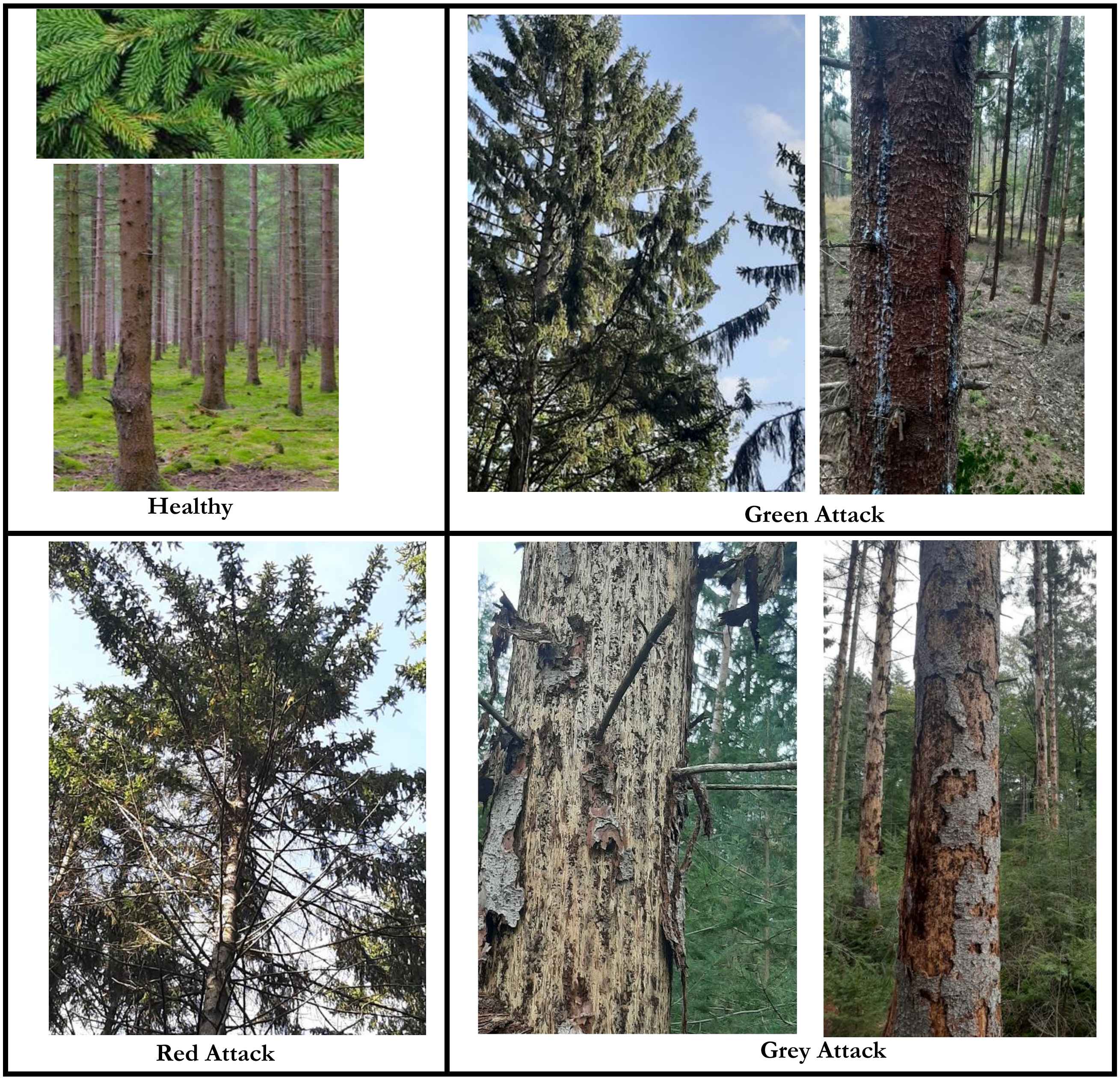
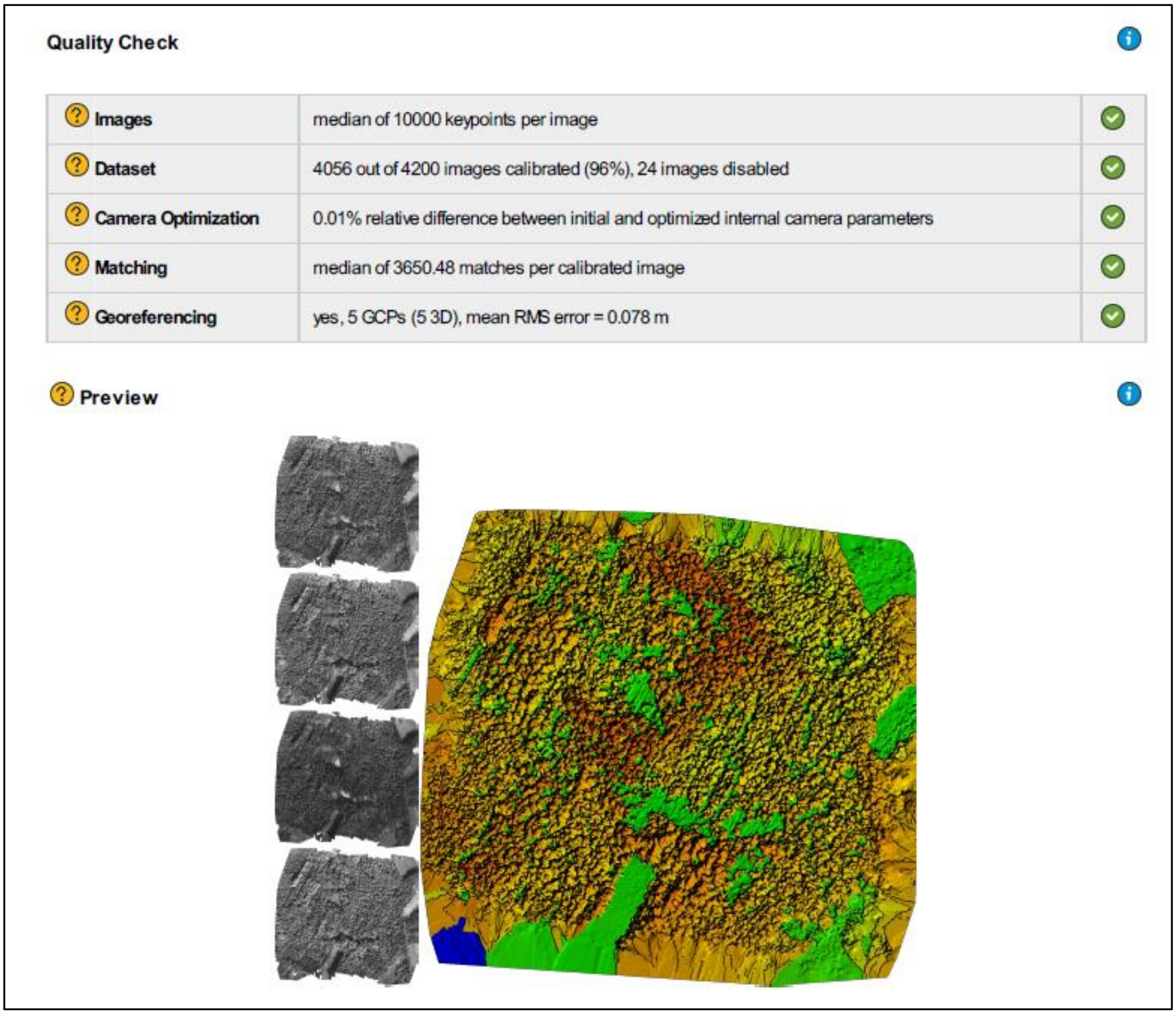


Figure 13: Samples of bark beetle infestation stages on European spruce captured within the study area

2.4 UAV data processing (orthophoto generation)

Pix4D software was used to process the UAV multispectral sequoia camera data based on the photogrammetric process of creating a three-dimensional scene from many overlapping two-dimensional UAV multispectral images. The output of the photogrammetric process is the orthophoto. For absolute orientation, a different number of ground control points (GCP's) were collected using the Real-Time Kinematic Global Navigation Satellite System (RTK GNNS). All three sampled blocks used a different number of checkpoints. Checkpoints are a different set of GCP's used to assess the accuracy of absolute orientation (Wingtra, 2021). According to the initial quality reports generated by Pix4D mapper after UAV data processing: among the three sampled blocks, block one had six (6) GCP's with two (2) checkpoints with the Root Mean Square Error (RMSE) of 0.7 cm. Block two had five (5) GCPs with three (3) checkpoints giving the RMSE of 0.78 cm; finally, block three had three (3) GCPs with three (3)checkpoints having the RMSE of 2.87 cm. According to Stöcker et al. (2020), the RMSE ranges between 5 and 20 cm when three to six GCPs are utilized for absolute orientation. As a result, the RMSE for the three blocks in this study is highly accurate. Figure 14 shows the initial quality report for block two generated by Pix4D Mapper for UAV Multispectral image. Block one and three initial quality reports are shown in appendix 1 and 2 for reference.


Quality Check

Images	median of 10000 keypoints per image	
Dataset	4056 out of 4200 images calibrated (96%), 24 images disabled	
Camera Optimization	0.01% relative difference between initial and optimized internal camera parameters	
Matching	median of 3650.48 matches per calibrated image	
Georeferencing	yes, 5 GCPs (5 3D), mean RMS error = 0.078 m	

Preview

Figure 14: Initial quality report for block two generated by Pix4D Mapper for UAV multispectral Image

2.5 UAV multispectral color composite generation

The generated UAV multispectral orthophotos had four bands. The bands include band_1 red (R), band_2 green (G), band_3 red edge (RE), and band_4 near-infrared (NIR). The generated bands were later combined to create a color composite using ArcGIS software. The color composite was essential for this study because the analysis required combining some spectrum of parrot sequoia sensors. Furthermore, the generated color composites were displayed using bands 1,3 and 4. The generated bands and color composites for blocks one, two, and three are shown in appendix 3 for reference.

2.6 European spruce overlayed with field data

The generated UAV multispectral image composites for the three sampled blocks were overlaid with manually digitized tree crowns, representing different bark beetle infestation stages using ArcGIS software. The overlaid digitized tree crowns were based on the sampled European spruce from the field. The digitized tree crown includes healthy, green, and red attacks, respectively, as shown in figure 15. Blocks one and three overlaid with field data are shown in appendixes 4 and 5 for reference. The total number of trees sampled at each stage from the three respective blocks is shown in table 5.

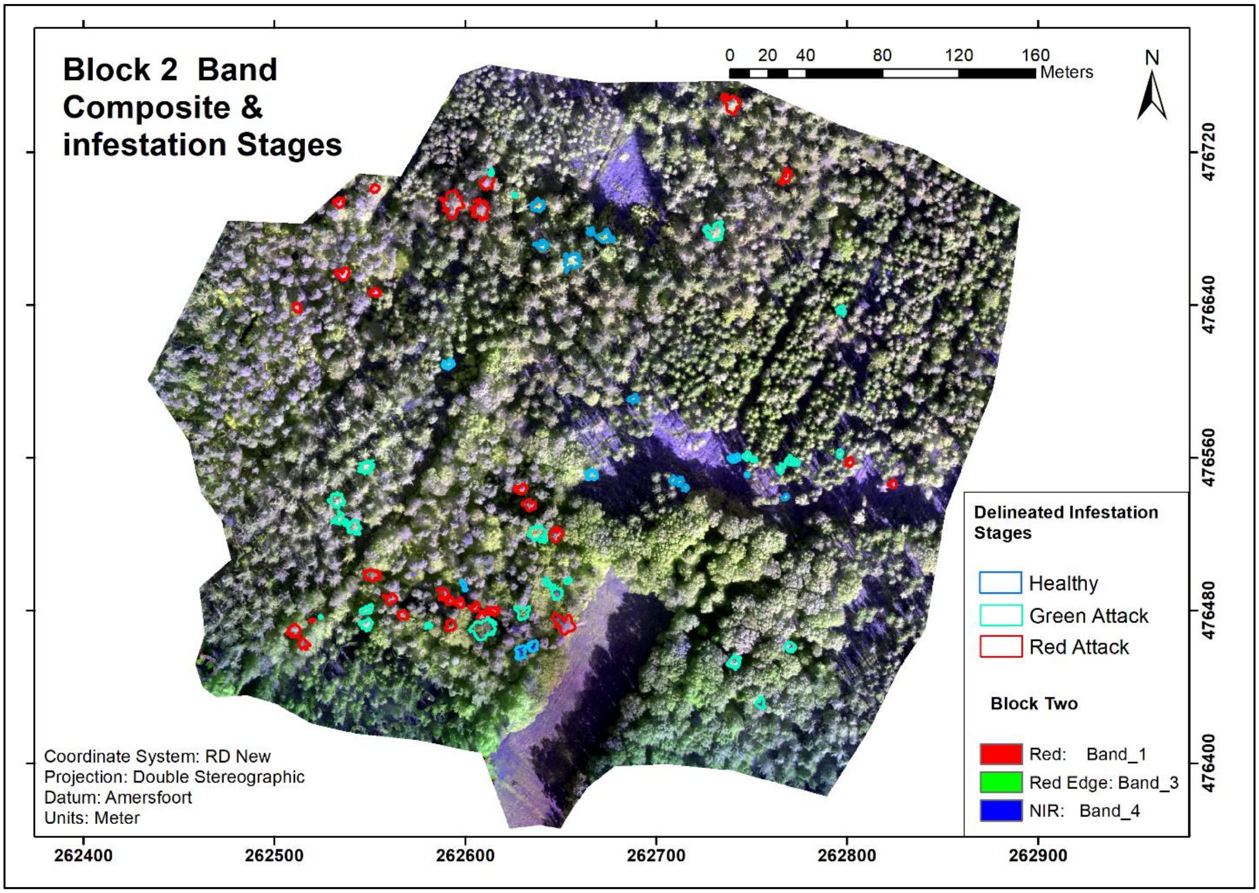


Figure 15: European spruce for block two overlaid with manually digitized tree crown polygons

2.7 Image segmentation on block one, two, and three

Image segmentation is a technique of splitting an image into non-overlapping categories. Segmentation uses Object-Based Image Analysis (OBIA). The process of image segmentation was carried out using eCognition Developer software. OBIA is one of the methods that is used in delineating images into meaningful image objects. OBIA evaluates image characteristics through spatial, spectral, and topological relationships (Ágnes, Eszter, & Ágnes, 2010). OBIA offers an operational structure for the machine-based interpretation of complex classes as well as hierarchical properties. In addition, OBIA analyses images as object spaced rather than pixel spaced, and images are used as primitive for classification and not pixels (Blaschke, Lang, & Hay, 2008).

2.7.1 Multi-resolution segmentation

For this research, multi-resolution segmentation, an eCognition Developer technique mainly applied in very high-resolution images, was used. The method is based on regional growing by joining each seed to neighboring pixels around it with similar characteristics such as color, shape, smoothness, and compactness. Multi-resolution segmentation was used in this study to reduce the heterogeneity of objects while enhancing the homogeneity, which results in significant expected objects. Before segmenting the image, the technique requires a few object parameter values to produce favorable results. The parameters include image layer weight, scale, shape, and compactness (Chen, Chen, & Jing, 2019).

Further, the object primitives for all three blocks were created to attain an accurate image segmentation process based on image weight, scale, shape, compactness, and classes. Figure 16 shows the ruleset of the primitives for blocks one and two. The ruleset for block three is shown in appendix 6 for reference.

Image weight

Image layer weight outlines the influence of each band in the segmentation procedure. For this study, the image used was a UAV multispectral with four bands. The layers were assigned 1.1.4.4. It had the red edge and near infra-red with the highest weight of 4 because of its highest reflectance in vegetation. On the contrary, other bands (Red and Green) were assigned the weight of 1 (Munyati, 2018).

Scale

The scale defines the limit of heterogeneity allowance in the segmented objects. A larger-scale enhances the production of larger objects with a great extent of heterogeneity. Besides, choosing the small-scale value optimizes the production of homogeneous segments. The scale parameter of 40 was used in this study to enhance homogeneous segments (Munyati, 2018). In addition, Jing, Hu, Noland, & Li (2012) states that the appropriate scale is also related to the size of image objects in a forest stand, as it defines the range and dominating size of crowns in a specific forest area.

Shape

The shape parameter specifies the spectral values of the image layers that affect the segmented object's heterogeneity. The higher the shape value, the more spatial uniformity and less spectral uniformity. Likewise, the lower the shape value, the higher the chances of having homogeneous spectral segments. In this study, the shape value of 0.9 was used to maximize segments' production with high heterogeneity because of different reflectance resulting from different bark beetle infestation stages and minimal weight spectral uniformity (Munyati, 2018).

Compactness

Compactness gives a relative weighting against the smoothness of the segmented image borders. The higher the compact value, the more compacted the segments are. In the case of trees, the compactness is low. Therefore, in this study, the default value of 0.5 was used to strike a balance between compact and non-compact object segments (Munyati, 2018).

Classes and shadow masking

Two classes were generated, which include shadow and trees. The shadow was separated from the trees using the near infra-red brightness value greater than or equal to 18000. Further, the trees were separated with a brightness value of greater than 18000 and less than 65500. Shadow masking was used to remove the shadow from the segmented image. In addition, the trees were merged to have a continuous tree object. Masking removes undesired objects (Belgiu, Drăguţ, & Strobl, 2014).

Morphology technique

The morphology technique has two operations which include open and closed image objects. Open image object, isolated pixels are deleted from image object as contrary to the closed image object. All surrounding pixels are added to the object. Under the morphology technique, there are two types of masks: square and circular; the circular technique enables circular masks (Srisha & Khan, 2013); for this study, the circular procedure was applied because of the tree crown's shape with a width size of 10. Usually, the width size depends on the expert knowledge; it's a try and error approach based on the shape and size of the mask you want.

Watershed transformation

Watershed transformation is the process of splitting clustered crowns into individual tree crowns using a splitting threshold. The threshold is determined by the operator using the expert's knowledge (Beucher, 2000). When water is introduced into the watershed system, the valleys collect water from the local minimum. The water will fill the valleys until it overflows into the adjacent valley. When individual watersheds are close enough to touch, the objects are divided to reduce crown clustering. When performing a watershed transformation in the forest, the clustered trees are treated as a watershed under the flooding hypothesis. The trees are divided into valleys that contact each other and become individual trees (Wang, Gong, & Biging, 2004). In this study, the threshold was set at 40 as the length factor.

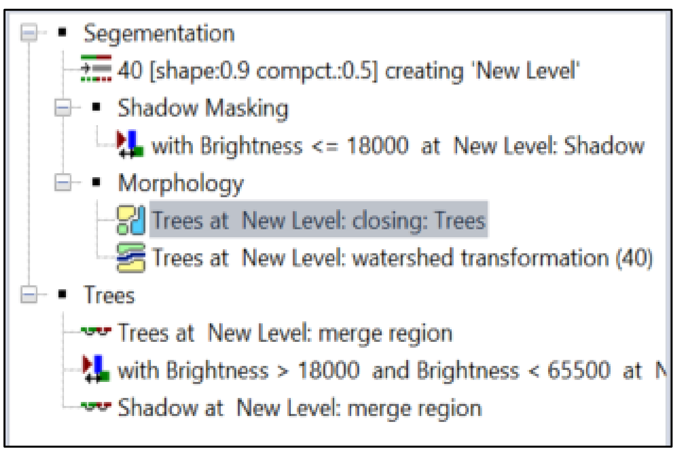


Figure 16: Ruleset under the process tree for block one and two used for image segmentation

2.8 Feature extraction

The classification of different bark beetle infestations on European spruce on the two proposed machine learning classifiers (RF and SVM) requires the extraction of objects' features. The extracted features distinguish different bark beetle infestation stages and other unnecessary objects from the image. Therefore, the features were extracted from the segmented objects. The feature extracted for this study was spectral (brightness) (Eshetae, 2020). Feature extraction identifies pixels with unique traits in the image; the unique trait for spectral features is color. Color space is the most fundamental spectral feature, and it is concerned with the color distribution over the image. The color histogram, which shows the joint probability of the intensities of color channels, is one of the primary techniques used to extract this information. In spectral feature extraction, the image is processed using a pixel matrix; each pixel signifies the brightness and color level found matching the image location (Lester, 1998; Kunaver & Tasič, 2005). Since the different bark beetle infestations have different reflectances, the spectral feature will help distinguish the infestation stages since it is concerned with the color distribution over the image.

Furthermore, an image-layer histogram from the eCognition Developer was utilized to generate the mean and standard deviation brightness values used to distinguish infested European spruce from other items. The mean and standard deviation brightness values for each band, red, green, red edge, and near-infrared, are among the variables extracted under features, as shown in table 6. The image layer brightness mean and standard deviation from the participating pixels in each segmented object was calculated to retrieve vector features from the image (Fazal-E-Malik & Baharudin, 2011). The features such as mean and standard deviation brightness values usually solve the challenge of disparity in the brightness of the features (Machala & Zejdová, 2014). The higher the mean and standard deviation brightness values, the higher the contrast among features (Fazal-E-Malik & Baharudin, 2011). Figure 17 shows the spatial distribution of extracted features from block three.

Table 6: Extracted features for distinguishing bark beetle infestation stages (mean and standard deviation values) from UAV multispectral bands red, green, red edge, and near-infrared

Bark beetle Infestation Stages	**Extracted Features**
Healthy, Green, Red, and Grey attack	Mean Red **(23930)**, Mean Green **(17240)**, Mean Red Edge **(24684)**, Mean Near Infrared **(21031)**, and Std. Dev Red **(7.25)**, Std. Dev Green **(5.15)**, Std. Dev Red Edge **(7.75)** and Std. Dev Near Infrared **(5.35)**.

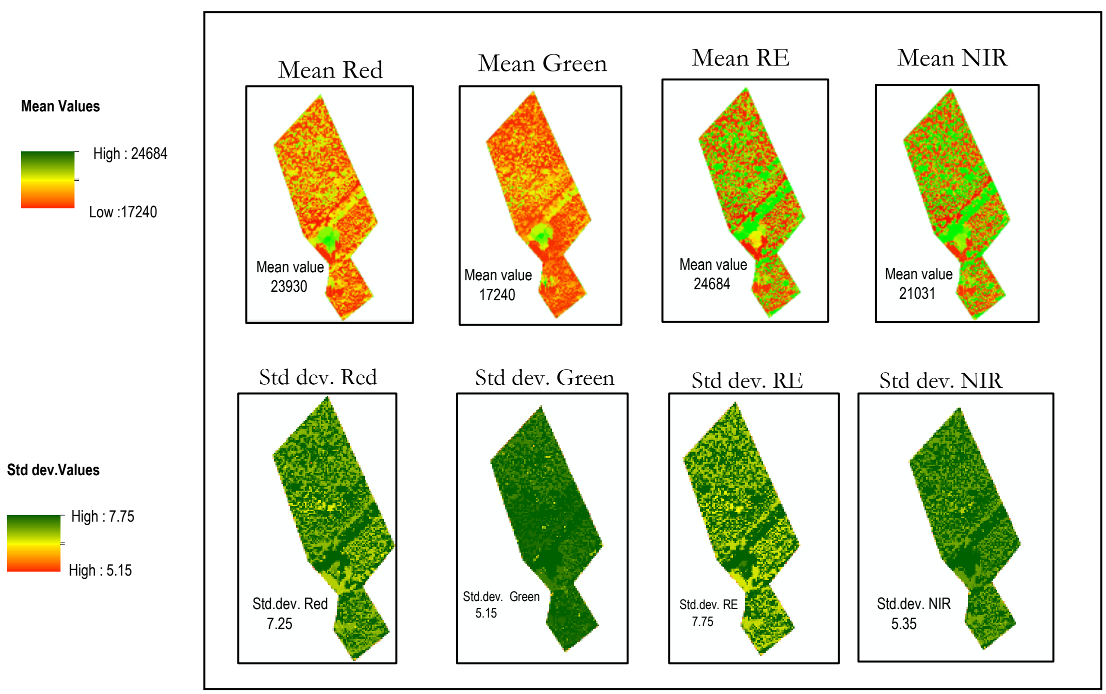


Figure 17: Spatial distribution of extracted mean and standard deviation feature values of image objects for block three

2.9 Segmented objects overlaid on European spruce for block two

The segmented shapefile layer from eCognition Developer overlaid on block two for further analysis, as shown in figure 18. Further, the overlaid segmented shapefile layer overlaid on blocks one and three are shown in appendixes 7 and 8 for reference.

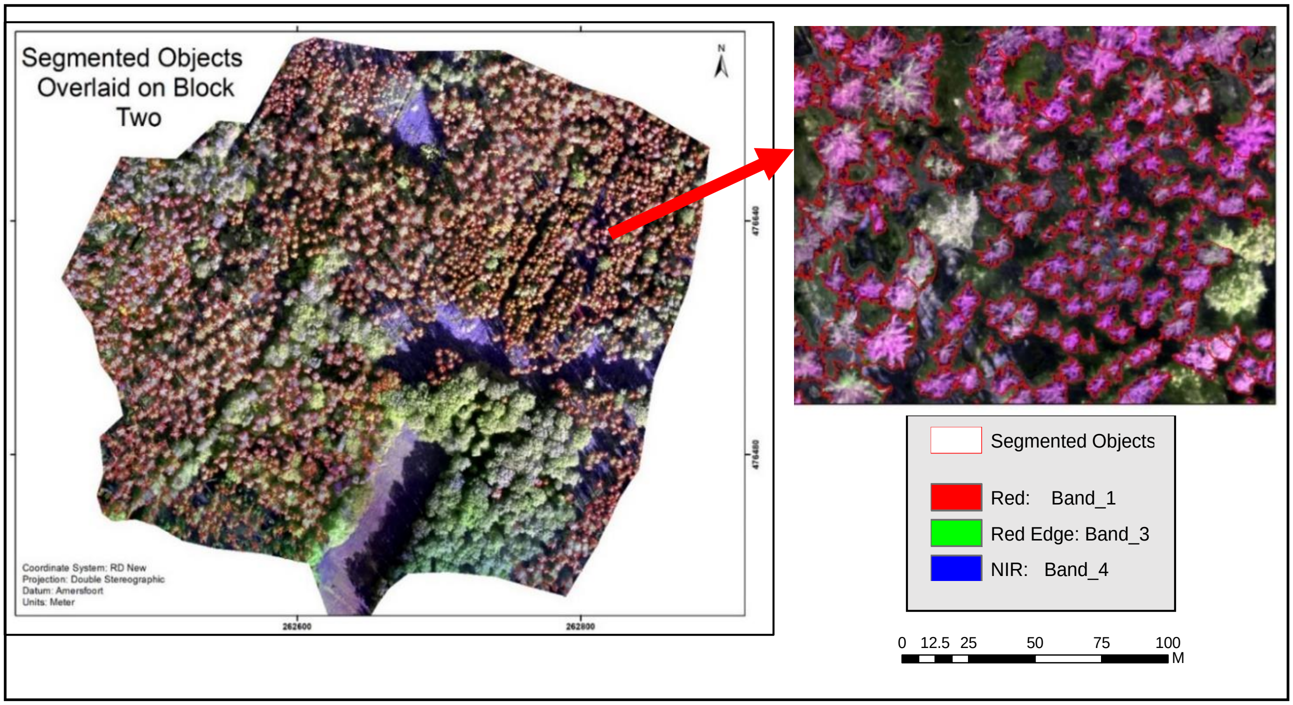


Figure 18: Segmented shapefile layer overlaid on European spruce on block two

2.9.1 Image segmentation accuracy assessment

Image segmentation accuracy assessment was carried out in ArcGIS. The analysis tools using intersect under overlay were applied to attain the results for image segmentation accuracy assessment. Image segmentation accuracy assessment is the process of evaluating segmented objects when compared to reference objects. It involves matching segmented and digitized polygons; the assessment employs over-segmentation and under-segmentation to find the total error. According to El-naggar (2018), over-segmentation is when image segments are divided into smaller segments than the assigned objects. In addition, under segmentation happens when the output segments are larger than the reference polygons within the image. Over segmentation and under segmentation lies within the range of [0,1): where over-segmentation = 0 and under segmentation = 0 if the reference objects are fitting perfectly (Clinton, Holt, Scarborough, Yan, & Gong, 2010). Equations one, two, and three were used in assessing image segmentation accuracy.

Over Segmentation ij = 1 – area (xi ∩ yj) / area(xi). ………Equation 1

Under Segmentation ij = 1 – area (xi ∩ yj) / area(yj) ....Equation 2

$$D_{ij} = \sqrt{\frac{OverSegmentation_{ij}^2 + UnderSegmentation_{ij}^2}{2}}.$$ ....Equation 3

Where;
Xi = area of segmented objects
Yj= Area of reference polygons
Area (Xi ∩ Yj) = Area of intersection of reference polygons and segmented objects
Dij = Total detected error
Accuracy = (1-D) *100

Source: (Nicholas Clinton, Ashley Holt, Li Yanb, 1999)

The manually digitized tree crown polygons for different bark beetle infestation stages collected from the field were used as reference polygons during the image segmentation accuracy assessment process. The healthy stage had 15 polygons, green and red attacks had 23 and 29, respectively, bringing the total number of polygons to 67. The intersection operation was carried out in ArcGIS to determine the intersection of polygons between the reference and segmented objects. The intersected area was then used in equations one and two to ascertain under segmentation and over-segmentation. The results obtained in equations one and two were later used to detect total error to find the overall segmentation accuracy using equation three. Figures 19 show the reference digitized polygon in (green) with segmented objects polygons in (red) used for image segmentation accuracy assessment. The detailed image segmentation accuracy assessment results are shown in table 7. The image segmentation accuracy assessment results for blocks one and three are shown in appendixes 9 and 10 for reference.

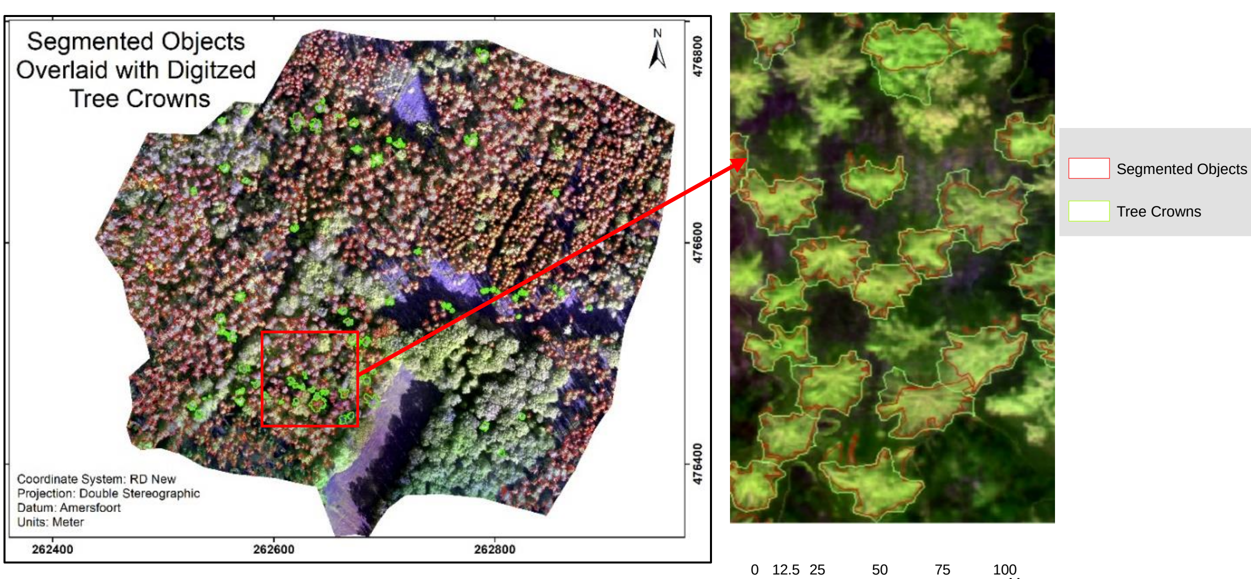


Figure 19: Segmented objects (Red) overlaid with digitized tree crown polygons (Green) in block two

Table 7: Image segmentation accuracy assessment results for block two

Block Two Image Segmentation Accuracy Assessment Results			
Variables	**No of Segments**	**Area/Results**	**Accuracy %**
Segmented Objects	90	2091.92	
Digitized Objects	67	1662.66	
Intersection	84	1658.62	
Under Segmentation	57	0.21	79%
Over Segmentation	20	0.003	99%
Total Detected Error		0.21	
Accuracy%			**79**

2.10 Image classification of bark beetle infestation stages on European spruce

The classification of bark beetle infestation on UAV multispectral images was performed using the three classifiers: RF, SVM, ML, and vegetation indices, NDVI, NDRE, and GNDVI using the Species Distribution Model (SDM) as input variables. Table 8 shows the total number of training samples used at each infestation stage from the three sampled blocks.

Table 8: Total number of samples collected at each infestation stage from the three sampled blocks split up into training and validation samples

Block No	Infestation Stage	Total Number of Trees Sampled	Training Samples	Validation Samples
1				
	Healthy	26	18	8
	Green attack (Stage 1)	33	23	10
	Red attack (Stage2)	45	31	14
Total		**104**	**72**	**32**
2				
	Healthy	15	10	5
	Green attack (Stage 1)	23	16	7
	Red attack (Stage2)	29	20	9
Total		**67**	**46**	**21**
3				
	Healthy	39	27	12
	Green attack (Stage 1)	32	25	7
	Red attack (Stage2)	21	15	6
Total		**92**	**67**	**25**
Total Number of Trees Sampled in the Selected Three Blocks		**263**	**185**	**78**

2.10.1 Random Forest classifier

The classification of RF was performed based on the extracted features from the eCognition Developer, as shown in table 6. QGIS software was used to carry out RF classification. The digitized polygons for bark beetle infestation stages were used as training data. A polygon is a circle or an irregular shape drawn over an object used to extract spectral information about the object's pixels. The process of determining each class's spectral cover based on a priori knowledge of the image data and objects to be mapped is known as training. The data used in this process is usually acquired through a field survey. The digitized polygons were employed as training data for the study because the classification algorithm classifies the remainder of the image based on the training locations, resulting in the most accurate classification results (Green et al., 2000). Furthermore, the extracted features (mean and standard deviation brightness values) were also employed as input variables to categorize distinct bark beetle infestation stages on European spruce, as illustrated in figure 17.

The classification of image objects produced from segmentation in RF was based on extracted features (mean and standard deviation brightness values). The extracted features were used as predictors to classify different bark beetle infestation stages on European spruce. The classes were divided into three classes which include healthy, green, and red attacks, respectively. The RF classifier constructs trees by replacing a subset of training samples. 70% of the total collected samples of different bark beetle infestation stages were used to constructs trees by replacing a subset of training samples to train the trees. The classifier requires two-dimensional data as an input variable; in this study, the X, Y coordinate tree location was used. While 30% of the total samples referred to as out-of-the-bag samples were used for internal cross-validation procedure for assessing how well the derived RF model performed. Averaging the class assignment probabilities calculated by all trees resulted in the final classification decision (Akayezu et al., 2019; Belgiu & Drăgu, 2016). Figure 20 shows data from a two-dimensional space input variable. A decision tree with a hierarchical structure of connected nodes is shown in figure 21. Figure 22 depicts how the RF classifier is applied in distinguishing particular classes. The bootstrapping procedure, which involves selecting and replacing random samples, was used to arrive at the class with the most votes to be chosen by the classifier (Liarokapis, Artemiadis, & Kyriakopoulos, 2013). The total number of samples for block three can be referred to in table 9, and the decision tree was set to N-10.

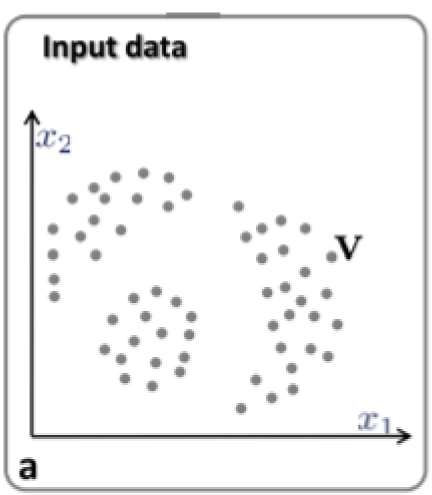
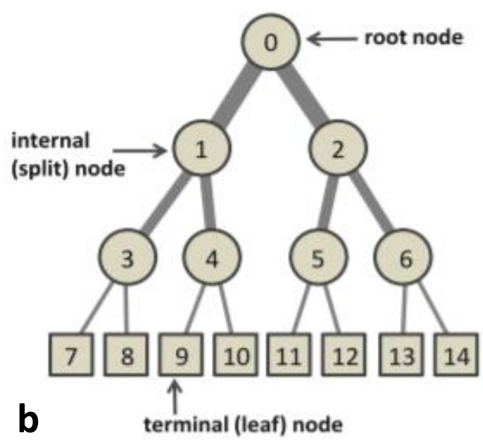


Figure 20: Input data, a collection of 2-dimensional space



Figgure 21: A decision tree showing the hierarchical structure of connected nodes

Source: (Criminisi, A., & Shotton, 2013)

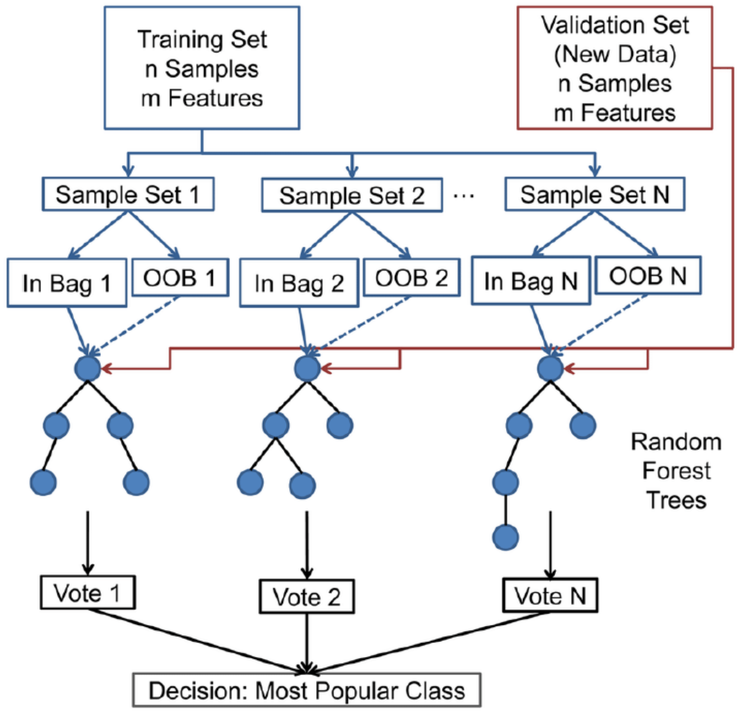


Figure 22: Random Forest classification technique for N trees using the out of bag (OOB) samples

Source: (Liarokapis et al., 2013)

2.10.2 Support vector machine classifier

The classification of image objects produced from segmentation in SVM was based on extracted features (mean and standard deviation brightness values). The extracted features illustrated in figure 17 were used as predictors to classify different bark beetle infestation stages on European spruce. The classes were divided into three classes which include healthy, green, and red attacks, respectively. SVM classification was carried out using ENVI software. The training regions for each class (healthy, green, and red attack) were imported into Region of Interest (ROI) since they were already digitized in ArcGIS. Regions of Interest (ROIs) are selected samples that are generally utilized for training the model using field data (Kiadtikornthaweeyot & Tatnall, 2016). The support vectors (bark beetle infestation stages), representing 70% of the total collected samples for block three, as shown in Table 9, were used as training data (Akayezu et al., 2019). The training samples were selected randomly from the total samples. The support vectors that lie closer to the hyperplane were selected from the training data to measure the distance between the line and samples. The distance between the samples and the line is called the margin. Therefore, the hyperplane with the long-distance was considered as the optimal hyperplane. The vector points facilitated the model development and separated the classes based on the linear function. SVM aims at searching for an optimal decision hyperplane in high dimension spaces, which enhances optimal separation of classes. Figure 23 (A) depicts various hyperplanes in SVM, whereas (B) depicts the best hyperplane for separating the different stages of bark beetle infestation on European spruce.

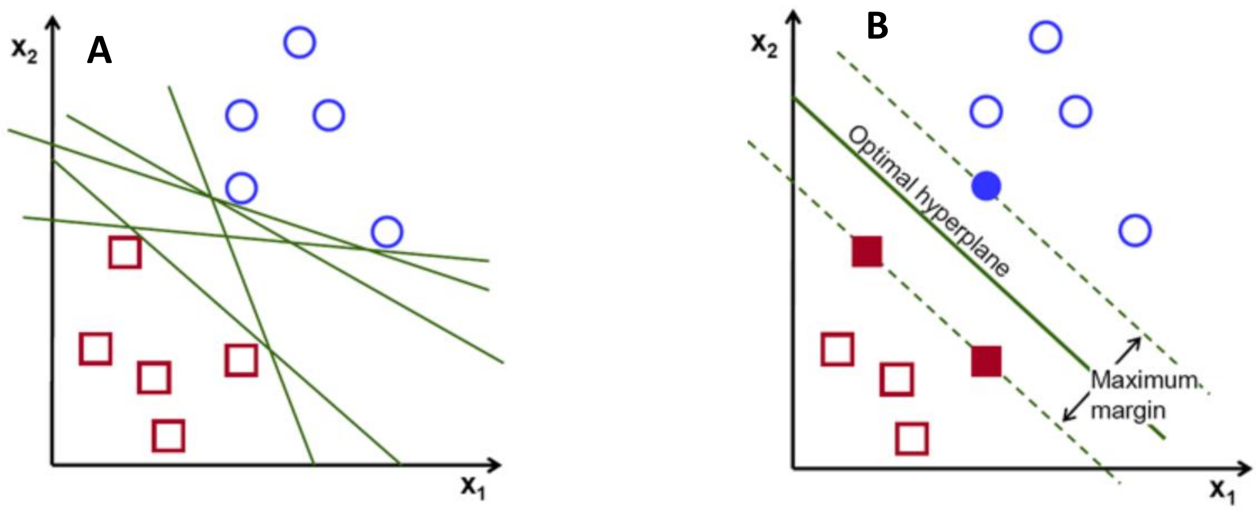


Figure 23: (A) Possible hyperplane in SVM *(B) Optimal hyperplane to separate the classes in SVM*

Source: (Rohith Gandhi, 2018).

2.10.3 Maximum Likelihood Classifier

Maximum likelihood classification was carried out using Erdas Imagine software. ML classification is a supervised classification that requires digitized polygons called Area of Interest (AOI). A digitized polygon is a circle or an irregular shape drawn over an object used to extract statistical information about the object's pixels (Green et al., 2000). Therefore, digitized polygons of European spruce tree crowns representing different bark beetle infestation stages were used as AOI layers. AOIs are selected samples that are generally employed for training the model during image classification using field data (Kiadtikornthaweeyot & Tatnall, 2016). The AOI was used to create the spectral signature file to train the UAV multispectral image. The spectral signature is the statistical information about the object's pixels used to distinguish objects based on spectral differentiation (Schuckman, 2020). Three classes under the spectral signature file for this study include healthy, green, and red attacks, respectively. In addition, 70% of the total collected samples for all three blocks were used as training samples (Akayezu et al., 2019). Furthermore, the training samples were selected randomly from the total samples. The total number of samples used for ML image classification at each infestation stage for block three is shown in table 9. During the ML classification, the decision rule parameters were set as follows: Non-parametric rule was set to non because ML is a parametric classifier. On the other hand, the parametric rule was set to maximum likelihood since the classifier is based on a bigger maximum chance that the pixels belong to a specific class, resulting in improved classification results (Kumar & Miller, 2006). The three methods (RF, SVM, and MLC) were applied on block three to determine which classifier gives higher results in detecting the different bark beetle infestation stages. Therefore, on the two remaining blocks (one and two), the best-performing classifier was used.

2.11 Image classification accuracy assessment

The confusion matrix is the focal point of image classification accuracy assessment. It is an essential part of supervised classification. The main components of image classification accuracy assessment include overall accuracy, kappa coefficient, user accuracy, producer accuracy (Foody, 2002). The confusion matrix is a technique that is used to determine the accuracy of supervised image classification. The confusion matrix is the table showing the relationship between the reference data and the classified image. To produce a confusion matrix, ground truth validation data is required (Lewis & Brown, 2001). In this study, 30% of the total collected samples represented by manually delineated tree crowns of different bark beetle infestation stages were used for validation to determine image classification accuracy assessment (Akayezu et al., 2019). Furthermore, the validation samples were selected randomly from the total samples.

Overall accuracy

The overall accuracy indicates the total amount of the reference data that were classified correctly. The overall accuracy is in percentage, with 100% perfect if the ground truth data were mapped and classified correctly (Gómez & Montero, 2011).

User accuracy

User accuracy is the accuracy of the classified map from the map user point of view: It also shows the class's frequency on the map concerning the study area. User accuracy is also known as a measure of reliability, expressed in percentage (Pavel, 2016).

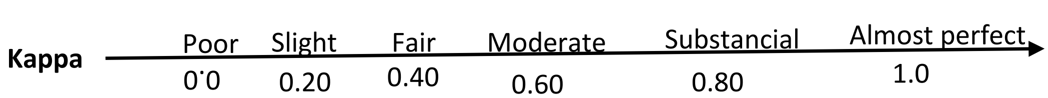
Producer accuracy

Producer accuracy is the accuracy of the classified map from the map maker's point of view. It shows the frequency of how accurately the features in the study area are shown on the classified image (Pavel, 2016).

Kappa statistics

According to Viera & Garrett (2005), Kappa statistics measure how closely the field data matches the classifiers. The kappa statistics also weigh and evaluates how well the classification has performed. The value range of kappa statistics is from 0 to 1 as shown in figure 7: <0 means the poor chance of agreement, 0.01 to 0.20 minor agreement, 0.21 to 0.40 fair agreement, 0.41 to 0.60 balanced agreement, 0.61 to 0.80 significant agreement, and 0.81 to 0.99 almost perfect.

Interpretation of Kappa



Source : (Viera & Garrett, 2005)

2.12 Performance of vegetation indices (GNDVI, NDRE, NDVI) in predicting the green attack of bark beetle infestation stage on European spruce

The vegetation indices' determination to predict the green attack of bark beetle infestation stage on European spruce requires the spectral band combination from the UAV multispectral image. Therefore, to determine the influence of the four bands (red, (green, red edge, and near-infrared) in predicting bark beetle infestation stage in the participating vegetation indices; spectral separability and reflectance curve for bark beetle infestation stages based on field data was generated

The spectral separability concept distinguished the major influencing bands in spectral separability among the four bark beetle infestation stages in the selected vegetation indices (GNDVI, NDRE, and NDVI) (Abdullah, Skidmore, Darvishzadeh, & Heurich, 2019a). The spectral reflectance curve was plotted based on different bark beetle infestation stages obtained from the field. The spectral separability reflectance curve is the graphical representation of the spectral response of different bark beetle infestations over different wavelengths of the electromagnetic spectrum (ITC, n.d.).The red and grey attack shows higher reflectance in red, red-edge, and near-infrared; this can be related to the loss of chlorophyll, a pigment that absorbs red as the leaves turn yellow-brown to grey due to bark beetle infestation (Heenkenda et al., 2015). Healthy and green attacks have a similar pattern; they both have higher reflectance in the green and NIR bands because of chlorophyll content in the needles. Besides, red edge and red had higher spectral separability as compared to green and near-infrared. The spectral separability reflectance curve was plotted using Excel software.

Nevertheless, the red and grey attack also shows high reflectance in the red edge and slightly lower in the NIR band. This trend is called the red-edge shift; this happens because the center wavelength of the red edge between the red and NIR bands makes a little shift to the shorter wavelength. The contributing factor for the red-edge shift is the decreased vitality due to plant stress (Immitzer & Atzberger, 2014). Figure 24 shows the spectral separability and reflectance curve of bark beetle infestation stages on European spruce using a UAV multispectral image

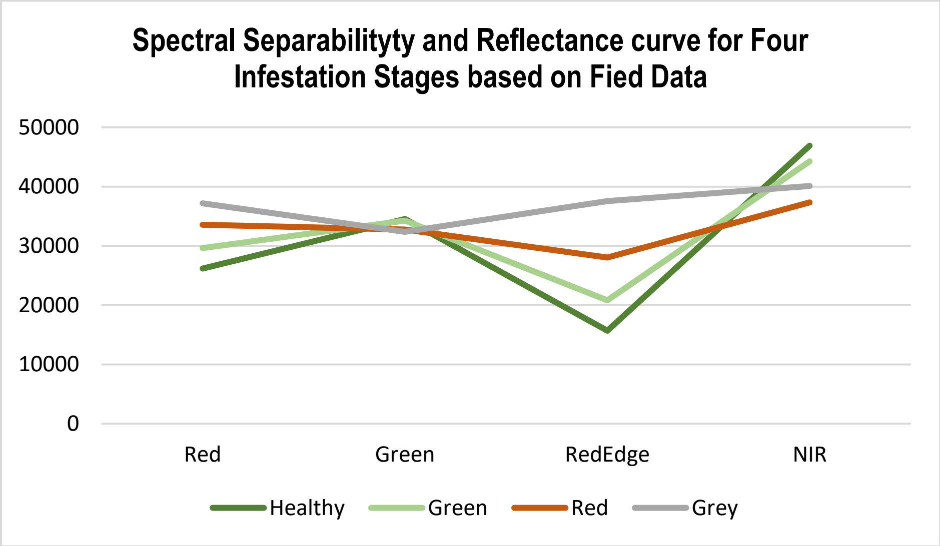


Figure 24: Spectral separability and reflectance curve generated based on sampled bark beetle infestation stages in the field

2.12.1 Vegetation indices calculation

The vegetation indices that will be applied in this study include; NDVI, GNDVI, and NDRE (da Silva et al., 2020). According to Abdullah (2019), the selected vegetation indices have been used in previous studies to assess stress-induced variations in water and chlorophyll content, which are essential parameters of tree healthy. They have proved to be sensitive to any slight change in chlorophyll content. The sensitivity has been used to evaluate tree stress resulting from stomata closure due to bark beetle infestation. The stomata closure reduces pigment absorption, leading to an increased reflectance from the leaves in visible wavelength and red edge and near-infrared bands.

Besides, vegetation indices such as GNDVI, NDRE, and NDVI have been used to measure vegetation's greenness to determine vegetation health or stress in the coniferous forest (Eitel et al., 2011; Ballester, Brinkhoff, Quayle, & Hornbuckle, 2019). GNDVI is similar to NDVI; the only difference is that it measures the green spectrum in the micron range, which is the photosynthetic indicator in vegetation. It is very sensitive to chlorophyll concentration and is mainly used in determining plant stress and aged vegetation (Abdullah et al., 2019b). NDRE is a good indicator of healthy vegetation that has accumulated high chlorophyll levels in the leaves. This is because the red edge light is very sensitive to abrupt small chlorophyll content changes (Oliech, 2019). Table 9 shows the equations for the three selected vegetation indices to be used in this study.

Table 9: The three equations for NDVI, GNDVI, and NDRE vegetation indices

(NIR-R)/(NIR+R)	The measure of greenness in vegetation	(Ballester, Brinkhoff, Quayle, & Hornbuckle, 2019)
(NIR- Green)/ (NIR+ Green)	Sensitive to chlorophyl concentration	(Abdullah et al., 2019b)
(NIR-Red Edge)/(NIR+Red Edge)	Very sensitive to abrupt small changes in chlorophyll content	(Abdullah et al., 2019b)

In this study, three vegetation indices were used to attain objective two: the calculated vegetation indices include GNDVI, NDRE, and NDVI. The three vegetation indices were calculated based on equations 4, 5, and 6 using ArcGIS software. The digitized European spruce tree crowns in blue were overlaid over the vegetation indices to check if the sampled data correctly represented the different back beetle infestation stages. According to Xiao & McPherson (2005), the high values indicate healthy trees regarding vegetation indices. In contrast, the lower values indicate unhealthy trees, bare ground, and water bodies. Therefore, the lower the vegetation index value, the higher the infestation stage (Xiao & McPherson, 2005). Refer to appendix 11 for the three vegetation indices (GNDVI, NDRE, and NDVI) generated from the UAV multispectral image

Green Normalized Difference Vegetation Index (GNDVI)

The GNDVI was calculated based on the reflectance of the red and green bands, as shown in equation 4.

$$GNDVI = \frac{\text{NIR - Green}}{\text{NIR + Green}} \qquad \text{Equation 4}$$

Normalized Difference Red Edge Index (NDRE)

The NDRE was calculated based on the reflectance of the red and red edge bands, as shown in equation 5.

$$NDRE = \frac{\text{NIR} - \text{Red Edge}}{\text{NIR + Red Edge}} \qquad \text{Equation 5}$$

Normalized Difference Vegetation Index (NDVI)

The NDVI was calculated based on the reflectance of the red and NIR bands, as shown in equation 6.

$$NDVI = \frac{\text{NIR - Red}}{\text{NIR + Red}} \qquad \text{Equation 6}$$

2.12.2 European spruce from vegetation indices using object segmented European spruce shapefile Layer

The European spruce trees were extracted from calculated vegetation indices (GNDVI, NDRE, and NDVI) using the object segmented European spruce shapefile layer exported from the eCognition Developer. The European spruce was extracted, leaving all the unwanted objects using ArcGIS software to enhance the performance of the classification results. See appendix 12 for the extracted European spruce.

2.12.3 Extracting vegetation indices values

The standard deviation spectral values were extracted from the three vegetation indices (NDVI, NDRE, and GNDVI) based on the manually digitized European spruce tree crowns for different bark beetle infestations. The zonal statistics as a table under the spatial analyst tool using ArcGIS software was used to extract standard deviation spectral values. The process was done by summarizing the raster values within the zones and report the results to the table. The standard deviation spectral values were used because they usually solve the challenge of disparity in the reflectance of the features (Machala & Zejdová, 2014). The higher the standard deviation value, the higher the contrast among features (Fazal-E-Malik & Baharudin, 2011). The extracted standard deviation spectral values were used as input variables together with different bark beetle infestation stages to predict the green attack in RStudio software for Species Distribution Modelling (SDM) development (Immitzer & Atzberger, 2014). Table 10 shows the input variables for SDM development.

Table 10: Input variables for the SDM

Variable group	Variables	Ranges
Indices Maps	NDVI_Indices values	-1 to 1
	NDRE_Indices values	-1 to 1
	GNDVI_Indices values	-1 to 1
Infestation Stages	Healthy	1
	Green Attack	2
	Red Attack	3

2.13 Predicting the green attack bark beetle infestation stages using SDM

The input data for the proposed model (SDM) in this study comprises the dependent variable, which is the different bark beetle infestation stages. The independent variables include the extracted standard deviation values of vegetation indices from the digitized polygons, which are predictor variables. Therefore, the DSM was used in predicting the presence of bark beetle infestation on European spruces because of its ability to discriminate the presence and absence of occurrence. The occurrence being looked at are the different bark beetle infestation stages, especially the green attack. In addition, vegetation indices were used as input data in the DSM to support the discrimination of the presence and absence of the bark beetle infestation, especially the green attack. However, very little research has been conducted using the SDM in the detection of bark beetle infestation.

The SDM was performed using the RStudio (R) software package. The standard deviation values extracted from the three vegetation indices (GNDVI, NDRE, and NDVI) for bark beetle infestation stages digitized tree crowns; tree location (X, Y) coordinates, and vegetation indices raster maps (GNDVI, NDRE, and NDVI) were the four most important variables involved in predicting the green attack in this study. The four parameters are required for the model to be processed, as shown in figure 27. The vegetation indices' standard deviation values were used as predictors. In contrast, the bark beetle infestation stages were used to indicate the presence (1) and absence (0) of the occurrence of the particular infestation stage (healthy, green, and red attack). The vegetation indices raster maps and tree location X, Y coordinate were used to determine the actual location of the infestation stage being predicted. The diagram representing the process involved in SDM, which includes preprocessing, processing, and post-processing, is shown in figure 25.

2.13.1 Prediction of green attack bark beetle infestation stage using SDM

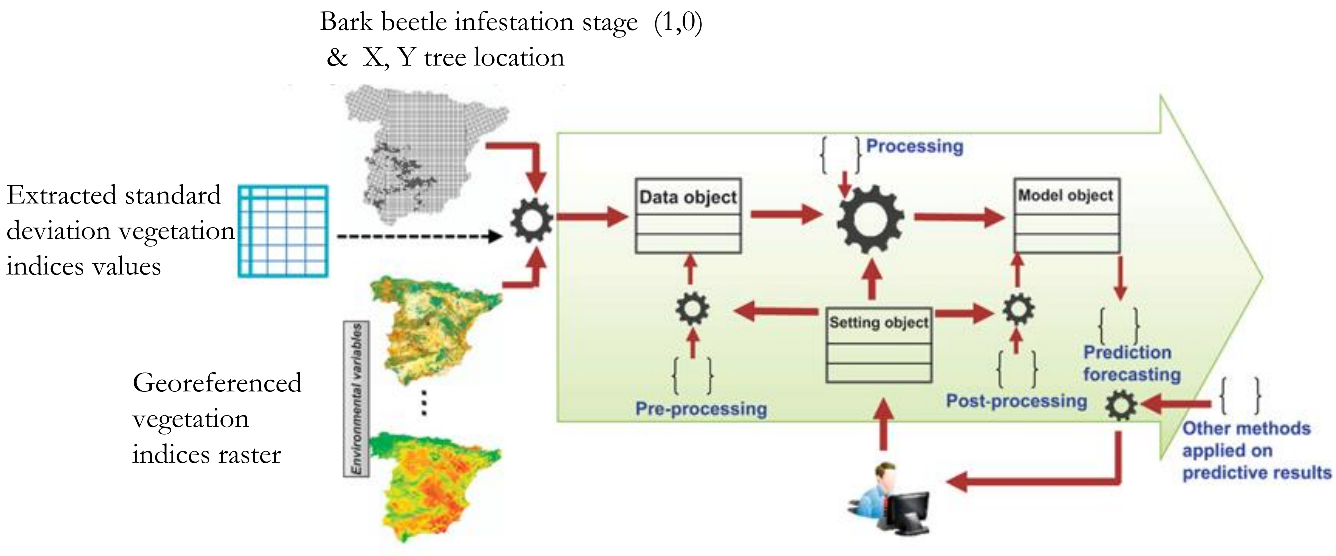


Figure 25: A visual representation of how a species distribution model works, and the essential processes

Adopted from : (Naimi & Araújo, 2016)

The three main steps involved in SDM include preprocessing, processing, and post-processing. Preprocessing is a technique for making all of the SDM's input datasets available for processing. The datasets available for this study were (vegetation indices raster maps, X, Y coordinate tree location, bark beetle infestation stages, and Extracted standard deviation vegetation indices values). Processing involves fitting the model to related response variables (bark beetle infestation distribution) to predictor environmental variables (extracted standard deviation vegetation indices values and vegetation indices raster maps). At this stage, the user selects the modeling procedure. Post-processing of model prediction was based on 70% of the total collected samples for block two; furthermore, the training samples were selected randomly from the total samples. At the same time, post-processing also includes a model evaluation which involves accuracy assessment. At this stage, another independent dataset (test data) accounting for 30 % of the total collected samples from block two was used to assess model predictions; further, the test data were selected randomly from the total samples (Akayezu et al., 2019). Furthermore, the optimal threshold was used to measure accuracy; the optimal threshold is the point at which the specificity is equal to sensitivity. The optimal threshold lies between 0 and 1; the optimal threshold is the value that best separates the presence from the absence. The optimal threshold takes the median of the predicted values of the presence. This means that any pixel value below the optimal threshold is not the green attack stage, except the pixel values greater than or equal to the optimal threshold are classified as the presence of green attack (Perez-Jaume, Pallarès, Skaltsa, & Carrasco, 2017). Further, sensitivity and specificity are statistical measurements of a binary classification test's performance. The accurately identified positives (i.e., the fraction of European spruce at the green attack stage) are measured by sensitivity. Meanwhile, the true negatives accurately detected (i.e., the fraction of European spruce not in the green attack stage) is the subject of specificity (Jiménez-Valverde, 2012). Figure 26 shows the thresholds for GNDVI, NDRE, and NDVI for block two, and table 11 shows the thresholds for the three indices (GNDVI, NDRE, and NDVI), respectively. For further details on how the script was written in RStudio to generate the SDM results, kindly refer to appendix 11. Block two was chosen for the prediction in the SDM because, based on the results generated under the classifiers, it seemed like almost the whole block was affected. Therefore, it was a way of validating the generated results.

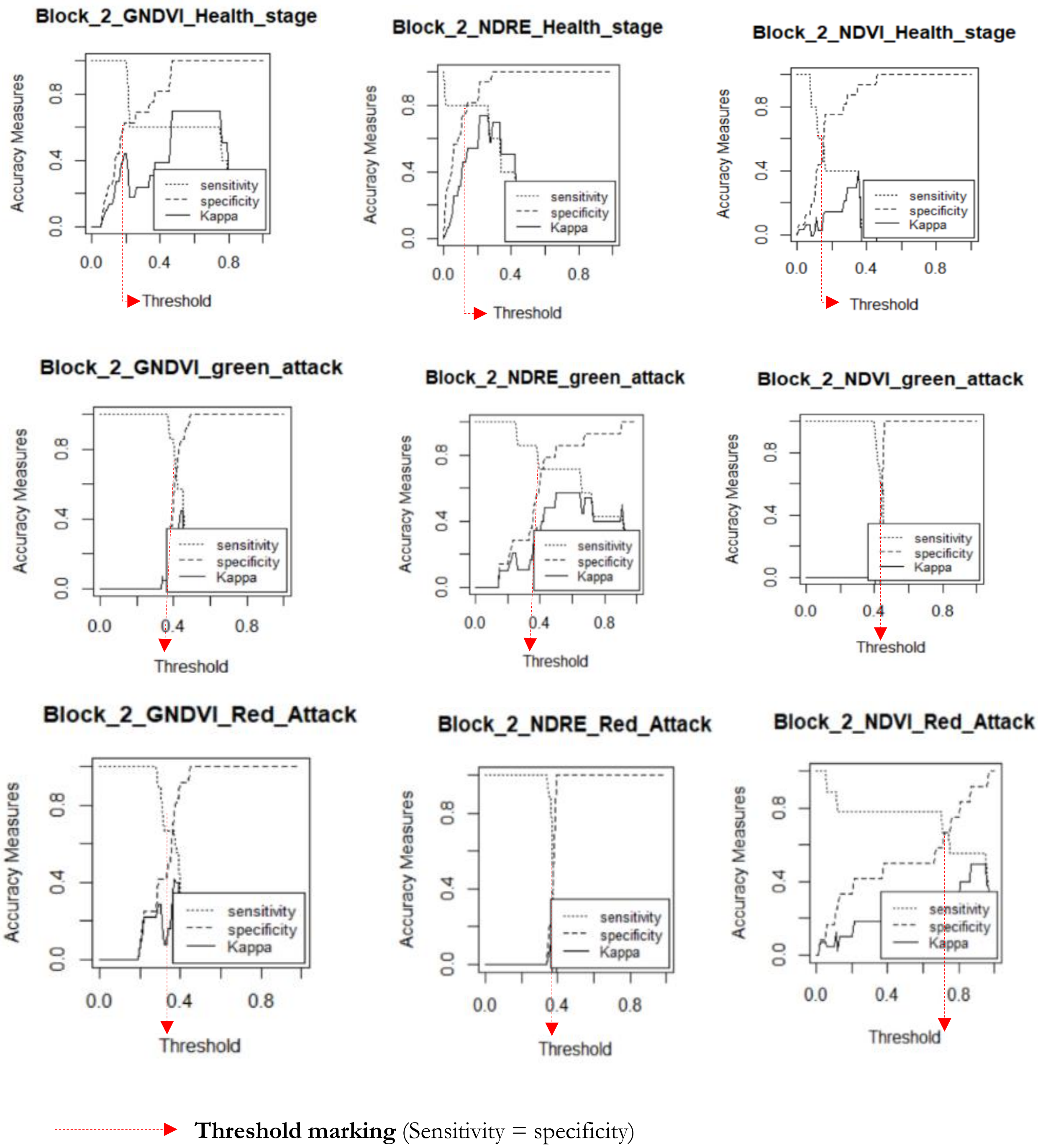


Figure 26: Thresholds for GNDVI, NDRE & NDVI for block two based on the median of the extracted vegetation indices values

Table 11: Thresholds based on the median of extracted vegetation indices values of infestation stages from GNDVI, NDRE &NDVI

Thresholds for Block Two			
	GNDVI	**NDRE**	**NDVI**
Healthy	0.23	0.17	0.35
Green attack	0.41	0.41	0.44
Red Attack	0.36	0.37	0.73

2.13.2 The Area Under the Curve (AUC)

The area under the curve (AUC) was applied to evaluate the accuracy of the performance or predictive power of SDM on the three vegetation indices in predicting the bark beetle infestation stages. The AUC is the graph that examines the model's ability to discriminate the presence and absence of the variable. The false positive (1- specificity) is plotted in the x-axis against the true positive (Sensitivity) in the y - axis. The range of AUC is 0.5 to 1, of which 0.5 is considered a random model, while 1 is a perfect discrimination capacity model. The higher predictive value usually is closer to the left top corner of the AUC graph (Hanley & McNeil, 1982). Figure 35 shows the predictive power of the SDM in detecting the bark beetle infestation stages using AUC.

2.13.3 Testing the significance of the developed models in comparison to field data

The model's statistical significance was based on the Akaike Information Criterion (AIC) and the Pr(>|z|). The AIC and the Pr(>|z|) were used to assess the models' quality and help out to compare the models. The two techniques use deviance as a measure of goodness of fit. The AIC and the Pr(>|z|) measures how well the model fits based on the difference between the fitted model and observed data. The Akaike information criterion (AIC) gives the prediction error and relative model quality for a given set of data. AIC measures the quality of each model in relation to the other models given a set of data models. The AIC is based on presence data only. In addition, the Pr(>|z|) is the test for whether the coefficient point estimate is significantly different from 0; both the (AIC) and the Pr(>|z|) is preferred when the values are low (De Rivera et al., 2017; Gutierrez & Heming, 2018). Thus, the AIC and Pr(>|z|) was used to evaluate the GNDVI, NDRE, and NDVI model's goodness of fit regarding how well the output matches the likelihood of the bark beetle infestation field data used to predict the model's parameters.

3 RESULTS

3.1 What classifier among RF, SVM, and ML gives the highest accuracy in detecting the green attack of the bark beetle infestation stage on European spruce from UAV multispectral images?

The accuracy of the classified images was evaluated through the image classification accuracy assessment procedure. The findings of the three classifiers (RF, SVM, and ML) were compared to see which one performed best in detecting the green attack stage of bark beetle infestation using UAV multispectral images. Figure 27 and 28 shows the classified images for the three image classification techniques: RF, SVM, and ML. In addition, table 12 shows the summary of the confusion matrix for RF, SVM, and ML image classification accuracy assessment for blocks three, two, and one.

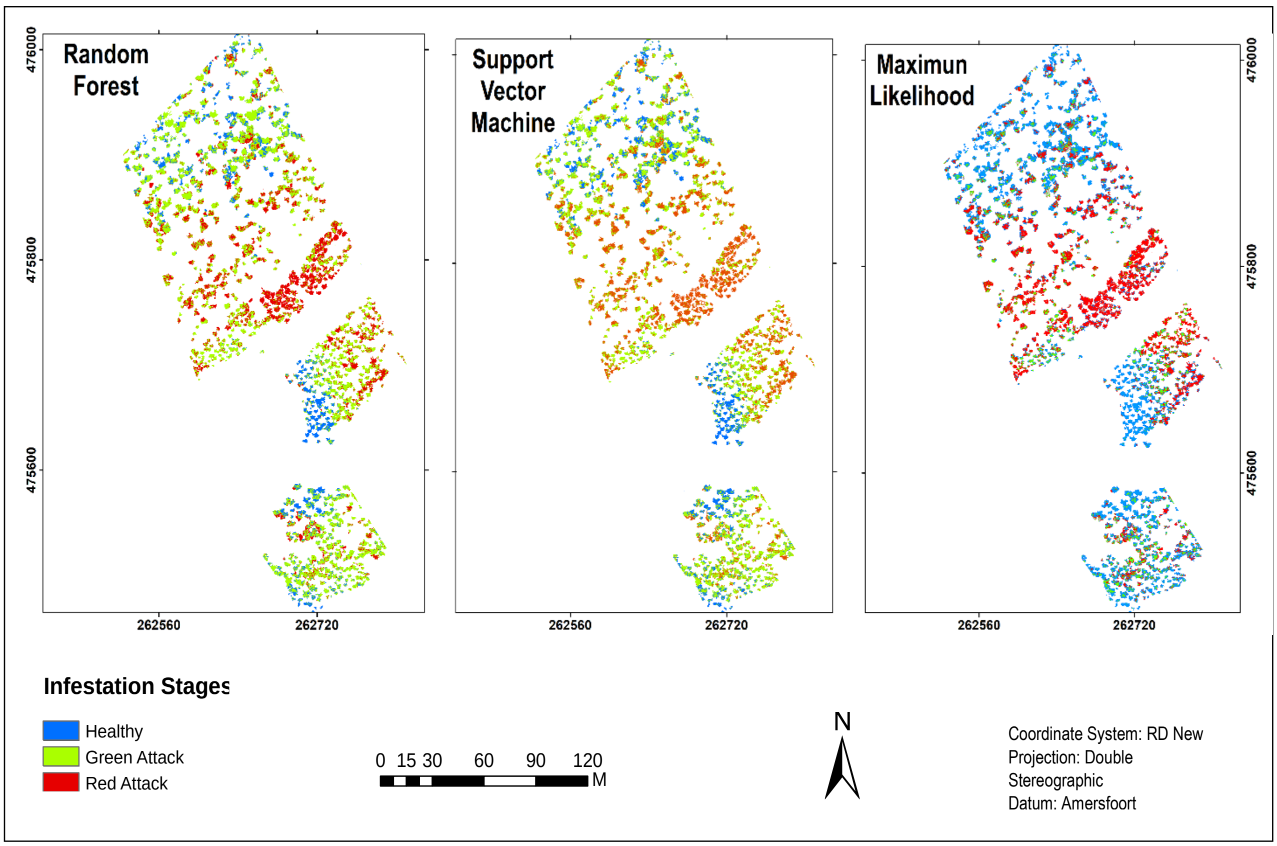


Figure 27: RF, SVM and ML classified maps for block three

Block three had some healthy European spruce in certain areas. The classification for RF and SVM did not perform well as it showed that almost all the trees were infested. ML performed well as the classification output is in line with the sampled field data. This can be confirmed by the image classification accuracy assessment results, as shown in table 12. The bark beetle infestations were presented as follows; blue (healthy), green (green attack), and red (red attack), respectively.

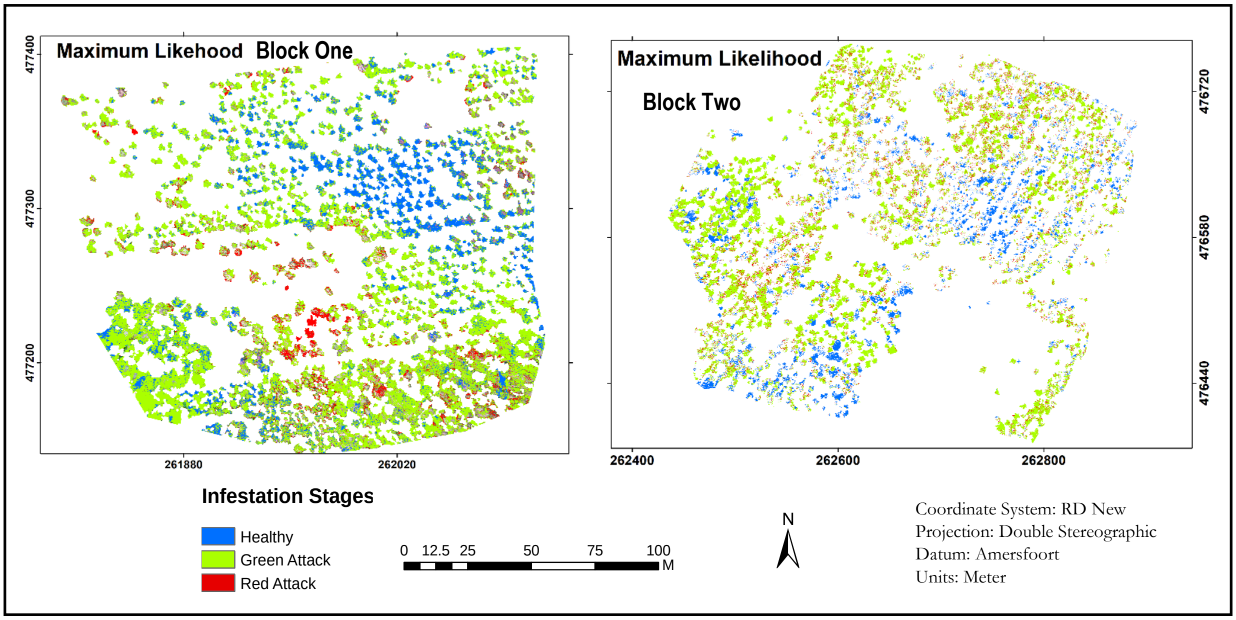


Figure 28: Maximum likelihood classified maps for block one and two

Table 12: Summary of a confusion matrix for RF, SVM, and ML image classification accuracy assessment for block one, two, and and three

Block No	Classifiers	Healthy		Green attack		Red attack		Overall accuracy measure	
		User Accuracy	Producers Accuracy	User Accuracy	Producers Accuracy	User Accuracy	Producers Accuracy	Overall Accuracy	Kappa
One	ML	94%	92%	88%	83%	91%	87%	89%	0.83
Two	ML	91%	84%	92%	90%	98%	66%	92%	0.87
Three	RF	87%	64%	73%	86%	85%	85%	78%	0.64
	SVM	87%	87%	50%	67%	84%	65%	74%	0.36
	ML	94%	76%	85%	94%	92%	65%	90%	0.86

The three approaches (RF, SVM, and MLC) were applied on block three to determine which classifier performed better in detecting the different stages of bark beetle infestation, especially the green attack. It was discovered that ML outperformed RF and SVM. As a result, the best-performing classifier was employed on the remaining two blocks (one and two). ML classification results were very much in line with the sampled field data for three classes.

3.2 Which vegetation indices (GNDVI, NDRE, NDVI) from UAV multispectral images have the highest probability in predicting the green attack of bark beetle infestation stage on European spruce using UAV multispectral images?

3.2.1 Predicted bark beetle infestation stages based on GNDVI, NDRE, and NDVI for block two

Figure 29 shows the predicted maps for the occurrence of bark beetle infestation stages in block two on the European spruce from three vegetation indices (GNDVI, NDRE, and NDVI) based on the input parameters shown in figure 25. According to the legend of the generated maps, the blue color represents (healthy), green color (green attack), and red color (red attack).

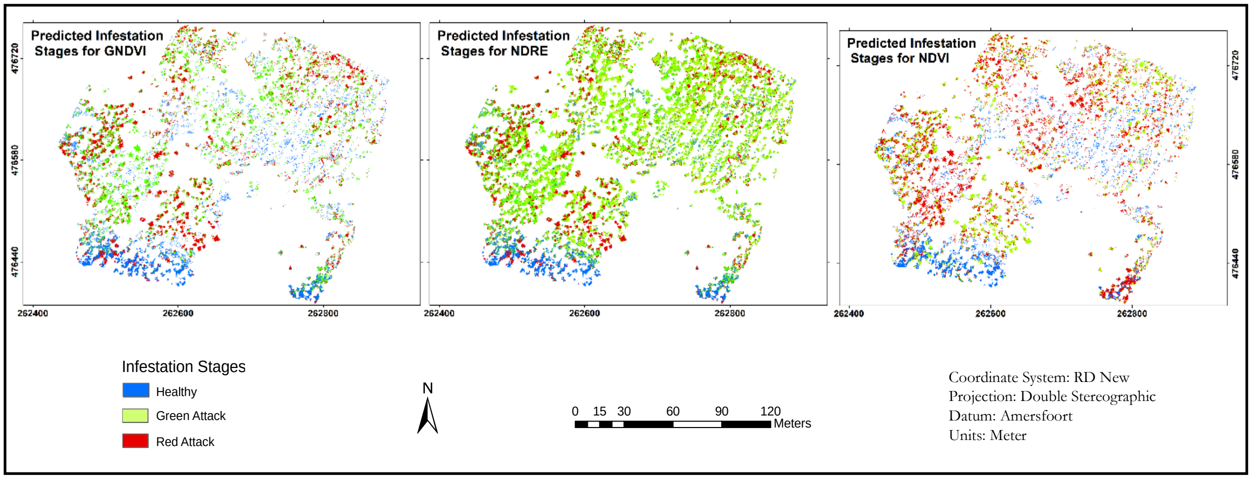


Figure 29: Predicted NDVI, NDRE & NDVI maps for block two based on sampled bark beetle infestation stages

The AUC was used to measure the accuracy of the three developed models (GNDVI, NDRE, and NDVI) using SDM. Figure 30 shows the probability of the SDM predictive power of the bark beetle infestation stages using AUC. Table 13 summarizes the three vegetation indices' performance in predicting the bark beetle infestation stages, especially the green attack.

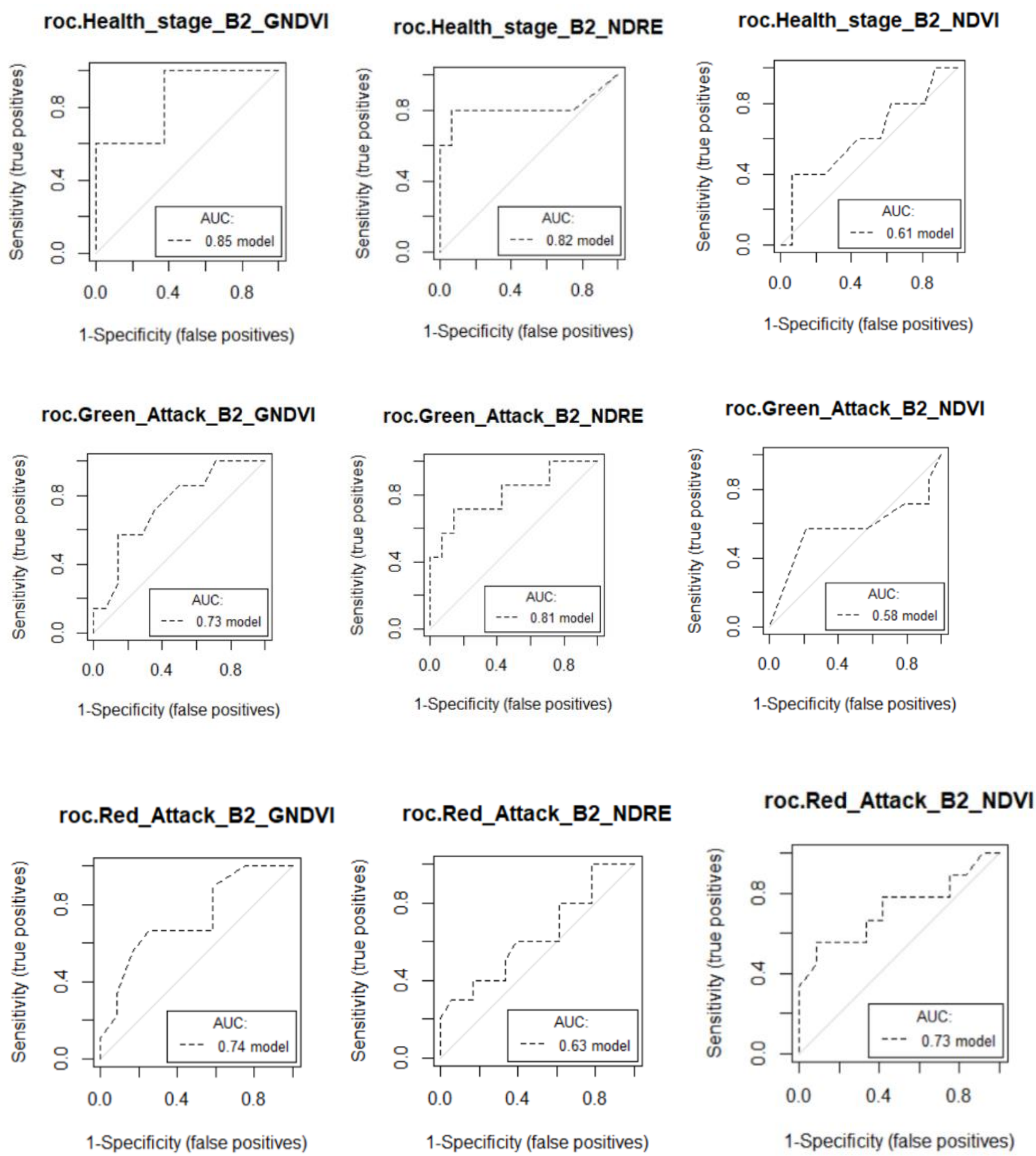


Figure 30: (AUC) Accuracy measure for the prediction of different bark beetle infestation on GNDVI, NDRE & NDVI

Based on the results generated by the SDM using AUC, NDRE had higher accuracy in predicting the green attack bark beetle infestation than GNDVI and NDVI, as shown in table 13.

Table 13: AUC Predicted results for GNDVI, NDRE & NDVI based on bark beetle infestation stages

AUC for Infestation Stages for Block Two			
	GNDVI	**NDRE**	**NDVI**
Healthy	0.85	0.82	0.61
Green attack	0.73	0.81	0.58
Red Attack	0.74	0.63	0.73

3.2.2 Testing the significance of the developed models when compared to field data

The model's statistical significance was based on the Akaike Information Criterion (AIC) and the Pr(>|z|). The AIC and the Pr(>|z|) were used to assess the models' quality and help out to compare the models. Table 14 shows the summary of the statistical significance of GNDVI, NDRE, and NDVI. It was discovered that NDRE was statistically significant in the prediction of the green attack.

Table 14: Summary of the model's statistical significance based on GNDVI, NDRE, and NDVI

Vegetation Indices	**Estimate Std. Error z value Pr(>|z|)**	**AIC**
GNDVI	0.604	54.34
NDRE	0.018 *	47.62
NDVI	0.544	53.77
	Significance. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1	

3.3 Which of the classifier (RF, SVM, ML) and vegetation indices (GNDVI, NDRE, and NDVI) has the highest accuracy in detecting and predicting different bark beetle infestations stages, especially the green attack using UAV multispectral images?

Objective three was to determine the classifier (RF, SVM, ML) with the highest accuracy in detecting the different bark beetle infestation stages using UAV multispectral images. Thus, among the three classifiers, ML outperformed RF and SVM. ML had 80% ability to detect the green attack stage with an overall accuracy of 90% and Kappa statistics 0.86. In contrast, the ability of RF in detecting the green attack was at 73%, with an overall accuracy of 78% and kappa statistics of 0.64. In addition, the SVM had 50% ability to detect the green attack with an overall accuracy of 74% and Kappa statistics of 0.36, as shown in table 12. In addition, the second part of the objective was about determining the vegetation indices (GNDVI, NDRE, NDVI) with the highest accuracy in predicting the different bark beetle infestation stages using UAV multispectral images. NDRE outperformed GNDVI and NDVI. The probability of the predictive power of the green attack using NDRE was at 0.81, GNDVI had 0.73, while NDVI had 0.58, respectively.

3.4 Model validation based field revisit

Four months after data collection, the classification and prediction results of the different bark beetle infestation stages for the three sampled blocks were ready. Visual model validation was conducted by revisiting blocks one, two, and three and comparing the generated results with what was prevailing on the ground. It was discovered that most European spruce were either green or red attack with very few healthy ones. The case of blocks two and three had more clear cuts. This can be evidenced by the pictures shown in figure 31. Based on the field revisit findings, it can be assumed that the clear cuts were probably a result of bark beetle infestation because most of the European spruce classified as infested were removed.

Figure 31: Clear cuts of infested European spruce from block two and three

Source: (Field data, 2020)

4 DISCUSSION

4.1 Comparison of the three classifiers (RF, SVM, and ML) and vegetation indices (GNDVI, NDRE, NDVI)

The study evaluated UAV multispectral images potential in detecting bark beetle infestation on European spruce, especially the green attack. The three classifiers (RF, SVM, ML) and vegetation indices (GNDVI, NDRE, NDVI) were used to detect bark beetle infestation. Consequently, the performance of the classifiers was examined separately from the vegetation indices in this study. The reason was that the two methodologies have different approaches and so cannot be compared. The classifiers used the spectral signatures of different bark beetle infestations captured in the field. The spectral signatures from the manually digitized European spruce tree crowns from different bark beetle infestations were used as training samples to classify the UAV multispectral color composites. On the contrary, the SDM used the standard deviation spectral values extracted from the three vegetation indices based on the manually digitized European spruce tree crowns for different bark beetle infestations. The process was done by summarizing the raster values for vegetation indices within the manually digitized European spruce tree crowns and used as predictors in the SDM. Additionally, the SDM used vegetation indices as input layers; the model requires some raster layers as environmental variables.

In addition, the use of UAV multispectral images with the spatial resolution of 10.49 cm x 10.49 cm and flight height of between 100 and 110 meters eliminated the effects that arise as the distance between sensors and foliage increases; this is in line with Seifert et al. (2019). The weather condition also favored the drone flight because the sky was fully cloudy with no wind and rain; this is in line with (Dandois et al., 2015). Also, because of the set parameters of the UAV multispectral images used in this study, it was easy to identify individual trees and even leaves due to their high spatial resolution, as this agrees with Klouček et al. (2019). Therefore, the two techniques (SDM and ML), which require few training samples and UAV multispectral images, show that there is potential in detecting the bark beetle infestation stages on time, especially the green attack to mitigate the spread. The two techniques have the potential of reduced data capturing period for model development hence contributing to the quick solution of the mitigation process.

A purposive sampling approach was used to collect data for this study because it required the researchers' judgment in selecting representative sample sites and training samples. The expert knowledge was applied in choosing the right infestation stages to ensure that the sample collected was correct and accurate, which agrees with (Hewson, 2007; Ramezan, Warner, & Maxwell, 2019).

Also, the use of OBIA in this study facilitated the enhancement of the classifiers and predictors accuracy because all the European spruce tree crowns were splitted into non-overlapping categories. The images were analyzed as object spaced and not pixel spaced to avoid overlap. The UAV multispectral images in this study were delineated and evaluated based on the spectral relationship of the objects. The spectral relationship was further used for spectral feature extraction, such as mean and standard deviation brightness values for each band, red, green, red edge, and near-infrared, as shown in table 7. The extracted features were used to distinguish different bark beetle infestation stages from the image during the classification of RF and SVM, as the two classifiers require the extraction of objects' features.

According to the results generated in table 12, section 3.1, the RF classifier achieved slightly higher accuracy than SVM but lower than ML. The lower accuracy of RF resulted from the small training data sample size, which had a negative effect on the classifier. According to Colditz (2015), the classifier requires a training sample size (at least 50 samples per class) to create classification tree algorithms. This study had 93 samples for block three; the healthy stage had 39, the green attack 32, and the red attack had 21, respectively. The limited number of samples resulted from the small sample sites comprised of European spruce in the coniferous forest, which hindered the collection of as many samples as possible for the study. Further, it was a challenging exercise to distinguish between the green attack from the healthy European spruce. Thus, the sampling of infestation stages was restricted to human vision. Kim (2009) found out that the performance of the RF model improves as the sample size increases. For example, when 200 samples were used to predict the model, the overall accuracy was 91%; furthermore, when the samples were increased to 300, the accuracy also improved by 95%. Thus the findings of Kim (2009) confirms why the RF classifier in this study had a lower accuracy when compared to ML. In addition, Hollaus & Vreugdenhil (2019b) found out that RF gives higher accuracy (overall accuracy of 94% with a kappa of 0.51 and 87% with a kappa of 0.23) when using RapidEye and TerraSAR-X imagery in their research of detecting bark beetle outbreaks in forests. This high accuracy can be aligned to the higher spatial resolution of both imageries used as it can efficiently resolve a branch, stem, and leaf structure and sufficient sample size, as alluded by (Kellner et al., 2019).

Additionally, the SVM had the lowest accuracy among the three classifiers, as shown in section 3.1, table 12. The lowest accuracy results can be attributed to the small training data sample size. The sample size used in SVM was the same as the one used in RF since it was the same block under study. Belgiu & Drăgu (2016) and Fassnacht et al. (2014) are of the view that SVM cannot promise higher accuracy with a small training sample size as it normally results in misclassification in most cases. The challenge with SVM is that the training samples of the classes are not always linearly separable from each other as it requires the use of hyperplanes of support vectors (refer to figure 24A under section 2.11.2). SVM uses outlier variables to limit the violation of the constraints imposed by the two supporting hyperplanes. The outlier variable measures how far the outlying sample is from the supporting hyperplane that holds the class's support vectors. The technique makes it difficult for SVM to perform well though it can still perform better with fewer training sample sizes than RF. For instance, Abdel-Rahman et al. (2014) used AISA Eagle hyperspectral data to evaluate the performance of RF and SVM classifiers in differentiating between healthy, sirex grey-attacked, and lightning-damaged pine trees. The researchers evaluated the performance of RF and SVM classifiers in their study. Their findings revealed that RF had a slightly higher accuracy of 74.5% than the SVM, which only had a 73.5 %. According to their findings, the two non-parametric classifiers performed similarly in classifying the eight classes, although the RF had a higher overall accuracy. Their findings are similar to those of this study, which found that RF achieved slightly higher overall accuracy than SVM with very narrow margins in overall accuracy. In their study, Stych et al. (2019) used SVM to compare the image classification of forests affected by the bark beetle outbreak using WorldView-2 and Landsat 8. The overall accuracies of 86% and 71%, and Kappa of 0.84 and 0.66, were attained for WorldView-2 and Landsat8. Their findings are in line with this study. The results of this study based on UAV multispectral image lie in between WorldView2 and Landsat8, which seem to be promising. The findings reveal that the higher the spatial resolution, the higher the overall accuracy and the kappa. In addition, the lower the spatial resolution, the lower the overall accuracy and kappa.

Furthermore, the ML classifier outperformed the two classifiers (RF and SVM), as shown in table 12, section 3.1. The sample size used under the ML classifier was the same as the one used in RF and SVM since it was still the same block under study. Thus, ML had the highest accuracy among the three classifiers because it is the least affected technique that can still produce relatively higher accuracy using a smaller training sample size (Li, Wang, Wang, Hu, & Gong, 2014; Sheykhmousa et al., 2020). In addition, Klouček et al. (2019a) found out that the ML classifier gave a higher overall accuracy in their study of UAV-mounted sensors for the precise detection of bark beetle infestation using the low-cost RGB. The overall accuracy was ranging from 78% to 94% for the results generated. The sample size in the study Klouček et al. was similar to this study as it had (forty) 40 healthy trees and ten (10) infested trees. Based on Klouček et al. findings, it truly confirms that the ML classifier still performs better even when the sample size is small. Further, ML can still perform better even on a low-cost RGB UAV sensor.

In addition, in a related study (Yang, 2019a) found out that ML classification gives higher accuracy in detecting bark beetle damage with Sentinel-2 multi-temporal data at 10-meter resolution. According to his findings, ML had an overall accuracy of 89%, with Kappa of 0.74, and RF had an overall accuracy of 85% and kappa of 0.62. Based on the findings from this study, it can be concluded that ML performs better regardless of spatial image resolution. Therefore, the findings of Yang (2019a) are in line with the present study that showed the overall accuracy of ML to be higher than RF, as shown in table 12. Furthermore, Meddens et al. (2013) indicate that the maximum likelihood classifier has been utilized successfully in previous studies to assess techniques for detecting bark beetle-caused tree mortality using Landsat images. The achieved overall accuracy was 91% and kappa of 0.88. The findings of Meddens et al. confirms that ML can still perform better on medium spatial resolution images. According to the literature review under this section, maximum likelihood performs better with a small sample size based on medium or high spatial resolution.

4.2 Comparison of the three vegetation indices (GNDVI, NDRE, and NDVI) based on the area under the curve

The performance of the three vegetation indices (GNDVI, NDRE, and NDVI) in predicting the bark beetle infestation on European spruce was evaluated based on the AUC concept in SDM. According to Kienast, Bolliger, & Zimmermann (2012), the SDM is an example of a generalized linear model (GLM) representing the extension of other linear regression techniques. The advantage of the SDM can be predicted and fitted even with the small sample size (Li et al. 2014; Sheykhmousa et al., 2020). Proosdij et al. (2016) state that the SDM can predict the occurrence with the sample size ranging from three to thirteen reference points because it only deals with presence data. Duračiová et al. (2020) used a similar model to predict new bark beetle infestation, and the predictive power between (AUC 0.56 and 0.64) was attained in their study of bark beetle infestation predictive model based on satellite data. Thus the results generated from this study are promising as they show that the UAV multispectral images have potential in predicting the bark beetle infestation stages, especially the green attack as captured in table 13 of section 3.2.2. In most literature reviewed, color composite raster layers have been used, and vegetation indices have not been exploited. Thus the use of vegetation indices seems to enhance the results in the detection of bark beetle infestation. Additionally, it is worth noting that very little research has been conducted using the SDM to detect bark beetle infestation. Therefore, the present study might contribute to the few existing knowledge in the use of SDM to detect bark beetle infestation, especially the green attack.

The performance of GNDVI in predicting the bark beetle infestation stages (healthy, green, and red attacks) in this study was based on the AUC using the SDM as indicated in table 13 of section 3.2.1. GNDVI had higher results in healthy trees as compared to green and red attacks. Healthy European spruce is expected to have high chlorophyll content than green and red attacks. This is because GNDVI is the photosynthetic indicator in vegetation and is particularly sensitive to chlorophyll concentration (Abdullah et al., 2019b). In the study of detection of flavescence doree grapevine disease using UAV multispectral, Albetis et al. (2017) indicate that GNDVI had a high predictive power ranging from 0.75 to 0.95 for the sampled vineyards in estimating plant productivity and leaf pigments such as chlorophyll. The findings of this study are in line with Albetis et al. (2017). In addition, Théau et al. (2020) discovered that GNDVI had the highest results of 100 % in the detection of crop stress using UAV imagery. Generally, it can be concluded that GNDVI performs better in the detection of healthy plants. Thus, the findings of this study are in line with the literature reviewed in this section, as the predictive power was at 0.85.

Also, the performance of NDRE in predicting the bark beetle infestation stages (healthy, green, and red attacks) in this study was also based on the AUC using the SDM. NDRE had high predictive power in both healthy and green attack with lower results in a red attack as captured in table 13 under section 3.2.1. The predictive power was high in healthy and green attack because NDRE is a good indicator of healthy vegetation that has accumulated high chlorophyll levels in the leaves. Also, the red edge band light in NDRE is very sensitive to abrupt small chlorophyll content changes caused by bark beetle infestation (Oliech, 2019). Additionally, Huang et al. (2019) are of the view that NDRE is a good measure of early forest stress detection. NDRE has been used before in related studies; it has proved to perform well in predicting bark beetle infestation using different techniques apart from the SDM technique. July (2020a) found out that NDRE gives higher results in detecting healthy and green attacks ranging from 0.4 to 0.7. In contrast, the results for the infected European spruce are lower, ranging from 0 to 0.6. The results were generated using the temporal profiles using sentinel-2 on the study of differentiating healthy and bark beetle-infected spruce trees. Thus, he findings of this study were a bit higher in reference to July (2020a). According to Abdullah et al. (2018), NDRE has been proven to give "accurate maps of green attack stage of European spruce using sentinel-2." NDRE has also been used to detect forest stress induced by drought and insect infestation in the early stage. The accuracy was ranging from 0.4 and <0.8 for the green attack. Similarly, the predictive power for NDRE in this study was at 0.81; the results are within the range of Abdullah et al.'s findings.

Furthermore, Ortiz et al. (2013a) are of the view that models with AUC values above 0.75 can be regarded as useful for predictions because they are considered to be very accurate as they are closer to 1. The NDRE model results for this study can be considered useful in the prediction of the bark beetle green attack since the results were above 0.75. It has also been observed that NDRE has the capacity to produce high accuracy in the prediction of healthy and green attack of both medium and high spatial resolution images used (sentinel-2, UAV multispectral, RapidEye). The results for NDRE are consistent with the results from the previous studies of (Ortiz et al. 2013a; Eitel et al. 2011; Abdullah et al. 2018; July 2020a; Huang et al. 2019; Oliech, 2019) based on the literature review under this section.

The performance of NDVI in predicting the bark beetle infestation stages (healthy, green, and red attacks) in this study was also based on the AUC using the SDM. NDVI had high predictive power in a red attack with lower results in a green attack and healthy stage, as shown under section 3.2.1 table 13. The higher results of NDVI in red attack can be related to the loss of chlorophyll, a pigment that absorbs red light as the leaves turn yellow-brown due to bark beetle infestation in accordance with Heenkenda et al. (2015). Consequently, NDVI is known to measure greenness in leaves (Yoder & Waring, 1994). According to Abdullah et al. (2019b). NDVI is the photosynthetic indicator in vegetation and is very sensitive to chlorophyll concentration. NDVI is mainly used to determine plant stress; hence higher results under red attack because of low chlorophyll content. This is similar to the findings of Yang (2019a), where NDVI was used to detect the infected trees, and overall accuracy of 83% was attained using sentinel-2 images with a spatial resolution of 10 meters. For example, Morgan (2017) found out that NDVI had an overall accuracy of 73% in detecting the red attack stage in his study of analyzing red and grey stages of bark beetle attack using Landsat. Walter & Platt (2013) also discovered that NDVI gives higher accuracy in detecting the red attack ranging between 85 and 91%, according to their study on multi-temporal analysis in the prediction of mountain pine beetle infestation. Based on this study, the overall accuracy of the red attack under NDVI was at 73%%, which is in line with the literature review for this section.

The significance of the three developed models is illustrated in table 14 under section 3.2.2, and the SDM significance level was set at 0.05. Thus, the AIC and the estimated standard error z value $Pr(>|z|)$ were the two measures of significance levels. De Rivera et al. (2017; Gutierrez & Heming (2018) are of the view that the lower the $Pr(>|z|)$ and AIC value and higher the AUC, the better. The statement given by the preceding authors are in agreement with this study's findings. According to Mezei et al. (2019), the AIC has been used before to determine the extent of the European spruce bark beetle's unusual tree death. The AIC is the most commonly used model selection method in statistics. It is used to determine the best fit for the data by calculating and comparing various AIC scores. The link between infestation spots and predictor variables was modeled. The AIC was used to rank the models; the lowest AIC value was chosen as the best fit. Further, Seidl et al. (2016) also state that AIC was utilized to estimate the likelihood of bark beetle infestation. The study evaluated which model formulations were most strongly supported by the data using the AIC. The model with the smallest number of AIC values was chosen as the best fit. In addition, Rumsey (2021) indicates that the z value is the standard score in statistics that shows the number of standard deviations by which the value of a raw score is higher than or lower than the mean value of what is being observed or measured. A specific normal distribution with a mean of 0 and a standard deviation of 1 is the standard normal distribution.

For this study, kappa coefficient, overall, user and producer's accuracy, and AUC were used for image classification accuracy assessment and evaluation of the predictive performance of the SDM. According to Ortiz et al. (2013b), Kappa Coefficient and area under the curve (AUC) have been used before in determining the classification accuracy of bark beetle infestation stages. Kappa Coefficient has been described as the best model because of its ability to avoid overfitting. The kappa 0.81 and above is considered to be very good. In this regard, according to the maximum likelihood accuracy assessment results for this study, the kappa statistics were between 0.83 and 0.87. At the same time, RF and SVM had 0.64 and 0.36, respectively. Thus the findings of the present study ML classifier produced a higher accuracy, and the results are in agreement with Ortiz et al. (2013b). In addition, Ortiz et al. (2013b) further said that models with AUC values above 0.75 are considered useful for the predictions as perfect models have AUC values close to 1.0. According to the results generated under this study using NDRE, the probability of detecting the green attack was at 0.81, GNDVI had 0.73, NDVI had 0.58. Above all, that NDRE performed better than the other two indices.

Therefore, based on the findings of this study, I would recommend the use of UAV multispectral images in detecting the bark beetle infestation because they are cheap, flexible, limited expertise required to operate. Also, when the bark beetles start manifesting, the weather condition is favorable for drone flight because the sky is fully cloudy or clear with no wind and rain. On the contrary, most satellites are usually commercial, and image collection requires placing an order ahead of time, leading to the great destruction of the forests by bark beetle outbreak. Satellites like RapidEye, Worldview, and HyMap have spatial resolutions ranging from 0.5-2m. They may only be stable when most infected trees are in small, discrete patches, and this is in line with Abdullah et al. (2019a) findings.

4.3 Limitation in the research

The constraint in this study was the limited number of training samples, especially the green attack. It was a challenging exercise to distinguish the green attack from the healthy European spruce. Thus, only the green attack samples was sure of were collected. This hindered the collection of as many samples as possible for the study.

5 CONCLUSION AND RECOMMENDATION

5.1 Conclusion

The findings of this study showed that a UAV multispectral image could be utilized to detect the bark beetle infestations stages on European spruce, especially the green attack, with the possibility of giving higher results. It was possible to classify, predict, and identify specific trees based on bark beetle infestation. The spatial resolution of the multispectral images acquired for this study was 10.49 cm x 10.49 cm high enough to improve branch, stem, and leaf structure. Using three classifiers (RF, SVM, and ML) and vegetation indices (GNDVI, NDRE, and NDV), the data captured by the UAV sequoia camera and the field data was used to classify the stages of bark beetle infection on European spruce. In addition, the three generated vegetation indices were used as input variables in the SDM to predict their performance in detecting the green attack presence on European spruce. The phases of the proposed approach in this study were as follows: fieldwork planning, UAV multispectral image acquisition, and field data collection, image processing and pre-processing, delineation of different bark beetle infestation tree crowns, object-based image segmentation, feature extraction, and image segmentation accuracy assessment, vegetation indices calculation and extraction of vegetation indices mean and standard deviation values, image classification and model development and accuracy evaluation of classified maps and model validation. The results of this study can be extrapolated to other areas because they are promising as they show that the UAV multispectral images have greater potential in detecting the bark beetle infestation stages, especially the green attack. The research had three objectives and three research questions; the specific conclusion is presented below:

Question 1: What classifier among RF, SVM, and ML gives the highest results in detecting the green attack of the bark beetle infestation stage on European spruce from UAV multispectral images?

The overall classification accuracy achieved for the three classifiers (RF, SVM, and ML) in detecting the bark beetle infestation on European spruce, especially the green attack for the block was 78%, 74%, and 90%; with the ability to detect the green attack stage at 73%, 50%, and 85% respectively.

Question 2: Which vegetation indices (GNDVI, NDRE, NDVI) from UAV multispectral images have the highest probability of detecting the green attack of bark beetle infestation stage on European spruce?

NDRE outperformed GNDVI and NVDI with the predictive power of 0.81, 0.73, and 0.58, respectively using UAV multispectral images. The results were based on AUC's probability of detecting the green attack stage on European spruce when the three vegetation indices were used as input variables in the SDM.

Question 3: Which classifier (RF, SVM, ML) and vegetation indices (GNDVI, NDRE, and NDVI) from UAV multispectral images gives the highest results in detecting the different bark beetle infestations stage using, especially the green attack?

Among the three classifiers (RF, SVM, and ML), maximum likelihood obtained the highest overall classification accuracy of 90% in detecting the bark beetle infestation stages on European spruce. The highest accuracy was attributed to the ability of MLC performing well even when the training sample size was small compared to RF, which had 78% and 74% for SVM.

Further, among the three vegetation indices, NDRE had the highest probability of 0.81 of detecting the green attack stage on European spruce because the red edge band in NDRE is very sensitive to abrupt small chlorophyll content changes as opposed to GNDVI and NDVI, which had 0.73 and 0.58 probability respectively. In addition, the red-edge band had higher spectral separability compared to red, green, and near-infrared, as shown in figure 23, making it possible for the high probability of green attack detection on a UAV multispectral image.

5.2 Recommendation

Thus, based on the results generated in table 12, section 3.1, and table 13 of section 3.2.2. I would recommend that the forest managers use the SDM and ML classifier to detect bark beetle infestation on European spruce, especially the green attack. The two techniques can work the small number of training samples and still give higher accuracies as compared to RF and SVM. Proosdij et al. (2016) state that the SDM can predict the occurrence with the sample size ranging from three to thirteen reference points because it only deals with the presence data only. Further, (Li, Wang, Wang, Hu, & Gong, 2014; Sheykhmousa et al., 2020) found out that ML classifier is the least affected technique that can still produce relatively higher accuracy using a smaller training sample size. Since, when using the SDM and ML techniques, only a few samples are required to classify and predict the bark beetle infestation stages. The two techniques can help mitigate the bark beetle spread as little time is required to collect the samples for model development.

The results from the research used few training samples for training and accuracy assessment, as shown in table 8. As a result, only ML and the SDM could produce higher accuracy than RF and SVM, which require more training samples. Therefore, future research related to this kind of study should consider increasing the number of collected training samples. This would possibly result in higher classification accuracies for RF and SVM. In addition, the samples for different bark beetle infestation stages on European spruce should never be collected in winter (rainy season) as it is practically impossible to distinguish the infestation stages. Furthermore, ML and SDM based on the NDRE can be used in similar studies as they both can still produce a higher accuracy even with a smaller training sample size. Future research in a related study should consider using the vegetation indices as input layers for the classifiers.This would possibly increase the classification accuracy.

LIST OF REFERENCES

Abdel-Rahman, E. M., Mutanga, O., Adam, E., & Ismail, R. (2014). Detecting Sirex noctilio grey-attacked and lightning-struck pine trees using airborne hyperspectral data, random forest and support vector machines classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, *88*, 48–59. https://doi.org/10.1016/j.isprsjprs.2013.11.013

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*(July 2017), 199–209. https://doi.org/10.1016/j.jag.2017.09.009

Abdullah, H., Skidmore, A. K., Darvishzadeh, R., & Heurich, M. (2018). 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. https://doi.org/10.1002/rse2.93

Abdullah, H., Skidmore, A. K., Darvishzadeh, R., & Heurich, M. (2019). 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*. https://doi.org/10.1016/j.jag.2019.101900

Ágnes, T., Eszter, T., & Ágnes, G. (2010). *Urban Tree Crown Projection Area Mapping With Object Based Image Analysis for Urban Ecosystem Service Indicator Development*. 415–422.

Akayezu, P., van Duren, I. C., Groen, T. A., Grueter, C. C., & Robbins, M. M. (2019). Abundance and spatial distribution of the main food species for mountain gorillas in the Virunga Massif, Rwanda. *Biodiversity and Conservation*, *28*(13), 3597–3620. https://doi.org/10.1007/s10531-019-01838-0

Al-doski, J., Mansor, S. B., Zulhaidi, H., & Shafri, M. (2013). *Image Classification in Remote Sensing*. *3*(10), 141–148. Retrieved from www.iiste.org

Albetis, J., Duthoit, S., Guttler, F., Jacquin, A., Goulard, M., Poilvé, H., … Dedieu, G. (2017). Detection of Flavescence dorée grapevine disease using Unmanned Aerial Vehicle (UAV) multispectral imagery. *Remote Sensing*, *9*(4). https://doi.org/10.3390/rs9040308

Alfons, A., Templ, M., & Filzmoser, P. (2010). An object-oriented framework for statistical simulation: The R package simframe. *Journal of Statistical Software*, *37*(3), 1–36. https://doi.org/10.18637/jss.v037.i03

Barbati, A., Marchetti, M., Chirici, G., & Corona, P. (2014). European Forest Types and Forest Europe SFM indicators: Tools for monitoring progress on forest biodiversity conservation. *Forest Ecology and Management*, *321*, 145–157. https://doi.org/10.1016/j.foreco.2013.07.004

Belgiu, M., & Drăgu, L. (2016, April 1). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 114, pp. 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Belgiu, M., Drăguţ, L., & Strobl, J. (2014). Quantitative evaluation of variations in rule-based classifications of land cover in urban neighbourhoods using WorldView-2 imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *87*(October), 205–215. https://doi.org/10.1016/j.isprsjprs.2013.11.007

Beucher, S. (2000). (PDF) The Watershed Transformation Applied To Image Segmentation. Retrieved February 23, 2021, from https://www.researchgate.net/publication/2407235_The_Watershed_Transformation_Applied_To_Image_Segmentation

Blanchard, Y., Pelon, J., Eloranta, E. W., Moran, K. P., Delanoë, J., & Sèze, G. (2014). A Synergistic analysis of cloud cover and vertical distribution from A-Train and ground-based sensors over the high Arctic station Eureka from 2006 to 2010. *Journal of Applied Meteorology and Climatology*, *53*(11), 2553–2570. https://doi.org/10.1175/JAMC-D-14-0021.1

Blaschke, T., Lang, S., & Hay, G. J. (2008). Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications. In G. H. Thomas Blaschke, Stefan Lang (Ed.), *LibTuDelftNet*. Retrieved from https://books.google.nl/books?hl=en&lr=&id=5ktbsV-aSBAC&oi=fnd&pg=PR15&dq=What+is+object+image+analysis&ots=bs8q2ozKl6&sig=9Tye9txhUIWQMOSR_oVuzb8NqrI&redir_esc=y#v=onepage&q=What is object image analysis&f=false

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bright, B. C., Hudak, A. T., McGaughey, R., Andersen, H. E., & Negrón, J. (2013). Predicting live and dead tree basal area of bark beetle affected forests from discrete-return lidar. *Canadian Journal of Remote Sensing*, *39*(SUPPL.1). https://doi.org/10.5589/m13-027

Caudullo, G., Commission, E., & Rigo, D. De. (2017). Picea abies in Europe: distribution, habitat, usage and threats. *Picea Abies in Europe: Distribution, Habitat, Usage and Threats*, (12).

Chazdon, R. L., Brancalion, P. H. S., Laestadius, L., Bennett-Curry, A., Buckingham, K., Kumar, C., … Wilson, S. J. (2016). When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration. *Ambio*, *45*(5), 538–550. https://doi.org/10.1007/s13280-016-0772-y

Chen, G., & Meentemeyer, R. K. (2016). *Remote Sensing of Forest Damage by Diseases and Insects*. (November), 145–162. https://doi.org/10.1201/9781315371931-9

Chen, Y., Chen, Q., & Jing, C. (2019). Multi-resolution segmentation parameters optimization and evaluation for VHR remote sensing image based on meanNSQI and discrepancy measure. *Journal of Spatial Science*, *00*(00), 1–26. https://doi.org/10.1080/14498596.2019.1615011

Christopoulou, O. G. (2011). Deforestation / Reforestation in Mediterranean Europe: The Case of Greece. *Soil Erosion Studies*, (Pearce). https://doi.org/10.5772/23466

Clinton, N., Holt, A., Scarborough, J., Yan, L. I., & Gong, P. (2010). Accuracy assessment measures for object-based image segmentation goodness. *Photogrammetric Engineering and Remote Sensing*, *76*(3), 289–299. https://doi.org/10.14358/PERS.76.3.289

Coggins, S. B., Coops, N. C., & Wulder, M. A. (2011). Estimates of bark beetle infestation expansion factors with adaptive cluster sampling. *International Journal of Pest Management*, *57*(1), 11–21. https://doi.org/10.1080/09670874.2010.505667

Colditz, R. R. (2015). An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms. *Remote Sensing*, *7*(8), 9655–

9681. https://doi.org/10.3390/rs70809655

Çöltekin, A., Griffin, A. L., Slingsby, A., Robinson, A. C., Christophe, S., Rautenbach, V., … Klippel, A. (2020). Geospatial Information Visualization and Extended Reality Displays. In *Manual of Digital Earth*. https://doi.org/10.1007/978-981-32-9915-3_7

Criminisi, A., & Shotton, J. (2013). Decision Forests for Computer Vision and Medical Image Analysis. In *Advances in Computer Vision and Pattern Recognition*. Retrieved from www.springer.com/series/4205

Dandois, J. P., Olano, M., & Ellis, E. C. (2015). Optimal altitude, overlap, and weather conditions for computer vision uav estimates of forest structure. *Remote Sensing*, *7*(10), 13895–13920. https://doi.org/10.3390/rs71013895

De Rivera, Ó. R., & López-Quílez, A. (2017). Development and comparison of species distribution models for forest inventories. *ISPRS International Journal of Geo-Information*, *6*(6). https://doi.org/10.3390/ijgi6060176

Deng, L., Mao, Z., Li, X., Hu, Z., Duan, F., & Yan, Y. (2018). UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras. *ISPRS Journal of Photogrammetry and Remote Sensing*, *146*, 124–136. https://doi.org/10.1016/j.isprsjprs.2018.09.008

Duračiová, R., Muňko, M., Barka, I., Koreň, M., Resnerová, K., Holuša, J., … Jakuš, R. (2020). A bark beetle infestation predictive model based on satellite data in the frame of decision support system tanabbo. *IForest*, *13*(3), 215–223. https://doi.org/10.3832/ifor3271-013

Eitel aM Lee A Vierling Marcy E Litvak DanS Long Urs Schulthess Alan A Ager Dan Krofcheck Leo Stoscheck, jan U. (2011). Remote Sensing of Environment. In *Remote Sensing of Environment* (Vol. 115). Retrieved from www.elsevier.com/locate/rse

Eitel, J. U. H., Vierling, L. A., Litvak, M. E., Long, D. S., Schulthess, U., Ager, A. A., … Stoscheck, L. (2011). Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland. *Remote Sensing of Environment*, *115*(12), 3640–3646. https://doi.org/10.1016/j.rse.2011.09.002

El-naggar, A. M. (2018, December 1). Determination of optimum segmentation parameter values for extracting building from remote sensing images. *Alexandria Engineering Journal*, Vol. 57, pp. 3089–3097. https://doi.org/10.1016/j.aej.2018.10.001

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Erdle, K., Mistele, B., & Schmidhalter, U. (2011, October 9). Comparison of active and passive spectral sensors in discriminating biomass parameters and nitrogen status in wheat cultivars. https://doi.org/10.1016/j.fcr.2011.06.007

Eshetae, M. A. (2020). *[July, 2020]*.

European Union. (2019). *THE EUROPEAN UNION AND FORESTS*.

Fahy, O. & C. J. (1999). COST Action E27 Protected Forest Areas in Europe – Analysis and Harmonisation

Country Report - Ireland. *Country Report Ireland*, (January), 173–186.

Fassnacht, F. E., Latifi, H., Ghosh, A., Joshi, P. K., & Koch, B. (2014). Assessing the potential of hyperspectral imagery to map bark beetle-induced tree mortality. *Remote Sensing of Environment*, *140*, 533–548. https://doi.org/10.1016/j.rse.2013.09.014

Fazal-E-Malik, & Baharudin, B. Bin. (2011). Mean and standard deviation features of color histogramusing laplacian filter for content-based image retrieval. *Journal of Theoretical and Applied Information Technology*, *34*(1), 1–7.

Felton, A., Lindbladh, M., Brunet, J., & Fritz, Ö. (2010, August 15). Replacing coniferous monocultures with mixed-species production stands: An assessment of the potential benefits for forest biodiversity in northern Europe. *Forest Ecology and Management*, Vol. 260, pp. 939–947. https://doi.org/10.1016/j.foreco.2010.06.011

Fettig, C. J., & Hilszczański, J. (2015). Management Strategies for Bark Beetles in Conifer Forests. *Bark Beetles: Biology and Ecology of Native and Invasive Species*, 555–584. https://doi.org/10.1016/B978-0-12-417156-5.00014-9

Food and Agriculture Organization of the United Nations(FAO). (2005a). Global Review of forest pests and disease. *Management*, *2*.

Food and Agriculture Organization of the United Nations(FAO). (2005b). Global Review of Forest pests and diseases. *Management*, *2*.

Foody, G. M. (2002, April 1). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, Vol. 80, pp. 185–201. https://doi.org/10.1016/S0034-4257(01)00295-4

Forestry Commission. (2018). *Eight-toothed spruce bark beetle (Ips typographus)-contingency plan*.

Gaden, K. (2020). *Assessing Potential of Uav*.

Gandhi, U. (2017). Digitizing Map Data — QGIS Tutorials and Tips. Retrieved May 21, 2021, from QGIS Tutorials and Tips website: https://www.qgistutorials.com/en/docs/digitizing_basics.html

Gartner, M. H., Veblen, T. T., Leyk, S., & Wessman, C. A. (2015). Detection of mountain pine beetle-killed ponderosa pine in a heterogeneous landscape using high-resolution aerial imagery. *International Journal of Remote Sensing*, *36*(21), 5353–5372. https://doi.org/10.1080/01431161.2015.1095369

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern Recognition Letters*, *27*(4), 294–300. https://doi.org/10.1016/j.patrec.2005.08.011

Gómez, D., & Montero, J. (2011). Determining the accuracy in image supervised classification problems. *Proceedings of the 7th Conference of the European Society for Fuzzy Logic and Technology, EUSFLAT 2011 and French Days on Fuzzy Logic and Applications, LFA 2011*, *1*(1), 342–349. https://doi.org/10.2991/eusflat.2011.103

Gove, R., & Faytong, J. (2012). Machine Learning and Event-Based Software Testing: Classifiers for Identifying Infeasible GUI Event Sequences. In *Advances in Computers* (Vol. 86, pp. 109–135). https://doi.org/10.1016/B978-0-12-396535-6.00004-1

Green, E.P., Clark, C.D., Edwards, A. J. (2000). Image classification and habitat mapping. *Remote Sensing Handbook for Tropical Coastal Management*, 141-154. Retrieved from https://www.researchgate.net/publication/261795926_Image_classification_and_habitat_mapping

Gutierrez, E. E., & Heming, N. M. (2018). *Introducing AIC model averaging in ecological niche modeling: a single-algorithm multi-model strategy to account for uncertainty in suitability predictions*. (2). Retrieved from http://arxiv.org/abs/1807.04346

Heenkenda, M. K., Joyce, K. E., Maier, S. W., & de Bruin, S. (2015). Quantifying mangrove chlorophyll from high spatial resolution imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *108*, 234–244. https://doi.org/10.1016/j.isprsjprs.2015.08.003

Heurich, M. (2009). Progress of forest regeneration after a large-scale Ips typographus outbreak in the subalpine Picea abies forests of the Bavarian Forest National Park. *Silva Gabreta*, *15*(1), 49–66.

Hewson, J. (2007). Field Methods in Remote Sensing. *Cartographic Perspectives*, (56), 52–54. https://doi.org/10.14714/cp56.307

Hirschmugl, M., Ofner, M., Raggam, J., & Schardt, M. (2007). Single tree detection in very high resolution remote sensing data. *Remote Sensing of Environment*, *110*(4), 533–544. https://doi.org/10.1016/j.rse.2007.02.029

Hofmann-Wellenhof et al. (2008). *Global navigation satellite systems (GNSS) are satellite-based positioning systems which provide three-dimensional geodetic coordinates of a measurement point*. Retrieved from https://www.sciencedirect.com/topics/earth-and-planetary-sciences/gnss/pdf

Hollaus, M., & Vreugdenhil, M. (2019, December 1). Radar Satellite Imagery for Detecting Bark Beetle Outbreaks in Forests. *Current Forestry Reports*, Vol. 5, pp. 240–250. https://doi.org/10.1007/s40725-019-00098-z

Huang, C. ying, Anderegg, W. R. L., & Asner, G. P. (2019). Remote sensing of forest die-off in the Anthropocene: From plant ecophysiology to canopy structure. *Remote Sensing of Environment*, *231*, 111233. https://doi.org/10.1016/j.rse.2019.111233

Immitzer, M., & Atzberger, C. (2014). Early detection of bark beetle infestation in Norway Spruce (Picea abies, L.) using worldview-2 data. *Photogrammetrie, Fernerkundung, Geoinformation*, *2014*(5), 351–367. https://doi.org/10.1127/1432-8364/2014/0229

Jiménez-Valverde, A. (2012). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, *21*(4), 498–507. https://doi.org/10.1111/j.1466-8238.2011.00683.x

Jing, L., Hu, B., Noland, T., & Li, J. (2012). An individual tree crown delineation method based on multi-scale segmentation of imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *70*, 88–98. https://doi.org/10.1016/j.isprsjprs.2012.04.003

July, P. (2020a). Differentiating Healthy and Bark Beetle Infected Spruce Trees With Sentinel-1 SAR.

Kanellakis, C., & Nikolakopoulos, G. (2017). Survey on Computer Vision for UAVs: Current Developments and Trends. *Journal of Intelligent and Robotic Systems: Theory and Applications*, *87*(1), 141–168.

https://doi.org/10.1007/s10846-017-0483-z

Kellner, J. R., Armston, J., Birrer, M., Cushman, K. C., Duncanson, L., Eck, C., … Zgraggen, C. (2019, July 15). New Opportunities for Forest Remote Sensing Through Ultra-High-Density Drone Lidar. *Surveys in Geophysics*, Vol. 40, pp. 959–977. https://doi.org/10.1007/s10712-019-09529-9

Kiadtikornthaweeyot, W., & Tatnall, A. R. L. (2016). Region of interest detection based on histogram segmentation for satellite image. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *41*, 249–255. https://doi.org/10.5194/isprsarchives-XLI-B7-249-2016

Kienast, F., Bolliger, J., & Zimmermann, N. E. (2012). *Species Distribution Modeling (SDM) with GLM, GAM and CART Dependent vs. independent variables: a conceptual ecological view*. *11*(1966), 1–16.

Kim, S. Y. (2009). Effects of sample size on robustness and prediction accuracy of a prognostic gene signature. *BMC Bioinformatics*, *10*, 147. https://doi.org/10.1186/1471-2105-10-147

Klouček, T., Komárek, J., Surový, P., Hrach, K., Janata, P., & Vašíček, B. (2019). The use of UAV mounted sensors for precise detection of bark beetle infestation. *Remote Sensing*, *11*(13). https://doi.org/10.3390/rs11131561

Kranjčić, N., & Rezo, M. (2018). Satellite-based hyperspectral imaging and cartographic visualization of bark beetle forest damage for the city of Čabar. *Tehnički Glasnik*, *12*(1), 39–43. https://doi.org/10.31803/tg-20171219085721

Kumar, D. N. (2017). Spectral reflectance curves. *Remote Sensing: Introduction and Basic Concepts*, (Gibson 2000), 1–13.

Kumar, M., & Miller, D. A. (2006). Anon-parametric classification strategy for remotely sensed images using both spectral and textural information. *Proceedings of the Third IASTED International Conference on Signal Processing, Pattern Recognition, and Applications*, *2006*, 81–89.

Lester, E. D. (1998). Feature Extraction , Image Segmentation , and Surface Fitting : the Development of a 3D Scene Reconstruction System. *Science*, (May).

Lewinsohn, E., Gijzen, M., & Croteau, R. (1991). Defense Mechanisms of Conifers. *Plant Physiology*, *96*(1), 44–49. https://doi.org/10.1104/pp.96.1.44

Lewis, H. G., & Brown, M. (2001). A generalized confusion matrix for assessing area estimates from remotely sensed data. *International Journal of Remote Sensing*, *22*(16), 3223–3235. https://doi.org/10.1080/01431160152558332

Li, C., Wang, J., Wang, L., Hu, L., & Gong, P. (2014). Comparison of classification algorithms and training sample sizes in urban land classification with landsat thematic mapper imagery. *Remote Sensing*, *6*(2), 964–983. https://doi.org/10.3390/rs6020964

Liarokapis, M. V., Artemiadis, P. K., & Kyriakopoulos, K. J. (2013). Task discrimination from myoelectric activity: A learning scheme for EMG-based interfaces. *IEEE International Conference on Rehabilitation Robotics*, 6. https://doi.org/10.1109/ICORR.2013.6650366

Lin, Chin TengLi, C. H., Kuo, B. C., & Huang, C. S. (2012). A spatial-contextual support vector machine for remotely sensed image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(3), 784–799. https://doi.org/10.1109/TGRS.2011.2162246

Lottering, R., Mutanga, O., & Peerbhay, K. (2018). Detecting and mapping levels of Gonipterus scutellatus-induced vegetation defoliation and leaf area index using spatially optimized vegetation indices. *Geocarto International*, *33*(3), 277–292. https://doi.org/10.1080/10106049.2016.1250823

Machala, M., & Zejdová, L. (2014). Forest mapping through Object-based image analysis of multispectral and LiDAR Aerial data. *European Journal of Remote Sensing*, *47*(1), 117–131. https://doi.org/10.5721/EuJRS20144708

Marshall, G. J., Rees, W. G., & Dowdeswell, J. A. (1993). Limitations imposed by cloud cover on multitemporal visible band satellite data sets from polar regions. *Annals of Glaciology*, *17*(1987), 113–120. https://doi.org/10.1017/s0260305500012696

Mauri, A., Strona, G., & San-Miguel-Ayanz, J. (2017). Unmanned Aerial Vehicles (UAVs) in environmental biology. *European Journal of Ecology .EU-Forest, a High-Resolution Tree Occurrence Dataset for Europe*, *4*, 1–8. https://doi.org/10.1038/sdata.2016.123

Meddens, A. J. H., Hicke, J. A., Vierling, L. A., & Hudak, A. T. (2013). Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date Landsat imagery. *Remote Sensing of Environment*, *132*, 49–58. https://doi.org/10.1016/j.rse.2013.01.002

Mezei, P., Potterf, M., Škvarenina, J., Rasmussen, J. G., & Jakuš, R. (2019). Potential solar radiation as a driver for bark beetle infestation on a landscape scale. *Forests*, *10*(7), 1–12. https://doi.org/10.3390/f10070604

Morgan, A. J. (2017). Analyzing red and gray stages of bark beetle attack in the San Bernardino National Forest using remote sensing. Retrieved from http://adsabs.harvard.edu/abs/2017PhDT........68M

Munyati, C. (2018). Optimising multiresolution segmentation: delineating savannah vegetation boundaries in the Kruger National Park, South Africa, using Sentinel 2 MSI imagery. *International Journal of Remote Sensing*, *39*(18), 5997–6019. https://doi.org/10.1080/01431161.2018.1508922

Naimi, B., & Araújo, M. B. (2016). Sdm: A reproducible and extensible R platform for species distribution modelling. *Ecography*, *39*(4), 368–375. https://doi.org/10.1111/ecog.01881

Nicholas Clinton, *, Ashley Holt, Li Yanb, P. G. (1999). An accuracy assessment measure for object based image. *Archives*, 1189–1194. Retrieved from http://berkenviro.com/berkeleyimgseg/

Nowak, M. M., Dziób, K., & Bogawski, P. (2019). Unmanned Aerial Vehicles (UAVs) in environmental biology: A review. *European Journal of Ecology*, *4*(2), 56–74. https://doi.org/10.2478/eje-2018-0012

Nyan, S., Kamarudin, N., Wahab, Z. A., & Ismail, M. H. (2018). Mobile phone and tablet assisted positioning device for trekkers: The case for sultan Idris shah forestry education centre (SISFEC). *Malaysian Forester*, *81*(2), 174–182. Retrieved from https://www.researchgate.net/publication/326413718

Oliech, I. O. (2019). *Geo-Spatial Assessment of Forest Health Status Using UAS Technology in Anopoli ,*. 73.

Oniga, E., Breaban, A.-I., & Statescu, F. (2018). Determining the optimum number of ground control points for obtaining high precision results based on UAS images. *Proceedings*, *2*(7), 5165.

https://doi.org/10.3390/ecrs-2-05165

Ortiz, S. M., Breidenbach, J., & Kändler, G. (2013a). Early detection of bark beetle green attack using terraSAR-X and rapideye data. *Remote Sensing*, *5*(4), 1912–1931. https://doi.org/10.3390/rs5041912

Ortiz, S. M., Breidenbach, J., & Kändler, G. (2013b). Early detection of bark beetle green attack using terraSAR-X and rapideye data. *Remote Sensing*, *5*(4), 1912–1931. https://doi.org/10.3390/rs5041912

Paneque-Gálvez, J., McCall, M. K., Napoletano, B. M., Wich, S. A., & Koh, L. P. (2014). Small drones for community-based forest monitoring: An assessment of their feasibility and potential in tropical areas. *Forests*, *5*(6), 1481–1507. https://doi.org/10.3390/f5061481

Pavel, U. (2016). Classification accuracy assessment (confusion matrix method. Retrieved February 17, 2021, from 23/09/2016 website: http://www.50northspatial.org/classification-accuracy-assessment-confusion-matrix-method/

Perez-Jaume, S., Pallarès, N., Skaltsa, K., & Carrasco, J. L. (2017). ThresholdROC: Optimum threshold estimation tools for continuous diagnostic tests in R. *Journal of Statistical Software*, *82*(4). https://doi.org/10.18637/jss.v082.i04

Primasatya, R. F., Hussin, Y. A., & Van Leeuwen, L. (2016). Terrestrial Laser scanning to support carbon estimation in nature conservation area: A case study of Haagse Bos and Snippert Forest, Netherlands. *37th Asian Conference on Remote Sensing, ACRS 2016*, *2*, 935–944.

Protection, F. (n.d.). Forest Protection, Bark Beetle Trap, Southern Bohemia, Czech. Retrieved May 25, 2021, from https://www.123rf.com/photo_134217366_forest-protection-bark-beetle-trap-southern-bohemia-czech-republic.html

Ramezan, C. A., Warner, T. A., & Maxwell, A. E. (2019). Evaluation of sampling and cross-validation tuning strategies for regional-scale machine learning classification. *Remote Sensing*, *11*(2). https://doi.org/10.3390/rs11020185

Rath, M., & Nandeesha D. Shekarappa. (2007). *Ground control station for UAV*. Retrieved from https://patentimages.storage.googleapis.com/a2/d8/a3/dd63327e9252b0/US20070244608A1.pdf

Robert Gove, J. F. (2012). support vector machines - an overview | ScienceDirect Topics. Retrieved January 4, 2021, from https://www.sciencedirect.com/topics/computer-science/support-vector-machines

Rohith Gandhi. (2018). Support Vector Machine — Introduction to Machine Learning Algorithms | by Rohith Gandhi | Towards Data Science. *Towards Data Science*, 1–9. Retrieved from https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47

Royo, S., & Ballesta-Garcia, M. (2019). An overview of lidar imaging systems for autonomous vehicles. *Applied Sciences (Switzerland)*, *9*(19). https://doi.org/10.3390/app9194093

Rumsey, D. J. (2021). *How z-Values Are Used in Statistics - dummies*. Retrieved from https://www.dummies.com/education/math/statistics/how-z-values-are-used-in-statistics/

Schuck, A., & van der Maaten, M. (2013). Integration of Nature Protection in Forest Policy in the Netherlands INTEGRATE Country Report. *European Forest Institute*.

Schuckman, K. (2020). Digital Image Classification: Exploring Imagery and Elevation Data in GIS Applications. Retrieved May 30, 2021, from https://www.e-education.psu.edu/geog480/node/496

Seidl, R., Müller, J., Hothorn, T., Bässler, C., Heurich, M., & Kautz, M. (2016). Small beetle, large-scale drivers: How regional and landscape factors affect outbreaks of the European spruce bark beetle. *Journal of Applied Ecology*, *53*(2), 530–540. https://doi.org/10.1111/1365-2664.12540

Seifert, E., Seifert, S., Vogt, H., Drew, D., van Aardt, J., Kunneke, A., & Seifert, T. (2019). Influence of drone altitude, image overlap, and optical sensor resolution on multi-view reconstruction of forest images. *Remote Sensing*, *11*(10). https://doi.org/10.3390/rs11101252

Sheykhmousa, Mohammadreza;Masoud, Mahdianpari;Hamid, G., Fariba, Mohammadimanesh; Pedram, G., & Saeid, H. (2020). Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING*, *13*. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9206124

Six, D. L., & Wingfield, M. J. (2011). The role of phytopathogenicity in bark beetle-fungus symbioses: A challenge to the classic paradigm. *Annual Review of Entomology*, *56*(May 2014), 255–272. https://doi.org/10.1146/annurev-ento-120709-144839

Srisha, R., & Khan, A. (2013). Morphological Operations for Image Processing : Understanding and its Applications. *NCVSComs-13*, (December), 17–19. Retrieved from https://www.researchgate.net/publication/272484795_Morphological_Operations_for_Image_Processing_Understanding_and_its_Applications

Stöcker, C., Nex, F., Koeva, M., & Gerke, M. (2020). High-quality uav-based orthophotos for cadastral mapping: Guidance for optimal flight configurations. *Remote Sensing*, *12*(21), 1–23. https://doi.org/10.3390/rs12213625

Sun, J., Yang, J., Zhang, C., Yun, W., & Qu, J. (2013). Automatic remotely sensed image classification in a grid environment based on the maximum likelihood method. *Mathematical and Computer Modelling*, *58*(3–4), 573–581. https://doi.org/10.1016/j.mcm.2011.10.063

Teillet, P. M., Staenz, K., & Williams, D. J. (1997). Effects of spectral, spatial, and radiometric characteristics on remote sensing vegetation indices of forested regions. *Remote Sensing of Environment*, *61*(1), 139–149. https://doi.org/10.1016/S0034-4257(96)00248-9

Théau, J., Gavelle, E., & Ménard, P. (2020). Crop scouting using UAV imagery: A case study for potatoes. *Journal of Unmanned Vehicle Systems*, *8*(2), 99–118. https://doi.org/10.1139/juvs-2019-0009

Tonks, D. (2020). Facts and Information. In *Teaching AIDS* (pp. 43–60). https://doi.org/10.4324/9780203406007-8

Tzotsos, A., & Argialas, D. (2008). Support vector machine classification for object-based image analysis. *Lecture Notes in Geoinformation and Cartography*, *0*(9783540770572), 663–677. https://doi.org/10.1007/978-3-540-77058-9_36

USDA Forest Service. (2014). Bark Beetles - US Forest Service Research & Development. Retrieved May 25, 2021, from USDA website: https://www.fs.fed.us/research/invasive-species/insects/bark-beetle.php van Proosdij, A. S. J., Sosef, M. S. M., Wieringa, J. J., & Raes, N. (2016). Minimum required number of specimen records to develop accurate species distribution models. *Ecography*, *39*(6), 542–552. https://doi.org/10.1111/ecog.01509

Viera, A. J., & Garrett, J. M. (2005). Anthony J. Viera, MD; Joanne M. Garrett, PhD (2005). Understanding interobserver agreement: the kappa statistic. Fam Med 2005;37(5):360-63. *Family Medicine*, *37*(5), 360–363. Retrieved from http://www1.cs.columbia.edu/~julia/courses/CS6998/Interrater_agreement.Kappa_statistic.pdf

Walter, J. A., & Platt, R. V. (2013). Multi-temporal analysis reveals that predictors of mountain pine beetle infestation change during outbreak cycles. *Forest Ecology and Management*, *302*, 308–318. https://doi.org/10.1016/j.foreco.2013.03.038

Wang, L., Gong, P., & Biging, G. S. (2004). Individual tree-crown delineation and treetop detection in high-spatial-resolution aerial imagery. *Photogrammetric Engineering and Remote Sensing*, *70*(3), 351–357. https://doi.org/10.14358/PERS.70.3.351

Wang, Z., Wang, T., Darvishzadeh, R., Skidmore, A. K., Jones, S., Suarez, L., … Hearne, J. (2016). Vegetation indices for mapping canopy foliar nitrogen in a mixed temperate forest. *Remote Sensing*, *8*(6), 491. https://doi.org/10.3390/rs8060491

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. https://doi.org/10.1016/j.foreco.2004.07.018

Wikiloc | Haagse Bos (Enschede/Losser) Trail. (2020). Retrieved December 20, 2020, from https://www.wikiloc.com/hiking-trails/haagse-bos-enschede-losser-14362344

Wingtra. (2021). Ground control points: How many do you need (and when are checkpoints enough)? | Wingtra. Retrieved May 29, 2021, from https://wingtra.com/ground-control-points-how-many-do-you-need-and-when-are-checkpoints-enough/

Wulder, M. A., Dymond, C. C., White, J. C., Leckie, D. G., & Carroll, A. L. (2006). Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities. *Forest Ecology and Management*, *221*(1–3), 27–41. https://doi.org/10.1016/j.foreco.2005.09.021

Wulder, M. A., White, J. C., Coops, N. C., & Butson, C. R. (2008). Multi-temporal analysis of high spatial resolution imagery for disturbance monitoring. *Remote Sensing of Environment*, *112*(6), 2729–2740. https://doi.org/10.1016/j.rse.2008.01.010

Xiao, Q., & McPherson, E. G. (2005). Tree health mapping with multispectral remote sensing data at UC Davis, California. *Urban Ecosystems*, *8*(3–4), 349–361. https://doi.org/10.1007/s11252-005-4867-7

Yang, S. (2019a). Detecting bark beetle damage with Sentinel-2 multi-temporal data in Sweden. In *Student thesis series INES; (2019)*. Retrieved from http://lup.lub.lu.se/student-papers/record/8989179

Yang, S. (2019b). *Student thesis series INES 490 Detecting bark beetle damage with Sentinel-2 multi-temporal data in Sweden*.

Yoder, B. J., & Waring, R. H. (1994). The normalized difference vegetation index of small Douglas-fir canopies with varying chlorophyll concentrations. *Remote Sensing of Environment*, *49*(1), 81–91.

https://doi.org/10.1016/0034-4257(94)90061-2

Yuan, Y., & Hu, X. (2016). Random forest and objected-based classification for forest pest extraction from uav aerial imagery. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *2016-Janua*, 1093–1098. https://doi.org/10.5194/isprsarchives-XLI-B1-1093-2016

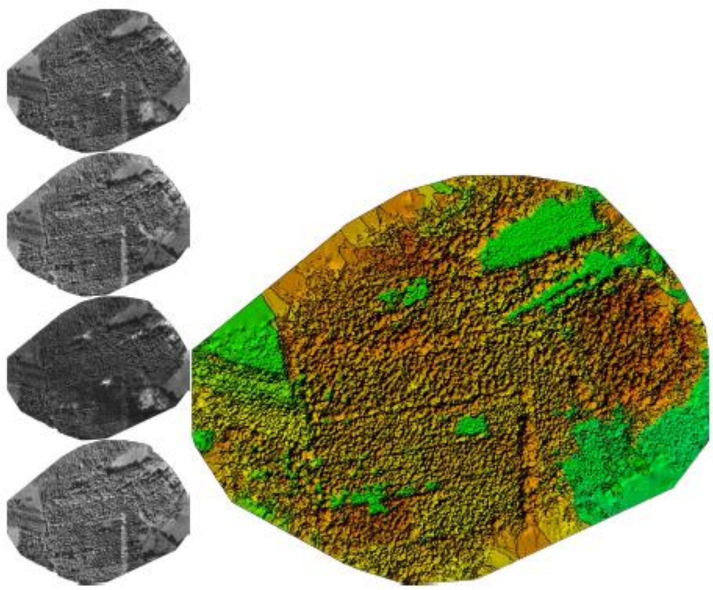
APPENDICES

Appendix 1: Summary of the initial quality report for block one generated by Pix4D Mapper for UAV multispectral image

Quality Check

Images	median of 10000 keypoints per image	✓
Dataset	3468 out of 3496 images calibrated (99%), all images enabled	✓
Camera Optimization	0.03% relative difference between initial and optimized internal camera parameters	✓
Matching	median of 4796.65 matches per calibrated image	✓
Georeferencing	yes, 6 GCPs (6 3D), mean RMS error = 0.07 m	✓

Preview

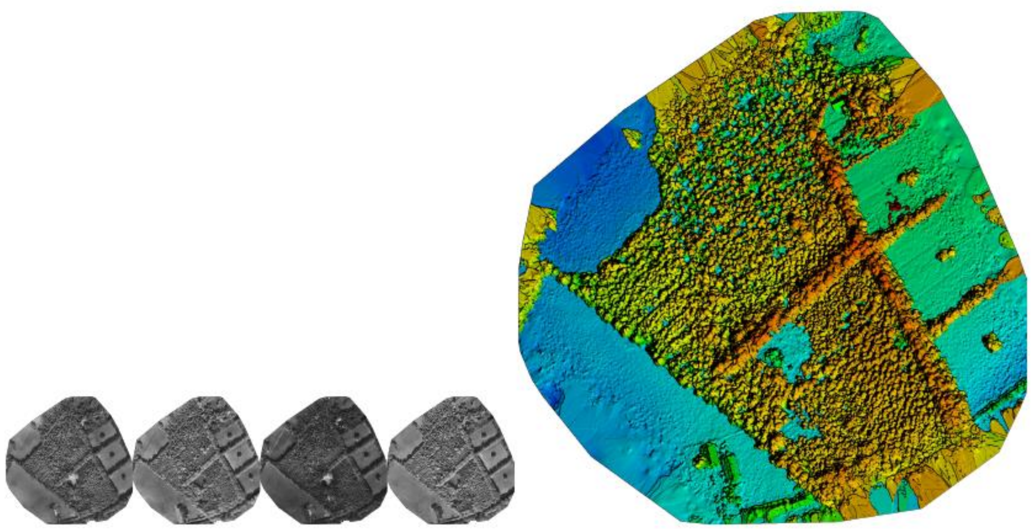


Appendix 2: Summary of the initial quality report for block three generated by Pix4D Mapper for UAV multispectral image

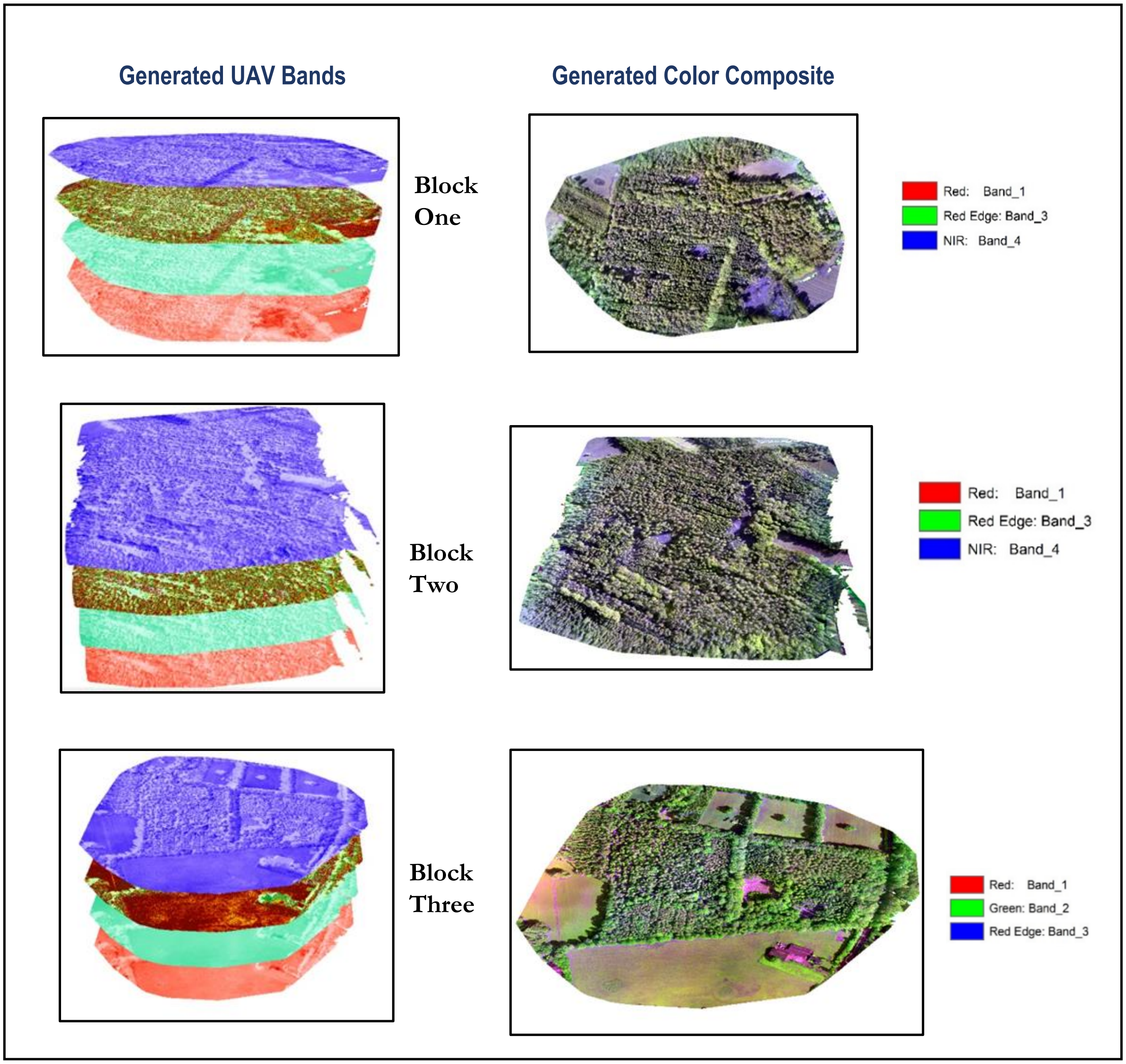
Quality Check

Images	median of 10000 keypoints per image	
Dataset	5588 out of 5788 images calibrated (96%), 28 images disabled, 3 blocks	
Camera Optimization	0% relative difference between initial and optimized internal camera parameters	
Matching	median of 3844.34 matches per calibrated image	
Georeferencing	yes, 3 GCPs (3 3D), mean RMS error = 0.287 m	

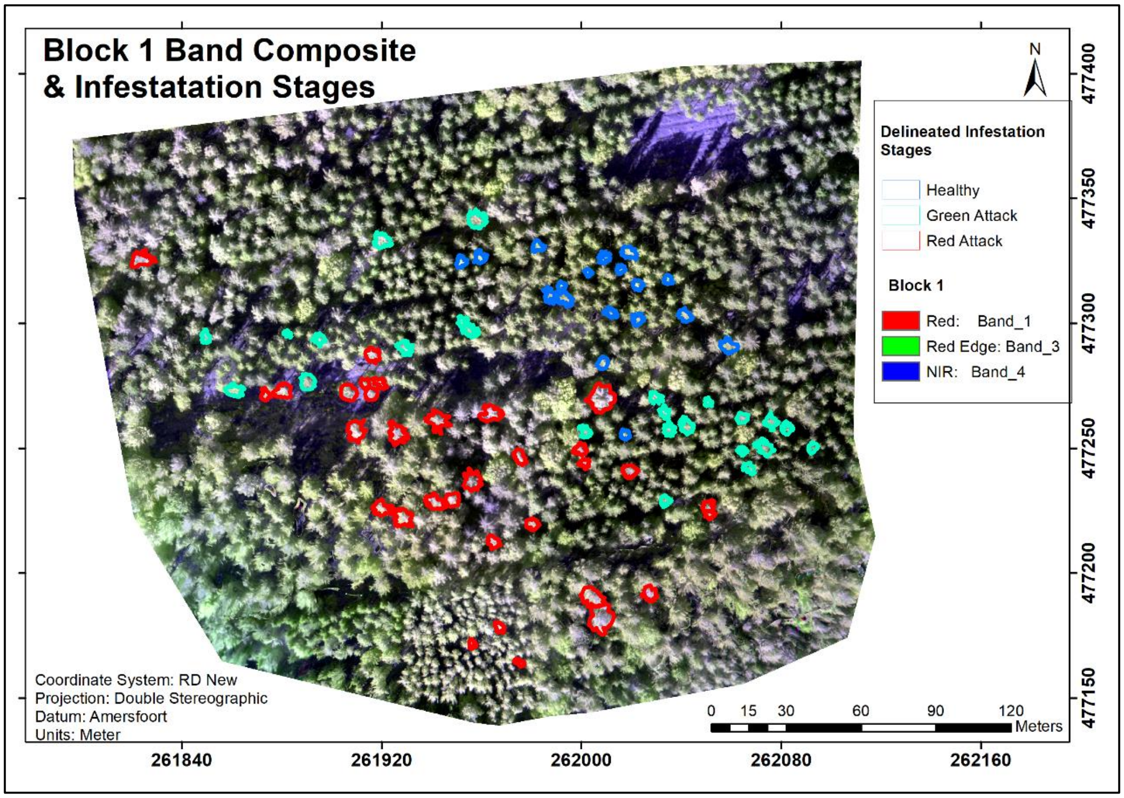
Preview



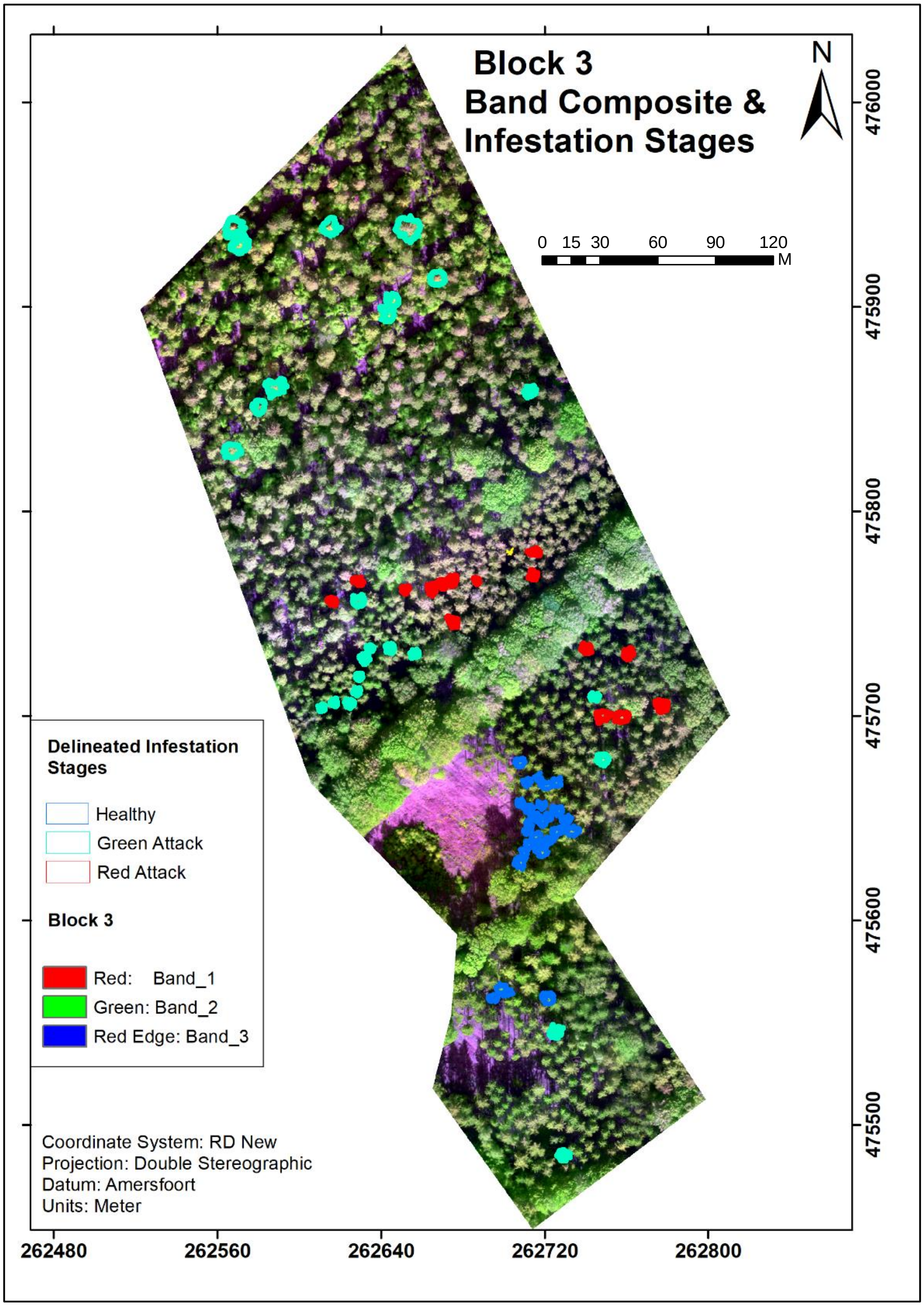
Appendix 3: Band composites of UAV Multispectral Image (Red, green, Red edge and Near Infra-Red) for block one and three



Appendix 4: Block one overlaid with manually digitized bark beetle infestation stages tree crowns



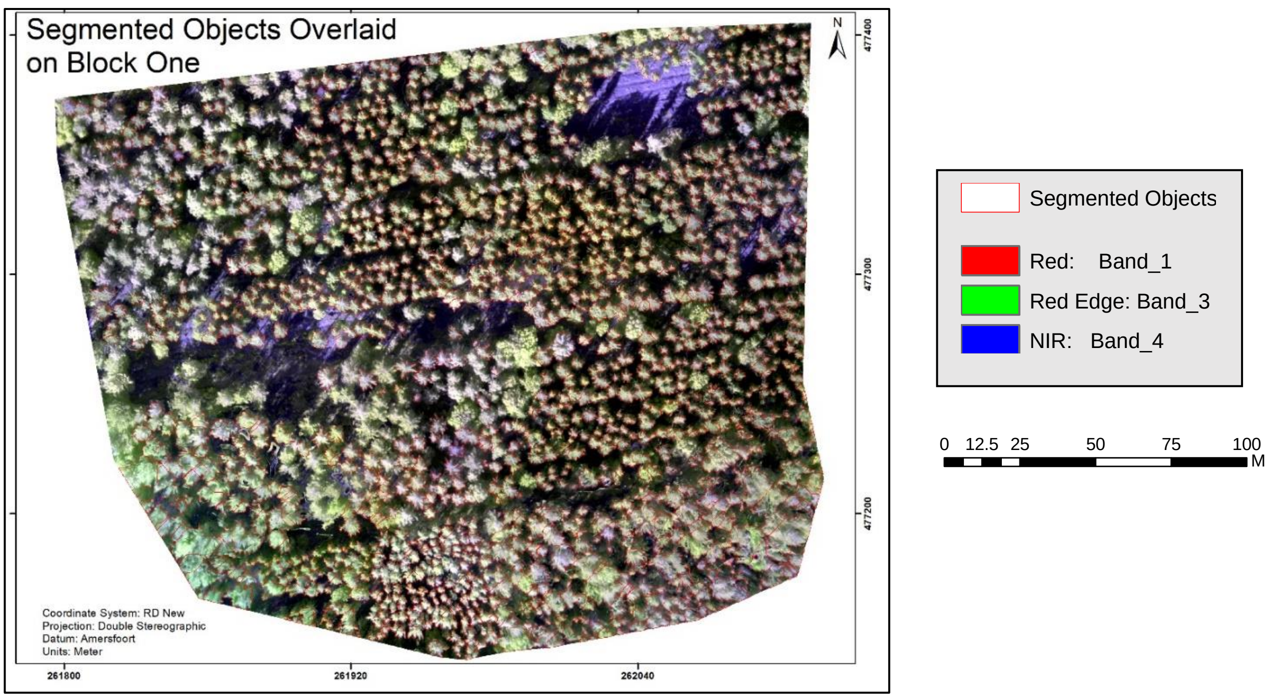
Appendix 5: Block three overlaid with manually digitized bark beetle infestation stages



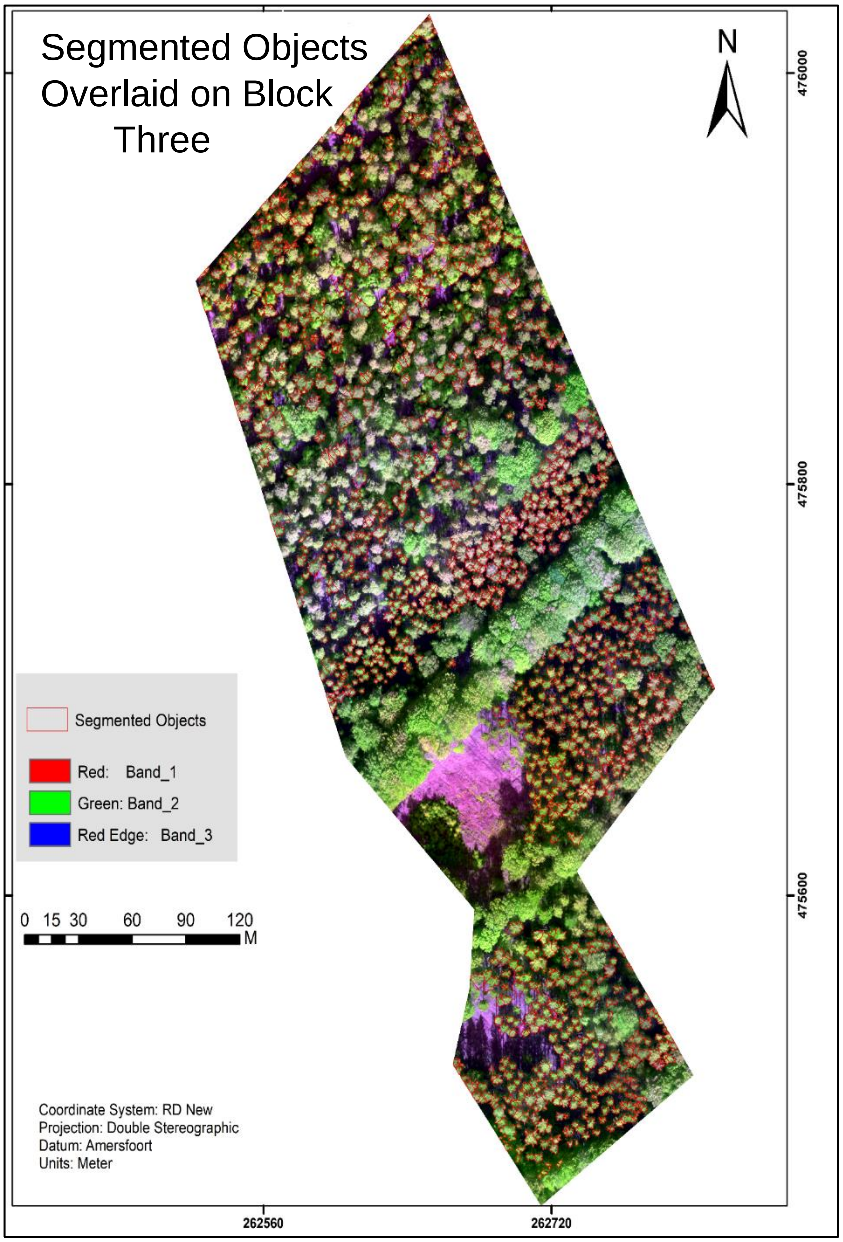
Appendix 6: Ruleset under the process tree used for image segmentation of block three

```
▪ Segementation
    36.297   40 [shape:0.9 compct.:0.5] creating 'New Level'
    ▪ 0.469   Shadow Masking
        0.062   with Brightness <= 20000  at  New Level: Shadow
    ▪ Morphology
        Trees1 at  New Level: closing: Trees1
        Trees1 at  New Level: watershed transformation (40)
▪ Trees
    Trees1 at  New Level: merge region
    0.094   with Brightness > 20000  and Brightness < 65500  at  New Level: Trees
    <0.001s   Shadow at  New Level: merge region
```

Appendix 7: Segmented objects overlaid on masked European spruce block one



Appendix 8: Segmented objects overlaid on masked European on block three



Appendix 9: Segmented objects (Red) overlaid with digitized tree polygons (Green) block one

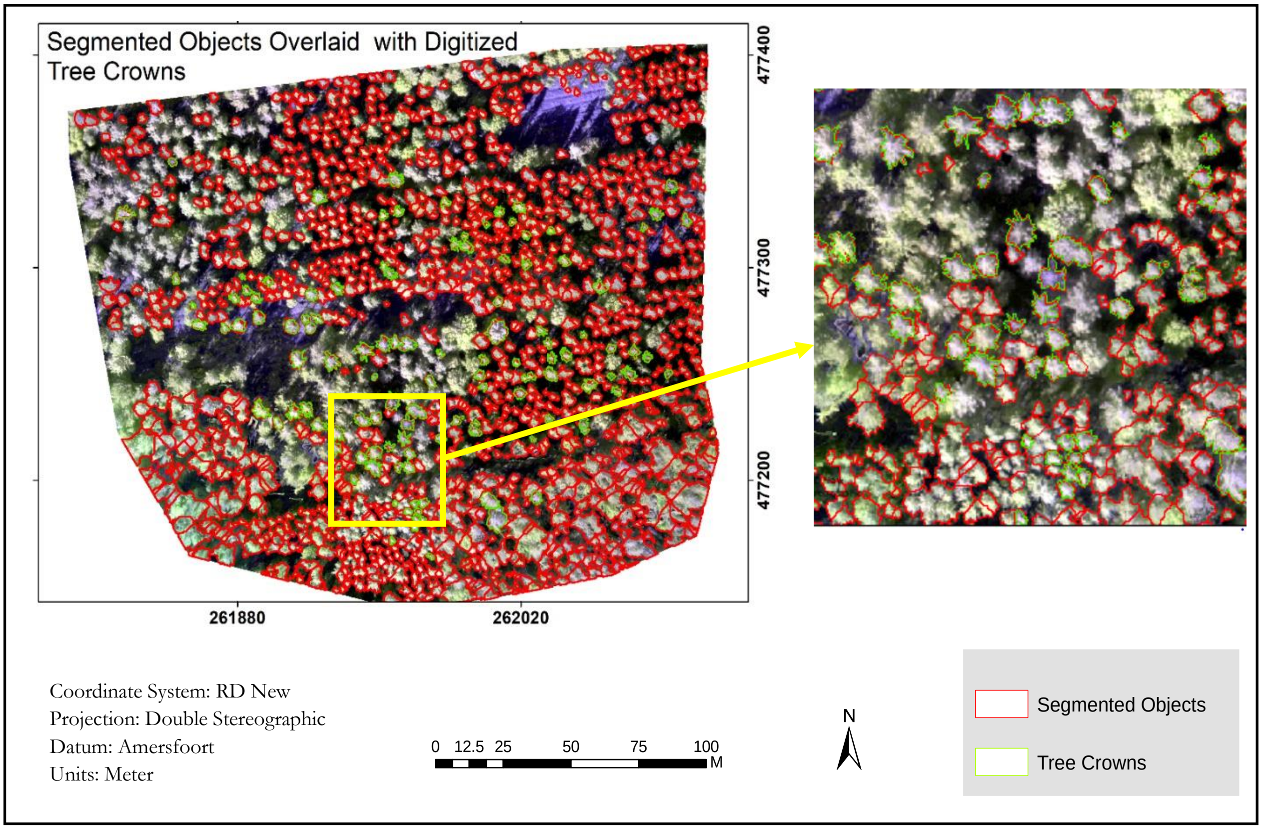


Image segmentation accuracy assessment results for block one

Block One Image Segmentation Accuracy Assessment Results			
Variables	**No of Segments**	**Area/Results**	**Accuracy %**
Segmented Objects	118	2404.37	
Digitized Objects	115	2246.66	
Intersection	111	2085.26	
Under Segmentation	77	0.07	93%
Over Segmentation	34	0.13	87%
Total Detected Error		0.15	
Accuracy %			**85%**

Appendix 10: Segmented objects (Red) overlaid with digitized tree polygons (Green) block three

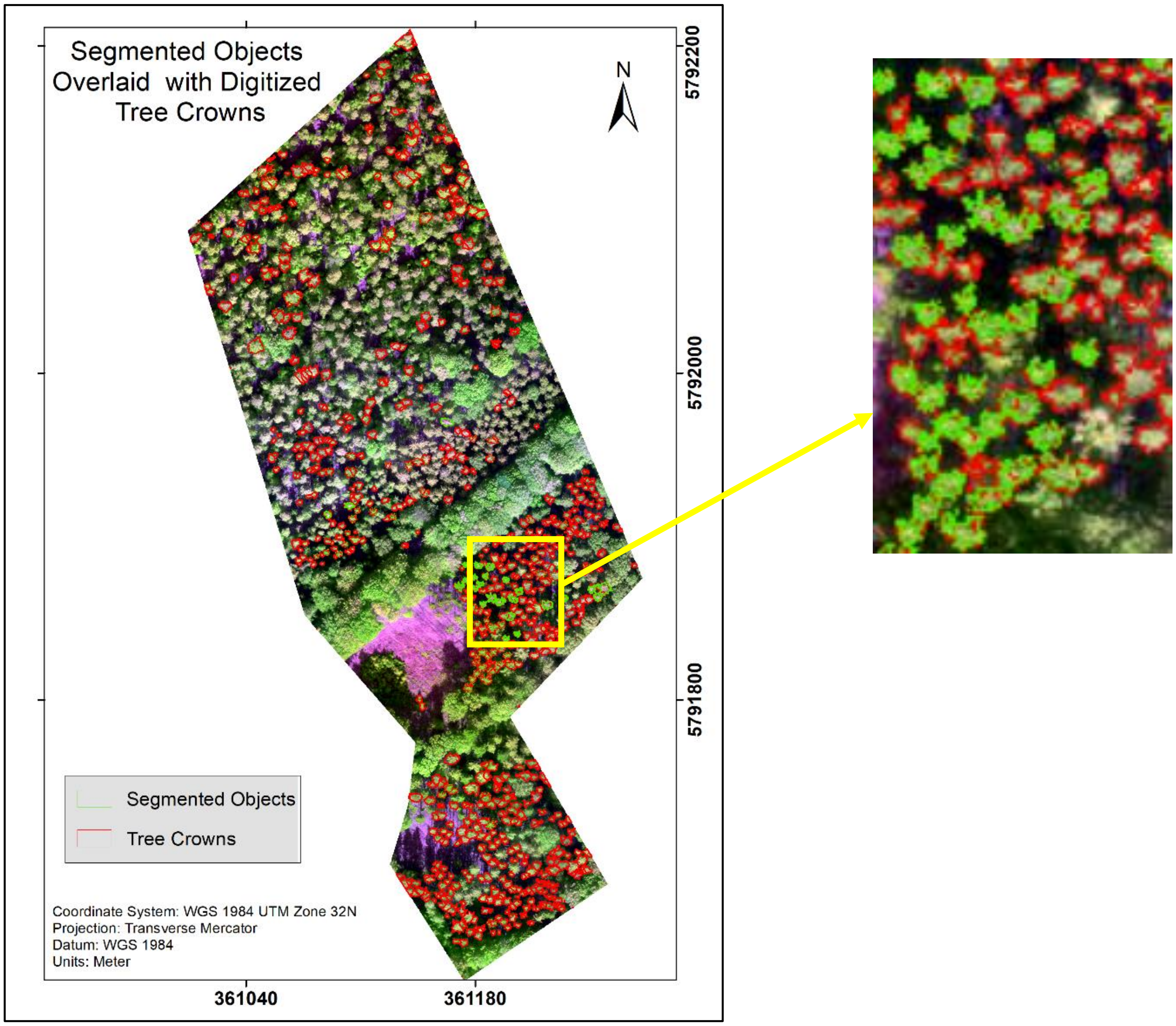
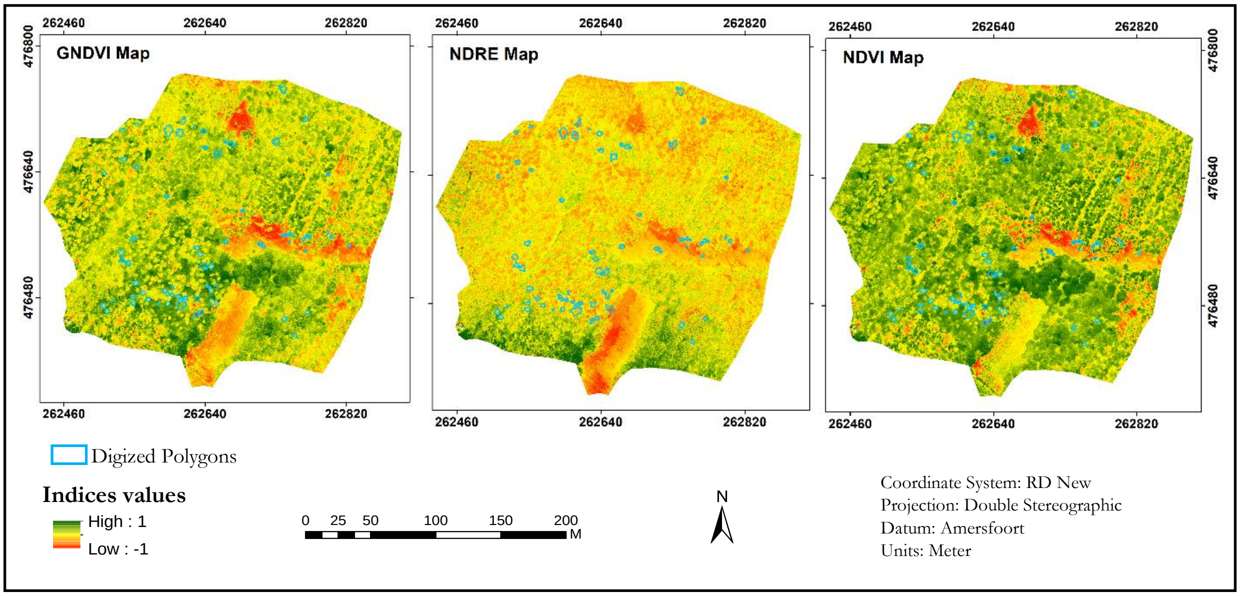


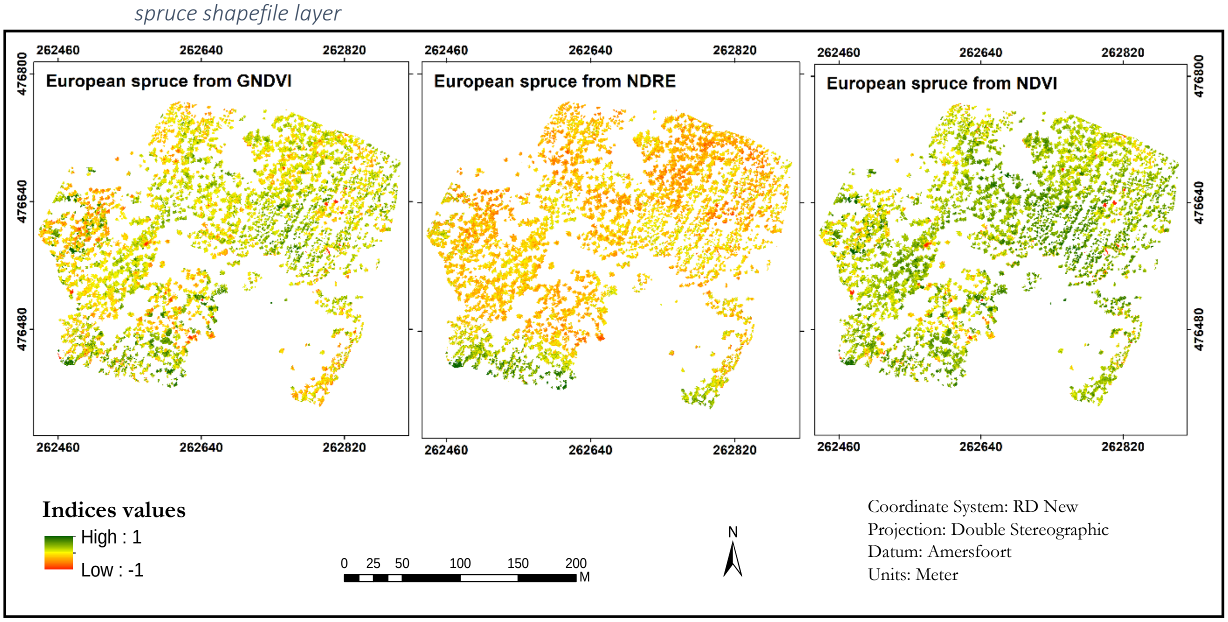
Image segmentation accuracy assessment results for block three

Block Three Image Segmentation Accuracy Assessment Results			
Variables	**No of Segments**	**Area/Results**	**Accuracy %**
Segmented Objects	142	1928.61	
Digitized Objects	94	1427.80	
Intersection	115	1224.34	
Under Segmentation	68	0.14	86
Over Segmentation	44	0.37	63
Total Detected Error		0.2	
Accuracy %			**80**

Appendix 11: GNDVI, NDRE and NDVI maps for block two overlaid with manually digitized field samples



Appendix 12: European spruce from GNDVI, NDRE and NDVI maps for block two based on object segmented European spruce shapefile layer



Appendix 13: SDM code for the prediction of green attack bark beetle stage on European spruce using GNDVI, NDRE & NDVI

```
library(raster)
#install.packages(rgdal)
library(rgdal)
library(sp)
library(spam)
library(dismo)
?dismo
#library(biomod2)
?biomod2
#roc('model_heathy_step')
library(PresenceAbsence)
#install.packages("sdm")
#library(sdm)
#library(gbm)

setwd("D:\\Second year\\Second year\\Thesis\\Research Data\\Second
Objective\\R\\Indices_files\\Indices")

#d<-read.table(file="GNDVI.txt", header = TRUE, sep=",")
d_GSD3_train<-read.csv(file="B3_TrainingSTD.csv", header = TRUE, sep=",")
d_GSD3_test<-read.csv(file="B3_TestingSTD.csv", header = TRUE, sep=",")

d_GSD2_train<-read.csv(file="B2_TrainingSTD.csv", header = TRUE, sep=",")
d_GSD2_test<-read.csv(file="B2_TestingSTD.csv", header = TRUE, sep=",")

d_GSD1_train<-read.csv(file="B1_TrainingSTD.csv", header = TRUE, sep=",")
d_GSD1_test<-read.csv(file="B1_TestingSTD.csv", header = TRUE, sep=",")

head(d)
B3_GNDVI<-raster("New_Maps\\img\\GNDVI.img")
B3_NDRE<-raster("New_Maps\\img\\NDRE.img")
B3_NDVI<-raster("New_Maps\\img\\NDVI.img")

B1_GNDVI<-raster("New_Maps\\img\\B1_GNDVI.img")
B1_NDRE<-raster("New_Maps\\img\\B1_NDRE.img")
B1_NDVI<-raster("New_Maps\\img\\B1_NDVI.img")

B2_GNDVI<-raster("New_Maps\\img\\B2_GNDVI.img")
B2_NDRE<-raster("New_Maps\\img\\B2_NDRE.img")
B2_NDVI<-raster("New_Maps\\img\\B2_NDVI.img")

plot(B3_GNDVI)
plot(B3_NDRE)
plot(B3_NDVI)

plot(B1_GNDVI)
plot(B1_NDRE)
plot(B1_NDVI)
```

```
plot(B2_GNDVI)
plot(B2_NDRE)
plot(B2_NDVI)

names(B3_GNDVI)
names(B3_NDRE)
names(B3_NDVI)

names(B1_GNDVI)
names(B1_NDRE)
names(B1_NDVI)

names(B2_GNDVI)
names(B2_NDRE)
names(B2_NDVI)

Model_green_attack_B3_GNDVI<-glm(Presence~GNDVI, data=d_GSD3_train, family = "binomial")
Model_green_attack_B3_NDRE<-glm(Presence~NDRE, data=d_GSD3_train, family = "binomial")
Model_green_attack_B3_NDVI<-glm(Presence~NDVI, data=d_GSD3_train, family = "binomial")

Model_green_attack_B1_GNDVI<-glm(Presence~B1_GNDVI, data=d_GSD1_train, family = "binomial")
Model_green_attack_B1_NDRE<-glm(Presence~B1_NDRE, data=d_GSD1_train, family = "binomial")
Model_green_attack_B1_NDVI<-glm(Presence~B1_NDVI, data=d_GSD1_train, family = "binomial")

Model_green_attack_B2_GNDVI<-glm(Presence~GNDVI, data=d_GSD2_train, family = "binomial")
Model_green_attack_B2_NDRE<-glm(Presence~NDRE, data=d_GSD2_train, family = "binomial")
Model_green_attack_B2_NDVI<-glm(Presence~NDVI, data=d_GSD2_train, family = "binomial")

summary(Model_green_attack_B3_GNDVI)
summary(Model_green_attack_B3_NDRE)
summary(Model_green_attack_B3_NDVI)

summary(Model_green_attack_B1_GNDVI)
summary(Model_green_attack_B1_NDRE)
summary(Model_green_attack_B1_NDVI)

summary(Model_green_attack_B2_GNDVI)
summary(Model_green_attack_B2_NDRE)
summary(Model_green_attack_B2_NDVI)

Predict_map_green_attack_B3_GNDVI<-predict(B3_GNDVI,Model_green_attack_B3_GNDVI,
                    type="response", file="Predict_green_attack_GNDVI.img")

Predict_map_green_attack_B3_NDRE<-predict(B3_NDRE,Model_green_attack_B3_NDRE,
                    type="response", file="Predict_green_attack_NDRE.img")

Predict_map_green_attack_B3_NDVI<-predict(B3_NDVI,Model_green_attack_B3_NDVI,
                    type="response", file="Predict_green_attack_NDVI.img")
```

```
Predict_map_green_attack_B1_GNDVI<-predict(B1_GNDVI,Model_green_attack_B1_GNDVI,
                    type="response", file="Predict_green_attack_B1_GNDVI.img")

Predict_map_green_attack_B1_NDRE<-predict(B1_NDRE,Model_green_attack_B1_NDRE,
                    type="response", file="Predict_green_attack_B1_NDRE.img")

Predict_map_green_attack_B1_NDVI<-predict(B1_NDVI,Model_green_attack_B1_NDVI,
                    type="response", file="Predict_green_attack_B1_NDVI.img")

Predict_map_green_attack_B2_GNDVI<-predict(B2_GNDVI,Model_green_attack_B2_GNDVI,
                    type="response", file="Predict_green_attack_B2_GNDVI.img")

Predict_map_green_attack_B2_NDRE<-predict(B2_NDRE,Model_green_attack_B2_NDRE,
                    type="response", file="Predict_green_attack_B2_NDRE.img")

Predict_map_green_attack_B2_NDVI<-predict(B2_NDVI,Model_green_attack_B2_NDVI,
                    type="response", file="Predict_green_attack_B2_NDVI.img")

plot(Predict_map_green_attack_B3_GNDVI, main="Predict_map_Green_Attack_B3_GNDVI")
plot(Predict_map_green_attack_B3_NDRE, main= 'Predict_map_Green_Attack_B3_NDRE')
plot(Predict_map_green_attack_B3_NDVI, main='Predict_map_Green_Attack_B3_NDVI')

plot(Predict_map_green_attack_B1_GNDVI, main="Predict_map_Green_Attack_B1_GNDVI")
plot(Predict_map_green_attack_B1_NDRE, main= 'Predict_map_Green_Attack_B1_NDRE')
plot(Predict_map_green_attack_B1_NDVI, main='Predict_map_Green_Attack_B1_NDVI')

plot(Predict_map_green_attack_B2_GNDVI, main="Predicted_Green_Attack_GNDVI")
plot(Predict_map_green_attack_B2_NDRE, main= 'Predicted_Green_Attack_NDRE')
plot(Predict_map_green_attack_B2_NDVI, main='Predicted_Green_Attack_NDVI')

d_GSD3_test$test_B3_GNDVI_predict<-predict(Model_green_attack_B3_GNDVI,type="response",
newdata=d_GSD3_test)

d_GSD3_test$test_B3_NDRE_predict<-predict(Model_green_attack_B3_NDRE,type="response",
newdata=d_GSD3_test)

d_GSD3_test$test_B3_NDVI_predict<-predict(Model_green_attack_B3_NDVI,type="response",
newdata=d_GSD3_test)

d_GSD1_test$test_B1_GNDVI_predict<-predict(Model_green_stage_B1_GNDVI,type="response",
newdata=d_GSD1_test)

d_GSD1_test$test_B1_NDRE_predict<-predict(Model_green_stage_B1_NDRE,type="response",
newdata=d_GSD1_test)

d_GSD1_test$test_B1_NDVI_predict<-predict(Model_green_stage_B1_NDVI,type="response",
newdata=d_GSD1_test)
```

```
d_GSD2_test$test_B2_GNDVI_predict<-predict(Model_green_attack_B2_GNDVI,type="response",
newdata=d_GSD2_test)

d_GSD2_test$test_B2_NDRE_predict<-predict(Model_green_attack_B2_NDRE,type="response",
newdata=d_GSD2_test)

d_GSD2_test$test_B2_NDVI_predict<-predict(Model_green_attack_B2_NDVI,type="response",
newdata=d_GSD2_test)

roc.green_attack_B3_GNDVI<-data.frame(ID=1:length(d_GSD3_test$Presence),
                    PA=d_GSD3_test$Presence,
                    model=d_GSD3_test$test_B3_GNDVI_predict)

roc.green_attack_B3_NDRE<-data.frame(ID=1:length(d_GSD3_test$Presence),
                      PA=d_GSD3_test$Presence,
                      model=d_GSD3_test$test_B3_NDRE_predict)

roc.green_attack_B3_NDVI<-data.frame(ID=1:length(d_GSD3_test$Presence),
                         PA=d_GSD3_test$Presence,
                         model=d_GSD3_test$test_B3_NDVI_predict)

roc.green_attack_B1_GNDVI<-data.frame(ID=1:length(d_GSD1_test$Presence),
                          PA=d_GSD1_test$Presence,
                          model=d_GSD1_test$test_B1_GNDVI_predict)

roc.green_attack_B1_NDRE<-data.frame(ID=1:length(d_GSD1_test$Presence),
                         PA=d_GSD1_test$Presence,
                         model=d_GSD1_test$test_B1_NDRE_predict)

roc.green_attack_B1_NDVI<-data.frame(ID=1:length(d_GSD1_test$Presence),
                         PA=d_GSD1_test$Presence,
                         model=d_GSD1_test$test_B1_NDVI_predict)

roc.green_attack_B2_GNDVI<-data.frame(ID=1:length(d_GSD2_test$Presence),
                          PA=d_GSD2_test$Presence,
                          model=d_GSD2_test$test_B2_GNDVI_predict)

roc.green_attack_B2_NDRE<-data.frame(ID=1:length(d_GSD2_test$Presence),
                         PA=d_GSD2_test$Presence,
                         model=d_GSD2_test$test_B2_NDRE_predict)

roc.green_attack_B2_NDVI<-data.frame(ID=1:length(d_GSD2_test$Presence),
                         PA=d_GSD2_test$Presence,
                         model=d_GSD2_test$test_B2_NDVI_predict)

roc.green_attack_B2_GNDVI<-data.frame(ID=1:length(d_GSD2_test$Presence),
                           PA=d_GSD2_test$Presence,
                           model=d_GSD2_test$test_B2_GNDVI_predict)
```

```
roc.green_attack_B2_NDRE<-data.frame(ID=1:length(d_GSD2_test$Presence),
                          PA=d_GSD2_test$Presence,
                          model=d_GSD2_test$test_B2_NDRE_predict)

roc.green_attack_B2_NDVI<-data.frame(ID=1:length(d_GSD2_test$Presence),
                          PA=d_GSD2_test$Presence,
                          model=d_GSD2_test$test_B2_NDVI_predict)

optimal.thresholds(roc.green_attack_B3_GNDVI)
error.threshold.plot(roc.green_attack_B3_GNDVI,main="Block_3_GNDVI_green_attack")

optimal.thresholds(roc.green_attack_B3_NDRE)
error.threshold.plot(roc.green_attack_B3_NDRE,main="Block_3_NDRE_green_attack")

optimal.thresholds(roc.green_attack_B3_NDVI)
error.threshold.plot(roc.green_attack_B3_NDVI,main="Bock_3_NDVI_green_attack")

optimal.thresholds(roc.green_attack_B1_GNDVI)
error.threshold.plot(roc.green_attack_B1_GNDVI,main="Block_1_GNDVI_green_attack")

optimal.thresholds(roc.green_attack_B1_NDRE)
error.threshold.plot(roc.green_attack_B1_NDRE,main="Block_1_NDRE_green_attack")

optimal.thresholds(roc.green_attack_B1_NDVI)
error.threshold.plot(roc.green_attack_B1_NDVI,main="Block_1_NDVI_green_attack")

optimal.thresholds(roc.green_attack_B2_GNDVI)
error.threshold.plot(roc.green_attack_B2_GNDVI,main="Block_2_GNDVI_green_attack")

optimal.thresholds(roc.green_attack_B2_NDRE)
error.threshold.plot(roc.green_attack_B2_NDRE,main="Block_2_NDRE_green_attack")

optimal.thresholds(roc.green_attack_B2_NDVI)
error.threshold.plot(roc.green_attack_B2_NDVI,main="Block_2_NDVI_green_attack")

PA_green_attack_GNDVI<-presence.absence.accuracy(roc.green_attack_B3_GNDVI, threshold = 0.20)
PA_green_attack_NDRE<-presence.absence.accuracy(roc.green_attack_B3_NDRE, threshold = 0.17)
PA_green_attack_NDVI<-presence.absence.accuracy(roc.green_attack_B3_NDVI, threshold = 0.18)

PA_green_attack_B1_GNDVI<-presence.absence.accuracy(roc.green_attack_B1_GNDVI, threshold = 0.44)
PA_green_attack_B1_NDRE<-presence.absence.accuracy(roc.green_attack_B1_NDRE, threshold = 0.21)
PA_green_attack_B1_NDVI<-presence.absence.accuracy(roc.green_attack_B1_NDVI, threshold = 0.55)

PA_green_attack_B2_GNDVI<-presence.absence.accuracy(roc.green_attack_B2_GNDVI, threshold = 0.41)
PA_green_attack_B2_NDRE<-presence.absence.accuracy(roc.green_attack_B2_NDRE, threshold = 0.41)
PA_green_attack_B2_NDVI<-presence.absence.accuracy(roc.green_attack_B2_NDVI, threshold = 0.44)

class(d_G_test)
```

```
plot(Predict_map_green_attack_B3_GNDVI>0.20,main="TH(0.20)_B3_GNDVI_green_attack")
plot(Predict_map_green_attack_B3_NDRE>0.17,main="TH(0.34)_B3_NDRE_green_attack")
plot(Predict_map_green_attack_B3_NDVI>0.18,main="TH(0.34)_B3_NDVI_green_attack")

plot(Predict_map_green_attack_B1_GNDVI>0.44,main="TH(0.44)_B1_GNDVI_green_attack")
plot(Predict_map_green_attack_B1_NDRE>0.23,main="TH(0.23)_B1_NDRE_green_attack")
plot(Predict_map_green_attack_B1_NDVI>0.55,main="TH(0.55)_B1_NDVI_green_attack")

plot(Predict_map_green_attack_B2_GNDVI>0.41,main="TH(0.41)_B2_GNDVI_green_attack")
plot(Predict_map_green_attack_B2_NDRE>0.41,main="TH(0.41)_B2_NDRE_green_attack")
plot(Predict_map_green_attack_B2_NDVI>0.44,main="TH(0.44)_B2_NDVI_green_attack")

auc.roc.plot(roc.green_attack_B3_GNDVI, main="roc.Green_Attack_B3_GNDVI")
auc.roc.plot(roc.green_attack_B3_NDRE, main="roc.Green_Attack_B3_NDRE")
auc.roc.plot(roc.green_attack_B3_NDVI, main="roc.Green_Attack_B3_NDVI")

auc.roc.plot(roc.green_attack_B1_GNDVI, main="roc.Green_Attack_B1_GNDVI")
auc.roc.plot(roc.green_attack_B1_NDRE, main="roc.Green_Attack_B1_NDRE")
auc.roc.plot(roc.green_attack_B1_NDVI, main="roc.Green_Attack_B1_NDVI")

auc.roc.plot(roc.green_attack_B2_GNDVI, main="roc.Green_Attack_B2_GNDVI")
auc.roc.plot(roc.green_attack_B2_NDRE, main="roc.Green_Attack_B2_NDRE")
auc.roc.plot(roc.green_attack_B2_NDVI, main="roc.Green_Attack_B2_NDVI")
```