2011

Estimation of scots pine defoliation by the common pine sawfly (*Diprion pini* L.) using multi-temporal radar data

Petri T. Latva-Käyrä
Michigan Technological University

Recommended Citation

ESTIMATION OF SCOTS PINE DEFOLIATION BY THE COMMON PINE SAWFLY
(*DIPRION PINI* L.) USING MULTI-TEMPORAL RADAR DATA

By

Petri T. Latva-Käyrä

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Forestry

MICHIGAN TECHNOLOGICAL UNIVERSITY

2011

This thesis, "Estimation of Scots pine defoliation by the Common pine sawfly (*Diprion pini* L.) using multi-temporal radar data" is hereby approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE IN FORESTRY.

School of Forest Resources and Environmental Science

Signatures:

Thesis Advisor   _____

Assistant Professor Michael Falkowski

Thesis Co-Advisor   _____

Adjunct professor Markus Holopainen

Dean   _____

Margaret Gale

Date   _____

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

LIST OF ABBREVIATIONS

ACS	Adaptive Cluster Sampling
ALS	Airborne Laser Scanning
ENVISAT	Environmental Satellite
ERS	European Remote Sensing
ESA	European Space Agency
ETM	Enhanced Thematic Mapper
glm	generalized linear model
H	Horizontal
JERS	Japanese Earth Resource Satellite
kNN	k-Nearest Neighbor
LAI	Leaf Area Index
LDA	Linear Discriminant Analysis
LiDAR	Light Detection and Ranging
MSS	Multi-Spectral Scanner
NDMI	Normalized Difference Moisture Index
NDVI	Normalized Difference Vegetation Index
PCA	Principal Component Analysis
Radar	Radio Detection and Ranging
RAR	Real Aperture Radar
RMSE	Root Mean Square Error
RVI	Ratio Vegetation Index
SAR	Synthetic Aperture Radar
SPOT	Systeme Pour l'Observation de la Terre
SRS	Simple Random Sampling
TM	Thematic Mapper
V	Vertical
WEKA	Waikato Environment for Knowledge Analysis

ABSTRACT

In 1998-2001 Finland suffered the most severe insect outbreak ever recorded, over 500,000 hectares. The outbreak was caused by the common pine sawfly (*Diprion pini* L.). The outbreak has continued in the study area, Palokangas, ever since. To find a good method to monitor this type of outbreaks, the purpose of this study was to examine the efficacy of multi-temporal ERS-2 and ENVISAT SAR imagery for estimating Scots pine (*Pinus sylvestris* L.) defoliation. Three methods were tested: unsupervised k-means clustering, supervised linear discriminant analysis (LDA) and logistic regression. In addition, I assessed if harvested areas could be differentiated from the defoliated forest using the same methods. Two different speckle filters were used to determine the effect of filtering on the SAR imagery and subsequent results. The logistic regression performed best, producing a classification accuracy of 81.6% (kappa 0.62) with two classes (no defoliation, >20% defoliation). LDA accuracy was with two classes at best 77.7% (kappa 0.54) and k-means 72.8 (0.46). In general, the largest speckle filter, 5 x 5 image window, performed best. When additional classes were added the accuracy was usually degraded on a step-by-step basis. The results were good, but because of the restrictions in the study they should be confirmed with independent data, before full conclusions can be made that results are reliable. The restrictions include the small size field data and, thus, the problems with accuracy assessment (no separate testing data) as well as the lack of meteorological data from the imaging dates.

1. INTRODUCTION

1.1. Background

Worldwide climate has experienced considerable changes over the past few decades (Intergovernmental Panel on Climate Change 2007). The rate of change seems to be more pronounced at higher latitudes, such as those in the Nordic countries (Regniere 2009, Kantola et al. 2010). The changes (e.g. rising temperatures) have already resulted in various effects on species distribution and phenology, ultimately causing pest damage in managed forests (e.g. Dale et al. 2001, Evans et al. 2002, Walther et al. 2002).

Forest insects, formerly regarded as harmless species, are now causing severe damage in Finnish forests (Lyytikäinen-Saarenmaa and Tomppo 2002, de Somviele et al. 2007). Economic losses caused by defoliating pest insects can be substantial (~300 - 1000 EUR ha^{-1} depending on intensity of needle loss and the length of outbreak period) (Lyytikäinen-Saarenmaa et al. 2006). Thus, effectively predicting changes in the intensity and distribution of insect-caused forest damages has become an important topic in the field of forest research.

Developing methods for the rapid assessment and monitoring of forest areas affected by hazardous events such as insect outbreaks, as well as understanding the risks of e.g. soil type, stand characteristics and management methods, has gained a high importance in forest management (Lyytikäinen-Saarenmaa et al. 2008, Kantola et al. 2010). Remote sensing can produce data for large areas of remote, inaccessible forestlands quickly and often at a much lower cost than traditional ground surveys (Ciesla 2000, Hall et al. 2007). In many cases, symptoms of forest insect damage are visible from far distances. Furthermore, forest damages such as crown discoloration, defoliation, and dieback are even more perceptible from aerial views as compared to ground-based assessments. Thus, the synoptic view provided by satellite and airborne remote sensing platforms can enhance the assessment of the pattern of dispersion and range extension of insect outbreaks (Ciesla 2000, Hall et al. 2007).

Employing remotely sensed data in the assessment of insect damage is not without limitations. Often there is limited access to data with appropriate characteristics (e.g., cloud-free Landsat images on correct timing). When appropriate data are accessible, often time they are expensive, ultimately limiting practical implementation for forest monitoring applications. For example, developing monitoring protocols relying on high spatial resolution imagery such as aerial photographs or data such as Light Detection and Ranging (LiDAR) (which is sensitive to three-dimensional forest structure) is often times not feasible from an economic standpoint. More affordable middle spatial resolution (10–50 m) multi-spectral satellite imagery (e.g. Landsat) have long been used in environmental mapping, forest inventory and other applications. However, the availability of cloud-free imagery is often a problem. However, the continual development of remote sensing systems that can provide inexpensive, cloud free data may eventually alleviate many of the aforementioned limitations.

European Space Agency (ESA) European Remote Sensing Satellite-2 (ERS-2) Synthetic Aperture Radar (SAR) and Environmental Satellite (ENVISAT) SAR are two radio detection and ranging (radar) systems that can reliably provide inexpensive, cloud-free imagery. However, the efficacy of ERS-2 and ENVISAT SAR data for the characterization and monitoring of insect damage has not been tested. Therefore, the primary objective of this study is to evaluate ERS-2 and ENVISAT SAR data for characterizing and monitoring defoliation. Specifically, I employ three different approaches (linear discriminant analysis, k-means clustering and logistic regression) to predict the defoliation level of the boreal Scots pine in Finland. For more details on research questions and objectives, see chapter 1.6.

1.2. Study object

The Common pine sawfly (*Diprion pini* L.)(Hymenoptera, Diprionidae) is a univoltine species in Finland. The female lays eggs in the early summer on the previous year's needles of Scots pine (*Pinus sylvestris* L.) (Pinaceae). After about 3-4 weeks the larvae

hatch and begin to consume the pine needles. The consumption of pine needles usually occurs during July and August. However, the consumption can continue through September if the mean temperature of summer months is low. During the winter *D. pini* pupates in the ground and the insects hatch out after the winter in May to July. *D. pini* is considered one of the most destructive needle feeding pest species in Central European forests (Viitasaari and Varama 1987). In Finland, the damage caused by *D. pini* are usually not as economically important as are damage caused by its relative, the European pine sawfly (*Neodiprion sertifer* Geoffr.) (Viitasaari and Varama 1987). However, future climate change might make the conditions in Finland better suited for *D. pini* and, thus, cause more severe and more frequent outbreaks. At present, the frequency of *D. pini* outbreaks in Finland is sparser as compared to Central Europe. This is primarily because the species is univoltine. Also, the diapause stage is more common for *D. pini* in northern conditions such as in Finland. The outbreak area is usually from hundreds to tens of thousands of hectares and the duration varies between 1-3 years (Viitasaari and Varama 1987.) The most recent outbreak of *D. pini* in Finland has been unusually large. It has also been long lasting in certain areas of Finland areas such as Palokangas, in Eastern-Finland. Furthermore, the area and duration of the outbreak is regulated by predators and parasites (Viitasaari and Varama 1987).

D. pini prefers commonly dry heath sites and other dry sites such as clifftops (Viitasaari and Varama 1987). Stand age is considered the main factor affecting the intensity of the defoliation. Secondary factors include tree species composition, stand area, stand openness, soil characteristics and age differences between surrounding stands. For instance, infertile, shallow and well-drained sites are more vulnerable to an outbreak than other sites (De Somviele et al. 2004.) Monoculture in tree species composition also increases the risk of infection (De Somviele et al. 2004). Therefore, the managed boreal pine forests of Finland are more vulnerable.

D. pini consumes both old and current year needles, which makes the species more harmful than its relative *N sertifer*. The latter species does not consume the current year's needles, giving the pines a greater chance of survival and recovery from needle losses

(Lyytikäinen-Saarenmaa and Tomppo 2002). Despite this, the current opinion is that *N. Sertifer* has the highest potential to cause damage in Finnish Scots pine forests and produce outbreaks (De Somviele et al. 2004). *D. pini* feeds on stands of all ages, but typically prefers maturing and mature stands. At outbreak densities, seedlings can also suffer from defoliation (Geri 1988, de Somviele et al. 2004, Lyytikäinen-Saarenmaa et al. 2006). Defoliation by *D. pini* causes considerable growth losses. The full recovery of a tree may take several years. Consuming the needles during several years may even kill the tree (Lyytikäinen-Saarenmaa and Tomppo 2002). Mortality in Scots pine stands after a single year outbreak is usually quite small. Single year outbreaks result in slower growth and, thus, some economical losses. More frequent defoliation in trees or an outbreak lasting several years can increase mortality rates or at least severe deceleration in tree growth (Lyytikainen-Saarenmaa et al. 2006). Studies have demonstrated that tree mortality in infested stands can range from 10% to 80%, depending on the defoliation level of a stand after an outbreak (Lyytikäinen-Saarenmaa and Tomppo 2002, Lyytikäinen-Saarenmaa et al. 2006). Defoliation can cause water stress in the trees, which can also result in tree mortality. Increased water stress also makes the trees more susceptible to secondary damages from other insects such as bark beetles (Viitasaari and Varama 1987).

During 1998-2001 Finland experienced its first large scaled outbreak of *D. pini*. This outbreak was largest insect outbreak recorded in Finland. Over 500,000 hectares of forest was affected and massive defoliation occurred in the affected area (Lyytikäinen-Saarenmaa and Tomppo 2002). The outbreak reached the Ilomantsi district in 1999 where sawfly densities have persisted and fluctuated since then, showing a chronic nature.

1.3. Assessment of defoliation

The intensity of pine sawfly outbreaks can be assessed by inventorying canopy defoliation, since, defoliation is an inevitable consequence of the larvae consuming needles. Forest defoliation has been traditionally inventoried via ocular assessment. The

defoliation level is often characterized by visually estimating the percentage of needles missing in relation to a reference tree. The reference tree is either a healthy tree near the site, a photograph of a healthy tree from the same area growing on the same site type and canopy level, or an imaginary healthy tree (United Nations Economic Commission for Europe 2006). The ocular method is a subjective method of inventory. The error can reach up to 25-75 % depending of the experience of the surveyor (Belanger and Anderson 1988). One less subjective method of measuring tree defoliation is terrestrial laser scanning (e.g. Solberg et al. 2006, Holopainen and Hyyppä 2010).

An important aspect of inventorying is the selection of the optimal inventory method. The method of inventory has even greater impact on the results when inventorying rare or clustered environmental phenomena. In conventional forest inventory methods it is difficult to obtain a representative sample from a phenomenon that is clustered and rare, especially when using Simple Random Sampling (SRS). However, sampling methods have been developed for estimating rare or clustered phenomena, such as defoliation by pest insects.

Adaptive Cluster Sampling (ACS) is a two-phased sampling method designed for sampling phenomena that are rare and clustered. In ACS, an initial sample is selected using a standard probability sampling method (eg. simple random, systematic, transect survey) from the area. After the initial sample, more sample plots are added to areas where the phenomenon is encountered. ACS produces a large amount of data about the phenomenon of interest (Thompson 1990). For example Talvitie et al. (2006) employed the ACS method successfully to assess drought damage in Helsinki, Finland. Also, Talvitie et al. (2011) used ACS successfully in the inventorying of forest damages by *D. pini* and they showed that ACS is an effective inventorying method for estimating the range and total amount of a rare and clustered phenomenon. ACS is suited for studies that require large amounts of data about the phenomenon of interest. If information about the extent of the phenomenon is desired then, for instance, transect survey, which is commonly applied in forest entomological studies, can be employed (see e.g. Lyytikäinen-Saarenmaa et al. 1999).

A predetermined sampling design was not used in this study. The sample plots were assigned in 2002 in a manner that ensured that data would be obtained from all damage classes, ranging from severely defoliated to healthy forests. Furthermore, the sample plots were assigned for monitoring health and tree growth responses only, not for remote sensing purposes, which can cause some problems. These sampling related problems will be discussed in subsequent sections of the study.

1.4. Remote sensing in insect damage recognition

Remote sensing is defined as the process of measuring objects without being in direct contact with the objects themselves. In Earth observation this process typically involves recording electromagnetic radiation which is reflected off or emitted from the surface of the earth (Campbell 2007). Remote sensing systems can be either active or passive. Passive remote sensing sensors detect and record solar energy that is either reflected from the surface of the earth (e.g. aerial photographs and optical satellite sensors), or solar energy that is absorbed by, and re-emitted from, the surface of the earth (e.g., thermal sensors). Active remote sensing systems also record reflected electromagnetic energy; however, in this case the sensor itself is the source of energy, rather than the sun. LiDAR and radar are the most common active remote sensing systems (Campbell 2007). Some of the most commonly used remote sensing systems employed in environmental research include aerial photographs, optical satellite images, radar, LiDAR, and hyperspectral imagery.

In conventional forestry operations, remote sensing is most commonly used to assist with forest resource assessment and inventory. Other typical forestry related applications of remote sensing include forest disturbance detection and assessment (Williams and Nelson 1986, Mattila 1998, Ranson et al. 2003, Gimeno et al. 2004, Holopainen et al. 2006) and mapping forest defoliation caused by insect outbreaks (Falkenström and Ekstrand 2002, Fraser and Latifovic 2005, Thomas et al. 2007, Coops et al. 2009b). The detection of disturbances via remote sensing is typically characterized by measuring changes in

vegetation surface reflectance that are caused by the disturbance of interest. For example, during an outbreak of *D. pini* there is a drastic change in the reflectance properties of a forest canopy due to defoliation. Such changes can be easily detected and monitored via the synoptic data provided by remote sensing instruments.

Medium spatial resolution, multi-spectral imagery is perhaps the most common type of remotely sensed data used in the assessment of forest disturbance. For example, optical satellite imagery has been widely used in the boreal zone to assess forest disturbances such as insect outbreaks and fire (Dottavio and Williams 1983, Williams and Nelson 1986, Ranson et al. 2003, Wulder et al. 2006, Coops et al. 2009b, Ilvesniemi 2009). Vegetation indices such as the Ratio Vegetation Index (RVI), Normalized Difference Vegetation Index (NDVI), and Normalized Difference Moisture Index (NDMI) have been successfully used in several studies to detect and monitor disturbances (Dottavio and Williams 1983, Falkenström and Ekstrand 2002, Goodwin et al. 2008). For example, Nelson (1981) used RVI from Landsat Multi-Spectral Scanner (MSS) images to classify defoliation at site in Pennsylvania, USA and achieved classification accuracies of 88.5%, 64.9% and 56.6% using three different defoliation classes, heavy defoliation, moderate defoliation, and healthy forest, respectively. In a separate study, Ranson et al. (2003) assessed Siberian silkmoth (*Dendrolimus sibiricus* Tschetverikov) defoliation with Landsat Enhanced Thematic Mapper (ETM) data and achieved classification accuracies of 96.2% and 95.6% for two classes (severe and moderately defoliated forests). Goodwin et al. (2008) used multi-temporal Landsat data to classify forests with mountain pine beetle (*Dendroctonus ponderosae* Hopkins) damages in Western Canada. They used a decision tree classification and achieved an overall classification accuracy of 86% for separating two defoliation classes. Ilvesniemi (2009) used two different methods to classify (maximum likelihood classification and k-means clustering) for two defoliation classes in a Scots pine forest using Landsat TM images. The maximum likelihood classification achieved an accuracy of 85.9%, while accuracy of the k-means clustering classification was 88.7%, with kappa values of 0.71 and 0.77, respectively.

Aerial photography has also proved to be an effective remote sensing technology in the assessment of insect related defoliation (Haara and Nevalainen 2002, Fukuda and Pearson 2006, Ilvesniemi 2009). The benefit in using aerial images is the precise spatial resolution. They even enable the single tree estimation. The usual pixel size is 0.25–1 m. Aerial images are widely used in forest inventory and planning because of their availability and reasonable price. Cloud cover is a problem in both aerial photography and optical satellite imagery, often preventing the use of such data in monthly based monitoring programs (Karjalainen et al. 2010). For example, Haara and Nevalainen (2002) employed aerial photography to classify, via a semiautomatic pattern recognition technique, inventory plots into three defoliation classes. They achieved an overall accuracy of 89.5%. In another study, Fukuda and Pearson (2006) classified mountain pine beetle damage from aerial photographs via a decision tree classifier produced with the data mining package, WEKA (Waikato Environment for Knowledge Analysis). The classification accuracies for three defoliation classes were 75% and 31–49% using cross validation and smaller training sets, respectively. Ilvesniemi (2009) classified two defoliation classes in Scots pine from aerial photographs using three classification different methods; Maximum likelihood classification, K-means clustering, and linear regression models attained classification accuracies of 87.3%, 88.7% and 87.3%, respectively. Kappa-values attained via the same methods were 0.73, 0.76 and 0.74, respectively.

Airborne laser scanning (ALS i.e. LiDAR) is becoming one of the preferred technologies for determining terrain elevation and estimating forest inventory parameters such as stand-level height and volume, individual tree heights and species, forest growth (via change detection), and, forest damage (Solberg et al. 2006, Hyyppä et al. 2008, Hyyppä et al. 2009, Holopainen et al. 2010). Research in Finland has demonstrated that ALS based forest inventories can produce more accurate results than in the conventional Finnish standwise inventory used in forest planning (which is on average 15-30% for volume per hectare) (Næsset et al. 2004). The results of ALS based forest inventories are so accurate that, since 2010, ALS has been started to use operationally in Finnish forest planning to estimate stand-level forest characteristics. In terms of insect outbreak

detection, ALS has been successfully employed to detect defoliation by monitoring changes in Leaf Area Index (LAI) (Solberg et al. 2006, Solberg 2008, Solberg et al. 2009, Coops et al. 2009a). Puolakka (2010) studied the plot level estimation of the Scots pine defoliation with area-based ALS data. The best classification accuracy with two classes was 74.0% (kappa value 0.47) and 60.0% (kappa value 0.37) with four defoliation classes. Kantola et al. (2010) predicted the needle losses of individual Scots pines by *D. pini* using high density ALS data combined with aerial images. Classification accuracy in test data with the RandomForest method was 88.1% (kappa value 0.76) with two classes. The downside in ALS is, as in aerial imagery, that it is fairly expensive to use in large-scale applications, ultimately limiting its use in intensive, large-area monitoring tasks. Other promising technologies for insect related forest damage recognition include hyperspectral imagery (Kankaanhuhta et al. 2000) and radar data (Pulliainen et al. 1992, Ranson et al. 2003, Karjalainen et al. 2010).

1.5. Radar images in forestry

Radar is an active remote sensing sensor that operates in the microwave region of the electromagnetic spectrum. Typical radar wavelengths range approximately between 1 cm and 1 m. Radar systems can be set to transmit and receive either vertically (V) or horizontally (H) polarized energy. A radar system can, in principle, produce four images; two like-polarized and two cross-polarized images. Like-polarized images are transmitted and received in the same polarization direction (e.g., HH & VV), whereas, cross-polarized is transmitted and received in different polarization directions (e.g., HV & VH) (Campbell 2007, Lillesand et al. 2007). There are two types of imaging radars, real aperture radars (RAR) and synthetic aperture radars (SAR), the latter being typically used in the modern day systems. The resolution of a radar image is divided into two different parts, the range and the azimuth resolution. The range resolution is the resolution perpendicular to the flight direction and it is dependent on the length of the radar pulse. The azimuth resolution, on the other hand, is the resolution parallel to the flight direction

and is dependent on the length of the antenna or the wavelength. With a given wavelength the azimuth resolution of the image can be controlled with changing the length of the antenna. Therefore, in RARs the physical length of the antenna is changed. The 'synthetic aperture' in SAR means that the azimuth resolution of the radar does not depend on the physical length of the antenna. When the SAR instrument is moving along the flight line, the effect of a much longer antenna is mathematically synthesized from several different positions of the shorter, actual antenna using the Doppler effect (Lillesand et al. 2007). New features in radar technology that have improved the quality of SAR imagery include SAR instruments with increased spatial accuracy as well as interferometric radars and polarimetry. The spatial resolution of the new SAR systems is typically ~1-3 m. In full polarimetric mode, the spatial resolution is slightly less (> 3 m). (Campbell 2007, Lillesand et al. 2007)

In the Nordic countries radar imagery has not been used as commonly as a source of remote sensing material in conventional forestry applications as compared to traditional optical remotely sensed imagery (i.e. aerial photographs or Landsat imagery). The difficulty of interpreting radar data is perhaps the main reason that it hasn't yet been used in forest resource assessment and, to date, satellite-based radar imagery doesn't have the requisite spatially accuracy for forest resource inventory applications (Holopainen et.al. 2009b). However, the aforementioned issue is changing as new satellites with spatially high resolution radar instrument are becoming available. Radar has several advantages compared to traditional optical imagery. For example, since long wave microwave energy can penetrate clouds, radar imagery can be acquired in almost any weather conditions. The fact that clouds don't reduce the quality of radar images is a major advantage when monitoring forest condition in cloud prone areas. Furthermore, radar images can also measure three-dimensional characteristics of forest canopies, which is difficult to achieve with conventional optical satellite imagery which are not overly sensitive to three-dimensional forest properties.

Forest inventory surveys using SAR imagery have typically concentrated on generating large area inventories (Rauste 1990) or the estimation of forest biomass (Le Toan et al.

1992, Rauste et al. 1994, Rauste 2005). However, SAR imagery has also, to some extent, been employed to detect forest disturbances (Rauste 1996, Ranson et al. 2003). In recent years the use of high spatial resolution SAR in forest inventory and assessment has increased (Holopainen et al. 2009a, Holopainen et al. 2010), and has been an accurate means of remote sensing based forest inventory. For example, Hyyppä et al. (2000) compared SAR-derived inventory with inventory data derived from Landsat TM imagery. The authors concluded that optical satellite imagery had still more explanatory power as compared to satellite SAR data; thus, it was considered more useful in forest inventory applications. Hyyppä et.al (2000) also stated that airborne profiling radar attained better accuracies than aerial photography, and was the only remote sensing method which made it possible to get the same level of accuracies as conventional Finnish standwise forest inventory, which is on average 25% root mean square error (RMSE) for volume. However, because the profiling radar has a very limited field of view, it is considered unpractical to implement in operational forest inventories (Holopainen et al. 2009b).

Le Toan et al. (1992) studied the use of SAR data to biomass estimation in a homogenous maritime pine (*Pinus pinaster* Ait.) plantation. They demonstrated that L- and P-bands (frequency = 1.225 and 0.44 GHz, respectively) were strongly correlated with biomass, while shorter frequencies (X- and C-bands; frequency = 9.6 and 5.3 GHz, respectively) were weakly correlated with biomass. This was said to be due to the fact that shorter wavelengths are primarily scattered by smaller canopy elements such as branches and needles. In a separate study, Manninen et al. (2005) employed ENVISAT ASAR images to estimate LAI in boreal forests. The authors compared the SAR derived LAI accuracies to Systeme Pour l'Observation de la Terre (SPOT) satellites High-Resolution Visible and Infrared (imaging instrument) (HRVIR) 1 derived ones. The mean estimation error using one SAR image was 0.28, whereas, SPOT Normalized Difference Vegetation Index (NDVI) derived was 0.39 and SPOT Reduced Simple Ratio (RSR) 0.32. Manninen et al. (2005) concluded that radar imagery is a good alternative to optical images in retrieving LAI. SAR has also been studied in the estimation of plot-level forest attributes. Holopainen et al. (2009a) employed TerraSAR-X imagery and kNN-method to estimate

forest stand attributes. For mean volume, basal area, mean height and mean diameter the authors obtained RMSEs of 47.4%, 39.3%, 20.3% and 22.4%, respectively.

Radar images have not been widely used in disturbance recognition, but some examples are found in the literature. Radar imagery has been used to assess and monitor disturbances such as insect outbreaks, fires, and floods, among other natural disturbances. For example, Gimeno et al. (2004) mapped burned areas in a Mediterranean forest environment using a time-series of European Remote Sensing-2 (ERS-2) SAR data and attained overall accuracies of 92.11% (via a neural network classifier) and 89.9% (via principal component analysis (PCA)). In another study, Pulliainen et al. (1992) attempted to determine the optimal frequencies and polarizations for detecting forest defoliation. To achieve this, they simulated natural defoliation by artificially removing needles from Norway spruces (Picea abies) and measured the trees with three different instruments with varying frequencies (5, 10, and 35 GHz). Helicopter-borne measurements were conducted using a 5 GHz instrument. Their findings demonstrated that 10 GHz was the optimal frequency for detecting defoliation. However, the helicopter-borne instrument with 5 GHz also produced reasonable results. In a separate study, Ranson et al. (2003) compared radar instruments and Landsat 7 for disturbance detection. They combined Japanese Earth Resource Satellite (JERS) and Radarsat SAR imagery to classify moderate and severe insect damage caused by Siberian silkmoth. They concluded that the combination was not useful for separating insect damage stands from undisturbed forest; classification accuracies were 29% and 46% for severe and moderate insect damage classes, respectively. Furthermore, the insect damage classes were misclassified mainly as deciduous forest. Karjalainen et al. (2010) employed multi-temporal ERS-2 SAR images to map two forest damage classes (> and < 20% defoliation) in Finnish pine forests. Using a 3-nearest neighbor classification method, they achieved accuracies of 75.2% via leave-one-out cross validation.

1.6. Research questions and objectives

Given the previously demonstrated potential of SAR imagery for the assessment of insect defoliation, the primary objective of this master's thesis was to study the efficacy of detecting forest damage caused by the Common pine sawfly, *Diprion pini*. Specifically, the accuracy of detecting defoliation was assessed via a change detection technique that employs multi-temporal ERS-2 and ENVISAT SAR imagery. In this study the linear discriminant analysis (LDA), k-means clustering and logistic regression were used as methods to estimate the plot level needle loss level. The following five research questions were addressed:

1) What level of accuracy can be achieved when classifying *Diprion pini* defoliation via multi-temporal ERS-2 and ENVISAT SAR imagery?

2) Is it possible to separate *Diprion pini* defoliation from forest clear cuts and thinnings?

3) How does simple speckle noise filtering of the SAR imagery influence classification accuracy?

4) What are the best SAR features to characterize defoliation in Scots pine forests?

5) Do different classification methods produce variation in classification accuracies?

2. METHODS

2.1. Study area

The study site is located in Palokangas area (62°52' N, 30°53' E), in Eastern Finland in Ilomantsi district (Figure 2.1). The study site is primarily comprised of even-aged Scots pine forest on relatively dry soils. The majority of the forests in the area are young or middle-aged. The sampling plots chosen in 2002 were located in mature and maturing stands for monitoring the recovery and possible mortality of trees and controlling factors of the sawfly population in the outbreak area the study area is owned by a forest products company, Tornator Ltd.
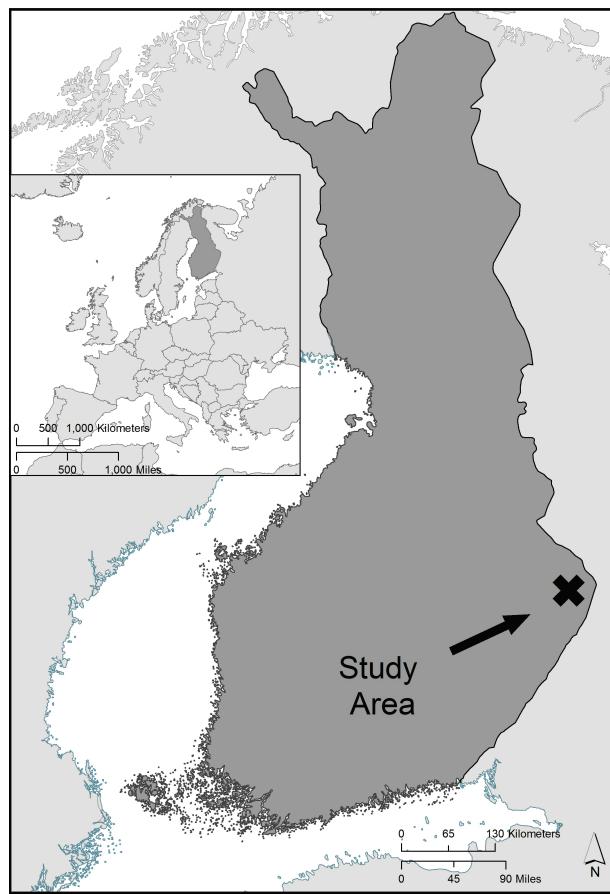
Figure 2.1: The location of the study area, Palokangas. Original data © Maanmittauslaitos (National Land Survey of Finland) 2010 and © Centro Internacional de la Papa -GIS database 2010.
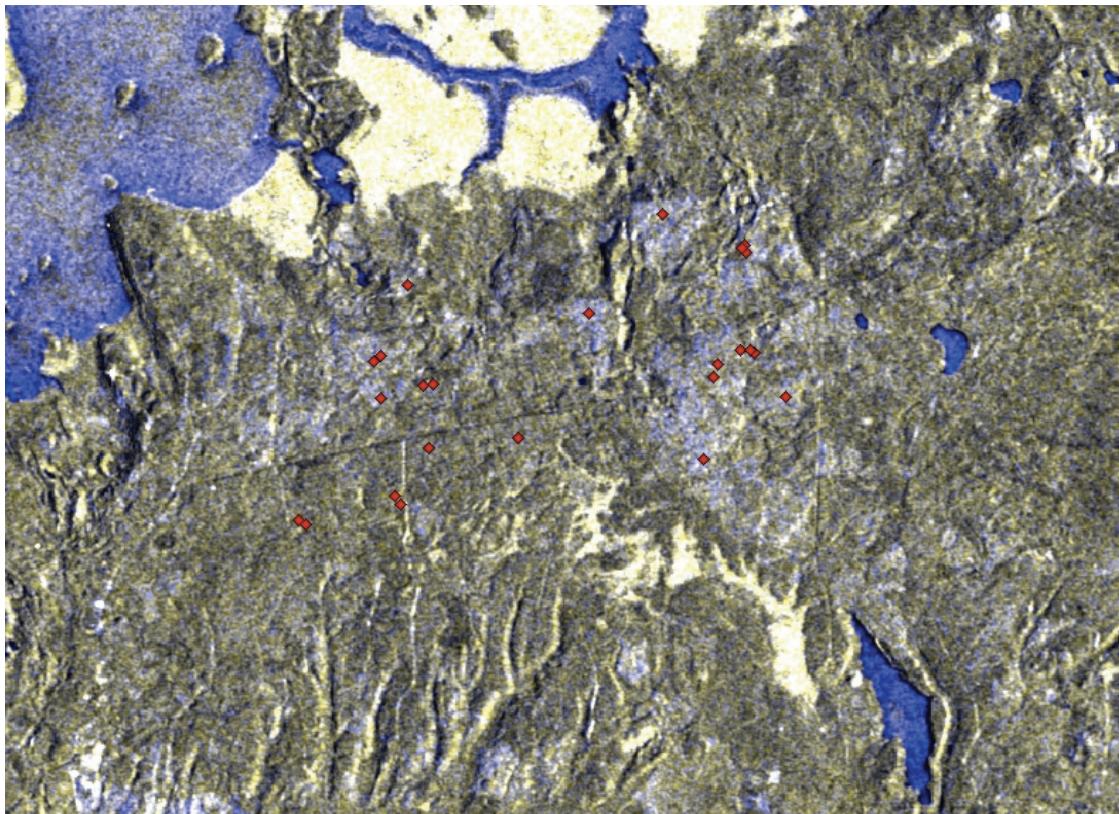
Figure 2.2: Composition image of all SAR images in Palokangas area and the sample plots. Red band in image: Average amplitude, Green band: Average amplitude and Blue band: Standard deviation of amplitude. If area has high standard deviation, i.e. blue color, there has most likely happened some kind of change during the years of the study. Coordinate system ETRS-TM35FIN/WGS84. Original Data © ESA 2009.

2.2. Field data

The field measurements were conducted at 16 permanent sample plots in 2002. The tree and plot-wise characteristics were measured. The defoliation level for each individual tree was also visually estimated. The plot centres were precisely located in the field with a Trimble Pro XH (Trimble Navigation Ltd., Sunnyvale, CA, USA), which can reach up to 30 cm precision. The follow-up estimation of the defoliation levels were conducted in years 2004-09 (Table 2.1). Measurements were carried out every year at approximately the same time between May and June before shoot elongation and needle consumption of the year in question. Five of the sample plots were harvested during the monitoring

period; thus, only 11 of the sample plots have measurements across the entire duration of the study. The data consists of 90 observations from the 16 sample plots during 8 years period. When the harvested class was added to the classification all the sample plots that had been clear-cut or thinned were added to the classification. Also, an additional eight sample plots (sample plots 17-24, Table 2.1) were selected to assure samples were evenly spread between no defoliation, defoliation, and felling classes.

The sampling plots were circular plots having varying radius. The Finnish Forest Research Institute's (Metla) forest fieldwork guide (Vuokila 1987) was used for developing the sampling protocol. The sample plots were selected so that minor, moderate, and severe defoliation classes were evenly represented. The sample plots were located in clusters of two or three (figure 2.2), and the radius of each plot was selected to ensure that every plot had a minimum of 20 trees. However, in one plot (plot 10) this minimum requirement was not fulfilled. The accuracy in the measurement of the radius was 0.5 m and, therefore, several plots had more than 20 trees (Table 2.1). Plot-level defoliation was calculated by summarizing the individual tree data within each plot. Forest planning information (provided by Tornator Ltd) was used to determine the location and timing of harvesting across the study area. The forest planning information contained information of the time of cutting, and whether a compartment was harvested, however, the type of cutting (thinning or clear-cut) was not included. Thus, the felled classes include both clear-cut and thinned forests.

Figure 2.3 shows the distribution of defoliation in individual sample plots in the years 2002-2009. Major portions of the trees and sample plots had defoliation less than 10%, but it should be noted that, for example, in 2002 only 1 out of 16 sample plots had 10% or less of defoliation (Table 2.1). In later years, when the outbreak had subsided, the number of plots with minor defoliation (defoliation < 10%) increased. As noted in Table 2.1, the defoliation was highest in either 2002 or 2005, depending on the plot. After the first gradation in 1999-2002, the second gradation was launched in 2004-2007. Variation in defoliation between sample plots is considerable; while some plots completely

recovered from the outbreak others did not and continue to exhibit a large amount of defoliation and tree mortality.

Table 2.1
Showing the sample plot characteristics from the year 2002 and the defoliation measurements for the years 2002 and 2004 to 2009. F= Clear cut or thinned sample plot.

Plot	Nbr of trees	Diameter of plot, m	Avg dbh[a], cm	Avg height, m	Average defoliation in years 2002, 2004-2009, %						
					02	04	05	06	07	08	09
1	21	13	23.6	18.6	26	22	9	F	F	F	F
2	24	13	22.0	18.2	30	25	16	F	F	F	F
3	21	12	22.3	19.4	51	40	24	20	12	10	10
4	20	12	21.9	17.6	40	23	14	8	8	7	6
5	21	10	18.5	15.5	24	16	F	F	F	F	F
6	21	10	15.5	12.4	26	14	F	F	F	F	F
7	22	9	18.9	17.6	28	20	7	5	6	4	4
8	21	9	17.4	16.3	24	16	10	7	8	6	4
9	21	12	17.7	15.4	45	49	99	97	92	84	80
10	18	11	19.5	16.5	49	43	67	50	44	36	32
11	26	12	20.5	16.9	58	46	68	59	52	44	40
12	20	11	20.0	17.8	59	35	76	F	F	F	F
13	20	11	19.6	17.8	13	5	4	3	4	4	3
14	24	10	18.9	18.0	15	8	8	4	4	4	3
15	24	8,5	15.1	18.3	10	5	1	1	2	3	2
16	20	12	21.1	17.6	63	62	97	95	86	80	82
AVG	21.5	11	195.3	17.1	35	27	36	32	29	25	24
17								F	F	F	F
18							F	F	F	F	F
19								F	F	F	F
20						F	F	F	F	F	F
21						F	F	F	F	F	F
22									F	F	F
23										F	F
24								F	F	F	F

[a] dbh stands for tree diameter at breast height
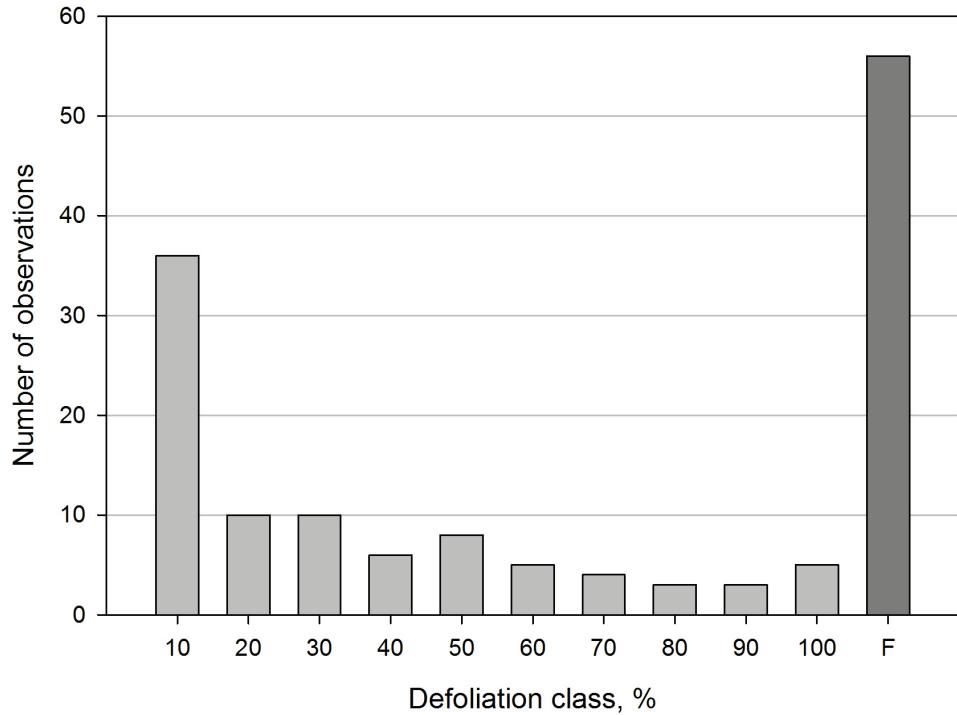
Figure 2.3: The distribution of total number of observations from the years 2002-2009. 'F' stands for felled sample plot.

The defoliation level was assessed in the field using an adapted measurement protocol developed and described by Eichhorn (1998). In this method, tree-level defoliation is quantified via a visual assessment. This is achieved by comparing a defoliated tree to non-defoliated reference tree. Defoliation is expressed as a relative needle loss (%) between the sample and reference tree. The reference tree is an imaginary tree with healthy and full foliage on a similar forest site type and prevailing conditions. Defoliation was partitioned into classes with 10% intervals (0, 10, ..., 90, 100%), with the exception of the 0.01% defoliation class which indicates minor defoliation (i.e. defoliation was present, but barely visible). Thus, in total there were 12 defoliation classes. Plot level defoliation was calculated by averaging the tree-level defoliation estimates. Mean defoliation from each year was also calculated as the mean of all trees.

Table 2.2

The class combinations used in supervised and unsupervised classifications and the distribution of observation to different classes. Class combinations 3 and 4 have in addition to defoliation plots also the field plots which have been felled during the sampling years.

Class combination 1 (D2)		Class combination 2 (D3)		Class combination 3 (DF3)		Class combination 4 (DF4)	
2 classes		3 classes		3 classes		4 classes	
Defol., %	N sample plots	Defol., %	N sample plots	Defol., %	N sample plots	Defol., %	N sample plots
<20	47	<20	47	<20	47	<20	47
≥20	56	20-50	29	≥20	56	20-50	29
		≥50	27	Felled	62	≥50	27
						Felled	62

Four different class combinations were created for this study (Table 2.2). These combinations were used in supervised and unsupervised classifications. The first two classifications had only defoliation classes and the two latter classifications had also the 'felled' class included. The threshold used to identify moderate defoliation existing in the area was 20%, meaning that 20% or more of the needles on that sample plot have been defoliated. The threshold for severe defoliation was 50%, meaning that at least half of the needles have been defoliated. The classifications are not portraying just the situation of one year only but they are done for all the observations, from different years, at the same time.

The field measurements were originally set up to study the recovery process and mortality of the stands and trees in the outbreak area. Due to the prolonged outbreak in the area, the sample plots have been also utilized in other ways than just in the follow-up study. Since the sample plots were not created specifically for this study, sampling system used may not be the best possible one to use for estimating defoliation via remote sensing.

2.3. Remote sensing data

Imagery for this study was acquired from two different European Space Agency (ESA) satellites with a SAR sensor on board; ERS-2 and ENVISAT (Table 2.3). Images were acquired on an annual basis in the months of September or October. Since the radar imagery is acquired during the fall, well after *D. pini* has ended its larval phase, it is assumed that the defoliation conditions are consistent between fall and the following spring. Therefore, an image from fall 2001 should correspond to the field data collected during the spring 2002, while an image from 2002 corresponds to the field data collected during the spring 2003, and so on. The radar imagery archive used for this study originally consisted of 11 SAR images taken between the years 1999-2008, not including years 2004 and 2007. Because the field data were acquired in 2002 and each year between 2004-2009, the images employed were acquired in 2001 and on an annual basis between 2003-2008 (Table 2.4).

Table 2.3

Parameters of the ERS-2 and ENVISAT satellites and their SAR instruments. The satellite sensors used in this study have similar characteristics and, therefore, images from two different sensors could be used together without introducing additional error.

Satellite	ERS-2	ENVISAT
Pass	Desc.	Desc.
Track	50	50
Frame	2331	2331
Polarization	VV	VV
Incidence angle	23	23
Spatial resolution on ground, m	15	15

Table 2.4

Acquisition dates of images used in this study and the corresponding satellite that took the images.

Image number	Satellite	Image date
3	ERS-2	19 Oct 2001
5	Envisat	19 Sep 2003
6	ERS-2	19 Sep 2003
7	Envisat	19 Aug 2005
8	ERS-2	13 Oct 2006
9	ERS-2	12 Sep 2008
10	ERS-2	17 Oct 2008
11	ERS-2	21 Nov 2008

The data was received from ESA and was preprocessed by the Finnish Geodetic Institute (FGI). All images were geo-registered to the first image acquired, and geo-referenced to ETRS-TM35FIN projection with a ground pixel size of 15 meters. The pixel values of the original images quantified the backscatter amplitude of the radar signal. Pixel by pixel average amplitude and standard deviation of the backscatter amplitudes were calculated from the original backscatter amplitudes. In addition, two difference images were created from the amplitude images. First, the other images were subtracted from the 2002 amplitude images and, second, the other images were subtracted from the last 2008 image to create two sets of amplitude difference images. Information about the soil or canopy moisture was not known from the image acquisitions dates. Because moisture has an effect on the backscatter recorded by the SAR instrument, not having in situ soil and canopy moisture measurements may impact the accuracy of the results (Karjalainen et al. 2010). Additional processing was conducted to reduce the impact of speckle noise. This was achieved by filtering the radar images with 3 x 3 and 5 x 5 speckle filters. Defoliation classifications derived from each speckle filter (3 x 3 and 5 x 5) were compared to classification derived from the unfiltered radar images. Speckle filtering assigns every pixel the mean value of all pixels within the window (3 x 3 or 5 x 5), which

degrades the spatial resolution of the images (Figures 2.4, 2.5 and 2.6). Thus, the purpose of this step is to determine the impacts different filtering levels on the accuracy of the defoliation classification.

2.4. Testing individual image features

The order of superiority of the SAR image features (original amplitude, average and standard deviation of the amplitude and two difference images of amplitude) was tested with linear regression. Each image feature was used individually to estimate the defoliation. The regression was done for each image window separately. For determining the superiority coefficient of determination, R^2, and correlation are calculated.

In linear regression the response variable is modeled with a linear line, by using linear regression model

$$y = a + \beta x \qquad (2.1)$$

Where y is response variable, a is intercept, β is slope and x is the explanatory variable.

The unknown variables were estimated from the data using function lm in the statistical program R (The R development core team 2009). The variables were estimated by minimizing the sum of squared residuals. The performance of the model was estimated by calculating the coefficient of determination, R^2, adjusted R^2 and the correlations between each variable and the defoliation. Formula for R^2 is written as (Ranta et al. 2005):

$$R^2 = 1 - \frac{SS_{err}}{SS_{tot}} \qquad (2.2)$$

Where SS_{err} is the residual sum of squares and SS_{tot} is the total sum of squares.
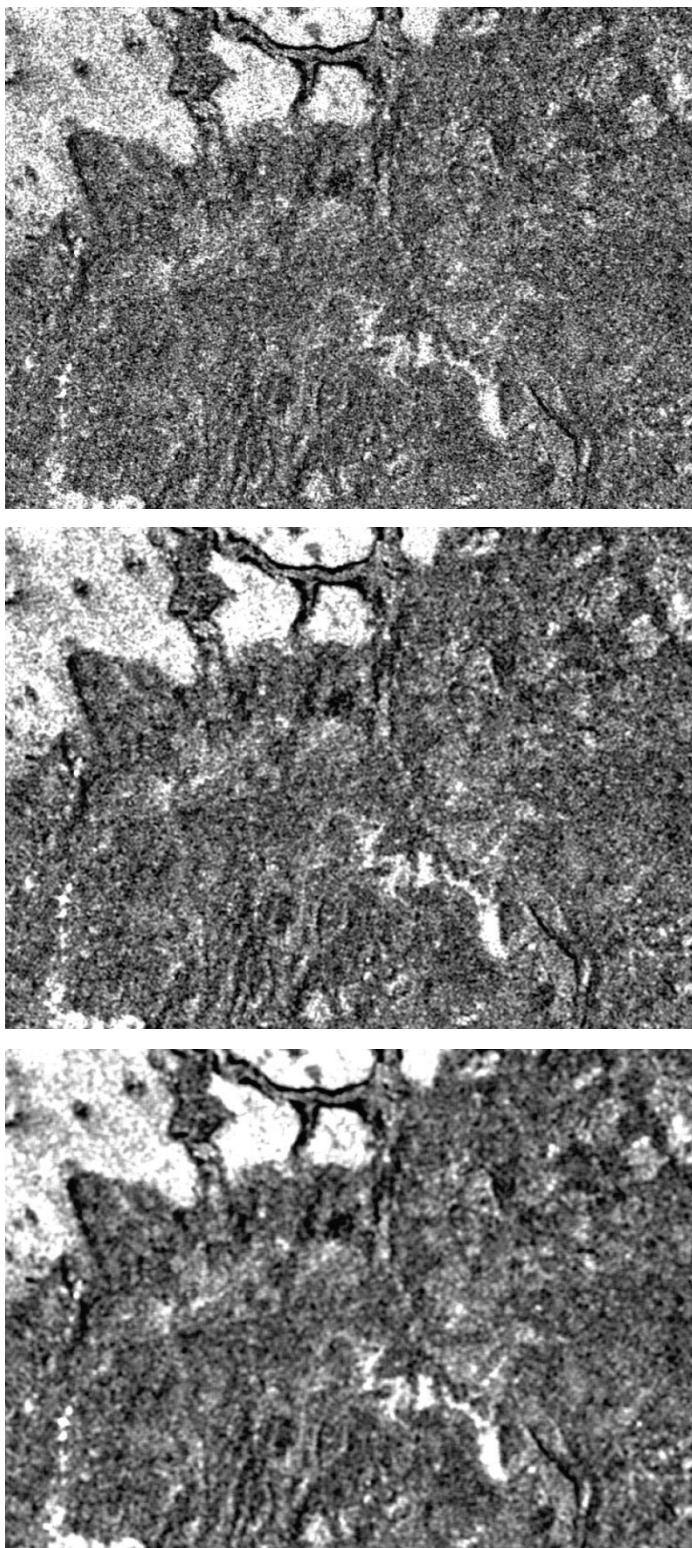
Figure 2.4: The effect of speckle filtering on SAR images. From up to bottom: the unfiltered image, 3 x 3 filtered image and 5 x 5 filtered image. In all figures original amplitude image from the year 2002 is presented.

2.5. Defoliation classification

2.5.1. Supervised classification

In supervised classification, training data was used to produce an empirical function, which was ultimately used to classify each image pixel into different categories. The classification function varies depending on the method that is used. In the current study, the field data is used as the training set and the linear discriminate analysis (LDA) is used to develop the classification.

LDA is a classification procedure that assigns each pixel to a class by determining the best linear combinations of predictor variables (e.g. remotely sensed data) for differentiating between each class in the entire suite of classes. This is achieved based on the following function:

$$g_i(x) > g_j(x), \; where \; j \neq i \tag{2.3}$$

where i is class 1,…, c and x is the feature values of the pixel. In other words, pixels are partitioned into a particular class when the linear discriminant function is maximized (Duda et al. 2001.)

The discriminant function is written as:

$$g_i(x) = P(\omega_i|x) \tag{2.4}$$

where P is the posterior probability of a pixel belonging to class i. Thus, the classifier assigns the pixel to a class where the class discriminant function gives maximum posterior probability. The posterior probabilities were calculated empirically from the class density function as characterized by the training data. In this study, the Maximum-likelihood procedure (MLE) is used to calculate the estimates for the mean and covariance matrix of the density function (Duda et al. 2001.) In this research, LDA was

implemented within the statistical program R via the lda function (The R development core team 2009). The R function produces discriminant functions for the different classes and uses cross validation to predict a class for each sample plot.

2.5.2. Unsupervised classification

In unsupervised classification, prior knowledge (in the form of training data) is not used to identify the classes. Instead, the classification algorithm employs statistical clustering to identify individual classes based upon image features. In this study the k-means clustering algorithm was utilized as an unsupervised classification method. K-means clustering is a common classification method, widely used in remote sensing classification applications. The k-means procedure finds homogenous clusters within image feature space in a manner that minimizes within-cluster variance and maximizes between-cluster variance (Hartigan 1975). The basic procedure is implemented as follows: Firstly, the algorithm selects initial cluster centers (for a user-defined number of clusters, k) from the feature space, so that the distance between the cluster centers is maximized. Secondly, all the observations (i.e., pixels) are assigned to the nearest cluster in terms of distance across the spectral feature space. The cluster means are recalculated, cluster centers are adjusted, and the pixels are reassigned to the nearest cluster. The procedure iterates until cluster centers do not change or when user-defined number of iterations has been done.

The R function kmeans was used to perform the k-means unsupervised classification (The R development core team, 2009). The number of clusters, k, was set to 10 after testing the performance of several different values. After the clustering procedure executes, each identified cluster was assigned a defoliation level by calculating the weighted mean of all observations in a cluster. The weight was based on the number of observations in each defoliation class in the cluster. The clusters were then assigned to the defoliation classes based on the mean values they have. For example, all the plots in a cluster were assigned the value 'felled' if a majority of the plots in that cluster have the

real (field observation) value 'felled'. The classifications are then compared to the field observations to evaluate the accuracy of the estimation.

2.5.3. Logistic regression

A logistic regression model is a linear regression model that predicts the probability of an event occurring. The model produces a distribution curve, the logit function. The mean of the distribution curve is the estimate for the variable explained. Logistic regression is a special case of generalized linear model (glm) (Hosmer and Lemeshow 2000). In glm the link function gives the relationship between the linear function of the explanatory parameters and the expected value of the dependent variable, i.e. the mean of the distribution function. The link function in the case of logistic regression is the logit transformation (2.5).

$$g(x) = ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p, \qquad (2.5)$$

where $g(x)$ is the logit of the multiple logistic regression model, $\pi(x)$ is the mean of the distribution function, β_0 is the intercept and $\beta_1, \beta_2, \dots, \beta_p$ are the coefficients for variables $x_1, x_2, \dots, x_p$ (Hosmer and Lemeshow 2000).

Thus, the logistic regression model, estimating the mean of the distribution function, $\pi(x)$, is achieved with the equation (2.6).

$$\pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}} \qquad (2.6)$$

Equation (2.6) produces a probability, a value between 0 and 1, of an event occurrence. In this case, a pixel is assigned to class 0 or 1 (Hosmer and Lemeshow 2000). In the current study the R function glm was used with a binomial logit link function to perform

logistic regression (The R development core team, 2009). The function produces coefficients that are used to predict the mean of the probability function for each sample plot. Two different estimations were made using the data. The first model estimated the presence and the absence of the defoliation. The defoliation level of 20% was used as a threshold value. The second model estimated the presence or the absence of fellings (i.e., harvests) in the area.

2.6. Accuracy assessment

Standard accuracy assessments were used to evaluate the quality of the various defoliation classifications. For the LDA classification, a leave-one-out cross-validation procedure was used. Leave-one-out cross-validation is one of the simplest and most widely used accuracy assessment methods used, and popular especially when data set is small. In leave-one-out cross-validation one data point at a time is left aside and the rest of the data is used in producing a model and an estimate for that point. Each data point gets and estimate the same way and those estimates can be then compared to the reference data (Diamantidis et al. 2000, Ranta et al. 2005). Leave-one-out cross-validation is known to produce slight overestimated results (Karjalainen et al. 2010). In k-means clustering the data was divided to the defoliation classes by means of feature values; therefore, cross-validation was not performed. In the case of logistic regression the same data was used for both producing the model and assessing the accuracy, because of the small size of the dataset. The standard error matrices were developed for each classification method. A standard error matrix displays correctly classified plots in the diagonal matrix elements and incorrectly classified values off of the diagonal. The overall classification accuracy was calculated from the matrices by dividing the amount of sample plots classified correctly by the total number of sample plots (2.7).

$$Percentage\ correct = 100 * \frac{sample\ plots\ on\ diagonal}{All\ sample\ plots}, \qquad (2.7)$$

where *sample plots on diagonal*= the count of sample plots on diagonal and *all sample plots*= the count of all sample plots.

Cohen's Kappa was also calculated for each classification. Cohen's Kappa was developed to account for the fact that randomly assigning pixels to different classes can often result in relatively high classification values (Campbell 2004). Cohen's kappa value (2.8) (Cohen 1960) adjusts predicted values for the amount of agreement that could be expected due to chance alone. Kappa values vary between -1 and +1. If kappa is between -1 and 0 it means that the classification is worse than classification arrived at by mere chance. If the kappa value is greater than 0 the model gives better results than what would be expected by mere chance (Table 2.5).

$$K = \frac{P_o - P_c}{1 - P_c},$$ (2.8)

where P_o = the amount of correctly classified observations and P_c = the amount of correctly classified observations expected by chance (Cohen 1960).

Table 2.5
Interpretation of the kappa statistic. If kappa statistic is larger than 0 classification was better than would be expected by mere chance. (Landis and Koch 1977).

Kappa statistic	Strength of agreement
<0.00	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost Perfect

3. RESULTS

3.1. Individual SAR features

The R^2 and correlations of the individual SAR features were tested with a linear regression individually (Table 3.1). In table 4.1, Difference 2002 and Difference 2008 refer to the difference values between the amplitudes from different years. The regression results indicate that the individual images feature explaining very little variation in defoliation levels. The best R^2 was less than 0.06 (The amplitude from 3 x 3 and 5 x 5 window sizes). The correlation column indicates the correlation between the feature in question and the defoliation level. Also the correlations are fairly weak between the individual image features and the defoliation, being at most 0.24 (the amplitude from 3 x 3 and 5 x 5 window sizes).

Table 3.1
Correlation and R^2 of linear regression model between individual SAR features of each filter window and defoliation.

1x1	R^2	Correlation
Amplitude	0.016	0.13
Difference 2002	0.011	0.1
Difference 2008	0.0001	-0.01
Average	0.051	0.23
Standard deviation	0.021	0.14
3x3	R^2	Correlation
Amplitude	0.058	0.24
Difference 2002	0.002	0.04
Difference 2008	0.004	-0.06
Average	0.051	0.23
Standard deviation	0.011	0.1
5x5	R^2	Correlation
Amplitude	0.058	0.24
Difference 2002	0.018	-0.13
Difference 2008	0.017	-0.13
Average	0.051	0.23
Standard deviation	0.048	0.22

Individual image features do not explain the defoliation well (Figure 3.1). When the defoliation level is less than 20% the amplitude is more likely to be lower than in the case of the more severe defoliation. When the defoliation level is high the amplitude has a wide variation between 250 and 500; there is not any degree of correlation. The relationship seems to be the similar between Difference 2002 and defoliation. For example, when the difference in amplitudes between two images is zero, the defoliation varies widely between 5 and 65%. The relationship between average and standard deviation to defoliation level seems similar to the relationship between amplitude and defoliation. When the defoliation level, the average and standard deviation image feature are also low. However, in terms of the high defoliation values the image feature values do not following a similar trend, ultimately make those less feasible for the prediction of defoliation level.

3.2. Supervised classification

The best classification accuracy (77.7%) was achieved via LDA using two defoliation classes with the 5 x 5 speckle filter (Table 4.2). In terms of the speckle filtering, the 5 x 5 filter produced higher accuracies than the other images filters. Adding a felled (harvested) class into the classification lowered the classification accuracy. In general, increasing the speckle filter window size had a greater impact on the accuracy when the felled class was included in the analysis. In most cases, increasing the speckle window filter size from 3 x 3 to 5 x 5 pixels increased the classification accuracy by approximately 10%.

The classification works fairly well using two classes, recognizing nearly 90% of the defoliated pixels in the best case (Table 3.2). When the 'felled' class is included, the accuracy of the '<20%' defoliation class remains stable; however the accuracy of the '≥20' and '≥50' defoliation classes are significantly reduced. When all the classes are included, the classes '≥20' and ≥50' have very low classification accuracies, while the other two defoliation classes ('<20%' and 'felled') have relatively high accuracies.
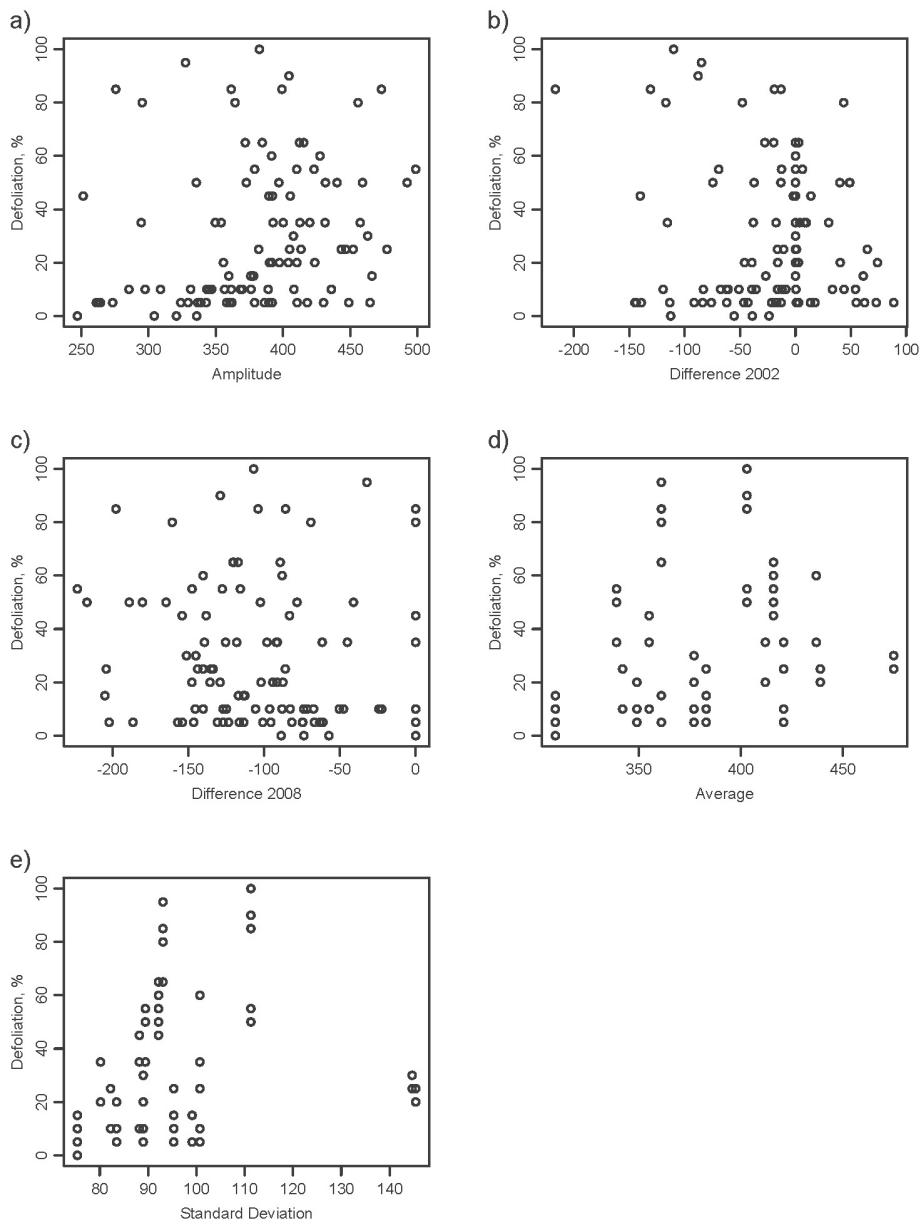
Figure 3.1. Individual 5 x 5 pixel window image features plotted against defoliation. Y-axis always defoliation and x-axis is values from feature: a) Amplitude, b) Difference 2002, c) Difference 2008, d) Average, e) Standard Deviation.

In most cases, the kappa values follow the same trend as the overall classification accuracies (Table 3.2). In some cases, however, the kappa value increased when more classes were added to the classification, while the overall classification accuracy decreases. This is particularly true when the 'felled' class is added; the kappa values increase. This implies that the chance of randomly correctly classifying defoliation decreases when the number of classes in the classification increases.

Table 3.2
Classification performance of LDA method. Classifications are made for each class combination (Table 2.2) and for 1 x 1, 3 x 3 and 5 x 5 speckle filter images. For each class combination classification accuracy for each class, total classification accuracy and kappa value are presented.

LDA		1x1	3x3	5x5
Class combination 1	<20	68.1	66.0	63.8
	>20	75.0	76.8	89.3
	Total	71.8	71.8	77.7
	Kappa	0.43	0.43	0.54
Class combination 2	<20	78.7	83.0	74.5
	≥20	51.7	48.3	55.2
	≥50	11.1	14.8	66.7
	Total	53.4	55.3	67.0
	Kappa	0.25	0.27	0.48
Class combination 3	<20	72.3	68.1	66.0
	≥20	30.4	55.4	70.4
	Felled	67.7	64.5	77.4
	Total	56.4	62.4	71.8
	Kappa	0.35	0.44	0.57
Class combination 4	<20	78.7	87.2	87.2
	≥20	0.0	0.0	10.3
	≥50	0.0	0.0	33.3
	Felled	75.8	71.0	77.4
	Total	50.9	51.5	61.2
	Kappa	0.27	0.29	0.44

3.3. Unsupervised classification

The best classification accuracy with the k-means method, 72.8%, came with two defoliation classes and with the original 1 x 1 pixel window size (Table 3.3). Using two and three defoliation classes the percentage of correctly classified pixels decreased when the pixel window size increased from 1 x 1 to 5 x 5 pixel size. The lowest classification accuracies were attained using the 3 x 3 pixel size window in both cases. When the 'felled' class was added the 5 x 5 pixel window produced the best classification accuracies with both two and three defoliation classes. However, the difference between window sizes was not large. The 1 x 1 pixel window produced better results with unsupervised than with supervised classification. The best classification accuracy with the 'felled' class included was 63.6%, with 2 defoliation classes and 5 x 5 pixel window.

Table 3.3
Classification performance of k-means method. Classifications are made for each class combination (Table 2.2) and for 1 x 1, 3 x 3 and 5 x 5 speckle filter images. For each class combination classification accuracy for each class, total classification accuracy and kappa are presented.

| k-means		1x1	3x3	5x5
Class combination 1	<20	83.0	59.6	68.1
	>20	64.3	69.6	67.9
	Total	72.8	65.0	68.0
	Kappa	0.46	0.29	0.36
Class combination 2	<20	83.0	59.6	68.1
	≥20	31.0	65.5	65.5
	≥50	59.3	0.0	18.5
	Total	62.1	45.6	54.4
	Kappa	0.39	0.15	0.28
Class combination 3	<20	46.8	19.1	46.8
	≥20	48.2	82.1	78.6
	Felled	66.1	59.7	62.9
	Total	54.5	55.8	63.6
	Kappa	0.31	0.32	0.45
Class combination 4	<20	70.2	55.3	46.8
	≥20	17.2	65.5	79.3
	≥50	11.1	0.0	0.0
	Felled	66.1	59.7	62.9
	Total	49.7	49.7	50.9
	Kappa	0.28	0.31	0.34
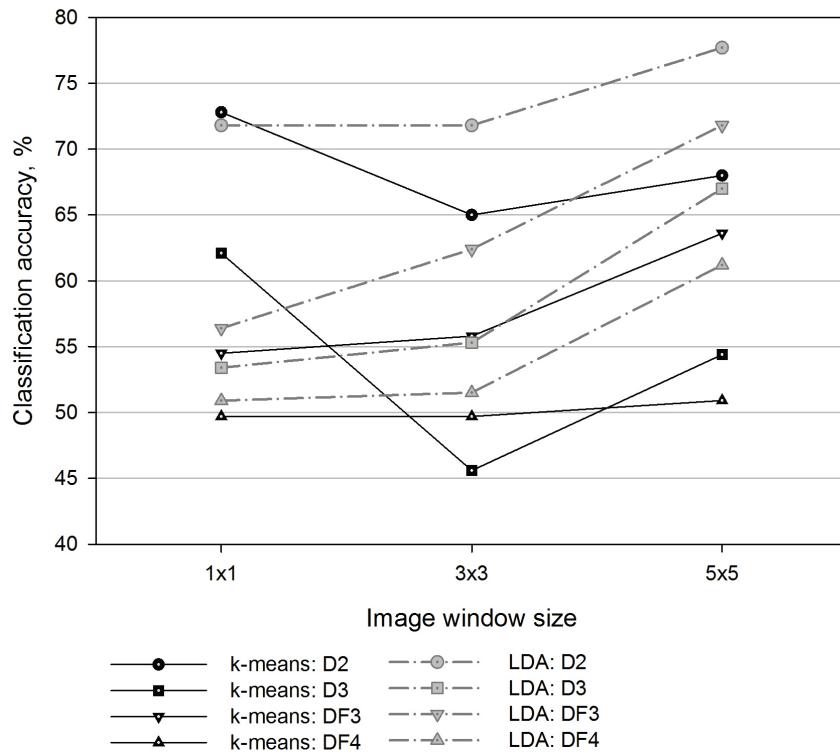
Figure 3.2: The accuracy of the estimation with K-means and LDA methods. Abbreviations stand for different class combination (Table 2.2)

When looking at the individual classes added to the classification, some differences can be found. Defoliation class '<20%' obtained the best classification accuracy usually with the 1 x 1 pixel window. However, when the defoliation class was ≥20% the best classification accuracy was received usually by using the 5 x 5 pixel window. Overall, the 1 x 1 pixel window attained the highest accuracies with the class combinations D2 and D3 and the 5 x 5 pixel window with class combinations DF3 and DF4 (Table 3.3). Similar to LDA, adding the 'felled' class to the k-means classification leads to increased kappa values. In general, the LDA algorithm attains higher classification accuracies as compared to the k-means classifier. The only exception was while using the 1 x 1 pixel window the k-means classifier attained a higher classification accuracy than the LDA algorithm (Figures 3.2 and 3.3). Neither LDA nor k-means attained a substantial kappa value (Table 2.1); all kappa values were under 0.60, regardless of the method.
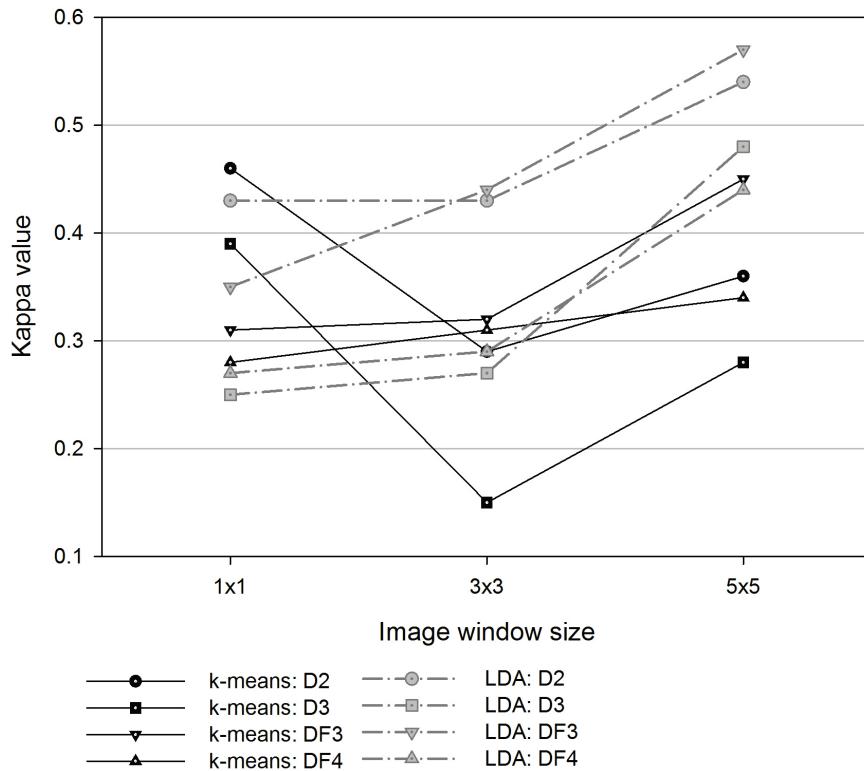
Figure 3.3: Kappa values of the estimation with K-means and Lda methods. Abbreviations stand for class combination (Table 2.2).

3.4. Logistic regression

Both classifications (defoliation and felling) achieved high classification accuracy with logistic regression (Table 3.4). The estimation accuracies were 81.6% and 84.2% for defoliation and felling, respectively. In both cases the highest classification accuracy was attained with the 5 x 5 window. The difference between them was the estimation accuracy of the different categories. In the case of defoliation, the logistic regression was more accurate when predicting the presence of defoliation, as compared to estimating the absence of defoliation. When felling is considered, the logistic regression was more accurate for estimation the absence of defoliation as compared the estimation of the absence of defoliation. The highest accuracy (92.9%) was attained when estimating the

presence of defoliation with 5 x 5 pixel window. The classification accuracy for the absence of felling was highest (91.3%) when the 3 x 3 pixel window was used.

The kappa values of the defoliation estimation tend to follow the same trends as the classification accuracies. Increasing the window size increased the accuracy of the felling class, while the defoliation category had nearly equal classification accuracies, regardless of window size. The estimation of defoliation with logistic regression and 5 x 5 pixel window and the estimation of felling with 3 x 3 or 5 x 5 pixel windows attained substantially high kappa values (Table 3.1). The highest kappa values were for defoliation and felling; 0.62 and 0.66, respectively. Figures 4.4 and 4.5 compare the difference between all the methods when estimating only two defoliation classes. The figures demonstrate that the accuracy of the logistic regression for estimating harvesting (i.e., felling) improves with an increasing window size.

Table 3.4

Classification accuracies and kappa values of the logistic regression method. Classification was done for two different phenomena, defoliation and felling. In the table there are classification accuracies listed for individual classes and total classification accuracy and kappa value.

| Logistic regression		1x1	3x3	5x5
Defoliation ≥20%	No	72.3	76.6	68.1
	Yes	83.9	82.1	92.9
	Total	78.6	79.6	81.6
	Kappa	0.57	0.59	0.62
Felling	No	83.5	91.3	88.3
	Yes	61.3	66.1	77.4
	Total	75.2	81.8	84.2
	Kappa	0.46	0.60	0.66
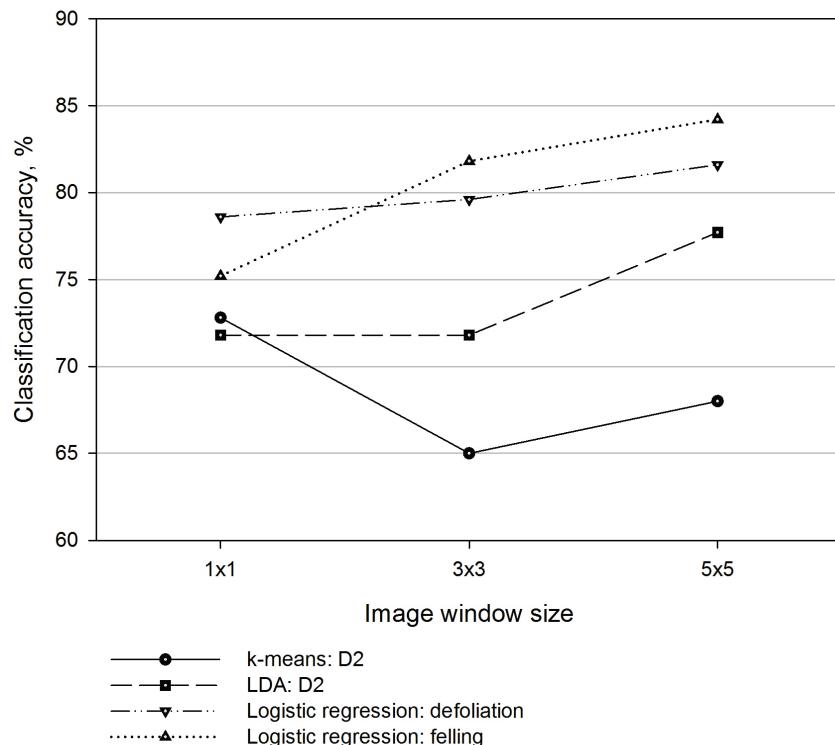
Figure 3.4: The accuracy of the estimation with different methods with 2 classes (i.e. class combination D2).
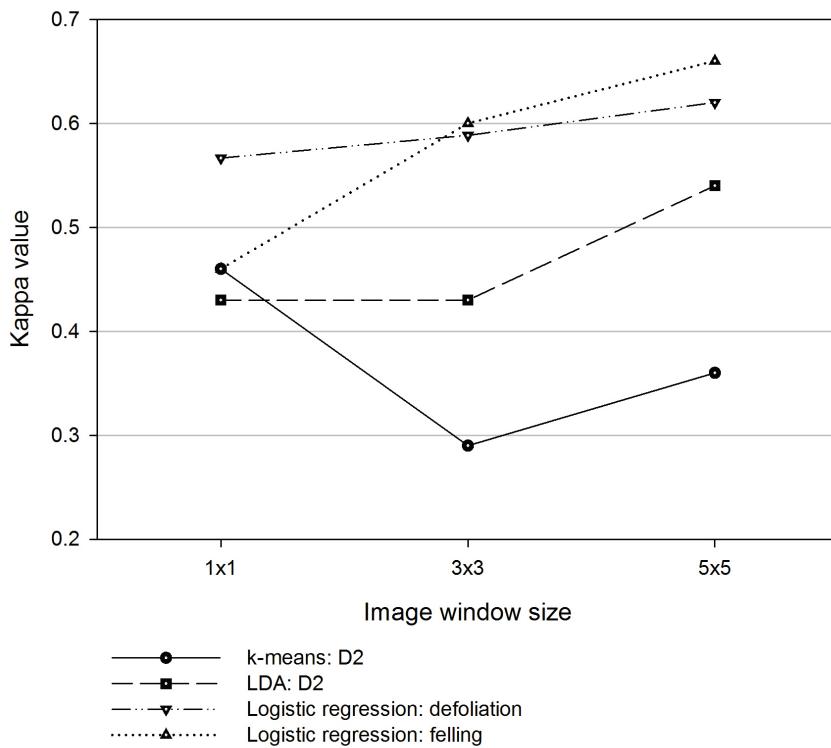
Figure 3.5. Kappa values of the estimation with different methods with 2 classes (i.e. class combination D2).

4. DISCUSSION

4.1. Starting point of the study

The objective of this study was to test the efficacy of mid-spatial resolution multi-temporal ERS-2 and ENVISAT SAR imagery in classification of Scots pine defoliation. Different combinations of individual classes and different classification procedures were tested to detemine the best classification methods. Two different speckle filters were also tested to determine their effect impact on classification accuracy. The results with the speckle filters were compared to the results without the filtering. In addition to classifying insect defoliation, the methods were also tested to determine how accurate the classifications were when considering clear-cut and thinning harvest classes together with the defoliation classes. The best SAR features predicting the defoliation level of the Scots pine stands were also investigated. The aim of this study was to find out the usability of SAR imagery in monitoring purposes of insect caused defoliation in conifer, and especially Scots pine, forests. When interpreting the results the small sample size and the correlation in the field data should be remembered.

4.2. Classification of defoliation and fellings

The classification of defoliation with the SAR imagery produced mostly moderate results and in the best case substantial classification accuracy. The best results in defoliation classification were accomplished using logistic regression for differentiating between two classes (presence and absence of defoliation). The highest classification accuracy using two classes was 81.6% (kappa value 0.62). The results are presumably too optimistic due to restrictions in field data, RS data, and accuracy assessment. These issues are discussed later. The most reliable results were received with k-means clustering, which uses only the spectral information of the RS images in the classification. However, k-means attained the lowest accuracies (the highest k-means classification accuracy was only 72.8% (kappa value 0.46)). The decrease of the accuracy with the addition of more

classes was expected due to the decreasing number of sample plots within each class. The effect of adding more categories to the classification followed the results of earlier studies (Ilvesniemi 2009, Puolakka 2010). In general, more classes lead to lower classification accuracies. This is especially true when the third defoliation class was added; the classification accuracy typically decreased by 10% or more when the defoliation class was added. The results are similar to Ilvesniemi's (2009), where classification accuracy decreased approximately by 5% to 15% when adding a third defoliation class. However, increasing the number of classes did not always decrease classification accuracy. In some situations, they type of class (defoliation or felled) had a varying effect on classification accuracy. The decrease in classification accuracy was on average smaller when the 'felled' class was added as compared to adding a third defoliation class. The class combination 4 (DF4) provided different results with the addition of the third defoliation class and the 'felled' class. Both the k-means and the LDA produced decent classification accuracies (over 50% of total classification accuracy) of 'no defoliation' and 'felled' classes at the expense classification accuracy in the moderate and severe defoliation classes. This is especially true when using LDA with 1 x 1 and 3 x 3 pixel windows; LDA did not correctly classify any in this situation. K-means did classify the moderate defoliation class ($\leq$20%) with a fairly high level of accuracy. However, the classification accuracy for the severe defoliation class was zero or near zero. However, k-means misclassified the plots suffering from severe defoliation mainly as moderate defoliation plots. This type misclassification was not serious because, in practice, it is more important to distinguish the defoliated forest from the healthy forest. Therefore, in the LDA classification the misclassification would be a more severe problem in practice, because the 'moderate defoliation' was mainly misclassified as 'felled' and the 'severe defoliation' to 'no defoliation'. When using the 5 x 5 pixel window, LDA was able to classify correctly at least some of the inventory plots.

The classification of defoliation attained with logistic regression can be compared to the two-class k-means and LDA defoliation classifications. In that respect, logistic regression produced the highest classification accuracies; the overall accuracy was 81.6% with a kappa value of 0.62. The high classification accuracy of the presence of defoliation (

92.9%) was surprising, especially considering the fact that the absence was classified with almost 70% of accuracy. Also, the kappa value was higher with logistic regression as compared to the 2-class classification with k-means or LDA. However, the difference in kappa values between logistic regression and LDA was not large. Nevertheless, logistic regression was the only method that produced substantial kappa result (Table 2.1). The results also demonstrate that, in defoliation classification with logistic regression, the pixel window size did not significantly influence classification accuracy. Although the accuracy increased with increasing window size, the difference in accuracy was small (only 3 percentage units). However, while the overall accuracy remains stable, the accuracy of individual classes changes. When the window size increased from 3 x 3 to 5 x 5 the accuracy of classifying presence of defoliation increased approximately 10 percent points. At the same time the classification of absence decreases, but not as much. One reason for this might be the fact that defoliated forest patches are closer in size (i.e., area) to a 5 x 5 pixel window (75 m x 75 m) than to either the 3 x 3 (45 m x 45 m) or the 1 x 1 (15 m x 15 m) pixel windows.

While the classification of defoliation worked quite well with the logistic regression, classifying the presence or absence or presence of felling worked even better. The 84.2% of classification accuracy and 0.66 kappa value are the best classification results attained in this study. The 77.4% accuracy of presence was slightly lower. This is largely due to the fact that the classification problem may be overly simple; determining the presence or absence of felling.

There were five different spectral features used in the classifications, the pixel-by-pixel amplitude, two amplitude differences, and the average and standard deviation between the SAR images. The individual explanatory power of the features is fairly poor (Table 3.1; Figure 3.1). When comparing the R^2 between different pixel windows (Table 4.1) the 5 x 5 pixel window produces the best results; however, the actual results are not very good in any case. Figure 4.1 demonstrates why the plots with low or no defoliation have the highest accuracies. The low defoliation plots are the only ones where image features

such as amplitude are correlated, to some degree, with defoliation; plots with high defoliation are not correlated with image features.

4.3. Speckle filtering

Another aspect of this study was concerned with determining the impact of speckle filtering on the classification results. The image windows employed for speckle filtering were 3 x 3 and 5 x 5 mean image filters. The classification accuracies attained via the two image filter sizes were compared to classification accuracies produced from non-filtered images. In addition to removing speckle noise, the image filters have other distinct advantages when classifying forest defoliation. For example, since the spatial extent of felled and defoliated areas is typically much larger than the spatial resolution of the radar data (15 x 15 m), and mean image filters essentially coarsen the spatial resolution of he radar data, using progressively larger image filters ultimate produce better classification results. However, the results don't unambiguously confirm this statement. K-means produced opposite results with class combinations D2 and D3, the unfiltered image produced the highest classification accuracies. This is especially true for with class combination D3 where the difference is classification was nearly 8 percent points higher than the next highest classification accuracy. For the class combination D2 the difference between the classification accuracies was small. Furthermore, the 3 x 3 filtered image produced the lowest classification accuracies in class combination D2. The reasons for this are unknown. In the other two class combinations (DF3 and DF4) the results were more as expected, the classification accuracy increases when the filter size increases. However, again, the differences in classification accuracies were not large.

The results were more like expected with the LDA and logistic regression methods. Both of the methods produced consistent results in terms of classification accuracy. Furthermore, when considering image filters, classification accuracy increased with increasing filter size. Following this logic, it would be interesting test the impacts of even larger window sizes on the accuracy of defoliation classifications. The classification

accuracy would perhaps increase, at least when not considering the felled class. Because the size of the felled compartments are between 0.5 and 2 hectares, the felling areas might get lost in a larger pixel window, likely decreasing the classification accuracy of this class. In this situation, the 3 x 3 and 5 x 5 pixel windows used in this study might be suited for identifying clear cuts and fellings.

4.4. Limitations and usability of the method

Answering the question, whether multi-temporal ERS-2 and ENVISAT SAR images are useful in the classification of defoliation in Scots pine forest, is not simple. Since the classification accuracies are fairly good, one could conclude that the methods employed herein to could be effectively employed to classify defoliation. This is especially true for the logistic regression methods, which produced the highest classification accuracies However, the restrictions of the study could actually inflate the classification accuracies attained in the current study. Particularly, the accuracy assessment for LDA and logistic regression methods were likely too optimistic. Cross validation is known to overestimate classification accuracies and model fit (Karjalainen et al. 2010). Because of the small size of the field data it was not possible to produce separate training and testing data to assess classification accuracy. The small sample size of the field data used in this study creates additional problems. A larger data would have likely given more robust models. Because of the small sample size it was not possible to test additional class combinations. Including additional classes with the current sample data would have lead to a paucity sample plots per class, and likely lead to poor classification results. The k-means method, being an unsupervised classification method, may perhaps give more realistic estimates of classification accuracy because it is based solely to the spectral properties of the SAR images. The best k-means classification produced a kappa value of approximately 0.45 and a classification accuracy of 70%. Thus, k-means produced, according to Landis and Koch (1977), a moderately accurate classification. The moderate classification accuracy achieved indicates that multi-temporal spatially mid-resolution SAR imagery could be

effective in the classification of, especially if remote sensing data from specific dates is needed. Still additional research must be conducted to determine the true capacity of classifying defoliation with multi-temporal SAR data. Indeed, including a more robust reference dataset that can support a rigorous accuracy assessment would support a more detailed evaluation and assessment of classifying defoliation with multi-temporal SAR data. Furthermore, although including both defoliation and felling classes in the same classification produced moderate accuracies, further research should be conducted to determine if there are other image features that are more effective for these two phenomena.

Usually two things impact the reflection intensity value recorded by the SAR instrument; the characteristics of the instrument itself and characteristics of the object being sensed (e.g., moisture content and texture). Specifically, the moisture content of an object has a large impact on its reflectivity in the microwave region of the electromagnetic spectrum. This is because moisture content alters an object's dielectric constant, a measure electrical permittivity of an object. In the current study, the characteristics of the SAR instruments are the same for each image, so sensor related differences are not likely influencing the classification results. However, due the multi-temporal nature of the SAR images used in this study, there is likely a large variation in surface moisture characteristics between each image acquisition, which could lead to significant differences in an objects microwave reflectivity between SAR acquisitions. This variation in microwave reflectivity likely degrades the accuracies of the defoliation classifications. Although it is possible to account for dielectric differences via the incorporation of meteorological information recorded during image acquisition times, we did not have access to such data. Because the meteorological properties are not known (i.e., we do not know if it has rained prior or during image acquisition) we do not have a means to correct for changes in the dielectric properties of the surface. The differenced SAR image feature performed poorer than was expected. However, they would have most likely been better if differences in moisture content could have been accounted for.

4.5. Comparing SAR and other remote sensing tools

Although the ERS-2 and ENVISAT SAR imagery employed in this study produced moderately accurate defoliation classifications, other studies have achieved higher accuracies with Landsat imagery (Table 4.1). For example, Ilvesniemi (2009) employed Landsat TM to classify defoliation and attained classification accuracy and kappa values of 88.7% and 0.76, respectively, which outperforms results of this study. Ranson et al. (2003) attained classification accuracies as high as 96.2% and 95.6% in the classification of severe and moderate insect damage classes, respectively, using Landsat TM. Puolakka (2010) tested the performance of ALS data in Scots pine defoliation mapping in the same study area (using a similar methodology) as in this study. LDA classification accuracies with two and four classes were 73% (kappa 0.42) and 53% (kappa 0.28), respectively, which is slightly weaker than in the current study. With k-means the best results in the study were 65% (kappa 0.26) and 13% (kappa 0.20), two and four classes, respectively, which are also lower than the current study, but fairly close. Furthermore, the classification accuracies in this study were better than in some other studies employing SAR data to classify defoliation. For instance, Ranson et al. (2003) used also SAR imagery in their defoliation classification and only achieved classification accuracies of 29% and 46% for severe and moderate insect defoliation classes, respectively. A study conducted by Karjalainen et al. (2010) attained with 3-Nearest Neighbor method accuracies of 75.2% and 67.8% using 400 x 400 meter pixel grid and the leave-one-out cross-validation and holdout accuracy assessment methods, respectively; results similar to those achieved in the current study. It should be noted that the results described by Karjalainen et al. (2010) are from the same area and using the same imagery as in the current study. However, the only difference was that the study by Karjalainen et al. (2010) was based on different field plots which were established in 2008.

Many new satellites are shot into space with modern remote sensing sensors on board, including high-end SAR sensor. For example new German TerraSAR-X satellite was successfully launched into space on 2007. TerraSAR-X has very high spatial resolution X-band SAR sensor on board, which could be one option in similar applications as was

done in this study. The sensors X-band would be suitable for defoliation estimation, because according to Le Toan et al. (1992) the shorter SAR frequencies (X and C-bands) are more correlated with the smaller canopy elements such as branches and needles. Also, the higher spatial resolution of the TerraSAR-X could give some benefit to the estimation process, though, increasing resolution always causes also increasing price which would be a drawback in a monitoring application. With TerraSAR-X it is also possible use the different images with different polarizations, which in this study was not possible. Using different polarizations can reveal new important aspects which could help in the estimation process. Also, the Japanese ALOS PALSAR is a fairly new SAR sensor. The PALSAR sensor uses the L-band in its imaging which according to Le Toan et al. (1992) is more correlated with the biomass, i.e. the tree trunks, than the smaller branches or needles. That could be a drawback with this sensor, but then again the possibility to use multiple polarizations. PALSAR does not produce as spatially accurate images as the TerraSAR-X, but ground resolution is still close to that of ERS-2 or ENVISAT. However, using multiple polarizations decreases the ground resolution. Also one interesting characteristic that should be looked into and could bring something new to the field is the use of interferometry, especially now after the launch of the TerraSAR-X sister satellite, TanDEM-X, in 2010. The satellites can produce together interferometric data by combining the data from each sensor into one image. TerraSAR-X and TanDEM-X are mainly to produce accurate digital elevation model of the world, but the use of interferometric data should be tested in other applications as well.

Table 4.1

Summary of studies made in classification of defoliation and list of the pros and cons of each remote sensing material.

Type of RS material	Author(s), year	Classification accuracy	Kappa value if applicable	Pros of RS material	Cons of RS material
Aerial photographs	Ilvesniemi 2009	88.7%	0.76	Very high spatial resolution, Visible/NIR bands and vegetation indexes bands, Planes can get in air almost anytime	Airplanes needs decent weather to fly, cloud coverage can be a problem, Can be expensive in large-area monitoring
Landsat TM/ETM+	Ilvesniemi 2009, Ranson et al. 2003	88.7%, 96.2% and 95.6% (for severe and moderate defoliation)	0.77, -	Decent spatial resolution, Easy data processing, Visible/NIR bands and vegetation indexes, Cheap	Cloud coverage makes images useless
SAR, single date	Ranson et al 2003	29% and 46% for severe and moderate defoliation	-	Cloud coverage does not matter, Price (depends on sensor) Temporal coverage, SAR features can describe characteristics that others cannot	Images not as easy to interpret as visible/NIR images, Single-date images may not give good enough accuracy, Price (depends on sensor)

Table 4.1 (*continued*)

Type of RS material	Author(s), year	Classification accuracy	Kappa value if applicable	Pros of RS material	Cons of RS material
SAR, multi-temporal	Karjalainen et al. 2010, Latva-Käyrä 2011	75.2% and 67.8% (Leave-One-Out CV and Holdout), 81.6%	- 0.62	Cloud coverage does not matter, Price (depends on sensor), Temporal coverage, Multi-temporality good in monitoring, SAR features can describe characteristics that others cannot	Images not as easy to interpret as visible/NIR images, Price (depends on sensor)
Area-based ALS	Puolakka 2010	73 %	0.42	Has worked really accurately in forest inventory and decently in defoliation applications	Expensive which limits use in large-scale monitoring
Tree-level ALS	Kantola et al. 2010 (ALS in combination with aerial photography)	88.1%	0.76	Works well in defoliation classifications on tree-level, Can also detect defoliation by monitoring changes in LAI	Even more expensive than area based ALS which limits use in large-scale monitoring

5. CONCLUSIONS

Although the study design may have lead to lower classification accuracies, the results descried herein are similar to those reported in other studies. This study also provides addition insight into the efficacy of employing multi-temporal SAR imagery for the assessment of defoliation. The plot-level defoliation estimates are likely too small to use with SAR imagery in its native spatial resolution (15 m). Smoothing the SAR data with image filters improved the classification results. In general, larger window sizes produced more accurate defoliation classifications. Thus, we recommend using larger image filters to smooth the SAR data prior to classification. Furthermore, the lack of meteorological information at the time of SAR acquisition likely degraded classification results. Thus, we stress the importance obtaining meteorological information during SAR image acquisitions. This data could be used to correct for moisture-related changes in dielectric constant, ultimately leading to more accurate defoliation classifications.

In the future mid-resolution SAR should be tested using larger image plots (tens to hundreds of pixels per plot), where the features have been averaged to a larger image window than in this study. Karjalainen et al. (2010) have done such an analysis and demonstrated that large 400 x 400 meter is a size which works in defoliation classification. Obtaining SAR data with a higher spatial resolution may also improve defoliation classification. Indeed, an increased spatial resolution and the possible use of full polarimetry and interferometry may enhance defoliation classification. In defoliation classification other remote sensing instruments have shown to be more viable than SAR (Table 5.1). Tree-level airborne laser scanning (Solberg et al. 2006, Kantola et al. 2010), aerial photographs (Haara and Nevalainen 2002, Ilvesniemi 2009), optical satellite imagery (Thomas et al. 2007, Ilvesniemi, 2009) and their combinations have been studied and proven effective in forest damage monitoring. However, it is still important to continue researching the efficacy of SAR imagery in defoliation studies, because the frequency of forest damages may rise due to continuing climate change, Thus, the demand for accurate and affordable remote sensing data will likely increase.

Furthermore, SAR has undisputed advantages in monitoring, most importantly the ability of SAR to image cloudy areas. In practice SAR would be advantageous many ways in helping in monitoring, predicting and/or in risk management of insect induced or other types of defoliation in boreal forests. Also, SAR could be used in other typical forestry applications, such as forest inventory and assessment, especially when combined with other remote sensing materials (e.g. ALS data).

6. REFERENCES

Belanger RP, Anderson RL. 1988. A guide for visually assessing crown densities of Loblolly and Shortleaf pines. Research Note SE-352. United States Department of Agriculture Forest Service, South-eastern Forest Experiment Station, Asheville. 4 p.

Campbell JB. 2007. Introduction to Remote Sensing. 4th ed. New York (NY): The Guildford Press. 626 p.

Ciesla W. 2000. Remote sensing in forest health protection. USDA Forest Service Remote Sensing Applications Center, Salt Lake City, UT and Forest Health Technology Enterprise Team, Fort Collins, CO. 266 p.

Cohen J. 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. 20(1):37-46.

Coops NC, Varhola A, Bater CW, Teti P, Boon S, Goodwin N, Weiler M. 2009a. Assessing differences in tree and stand structure following beetle infestation using lidar data. Canadian Journal of Remote Sensing. 35(6):497-508.

Coops NC, Waring RH, Wulder MA, White JC. 2009b. Prediction and assessment of bark beetle-induced mortality of lodgepole pine using estimates of stand vigor derived from remotely sensed data. Remote Sensing of Environment. 113:1058-1066.

Dale VH, Joyce LA, McNulty S, Neilson RP, Ayres MP, Flannican MD, Hanson PJ, Irland LC, Lugo AE, Peterson CJ, Simberloff D, Swanson FJ, Stocks BJ, Wotton BM. 2001. Climate change and forest disturbances. Bioscience. 51:723-734.

Diamantidis NA, Karlis D, Giakoumakis EA. 2000. Unsupervised stratification of cross-validation for accuracy estimation. Artificial Intelligence. 116(1-2):1-16.

De Somviele B, Niemelä P, Lyytikäinen-Saarenmaa P. 2004. Sawfly (Hym., Diprionidae) outbreaks on Scots pine: effect of stand structure, site quality and relative tree position on defoliation intensity. Forest Ecology and Management. 194:305-317.

De Somviele B, Lyytikäinen-Saarenmaa P, Niemelä P. 2007. Stand edge effects on distribution and condition of Diprionid sawflies. Agricultural and Forest Entomology. 9:17-30.

Dottavio, CL, Williams DL. 1983. Satellite technology: An improved means for monitoring forest insect defoliation. Journal of Forestry. 81(1):30-34.

Duda RO, Hart PE, Stork DG. 2001. Pattern Classification, 2nd ed. New York (NY): Wiley & Sons. 654 p.

Evans H, Straw N, Watt A. 2002. Climate change: Implication for insect pests. In: Broadmeadow M. (ed.) Climate Change: Impact on UK Forests. Forestry Commission Bulletin. 125:99-118.

Falkenström H, Ekstrand S. 2002. Evaluation of IRS-1c satellite data for defoliation assessment on Norway spruce and Scots pine. Remote Sensing of Environment. 82:208-223.

Fraser RH, Latifovic R. 2005. Mapping insect-induced tree defoliation and mortality using coarse spatial resolution satellite imagery. International Journal of Remote Sensing. 26(1):193-200.

Fukuda K, Pearson PA. 2006. Data mining and image segmentation approaches for classifying defoliation in aerial forest imagery. 3rd biennial meeting of the IEMSs, Burlington, VT, July 2006.

Geri G. 1988. The pine sawfly in central France. In: Berryman, A.A. (ed.). Dynamics of Forest Insect Populations: Patterns, Causes, Implications. New York (NY): Plenum Press. p. 377-405.

Gimeno M, San-Miguel-Ayanz J, Schmuck G. 2004. Identification of burnt areas in Mediterranean forest environments from ERS-2 SAR time series. International Journal of Remote Sensing. 25(22):4873-4888.

Goodwin NR, Coops NC, Wulder MA, Gillanders S, Schroeder TA, Nelson T. 2008. Estimation of insect infestation dynamics using a temporal sequence of Landsat data. Remote Sensing of Environment. 112(9):3680-3689.

Haara A, Nevalainen S. 2002. Detection of dead or defoliated spruces using digital aerial data. Forest Ecology and Management. 160:97-107.

Hall RJ, Skakun RS, Arsenault EJ. 2007. Remotely sensed data in the mapping of insect defoliation. In: Wulder MA, Franklin SE (eds.). Understanding Forest Disturbance and Spatial Pattern. Remote Sensing and GIS Approaches. Boca Raton (FL): CRC Press, Taylor & Francis Group. p. 85-111.

Hartigan JA. 1975. Clustering algorithms. New York (NY): Wiley. 322 p.

Holopainen M, Leino O, Kämäri H, Talvitie M. 2006. Drought damage in the park forests of the City of Helsinki. Urban Forestry and Urban Greening. 4:75-83.

Holopainen M, Haapanen R, Karjalainen M, Vastaranta M, Hyyppä J, Yu X, Tuominen S, Hyyppä H. 2009a. Combination of low-pulse ALS data and TerraSAR-X radar images in the estimation of plot-level forest variables. In: Bretar F, Pierrot-Deseilligny M, Vosselman G. (eds.). Proceedings Laserscanning 09. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science. 38(3):135-140.

Holopainen M, Tuominen S, Karjalainen M, Hyyppä J, Hyyppä H, Vastaranta M, Hujala T, Tokola T. 2009b. Korkearesoluutioisten E-SAR-tutkakuvien tarkkuus puustotunnusten koealatason estimoinnissa. Metsätieteen aikakauskirja. 4/2009:309-323. (In Finnish).

Holopainen M, Hyyppä J. 2010. Precision forestry by means of advanced laser measurements. In: Hetemäki L, Mary G. (Eds.). Future of the Forests – Responding to Global Changes. IUFRO-WFSE. In press.

Holopainen M, Haapanen R, Karjalainen M, Vastaranta M, Hyyppä J. Yu X, Tuominen S, Hyyppä H. 2010. Comparing accuracy of airborne laser scanning and TerraSAR-X radar images in the estimation of plot-level forest variables. Remote Sensing. 2(2):432-445.

Hosmer DW, Lemeshow S. 2000. Applied logistic regression, 2nd ed. New York (NY): Wiley & Sons. 392 p.

Hyyppä J, Hyyppä H, Inkinen M, Engdahl M, Linko S, Zhu Y-H. 2000. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. Forest Ecology and Management. 128:109-120.

Hyyppä J, Hyyppä H, Leckie D, Gougeon F, Yu X, Maltamo M. 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. International Journal of Remote Sensing. 29(5):1339-1366.

Hyyppä J, Hyyppä H, Yu X, Kaartinen H, Kukko H, Holopainen M. 2009. Forest inventory using small-footprint airborne lidar, In: Shan J, Toth C (Eds). Topographic Laser Ranging and Scanning: Principles and Processing. p. 335-370.

Ilvesniemi S. 2009. Numeeriset ilmakuvat ja Landsat TM -satelliittikuvat männyn neulaskadon arvioinnissa. Master's thesis. University of Helsinki. Department of Forest Sciences. 60 p. (In Finnish).

Intergovernmental Panel on Climate Change. 2007. Summary for policymakers. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Avyret KB, Tignor M, Miller HL (eds.). Climate Change 2007: The Physical Science basis. Contribution of Working Group 1 to the Fourth Assessment Report of the Intercovernmental Panel on Climate Change. Cambridge, UK and New York, (NY): Cambridge University Press.

Kankaanhuhta V, Mäkisara K, Tomppo E, Piri T, Kaitera J. 2000. Monitoring of diseases caused by *Heterobasidion annosum* and *Peridermium pini* in Norway spruce and Scots pine stands by airborne imaging spectrometry (AISA). In: Ukonmaanaho L, Raitio H (eds.). Forest condition monitoring in Finland - National report 1999. Metsäntutkimuslaitoksen tiedonantoja - The Finnish Forest Research Institute, Research Papers. 782:113-131.

Kantola T, Vastaranta M, Yu X, Lyytikäinen-Saarenmaa P, Holopainen M, Talvitie M, Kaasalainen S, Solberg S, Hyyppä J. 2010. Classification of defoliated trees using tree-

level airborne laser scanning data combined with aerial images. Remote Sensing. 2:2665-2679.

Karjalainen M, Kaasalainen S, Hyyppä J, Holopainen M, Lyytikäinen-Saarenmaa P, Krooks A, Jaakkola A. 2010. SAR satellite images and terrestrial laser scanning in forest damages mapping in Finland. The Proceedings of the ESA Living Planet Symposium, 28 June – 2 July 2010, Bergen, Norway. p. 6.

Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. Biometrics. 33(1):159-174.

Le Toan T, Beaudoin A, Riom J, Guyon D. 1992. Relating Forest Biomass to SAR Data. IEEE Transactions on Geoscience and Remote Sensing. 30(2):403-411.

Lillesand T, Kiefer R, Chipman J. 2007. Remote sensing and image interpretation. 6th ed. New York (NY): John Wiley & Sons. 804 p.

Lyytikäinen-Saarenmaa P, Anderbrant O, Löfqvist J, Hedenström E, Högberg HE. 1999. Monitoring the European pine sawfly population densities with pheromone traps in young pine plantations. Forest Ecology and Management. 124:113-121.

Lyytikäinen-Saarenmaa P, Tomppo E. 2002. Impact of sawfly defoliation on growth of Scots pine Pinus sylvestris (Pinaceae) and associated economic losses. Bulletin of Entomological Research. 92:137-140.

Lyytikäinen-Saarenmaa P, Niemelä P, Annila E. 2006. Growth Responses and Mortality if Scots Pine (Pinus Sylvestris L.) after a Pine Sawfly Outbreak. In report: Kamata N, Liebhold S, Quiring D, Clancy K. (Eds.). Proceedings: IUFRO Kanazawa 2003. Forest Insect Population Dynamics and Host Influences, Sept. 14-19, Kanazawa Japan. p. 81-85.

Lyytikäinen-Saarenmaa P, Holopainen M, Ilvesniemi S, Haapanen R. 2008. Detecting pine sawfly defoliation by means of remote sensing and GIS. Forstschutz Aktuell. p. 14-15.

Manninen T, Stenberg P, Rautiainen M, Voipio P, Smolander H. 2005. Leaf Area Index Estimation of Boreal Forest Using ENVISAT ASAR. IEEE Transactions on Geoscience and Remote Sensing. 43(11):2627-2635.

Mattila E. 1998. Use of satellite and field information in a forest damage survey of Eastern Finnish Lapland in 1993. Silva Fennica. 32(2):141-162.

Nelson R. 1981. ASSESS2 analysis of four methods for classifying forest defoliation. Goddard Earth Resource Branch, Internal Report. p. 11.

Næsset E, Gobakken T, Holmgren J, Hyyppä H, Hyyppä J. Maltamo M, Nilsson M, Olsson H, Persson Å, Söderman U. 2004. Laser scanning of forest resources: the Nordic experience. Scandinavian Journal of Forest Research. 19(6):482-499.

Pulliainen J, Heiska K, Hyyppä J, Hallikainen M. 1992. Detection of Spruce Defoliation with Microwave Radar Techniques. 22nd European Microwave Conference, Sept. 5-9 1992, vol.2:939-944.

Puolakka P. 2010. Männyn neulaskadon arviointi laserkeilauksella. Master's thesis. University of Helsinki. Department of Forest Sciences. 64 p. (In Finnish).

Ranson KJ, Kovacs K, Sun G, Kharuk VI. 2003. Disturbance recognition in the boreal forest using radar and Landsat-7. Canadian Journal of Remote Sensing. 29(2):271-285.

Ranta E, Rita H, Kouki J. 2005. Biometria - Tilastotiedettä ekologeille. 9th ed. Helsinki: Yliopistopaino. 569 p.

Rauste Y. 1990. Incidence-angle dependence in forested and non-forested areas in Seasat SAR data. International Journal of Remote Sensing. 11:1267-1276.

Rauste Y, Häme T, Pulliainen J, Heiska K, Hallikainen M. 1994. Radar-based forest biomass estimation. International Journal of Remote Sensing. 15:2797-2808.

Rauste Y. 1996. Detection of forest damages with multitemporal ERS-1 SAR data. Publications of the Academy of Finland. Finnish research programme on climate change, Final report:427-432.

Rauste Y. 2005. Multi-temporal JERS SAR data in boreal forest biomass mapping. Remote Sensing of Environment. 97:263-275.

The R development core team. 2009. R: A language and Environment for Statistical Computing. Reference Index. Version 2.10.0. 2975 p. Available from: http://www.r-project.org/.

Régniére J. 2009. Predicting insect continental distributions from species physiology. Unasylva. 60:37-42.

Solberg S, Næsset E, Hanssen KH, Christiansen E. 2006. Mapping defoliation during a severe insect attack on Scots pine using airborne laser scanning. Remote Sensing of Environment. 102:364-376.

Solberg S. 2008. Mapping gap fraction, LAI and defoliation using various ALS penetration variables. International Journal of Remote Sensing. 31(5):1227-1244.

Solberg S, Brunner A, Hanssen K, H, Lange H, Næsset E, Rautiainen M, Stenberg P. 2009. Mapping LAI in a Norway spruce forest using airborne laser scanning. Remote Sensing of Environment, 113(11):2317-2327.

Talvitie M, Leino O, Holopainen M. 2006. Inventory of sparse forest populations using adaptive cluster sampling. Silva Fennica. 40(1):101-108.

Talvitie M, Kantola T, Holopainen M, Lyytikäinen-Saarenmaa P. 2011. Adaptive Cluster Sampling in Inventorying Forest Damage by the Common Pine Sawfly. Journal of forest planning. 16:1-7.

Thomas SJ, Deschamps A, Landry R, van der Sanden JJ. 2007. Mapping insect defoliation using multi-temporal Landsat data. CRSS/ASPRS 2007 Specialty conference. Our common borders - safety, security and the environment through remote sensing. Conference proceedings.

Thompson SK. 1990. Adaptive cluster sampling. Journal of the American Statistical Association. 85(412):1050-1059.

United Nations Economic Commission for Europe. 2006. Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. Part II: Visual assessment of crown conditions. [cited: 2011 Feb 21] Available from: http://www.icp-forests.org/pdf/Chapt2_compl06.pdf .

Viitasaari M, Varama M. 1987. Sahapistiäiset 4. Havupistiäiset (Diprionidae). University of Helsinki, Department of Agricultural and Forest Zoology, Reports 10. 79 p. (In Finnish).

Vuokila Y. (Ed). 1987. Metsikkökokeiden maastotyöohjeet. Metsäntutkimuslaitoksen tiedonantoja. 257. Helsinki: Hakapaino Oy. 237 p. (In Finnish).

Walther GR, Post E, Convey P, Menzel A, Parmesan C, Beebee TJC, Fromentin J-M, Hoegh-Guldberg O, Bairlein F. 2002. Ecological responses to recent climate change. Nature. 416:389-395.

Williams DL, Nelson RF. 1986. Use of remotely sensed data for assessing forest stand conditions in the Eastern United States. IEEE Transactions on Geoscience and Remote Sensing. 24(1):130-138.

Wulder MA, Dymond CC, White JC, Leckie DG, Caroll AL. 2006. Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities. Forest Ecology and Management. 221:27-41.