Norwegian University of Life Sciences • Universitetet for miljø- og biovitenskap
Department of Ecology and Natural Resource Management
Philosophiae Doctor (PhD) Thesis 2012:44

Detection of Small Single Trees in the Forest-Tundra Ecotone Using Airborne Laser Scanning

Deteksjon av små enkelttrær i tregrensa ved hjelp av flybåren laserscanning

Nadja Stumberg

Detection of Small Single Trees in the Forest-Tundra Ecotone Using Airborne Laser Scanning

Deteksjon av små enkelttrær i tregrensa ved hjelp av flybåren laserscanning

Philosophiae doctor (PhD) Thesis

Nadja Stumberg

Department of Ecology and Natural Resource Management
Norwegian University of Life Sciences

Ås 2012

Thesis number 2012:44
ISSN 1503-1667
ISBN 978-82-575-1081-7

PhD supervisors

Professor Erik Næsset
Department of Ecology and Natural Resource Management
Norwegian University of Life Sciences
P.O. Box 5003, NO-1432 Ås, Norway

Professor Terje Gobakken
Department of Ecology and Natural Resource Management
Norwegian University of Life Sciences
P.O. Box 5003, NO-1432 Ås, Norway

Dr. Ole Martin Bollandsås
Department of Ecology and Natural Resource Management
Norwegian University of Life Sciences
P.O. Box 5003, NO-1432 Ås, Norway

Evaluation committee

Professor Benoît St-Onge
Département de géographie
Université du Québec à Montréal
400, rue Sainte-Catherine Est, Montréal (Québec) H2L 2C5, Canada

Professor Håkan Olsson
Deptartment of Forest Resource Management
Swedish University of Agricultural Sciences
Skogsmarksgränd, SE-901 83 Umeå, Sweden

Associate Professor Kari Klanderud
Department of Ecology and Natural Resource Management
Norwegian University of Life Sciences
P.O. Box 5003, NO-1432 Ås, Norway

Preface

This thesis has been submitted in completion of the requirements for the degree Philosophiae doctor (PhD) at the Department of Ecology and Natural Resource Management (INA) at the Norwegian University of Life Sciences (UMB). The project has been financed by the Research Council of Norway (project #184636/S30) and a stay at the Pacific Forestry Centre (PFC) was funded by their "Utviklingsfondet for skogbruk".

I would like to thank my supervisor, Professor Erik Næsset, and my co-supervisor, Professor Terje Gobakken, for giving me the opportunity to apply my scientific background as a cartographer to the problems and questions of forestry. Thank you for giving me enough time to get familiar with forest inventory and the freedom to follow ideas from my perspective. In January 2010, Dr. Ole Martin Bollandsås joined the team of supervisors. Thanks to the three of you for all your supervision, suggestions, constructive criticism, comments on my work and your patience to let me find my own way. Terje and Ole Martin, thank you for always being around and your open doors and ears.

I would also like to thank Dr. Michael A. Wulder at the PFC for hosting me in autumn/winter 2010 and for encouraging me in my work. Thanks to all my Canadian colleagues at the PFC, especially Dr. Meg Andrew and "the lab", for making my stay not just a scientific but also a personal enrichment.

Thanks to my fellow colleagues at INA, especially Dr. Hans Ole Ørka. Thank you all for always being around, supportive and encouraging, thanks for the coffee breaks, fruitful discussions, motivation, chocolate, inspiration, ice cream, ideas, cookies and your open doors and ears.

A special thanks to my friends and my family in particular for all their motivation and support and for being there at any time. Thank you, my dear Jon, for always being there for me during the last two years. Your patience, moral support and "tasks" such as chopping firewood and washing the tractor and harvester to clear my head were invaluable. And thanks to the three furry members of my home-office crew who always knew when I needed a hug.

Ås, August 2012

Contents

Appendix: Papers I to IV

Abstract

Alpine and arctic tree lines are expected to advance to higher altitudes and further north due to global warming. The forest-tundra ecotone in particular, is highly sensitive to climatic changes since many of the species found there are at their tolerance limits. Thus, the development of suitable methods for monitoring these changes is of great importance and interest. For the monitoring of such vast areas as the forest-tundra ecotone, airborne laser scanning (ALS) may provide a well-suited tool because of its capability to estimate biophysical parameters on single tree level at different geographical scales. The main objective of this thesis was to investigate the potential of using high-density ALS data for detection of small individual trees located in the forest-tundra ecotone. The specific parts of the thesis focus on (1) single tree detection using ALS height values in combination with variables describing tree characteristics and the site, (2) laser echo classification using laser height and intensity, geospatial and terrain variables, as well as geostatistics and statistical measures. Furthermore, (3) the potential of an unsupervised classification of raster cells for automated monitoring programs of small single trees was assessed. Along a 1,500 km long transect stretching from northern Norway (66°19' N) to the southern part of the country (58°3' N) field measurements of 744 small individual trees as well as ALS data were collected. Generalised linear models (GLM), a generalised linear mixed model (GLMM), support vector machines (SVM), and a raster-based algorithm concept were employed for the detection and classification of both trees as well as tree and non-tree laser echoes using different variables. Successful single tree detection using laser height values in combination with tree characteristics and spatial influences as latitude and region was verified for trees exceeding a height of 1 m using GLM and GLMM models. The results form a solid basis for generalisation and inference that goes far beyond previous research because of the huge geographical extension of the dataset. Secondly, the capability of the ALS data for classification into tree and non-tree echoes using laser measurements, geospatial and terrain variables was confirmed using the two different modelling techniques GLM and SVM. Furthermore, an extension of the classification models with geostatistical and statistical measures employing GLM and SVM revealed a significant improvement. Finally, the suitability of an unsupervised classification approach for the automatic detection of small single trees was verified for parameter values ensuring a justifiable trade-off between rates of successful detection and commission errors.

Sammendrag

De alpine og arktiske tregrensene er forventet å flytte seg høyere opp i fjellet og lengre mot nord som følge av global oppvarming. Trærne i disse områdene lever nær sin toleransegrense og er derfor følsomme for klimatiske endringer. Utvikling av egnede metoder for overvåking av disse endringene er av stor betydning og interesse. For overvåking av slike utstrakte områder som tregrensa, kan flybåren laserscanning (ALS) være et godt egnet redskap for å samle data som kan benyttes til å estimere biofysiske parametre på enkelttrenivå på ulike geografiske skalaer. Hovedmålet for denne avhandlingen var å undersøke potensialet av bruk av ALS-data med høy punkttetthet for detektering av små enkelttrær i tregrensa. De spesifikke delene av avhandlingen fokuserer på (1) deteksjon av enkelttrær ved hjelp av høydeverdier fra ALS-data i kombinasjon med variabler som beskriver tre-egenskaper og geografiske områder, (2) klassifikasjon av de enkelte ALS-registreringene (laserekko) ved hjelp av laserhøyde og -intensitet, informasjon om ALS-registreringenes romlige fordeling, terrengvariabler, samt geostatistiske og statistiske mål. Videre ble (3) potensialet av en ikke-styrt klassifikasjon av rasterceller for automatiserte overvåkingsprogrammer av små enkelltrær vurdert. Langs et 1500 km langt transekt, som strekker seg fra Nord-Norge (66°19’ N) til den sørlige delen av landet (58°3’ N), ble det samlet inn feltobservasjoner av 744 små enkelttrær og ALS-data. Både generaliserte lineære modeller (GLM), en generalisert blanded lineær modell (GLMM), support vector machines (SVM) og en rasterbasert algoritme ble brukt for deteksjon av trærne og for å klassifisere laserekkoene ut fra om de var returnert fra trær eller ikke-trær. Deteksjonen av enkelttrær ved hjelp av verdier fra ALS i kombinasjon med treegenskaper og romlige påvirkninger som breddegrad og region ble verifisert for trær over en høyde på 1 m ved hjelp av GLM og GLMM modeller. Resultatene danner et solid grunnlag for generalisering og slutning som går langt utover tidligere forskning på grunn av den betydelige geografiske utbredelsen av datasettet. Dernest ble evnen til å klassifisere de ulike ekko, hvorvidt de var returnert fra trær eller ikke-trær, ved hjelp av variable som beskriver terrenget og laserekkoenes romlige fordeling, bekreftet ved hjelp av de to ulike modelleringsteknikkene GLM og SVM. Videre ble det påvist en betydelig forbedring av GLM og SVM modellene når de ble utvidet med geostatistiske og andre statistiske mål. Avslutningsvis ble egnetheten til en ikke-styrt klassifikasjon for en automatisk deteksjon av små enkelttrær verifisert for parameterverdier som sikrer en forsvarlig avveining mellom vellykket deteksjon og inkluderingsfeil.

List of Papers

Paper I

Thieme, N., Bollandsås, O.M., Gobakken, T. & Næsset E. 2011. Detection of small single trees in the forest-tundra ecotone using height values from airborne laser scanning. *Canadian Journal of Remote Sensing*, Vol. 37, No. 3, pp. 264–274.

Paper II

Thieme, N., Ørka, H.O., Bollandsås, O.M., Gobakken, T. & Næsset E. 2012. Classifying tree and non-tree echoes from airborne laser scanning in the forest-tundra ecotone. Manuscript submitted for publication to *Canadian Journal of Remote Sensing*.

Paper III

Thieme, N., Bollandsås, O.M., Gobakken, T. & Næsset E. 2012. Improving classification of airborne laser scanning echoes in the forest-tundra ecotone using geostatistical and statistical measures.

Paper IV

Thieme, N., Bollandsås, O.M., Gobakken, T. & Næsset E. 2012. Automatic detection of small single trees in the forest-tundra ecotone using airborne laser scanning.

SYNOPSIS

1. Introduction

Shifts of the climatic treelines to higher altitudinal and latitudinal areas have been observed since the end of the Little Ice Age in the late 1880s (Holtmeier and Broll, 2007). Besides the favourable climatic period from 1925 to 1945, the annual average temperature in arctic regions increased at almost twice the global rate over the past decades albeit regional variations (ACIA, 2004; IPCC, 2007). Such a temperature increase may influence the prevailing tree limit not just by advancing into greater altitudes and higher latitudes (ACIA, 2004; Kullman and Öberg, 2009), but also by a densification (Danby and Hik, 2007; Batllori and Gutiérrez, 2008) and increased height growth (Kullman, 2002) of the current sparsely distributed pioneer trees.

The United Nations Framework Convention on Climate Change and the Kyoto protocol involve reports on land use change in respect of deforestation, afforestation and reforestation (UNFCCC, 2008) implying an important need for data collection in low biomass areas by means of carbon accounting. Since National Forest Inventories (NFI) or other monitoring systems commonly do not prioritise field plots in areas where the shifts in the climatic treeline may occur because of the high costs for data acquisition, efficient monitoring systems with the capability to cover vast areas with a high degree of detail at small scales are required.

1.1 Ecological aspects of the forest-tundra ecotone

Alpine and arctic treelines are seldom distinct demarcation lines, but are rather represented by transition zones (Callaghan et al., 2002; Holtmeier and Broll, 2005) covering the area between the mountain forest and the alpine and arctic zones. Clements (1905) referred to such transitions as ecotones and Harper et al. (2011) define the forest-tundra ecotone as “the transition between forest and tundra at high elevation or latitude”. Its location involves high sensitivity to climatic changes, and an advance of the alpine and arctic treelines to higher altitudes and northwards is expected caused by an increase in mean temperature and changes in precipitation as well as snow coverage affecting the length of the growing season (Callaghan, 2002; ACIA, 2004). Furthermore, an increased temperature may result in a densification, increased height growth as well as migration further north and to greater altitudes of the current tree limit (Kullman, 2002; ACIA, 2004; Danby and Hik, 2007; Batllori and Gutiérrez, 2008; Kullman and Öberg, 2009). However, a successful colonisation of formerly treeless areas require long-term survival of seedlings and saplings into trees

(Aune et al., 2011) as well as an increment in height growth of the present tree layer (Kullman, 2002). For this purpose, the production, dispersal, and germination of seeds but also the interplay of abiotic and biotic drivers are essential (Cairns and Moen, 2004; Holtmeier and Broll, 2005; Sturm et al., 2005; Aune et al., 2011). In addition, anthropogenic factors as herbivore activity by domestic animal and pastoral economy have an effect on the tree limit and may inhibit the climatic responses (Callaghan et al., 2002; Holtmeier and Broll, 2005; Post and Pedersen, 2008; Olofsson et al., 2009; Hofgaard et al, 2010; Aune et al., 2011). Furthermore, the changed properties of the forest-tundra ecotone will affect the biodiversity, biomass and carbon pools of the vegetation zones adjacent to the forest-tundra ecotone, i.e., the mountain forest and the tundra. This may for example be reflected in improved growth conditions in the forest-tundra ecotone caused by a denser mountain forest as a result of an increased occurrence of biomass in the forest-tundra ecotone that provided better protection for the mountain forest.

1.2 Airborne laser scanning in the forest-tundra ecotone

In Norway, the forest-tundra ecotone covers a large proportion of the total land area. For monitoring such vast areas, different remote sensing techniques could provide objective wall-to-wall data for land cover assessment. However, air- or spaceborne optical instruments are limited by their spatial resolutions which are not sufficient enough for the detection of small-sized trees and therewith the changes in their biophysical properties and spatial distribution. Assuming a height growth of 1 to 10 cm per year depending on locality and the prevailing microclimate, a remote sensing technique capable to observe subtle changes in growth and migration patterns is required. Airborne laser scanning (ALS) is a well suited data source as documented by several studies on the precise estimation of biophysical parameters on a single-tree level (e.g. Hyyppä et al., 2001; Persson et al., 2002; Solberg et al., 2006; Næsset and Nelson, 2007). Furthermore, ALS has proven its ability to discriminate small individual trees in the forest-tundra ecotone in several studies using different laser point densities (e.g. Næsset and Nelson, 2007; Rees, 2007). Employing ALS data with a point density of approximately 0.25 m^{-2}, Rees (2007) discriminated individual trees provided a minimum tree height of 2 m over vast areas covering hundreds of square kilometres. Næsset and Nelson (2007), however, used high-density ALS data with a point density of 7.7 m^{-2} to detect small single trees irrespective of their tree height. In this study, positive laser height values inside field-measured tree crown polygons were used as criterion for successful tree detection. Thereby, success rates of over 90% were reported for trees with tree heights exceeding 1 m

(Næsset and Nelson, 2007) implying an adequate reliability of the method used for trees larger than 1 m in height. The success rates for trees lower than 1 m were significantly lower when utilising positive laser height values as criterion for the successful discrimination of a tree caused by severe commission errors (Næsset and Nelson, 2007; Næsset, 2009b). The magnitude of positive laser height values emerging from non-tree objects is not just dependent on the occurrence of for instance rocks, hummocks, and other terrain structures, but also on the properties related to the terrain model, the sensor, and the flight settings (Næsset, 2009b). Commission errors of up to 490% were observed by Næsset and Nelson (2007) using a dataset based on a terrain model computed with commonly adopted smoothing criteria. Thus, the reliability of tree detection solely employing positive laser height values is to a high degree affected by these commission errors, especially for trees lower than 1 m in height. However, for monitoring purposes, high rates of commission errors are negligible because of the multi-temporal context involving a change in size and number of trees whereas terrain and terrain objects remain stable.

Furthermore, in a monitoring context but also for data acquisitions over vast areas, different sensors and acquisitions settings may be used over time and areas. For larger trees, it has been reported that variations in the vegetations returns and the properties of the laser point clouds may influence on estimated biophysical properties (Næsset, 2005, 2009a; Chasmer et al., 2006). However, previous research on small individual trees indicated that the utilisation of different instruments and configurations were equally well suited for single tree detection, provided a tree height exceeding 1 m (Næsset, 2009b).

1.3 Potential discriminators for single tree detection in the forest-tundra ecotone

Typically, studies utilising ALS data for forest inventory purposes solely employ the height information of the laser echoes rather than the full suite of available information as for instance spectral data, i.e., the backscatter intensity of the laser echoes. Intensity is an often neglected laser metric albeit it may be useful for the discrimination of tree and non-tree objects or laser echoes. Schreier et al. (1985) employed laser intensity for tree species classification already in 1985, but during the following years there was little focus on the usage of this laser metric because of the lack of radiometric calibration methods (Kaasalainen et al., 2005). During the last decade, however, several studies were conducted to classify tree species (e.g. Brandtberg et al., 2003; Holmgren et al., 2008; Ørka et al., 2009), age (Farid et al., 2006a,b), and land-cover (Brennan and Webster, 2006). Furthermore, an experimental study by Thieme et al. (2011) used normalised intensity values to investigate the spatial

pattern of tree and non-tree objects based on laser height and intensity. They found normalised intensity values and height information useful to separate between tree and non-tree objects in a forest-tundra ecotone environment when combined with the spatial point pattern of the individual laser echoes with positive height values.

Hence, the spatial point pattern, i.e., the spatial structure and distribution of the individual laser echoes may be conducive for the discrimination between different types of objects located on the terrain surface. The spatial variation of individual laser echoes may differ around tree and non-tree objects since a variety of biological phenomena demonstrate spatial correlation or dependency (Rossi et al., 1992) and often emerging in patches (Fry and Stephens, 2010). Voronoi polygons are a commonly used technique in point pattern analysis to investigate the spatial distribution of point data in numerous disciplines (Boots and Getis, 1988) and Thieme et al. (2011) showed promising results employing small-sized Voronoi polygons to recognise field-measured trees and non-tree objects selected from aerial imagery. Furthermore, Thieme et al. (2011) also employed experimental variograms and cross-variograms based on laser height and intensity values in a geostatistical analysis investigating differences in the pattern between tree and non-tree objects. Geostatistics are a common image processing technique in optical remote sensing where standard statistical measures such as mean and standard variation or the variogram-derived mean semivariance are calculated for each pixel using a moving window for image classification purposes (Wulder et al., 1998; Jakomulska and Clarke, 2001). Wulder et al. (1998) used first- and second-order texture as well as semivariance moment texture for textural image classification to increase the accuracy of leaf area index estimation. Thus, we hypothesize that Voronoi polygons representing the spatial point pattern of the ALS data, as well as standard statistical measures and a geostatistical component may be suitable co-discriminators for the classification of trees and non-trees in the forest-tundra ecotone.

Furthermore, several studies state that different terrain characteristics such as aspect and slope influence the potential presence and height growth of small pioneer trees (Mast et al., 1997; Boisvenue et al., 2004; Danby and Hik, 2007). Danby and Hik (2007), for instance, reported a difference in tree invasion patterns for north and south-facing slopes in a forest-tundra ecotone environment, primarily caused by the differential presence of permafrost. They further demonstrated the partial dependency of regional, landscape and local scale variability in the tree population on variations in the terrain, landscape setting and existing vegetation (Danby and Hik, 2007). Thus, the terrain parameters aspect and slope may be contributing to the discrimination between trees and non-trees.

1.4 Supervised and unsupervised classification

In general, statistical classification is separated into supervised and unsupervised classification methods. In supervised classification techniques, classes are built by using training data for the parametric or non-parametric characterisation of these classes and elements are assigned based on their characteristics. Such decision rules are not provided in unsupervised classification techniques where classes are built without any usage of training data and without any previous knowledge of the thematic content. Unsupervised classification methods generally embody cluster analysis by aggregating elements into clusters where each cluster represents a homogeneous class.

For the classification of trees and non-trees, both supervised and unsupervised classification methods are still a little utilised approach. However, these methods may represent a useful tool with a yet unknown potential for inventory and monitoring of small individual trees in vegetation zones as the forest-tundra ecotone.

1.5 Research objectives

The main objective of the thesis was to investigate the potential of high-density ALS data to detect small individual trees located in the forest-tundra ecotone using different methods and approaches. The specific parts of the thesis and the relationship of the research papers included are illustrated in Figure 1 and described as follows:

- *Paper I – "Detection of small single trees in the forest-tundra ecotone using height values from airborne laser scanning"* – studied the effects of sensor influences on small tree detection, the probability of small tree detection using ALS height values in combination with tree characteristics and site and further assessed the accuracy of laser-derived tree height estimation.
- *Paper II – "Classifying tree and non-tree echoes from airborne laser scanning in the forest-tundra ecotone"* – the analysis from Paper I was brought a step further and supervised laser echo classifications using laser measurements, geospatial, and terrain variables were tested.
- *Paper III – "Improving classification of airborne laser scanning echoes in the forest-tundra ecotone using geostatistical and statistical measures"* – employed the best classification models from Paper II and tested geostatistics and statistical measures derived from laser height and intensity values to improve the classification of tree and non-tree echoes.

- *Paper IV – "Automatic detection of small single trees in the forest-tundra ecotone using airborne laser scanning"* – took the results from Paper I a step further and assessed, contrary to Paper II and III, the potential of an unsupervised classification for automated monitoring programs of small single trees.

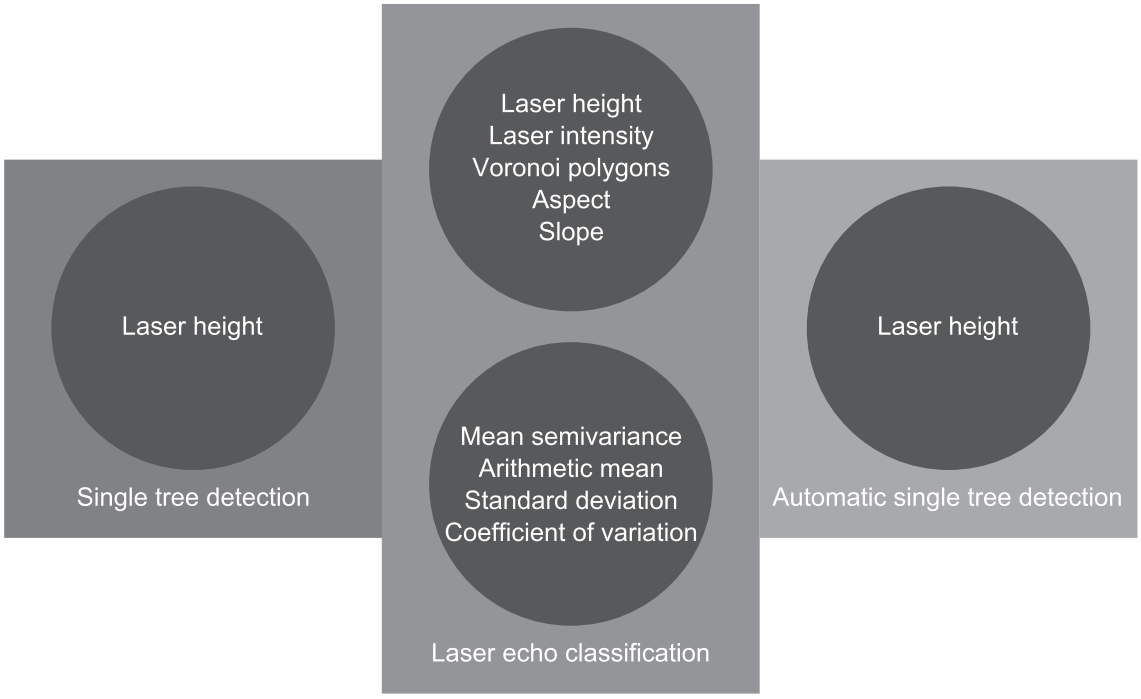

Figure 1 – Specific parts and relationship of the research papers included in the thesis.

2. Study area and materials

2.1 Study area

The study area is located along a 1,500 km long and approximately 180 m wide longitudinal transect and encompasses hundreds of mountain forest and alpine elevation gradients. The transect stretches from Tromsø in northern Norway (69°3' N 17°5' E) to Tvedestrand in the southern part of the country (58°3' N 9°0' E) (Figure 2) and covers sample plots in the forest-tundra ecotone at elevations between 350 and 1200 m a.s.l. In these transitions between the mountain forest and the alpine zone, the terrain is often characterised by rounded forms with occurrences of hummocks, rocks and boulders, and steep slopes. The prevalent tree species are Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and mountain birch (*Betula pubescens* ssp *czerepanovii*).

Figure 2 - Overview of the study area with the 35 specific field sites (black and white points). The 1,500 km long transect (black line) stretches from to 69°3' N 17°5' E to 58°3' N 9°0' E. Three field sites are located in the overlap zone between ALS acquisitions conducted with two different instruments (white points). (Reproduced with kind permission from CASI)

2.2 Field data

The field work was conducted in summer 2008 at 35 different field sites selected along the transect to provide *in situ* tree data for analyses. For this purpose, 56 evenly distributed sites were initially identified for the potential establishment of sample plots prior to field work. These sites were chosen based on aerial images and maps from the official Economic Map Series of the Norwegian Mapping Authority. The proximity to roads and the existence of sample plots of the NFI nearby for potential comparison were taken into account. The final selection of the specific sites, however, was made in field.

A field site typically consisted of two to four sample plots with a radius of 25 m to cover the entire width of the forest-tundra ecotone, i.e., the area between the mountain forest and the treeless alpine zone. The number of sample plots in a field site was determined in field based on visual judgement because of the variation in width of the forest-tundra ecotone between the different locations. Furthermore, the sample plots were laid out with an interdistance of 50 m within a field site to avoid overlap. In total, 111 sample plots were established at 35 different field sites located along the entire transect.

For precise navigation and positioning, real-time kinematic differential Global Navigation Satellite System (dGNSS) was utilised. Two Topcon Legacy E+ 20-channel dual-frequency receivers were employed as base and rover receivers observing pseudo range and carrier phase of both Global Positioning System and Global Navigation Satellite System satellites. A base station was established for each field site using the closest suitable reference point of the Norwegian Mapping Authority. The expected accuracy of the sample plot centre points was 3–4 cm provided an expected horizontal accuracy of about 2 cm for the field recordings relative to the base station and an expected accuracy of 3 cm for the reference points.

Individual trees were selected for measurement on each plot. The individual sample trees were selected using a modified version of the point-centred quarter sampling method (PCQ) (Cottam and Curtis, 1956; Warde and Petranka, 1981) with a maximum search distance of 25 m. For this purpose, the sample plot was divided into four quadrants defined by the cardinal directions from the sample plot centre using a Suunto compass. In each quadrant, trees that were closest to the plot centre in a respective tree height class were sampled independent of tree species. Three tree heights classes were defined as: (1) lower than 1 m, (2) between 1 and 2 m, and (3) taller than 2 m, resulting in a maximum of potentially 12 trees in a sample plot. In cases of doubt, the maximum search limit and the closest tree were determined using a surveyor's tape measure.

For each sample tree, the precise position was determined using dGNSS and several tree parameters were recorded individually. Tree height was measured using a steel tape measure or a Vertex III hypsometer for tall trees and stem diameter was callipered at root collar. Tree species was determined and crown diameters were measured in the cardinal directions using a steel tape measure.

In total, 744 trees were measured including 623 mountain birch, 68 Norway spruce, and 53 Scots pine. Tree heights ranged from 0.02 to 7.80 m, and crown areas from 0.001 to 19.54 m^2, computed as the ellipse defined by the crown diameters as the major and minor axes.

In Paper I, the entire dataset was used. In Paper IV, however, ten trees had to be discarded from the dataset because of their tree crowns being completely overlapped by tree crowns of taller trees which was regarded as invalid for the analyses in this study. Furthermore, a subset of the dataset with 524 trees, i.e. 404 mountain birch, 67 Norway spruce and 53 Scots pine, was used in Papers II and III. Summaries of the tree parameters for the dataset and the two subsets are given in Table 1.

2.3 Laser data

Airborne laser scanner data were acquired on 23 and 24 July 2006 with an Optech ALTM 3100C laser scanner system covering southern and central Norway and on 1 July 2007 with a Gemini upgraded version of the Optech ALTM 3100C laser scanner system, denoted as ALTM Gemini, in northern Norway. The laser data collection had to be separated in two acquisitions because of the large geographical extent of the study area and difficult weather conditions. An overlap zone of approximately 80 km was scanned in the county of Nordland (65°53' N 13°27' E) with both systems to facilitate sensor comparisons and thus control for any potential sensor effects.

Both laser scanner systems were carried by a Piper PA-31 Navajo aircraft with a flight speed of approximately 75 ms^{-1} at an average flying altitude of 800 m above ground level. Furthermore, the scan frequency was 70 Hz and the maximum half scan angle was 7° in both acquisitions resulting in an estimated average footprint diameter of 20 cm. For the ALTM 3100C laser scanner system, pulse repetition frequency (PRF) was 100 kHz resulting in a mean pulse density of 6.8 m^{-2}. A test flight in May 2007 conducted in another area suggested a PRF of 125 kHz for the ALTM Gemini laser scanner system to obtain laser point clouds as similar as possible for the two data acquisitions. Thus, the laser data in northern Norway was collected with a mean pulse density of 8.5 m^{-2} in 2007 using the ALTM Gemini system. The

1,500 km long transect was split into 147 individual flight lines to keep the flying altitude above the terrain and hence the pulse density as constant as possible.

Pre-processing of the laser data was accomplished by the contractor (Blom Geomatics, Norway). For all laser echoes, planimetric coordinates (*x* and *y*) and the ellipsoidal height values were computed.

For the derivation of the terrain model, laser echoes labelled as "last-of-many" and "single" (LAST) were used, whereas laser echoes labelled as "first-of-many" and "single" (FIRST) were used for the analyses in the current thesis. The planimetric coordinates and the corresponding height values of the LAST echoes were used to classify ground echoes based on an iteration distance of 1.0 m with the TerraScan software (Terrasolid, 2011) and a triangulated irregular network (TIN) was computed with an iteration angle of 9°. Furthermore, the FIRST echoes were projected onto the TIN surface to interpolate the corresponding terrain height values on these locations. For the analyses, the height differences between the height values of the FIRST echoes and the corresponding interpolated terrain height values were computed and stored. For the analysis of tree detection in Paper I all FIRST echoes were included, whereas only FIRST echoes with height values greater than zero were included in the classification analyses of Papers II, III and IV since this criterion represents the sole indicator for the presence of objects on the terrain surface.

Both the ALTM 3100C and the ALTM Gemini systems may record up to four echoes per laser pulses with a minimum vertical distance of 2.1 m between two subsequent echoes for the ALTM 3100C. The minimum vertical distance is assumed to be larger for the ALTM Gemini because the vertical resolution is a function of pulse width (cf. Baltsavias, 1999). In combination with low vegetation, this instrument property may result in potentially very few pulses with more than a single echo. Hence, the LAST and FIRST datasets were almost identical for many of the sample plots.

In addition to the field data involving the 111 sample plots allocated along the transect, 54 sample plots were established in the overlap zone for an area-based analysis of potential sensor effects in Paper I. Based on the visual inspection of digital aerial imagery, the 54 sample plots were purposefully selected, each with an area of 1,000 m^2. Three different categories were used, i.e., (1) solitary distributed small-sized trees, (2) sparsely distributed medium-sized trees, and (3) large trees in dense forest stands, to provide are more detailed assessment that is comparable to the dataset of the entire transect.

Table 1 – Summary of field measurements of trees. (* Reproduced with kind permission from CASI)

Tree species	Characteristics	Original dataset Paper I*				Subset Paper IV				Subset Paper II*, III			
		n	Mean	Min	Max	*n*	Mean	Min	Max	*n*	Mean	Min	Max
Mountain birch	Height (m)	623	1.28	0.02	7.80	614	1.27	0.02	7.80	404	1.41	0.04	7.80
	Diameter (cm)	622[a]	3.65	0.10	34.00	613[a]	3.65	0.10	34.00	404	4.24	0.10	34.00
	Crown area (m^2)	623	0.90	0.001	19.54	614	0.91	0.001	19.54	404	1.13	0.001	19.54
Norway spruce	Height (m)	68	1.66	0.07	7.00	67	1.67	0.07	7.00	67	1.67	0.07	7.00
	Diameter (cm)	66[a]	6.53	0.20	19.10	65[a]	6.54	0.20	19.10	65[a]	6.54	0.20	19.10
	Crown area (m^2)	68	1.43	0.006	5.69	67	1.45	0.006	5.69	67	1.45	0.006	5.69
Scots pine	Height (m)	53	1.33	0.10	5.10	53	1.33	0.10	5.10	53	1.33	0.10	5.10
	Diameter (cm)	53	5.00	0.30	18.90	53	5.00	0.30	18.90	53	5.00	0.30	18.90
	Crown area (m^2)	53	0.81	0.002	7.28	53	0.81	0.002	7.28	53	0.81	0.002	7.28

Note: [a] Missing values due to tree properties.

3. Methods

3.1 Computations

For addressing the different research objectives in the four papers, a sequence of computations had to be conducted prior to analyses. Table 2 gives an overview over the different parameters computed and used in the different papers.

Field-measured crown diameters were used for the computation of elliptical tree crown polygons that were involved in the analyses of all four papers. For trees with crown diameter values less than 1.0 m in at least one cardinal direction, tree crown polygons with a fixed radius of 0.5 m were assigned because of a positioning error of the laser data of up to 0.5 m as reported by the contractor.

Furthermore, non-tree polygons were generated for the analyses in Papers II, III and IV utilising the basic properties of the PCQ method, which implicitly provide us with full control of some of the areas without any trees. The sampling design of the PCQ method resulted in a maximum of three sample trees per quadrant and the tree closest to the respective plot centre was selected irrespective of tree height class. Thus, the area defined by the distance between the closest tree in a quadrant and the plot centre was used to compute the non-tree polygons. Finally, the tree crown polygons of the selected trees were erased from the non-tree polygons to ensure full control over the treeless areas.

For Papers II and III, FIRST echoes with laser height values larger than zero were overlaid with the tree and non-tree polygons and the echoes falling inside the particular polygons were classified as tree and non-tree echoes, respectively. This procedure resulted in 2,323 tree and 27,487 non-tree echoes for the supervised classifications.

In Papers II and III, laser intensity values were used as discriminator for the classification analyses. For this purpose, the raw intensity (I_{Raw}) had to be normalised for the range R according to a formula suggested by Korpela et al. (2010).

Moreover, a digital elevation model (DEM) was generated using LP360 (QCoherent Software, 2010) with a cell size of 0.25 m for Papers II and III using LAST echoes classified as ground returns. The terrain-related variables (1) aspect (Paper II) which was divided into eight categories because of computational reasons, and (2) slope (Papers II and III) were derived from the DEM raster surface (Burrough and McDonald, 1998) and their values were assigned to the corresponding FIRST echoes.

To include a geospatial facet in the classification analysis of Paper II, Voronoi polygons were computed using FIRST echoes. For this purpose, adequate areas were defined for the

respective field sites to avoid edge effects at the sample plot borders and the Voronoi polygons were generated individually. The calculated area of each Voronoi polygon was assigned to the corresponding FIRST echo.

For the classification analysis in Paper III, semivariograms were employed as a geostatistical discriminator. By using the mean value of the semivariances of an experimental variogram the differences in the behaviour of spatial correlation of laser height and intensity values for tree and non-tree echoes were characterised.

Furthermore, statistical summary measures were derived for the classification study of Paper III. For this purpose, the arithmetic mean, the standard deviation, as well as the coefficient of variation were derived from both laser height and intensity values, respectively.

Table 2 – Overview of parameters computed and used in the different papers.

Parameter	Used in Paper
Tree crown polygons	I, II, III, IV
Non-tree polygons	II, III, IV
Intensity	II, III
DEM	II, III
Voronoi polygons	II
Mean semivariance from laser height and intensity	III
Arithmetic mean from laser height and intensity	III
Standard deviation from laser height and intensity	III
Coefficient of variation from laser height and intensity	III

3.2 Single tree detection

To investigate the detection success of small individual trees and the effects of different factors influencing the single tree detection, generalised linear models (GLM) and generalised linear mixed models (GLMM) were employed in Paper I. Because of the utilisation of two different laser scanner instruments during data acquisition, the potential sensor effects on detection of small individual trees were assessed prior to analysis employing data from the 80 km overlap zone. Both the tree crown polygons derived from the field-measured crown diameters of the trees located in the overlap zone and the 54 sample plots established in the overlap zone were used to calculate maximum laser-derived tree heights for the single-tree and the area-based approaches, respectively, from both data

acquisitions. Equivalence tests were applied for detected trees with maximum laser heights greater than zero and trees with maximum laser heights greater than or equal to zero. The maximum tree height was calculated for each of the 54 plots of the three different categories and equivalence test were applied for the respective categories testing the null hypothesis for a significant difference between the two data acquisitions.

In the first step of the detection analysis, FIRST echoes were overlaid with the tree crown polygons for evaluation of the echoes located inside the individual polygons. Tree crown polygons with at least one FIRST echo with a maximum laser height greater than zero were classified as being successfully detected. Furthermore, tree crown polygons including FIRST echoes with a maximum laser height equal to zero or no FIRST echo at all were regarded as not detected.

Different independent variables were included in the GLM models representing effects of (1) size of the tree as expressed by tree height and tree crown area, (2) tree species, and (3) geographic location according to latitude and region. Different GLM models were fitted and a likelihood-ratio test was performed to test the null hypothesis that tree species does not significantly contribute to the model. Moreover, another likelihood-ratio test was employed for an additional indicator variable representing the 35 different field sites. The GLM models were evaluated by leave-one-out cross validation and the Hosmer-Lemeshow test statistics (Hosmer and Lemeshow, 2000). Because of a potential contribution of regional effects found for the GLM model, a GLMM model was estimated including the variables used in the GLM analysis as fixed effects and the region parameter representing the 35 field sites as a random effect.

Finally, the accuracy of tree height estimation using laser data was assessed by comparing tree heights estimated from the laser data with the corresponding field-measured tree heights from the detected trees. The differences between the laser-derived and the field-measured tree heights were calculated and assessed by estimating the standard deviation for these differences for the three tree species and the three tree height classes individually.

3.3 Laser echo classification

In Papers II and III, individual FIRST echoes were classified into tree and non-tree echoes using GLM and support vector machines (SVM). Instead of using the maximum laser height in a tree crown polygon as criterion for successful tree detection as in Paper I, the potential of classifying individual FIRST echoes for further usage in a detection analysis was assessed. Different discriminators were tested for their contribution to the classification into

tree and non-tree echoes in Paper II and the best models were expanded with a new suite of discriminators in Paper III to potentially improve the classification in Paper II.

The approach from Paper I employing GLM to assess successful tree detection was pursued by using GLM as a classification method. For this purpose, the models consisting of different combinations of the discriminators were fitted and the probabilities of the FIRST echoes being non-tree echoes were predicted resulting in a classification of tree and non-tree echoes. Furthermore, SVM were introduced to the problem of FIRST echo classification to assess the potential of a relatively new classification method, at least in this research area. The different models were fitted and the FIRST echoes were classified into tree and non-tree echoes.

Leave-one-out cross-validation was employed for the assessment of classification performance of modelling with GLM and SVM. For both classification methods, the Cohen's kappa coefficient (Cohen, 1960) was estimated for each model to assess the classification performance.

In Paper II, kappa coefficients were compared both for the different models using the respective classification method and for the analogue models using the two classification methods (GLM versus SVM). Kappa coefficients were evaluated quantitatively according to the grading suggested by Landis and Koch (1977) and the comparison was conducted using test statistics suggested by Cohen (1960) that evaluates the normal curve deviate to assess the significance of the difference between two independent kappa coefficients. The models consisted of different combinations of laser height and intensity variables, a geospatial variable represented by the area of the Voronoi polygons and the terrain-related variables aspect and slope. The laser height and intensity values were included in all models because of their direct relation to the FIRST echoes.

For a potential improvement of classification performance compared to what was obtained in Paper II, the respective best model of the GLM and SVM classifications was extended in Paper III with the geostatistical parameter represented by the mean semivariance, and statistical measures, i.e., the arithmetic mean, the standard deviation and the coefficient of variation. Thereby, discriminators that revealed a significant improvement when included individually were subsequently combined in a new extended model with all possible combinations to asses a further potential contribution to the discrimination between tree and non-tree echoes. Furthermore, the differences between the independent kappa coefficients were estimated to investigate a superior performance of the models from Paper III in comparison to the simple models from Paper II.

3.4 Automatic detection of small single trees

Originating from the studies conducted in the three previous papers, Paper IV investigated the suitability of an unsupervised classification for automatic detection of small single trees for a potential further utilisation in a monitoring programme. In this context, a concept for a raster-based algorithm was developed using different raster grids with decreasing cell sizes as provided by a region quadtree approach for the classification of tree and non-tree raster cells. To assess the capability of this concept, binary raster grids with different cell sizes adapted to the sample plot size were computed and overlaid with FIRST echoes using six different laser height thresholds for the laser echoes included (0 cm, 10 cm, 20 cm, 30 cm, 40 cm, and 50 cm). Table 3 gives an overview over the different raster cell sizes used with the radius of the sample plot as the initial raster cell side length. Raster cells containing at least one FIRST echo were assigned the value 1 and empty cells the value 0. The grids were intersected with the tree crown polygons in three different categories for assessing the classification performance for tree pixels: (1) all tree crown polygons irrespective of tree height (*I*), (2) tree crown polygons with a tree height equal to or higher than the laser height threshold (*II*), and (3) tree crown polygons with tree heights exceeding 1 m (*III*). For the two laser height thresholds of 20 cm and 30 cm, additional classifications (*IV* and *V*) with different tree heights for the tree crown polygons were conducted because of an underestimation of laser-derived tree heights compared to the corresponding field-measured heights as reported by Næsset and Nelson (2007), Næsset (2009b), and Paper I. Furthermore, the different grids were intersected with the non-tree polygons to obtain the rate of non-tree pixels classified as tree pixels, i.e., commission errors. Finally, commission errors were investigated by intersecting the different grids with the non-tree polygons.

Table 3 - Summary of grid cell sizes.

Number of cells	Cell side length (m)	Cell size (m^2)	Labelling
4	25	625.000	1
16	12.5	156.250	2
64	6.25	39.063	3
256	3.125	9.766	4
1024	1.5625	2.441	5
4096	0.78125	0.610	6
16384	0.390625	0.153	7

4. Results and discussion

4.1 Single tree detection

The study reported in Paper I demonstrated the suitability of high-density ALS data to successfully detect small individual trees in the forest-tundra ecotone and confirmed the findings from previous studies (Næsset and Nelson, 2007; Næsset, 2009b). The analyses of potential sensor effects revealed no significant differences between the two laser scanner systems using the laser scanner data from the 80 km overlap zone in the county of Nordland. The equivalence tests for the small individual tree analysis based on the field-measured tree heights indicated no significant difference in maximum tree height between the two data acquisitions. Furthermore, the equivalence tests for the area-based analysis using the 54 established sample plots confirmed that there is no significant difference between the two sensors for the small and medium-sized trees. Albeit a significant difference was found for the large trees, a potential sensor effect caused by the utilisation of two different laser scanner systems could be neglected in this case because the small individual trees of the dataset are comparable to the trees of the small and medium-sized categories of the area-based analysis. This result was also in line with previous studies using different flight and sensor configurations (Næsset, 2009b). Hence, no data adjustment was performed to calibrate the laser data from the two different sensors in Papers I and IV which employed data from both acquisitions.

For the single tree detection in Paper I, an overall tree detection success rate of 71% was found irrespective of tree height class and tree species (Table 4). Of the 29% of the trees that were not successfully detected, 3% were not hit by any laser pulse at all. Furthermore, the tree species with highest success rate irrespective of tree height class was represented by spruce (81%) and over 90% of the trees exceeding a tree height of 1 m were successfully detected. Less than 11% of all trees >1 m were hit by at least one laser pulse with a laser height equal to zero and the tree species with the lowest amount irrespective of tree height was represented by spruce with 18%. Moreover, only spruce (1%) and birch trees (4%) were not hit by any laser pulse at all, i.e. all pine trees were hit by at least one laser pulse.

Table 4 - Success in tree detection for different trees species and tree height classes. (Reproduced with kind permission from CASI)

Tree species	Height class	*n*	Detected	$h_{max} = 0$	No hits
			n (%)	*n* (%)	*n* (%)
Mountain birch	< 1 m	293	147 (50)	134 (46)	12 (4)
	1-2 m	199	167 (84)	21 (11)	11 (5)
	> 2m	131	127 (97)	3 (2)	1 (1)
	total	623	441 (71)	158 (25)	24 (4)
Norway spruce	< 1 m	20	9 (45)	11 (55)	0 (0)
	1-2 m	24	22 (92)	1 (4)	1 (4)
	> 2m	24	24 (100)	0 (0)	0 (0)
	total	68	55 (81)	12 (18)	1 (1)
Scots pine	< 1 m	30	11 (37)	19 (63)	0 (0)
	1-2 m	12	12 (100)	0 (0)	0 (0)
	> 2m	11	10 (91)	1 (9)	0 (0)
	total	53	33 (62)	20 (38)	0 (0)
Total	total	744	529 (71)	190 (26)	25 (3)

The GLM analyses indicated that the probability of a tree being successfully detected increased with an increase in tree height and tree crown area in particular. Although being highly correlated, those two independent variables showed high significance in all models. A slightly higher probability for successful tree detection was found for increasing latitude accompanied by a slight inter-correlation with tree species involving a potential mixed effect of those two. This potential mixed-effect of latitude and tree species might result from the lack of spruce and pine trees north of 66° N as reflected in the field data. This together with the assumption that local terrain might influence the terrain model and therewith the success in tree detection, led to the creation of a subset of the field and laser data. Only areas covered by the ALTM 3100C sensor were included to obtain a more balanced dataset for the classification of tree and non-tree echoes in Papers II and III. Furthermore, spruce and pine trees revealed slightly higher probabilities of being detected than birch. Employing the predicted probabilities of being successfully detected, almost all conifer trees were found provided a minimum tree height of 1.3–1.4 m, whereas a birch tree had to reach a tree height of at least 1.9 m. For tree crown areas, almost all pine trees were detected when reaching a size of 0.5 m^2, and almost all spruce and birch trees were found provided a tree crown area of 1.1 m^2. Figure 3 illustrates the predicted probabilities of a tree for being hit by at least one laser pulse for tree height and tree crown area for the three tree species. This result from the

GLM analysis that almost every tree exceeding a height of 1 m was successfully detected could be integrated in the raster-based detection concept of Paper IV for verification.

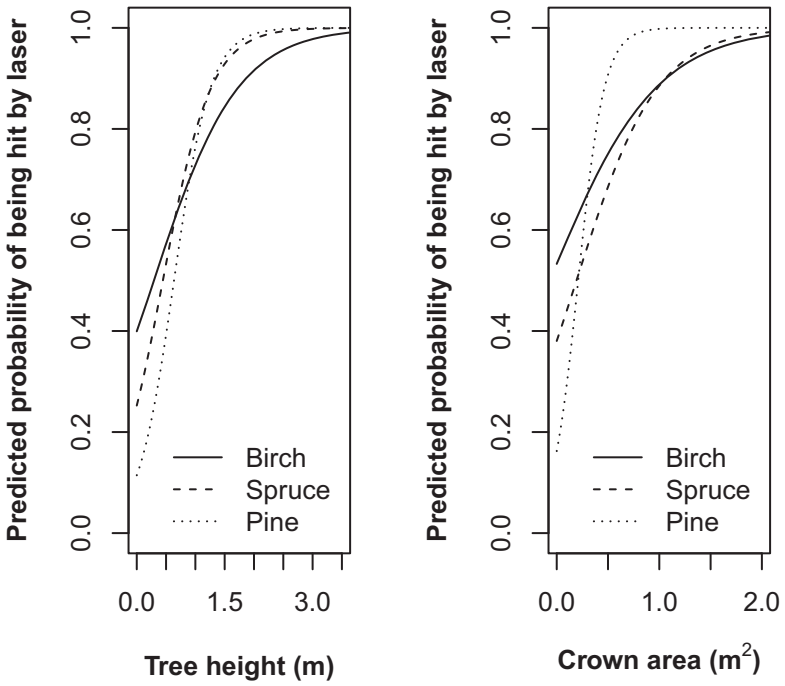

Figure 3 - Predicted probability of a tree being hit by at least one laser pulse as a function of tree height (left) and tree crown area (right) for birch, spruce, and pine. (Reproduced with kind permission from CASI)

Furthermore, a likelihood-ratio test indicated no significant difference in successful tree detection between pine and spruce trees, however, no simplification of the model was conducted because all indicator variables representing tree species were statistically significant. Another likelihood-ratio test revealed a significant difference for the probability of tree detection between the 35 field sites leading to an extension of the basic model. The Hosmer-Lemeshow test statistics showed no lack of fit for any of the models and the leave-one-out cross validation indicated an overall agreement of over 80% between observed and predicted values for both models as well.

The GLMM analysis revealed the same results for the fixed effects as in the GLM analysis with only minor changes in the levels of significance for most of the independent variables. Furthermore, the random effect indicated some variations in the models that were

caused by locality of the respective field sites. The presence of this regional effect also corroborated a reduction of the dataset for the analyses conducted in Papers II and III.

The accuracy assessment of the laser-derived tree heights revealed a general trend of greater underestimations for larger trees with a concurrent lower precision. The greatest underestimation errors were found for pine trees, however, coinciding with the lowest variation in the estimated heights. These underestimations found for the laser-derived tree heights were concordant with previous studies using small pioneer trees (Næsset and Nelson, 2007; Næsset, 2009b), and were accounted for in the analyses in Paper IV. Overestimations were found for few individual trees and may be traced back to the special characteristics of small birch and spruce trees that tend to grow in groups with several trees of different heights.

4.2 Laser echo classification

Papers II and III verified the capability of high-density ALS data to classify tree and non-tree echoes directly from the laser point cloud using different variables and classification methods. In both studies, tree and non-tree echoes were classified with a total accuracy of at least 93% and a moderate fit irrespective of the classification method or the model used. This overall result is in accordance with other studies detecting small trees on an individual tree basis in a forest-tundra environment, albeit using individual laser echoes in Papers II and III. Paper I, as well as the studies conducted by Næsset and Nelson (2007) and Næsset (2009b) reported success rates of at least 90% for trees exceeding 1 m in height.

The classification in Paper II of the FIRST echoes into tree and non-tree echoes using GLM and SVM with variables represented by laser height and intensity, the area of Voronoi polygons, and the terrain variables aspect and slope revealed total accuracies of at least 93.6% and 94.8%, respectively (Table 5). Kappa coefficients of at least 0.515 for GLM and 0.560 for SVM indicated moderate fits for the estimated models (Table 5).

For the GLM classification in Paper II, a maximum difference of 1.3 percentage points was found for the eight different models, of which models not including the terrain variable slope had slightly higher accuracies. The model with the highest accuracy (94.9%) included only laser height and intensity. The corresponding kappa coefficient was 0.573, which was the highest as well. Furthermore, no significant effect on the classification performance could be found for the geospatial variable and the terrain variables.

The classification of the FIRST echoes using SVM had a maximum difference of 0.5 percentage points between the eight models. Models not including the terrain variable slope

had slightly lower accuracies, however, with no serious inferior performance. The model consisting of laser height and intensity as well as the terrain variable slope was the model with the highest accuracy of 95.3% accompanied by the highest kappa coefficient of 0.600. Significantly higher accuracies were found for models including the slope variable, whereas the geospatial variable and the terrain variable aspect had no significant influence on the classification performance.

Laser height and intensity values were found to be important discriminators for the classification of tree and non-tree echoes in Paper II, especially when included as the two sole discriminators using GLM. The geospatial variable represented by the Voronoi polygons, however, did not contribute significantly to classification performances both with regard to GLM and SVM. Therewith, this variable can be disregarded as discriminator for FIRST echo classification, albeit Thieme et al. (2011) reported promising results for this geospatial variable in combination with laser height and intensity values for the detection of tree and non-tree objects in a forest-tundra ecotone environment. For the terrain variables aspect and slope, no significant contribution was found using GLM. In other studies, the terrain variable aspect was found to have an important influence on tree invasion (Mast et al., 1997; Danby and Hik, 2007), however, using the variable on a coarser scale compared to Paper II. Furthermore, the terrain variables aspect and slope have a similar ecological meaning describing amongst others the solar radiation and moisture (Bader and Ruijten, 2008). These essential factors for the occurrence of trees may already be covered by the aspect variable in this linear classification method.

When comparing the classification performances of GLM and SVM for the different models in Paper II, no significant differences were found between models including different combinations of laser height and intensity, the geospatial variables and the terrain variable aspect (Table 5). However, the *p*-values for all models comprising the terrain variable slope indicated a significantly better classification performance for SVM. Thus, slope represented a significant discriminator for the classification of tree and non-tree echoes using SVM and might be a significant contributor to FIRST echo classification when utilising a nonlinear kernel to construct hyperplanes during SVM classification. Furthermore, it is reasonable to assume that a nonlinear classification is sensitive to the slope variable.

Table 5 – Performance and comparison of the different models used for laser echo classification with GLM and SVM. (Reproduced with kind permission from CASI)

Model[a]	GLM		SVM		GLM versus SVM
	Accuracy	Kappa	Accuracy	Kappa	Z^b
HI	0.949	0.573	0.949	0.568	0.428
HIP	0.948	0.569	0.949	0.569	0.004
HIA	0.948	0.568	0.948	0.560	0.570
HIS	0.942	0.550	0.953	0.600	3.719 **
HIPA	0.946	0.559	0.948	0.563	0.268
HIPS	0.943	0.546	0.952	0.594	3.514 **
HIAS	0.939	0.523	0.952	0.594	5.165 **
HIPAS	0.936	0.515	0.952	0.596	5.930 **

Note: Level of significance: *<.05. **<.005

[a] *H*=Laser height, *I*=Laser intensity, *P*=Voronoi polygons, *A*=Aspect, *S*=Slope.

[b] As received by the comparison between two independent kappa coefficients.

The results of Paper III revealed significant improvements by extending the best simple GLM and SVM models from Paper II with geostatistical and statistical measures. For the classification of FIRST echoes into tree and non-tree echoes using GLM and SVM models consisting of geostatistical and statistical measures in Paper III, total accuracies of at least 93.6% and 94.7% were found, respectively. Compared to the classification performances in Paper II, the kappa coefficients were improved in Paper III by at least 0.032 and 0.034 using GLM and SVM, respectively (Table 6).

Using GLM, a maximum difference of 1.3 percentage points was found between the total accuracies of the different models consisting of geostatistical and statistical measures derived from the laser height and intensity values. However, models including the geostatistical and/or the statistical measures represented by the arithmetic mean and the standard deviation derived from laser height values revealed higher classification performances as expressed by the kappa coefficients. The highest kappa coefficient of 0.606 was revealed by the two models consisting of the arithmetic mean, and the mean semivariance and the arithmetic mean, respectively, indicating a moderate fit. Furthermore, no significant contribution was found for the geostatistical and statistical measures derived from laser intensity values when compared to the simple model of Paper II that involved laser height and intensity values. For the geostatistical and statistical measures derived from laser height values, however, a significant improvement was found for all models that included the mean semivariance and the arithmetic mean individually and in combination compared to the simple model from Paper II using the Cohen's test statistics (Table 6).

For SVM, the total accuracies differed by 1.0 percentage points between the twelve models. Most models including geostatistical and statistical measures derived from the laser height values revealed slightly higher accuracies. Kappa coefficients for models consisting of the mean semivariance, the arithmetic mean, and the standard deviation derived from the laser height values were higher than for the other models, both when included individually or in combination with one another. The model merely consisting of the mean semivariance in addition to the simple model from Paper II revealed the highest kappa coefficient of 0.666 indicating a substantial fit. By comparing the kappa coefficients of the models to the simple model from Paper II, no significant contribution was found for any of the models with geostatistical and statistical measures derived from the laser intensity values. Using the laser-derived geostatistical and statistical measures, six models including the mean semivariance, the arithmetic mean, and the standard deviation individually or in combination with one another, the simple model of Paper II was improved (Table 6). The classification performance was improved by at least 0.034.

Geostatistical and statistical measures derived from laser intensity values revealed no significantly contribution to any model using GLM or SVM. This can be traced back to the respective distributions of the different measures that were similar for tree and non-tree echoes resulting in no or a very weak discriminating effect. Furthermore, standard deviations and coefficient of variations derived from laser height values revealed similar distributions and did therewith not suggest a discriminating effect. A positive effect of the nonlinear classification method SVM was found for the laser height-derived standard deviation when used individually or in combination with the mean semivariance or the arithmetic mean from laser height values. Moreover, obvious differences in the value distributions of the arithmetic mean and the mean semivariance derived from laser height values were found for tree and non-tree echoes. This was reflected in the significant improvements of the simple models from Paper II both using GLM and SVM when extended with these two discriminators individually or in combination with one another. The superior performance of the mean semivariance both using GLM and SVM is in line with Thieme et al. (2011) who employed experimental variograms to characterise and distinguish between tree and non-tree objects in the forest-tundra ecotone. Furthermore, variogram-based measures were found to be beneficial for the classification of vegetation classes including grassland, rocks and woodland, however, using optical airborne imagery (Jakomulska and Clarke, 2001).

Table 6 – GLM and SVM models from Paper III that significantly improved the simple GLM and SVM models from Paper II.

Model[a]		Accuracy	Kappa	Z[b]
GLM	HI_H_{SV}	0.947	0.605	2.333 *
	HI_H_{AM}	0.946	0.606	2.482 *
	$HI\_H_{SV}\_H_{AM}$	0.946	0.606	2.480 *
SVM	HIS_H_{SV}	0.957	0.666	4.995 **
	HIS_H_{AM}	0.956	0.655	4.183 **
	HIS_H_{SD}	0.957	0.660	4.539 **
	$HIS\_H_{SV}\_H_{AM}$	0.955	0.643	3.186 **
	$HIS\_H_{SV}\_H_{SD}$	0.957	0.664	4.875 **
	$HIS\_H_{AM}\_H_{SD}$	0.954	0.634	2.556 *

Note: Level of significance: *<.05. **<.005.
[a] *H*=Laser height, *I*=Laser intensity, *S*=Slope, H_{SV}=Mean semivariance derived from laser height, H_{AM}=Arithmetic mean derived from laser height, *HIS*=Laser height and intensity and slope, H_{SD}=Standard deviation derived from laser height.
[b] As received by the comparison between two independent kappa coefficients.

4.3 Automatic single tree detection

In Paper IV, the suitability of the unsupervised classification for automatic detection of small individual trees revealed promising results for the classification into tree and non-tree raster cells. The different raster cell sizes demonstrated different usages and suitability in an algorithmic context based on the detection success rates of the tree raster cells. In all classifications, the three largest raster cell sizes (Table 3) were found to be well suited for the exclusion of areas without any trees, however, being too large in size for reliable tree raster cell detection because of the very imprecise description of positioning for the small trees. Raster cell size *6* (Table 3) that corresponds to approximately half the size of the mean tree crown area, however, represented an optimal raster cell size in terms of a relatively precise tree positioning and a still satisfying detection success rate (Figure 4). Furthermore, raster cell size *7* revealed a significant decrease in the rate of successfully detected tree raster cells (Figure 4) indicating the inadequacy of raster cell sizes considerably smaller than half of the mean tree crown area for tree raster cell detection. For the classifications *I*, *II*, and *III*, large raster cell sizes generally resulted in high detection success rates for the tree raster cells which decreased with decreasing raster cell size. The success rates were further inversely proportional to the laser height thresholds and a decrease in raster cell size was accompanied by a decrease in differences of the success rates between the different laser height thresholds (Figure 3). Also the classifications accounting for the underestimation of laser-derived tree heights (*IV* and *V*) revealed a decrease in success rate when decreasing the raster cell size.

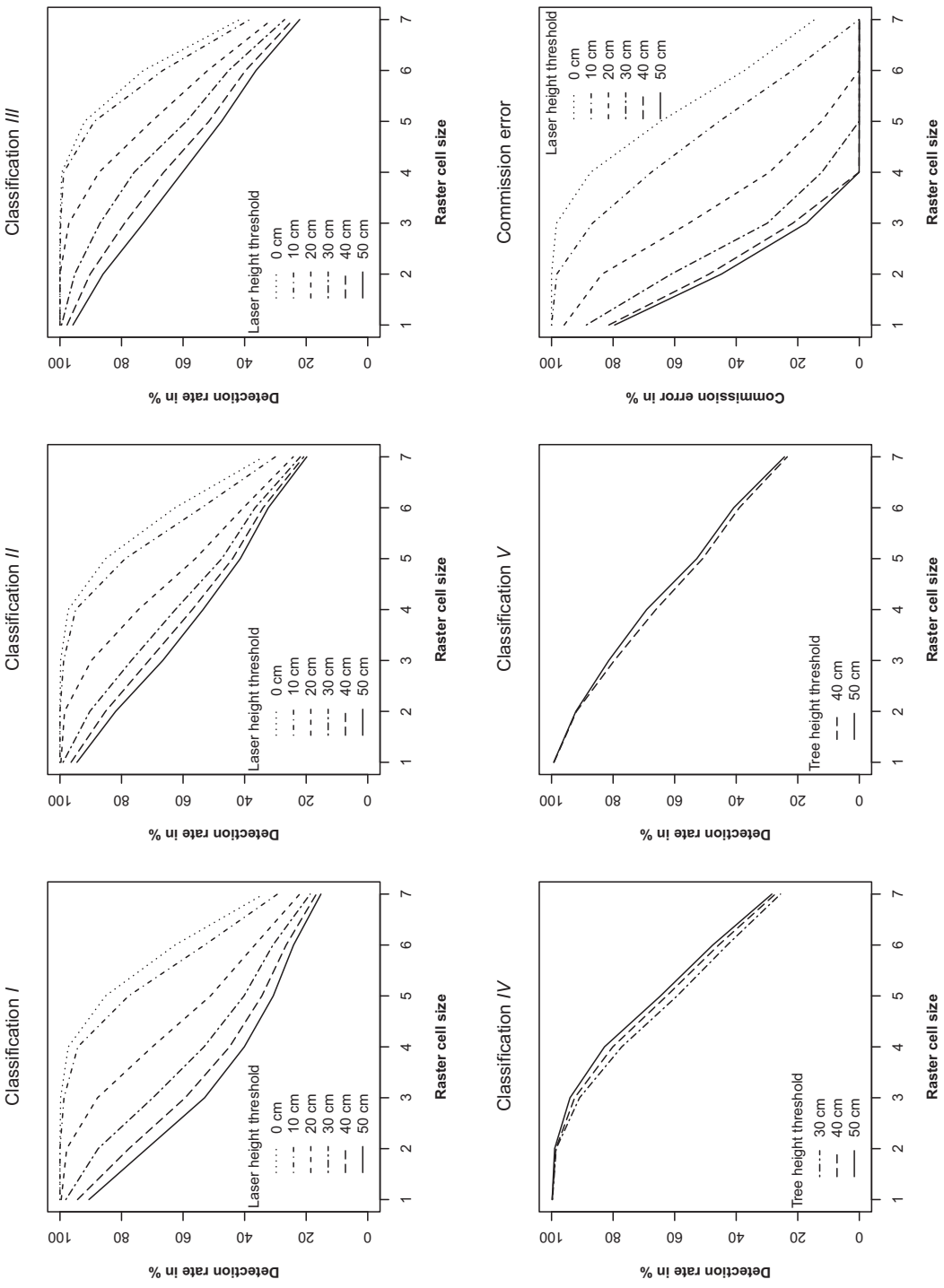

Figure 4 – Detection rate for different raster cell sizes for classification I-V and commission errors for different laser height thresholds.

Rates for non-tree raster cells classified as tree raster cells for the optimal raster cell size *6* ranged between 0.01% and 37.3% depending on the laser height threshold (Figure 4). Higher rates of commission error were found for the lower laser height thresholds and decreased with an increase in laser height threshold. For the laser heights thresholds of 20 cm and 30 cm, the sudden decrease in commission errors for the respective raster cell sizes *3* to *6* (Figure 4) suggested a diminution in the laser data noise. Together with almost non-existent decrease in the rate of commission errors for higher laser thresholds, this result confirmed the assumption of such an upper limit that was observed during data processing.

Depending on the laser height threshold and using the optimal raster cell size *6*, at least 24.1% and up to 62.6% of the tree pixels were detected including all tree crown polygons irrespective of tree height (classification *I*) and 32.3% up to 62.6% when tree crown polygons with tree heights equal to or higher than the laser threshold (classification *II*) were employed (Figure 3). The rates of successful tree detection ranged between 36.3% and 72.9% for tree crown polygons with tree heights exceeding 1 m (classification *III*) using different thresholds based on the optimal raster cell size *6*. These results revealed slightly lower detection rates than in Paper I or other studies on small single tree detection in such an environment as the forest-tundra ecotone (Næsset and Nelson, 2007; Næsset, 2009b). For the classifications *IV* and *V* that were taking into account the underestimation of laser-derived tree heights, at least 39.0% of the tree pixels were successfully classified using the optimal raster cell size *6* and higher success rates were found for higher tree height thresholds. The differences between the success rates for the different tree height thresholds were small or almost equal to zero for the classifications, albeit a decrease in raster cell size resulted in a decreasing success rate. Moreover, the influence of potential underestimation of field-measured tree height using ALS data as demonstrated in Paper I and in other studies (Næsset and Nelson, 2007; Næsset, 2009b) was reflected in an increase of successful tree raster cell detection when increasing the height of trees included in the classification.

For monitoring purposes, parameters values for raster cell sizes, laser height thresholds, and a potential lower tree limit are critical with regard to a justifiable trade-off between rates of successful detection and commission errors as reflected by the results of the different classifications. Also the usage of different sensors and acquisition settings over time is challenging from a monitoring perspective (Næsset, 2009b) and may result in varying laser point densities. Since the probability of a tree being hit by at least one laser pulse generally is a function of laser point density, low densities may not be suitable for single tree detection in the forest-tundra ecotone, whereas almost all trees ≥1 m in height were hit by at least one

laser pulse in the studies conducted in Paper I and by Næsset and Nelson (2007) using ALS data with high laser point densities (6.8-8.5 m^2). In the analysis in Paper IV, the sudden decrease of commission errors for the laser height thresholds of 20 cm and 30 cm suggested a relatively large proportion of data noise for laser echoes lower than these thresholds. This may result in severe commission errors for unsupervised classifications only using laser height values because of the limited capability to distinguish between the smallest trees and other types of vegetation equal in height, such as for example shrubs. Furthermore, the detection of regeneration and mortality of small individual trees is strongly dependent on a sufficient time span between data acquisitions. The limited height growth of such trees, i.e., for example between 1 and 10 cm depending on locality and the prevailing microclimate, may require longer time spans for a raster-based automatic detection algorithm, especially with regard to tree establishment.

5. Conclusion

To conclude, this thesis provides an in-depth investigation of the potential of high-density ALS data to assist in detection of small individual trees in the forest-tundra ecotone. First, successful single tree detection using laser height values in combination with tree characteristics and spatial influences was verified for trees exceeding a height of 1 m in Paper I. The results form a solid basis for generalisation and inference that goes far beyond previous research because of the huge geographical extension of the dataset. Secondly, the capability of the ALS data for classification into tree and non-tree echoes using laser measurements, geospatial and terrain variables was confirmed using two different modelling techniques in Paper II. Furthermore, an extension of the classification models with geostatistical and statistical measure revealed a significant improvement in Paper III. Concerning the automatic detection of small single trees using unsupervised classification in Paper IV, suitable initial values for raster cell sizes resulting in the exclusion of large treeless areas as well as an optimal raster cell size with a still satisfying detection success rate were determined and a laser height threshold revealing a significant decrease in the rate of commission error was identified.

For the utilisation of the supervised and unsupervised classifications presented in Papers II, III and IV for monitoring purposes, different challenges have to be met and investigated in future research. Depending on the time span, the usage of different sensors and acquisition settings resulting in varying laser point densities has to be analysed involving field surveys

that have to be obtained for all ALS acquisitions (Næsset, 2009b). Employing the methods suggested in Paper IV, raster grids can be applied as map products presenting the variation in tree occurrence and distribution over time provided a sufficient time span. Furthermore, the selection of a sufficient time span between the different inventories is entailed by tree growth, regeneration and mortality as well as positional accuracy of individual echoes with regard to the methods used in Papers II and III. The aggregation of the individual classified echoes to another scale, e.g. individual tree level or a raster in different scales, is depending on the specific needs of a monitoring programme. Finally it should be mentioned that the basic techniques used in this study may be implemented in a sampling framework as in Falkowski et al. (2009) and Gobakken et al. (2012) using the laser as a sampling device to provide a time- and cost-efficient monitoring tool for regional estimates of changes in the tree line and abundance of small trees in the forest-tundra ecotone or to provide map products as in Ørka et al. (2012).

References

ACIA. 2004. *Impacts of a warming Arctic: Arctic Climate Impact Assessment*. Cambridge University Press, Cambridge, UK. 146 pp.

Aune, S., Hofgaard, A., and Söderström, L. 2011. Contrasting climate- and land-use-driven tree encroachment patterns of subarctic tundra in northern Norway and the Kola Peninsula. *Canadian Journal of Forest Research*, Vol. 41, pp. 437–449.

Bader, M.Y., and Ruijten, J.J.A. 2008. A topography-based model of forest-cover at the alpine tree line in the tropical Andes. *Journal of Biogeography*, Vol. 35, pp. 711–723.

Baltsavias, E.P. 1999. Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 54, pp. 199–214.

Battlori E., and Gutiérrez E. 2008. Regional tree line dynamics in response to global change in the Pyrenees. *Journal of Ecology*, Vol. 96, No. 6, pp. 1275–1288.

Boisvenue, C., Temesgen, H., and Marshall, P. 2004. Selecting a small tree height growth model for mixed-species stands in the southern interior of British Columbia, Canada. *Forest Ecology and Management*, Vol. 202, pp. 301–312.

Boots, B.N., and Getis, A. 1988. *Point Pattern Analysis*. Scientific Geography Series, Volume 8, Sage Publications, Newbury Park. 93 pp.

Brandtberg, T., Warner, T.A., Landenberger, R.E., and McGraw, J.B. 2003. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density LIDAR data from the eastern deciduous forest in North America. *Remote Sensing of Environment*, Vol. 85, No. 3, pp. 290–303.

Brennan, R., and Webster, T.L. 2006. Object-oriented land cover classification of lidar-derived surfaces. *Canadian Journal of Remote Sensing*, Vol. 32, No.2, pp. 162–172.

Burrough, P.A., and McDonell, R.A. 1998. *Principles of Geographical Information Systems*. Oxford University Press, New York. 356 pp.

Callaghan, T.V., Werkman B.R., and Crawford, R.M.M. 2002. The tundra-taiga interface and its dynamics: Concepts and applications. *Ambio Special Report*, Vol. 12, pp. 6–14.

Cairns, D.M., and Moen, J. 2004. Herbivory influences tree lines. *Journal of Ecology*, Vol. 92, No. 6, pp. 1019–1024.

Chasmer, L., Hopkinson, C., Smith, B., and Treitz, P. 2006. Examining the influence of changing laser pulse repetition frequencies on conifer forest canopy returns. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 12, pp. 1359–1367.

Clements, F.E. 1905. *Research methods in ecology*. University Publishing Company, Lincoln, Nebraska. 334 pp.

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Management*, Vol. 20, No. 1, pp. 37–46.

Cottam, G., and Curtis, J.T. 1956. The use of distance measures in phytosociological sampling. *Ecology*, Vol. 37, No. 3, pp. 451–460.

Danby, R.K., and Hik, D.S. 2007. Variability, contingency and rapid change in recent subarctic alpine tree line dynamics. *Journal of Ecology*, Vol. 95, pp. 352– 363.

Falkowski, M.J., Wulder, M.A., White, J.C., and Gillis, M.D. 2009. Supporting large-area, sample-based forest inventories with very high spatial resolution satellite imagery. *Progress in Physical Geography*, Vol. 33, No. 3, pp. 403-423.

Farid, A., Goodrich, D.C., and Sorooshian, S. 2006a. Using airborne LIDAR to discern age classes of cottonwood trees in a riparian area. *Western Journal of Applied Forestry*, Vol. 21, No. 3, pp. 149–158.

Farid, A., Rautenkranz, D., Goodrich, D.C., Marsh, S.E., and Sorooshian, S. 2006b. Riparian vegetation classification from airborne laser scanning data with an emphasis on cottonwood trees. *Canadian Journal of Remote Sensing*, Vol. 32, No. 1, pp. 15−18.

Fry, D.L., and Stephens, S.L. 2010. Stand-level spatial dependence in an old-grwoth Jeffrey pine – mixed conifer forest, Sierra San Pedro Mártir, Mexico. *Canadian Journal of Forest Research*, Vol. 40, pp. 1803–1814.

Gobakken, T., Næsset, E., Nelson, R., Bollandsås , O.M., Gregoire, T.G., Ståhl, G., Holm, S., Ørka, H.O., and Astrup, R. 2012. Estimating biomass in Hedmark County, Norway using national forest inventory field plots and airborne laser scanning. *Remote Sensing of Environment*. Vol. 123, pp. 443–456.

Harper, K.A., Danby, R.K., De Fields, D.L., Lewis, K.P., Trant, A.J., Starzomski, B.M., Savidge, R., and Hermanutz, L. 2011. Tree spatial pattern within the forest-tundra ecotone: a comparison of sites across Canada. *Canadian Journal of Forest Research*, Vol. 41, pp. 479–489.

Hofgaard, A., Løkken, J.O., Dalen, L., and Hytteborn, H. 2010. Comparing warming and grazing effects on birch growth in an alpine environment – a 10-year experiment. *Plant Ecology & Diversity*, Vol. 3, No. 1, pp. 19–27.

Holmgren, J., Persson, A., and Söderman, U. 2008. Species identification of individual trees by combining high resolution LIDAR data with multi-spectral images. *International Journal of Remote Sensing*, Vol. 29, No. 5, pp. 1537−1552.

Holtmeier, F.-K., and Broll, G. 2005. Sensitivity and response of northern hemisphere altitudinal and polar treelines to environmental changes at landscape and local scales. *Global Ecology and Biogeography*, Vol. 14, pp. 395–410.

Holtmeier, F.-K., and Broll, G. 2007. Treeline advance – driving processes and adverse factors. *Landscape Online*, Vol. 1, pp. 1–33.

Hosmer, D.W., and Lemeshow, S. 2000. *Applied Logistic Regression*. Wiley & Sons, Inc., Hoboken, New Jersey, USA. 392 pp.

Hyyppä, J., Kelle, O., Lehikoinen, M., and Inkinen, M. 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 5, pp. 969–975.

IPCC. 2007. *Climate Change 2007: The Physical Science Basis – Summary for Policymakers*. 10th session of Working group 1 of the IPCC. IPCC, Paris, France.

Jakomulska, A., and Clarke, K. C. 2001. Variogram-derived measured of textural image classification: Application to large-scale vegetation mapping. In P. Monestiez, D. Allard & R. Froidevaux (Eds.), *geoENV III – Geostatistics for Environmental Applications* (pp. 345–355). Kluwer Academic Publishers, Dordrecht.

Kaasalainen, S., Ahokas, E., Hyyppa, J., and Suomalainen, J. 2005. Study of surface brightness from backscattered laser intensity: Calibration of laser data. *IEEE Geoscience and Remote Sensing Letters*, Vol. 2, No. 3, pp. 255–259.

Korpela, I., Ørka, H.O., Maltamo, M., Tokola, T., and Hyyppä, J. 2010. Tree species classification using airborne LiDAR – effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica*, Vol. 44, No. 2, pp. 319–339.

Kullman, L. 2002. Rapid recent range-margin rise of tree and shrub species in the Swedish Scandes. *Journal of Ecology*, Vol. 90, pp. 68–77.

Kullman, L., and Öberg, L. 2009. Post-Little Ice Age tree line rise and climate warming in the Swedish Scandes: a landscape ecological perspective. *Journal of Ecology*, Vol. 97, No. 3, pp. 415–429.

Landis, J.R., and Koch, G.G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, Vol. 33, No. 1, pp. 159–174.

Mast, J.N., Veblen, T.T., and Hodgson, M.E. 1997. Tree invasion within a pine/grassland ecotone: an approach with historic aerial photography and GIS modeling. *Forest Ecology and Management*, Vol. 93, pp. 181–194.

Næsset, E. 2005. Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*, Vol. 98, pp. 356–370.

Næsset, E. 2009a. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*, Vol. 113, pp. 148–159.

Næsset, E. 2009b. Influence of terrain model smoothing and flight and sensor configurations on detection of small pioneer trees in the boreal-alpine transition zone utilizing height metrics derived from airborne scanning lasers. *Remote Sensing of Environment*, Vol. 113, pp. 2210–2223.

Næsset, E., and Nelson, R. 2007. Using airborne laser scanning to monitor tree migration in the boreal-alpine transition zone. *Remote Sensing of Environment*, Vol. 110, pp. 357–369.

Olofsson, J., Oksanen, L., Callaghan, T., Hulme, P.E., Oksanen, T., and Suominen, O. 2009. Herbivores inhibit climate-driven shrub expansion on the tundra. *Global Change Biology*, Vol. 15, No. 11, pp. 2681–2693.

Ørka, H.O., Næsset, E., Bollandsås, O.M. 2009. Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. *Remote Sensing of Environment*, Vol. 113, pp. 1163–1174.

Ørka, H.O., Wulder, M.A., Gobakken, T., and Næsset, E. 2012. Subalpine zone delineation using LiDAR and Landsat imagery. *Remote Sensing of Environment*, Vol. 119, pp. 11–20.

Persson, Å., Holmgren, J., and Söderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, Vol. 68, No. 9, pp. 925–932.

Post, E., and Pedersen, C. 2008. Opposing plant community responses to warming with and without herbivores. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 105, No. 34, pp. 12353–12358.

QCoherent Software 2010. *Getting started with LP360*. QCoherent Software, Colorado Springs, USA. 12pp. Available from www.qcoherent.com [accessed 08 February 2012].

Rees, W.G. 2007. Characterisation of arctic treelines by LiDAR and multispectral imagery. *Polar Record*, Vol. 43, No. 227, pp. 345–352.

Rossi, R.E., Mulla, D.J., Journel, A.G., and Franz, E.H. 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecological Monographs*, Vol. 62, No. 2, pp. 277–314.

Schreier, H., Lougheed, J., Tucker, C., and Leckie, D. 1985. Automated measurements of terrain reflection and height variations using an airborne infrared laser system. *International Journal of Remote Sensing*, Vol. 6, No. 1, pp. 101−113.

Solberg, S., Næsset, E., and Bollandsås, O.M. 2006. Single tree segmentation using airborne laser scanner data in a heterogeneous spruce forest. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 12, pp. 1369−1378.

Sturm, M., Schimel, J., Michaelson, G., Welker, J.M., Oberbauer, S.F., Liston, G.E., Fahnestock, J., and Romanovsky, V.E. 2005. Winter biological processes could help convert arctic tundra to shrubland. *Bioscience*, Vol. 55, No. 1, pp. 17–26.

Terrasolid. 2011. *TerraScan User's Guide*. Terrasolid Ltd., Jyväskylä, Finland. 311 pp. Available from www.terrasolid.fi [accessed 26 September 2011].

Thieme, N., Bollandsås, O.M., Gobakken, T., and Næsset, E. 2011. Assessing spatial variation for tree and non-tree objects in a forest-tundra ecotone in airborne laser scanning data. In *SilviLaser 2011: 11th International Conference on LiDAR Applications for Assessing Forest Ecosystems*. 16-20 October 2011, Hobart, Australia. pp. 325–332. Available from http://www.iufro.org/science/divisions/division-4/40000/40200/40205/publications/ [accessed 1 February 2012].

UNFCCC. 2008. *Kyoto protocol reference manual on accounting of emissions and assigned amount*.

Warde, W., and Petranka, J.W. 1981. A correction factor table for missing point-center quarter data. *Ecology*, Vol. 62, No. 2, pp. 491–494.

Wulder, M. A., LeDrew, E. F., Franklin, S. E. & Lavigne, M. B. 1998. Aerial image texture information in the estimation of northern deciduous and mixed wood forest leaf area index (LAI). *Remote Sensing of Environment*, Vol. 64, pp. 64–76.

PAPER I

Copyright Notice

Can. J. Remote Sensing, Vol. 37, No. 3, pp. 264–274, 2011

Detection of small single trees in the forest–tundra ecotone using height values from airborne laser scanning

Nadja Thieme, Ole Martin Bollandsås, Terje Gobakken, and Erik Næsset

Abstract. Because of global warming, it is assumed that the arctic and alpine tree lines will advance northwards into the tundra and upwards into mountainous regions. Methods are needed to monitor these advances. Airborne laser scanning has recently been introduced for detection of small pioneer trees that form the advanced alpine tree line. The objective of this study was to analyze the capability of high-density airborne laser scanning data used for detecting such individual small trees in the transition between the mountain forest and the alpine zone, the forest–tundra ecotone. The study used field and laser data collected along a 1500 km transect stretching from northern Norway (69°3′ N) down to the southern part of the country (58°3′ N). In the field, 744 trees of mountain birch, Norway spruce, and Scots pine were geolocated with centimetre accuracy, and they were measured for height, root collar diameter, and crown diameter. Tree heights ranged between 0.02 and 7.80 m. The laser data were acquired in two separate acquisitions with mean pulse densities of 6.8 m^{-2} and 8.5 m^{-2}, respectively. Laser echoes with relative height values greater than zero within the individual tree crown polygons were used as a criterion for a successful tree detection. The detection success for trees taller than 1 m was 90%; however, for trees shorter than 1 m, the corresponding value was 49%. The highest detection success was found for spruce. Generalized linear models and a generalized linear mixed model with binary responses (detected/not detected) were applied to evaluate the effects of tree height, tree crown area, tree species, geographic location along the latitude gradient, and region on successful detection. Although they were highly correlated, tree height and tree crown area turned out to be the variables showing high significance ($p \leq 0.001$) in all of these models.

Resume. En raison du réchauffement climatique, les limites forestières arctique et alpine devraient progresser vers le nord dans la toundra et en plus haute altitude dans les régions montagneuses. Il est donc nécessaire de développer des méthodes afin de pouvoir faire le suivi de ces avancées. La technique de balayage laser aéroporté a été introduite récemment pour détecter ces petits arbres pionniers qui forment la frange avancée de la limite forestière alpine. L'objectif de cette étude était d'analyser la capacité des données laser aéroporté haute densité utilisées pour détecter de tels petits arbres individuels dans la zone de transition entre la forêt de montagne et la zone alpine appelée l'écotone forêt–toundra. Dans cette étude, on a utilisé des données de terrain ainsi que des données laser acquises le long d'un transect de 1500 km de longueur s'étendant du nord de la Norvège (69°3' N) jusqu'au sud du pays (58°3' N). Sur le terrain, 744 bouleaux montagne épinettes de Norvège et pins sylvestres ont été géolocalisés avec une précision au centimètre et mesurés pour la hauteur, le diamètre du collet et le diamètre de la couronne. Les hauteurs d'arbres variaient de 0,02 à 7,80 m. Les données laser ont été acquises au cours de deux campagnes d'acquisition avec des densités d'impulsion moyennes respectivement de 6,8 m^{-2} et de 8,5 m^{-2}. Les échos laser avec des valeurs relatives de hauteur plus grands que zéro à l'intérieur des polygones individuels de couronnes d'arbre ont été utilisés comme critère pour assurer le succès de la détection d'arbres. Le taux de succès pour la détection d'arbres d'une taille supérieure à un mètre était de 90 % tandis que pour les arbres de moins d'un mètre, la valeur correspondante était de 49 %. Le taux de succès de détection le plus élevé était observé pour l'épinette. Des modèles linéaires généralisés et un modèle linéaire généralisé mixte avec des réponses binaires (détecté/non détecté) ont été appliqués afin d'évaluer les effets de la hauteur des arbres, de la surface de la couronne, des espèces d'arbres, de la localisation géographique le long du gradient de latitude ainsi que de la région sur le taux de succès de détection. Quoique fortement corrélées, la hauteur des arbres et la surface de la couronne se sont avérées les variables montrant la plus forte signification ($p \leq 0{,}001$) dans tous ces modèles. [Traduit par la Rédaction]

Introduction

Alpine and arctic tree lines are expected to advance to higher altitudes and further north because of global warming, which involves changes in temperature, precipitation, and snow coverage (ACIA, 2004). Several studies show that there have been upward, or at least stationary, trends of tree limits in both latitude and elevation over the past century as a result of climate change (Kullman, 2001, 2002; Grace

Received 17 March 2011. Accepted 21 July 2011. Published on the Web at http://pubs.casi.ca/journal/cjrs on 15 December 2011.

N. Thieme[1], O.M. Bollandsås, T. Gobakken, E. Næsset. Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway.

[1]Corresponding author (e-mail: nadja.thieme@umb.no).

et al., 2002; ACIA, 2004; Harsch et al., 2009). Particularly, boreal regions are expected to be highly affected by global warming (Kirschbaum and Fischlin, 1996), and northern Scandinavia has already experienced large changes in forest coverage as a result of warming events (ACIA, 2004). In addition to these natural causes, tree lines are also responding to anthropogenic factors such as grazing and pastoral economy (Callaghan et al., 2002; Holtmeier and Broll, 2005).

Tree lines are often not represented by definite demarcation lines, but rather a transition zone (Callaghan et al., 2002; Holtmeier and Broll, 2005) that can be labelled as the forest–tundra ecotone. Harper et al. (2011) defined the forest–tundra ecotone as "the transition zone between forest and tundra at high elevation or latitude". Its location between the mountain forest and the treeless alpine zone involves a high sensitivity to climatic changes and implies an important need for monitoring (Callaghan et al., 2002). For trees located in such areas, Aune et al. (2011) reported significant differences for tree parameters among multiple study sites, such as crown area and basal stem diameter. A declining trend in tree height from southern to northern sampling areas was found by Hofgaard et al. (2009). Therefore, the geographic region plays an important role in the monitoring of small trees in the forest–tundra ecotone as local patterns of tree migration and tree height growth caused by climate warming can be expected (Næsset and Nelson, 2007).

Many remote sensing techniques provide wall-to-wall coverage for vast areas. However, the spatial resolutions of optical remote sensing instruments limit monitoring studies in areas like the forest–tundra ecotone considerably, especially when it comes to small trees with a height growth of 0–5 cm per year (Næsset and Nelson, 2007). To monitor climatic changes in the growth of such small trees and the colonization of areas both further north and to higher altitudes, airborne laser scanning may be a well-suited tool, as shown in studies where biophysical parameters on single tree levels at different scales have been predicted from laser data (e.g., Hyyppä et al., 2001; Persson et al., 2002; Solberg et al., 2006; Næsset and Nelson, 2007). Tree height growth over relatively short time spans, between 2 and 5 years, is detectable using both high- and low-density airborne laser scanning data (Næsset and Gobakken, 2005; Yu et al., 2006). Covering a 5-year period using high-density airborne laser scanning data with a point density of approximately 10 m^{-2} in an individual tree approach, Yu et al. (2006) reported a root-mean square error (RMSE) of 0.43 m for the estimation of tree height growth on pine trees having reached a height of about 4.14 m in a boreal forest. Næsset and Gobakken (2005) found that even low-density airborne laser scanning data (1 m^{-2}) was applicable for detecting a significant mean tree height growth for spruce and pine trees in young and mature forest stands over a 2-year period (RMSE of 0.53 m). Furthermore, empirical studies conducted by Næsset and Nelson (2007) and Rees (2007) verified the suitability of airborne laser scanning using different densities of laser scanning data for the discrimination of small trees located in the forest–tundra ecotone. Concerning low-density data with a laser point density of approximately 0.25 m^{-2}, airborne laser scanning has proven to be a useful tool to distinguish individual trees over large areas that cover hundreds of square kilometres (Rees, 2007). However, Rees (2007) defined a tree by heights greater than 2 m, whereas in their study, Næsset and Nelson (2007) used a positive laser height value inside a tree polygon as a criterion for tree detection using high-density airborne laser scanning data (7.7 m^{-2}). This approach seems to be applicable for trees with heights greater than or equal to 1 m as there are generally few other objects in a forest–tundra ecotone environment that are expected to exceed an elevation of 1 m above ground surface. Næsset and Nelson (2007) also reported a detection success rate of over 90% for trees with heights greater than or equal to 1 m compared with significantly lower numbers for trees below 1 m. Thus, an adequate reliability of height values for tree detection is assumable, given a tree height exceeding 1 m. Using laser height values as the sole criterion for tree detection may introduce severe commission errors (Næsset and Nelson, 2007; Næsset, 2009b), especially for laser echoes received from the ground surface or objects close to it. Moreover, laser echoes close to the ground are highly sensitive to terrain modelling. Dependent on the properties of the terrain model, the sensor, and flight settings a large number of positive height values are actually representing nontree objects such as rocks, hummocks, and other terrain structures. For example, Næsset and Nelson (2007) found a commission error of 490% using a dataset where the terrain model was derived with commonly adopted smoothing criteria. Such commission errors have a significant influence on the reliability of tree detection analysis using laser height values. However, in multitemporal analysis terrain and terrain objects will not change, whereas when provided a sufficient time span, trees will increase in height and will change in number.

Moreover, different sensors with different properties may be used during data acquisition because of the vast areas involved. Variations in the vegetation returns and properties of the laser point clouds influencing the estimated biophysical properties from these data are reported for larger trees (Næsset, 2005, 2009a; Chasmer et al., 2006). For small single trees, previous research seems to indicate that different instruments and configurations are equally suited for detection, provided that the tree height is greater than 1 m (Næsset, 2009b).

The main objective of this study was to assess the capability of high-density airborne laser scanner data to detect small single trees located in the forest–tundra ecotone. For this purpose, sensor influences, tree characteristics, and spatial influences were studied to determine their effects on the probability of small tree detection. Finally, the accuracy of laser-derived tree height estimation was assessed.

Study area and data

Study area

The study area comprises a 1500 km long and approximately 180 m wide north–south transect stretching from Tromsø in northern Norway (69°3′ N 17°5′ E) to Tvedestrand in the southern part of the country (58°3′ N 9°0′ E) (**Figure 1**). The transect encompasses hundreds of mountain forest and alpine elevation gradients and covers sample plots in the forest–tundra ecotone at elevations ranging from 350 to 1200 m above sea level. In many of these localities the terrain surface is often characterized by rounded forms with occurrences of hummocks, rocks and boulders, and steep slopes. The prevalent tree species are Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and mountain birch (*Betula pubescens* ssp *czerepanovii*).

Field data

The field work in the transect was carried out during summer 2008 to provide in situ tree data for the detection analysis.

Anterior to field work, 56 potential field sites evenly distributed along the transect were selected for the establishment of sample plots. Based on aerial images and maps from the official Economic Map Series of the Norwegian Mapping Authority, these sites were chosen depending on their proximity to roads and the existence of nearby sample plots of the National Forest Inventory (NFI) for potential comparison. However, the final selection of specific field sites was taken in field which resulted in a total of 36 field sites allocated along the transect. One field site typically consisted of two to four sample plots covering the entire transition zone. Because this extent, from the mountain forest to the alpine zone, varies for different locations, the number of sample plots was visually adjusted at each field site depending on the altitudinal dimension of the transition zone. This yielded a total of 114 sample plots. Furthermore, sample plots within field sites were laid out with a 50 m interdistance to avoid overlap. However, one field site and therefore three sample plots had to be discarded because of erroneous coordinates resulting in a total number of 111 sample plots located at 35 field sites.

For precise navigation and positioning, real-time kinematic differential Global Navigation Satellite Systems

Figure 1. Overview of the study area with the 35 specific field sites (black and white points). The 1500 km long transect (black line) stretches from to 69°3′ N 17°5′ E to 58°3′ N 9°0′ E. Three field sites are located in the overlap zone (white points).

(dGNSS) were used with two Topcon Legacy E+ 20-channel dual-frequency receivers observing the pseudo range and carrier phase of both the Global Positioning System and Global Navigation Satellite System satellites. The horizontal positioning accuracy of this system is about 2 cm, provided that the reference point is without error. For each field site, the closest suitable reference point of the Norwegian Mapping Authority was used to establish the base station. The reference points used had an accuracy of 3 cm. The expected precision for the centres of sample plots measured in the field sites was 3–4 cm.

The small individual trees were selected according to a modified version of the point-centred quarter sampling method (PCQ) (Cottam and Curtis, 1956; Warde and Petranka, 1981). Emanating from the sample point as the centre, the surrounding area was divided into four quadrants, defined by the cardinal directions using a Suunto compass. In each quadrant one tree from three different tree height classes was sampled independent of tree species. The height classes were (*i*) lower than 1 m, (*ii*) between 1 m and 2 m, and (*iii*) taller than 2 m. Following the scheme of the PCQ sampling, each tree that was closest to the plot centre in their respective height class in each quadrant was sampled. This design led to a selection of a maximum of 12 trees per plot. Contrary to the original version of the PCQ method where an object is recorded irrespective of its distance to the centre, a maximum search limit was set to 25 m. A surveyor's tape measure was used to determine the closest tree and the maximum search distance in cases of doubt.

In the field, tree height, stem diameter at root collar, crown diameter, and tree species were recorded individually for each of the sample trees. Tree heights were measured using a steel tape measure for smaller trees and a Vertex III hypsometer for taller trees. Stem diameters were callipered and crown diameters were measured in the cardinal directions using a steel tape measure. For each tree, the precise position was determined using dGNSS according to the procedure outlined previously.

The original dataset included 771 measured trees. However, 26 trees had to be discarded because of missing or erroneous coordinates. Another tree had to be excluded from the dataset as well, because its distance to the centre of the sample plot exceeded the set limit of 25 m. This led to a total number of 744 trees that included 623 mountain birch, 68 Norway spruce, and 53 Scots pine trees used for subsequent analysis. However, no Norway spruce or Scots pine trees were recorded north of 66° N. Tree heights ranged from 0.02 m to 7.80 m (**Table 1**). Furthermore, the crown area of each tree was computed as the ellipse enfolding the tree defined by the crown diameters as the major and minor axes.

Table 1. Summary of field measurements of trees.

Tree species	Characteristics	*n*	Mean	Min.	Max.
Mountain birch	Height (m)	623	1.28	0.02	7.80
	Diameter (cm)	622	3.65	0.10	34.00
	Crown area (m^2)	623	0.90	0.001	19.54
Norway spruce	Height (m)	68	1.66	0.07	7.00
	Diameter (cm)	66	6.53	0.20	19.10
	Crown area (m^2)	68	1.43	0.006	5.69
Scots pine	Height (m)	53	1.33	0.10	5.10
	Diameter (cm)	53	5.00	0.30	18.90
	Crown area (m^2)	53	0.81	0.002	7.28

Laser data

Airborne laser scanner data were acquired in two separate acquisitions because of the large longitudinal extension of the study area and difficult weather conditions. The first acquisition was conducted in Southern and Central Norway on 23 and 24 July 2006 with an Optech ALTM 3100C laser scanning system. The remaining part of the transect in Northern Norway was scanned on 1 July 2007 with a Gemini upgraded version of the Optech ALTM 3100C laser scanner, denoted as ALTM Gemini. For comparison reasons, an overlap zone of approximately 80 km was flown in the county of Nordland (65°53′ N 13°27′ E) with both systems.

For both acquisitions, a Piper PA-31 Navajo aircraft with an average flying altitude of 800 m above ground level carried the laser scanning systems. Furthermore, in both acquisitions the scan frequency was 70 Hz, maximum half scan angle was 7°, and the flight speed was approximately 75 ms^{-1}. The average footprint diameter was 20 cm in both cases. However, pulse repetition frequency (PRF) was 100 kHz in 2006 and 125 kHz in 2007, which resulted in mean pulse densities of 6.8 m^{-2} and 8.5 m^{-2} for the two acquisitions, respectively. The unequal PRFs were chosen to obtain laser point clouds of trees as similar as possible for the two acquisitions. A test flight in May 2007, with several different PRFs in another study area where laser data with the same system had been acquired prior to a system upgrade, indicated that a PRF of 125 kHz after the system upgrade produced similar data to the 100 kHz prior to the upgrade. Different energy output levels of the laser pulses of the sensor prior to and after the system upgrade were the most likely reason for the somewhat different properties of the point clouds over vegetated areas for a given PRF.

The routing and configuration of the transect were planned in accordance with the fixed 3 km × 3 km grid used by the NFI for its permanent field sample plots. Thereby, 50 NFI plots located close to the mountain forest were covered. Furthermore, the 1500 km long transect was split up into 147 individual flight lines to keep the flying altitude above ground level as constant as possible.

Preprocessing of the laser data was conducted by a contractor (Blom Geomatics, Norway). For all laser points, planimetric coordinates (*x* and *y*) and ellipsoidal height values were computed.

Laser echoes labelled "last-of-many" and "single" (LAST), and laser echoes labelled "first-of-many" and "single" (FIRST), were used for the derivation of the terrain model and the analysis of tree detection, respectively.

Both the ALTM 3100C and the ALTM Gemini are instruments recording up to four echoes per pulse with a minimum vertical distance of 2.1 m between echoes for the ALTM 3100C. For the ALTM Gemini, the minimum vertical distance is assumed to be larger because of the pulse width's influence on the vertical resolution (Baltsavias, 1999; Næsset, 2009b). Unfortunately, the system producer was unwilling to release such instrument-specific information. Based on these instrument properties, potentially very few pulses have more than one echo because of the low vegetation. From the planimetric coordinates and the corresponding height data of the LAST returns, ground points were classified based on an iteration distance of 1.0 m with the TerraScan software (Terrasolid, 2010). A triangulated irregular network (TIN) was derived with an iteration angle of 9° representing a typical value for the generation of terrain models using airborne laser scanning data.

For the analysis of tree detection, returns labelled as FIRST were projected onto the TIN surface and the corresponding terrain height values of these locations were interpolated. The differences between the FIRST return echo heights and the corresponding interpolated terrain height values were computed and stored for the tree detection analysis.

Methods

Sensor analysis

Because two different laser scanning systems were employed during data acquisition, data from the 80 km overlap zone were used to assess the potential sensor effects on the small individual tree detection. For this purpose, the maximum laser-derived tree height from both data acquisitions was estimated for each tree located in the overlap zone. In this individual tree approach, equivalence tests were performed using the *tost* procedure of the *equivalence* package (Robinson, 2010) in the statistical computing software R (R Development Core Team, 2007) for the detected trees with $h_{\max} > 0$, and all trees with $h_{\max} \geq 0$, respectively, where $h_{\max}$ is the maximum laser height. To test if the difference between the mean values of the two data acquisitions is larger than a negligible upper and lower limit, equivalence tests are more appropriate to use than one-sample *t* tests (Kaminski et al., 2010).

In addition to the 111 sample plots along the transect, 54 sample plots were established in the overlap zone for an area-based analysis. With an area of 1000 m^2 each, they were purposefully selected by visual inspection of digital aerial imagery and divided into three different categories: (*i*) solitary distributed small-sized trees, (*ii*) sparsely distributed medium-sized trees, and (*iii*) large trees in dense forest stands. Maximum tree height was calculated for each plot and equivalence tests were applied for the respective categories testing the null hypothesis for a significant difference between the two data acquisitions.

Tree detection

In this study, the detection of small individual trees in the forest–tundra ecotone was based on laser echoes with height values greater than the classified ground surface. These laser echoes represent the only indicators for the presence of objects on the terrain surface and are therefore the sole candidates for tree detection. Instead of detecting trees automatically from laser point clouds in individual tree detection for larger trees as introduced by Hyyppä et al. (2001), Persson et al. (2002), and Solberg et al. (2006), the small trees were defined by the field measurements of the vertical projection of their tree crowns. For this purpose, the field-measured crown diameters were used to compute elliptical tree crown polygons. Trees with a crown diameter value less than1.0 m in at least one of the cardinal directions were assigned a tree crown polygon with a constant radius of 0.5 m. This procedure was implemented to compensate for a positioning error of the laser data of up to 0.5 m as reported by the contractor. The FIRST return laser data were subsequently overlaid with the tree crown polygons for the evaluation of laser echoes falling inside the individual crown polygons. A tree hit by at least one FIRST return laser echo with $h_{\max} > 0$ was regarded as successfully detected. However, trees hit by at least one laser echo with $h_{\max} = 0$ and trees that were not hit by any laser echo at all were classified as not having been detected.

Generalized linear models (GLM) were employed for analyzing the effects of different factors influencing the probability of a tree being successfully detected. A GLM consists of three elements: the random component identifying the dependent variable *Y*, the systematic component specifying the independent variables $\{x_j\}$, and the link function relating those two components (Agresti, 2007). In this case, a logit link function connects the binary response *Y* (detected/not detected) to the independent variables $\{x_j\}$ representing the effects of (*i*) the size of the tree as expressed by tree height and tree crown area, (*ii*) the tree species, and (*iii*) the geographic location concerning latitude (GLM 1) and region (GLM 2) (**Table 2**). Thus, the following model was estimated:

$$\log\left[\frac{\pi(\text{detected})}{1-\pi(\text{detected})}\right] = \alpha + \beta_1 \underline{x}_1 + \cdots + \beta_k \underline{x}_k \qquad (1)$$

The GLM models were fitted using the *glm* (Dalgaard, 2008) procedure of the *stats* package (R Development Core Team, 2007) in R. Likelihood-ratio tests were employed to compare the fit of the respective model with a more complex model (Agresti, 2007).

To test the null hypothesis that tree species does not significantly contribute to the model, a likelihood-ratio test was performed. Another likelihood-ratio test was performed for an additional indicator variable representing the 35 different sample sites. Thus, a potential contribution of regional effects to the model was investigated. Both GLM models were evaluated by leave-one-out cross validation and

Table 2. Independent variables included in the generalized linear models (GLM 1 and GLM 2) and the generalized linear mixed model (GLMM 1).

Characteristic	Model
Tree height measured in field (m) (*H*)	GLM 1, GLM 2, GLMM 1
Estimated tree crown area (m^2) (*CA*)	GLM 1, GLM 2, GLMM 1
Latitude (m) (*LAT*)	GLM 1, GLM 2, GLMM 1
Indicator variable for tree species (*TS*)	GLM 1, GLM 2, GLMM 1
Indicator variable for region (*R*)	GLM 2, GLMM 1

the Hosmer–Lemeshow test statistics (Hosmer and Lemeshow, 2000) testing the goodness-of-fit by comparing observed and fitted event rates (Agresti, 2007).

A generalized linear mixed model (GLMM) was applied to further investigate a potential regional effect. Contrary to classical statistics, where all observations are assumed to be independent, mixed-effects models are based on the assumption that observations within a cluster are actually dependent, preserving the independency between clusters (Demidenko, 2004). Such variables are included as random effects Z whereas the remaining variables X are used as fixed-effects parameters (Bates, 2010). The following GLMM was therefore estimated:

$$\log\left[\frac{\pi(\text{detected})}{1-\pi(\text{detected})}\right] = X\beta + Zv \quad (2)$$

The GLMM model was estimated with the variables used in the GLM analysis employing the region parameter representing the 35 field sites as a random effect. A Laplace approximation was performed for fitting the GLMM using the *glmer* procedure provided by the *lme4* package (Bates and Maechler, 2010) in R.

Assessment of laser tree height estimation

Tree growth is an essential factor contributing to change and is thus also essential for change detection. The ability to estimate tree growth with laser data will depend on the accuracy of tree height estimation with the same data. It is therefore appropriate to compare tree heights estimated from laser data with in situ observations. For this purpose, the mean differences between the h_{max} and the corresponding field-measured tree heights were calculated for the detected trees. The precision of the tree heights derived from the laser data was assessed for the three tree species and three tree height classes individually by estimating the standard deviation for the differences.

Results and discussion

Sensor analysis

The analyses of potential sensor effects using the laser scanner data from the 80 km overlap zone indicated no significant differences between the two laser scanning systems. Both the individual tree analysis using the field-measured trees and the area-based analysis using the 54 established sample plots confirmed this finding.

For the individual trees located in the overlap zone, no significant difference in maximum tree height could be found between the two sensors. The equivalence tests resulted in $p < 0.001$ both for trees with $h_{max} > 0$ and $h_{max} \geq 0$. The mean differences were -0.042 m and 0.020 m for trees with $h_{max} > 0$ and $h_{max} \geq 0$, respectively.

For the small and medium-sized trees used in the area-based analysis, the equivalence test revealed $p < 0.001$ and mean differences of 0.073 m and 0.004 m, respectively. For the large trees class, the equivalence tests resulted in $p = 0.122$ and a mean difference of -0.390 m. Thus, the null hypothesis that there is a significant difference between the two data acquisitions within an upper and lower limit could be rejected for small and medium-sized trees, but not for large trees. However, all individual trees measured in the field are comparable with the small and medium-sized tree class in the area-based analysis. It was therefore considered unnecessary to perform any data adjustment to calibrate the laser data from the two acquisitions against each other prior to the subsequent analysis.

The probability of being hit by at least one laser pulse also is related to the laser pulse density and not solely dependent on tree size as used in these analyses. However, Næsset (2009b) noted that the configuration of the ALTM 3100C at 100 kHz and the ALTM Gemini at 125 kHz (as used in the current study) revealed the best corresponding canopy height distributions in a preliminary investigation with old-growth forest. Using the two laser pulse densities in this study, the majority of trees taller than 1 m and almost all trees taller than 2 m were hit by at least one laser pulse (**Table 3**). Previous studies on small single tree detection that also included different flight and sensor configurations (Næsset, 2009b), showed corresponding results concerning the suitability of different instruments provided laser pulse densities were higher than 7 m^{-2}.

Tree detection

The overall tree detection success rate was 71% irrespective of tree height classes and tree species (**Table 3**). With regard to the different tree height classes, success rates for different tree species ranged between 37% and 50% for trees smaller than 1 m. For trees between 1 m and 2 m and trees taller than 2 m the success rates were $> 84\%$ and $> 91\%$, respectively, for all three tree species. The best overall success rate irrespective of height class was found for spruce (81%), whereas the corresponding values for mountain birch and pine were 71% and 62%, respectively.

Trees that were not successfully detected by the airborne laser scanner were divided into trees that were hit by at least one laser pulse with $h_{max} = 0$ and trees that were not hit by any laser pulse at all. In total, 29% of the trees were not

Table 3. Success in tree detection for different species and height classes.

Tree species	Height class	*n*	Detected *n* (%)	$h_{max} = 0$ *n* (%)	No hits *n* (%)
Mountain birch	<1 m	293	147 (50)	134 (46)	12 (4)
	1–2 m	199	167 (84)	21 (11)	11 (5)
	>2 m	131	127 (97)	3 (2)	1 (1)
Total		**623**	**441 (71)**	**158 (25)**	**24 (4)**
Norway spruce	<1 m	20	9 (45)	11 (55)	0 (0)
	1–2 m	24	22 (92)	1 (4)	1 (4)
	>2 m	24	24 (100)	0 (0)	0 (0)
Total		**68**	**55 (81)**	**12 (18)**	**1 (1)**
Scots pine	<1 m	30	11 (37)	19 (63)	0 (0)
	1–2 m	12	12 (100)	0 (0)	0 (0)
	>2 m	11	10 (91)	1 (9)	0 (0)
Total		**53**	**33 (62)**	**20 (38)**	**0 (0)**
Total (all trees)		**744**	**529 (71)**	**190 (26)**	**25 (3)**

detected irrespective of tree height class and tree species, and 3% of those were not hit by any laser pulse (**Table 3**). For trees smaller than 1 m, between 46% and 63% had $h_{max} = 0$, whereas the amount was less than 11% for all trees greater than or equal to 1 m. Regarding the different tree species, there are major differences in tree detection rates for the smallest trees in particular (**Table 3**). Spruce had the lowest amount of $h_{max} = 0$ (18%), followed by birch (25%), and pine (38%). Further, pine was the tree species where all trees measured in the field were hit by at least one laser pulse, and merely 1% of the spruce trees and 4% of the birch trees were not hit by any laser pulse at all.

Multiple logistic regression analysis using GLM 1 showed that an increase in tree height and crown area in particular resulted in an increasing probability of a tree being successfully detected. This result fits well with the fact that an increase in tree size, especially in the area of the vertical tree crown projection, results in an increasing probability of being hit by at least one laser pulse (Næsset and Nelson, 2007). **Figure 2** illustrates the predicted probability of a tree being hit by at least one laser pulse for the three different tree species as a function of tree height and tree crown area. Furthermore, tree height and crown area were highly significant explanatory variables with *z* values of 5.34 ($p < 0.001$) and 3.41 ($p = 0.001$), respectively (**Table 4**). A *z* value represents the respective regression coefficient divided by the estimated standard error (R Development Core Team, 2007). However, a strong intercorrelation ($r = 0.76$) was found between those two variables including all tree species.

The positive regression coefficient for latitude ($p < 0.001$) indicated a slightly higher probability of a successful detection with increasing latitude. A slight intercorrelation ($r = 0.35$) was found between the latitude and the tree species variable that might result in a mixed effect of those two. This would be reasonable because of a lack of spruce and pine trees north of 66° N, which is also reflected in the field data. Another facet is attributed to the terrain model and its parameters as it may be the determining factor for detecting small individual trees in the forest–tundra ecotone (Næsset, 2009b; Næsset and Nelson, 2007). A 1500 km long transect involves changes in topography in the mountainous areas along the latitude gradient, which are reflected in the varying suitability of the used terrain model.

There was a tendency of slightly higher probabilities of more successful detection of spruce ($p = 0.085$) and pine ($p = 0.097$) trees than birch trees (**Table 4**) as indicated by the positive regression coefficients of the two indicator variables TS_{spruce} and TS_{pine} for tree species (*TS*). Conifer trees had a higher probability of being detected than birch, which may be due to the low foliage density of small birch trees compared with the more compact and dense conifers. The highest probability of being hit by at least one laser pulse was found for spruce trees, closely followed by pine trees (**Table 4**), given the same tree crown area. According to the predicted probabilities of being successfully detected (**Figure 2**), almost all trees can be found using an airborne

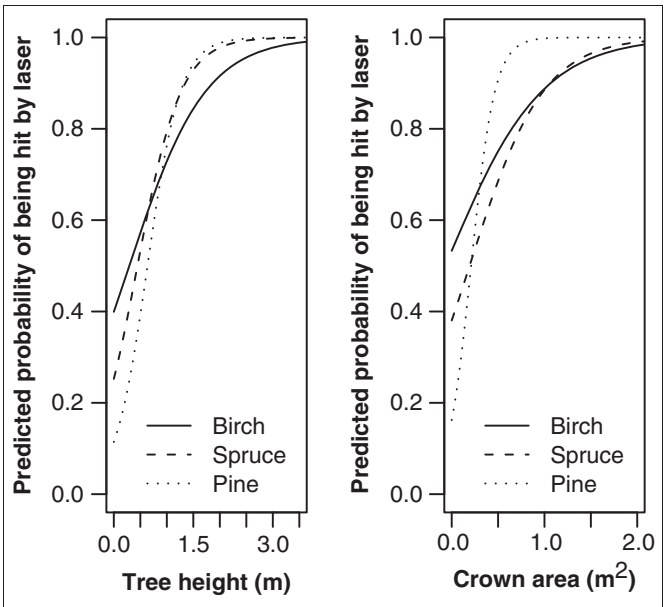

Figure 2. Predicted probability of a tree being hit by at least one laser pulse as a function of tree height (left) and tree crown area (right) for birch, spruce, and pine.

Table 4. Fitted generalized linear model (GLM 1).

Coefficient	Estimate	*z* value	*p* value
Intercept	−2.05E+01	−7.84	<0.001
H	1.21E–02	5.34	0.001
CA	1.15E+00	3.41	<0.001
LAT	2.74E–06	7.73	<0.001
TS_{pine}[a]	6.31E–01	1.66	0.097
TS_{spruce}[b]	7.34E–01	1.72	0.085

[a]Indicator variable for pine.
[b]Indicator variable for spruce.

laser scanner with a laser pulse density greater than 7 m^{-2} when they have reached heights of 1.3–1.4 m and 1.9 m for conifer and birch trees, respectively. Regarding tree crown area, almost all trees were detected after reaching a size of 0.5 m^2 for pine trees, and 1.1 m^2 for spruce and birch trees, respectively. Moreover, a likelihood-ratio test revealed no significant detection difference between pine and spruce trees ($p = 0.844$) (**Table 5**), that could justify a simplification of the model by merging the coniferous tree species to a common class. However, because all indicator variables representing tree species were close to being statistically significant in the original model at a 5% level, we decided to keep all tree species in the model.

A likelihood-ratio test indicated that the probability of tree detection differed significantly between the 35 field sites ($p < 0.001$) (**Table 5**). Therefore, it seemed reasonable to extend the GLM 1 model with an indicator variable representing the 35 different field sites (GLM 2).

For the GLM 1 model, the Hosmer–Lemeshow test statistics (Hosmer and Lemeshow, 2000) revealed no lack of fit ($p = 0.108$) (**Table 5**). The *p* value for the Hosmer–Lemeshow test for the model fitted, including the indicator variable for field site (GLM 2), revealed a good fit as well ($p = 0.310$) (**Table 5**). Leave-one-out cross validation indicated an overall agreement of 80.9% and 80.1% between observed and predicted values by the original model GLM 1 and the GLM 2 model including regional effects, respectively.

The GLMM 1 revealed the same results for the fixed effects as in the GLM analysis (**Table 6**). The *z* values were marginally lower for the coefficients of the GLMM 1. However, the levels of significance did not change for most of the explanatory variables. The *p* value for the indicator variable for pine was somewhat lower. The random effect represented by the field site indicator variable had a variation of 0.125 and a standard deviation of 0.35, indicating some variations in the model caused by locality of the respective sample sites. However, there is no clear interpretation and quantification of the proportion of the regional factor having influence on the GLMM (D. Bates, pers. comm.).

Table 5. Likelihood-ratio test and Hosmer–Lemeshow tests for GLM 1 and GLM 2.

Likelihood-ratio	Deviance	*p* value
H_0: $TS_{pine} = TS_{spruce}$[a]	0.04	0.844
H_0: $R_1 = R_2 = \ldots R_{35}$[b]	79.42	<0.001
Homser–Lemeshow	**Wald chi-square**	***p* value**
GLM 1	13.11	0.108
GLM 2	9.39	0.310

[a]Indicator variable for tree species pine and spruce.
[b]Indicator variable for the 35 field sites R_1 to R_{35}.

Table 6. Fitted generalized linear mixed model (GLMM 1).

Random effects	Variance	SD	
R (Intercept)	1.25E–01	0.35	
Fixed effects			
Coefficient	Estimate	*z* value	*p* value
Intercept	−2.05E+01	−6.92	<0.001
H	1.21E–02	5.28	<0.001
CA	1.15E+00	3.41	0.001
LAT	2.74E–06	6.78	<0.001
TS_{pine}[a]	6.31E–01	1.61	0.107
TS_{spruce}[b]	7.34E–01	1.66	0.096

[a]Indicator variable for pine.
[b]Indicator variable for spruce.

Assessment of laser tree height estimation

The assessment of laser-derived tree heights was based on the corresponding tree heights measured in the field for those trees that were successfully detected ($h_{max} > 0$). A general trend of greater underestimations for larger trees was found with a concurrent lower precision in these laser-derived tree heights. The greatest underestimation errors were observed for pine trees, however, which coincided with the lowest variation in the estimated heights.

Underestimations of the field-measured tree heights ($p < 0.01$) were revealed in the interval between 0.20 m and 1.08 m with a standard deviation of 0.22–0.76 m (**Table 7**). For the coniferous tree species, the underestimation was largest for trees between 1 m and 2 m tall, whereas the largest underestimation for mountain birch was found for trees taller than 2 m. Including all tree species, the smallest trees (<1 m) were underestimated by 0.20–0.47 m, and trees taller than 1 m were underestimated by 0.69–1.08 m. Furthermore, the standard deviation for the differences between tree heights derived from airborne laser scanning and field measurements ranged between 0.22 m and 0.35 m for trees smaller than 1 m and from 0.27 m to 0.76 m for the taller trees (≥1 m). However, heights of a few individual trees were greatly overestimated in the laser data. Such overestimations were found for trees in all height classes of mountain birch, but only for the smallest and the largest Norway spruce trees (**Figure 3**). Great overestimations of heights may be traced back to special characteristics of how the trees appear in their natural environment and to specific features of the applied sampling method. At certain locations, small birch and spruce trees tend to grow in groups consisting of several individual trees with different heights.

Table 7. Mean of field-measured tree height, differences (D) between field-measured and laser-derived tree height, and standard deviation of differences (SD).

				D (m)		
Tree species	Height class	n	Observed mean (m)	Range	Mean	SD
Mountain birch	<1 m	147	0.38	−0.85 to 2.85	−0.20*	0.35
	1–2 m	167	1.34	−1.68 to 2.00	−0.85*	0.56
	>2 m	127	3.22	−3.40 to 1.29	−1.03*	0.74
Norway spruce	<1 m	9	0.48	−0.62 to 0.04	−0.29†	0.24
	1–2 m	22	1.33	−1.53 to −0.34	−0.79*	0.30
	>2 m	24	3.01	−1.56 to 1.37	−0.69*	0.76
Scots pine	<1 m	11	0.67	−0.72 to −0.06	−0.47*	0.22
	1–2 m	12	1.33	−1.61 to −0.68	−1.08*	0.27
	>2 m	10	3.83	−1.41 to −0.44	−0.97*	0.33

Note: Level of significance: *<.001; †<.01

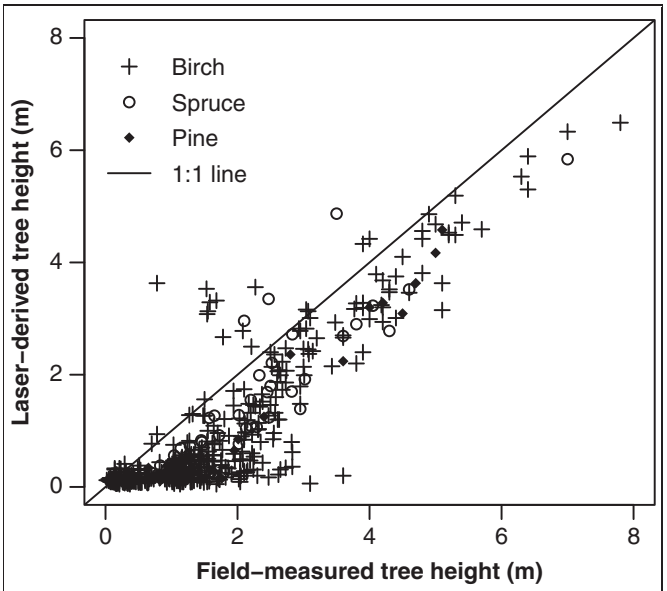

Figure 3. Field-measured tree height plotted against laser-derived tree height.

In such cases, there is a high potential of sampling trees in different tree height classes with small distances in between. This situation may lead to a false interpretation of laser pulses because branches belonging to the larger tree are misleadingly assigned to the smaller tree resulting in an overestimation.

These results revealed findings that are concordant with results from previous analyses, both related to small pioneer trees advancing the tree line further up and northwards (Næsset and Nelson, 2007; Næsset, 2009b) and larger trees (Hyyppä et al., 2001; Persson et al., 2002; Solberg et al., 2006).

Conclusion

This study has demonstrated that high-density airborne laser scanner data are useful for the detection of small individual trees located in the forest–tundra ecotone. Thus, it confirms findings from previous studies conducted in small individual study sites (Næsset and Nelson, 2007; Næsset, 2009b). The current study, with its huge geographical extension, forms a solid basis for generalization and inference that goes far beyond previous research. Almost every tree exceeding a height of 1 m was successfully detected by using laser pulse densities greater than 6.8 m^{-2}. No significant differences in the probability of a tree being hit by at least one laser pulse for laser point densities ranging from 6.8 m^{-2} to 8.5 m^{-2} were found. Therewith, the variation caused by differing instruments during data acquisition could be disregarded based on the definition that a tree is successfully detected if it was hit by at least one laser pulse with $h_{max} > 0$. For trees smaller than 1 m in height, a decrease in the success rate of tree detection was found. For these small trees, tree species was an important explanatory variable for the probability of being detected. However, only very few pine trees were included in the study which should be borne in mind by transferring the findings of this study to similar problems. Furthermore, the probability of detecting small individual trees in the forest–tundra ecotone revealed a latitude effect. This might be influenced by the occurrence of the different tree species along the latitude gradient that is highly dominated by birch in the northern part. Also the local topography might have a certain influence on the terrain model resulting in a significant difference in successful tree detection along a latitude gradient. The precision of the terrain model depends on the structure of the terrain and the applied algorithm (Reutebuch et al., 2003; Hodgson and Bresnahan, 2004; Peng and Shih, 2006; Su and Bork, 2006). Misclassification of trees occurs because they cannot be separated from terrain objects with positive laser height values without additional information. The probability of a laser pulse hitting a tree decreases with decreasing crown width. Furthermore, the probability of a laser pulse being returned with a positive height value decreases with decreasing tree height. Laser echoes from small trees also have a higher probability of being misclassified as ground returns.

Thus, the procedures used for classification of laser echoes and modelling of the terrain are essential for distinguishing between small terrain objects and small trees. Therefore, to enhance the success rate, future research on small individual tree detection should involve more aspects of local trends both concerning topography and the composition of tree species.

According to Næsset and Nelson (2007), significant changes in mean height should be detectable over short time periods such as one year for small areas, assuming an annual height growth of 5 cm in conjunction with a bias control across laser data acquisitions. However, for regions such as the forest–tundra ecotone, longer time spans for change detection are expected to be required because of the low and varying rates in growth in these areas. High-density airborne laser scanner data is capable of detecting small individual trees over large areas and time periods on the order of a few years. It is therefore a useful tool for monitoring tree migration changes in the forest tundra ecotone.

During the last five years, airborne laser scanning has become more and more important in regard to change detection and monitoring in areas like the forest–tundra ecotone. Campaigns in Canada (Boudreau et al., 2008) and Norway (Næsset and Nelson, 2007) focused on biomass estimation in remote areas not covered by the NFI plots and change detection in the forest–tundra ecotone, respectively. Furthermore, the Canadian Forest Service initiated a campaign of wide-area airborne laser data collection covering Canada's northern boreal forests to investigate ecosystem risks related to climate change (M. Wulder, pers. comm.). Such initiatives emphasize the expanding use of airborne laser scanning in remote areas and the increasing demand for well-suited methods for such areas.

Acknowledgements

This research was funded by the Research Council of Norway (project #184636/S30). We wish to thank Blom Geomatics AS, Norway, for the collection and processing of the airborne laser scanner data. Thanks also appertain to Mr. Vegard Lien at the Norwegian University of Life Sciences, who was responsible for the fieldwork. Furthermore, Nadja Thieme would like to thank Mr. Hans Ole Ørka at the Norwegian University of Life Sciences for valuable remarks and encouragement during the analysis process.

References

ACIA. 2004. *Impacts of a warming Arctic: Arctic Climate Impact Assessment*. Cambridge University Press, Cambridge, UK. 146 pp.

Agresti, A. 2007. *An introduction to categorical data analysis.* Wiley & Sons, Inc., Hoboken, New Jersey, USA. 400 pp.

Aune, S., Hofgaard, A., and Söderström, L. 2011. Contrasting climate- and land-use-driven tree encroachment patterns of subarctic tundra in northern Norway and the Kola Peninsula. *Canadian Journal of Forest Research*, Vol. 41, No. 3, pp. 437–449. doi:10.1139/X10-086.

Baltsavias, E.P. 1999. Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 54, pp. 199–214. doi:10.1016/S0924-2716(99)00015-5.

Bates, D.M. 2010. *lme4: Mixed-effects modelling with R*. Draft. Available from http://lme4.r-forge.r-project.org/book/ [accessed 7 July 2010].

Bates, D., and Maechler, M. 2010. *Package 'lme4'*. 29 pp. Available from http://lme4.r-forge.r-project.org/ [accessed 7 July 2010].

Boudreau, J., Nelson, R.F., Margolis, H.A., Beaudoin, A., Guindon, L., and Kimes, D.S. 2008. Regional aboveground biomass using airborne and spaceborne LiDAR in Québec. *Remote Sensing of Environment*, Vol. 112, pp. 3876–3890. doi:10.1016/j.rse.2008.06.003.

Callaghan, T.V., Werkman, B.R., and Crawford, R.M.M. 2002. The tundra–taiga interface and its dynamics: Concepts and applications. *Ambio Special Report*, Vol. 12, pp. 6–14.

Chasmer, L., Hopkinson, C., Smith, B., and Treitz, P. 2006. Examining the influence of changing laser pulse repetition frequencies on conifer forest canopy returns. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 12, pp. 1359–1367.

Cottam, G., and Curtis, J.T. 1956. The use of distance measures in phytosociological sampling. *Ecology*, Vol. 37, No. 3, pp. 451–460. doi:10.2307/1930167.

Dalgaard, P. 2008. *Introductory Statistics with R*. Springer Science + Busi-Business Media, LLC., New York, USA. 364 pp.

Demidenko, E. 2004. *Mixed Models: Theories and Applications.* Wiley & Sons, Inc., Hoboken, New Jersey, USA. 736 pp.

Grace, J., Berninger, F., and Nagy, L. 2002. Impacts of climate change on the tree line. *Annals of Botany*, Vol. 90, pp. 537–544. doi: 10.1093/aob/mcf222.

Harper, K.A., Danby, R.K., De Fields, D.L., Lewis, K.P., Trant, A.J., Starzomski, B.M., Savidge, R., and Hermanutz, L. 2011. Tree spatial pattern within the forest–tundra ecotone: a comparison of sites across Canada. *Canadian Journal of Forest Research*, Vol. 41, pp. 479–489. doi:10.1139/X10-221.

Harsch, M.A., Hulme, P.E., McGlone, M.S., and Duncan, R.P. 2009. Are treelines advancing? A global meta-analysis of treeline response to climate warming. *Ecology Letters*, Vol. 12, pp. 1–10. doi: 10.1111/j.1461-0248.2009.01355.x.

Hodgson, M.E., and Bresnahan, P. 2004. Accuracy of airborne lidar-derived elevation: Empirical assessment and error budget. *Photogrammetric Engineering and Remote Sensing*, Vol. 70, No. 3, pp. 331–339.

Hofgaard, A., Dalen, L., and Hytteborn, H. 2009. Tree recruitment above the treeline and potential for climate-driven treeline change. *Journal of Vegetation Science*, Vol. 20, pp. 1133–1144. doi: 10.1111/j.1654-1103.2009.01114.x.

Holtmeier, F.-K., and Broll, G. (2005). Sensitivity and response of northern hemisphere altitudinal and polar treelines to environmental changes at landscape and local scales. *Global Ecology and Biogeography*, Vol. 14, pp. 395–410. doi: 10.1111/j.1466-822X.2005.00168.x.

Hosmer, D.W., and Lemeshow, S. 2000. *Applied Logistic Regression*. Wiley & Sons, Inc., Hoboken, New Jersey, USA. 392 pp.

Hyyppä, J., Kelle, O., Lehikoinen, M., and Inkinen, M. 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 5, pp. 969–975. doi:10.1109/36.921414.

Kaminski, L., Schepers, U., and Wätzig, H. 2010. Analytical method transfer using equivalence test with reasonable acceptance criteria and appropriate effort: Extension of the ISPE concept. *Journal of Pharmaceutical and Biomedical Analysis*, Vol. 53, pp. 1124–1129. doi: 10.1016/j.jpba.2010.04.034.

Kirschbaum, M., and Fischlin, A. 1996. Climate change impacts on forests. In *Climate change 1995–Impacts, adaptations and mitigation of climate change: scientific-technical analysis. Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel of Climate Change*. R. Watson, M.C. Zinyowera, and R.H. Moss. Cambridge University Press, Cambridge, UK. pp. 99–129.

Kullman, L. 2001. 20th century climate warming and tree-limit rise in the Southern Scandes of Sweden. *Ambio*, Vol. 30, No. 2, pp. 72–80.

Kullman, L. 2002. Rapid recent range-margin rise of tree and shrub species in the Swedish Scandes. *Journal of Ecology*, Vol. 90, pp. 68–77. doi: 10.1046/j.0022-0477.2001.00630.x.

Næsset, E. 2005. Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*, Vol. 98, pp. 356–370. doi: 10.1016/j.rse.2005.07.012.

Næsset, E. 2009a. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*, Vol. 113, pp. 148–159. doi: 10.1016/j.rse.2008.09.001.

Næsset, E. 2009b. Influence of terrain model smoothing and flight and sensor configurations on detection of small pioneer trees in the boreal–alpine transition zone utilizing height metrics derived from airborne scanning lasers. *Remote Sensing of Environment*, Vol. 113, pp. 2210–2223.

Næsset, E., and Gobakken, T. 2005. Estimating forest growth using canopy metrics derived from airborne laser scanner data. *Remote Sensing of Environment*, Vol. 96, pp. 453–465. doi: 10.1016/j.rse.2009.06.003.

Næsset, E., and Nelson, R. 2007. Using airborne laser scanning to monitor tree migration in the boreal–alpine transition zone. *Remote Sensing of Environment*, Vol. 110, pp. 357–369. doi: 10.1016/j.rse.2005.04.001.

Peng, M.-H., and Shih, T.-Y. 2006. Error assessment in two lidar-derived TIN datasets. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 8, pp. 933–947. doi: 10.1016/j.rse.2007.03.004.

Persson, Å., Holmgren, J., and Söderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, Vol. 68, No. 9, pp. 925–932.

R Development Core Team. 2007. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rees, W.G. 2007. Characterisation of arctic treelines by LiDAR and multispectral imagery. *Polar Record*, Vol. 43, No. 227, pp. 345–352.

Reutebuch, S.E., McGaughey, R.J., Andersen, H.-E., and Carson, W.W. 2003. Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing*, Vol. 29, No. 5, pp. 527–535. doi: 10.5589/m03-022.

Robinson, A. 2010. Package 'equivalence'. 19 pp. Available from http://cran.r-project.org/web/packages/equivalence/equivalence.pdf [accessed 28 June 2011].

Solberg, S., Næsset, E., and Bollandsås, O.M. 2006. Single tree segmentation using airborne laser scanner data in a heterogeneous spruce forest. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 12, pp. 1369–1378.

Su, J., and Bork, E. 2006. Influence of vegetation, slope, and lidar sampling angle on DEM accuracy. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 11, pp. 1265–1274.

Terrasolid. 2010. *TerraScan User's Guide*. Terrasolid Ltd., Jyväskylä, Finland. 308 pp. Available from www.terrasolid.fi [accessed 6 May 2010].

Warde, W., and Petranka, J.W. 1981. A correction factor table for missing point-center quarter data. *Ecology*, Vol. 62, No. 2, pp. 491–494. doi: 10.2307/1936723.

Yu, X., Hyyppä, J., Kukko, A., Maltamo, M., and Kaartinen, H. 2006. Change detection techniques for canopy height growth measurements using airborne laser scanner data. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 12, pp. 1339–1348.

PAPER II

Classifying tree and non-tree echoes from airborne laser scanning in the forest-tundra ecotone

Authors

Nadja Thieme*, Ole Martin Bollandsås, Terje Gobakken, Erik Næsset

Address

Norwegian University of Life Sciences

Department of Ecology and Natural Resource Management

P.O. Box 5003, N-1432 Ås, Norway

*** Corresponding author**

Nadja Thieme

E-mail: nadja.thieme@umb.no

Phone: +47 64 96 57 20

Submitted for publication in the Canadian Journal of Remote Sensing/Journal canadien de télédétection**.**

Abstract

Temperature-sensitive ecosystems such as the forest-tundra ecotone are expected to be particularly affected by changing climate. A large proportion of the total land area in Norway is represented by the forest-tundra ecotone requiring effective monitoring techniques for these areas. Airborne laser scanning (ALS) has been proposed for detection of small pioneer trees and its height and intensity data may hold potentials for monitoring tasks. The main objective of the present study was to assess the capability of high-density ALS data to classify tree and non-tree echoes directly from the laser point cloud. For this purpose, the laser height and intensity, a geospatial variable represented by the area of Voronoi polygons, and the terrain variables aspect and slope were used in order to distinguish between tree and non-tree laser echoes along a 1,000 km long transect stretching from northern Norway (66°19' N) to the southern part of the country (58°3' N). Generalised linear models (GLM) and support vector machines (SVM) were employed for the classification using different combinations of the aforementioned variables. Total accuracy and the Cohen's kappa coefficient were used for performance assessment for the different models. A total accuracy of at least 93% was found irrespective of classification method or model and Cohen's kappa coefficients indicated moderate fits for all models using both classification methods. Comparisons of Cohen's kappa coefficients revealed equivalent performances for the GLM and SVM classification methods for models consisting of different combinations of the laser height, intensity, the geospatial variable, and aspect. However, SVM was superior when laser height and intensity were used together with slope. In summary, the capability of high-density ALS data for the classification of tree and non-tree echoes directly from the laser point cloud could be verified irrespective of the classification method.

1. Introduction

Arctic and alpine tree lines are expected to advance further north and to higher altitudes due to changing climate (ACIA, 2004). Changes in temperature, precipitation, and snow coverage will affect numerous ecosystems and their interaction (Stenseth et al., 2002; ACIA, 2004; Woodall et al., 2009) and forest ecosystems are expected to be highly affected by increasing temperatures, particularly in boreal regions (Kirschbaum and Fischlin, 1996). The forest-tundra ecotone as "the transition zone between forest and tundra at high elevation or latitude" (Harper et al., 2011) involves a high sensitivity to climatic changes. Therewith, the development of suitable methods for monitoring these changes is of great importance and interest (Callaghan et al., 2002).

In Norway, the forest-tundra ecotone represents a large proportion of the total land area, which implies that remote sensing techniques have to be employed for monitoring. However, monitoring studies in such areas are limited by the spatial resolutions of optical remote sensing instruments because of the small-sized and sparsely distributed objects of interest. Assuming that trees located in the forest-tundra ecotone have a height growth of 0–5 cm per year (Næsset and Nelson, 2007), a remote sensing technique is required that is capable of detecting subtle changes in growth and colonisation patterns both further north and to higher elevations. In this context, airborne laser scanning (ALS) may provide a well-suited tool due to its capability of predicting biophysical parameters on single tree level at different scales (e.g. Hyyppä et al., 2001; Persson et al., 2002; Solberg et al., 2006; Næsset and Nelson, 2007). Concerning the discrimination of small pioneer trees in the forest-tundra ecotone, studies by Næsset and Nelson (2007), Rees (2007), and Thieme et al. (2011b) verified the suitability of ALS using different laser point densities. Provided a minimum height of 2 m, Rees (2007) found low-density laser data with a point density of ~0.25 m^{-2} useful to discern individual trees over large areas covering hundreds of square kilometres. Næsset and Nelson (2007) and Thieme et al. (2011b), however, used positive laser height values inside a field-measured tree crown polygon as criterion for successful tree detection irrespective of height employing high-density laser data (6.8–8.5 m^{-2}). Both studies report detection success rates of over 90% and at least 84% for coniferous and mountain birch trees with heights ≥1 m, respectively (Næsset and Nelson, 2007; Thieme et al., 2011b). For trees lower than 1 m, the numbers were significantly lower due to the fact that there are generally few other objects in such environments that are expected to be higher than 1 m above the ground surface. This implies an adequate reliability of successful tree detection when tree heights exceed 1 m.

However, severe commission errors may be introduced using laser height values as the sole criterion for tree detection (Næsset and Nelson, 2007; Næsset, 2009). A large number of laser echoes above the ground surface will often be reflections from non-tree objects such as rocks, hummocks, and other terrain structures. Thereby, the proportion of pulses reflecting off of tree and non-tree objects will be highly dependent on the structure of terrain, i.e. the more rugged the terrain the larger the number of non-tree laser echoes. Furthermore, properties of the terrain model, the sensor, and various flight settings will affect the magnitude of echoes from non-tree objects with positive height values (Næsset, 2009). Næsset and Nelson (2007) reported a commission error of 490% for a dataset with a terrain model derived with an iteration angle of 9°. Therefore, the reliability of tree detection analysis based on laser height values is highly influenced by such commission errors. In a multi-temporal context, however, trees may change in height and number, whereas terrain and terrain objects will remain stable. Thus, the reliability of statistical estimates of change for a given area (e.g. change in mean tree height or change in tree numbers) may not be compromised by commission errors.

Most studies applying ALS data for forest inventory have merely utilised the height information of the individual echoes of the laser point cloud rather than the full suite of available information, including spectral data, i.e., the intensity value of each echo. However, intensity may be a useful parameter in order to distinguish between tree and non-tree echoes. The usage of laser intensity has rarely been investigated since a study on tree species classifications in 1985 (Schreier et al., 1985) because of the lack of radiometric calibration methods (Kaasalainen et al., 2005). During the last decade, however, several studies were conducted to classify tree species (e.g. Brandtberg et al., 2003; Holmgren et al., 2008; Ørka et al., 2009), age (Farid et al., 2006a,b) and land-cover (Brennan and Webster, 2006). Thieme et al. (2011a) used normalised intensity values in an experimental study investigating the spatial pattern of tree and non-tree objects both related to laser height and intensity. Thieme et al. (2011a) showed promising results regarding the use of normalised intensity values and height information to separate between tree and non-tree objects.

Thieme et al. (2011a) also suggested that the spatial point pattern of the ALS data may be a potential co-discriminator for tree and non-tree objects by employing Voronoi polygons. In point pattern analysis, Voronoi polygons are commonly used to analyse the spatial distribution of point data in numerous disciplines (Boots and Getis, 1988). Furthermore, a variety of biological phenomena demonstrate spatial correlation or dependency (Rossi et al., 1992), often emerging in patches (Fry and Stephens, 2010). Thus, the spatial variation around tree and non-tree objects may differ for laser echoes classified as vegetation.

Different terrain characteristics as represented by aspect and slope have an influence on the potential presence and height growth of small pioneer trees (Mast et al., 1997; Boisvenue et al., 2004; Danby and Hik, 2007). Concerning a forest-tundra ecotone environment, Danby and Hik (2007) reported differing tree invasion patterns for north- and south-facing slopes that were primarily caused by the differential presence of permafrost. Furthermore, their study demonstrated that regional, landscape and local scale variability in tree population was partially depending on variations in the terrain, landscape setting and existing vegetation (Danby and Hik, 2007). Thus, aspect and slope parameters may give a contribution to discriminate between tree and non-tree laser echoes.

The main objective of this study was to assess the capability of high-density ALS data to classify tree and non-tree echoes directly from the laser point cloud using different ALS-derived variables individually and in combination. For this purpose, the following classification model variables were tested as discriminators: (1) laser height and intensity, (2) a geospatial facet represented by the respective areas of Voronoi polygons, and (3) the terrain variables aspect and slope. Finally, the accuracy and performance of the different models were assessed based on two different classification methods.

2. Study area and data

2.1 Study area

The study was carried out along a 1,000 km long and approximately 180 m wide longitudinal transect encompassing hundreds of mountain forest/alpine elevation gradients. The transect stretches from Mo i Rana in northern Norway (66°19' N 14°9' E) to Tvedestrand in the southern part of the country (58°3' N 9°0' E) (Figure 1) and covers sample plots in the transitions between mountain forest and the alpine zone, the forest-tundra ecotone. The terrain in such localities is often characterised by rounded forms with occurrences of hummocks, rocks and boulders, but also steep slopes. The prevalent tree species are Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and mountain birch (*Betula pubescens* ssp *czerepanovii*).

Figure 1 **-** Overview of the study area with the 25 specific field sites (black points). The 1,000 km long transect (black line) stretches from 66°19' N 14°9' E to 58°3' N 9°0' E.

2.2 Field data

The field work was conducted during summer 2008 in order to provide *in situ* tree data for analysis. Field data were collected at 25 different field sites allocated along the transect. Each field site covers the entire ecotone with two to four sample plots laid out and spaced 50 m apart in order to avoid overlap. The geographical extent of the area between the mountain forest and the alpine zone varies between the different locations and thus the number of plots in each site was determined in field according to a visual and practical judgement of the altitudinal range of the forest-tundra ecotone in each case. This resulted in a total number of 77 sample plots.

Real-time kinematic differential Global Navigation Satellite Systems (dGNSS) was used for precise navigation and positioning. This involved two Topcon Legacy E+ 20-channel dual-frequency receivers that observe pseudo range and carrier phase of both GPS (Global Positioning System) and GLONASS (Global Navigation Satellite System) satellites with an expected horizontal position accuracy of about 2 cm. For each field site, a base station was established at the closest suitable reference point of the Norwegian Mapping Authority. These reference points have an accuracy of ~3 cm and the expected accuracy of the centre points of the sample plots is 3-4 cm.

In the field, a modified version of the point-centred quarter sampling method (PCQ) (Cottam and Curtis, 1956; Warde and Petranka, 1981) was used to select the individual sample trees. For this purpose, each sample plot was divided into four quadrants by defining the cardinal directions from the sample plot's centre using a Suunto compass. Furthermore, three tree height classes were used during sampling: (1) lower than 1 m, (2) between 1 and 2 m, and (3) taller than 2 m. According to the scheme of PCQ sampling, in each quadrant the trees that were closest to the plot centre in their respective tree height classes were sampled. This procedure was implemented independent of tree species and with a maximum search distance of 25 m. For determining the maximum search limit as well as the closest tree a surveyor's tape measure was used in cases of doubt.

Several tree parameters were recorded individually for each tree. Tree height was measured using a steel tape measure or a Vertex III hypsometer for tall trees. Stem diameter was callipered at the root collar and crown diameters were measured in the cardinal directions with a steel tape measure. Furthermore, the precise tree position was determined using dGNSS.

In total, 533 trees were measured in this study. However, nine trees had tree crown areas that were completely overlapped by tree crown areas of taller trees. These nine trees were regarded as non-beneficial for the analyses and therewith discarded from the dataset. This resulted in a total number of 524 trees, i.e., 404 mountain birch, 67 Norway spruce, and 53 Scots pine. For these trees, tree heights ranged from 0.04 to 7.80 m, and crown areas from 0.001 to 19.54 m^2, computed as the ellipse defined by the crown diameters as the major and minor axes. A summary of the biophysical properties of the measured trees is given in Table 1.

Table 1 - Summary of field measurements of trees.

Tree species	Characteristics	*n*	Mean	Min.	Max.
Mountain birch	Height (m)	404	1.41	0.04	7.80
	Diameter (cm)	404	4.24	0.10	34.00
	Crown area (m²)	404	1.13	0.001	19.54
Norway spruce	Height (m)	67	1.67	0.07	7.00
	Diameter (cm)	65	6.54	0.20	19.10
	Crown area (m²)	67	1.45	0.006	5.69
Scots pine	Height (m)	53	1.33	0.10	5.10
	Diameter (cm)	53	5.00	0.30	18.90
	Crown area (m²)	53	0.81	0.002	7.28

2.3 Laser data

Airborne laser data were acquired on 23 and 24 July 2006 with an Optech ALTM 3100C laser scanning system carried by a Piper PA-31 Navajo aircraft. Average flying altitude was 800 m above ground level, and the flight speed was approximately 75 ms^{-1}. Furthermore, scan frequency was 70 Hz, maximum half scan angle was 7°, and the average footprint diameter was estimated to 20 cm. Pulse repetition frequency was 100 kHz resulting in a mean pulse density of 6.8 m^{-2}. In order to keep the flying altitude and therewith the pulse density as constant as possible, the 1,000 km long transect was split up into 98 individual flight lines.

Pre-processing of the laser data was conducted by a contractor (Blom Geomatics, Norway) computing planimetric coordinates (*x* and *y*) and ellipsoidal height values for all laser points.

Laser echoes labelled “last-of-many” and “single”, hereafter denoted as LAST, were used for the derivation of the terrain model, whereas laser echoes labelled as “first-of-many” and “single”, hereafter denoted as FIRST, were used in the analyses of the current study. The ALTM 3100C records up to four echoes per laser pulse with a minimum vertical distance of 2.1 m between these echoes. This instrument property in combination with low vegetation results in potentially very few pulses having more than a single echo. Therefore the LAST and FIRST datasets will for many of the sample plots be identical. Based on an iteration distance of 1.0 m and an iteration angle of 9°, ground echoes were classified from the planimetric coordinates and the corresponding height data of the LAST echoes using the TerraScan software (Terrasolid, 2011), and a triangulated irregular network (TIN) was derived. Echoes labelled as FIRST were projected onto the TIN surface in order to interpolate the corresponding terrain height on these locations. The differences between the FIRST echo

heights and the corresponding interpolated terrain height values were computed and stored. However, merely FIRST echoes with height values greater than zero were included in the subsequent analyses since this criterion is the sole indicator for the presence of objects on the terrain surface.

3. Methods

3.1 Computations

In order to assess the ability to classify FIRST laser echoes into tree and non-tree echoes a broad suite of variables were considered. Variables related to the direct laser measurements, a geospatial variable, and terrain variables were employed.

Prior to the computation, areas were defined for the extraction of tree and non-tree FIRST echoes. Concerning tree echoes, elliptical tree crown polygons were estimated from the field-measured crown diameters and all FIRST echoes falling inside the individual tree crown polygons were classified as tree echoes. Due to a positioning error of the laser data of up to 0.5 m as reported by the contractor, trees with a crown diameter value less than 1.0 m in at least one of the cardinal directions were assigned a tree crown circle with a diameter of 1.0 m.

Non-tree echoes were classified using full control areas without any trees as received by the sampling design of the PCQ method. Using the PCQ method, the trees closest to the respective plot centre were sampled for the different height classes in each quadrant resulting in a maximum of three sampled trees per quadrant. In order to receive full control areas merely consisting of FIRST echoes emerging from non-tree objects, only the distance to the closest tree in each quadrant irrespective of tree height class was used for the computation of polygons without any trees. Since the distance includes parts of the tree crown, the tree crown polygons were erased from the computed non-tree polygons. FIRST echoes falling inside these full control polygons were classified as non-tree echoes. This procedure resulted in 2,323 tree and 27,487 non-tree echoes.

From all FIRST echoes, the laser height and intensity values were used as variables for the classification. Laser height could be extracted directly, whereas the raw intensity (I_{Raw}) had to be normalised for the range R according to the following formula suggested by Korpela et al. (2010):

$$I_{Ran} = \left(\frac{R}{R_{Ref}}\right)^{2.4} \cdot I_{Raw} \quad (1)$$

where R_{Ref} is an average reference range that was set to 800 m in this study.

In order to include a geospatial facet into the classification analysis, Voronoi polygons were employed using the FIRST echoes. The computation of Voronoi polygons is based on a continuous space that is divided into regions ensuring that any other location in the space is concatenated with the closest point of the point pattern (Okabe et al., 2000). This means "that each Voronoi polygon consists of an area that is closer to a given point than any other point" (Wulder et al., 2006).

As mentioned earlier, Voronoi polygons are a commonly used technique in point pattern analysis and are applied to problems in numerous disciplines (Boots and Getis, 1988). The spatial distribution and correlation of a point pattern is reflected in the area of the respective Voronoi polygons, i.e. the smaller the Voronoi polygons the denser the point pattern. In the present study, small-sized polygons are expected to indicate trees, whereas larger polygons are assumed to merely represent noise in the ALS data (Thieme et al., 2011a). Although small-sized polygons also might suggest the presence of non-tree objects such as rocks and hummocks of a certain size, the combination with laser height and intensity values is expected discern between tree and non-tree echoes.

For the computation of the Voronoi polygons, an adequate area was defined for each field site in order to avoid edge effects at the sample plot borders. Voronoi polygons were then computed for the FIRST echoes and the area of each polygon was calculated. The FIRST echoes were subsequently overlaid with the Voronoi polygons and the estimated area of each polygon was assigned to the corresponding FIRST echo.

Concerning the terrain-related variables, a digital elevation model (DEM) was computed (QCoherent Software, 2010) using LAST echoes classified as ground returns in order to compute aspect and slope values (Burrough and McDonald, 1998). The DEM had a cell size of 0.25 m due to the small size of the objects in question, i.e., the small pioneer trees. Aspect and slope were derived from the DEM raster surface (Burrough and McDonald, 1998) and the values assigned to the corresponding FIRST echoes. Furthermore, aspect was divided into eight categories because of computational reasons (Table 2).

A summary of the used variables is given in Table 2. For tree echoes, the prevalent aspects were northeast and southwest-facing slopes, whereas for non-tree echoes they were facing northeast and north.

Table 2 - Summary of the discriminator variables.

Class	Variable	Mean	Min.	Max.
Tree	Height (m)	1.59	0.04	6.49
	Intensity	51.62	4.24	90.95
	Voronoi Polygon (m^2)	0.61	0.02	19.81
	Aspect	-	-	-
	Slope (°)	16.49	1.05	49.89
Non-tree	Height (m)	0.17	0.01	4.72
	Intensity	56.22	0.51	110.82
	Voronoi Polygon (m^2)	1.50	0.03	62.67
	Aspect	-	-	-
	Slope (°)	16.54	0.005	79.68

Note: Aspect was divided into eight categories: N (North), NE (Northeast), E (East), SE (Southeast), S (South), SW (Southwest), W (West), and NW (Northwest).

3.2 Analyses

Generalised linear models (GLM) and support vector machines (SVM) were employed for the classification of FIRST echoes into tree and non-tree echoes.

Besides regression analysis, GLM is also used for binary classification by prediction of probabilities on a transformed scale (Dalgaard, 2008). SVM were developed by Cortes and Vapnik (1995) based on statistical learning theory and represent a tool for classification, regression, and novelty detection (Karatzoglou et al., 2006; Meyer, 2011). For both classification methods, a leave-one-out cross-validation procedure was implemented separating the dataset into a training set and a testing set. That is, the classes of the FIRST echoes for a field site were predicted using a model calculated from the FIRST echoes of the other remaining field sites.

The models consisted of different combinations of the laser height and intensity variables, the geospatial variable represented by the area of the Voronoi polygons and the terrain-related variables aspect and slope in order to find the best suited predictors for both classification methods of tree and non-tree echoes. We decided to include the laser height and intensity variables in all possible combinations due to their direct relation to the FIRST echoes. Table 3 gives an overview of the variable combinations tested.

Table 3 - Combinations of laser metrics, geospatial and terrain variables used for classification.

Model	Variables
HI	Laser *H*eight + *I*ntensity
HIP	Laser *H*eight + *I*ntensity + Voronoi *P*olygons
HIA	Laser *H*eight + *I*ntensity + *A*spect
HIS	Laser *H*eight + *I*ntensity + *S*lope
HIPA	Laser *H*eight + *I*ntensity + Voronoi *P*olygons + *A*spect
HIPS	Laser *H*eight + *I*ntensity + Voronoi *P*olygons + *S*lope
HIAS	Laser *H*eight + *I*ntensity + *A*spect + *S*lope
HIPAS	Laser *H*eight + *I*ntensity + Voronoi *P*olygons + *A*spect + *S*lope

3.3 GLM

GLM are characterised by three elements: the random component identifying the response variable *Y* and its probability distribution, the systematic component specifying the independent variables *X*, and the link function connecting those two components (Agresti, 2007; Dalgaard, 2008). In this study, the binomial random component (tree/non-tree) was related to the different combinations of the independent variables *X* using a logit link function. Thus, the following model was estimated:

$$log\left(\frac{\pi(tree)}{1-\pi(tree)}\right) = \alpha + \beta_1 x_1 + \cdots + \beta_k x_k \quad (2)$$

The GLM model was fitted using the *glm* function of the *stats* package (R Development Core Team, 2007) for the eight different models (Table 3) in the statistical computing software R (R Development Core Team, 2007). Furthermore, from the fitted models the probabilities of the FIRST echoes being a non-tree echo were predicted. For each model, different thresholds (from p=0.05 to p=0.95 in 0.05 steps) for these probabilities were used to classify the FIRST echoes into tree and non-tree echoes. The Cohen's kappa coefficient (Cohen, 1960) was estimated for each threshold used and the classification with the highest Cohen's kappa coefficient for each model was selected.

3.4 SVM

In essence, SVM finds the hyperplane with the maximal margin of separation between the two classes using the training set by solving a quadratic optimisation problem. Thereby, the so called support vectors contain the relevant information used for the classification and are represented by the points located on the margin boundaries. Concerning overlapping

classes, points lying on the opposite side of the margin are reduced in influence by weighting. *C* represents a cost or penalty parameter controlling the error term. Furthermore, the hyperplane is defined by a kernel function which allows for a nonlinear separator. In this study, the *C*-support vector classification was used with the radial basis function as the kernel, where γ represents a parameter regulating the radial basis function.

The *svm* function of the *e1071* package (Dimitriadou et al., 2011) was used for fitting the eight different models (Table 3). From the fitted models prediction of the FIRST echoes being a tree or non-tree echo was performed. Prior to classification and outside the leave-one-out cross-validation procedure, the best hyperparameters *C* and γ were found using a grid search implemented in the *tune.svm* function of the *e1071* package (Dimitriadou et al., 2011; Karatzoglou et al., 2006).

3.5 Accuracy assessment and classification performance

For the assessment of the classification performances, the total percentage of correct prediction and the Cohen's kappa coefficient (Cohen, 1960) for each combination of prediction parameters were estimated for both classification methods. Furthermore, the kappa coefficients of the different models and classification methods (GLM versus SVM) were compared using a formula suggested by Cohen (1960) that evaluates the normal curve deviate to assess the significance of the difference between two independent kappa coefficients:

$$z = \frac{\kappa_1 - \kappa_2}{\sqrt{\sigma_{\kappa_1}^2 + \sigma_{\kappa_2}^2}} \quad (3)$$

Kappa coefficients were evaluated according to the grading suggested by Landis and Koch (1977).

4. Results and discussion

The classification of the FIRST echoes into tree and non-tree echoes using GLM and SVM revealed total accuracies of at least 93.6% (Table 4) and 94.8% (Table 6), respectively. These accuracies are in line with other studies on detection of small single trees in the forest-tundra ecotone by taking into account that the present study is based on the individual echoes and not on an individual tree basis. Success rates of at least 90% for trees taller than 1 m are reported by Næsset (2009), Næsset and Nelson (2007), and Thieme et al. (2011b) which are consistent with the results of this study. Kappa coefficients of at least 0.515 (Table 4) and

0.560 (Table 6) for the estimated models indicated moderate fits using the GLM and SVM classifications, respectively.

4.1 GLM

For the GLM classification, total accuracies ranged between 93.6% and 94.9% (Table 4) for the eight different models. Furthermore, kappa coefficients ranged from 0.515 to 0.573 indicating moderate fits for all the estimated models (Table 4).

With regard to the total accuracies, a maximum difference of 1.3 percentage points was found for the eight models (Table 4). In this case, models not including the terrain variable slope (*HI*, *HIP*, *HIA* and *HIPA*) had slightly higher accuracies, and the model with the highest accuracy of 94.9% was represented by the basic model *HI* merely including the laser height and intensity variables. Concerning the corresponding kappa coefficients, models not including the terrain variable slope (*HI*, *HIP*, *HIA* and *HIPA*) featured a generally higher performance and model fit. The basic model *HI* with the highest accuracy also revealed the highest kappa coefficient.

Table 4 - Performance of the different models used for classification with GLM.

Model	*p*	Accuracy	Kappa
HI	0.75	0.949	0.573
HIP	0.75	0.948	0.569
HIA	0.75	0.948	0.568
HIS	0.80	0.942	0.550
HIPA	0.75	0.946	0.559
HIPS	0.75	0.943	0.546
HIAS	0.75	0.939	0.523
HIPAS	0.75	0.936	0.515

Comparing the kappa coefficients of the eight models, no significant contribution of the geospatial variable or the terrain variables on the classification performance could be found (Table 5 and Figure 2). Albeit the kappa coefficients of models including Voronoi polygons and/or aspect (*HIP*, *HIA* and *HIPA*) implied similar classification performances as the basic model *HI*, no improvement in classification was gained from the inclusion of the additional prediction variables. Concerning the terrain variable aspect, this result is different from findings reported in other studies on the potential presence of small pioneer trees, where especially aspect had an important influence on tree invasion (Mast et al. 1997; Danby and

Hik, 2007). However, both studies used aspect on a coarser scale detecting a more generalised trend than in the present study. Moreover, the geospatial variable based on Voronoi polygon sizes in combination with the laser height and intensity variables did not contribute significantly, even though Thieme et al. (2011a) showed promising results in regard to the indication of objects located above the terrain surface.

Table 5 - Test of significance between independent kappa indices for GLM.

Model 1	**Model 2**	***Z***	
HI	*HIP*	0.322	
	HIA	0.388	
	HIS	1.731	'
	HIPA	1.044	
	HIPS	1.974	*
	HIAS	3.647	**
	HIPAS	4.265	**
HIP	*HIA*	0.066	
	HIS	1.407	
	HIPA	0.721	
	HIPS	1.652	'
	HIAS	3.324	**
	HIPAS	3.941	**
HIA	*HIS*	1.340	
	HIPA	0.655	
	HIPS	1.586	
	HIAS	3.258	**
	HIPAS	3.874	**
HIS	*HIPA*	0.681	
	HIPS	0.256	
	HIAS	1.938	'
	HIPAS	2.554	*
HIPA	*HIPS*	0.931	
	HIAS	2.604	*
	HIPAS	3.218	**
HIPS	*HIAS*	1.672	'
	HIPAS	2.283	*
HIAS	*HIPAS*	0.605	

Note: Level of significance: '<.1. *<.05. **<.005

Furthermore, three of the models including the terrain variable slope (*HIPS*, *HIAS* and *HIPAS*) had significantly worse performances than the basic model *HI*, and equally or significantly worse performances than all other remaining models (*HIP*, *HIA* and *HIPA*) (Table 5 and Figure 2). This result may be attributed to the similar ecological meaning of the terrain variables aspect and slope. Both variables describe amongst others the solar radiation and moisture (Bader and Ruijten, 2008), which are essential factors for the occurrence of trees. In regard to the classification performances of the three aforementioned models, the terrain variable aspect may already cover these characteristics in this linear classification method.

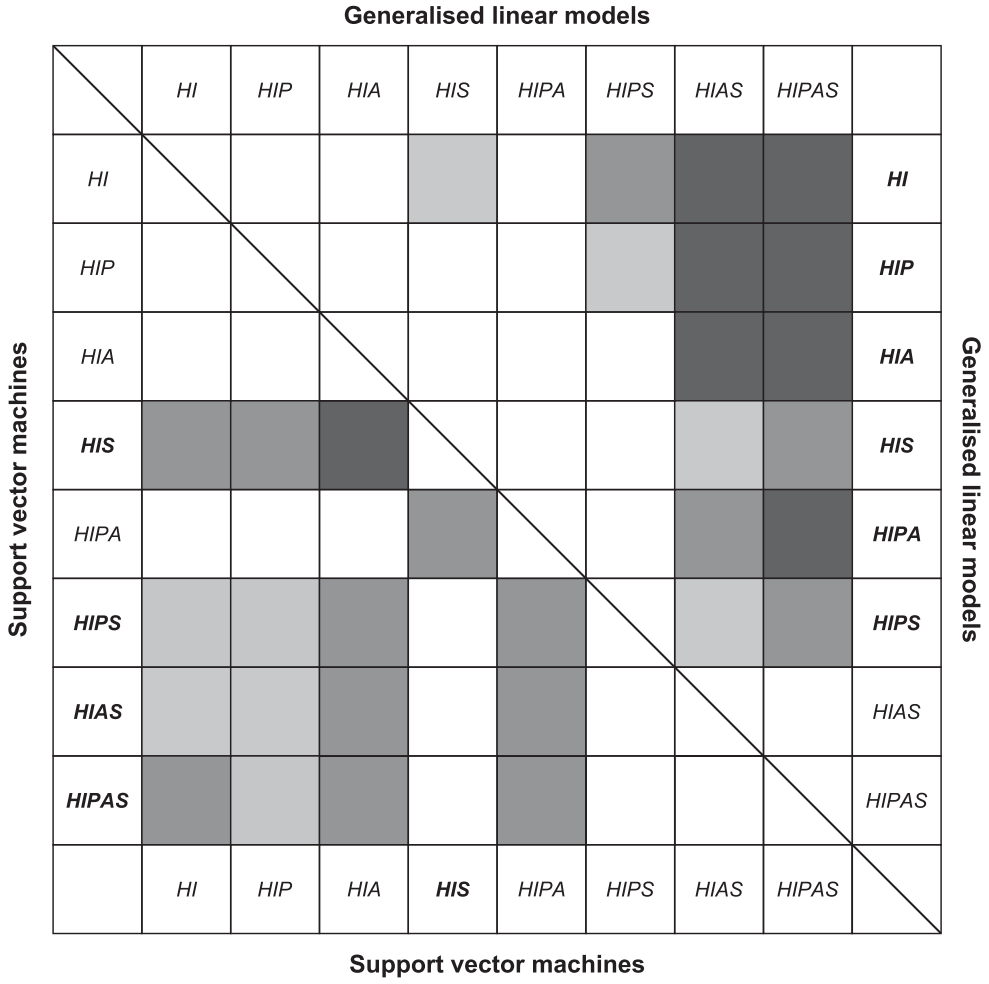

Figure 2 **-** Classification performances for the eight different models using GLM and SVM. The dark-grey areas indicate a significant difference in performance between two models at a $p<0.005$ level, medium grey areas at a $p<0.05$ level, and light grey areas at a $p<0.1$ level. White areas indicate no difference in performance. Models highlighted in bold indicate the superior models in the respective classification method. NOTE: For each classification method, the performances of the models on the vertical axes were compared to the performance of the models on the horizontal axes.

4.2 SVM

The classification of the FIRST echoes into tree and non-tree echoes using SVM revealed total accuracies ranging from 94.8% to 95.3% (Table 6) for the eight different models. Kappa coefficients ranged between 0.560 and 0.600 indicating moderate fits for all the models (Table 6).

Concerning the total accuracies for SVM, the eight models had a maximum difference of 0.5 percentage points (Table 6). Thereby, models not including the terrain parameter slope (*HI*, *HIP*, *HIA*, and *HIPA*) had slightly lower accuracies. However, a serious inferior performance of these models could not be found. The model with the highest accuracy of 95.3% was represented by the model including the variables laser height and intensity as well as the terrain variable slope (*HIS*). With regard to the corresponding kappa coefficients, models including the terrain parameter slope (*HIS*, *HIPS*, *HIAS*, and *HIPAS*) revealed a generally better performance for the classification of tree and non-tree echoes. The best kappa coefficient was found for the model *HIS* which featured the best total accuracy as well.

Table 6 - Performance of the different models used for classification with SVM.

Model	*C*	γ	Accuracy	Kappa
HI	1000	0.1	0.949	0.568
HIP	1000	0.1	0.949	0.569
HIA	1000	0.1	0.948	0.560
HIS	1000	0.1	0.953	0.600
HIPA	1000	0.1	0.948	0.563
HIPS	1000	0.1	0.952	0.594
HIAS	100	0.1	0.952	0.594
HIPAS	100	0.1	0.952	0.596

Furthermore, a significantly higher accuracy for classifications including the slope variable was obtained, whereas Voronoi polygons and aspect did not have a significant influence on the classification performance (Table 7 and Figure 2). For instance, for the basic model *HI* merely consisting of the laser height and intensity variables, the error matrices were not significantly different from models including Voronoi polygons and/or aspect in addition (*HIP*, *HIA* and *HIPA*) (Table 7 and Figure 2). As mentioned before, this finding contrasts to the studies conducted by Danby and Hik (2007) and Mast et al. (1997), who are both reporting an importance of terrain aspect on the potential presence of small pioneer trees, though at a coarser scale. However, all models including the terrain variable slope (*HIS*,

HIPS, *HIAS*, and *HIPAS*) revealed significantly better performances than the basic model *HI* (Table 7 and Figure 2). This fact implies the assumption, that slope is significantly contributing to the classification performances when utilizing a nonlinear kernel to construct the hyperplanes.

Table 7 - Test of significance between independent kappa indices for SVM.

Model 1	Model 2	*Z*	
HI	*HIP*	0.111	
	HIA	0.526	
	HIS	2.387	*
	HIPA	0.341	
	HIPS	1.953	'
	HIAS	1.927	'
	HIPAS	2.058	*
HIP	*HIA*	0.638	
	HIS	2.279	*
	HIPA	0.453	
	HIPS	1.845	'
	HIAS	1.819	'
	HIPAS	1.950	'
HIA	*HIS*	2.910	**
	HIPA	0.187	
	HIPS	2.478	*
	HIAS	2.451	*
	HIPAS	2.582	*
HIS	*HIPA*	2.731	*
	HIPS	0.438	
	HIAS	0.458	
	HIPAS	0.330	
HIPA	*HIPS*	2.297	*
	HIAS	2.271	*
	HIPAS	2.402	*
HIPS	*HIAS*	0.022	
	HIPAS	0.107	
HIAS	*HIPAS*	0.128	

Note: Level of significance: '<.1. *<.05. **<.005

4.3 GLM versus SVM

Concerning the comparison of the classification performances of SVM and GLM for the eight different models, both significant and non-significant differences were found (Table 8). For all models consisting of combinations of the laser height and intensity variables, the geospatial variable and the terrain variable aspect (*HI*, *HIP*, *HIA* and *HIPA*) *p*-values ranged between 0.569 and 0.997 when comparing the respective kappa coefficients (Table 8). This implies no significant difference between the two different classification methods. However, for all models including the terrain variable slope (*HIS*, *HIPS*, *HIAS* and *HIPAS*), all *p*-values were smaller than 0.001 indicating significantly better classification performances for SVM in all the four cases (Table 8). These results correspond well with the findings made for the respective classification methods, where the terrain variable slope did not improve the GLM classification but had a significant positive influence on the SVM classification. The basic model including the laser height and intensity variables and models consisting of the geospatial and/or the terrain variable aspect in addition, revealed similar classification performances for the two classification methods. Therewith, the terrain variable slope is discriminating the linear (GLM) and the nonlinear (SVM) classification.

Both the similarity in performance between the two classification methods and the dissimilarity seem reasonable due to the fact that SVM is similar to logistic regression under certain constrains, however, extending it to the nonlinear case (Venables and Ripley, 2002). Based on the results from the comparisons of classification performances within and between the classification methods, it is reasonable to assume that a nonlinear classification is sensitive to the slope variable.

Table 8 - Test of significance between independent kappa indices for SVM versus GLM.

SVM	GLM	*Z*	
HI	*HI*	0.428	
HIP	*HIP*	0.004	
HIA	*HIA*	0.570	
HIS	*HIS*	3.719	**
HIPA	*HIPA*	0.268	
HIPS	*HIPS*	3.514	**
HIAS	*HIAS*	5.165	**
HIPAS	*HIPAS*	5.930	**

Note: Level of significance: '<.1. *<.05. **<.005

5. Conclusion

To conclude, the capability of high-density ALS data to classify tree and non-tree echoes directly from the laser point cloud be verified. The classification of tree and non-tree echoes both employing GLM and SVM using different combinations of variables representing laser height and intensity, spatial pattern and terrain characteristics revealed promising results in this study. Tree and non-tree echoes could be classified with a total accuracy of at least 93% and a moderate fit irrespective of the classification method or model used.

With regard to laser height and intensity, these two variables represented important discriminators for the classification of tree and non-tree echoes. Using GLM, the basic model merely including height and intensity revealed both the best accuracy and the highest kappa coefficient.

Concerning the spatial variable represented by Voronoi polygons, this discriminator did not have a significant influence on the classification performances both in regard to the GLM and the SVM classification method.

Pertaining to the terrain variables aspect and slope, no significant contribution could be found using GLM. However, with regard to the SVM classification method, slope revealed as a significant discriminator for the classification performances.

Total accuracies were highest for models consisting of combinations of the laser height and intensity, the spatial variable as well as aspect using the GLM classification method, whereas models including the terrain variable slope had slightly higher accuracies using SVM.

Kappa coefficient analysis indicated moderate fits for all models using GLM, however, three of the models including the terrain variable slope had significantly worse kappa coefficients compared to the basic model. Concerning SVM, all models performed with a moderate fit as indicated by the kappa coefficient with significantly better values for models including the slope variable, reflecting the same tendency as observed for the total accuracies.

Comparing GLM and SVM classification performances, the two classification methods do not significantly differ for models consisting of different combinations of the laser height and intensity variables, the geospatial variable and the terrain variable aspect. However, concerning models including the terrain variable slope, SVM revealed a superior performance for the classification of tree and non-tree echoes using high-density airborne laser scanning data.

In summary, high-density ALS data is well-suited for the classification of tree and non-tree echoes in the forest-tundra ecotone. Depending on the classification method, satisfying results can be obtained employing relatively few and straightforward variables derived from the ALS data. However, to utilise such a classification for monitoring purposes, additional challenges using different sensors and acquisition settings is present (Næsset, 2009). Thus, field surveys are necessary to obtained for all ALS acquisitions. Furthermore, the time between individual inventories are limited by the tree growth, regeneration and mortality, together with the positional accuracy of individual echoes. Then, this technique provides a time- and cost-efficient monitoring tool to detect small single trees in the forest-tundra ecotone. Another question is aggregating the individual classified echoes to another scale e.g. individual tree level, a raster in different scale, depending on the monitoring needs. Furthermore, the technique may be utilized in a sampling frame work providing regional estimates (Falkowski et al., 2009; Gobakken et al., 2012) or map products (Ørka et al., 2012).

Acknowledgments

This research has been funded by the Research Council of Norway (project #184636/S30). We wish to thank Blom Geomatics AS, Norway, for collection and processing of the airborne laser scanner data. Thanks also appertain to Mr. Vegard Lien at the Norwegian University of Life Sciences, who was responsible for the fieldwork.

References

ACIA. 2004. *Impacts of a warming Arctic: Arctic Climate Impact Assessment*. Cambridge University Press, Cambridge, UK. 146 pp.

Agresti, A. 2007. *An introduction to categorical data analysis*. Wiley & Sons, Inc., Hoboken, New Jersey, USA. 400 pp.

Bader, M.Y., and Ruijten, J.J.A. 2008. A topography-based model of forest-cover at the alpine tree line in the tropical Andes. *Journal of Biogeography*, Vol. 35, pp. 711–723.

Boisvenue, C., Temesgen, H., and Marshall, P. 2004. Selecting a small tree height growth model for mixed-species stands in the southern interior of British Columbia, Canada. *Forest Ecology and Management*, Vol. 202, pp. 301–312.

Boots, B.N., and Getis, A. 1988. *Point Pattern Analysis*. Scientific Geography Series, Volume 8, Sage Publications, Newbury Park. 93 pp.

Brandtberg, T., Warner, T.A., Landenberger, R.E., and McGraw, J.B. 2003. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density LIDAR data from the eastern deciduous forest in North America. *Remote Sensing of Environment*, Vol. 85, No. 3, pp. 290–303.

Brennan, R., and Webster, T.L. 2006. Object-oriented land cover classification of lidar-derived surfaces. *Canadian Journal of Remote Sensing*, Vol. 32, No.2, pp. 162–172.

Burrough, P.A., and McDonell, R.A. 1998. *Principles of Geographical Information Systems*. Oxford University Press, New York. 356 pp.

Callaghan, T.V., Werkman B.R., and Crawford, R.M.M. 2002. The tundra-taiga interface and its dynamics: Concepts and applications. *Ambio Special Report*, Vol. 12, pp. 6–14.

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Management*, Vol. 20, No. 1, pp. 37–46.

Cortes C., and Vapnik, V. 1995. Support-vector networks. *Machine Learning*, Vol. 20, pp. 273–297.

Cottam, G., and Curtis, J.T. 1956. The use of distance measures in phytosociological sampling. *Ecology*, Vol. 37, No. 3, pp. 451–460.

Dalgaard, P. 2008. *Introductory Statistics with R*. Springer Science+Business Media, LLC., New York, USA. 364 pp.

Danby, R.K., and Hik, D.S. 2007. Variability, contingency and rapid change in recent subarctic alpine tree line dynamics. *Journal of Ecology*, Vol. 95, pp. 352– 363.

Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D., and Weingessel, A. 2011. *e1071: Misc Functions of the Department of Statistics (e1071), TU Wien*. Available from http://cran.r-project.org/web/packages/e1071/ [accessed 26 September 2011].

Falkowski, M.J., Wulder, M.A., White, J.C., and Gillis, M.D. 2009. Supporting large-area, sample-based forest inventories with very high spatial resolution satellite imagery. *Progress in Physical Geography,*Vol. 33, No. 3, pp. 403-423.

Farid, A., Goodrich, D.C., and Sorooshian, S. 2006a. Using airborne LIDAR to discern age classes of cottonwood trees in a riparian area. *Western Journal of Applied Forestry*, Vol. 21, No. 3, pp. 149−158.

Farid, A., Rautenkranz, D., Goodrich, D.C., Marsh, S.E., and Sorooshian, S. 2006b. Riparian vegetation classification from airborne laser scanning data with an emphasis on cottonwood trees. *Canadian Journal of Remote Sensing*, Vol. 32, No. 1, pp. 15−18.

Fry, D.L., and Stephens, S.L. 2010. Stand-level spatial dependence in an old-grwoth Jeffrey pine – mixed conifer forest, Sierra San Pedro Mártir, Mexico. *Canadian Journal of Forest Research*, Vol. 40, pp. 1803–1814.

Gobakken, T., Næsset, E., Nelson, R., Bollandsås , O.M., Gregoire, T.G., Ståhl, G., Holm, S., Ørka, H.O., and Astrup, R. 2012. Estimating biomass in Hedmark County, Norway using national forest inventory field plots and airborne laser scanning. *Remote Sensing of Environment*. Vol. 123, pp. 443–456.

Harper, K.A., Danby, R.K., De Fields, D.L., Lewis, K.P., Trant, A.J., Starzomski, B.M., Savidge, R., and Hermanutz, L. 2011. Tree spatial pattern within the forest-tundra ecotone: a comparison of sites across Canada. *Canadian Journal of Forest Research*, Vol. 41, pp. 479–489.

Holmgren, J., Persson, A., and Söderman, U. 2008. Species identification of individual trees by combining high resolution LIDAR data with multi-spectral images. *International Journal of Remote Sensing*, Vol. 29, No. 5, pp. 1537−1552.

Hyyppä, J., Kelle, O., Lehikoinen, M., and Inkinen, M. 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 5, pp. 969–975.

Kaasalainen, S., Ahokas, E., Hyyppa, J., and Suomalainen, J. 2005. Study of surface brightness from backscattered laser intensity: Calibration of laser data. *IEEE Geoscience and Remote Sensing Letters*, Vol. 2, No. 3, pp. 255–259.

Karatzoglou, A., Meyer, D., and Hornik, K. 2006. Support Vector Machines in R. *Journal of Statistical Software*, Vol. 15, No. 9, pp. 1–28.

Kirschbaum, M., and Fischlin, A. 1996. Climate change impacts on forests. In *Climate change 1995 - Impacts, adaptations and mitigation of climate change: scientific-technical analysis. Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel of Climate Change*. Edited by Watson, R., Zinyowera, M.C., and Moss, R.H. Cambridge University Press, Cambridge, UK. pp. 99–129.

Korpela, I., Ørka, H.O., Maltamo, M., Tokola, T., and Hyyppä, J. 2010. Tree species classification using airborne LiDAR – effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica*, Vol. 44, No. 2, pp. 319–339.

Landis, J.R., and Koch, G.G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, Vol. 33, No. 1, pp. 159–174.

Mast, J.N., Veblen, T.T., and Hodgson, M.E. 1997. Tree invasion within a pine/grassland ecotone: an approach with historic aerial photography and GIS modeling. *Forest Ecology and Management*, Vol. 93, pp. 181–194.

Meyer, D. 2011. *Support Vector Machines: The Interface to libsvm in package e1071*. Available from http://cran.r-project.org/web/packages/e1071/ [accessed 26 September 2011].

Næsset, E. 2009. Influence of terrain model smoothing and flight and sensor configurations on detection of small pioneer trees in the boreal-alpine transition zone utilizing height metrics derived from airborne scanning lasers. *Remote Sensing of Environment*, Vol. 113, pp. 2210–2223.

Næsset, E., and Nelson, R. 2007. Using airborne laser scanning to monitor tree migration in the boreal-alpine transition zone. *Remote Sensing of Environment*, Vol. 110, pp. 357–369.

Okabe, A., Boots, B., Sugihara, K., and Chiu, S.N. 2000. *Spatial Tesselations: Concepts and applications of Voronoi diagrams*. John Wiley & Sons, Chichester. 696 pp.

Ørka, H.O., Næsset, E., Bollandsås, O.M. 2009. Classifying species of individual trees by intensity and structure features derived from airborne laser skanner data. *Remote Sensing of Environment*, Vol. 113, pp. 1163–1174.

Ørka, H.O., Wulder, M.A., Gobakken, T., and Næsset, E. 2012. Subalpine zone delineation using LiDAR and Landsat imagery. *Remote Sensing of Environment,* Vol. 119, pp. 11–20.

Persson, Å., Holmgren, J., and Söderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, Vol. 68, No. 9, pp. 925–932.

QCoherent Software 2010. *Getting started with LP360*. QCoherent Software, Colorado Springs, USA. 12pp. Available from www.qcoherent.com [accessed 08 February 2012].

R Development Core Team. 2007. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rees, W.G. 2007. Characterisation of arctic treelines by LiDAR and multispectral imagery. *Polar Record*, Vol. 43, No. 227, pp. 345–352.

Rossi, R.E., Mulla, D.J., Journel, A.G., and Franz, E.H. 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecological Monographs*, Vol. 62, No. 2, pp. 277–314.

Schreier, H., Lougheed, J., Tucker, C., and Leckie, D. 1985. Automated measurements of terrain reflection and height variations using an airborne infrared laser system. *International Journal of Remote Sensing*, Vol. 6, No. 1, pp. 101–113.

Solberg, S., Næsset, E., and Bollandsås, O.M. 2006. Single tree segmentation using airborne laser scanner data in a heterogeneous spruce forest. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 12, pp. 1369–1378.

Stenseth, N.C., Mysterud, A., Ottersen, G., Hurrell, J.W., Chan, K.S., and Lima, M. 2002. Ecological effects of climate fluctuations. *Science*, Vol. 297, pp. 1292–1296.

Terrasolid. 2011. *TerraScan User's Guide*. Terrasolid Ltd., Jyväskylä, Finland. 311 pp. Available from www.terrasolid.fi [accessed 26 September 2011].

Thieme, N., Bollandsås, O.M., Gobakken, T., and Næsset, E. 2011a. Assessing spatial variation for tree and non-tree objects in a forest-tundra ecotone in airborne laser scanning data. In *SilviLaser 2011: 11th International Conference on LiDAR Applications for Assessing Forest Ecosystems*. 16-20 October 2011, Hobart, Australia. pp. 325–332. Available from http://www.iufro.org/science/divisions/division-4/40000/40200/40205/publications/ [accessed 1 February 2012].

Thieme, N., Bollandsås, O.M., Gobakken, T., and Næsset E. 2011b. Detection of small single trees in the forest-tundra ecotone using height values from airborne laser scanning. *Canadian Journal of Remote Sensing*, Vol. 37, No. 3, pp. 264–274.

Venables W.N., and Ripley, B.D. 2002. *Modern applied statistics with S.* Springer Science+Business Media, Inc., New York, USA. 512 pp.

Warde, W., and Petranka, J.W. 1981. A correction factor table for missing point-center quarter data. *Ecology*, Vol. 62, No. 2, pp. 491–494.

Woodall, C.W., Oswalt, C.M., Westfall, J.A., Perry, C.H., Nelson, M.D., and Finley, A.O. 2009. An indicator of tree migration in forests of the eastern United States. *Forest Ecology and Management*, Vol. 257, pp. 1434–1444.

Wulder, M.A., White, J.C., Dymond, C.C., Nelson, T., Boots, B., and Shore, T.L. 2006. Calculating the risk of mountain pine beetle attack: a comparison of distance- and density-based estimates of beetle pressure. *Journal of Environmental Informatics*, Vol. 8, No. 2, pp. 58–69.

PAPER III

Improving classification of airborne laser scanning echoes in the forest-tundra ecotone using geostatistical and statistical measures

Authors

Nadja Thieme*, Ole Martin Bollandsås, Terje Gobakken, Erik Næsset

Address

Norwegian University of Life Sciences

Department of Ecology and Natural Resource Management

P.O. Box 5003, N-1432 Ås, Norway

*** Corresponding author**

Nadja Thieme

E-mail: nadja.thieme@umb.no

Phone: +47 64 96 57 20

Abstract

The forest-tundra ecotone occupies a large proportion of the land area of Norway. Because of its temperature sensitivity, the vegetation in this transition between the mountain forest and the alpine zone is expected to be highly affected by climate change and requires effective monitoring techniques. Airborne laser scanning (ALS) has been proposed as a tool for the detection of small pioneer trees for such vast areas using laser height and intensity data. The main objective of the present study was to assess a possible improvement in the performance of classifying laser echoes into tree and non-tree echoes from high-density ALS data. The data were collected along a 1,000 km long transect stretching from southern Norway (58°3' N) to the northern part of the country (66°19' N). Different geostatistical and statistical measures derived from laser height and intensity values, i.e. the mean semivariance, the arithmetic mean, the standard deviation, and the coefficient of variation, were used to extent and potentially improve more simple models ignoring the spatial context. Generalised linear models (GLM) and support vector machines (SVM) were employed as classification methods. Total accuracies and Cohen's kappa coefficients were calculated and compared to those of simpler models from a previous study. For both classification methods, all models revealed total accuracies similar to the results of the simpler models. Concerning classification performance, however, the comparison of the kappa coefficients indicated a significant improvement for some models both using GLM and SVM. The highest kappa coefficient was found for SVM models including the mean semivariance derived from the laser height values. This implied a high potential as discriminator for tree and non-tree laser echoes, which was further supported by the results of the GLM classification.

Keywords: *ALS, classification, forest-tundra ecotone, GLM, SVM.*

1. Introduction

Particularly in the boreal regions, forest ecosystems are expected to be highly affected by increasing temperatures caused by climatic changes (Kirschbaum and Fischlin, 1996). As "the transition zone between forest and tundra at high elevation or latitude" (Harper et al., 2011), the forest-tundra ecotone entails a high sensitivity to these climatic changes, and alpine and arctic tree lines are expected to advance both to higher altitudinal and latitudinal areas because of changes in temperature, precipitation, and snow coverage (ACIA, 2004). Furthermore, anthropogenic factors in terms of herbivore activity and pastoral economy affect the tree limit beside the natural causes (Callaghan et al., 2002; Holtmeier & Broll, 2005). To monitor these abiotic and biotic changes, the development of suitable methods is essential (Callaghan et al., 2002).

A large proportion of the total land area in Norway is constituted by the forest-tundra ecotone. For such vast areas, cost-efficient motoring will most likely have to involve remote sensing techniques. However, the small size and sparse distribution of the objects of interest limit the monitoring capabilities of most optical remote sensing instruments because of their limited spatial resolutions. Trees located in the forest-tundra ecotone have an assumed height growth of 0–5 cm per year (Næsset and Nelson, 2007) and a remote sensing technique with the capability to detect subtle changes in growth and colonisation patterns in the forest-tundra ecotone is therefore needed. In this context, airborne laser scanning (ALS) may be a well-suited tool for monitoring changes regarding tree migration both further north and to higher altitudes. Several studies on the prediction of biophysical parameters have documented the suitability of ALS on a single-tree level at different scales (e.g. Hyyppä et al., 2001; Persson et al., 2002; Solberg et al., 2006; Næsset and Nelson, 2007). Furthermore, Næsset and Nelson (2007), Rees (2007), and Thieme et al. (2011b) verified the capability of ALS to discriminate small pioneer trees in the forest-tundra ecotone using different laser point densities. Rees (2007) demonstrated the utility of low-density laser data over hundreds of square kilometres with a point density of ~0.25 m^{-2} for the discrimination of individual trees with a minimum tree height of 2 m. Based on positive laser height values as a criterion for successful tree detection inside field-measured tree crown polygons, Næsset and Nelson (2007) and Thieme et al. (2011b) verified the suitability of high-density laser data (6.8–8.5 m^{-2}) for the detection of small pioneer trees irrespective of tree height. Detection success rates of over 90% for coniferous and at least 84% for mountain birch trees were reported for trees with tree heights ≥1 m (Næsset and Nelson, 2007; Thieme et al., 2011b), implying an adequate reliability of

successful tree detection for tree heights exceeding 1 m. However, severe commission errors may occur using laser height values as the sole criterion for tree detection (Næsset and Nelson, 2007; Næsset, 2009), which is reflected in the significantly lower detection success rates for trees lower than 1 m (Næsset and Nelson, 2007; Thieme et al., 2011b). Non-tree objects such as rocks, hummocks, and other terrain structures account for a large number of laser echoes above the ground surface, but the magnitude of non-tree echoes with positive laser height values also depends on the properties of the terrain model, the sensor, and flight settings (Næsset, 2009). For a dataset with a terrain model that was derived with commonly adopted smoothing criteria, Næsset and Nelson (2007) reported a commission error of 490%. Thus, the reliability of tree detection analysis using laser height values is highly dependent on these commission errors. However, in a multi-temporal context, terrain and terrain objects will remain stable while trees may increase in height and number over a sufficient time span. Thus, for monitoring the high rates of commission errors may not necessarily undermine the potentials of the technology.

With regard to forest inventory utilising ALS data, it is more common to merely employ the height information of the laser echoes instead of using the full suite of available information. Spectral data, i.e., the intensity values of the laser echoes, are often neglected, however, this additional information may be useful to discriminate between tree and non-tree echoes. Furthermore, the spatial structure and distribution of the individual laser echoes may be conducive to distinguish between different types of objects located on the terrain surface. Rossi et al. (1992) stated that a variety of biological phenomena demonstrate spatial correlation or dependency, often emerging in patches (Fry and Stephens, 2010). Hence, the spatial variation of laser echoes classified as vegetation may differ around tree and non-tree objects. For example, Thieme et al. (2011a) were able to recognise field-measured trees and non-tree objects identified using aerial imagery by investigating the spatial pattern of laser height and intensity values for small-sized Voronoi polygons and their neighbourhood in an empirical study. Also a geostatistical analysis employing experimental variograms and cross-variograms revealed differences in the pattern for tree and non-tree objects in that study (Thieme et al., 2011a). In optical remote sensing, geostatistics are a common image processing technique. For instance, standard statistical measures such as mean and standard deviation, and the variogram-derived mean semivariance are calculated for each pixel based on a moving window and further used for image classification purposes (Wulder et al., 1998; Jakomulska and Clarke, 2001). Wulder et al. (1998) used first- and second-order texture as well as semivariance moment texture for textural image classification to increase leaf area

index estimation. Thus, we hypothesize that standard statistical measures as well as a geostatistical component may have the potential to improve the classification of tree and non-tree laser echoes in the forest-tundra ecotone.

The main objective of this study was to assess the capability of geostatistical and standard statistical measures derived directly from high-density ALS data to improve the classification of tree and non-tree echoes. For this purpose, the following variables were derived from laser height and intensity values using a moving window and tested as discriminators in different classification models: (1) a geostatistical measure represented by the variogram-derived mean semivariance, and (2) standard statistical measures represented by the arithmetic mean, the standard deviation and the coefficient of variation. Based on two different classification methods, the accuracy and performance of the diverse models were assessed and finally compared to simpler models from a previous study (Thieme et al., 2012).

2. Study area and data

2.1 Study area

The study area covered a 1,000 km long and approximately 180 m wide longitudinal transect encompassing hundreds of mountain forest and alpine elevation gradients. The transect stretches from Mo i Rana in northern Norway (66°19' N 14°9' E) to Tvedestrand in the southern part of the country (58°3' N 9°0' E) (Figure 1). Sample plots were established in the forest-tundra ecotone, that is the transition between the mountain forest and the alpine zone. In most of the localities along the transect, the terrain was characterised by rounded forms with occurrences of hummocks, rocks and boulders, but also some steep slopes. The prevalent tree species were Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and mountain birch (*Betula pubescens* ssp *czerepanovii*).

2.2 Field data

The field work in the transect was carried out at 26 different field sites allocated along the transect during summer 2008 in order to provide *in situ* tree data for analysis.

Each field site consists of two to four sample plots to cover the width of the forest-tundra ecotone. Because the width of the forest-tundra ecotone varies between different locations, the number of sample plots in each site was determined in field based on both visual and practical judgement of the altitudinal range of the ecotone in each case. Furthermore, sample plots within field sites were laid out with 50 m interdistance to avoid overlap. These

procedures resulted in a total number of 80 sample plots. However, one field site and therefore three sample plots had to be discarded from the dataset because of erroneous coordinates. This resulted in a total number of 77 sample plots located at 25 different field sites.

Figure 1 - Overview of the study area with the 25 specific field sites (black points). The 1,000 km long transect (black line) stretches from to 66°19' N 14°9' E to 58°3' N 9°0' E.

Two Topcon Legacy E+ 20-channel dual-frequency receivers observing pseudo range and carrier phase of both Global Positioning System and Global Navigation Satellite System satellites were used as base and rover receivers for real-time kinematic differential Global Navigation Satellite Systems (dGNSS) navigation and positioning. For each field site, the closest suitable reference point of the Norwegian Mapping Authority was selected to establish the base station. With an expected accuracy of the reference points of 3 cm and an

expected horizontal accuracy of the field recordings relative to the base station of about 2 cm, the expected accuracy of the centre points of the sample plots was 3-4 cm.

For the selection of the sample trees in the field, a modified version of the point-centred quarter sampling method (PCQ) (Cottam and Curtis, 1956; Warde and Petranka, 1981) was used with a maximum search distance of 25 m. This sampling method involves the division of a sample plot into four quadrants defined by the cardinal directions from the centre of the sample plot. In each quadrant the tree that was closest to the plot centre in a specific tree height class was sampled independent of tree species. The tree height classes were defined as: (1) less than 1 m, (2) between 1 m and 2 m, and (3) taller than 2 m. Thus, a maximum of 12 trees could potentially be sampled in each plot. The cardinal directions were defined by using a Suunto compass, and both the closest tree and the maximum search limit were determined by using a surveyor's tape measure in cases of doubt.

For each sample tree, several tree metrics were recorded individually. Tree species was determined and tree height was measured using a steel tape measure for smaller trees and a Vertex III hypsometer for tall trees. Stem diameter was callipered at root collar and crown diameters were measured in the cardinal directions with a steel tape measure. Finally, the precise position for each tree was determined using the dGNSS-based procedure described above.

In this study, a total of 524 trees were used, i.e., 404 mountain birch, 67 Norway spruce and 53 Scots pine. Tree heights ranged from 0.04 m to 7.80 m, and crown areas, computed as the ellipse defined by the crown diameters as the major and minor axes, from 0.001 m^2 to 19.54 m^2. A summary of the tree metrics is given in Table 1.

2.3 Laser data

Airborne laser scanner data were acquired on 23 and 24 July 2006 with an Optech ALTM 3100C laser scanning system.

A Piper PA-31 Navajo aircraft carried the laser scanning system at an average flying altitude of 800 m above ground level. The flight speed was approximately 75 ms^{-1}. The scan frequency was 70 Hz, the maximum half angle was 7°, and the average footprint diameter was estimated to 20 cm. Furthermore, the pulse repetition frequency was 100 kHz and resulted in a mean pulse density of 6.8 m^{-2}. The 1,000 km long transect was split into 98 individual flight lines to keep the flying altitude across the mountains and hence the pulse density as constant as possible.

Table 1 - Summary of field measurements of trees.

Tree species	Characteristics	*n*	Mean	Min.	Max.
Mountain birch	Height (m)	404	1.41	0.04	7.80
	Diameter (cm)	404	4.24	0.10	34.00
	Crown area (m²)	404	1.13	0.001	19.54
Norway spruce	Height (m)	67	1.67	0.07	7.00
	Diameter (cm)	65[a]	6.54	0.20	19.10
	Crown area (m²)	67	1.45	0.006	5.69
Scots pine	Height (m)	53	1.33	0.10	5.10
	Diameter (cm)	53	5.00	0.30	18.90
	Crown area (m²)	53	0.81	0.002	7.28

Note: [a] Missing values due to tree properties.

Pre-processing of the laser scanning data was conducted by the contractor (Blom Geomatics, Norway). For all laser points, planimetric coordinates (x and y) and ellipsoidal height values were computed.

For the derivation of the terrain model, laser echoes labelled as “last-of-many” and “single” (LAST) were used. Ground echoes were classified from the planimetric coordinates and the corresponding height values of the LAST echoes, and based on an iteration distance of 1.0 m and an iteration angle of 9°, a triangulated irregular network (TIN) was derived using the TerraScan software (Terrasolid, 2011). Moreover, a digital elevation model (DEM) was computed (QCoherent Software, 2010) using the LAST echoes classified as ground returns to compute the terrain-related variable slope (Burrough and McDonald, 1998). Because of the small-sized objects in question, the DEM was derived with a cell size of 0.25 m.

Laser echoes labelled as “first-of-many” and “single” (FIRST) were used for the analyses. For this purpose, FIRST echoes were projected onto the TIN surface to interpolate the corresponding terrain height values on these locations. Furthermore, the differences between the FIRST echo height values and the corresponding interpolated terrain heights were computed and stored. In this study, merely the FIRST echoes, hereafter referred to as laser echoes, with height values greater zero were included because this criterion represents the sole indicator for the presence of objects on the terrain surface.

The ALTM 3100C instrument may record up to four echoes per laser pulse with a minimum vertical distance of 2.1 m between two subsequent echoes of an individual pulse. However, this instrument property in combination with low vegetation in the present study resulted in very few pulses with more than a single echo. Hence, the LAST and FIRST datasets were almost identical for most of the sample plots.

2.4 Computations

For assessing the capability of discriminators represented by geostatistical and standard statistical measures derived from the laser echoes to improve the classification of tree and non-tree echoes, a sequence of computations had to be conducted prior to analysis.

First, the field-measured crown diameters were used to compute elliptical tree crown polygons to select the tree echoes. Trees with a crown diameter value less than 1.0 m in at least one cardinal direction were assigned a tree crown polygon with a constant radius of 0.5 m because of a positioning error of the laser data of up to 0.5 m as reported by the contractor.

Furthermore, areas within the sample plots where it was ensured that there were no trees because of the basic properties of the PCQ sampling method were identified in order to find and select non-tree laser echoes. These areas were those sectors of the four quadrants that were closer to the plot center than the closest recorded tree irrespective of tree size class. In this process, the crown polygon of the closest tree was erased from the non-tree sector to ensure that only laser echoes emerging from non-tree objects were included.

The laser height and intensity values from the laser echoes were used for the computation of discriminators for the classification analyses. Concerning the laser height, the numerical height values were used directly. For laser intensity, the raw intensity values (I_{Raw}) had to be normalised for the range R according to the following formula suggested by Korpela et al. (2010):

$$I_{Ran} = \left(\frac{R}{R_{Ref}}\right)^{2.4} \cdot I_{Raw} \quad (1)$$

where R_{Ref} is an average reference range that was set to 800 m in this study.

For the computation of the geostatistical and statistical measures, each of the 77 sample plots was overlaid with equally spaced grid points with an interdistance of 1 m. A moving window consisting of a circular buffer with a radius of 3 m was employed to select laser echoes for the estimation of the different geostatistical and statistical measures at each grid point both based on the laser height and intensity values. Thereafter, each laser echo was assigned the computed measures of its closest grid point (Figure 2).

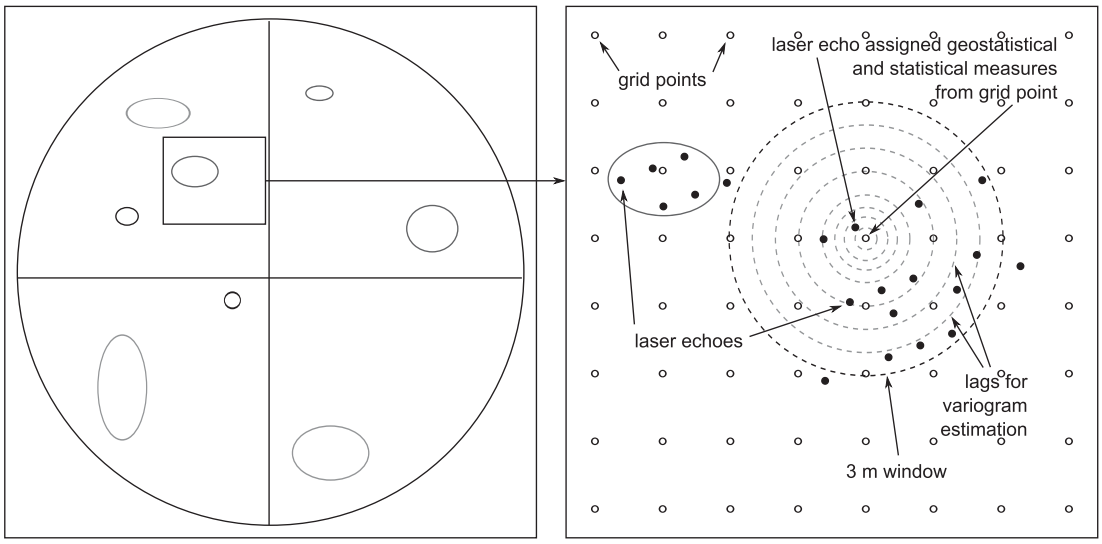

Figure 2 - Illustration of a PCQ sample plot (left) and a detailed demonstration of the computation of the geostatistical and statistical measures (right). Tree locations and the respective crown areas are represent in the three tree height classes: < 1 m (black ellipses), 1–2 m (dark grey ellipses), and >2 m (light grey ellipses). Using a 3 m moving window (black dashed circle), laser echoes (black points) were selected for the computation of the geostatistical and statistical measures for each grid point (white points). The geostatistical measure was estimated using different lags (light grey dashed circles).

Semivariograms were employed as the geostatistical discriminator. Semivariograms were used in the analyses as a mean to characterise differences in the behaviour of spatial correlation of laser height and intensity values for those tree and non-tree echoes with positive height values.

A measure for the spatial correlation of a variable is derived from the calculation of the semivariances of multiple pairs of observations as a function of their separation distance (Isaaks and Srivastava, 1989) and is referred to as an experimental variogram. The separation

distances used for estimation are represented by various distance classes which are referred to as lags. The semivariances of a dataset are computed as

$$\hat{\gamma}(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [z(x_i) - z(x_i + h)]^2 \quad (2)$$

where $\hat{\gamma}(h)$ is the estimated semivariance for lag h and $n(h)$ is the number of data points separated by h (Rossi et al., 1992). Furthermore, the semivariances and hence the spatial variability of a variable can be illustrated by a semivariogram, which is usually referred to as a variogram. In case of spatial dependence, a univariate experimental variogram is characterised by an increase in semivariance with distance h which may level off at the so called sill or increase *ad infinitum*. Furthermore, the so called nugget effect may occur in the univariate experimental variogram. This is characterised by a semivariance value greater/less than zero at the origin and represents spatial variability that is caused by measurement errors or distances shorter than the sample spacing. In this study, the mean value of the semivariances of an experimental variogram was used in the analyses. This mean value was denoted *SV* (Table 2).

For computation of the experimental variograms specifically, variograms were calculated individually for each grid point of the 77 sample plots using the *gstat* spatial package (Pebesma, 2004) in the statistical computing software R (R Development Core Team, 2007). The distance classes used for computation were defined to reflect the fact that lags closer to zero are expected to provide more information than lags further away. These lags were used: 0 m, 0.25 m, 0.5 m, 0.75 m, 1 m, 1.5 m, 2 m, 2.5 m, and 3 m. Furthermore, second-order stationarity was assumed which implies a constant mean, variance and covariances depending on separation only (Webster and Oliver, 2001). Isotropy was assumed for the spatial distributions of the laser height and intensity.

In addition to the geostatistical discriminator, statistical summary measures were employed. The arithmetic mean (*AM*) as the sum of values of a set of observations divided by the number of observations, the standard deviation (*SD*) as the square root of the averaged squares of the observations' deviations from their mean, and the coefficient of variation (*CV*) as the ratio between the arithmetic mean and the standard deviation were derived both from laser height and intensity values respectively (Table 2).

Table 2 - Geostatistical and statistical measures used for classification.

Based on	Discriminator	Abbreviation
Laser *H*eight	Mean Semivariance	H_{SV}
	Arithmetic Mean	H_{AM}
	Standard Deviation	H_{SD}
	Coefficient of Variation	H_{CV}
Laser *I*ntensity	Mean Semivariance	I_{SV}
	Arithmetic Mean	I_{AM}
	Standard Deviation	I_{SD}
	Coefficient of Variation	I_{CV}

3. Methods

3.1 Analyses

Generalised linear models (GLM) and support vector machines (SVM) were employed as classification methods in the analyses. Simple models (Table 3) from a study conducted by Thieme et al. (2012) were extended with the geostatistical and statistical measures to evaluate their potential for an improved classification performance. The two simple models included the laser height and intensity values for the GLM and the additional terrain variable slope for the SVM. A summary of the different discriminating geostatistical and statistical variables is given in Table 4.

Geostatistical and statistical measures that revealed a significant improvement of the model compared to the simple model when used individually were subsequently combined in extended models using all possible combinations (Table 3) to assess a potential contribution of these combinations for the discrimination between tree and non-tree echoes.

Table 3 - Models used classification with GLM and SVM.

Classification	Models[a]
Basic models GLM	HI_H_{SV}, HI_H_{AM}, HI_H_{SD}, HI_H_{CV}, HI_I_{SV}, HI_I_{AM}, HI_I_{SD}, HI_I_{CV}
Additional models GLM	$HI\_H_{SV}\_H_{AM}$
Basic models SVM	HIS_H_{SV}, HIS_H_{AM}, HIS_H_{SD}, HIS_H_{CV}, HIS_I_{SV}, HIS_I_{AM}, HIS_I_{SD}, HIS_I_{CV}
Additional models SVM	$HIS\_H_{SV}\_H_{AM}$, $HIS\_H_{SV}\_H_{SD}$, $HIS\_H_{AM}\_H_{SD}$, $HIS\_H_{SV}\_H_{AM}\_H_{SD}$

[a] *HI* and *HIS* indicate the simple models for GLM and SVM, respectively. Further abbreviations see Table 2.

3.2 GLM

GLM are commonly used in regression analysis, however, GLM also represent a suitable tool for binary classification problems predicting probabilities on a transformed scale (Dalgaard, 2008). GLM are defined by three elements consisting of the random component identifying the response variable Y and its probability distribution, the link function connecting the random component to the systematic component that is again specifying the independent variables X (Agresti, 2007; Dalgaard, 2008). In the present study, a logit link function was employed to relate the different combinations of the independent variables X to the binary response variable Y (tree/non-tree). Thus, the following model was fitted:

$$log\left(\frac{\pi(tree)}{1-\pi(tree)}\right) = \alpha + \beta_1 x_1 + \cdots + \beta_k x_k \quad (3)$$

In the statistical computing software R (R Development Core Team, 2007), the different GLM models (Table 3) were fitted using the *glm* function of the *stats* package (R Development Core Team, 2007). In the next step, the probabilities of the laser echoes for being a non-tree echo were predicted from the fitted models. Finally, different thresholds (from p=0.05 to p=0.95 in 0.05 steps) for these probabilities were employed to classify the laser echoes into tree and non-tree echoes for each model. For each threshold used during classification, the Cohen's kappa coefficient (Cohen, 1960) was estimated to identify the classification with the highest kappa coefficient.

3.3 SVM

SVM, which were developed by Cortes and Vapnik (1995), are a suitable tool for classification, regression, and novelty detection (Karatzoglou et al., 2006; Meyer, 2011). By solving a quadratic optimisation problem using a training set, SVM determine the hyperplane with the maximal margin of separation between two classes. In the process, the relevant information used during classification is comprised by the support vectors representing points located on the margin boundaries. Points located on the opposite side of the margin indicate overlapping classes and are reduced in influence by weighting. The error term is controlled by a so called cost or penalty parameter C and a kernel function allowing for a nonlinear separator defines the hyperplane. In the present study, the C-support vector classification was used with the radial basis function as the kernel, where γ represents a parameter regulating the radial basis function.

Table 4 - Summary of the discriminator variables.

Class	Variable	Mean	Min.	Max.
Tree	Height (m)	1.59	0.04	6.49
	Mean semivariance	0.95	0.00	6.28
	Mean	1.25	0.08	4.24
	Standard deviation	0.91	0.00	2.58
	Coefficient of variation	0.80	0.00	2.24
	Intensity	51.62	4.24	90.95
	Mean semivariance	114.36	0.00	603.08
	Mean	53.80	34.21	76.58
	Standard deviation	10.86	0.00	22.80
	Coefficient of variation	0.21	0.00	0.48
	Slope (°)	16.49	1.05	49.89
Non-tree	Height (m)	0.17	0.01	4.72
	Mean semivariance	0.04	0.00	4.02
	Mean	0.19	0.04	4.17
	Standard deviation	0.12	0.00	2.46
	Coefficient of variation	0.51	0.00	2.64
	Intensity	56.22	0.51	110.82
	Mean semivariance	60.14	0.00	1462.73
	Mean	56.10	10.65	94.01
	Standard deviation	7.56	0.00	38.26
	Coefficient of variation	0.14	0.00	1.04
	Slope (°)	16.54	0.005	79.68

The different models (Table 3) were fitted with the *svm* function of the *e1071* package (Dimitriadou et al., 2011) and a prediction of the laser echoes being a tree or non-tree echo was performed for each. Using the *tune.svm* function of the e1071 package (Karatzoglou et al., 2006; Dimitriadou et al., 2011), the best hyperparameters C and γ were determined prior to classification and outside the leave-one-out cross-validation procedure.

3.4 Accuracy assessment and classification performance

A leave-one-out cross-validation was used to assess the classification performance of the modelling with GLM and SVM. In the validation, each entire field site (i.e. several individual plots) was treated as either being part of the training dataset or the validation dataset. Thus, in each sequence of the cross-validation, models were fit with data from all sites apart from one of the sites, and the fitted models were used for classification on the single site that was excluded from the model fitting.

For each model fitted for prediction irrespective of the classification method, the total percentage of correct prediction and the Cohen's kappa coefficient (Cohen, 1960) were estimated to assess the classification performances. In the comparison between the simple models, i.e., *HI* for the GLM and *HIS* for the SVM (Table 3), and the respective extended models, the difference between two independent kappa coefficients was estimated using a statistics suggested by Cohen (1960) that evaluates the normal curve deviate to assess the significance of such a difference:

$$z = \frac{\kappa_1 - \kappa_2}{\sqrt{\sigma_{\kappa_1}^2 + \sigma_{\kappa_2}^2}} \quad (4)$$

where κ_1 and κ_2 are the two independent kappa coefficients, and σ_{κ_1} and σ_{κ_2} represent the respective standard errors. Kappa coefficients were evaluated quantitatively according to the grading suggested by Landis and Koch (1977).

4. Results

Classifications of the laser echoes into tree and non-tree echoes using GLM and SVM models including geostatistical and statistical measures revealed total accuracies of at least 93.6% (Table 5) and 94.7% (Table 6), respectively.

Furthermore, kappa coefficients were improved by at least 0.032 (Table 5) and 0.034 (Table 6) using GLM and SVM, respectively, compared to the results of the precedent classification study conducted by Thieme et al. (2012).

4.1 GLM

The classifications of the laser echoes using GLM revealed total accuracies between 93.6% and 94.9% (Table 5). The corresponding kappa coefficients ranged from 0.526 to 0.606 indicating moderate fits for all the estimated models (Table 5).

The total accuracies differed with 1.3 percentage points between models (Table 5). Models including geostatistical or statistical measures derived from the laser intensity values (*HI_I*$_{SV}$, *HI_I*$_{AM}$, *HI_I*$_{SD}$, and *HI_I*$_{CV}$) had slightly higher accuracies, of which the models including the standard deviation or the coefficient of variation (*HI_I*$_{SD}$ and *HI_I*$_{CV}$) had the highest accuracies of 94.9%.

Assessing the corresponding kappa coefficients, higher kappa coefficients were found for models including the geostatistical measure and/or the statistical measures represented by the

arithmetic mean and the standard deviation derived from the laser height values (*HI_H$_{SV}$*, *HI_H$_{AM}$*, *HI_H$_{SD}$*, and *HI_H$_{SV}$_H$_{AM}$*). The two models including the arithmetic mean, (*HI_H$_{AM}$*) and the mean semivariance and the arithmetic mean (*HI_H$_{SV}$_H$_{AM}$*), respectively, revealed the highest kappa coefficient of 0.606 (Table 5).

Comparing the kappa coefficients of the nine estimated models to the simple model (*HI*) that revealed the best classification performance using GLM in the study conducted by Thieme et al. (2012), no significant contribution was found for the geostatistical and statistical measures derived from the laser intensity values (Table 5). All kappa coefficients indicated equivalent classification performances for these models, however, neither suggesting significantly worse performances.

Using the geostatistical and statistical measures derived from the laser height values, a significant contribution could be found for the mean semivariance and the arithmetic mean (Table 5). All three models including these two discriminators individually or in combination (*HI_H$_{SV}$*, *HI_H$_{AM}$*, and *HI_H$_{SV}$_H$_{AM}$*) revealed significantly improved classification performances compared to the simple model *HI*. Furthermore, the inclusion of the standard deviation or the coefficient of variation, respectively, showed a similar or significantly worse classification performance than the simple model *HI* (Table 5).

Table 5 - Performance of the different models used for classification with GLM.

Model[a]	*p*	Accuracy	Kappa	*Z*[b]	
HI_H$_{SV}$	0.85	0.947	0.605	2.333	*
HI_H$_{AM}$	0.85	0.946	0.606	2.482	*
HI_H$_{SD}$	0.80	0.943	0.590	1.255	
HI_H$_{CV}$	0.75	0.936	0.526	3.469	**
HI_I$_{SV}$	0.75	0.948	0.570	0.285	
HI_I$_{AM}$	0.70	0.948	0.565	0.626	
HI_I$_{SD}$	0.65	0.949	0.573	0.029	
HI_I$_{CV}$	0.70	0.949	0.565	0.577	
HI_H$_{SV}$_H$_{AM}$	0.85	0.946	0.606	2.480	*
HI	0.75	0.949	0.573		

Note: Level of significance: '<.1. *<.05. **<.005.

[a] *HI* indicates the simple model. Further abbreviations see Table 2.

[b] As received by the comparison between two independent kappa coefficients, i.e. the simple model *HI* and the respective extended model.

4.2 SVM

For the SVM classification method, the twelve different models revealed total accuracies ranging from 94.7% to 95.7% (Table 6). Furthermore, the kappa coefficients ranged between 0.576 and 0.666, indicating moderate fits for four models and substantial fits for eight models, respectively (Table 6).

The twelve models had a maximum difference in total accuracy of 1.0 percentage points (Table 6), where most models consisting of geostatistical or statistical measures derived from the laser height values (*HIS_*H_{SV}, *HIS_*H_{AM}, *HIS_*H_{SD}, *HIS_*H_{SV}*_*H_{AM}, *HIS_*H_{SV}*_*H_{SD}, and *HIS_*H_{AM}*_*H_{SD}) revealed slightly higher accuracies. The highest accuracy of 95.7% was found for models including the mean semivariance and/or the standard deviation (*HIS_*H_{SV}, *HIS_*H_{SD}, and *HIS_*H_{SV}*_*H_{SD}).

Furthermore, the corresponding kappa coefficients were higher for models including the mean semivariance, the arithmetic mean, and the standard deviation derived from the laser height values, both individually and in combination with one another (*HIS_*H_{SV}, *HIS_*H_{AM}, *HIS_*H_{SD}, *HIS_*H_{SV}*_*H_{AM}, *HIS_*H_{SV}*_*H_{SD}, and *HIS_*H_{AM}*_*H_{SD}). The highest kappa coefficient of 0.666 was found for the model only including the mean semivariance, indicating a substantial fit (Table 6).

The comparison between the kappa coefficients the simple model *HIS* revealing the best classification performance in the study carried out by Thieme et al. (2012) and the twelve different models was used to assess the capability of the different geostatistical and statistical measures to improve previous classification.

No significant contribution could be found for any of the models consisting of the geostatistical and statistical measures derived from the laser intensity values (Table 6). The kappa coefficients for the models consisting of the mean semivariance, the standard deviation or the coefficient of variation (*HIS_*I_{SV}, *HIS_*I_{SD}, and *HIS_*I_{CV}) indicated equivalent classification performances for the models. However, the kappa coefficient of the model including the arithmetic mean (*HIS_*I_{AM}) suggested a significantly worse performance compared to the simple model *HIS*.

For the laser height derived geostatistical and statistical measure, a significant contribution was found for six models including the mean semivariance, the arithmetic mean, and the standard deviation individually or in combination with one another (Table 6). All these models (*HIS_*H_{SV}, *HIS_*H_{AM}, *HIS_*H_{SD}, *HIS_*H_{SV}*_*H_{AM}, *HIS_*H_{SV}*_*H_{SD}, and *HIS_*H_{AM}*_*H_{SD}) had kappa coefficients of at least 0.634 improving the simple model *HIS* by at least 0.034 and ameliorating the moderate fit into a substantial fit. Merely the two models

including the coefficient of variation or the combination of the mean semivariance, the arithmetic mean, and the standard deviation revealed no significant contribution to the basic model *HIS*, however, neither indicating a significantly worse classification performance. Furthermore, the mean semivariance represented the discriminator with the highest significant contribution to the basic model *HIS*.

Table 6 - Performance of the different models used for classification with SVM.

Model[a]	C[b]	γ[c]	Accuracy	Kappa	Z[d]	
HIS_H$_{SV}$	100	0.1	0.957	0.666	4.995	**
HIS_H$_{AM}$	1000	0.1	0.956	0.655	4.183	**
HIS_H$_{SD}$	100	0.1	0.957	0.660	4.539	**
HIS_H$_{CV}$	100	0.1	0.951	0.605	0.352	
HIS_I$_{SV}$	1000	0.1	0.953	0.613	0.901	
HIS_I$_{AM}$	1000	0.1	0.947	0.576	1.772	'
HIS_I$_{SD}$	100	0.1	0.953	0.608	0.570	
HIS_I$_{CV}$	1000	0.1	0.950	0.605	0.353	
HIS_H$_{SV}$_H$_{AM}$	100	0.1	0.955	0.643	3.186	**
HIS_H$_{SV}$_H$_{SD}$	100	0.1	0.957	0.664	4.875	**
HIS_H$_{AM}$_H$_{SD}$	100	0.1	0.954	0.634	2.556	*
HIS_H$_{SV}$_H$_{AM}$_H$_{SD}$	1000	0.1	0.952	0.621	1.552	
HIS	1000	0.1	0.953	0.600		

Note: Level of significance: '<.1. *<.05. **<.005.
[a] *HIS* indicates the simple model. Further abbreviations see Table 2.
[b] Cost or penalty parameter
[c] Parameter regulating the radial basis function
[d] As received by the comparison between two independent kappa coefficients, i.e. the simple model *HIS* and the respective extended model.

5. Discussion and conclusion

The classification into tree and non-tree echoes including geostatistical and statistical measures revealed total accuracies that are equivalent to the results obtained by Thieme et al. (2012) for both GLM and SVM. Furthermore, the accuracies of the GLM and SVM classifications are in accordance with other studies on the discrimination of small individual trees in an environment as the forest-tundra ecotone. On an individual tree basis, these studies reported success rates of at least 90% for trees exceeding a height of 1 m (Næsset and Nelson, 2007; Næsset, 2009; Thieme et al., 2011b). These rates are comparable to the results of the present study even though individual laser echoes were used in this case.

Kappa coefficients indicated a significant improvement when including geostatistical and statistical measures for some models in comparison to the classification performances reported by Thieme et al. (2012) both using GLM and SVM. However, geostatistical and statistical measures derived from laser intensity values revealed no significant contribution to any GLM or SVM model and actually a significantly worse performance for the SVM model including the arithmetic mean was obtained. By investigating the respective distributions of values of the different measures for tree and non-tree echoes (Table 4), these results seem reasonable. Particularly the summary values of the arithmetic mean and the coefficient of variation based on laser intensity values do not differ considerably, suggesting a relatively similar behaviour for both tree and non tree-echoes or even indicating an unprofitable effect of this discriminator on the classification performance. Also, for the laser height derived standard deviation and coefficient of variation, similar distributions of values of the different measures were found for tree and non-tree echoes, thus suggesting almost no discriminating effect for the coefficient of variation in particular (Table 4). These findings are reflected in the similar or significantly worse classification performances of both GLM and SVM models including these discriminators. However, regarding the standard deviation in context with SVM, this measure reveals a significant contribution individually or in combination with the mean semivariance or the arithmetic mean indicating a positive effect of a nonlinear classification method on this specific measure. The values distributions for the arithmetic mean and the mean semivariance (Table 4) show obvious differences for tree and non-tree echoes. This behaviour supports the significant improvement of the simple models extended with these discriminators individually or in combination with each other for both classification methods. Furthermore, the superior performance of the geostatistical measure represented by the mean semivariance for both the GLM and SVM classification methods is in line with results obtained by Thieme et al. (2011a). They found experimental variograms helpful to characterise and distinguish between tree and non-tree object in a forest-tundra ecotone environment. Also Jakomulska and Clarke (2001) reported a beneficial contribution of variogram-based measures for the classification of vegetation classes including grassland, rocks and woodland, however, based on optical airborne imagery.

To conclude, the classification of tree and non-tree echoes based on previous models from the study conducted by Thieme et al. (2012) that were extended with geostatistical and statistical measures using both GLM and SVM revealed a significant contribution of the majority of the laser height-derived measures.

Adding geostatistical measure represented by the mean semivariance derived from the laser height values significantly improved the results compared to the basic model of both the GLM and the SVM classification methods, respectively. For this discriminator, total accuracies of at least 94.6% could be obtained irrespective of the classification method or being used individually or in combination with other statistical measures. The mean semivariance estimated from the laser intensity values, however, did not reveal a significant contribution to the classification performances.

With regard to the statistical measures, the arithmetic mean derived from the laser height had a significantly positive effect on the classification performances for both classification methods when being used individually and in most combinations with other measures. The laser intensity-derived arithmetic mean, however, revealed an equivalent performance for GLM and a worse performance using SVM. Concerning the standard deviation, no significant contribution could be found using GLM for neither the laser height nor intensity-derived values. Employing SVM, a significant improvement was merely obtained for the discriminator derived from the laser height. The coefficient of variation revealed no significant contribution to neither of the basic models *HI* and *HIS*. With regard to the laser height-derived coefficient of variation used in GLM, the classification performance was worse than the basic model *HI*.

In general, the highest improvement of a basic model was found for the *HIS* model using SVM extended by the mean semivariance. This result in combination with the supporting outcome of the GLM classification suggests a high potential of the mean semivariance as a geostatistical discriminator for tree and non-tree echoes. However, further investigation into the characteristics the geostatistical measure as well as its capability is needed for being able to fully understand and utilise the power of this discriminator.

Acknowledgments

This research has been funded by the Research Council of Norway (project #184636/S30). We wish to thank Blom Geomatics AS, Norway, for collection and processing of the airborne laser scanner data. Thanks also appertain to Mr. Vegard Lien at the Norwegian University of Life Sciences, who was responsible for the fieldwork. Furthermore, Nadja Stumberg would like to thank Dr. Hans Ole Ørka and Dr. Liviu Ene at the Norwegian University of Life Sciences for valuable remarks during the analysis process.

References

ACIA. 2004. *Impacts of a warming Arctic: Arctic Climate Impact Assessment*. Cambridge University Press, Cambridge, UK. 146 pp.

Agresti, A. 2007. *An introduction to categorical data analysis*. Wiley & Sons, Inc., Hoboken, New Jersey, USA. 400 pp.

Boots, B.N., and Getis, A. 1988. *Point Pattern Analysis*. Scientific Geography Series, Volume 8, Sage Publications, Newbury Park. 93 pp.

Burrough, P.A., and McDonell, R.A. 1998. *Principles of Geographical Information Systems*. Oxford University Press, New York. 356 pp.

Callaghan, T.V., Werkman B.R., and Crawford, R.M.M. 2002. The tundra-taiga interface and its dynamics: Concepts and applications. *Ambio Special Report*, Vol. 12, pp. 6–14.

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Management*, Vol. 20, No. 1, pp. 37–46.

Cortes C., and Vapnik, V. 1995. Support-vector networks. *Machine Learning*, Vol. 20, pp. 273–297.

Cottam, G., and Curtis, J.T. 1956. The use of distance measures in phytosociological sampling. *Ecology*, Vol. 37, No. 3, pp. 451–460.

Dalgaard, P. 2008. *Introductory Statistics with R*. Springer Science+Business Media, LLC., New York, USA. 364 pp.

Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D., and Weingessel, A. 2011. *e1071: Misc Functions of the Department of Statistics (e1071), TU Wien*. Available from http://cran.r-project.org/web/packages/e1071/ [accessed 26 September 2011].

Fry, D.L., and Stephens, S.L. 2010. Stand-level spatial dependence in an old-grwoth Jeffrey pine – mixed conifer forest, Sierra San Pedro Mártir, Mexico. *Canadian Journal of Forest Research*, Vol. 40, pp. 1803–1814.

Harper, K.A., Danby, R.K., De Fields, D.L., Lewis, K.P., Trant, A.J., Starzomski, B.M., Savidge, R., and Hermanutz, L. 2011. Tree spatial pattern within the forest-tundra ecotone: a comparison of sites across Canada. *Canadian Journal of Forest Research*, Vol. 41, pp. 479–489.

Holtmeier, F.-K., and Broll, G. 2005. Sensitivity and response of northern hemisphere altitudinal and polar treelines to environmental changes at landscape and local scales. *Global Ecology and Biogeography*, Vol. 14, pp. 395–410.

Hyyppä, J., Kelle, O., Lehikoinen, M., and Inkinen, M. 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 5, pp. 969–975.

Isaaks, E.H., and Srivastava, R.M. 1989. *An introduction to applied geostatistics*. Oxford University Press, New York. 592 pp.

Jakomulska, A., and Clarke, K.C. 2001. Variogram-derived measured of textural image classification: Application to large-scale vegetation mapping. In *geoENV III – Geostatistics for Environmental Applications*. Edited by Monestiez, P., Allard, D., and Froidevaux, R., Kluwer Academic Publishers, Dordrecht, Netherlands. pp. 345–355.

Karatzoglou, A., Meyer, D., and Hornik, K. 2006. Support Vector Machines in R. *Journal of Statistical Software*, Vol. 15, No. 9, pp. 1–28.

Kirschbaum, M., and Fischlin, A. 1996. Climate change impacts on forests. In *Climate change 1995 - Impacts, adaptations and mitigation of climate change: scientific-technical analysis. Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel of Climate Change*. Edited by Watson, R., Zinyowera, M.C., and Moss, R.H. Cambridge University Press, Cambridge, UK. pp. 99–129.

Korpela, I., Ørka, H.O., Maltamo, M., Tokola, T., and Hyyppä, J. 2010. Tree species classification using airborne LiDAR – effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica*, Vol. 44, No. 2, pp. 319–339.

Landis, J.R., and Koch, G.G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, Vol. 33, No. 1, pp. 159–174.

Meyer, D. 2011. *Support Vector Machines: The Interface to libsvm in package e1071*. Available from http://cran.r-project.org/web/packages/e1071/ [accessed 26 September 2011].

Næsset, E. 2009. Influence of terrain model smoothing and flight and sensor configurations on detection of small pioneer trees in the boreal-alpine transition zone utilizing height metrics derived from airborne scanning lasers. *Remote Sensing of Environment*, Vol. 113, pp. 2210–2223.

Næsset, E., and Nelson, R. 2007. Using airborne laser scanning to monitor tree migration in the boreal-alpine transition zone. *Remote Sensing of Environment*, Vol. 110, pp. 357–369.

Okabe, A., Boots, B., Sugihara, K., and Chiu, S.N. 2000. *Spatial Tesselations: Concepts and applications of Voronoi diagrams*. John Wiley & Sons, Chichester. 696 pp.

Pebesma, E.J. 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, Vol. 30, pp. 683–691.

Persson, Å., Holmgren, J., and Söderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, Vol. 68, No. 9, pp. 925–932.

QCoherent Software 2010. *Getting started with LP360*. QCoherent Software, Colorado Springs, USA. 12pp. Available from www.qcoherent.com [accessed 08 February 2012].

R Development Core Team. 2007. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rees, W.G. 2007. Characterisation of arctic treelines by LiDAR and multispectral imagery. *Polar Record*, Vol. 43, No. 227, pp. 345–352.

Rossi, R.E., Mulla, D.J., Journel, A.G., and Franz, E.H. 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecological Monographs*, Vol. 62, No. 2, pp. 277–314.

Solberg, S., Næsset, E., and Bollandsås, O.M. 2006. Single tree segmentation using airborne laser scanner data in a heterogeneous spruce forest. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 12, pp. 1369−1378.

Stenseth, N.C., Mysterud, A., Ottersen, G., Hurrell, J.W., Chan, K.S., and Lima, M. 2002. Ecological effects of climate fluctuations. *Science*, Vol. 297, pp. 1292–1296.

Terrasolid. 2011. *TerraScan User's Guide*. Terrasolid Ltd., Jyväskylä, Finland. 311 pp. Available from www.terrasolid.fi [accessed 26 September 2011].

Thieme, N., Bollandsås, O.M., Gobakken, T., and Næsset, E. 2011a. Assessing spatial variation for tree and non-tree objects in a forest-tundra ecotone in airborne laser scanning data. In *SilviLaser 2011: 11th International Conference on LiDAR Applications for Assessing Forest Ecosystems*. 16-20 October 2011, Hobart, Australia. pp. 325–332. Available from http://www.iufro.org/science/divisions/division-4/40000/40200/40205/publications/ [accessed 1 February 2012].

Thieme, N., Bollandsås, O.M., Gobakken, T., and Næsset E. 2011b. Detection of small single trees in the forest-tundra ecotone using height values from airborne laser scanning. *Canadian Journal of Remote Sensing*, Vol. 37, No. 3, pp. 264–274.

Thieme, N., Ørka, H.O., Bollandsås O.M., Gobakken, T., and Næsset E. 2012. Classifying tree and non-tree echoes from airborne laser scanning in the forest-tundra ecotone. Manuscript submitted for publication to *Canadian Journal of Remote Sensing*.

Warde, W., and Petranka, J.W. 1981. A correction factor table for missing point-center quarter data. *Ecology*, Vol. 62, No. 2, pp. 491–494.

Webster, R., and Oliver, M.A. 2001. *Geostatistics for environmental scientists*. John Wiley & Sons Ltd., New York, USA. 271 pp.

Wulder, M.A., LeDrew, E.F., Franklin, S.E., and Lavigne, M.B. 1998. Aerial image texture information in the estimation of northern deciduous and mixed wood forest leaf area index (LAI). *Remote Sensing of Environment*, Vol. 64, pp. 64–76.

PAPER IV

Automatic detection of small single trees in the forest-tundra ecotone using airborne laser scanning

Authors

Nadja Thieme*, Ole Martin Bollandsås, Terje Gobakken, Erik Næsset

Address

Norwegian University of Life Sciences

Department of Ecology and Natural Resource Management

P.O. Box 5003, N-1432 Ås, Norway

*** Corresponding author**

Nadja Thieme

E-mail: nadja.thieme@umb.no

Phone: +47 64 96 57 20

Abstract

A large proportion of Norway's land area is occupied by mountains and tree-less areas above the alpine treeline. Thus, the forest-tundra ecotone, i.e., the transition zone between the mountain forest and the alpine zone constitute a significant part of the land surface. The vegetation of this temperatur-sensitive ecosystem is expected to be highly affected by climate change and effective monitoring techniques are required to detect potential changes. For the detection of small pionees trees in such an environment, airborne laser scanning (ALS) has been proposed as a useful tool for inventory tasks employing laser height data. The main objective of the present study was to assess the capability of an unsupervised classification for automated monitoring programs of small individual trees using high-density ALS data. For this purpose, field and ALS data were collected along a 1,500 km long transect stretching from northern Norway (69°3' N) to the southern part of the country (58°3' N). A concept for a raster-based algorithm was developed for automatic detection of small single trees. Different height thresholds for the laser echoes included (0 cm, 10 cm, 20 cm, 30 cm, 40 cm and 50 cm) were tested in various combinations in an unsupervised classification using different raster cell sizes (625 m^2, 156 m^2, 39.1 m^2, 9.77 m^2, 2.44 m^2, 0.61 m^2 and 0.15 m^2) for the classification of raster cells where trees were present, i.e. tree raster cells. Suitable initial values for the exclusion of large treeless areas were determined by employing the different raster cell sizes and an optimal raster cell size was recognised representing a lower limit with a still satisfying rate of successfully detected tree raster cells. Furthermore, a high rate of successful detection involved a high level of commission errors for the lower thresholds. At a threshold of 20 cm, however, the rate of non-tree raster cells classified as tree raster cells decreased significantly accompanied by a still satisfying rate of successfully detected tree raster cells.

1. Introduction

Alpine and arctic treelines are seldom distinctly demarcated, but rather represented by transition zones (Callaghan et al., 2002; Holtmeier and Broll, 2005) between the mountain forest and the alpine and arctic zones. Such transitions are referred to as ecotones (Clements, 1905) and the forest-tundra ecotone can be defined as "the transition between forest and tundra at high elevation or latitude" (Harper et al., 2011). Its location entails high sensitivity to climatic changes, notably to increasing temperatures and changes in precipitation as well as changes in snow coverage that affect the length of the growing season (Callaghan, 2004; ACIA, 2004).

Increased temperature in particular may influence the prevailing tree limit by a densification (Danby and Hik, 2007; Batllori and Gutiérrez, 2008) and increased height growth (Kullman, 2002) of the current sparsely distributed pioneer trees, and by an advance of trees into higher altitudinal and latitudinal areas (ACIA, 2004; Kullmann and Öberg, 2009). Beside the increment in height growth of the existing tree layer (Kullman, 2002), a successful colonisation of previous treeless areas involving a long-term survival of seedlings and saplings into trees is required (Aune et al., 2011). Thereby, factors as the production, dispersal, and germination of seeds (Aune et al., 2011) as well as the interplay of abiotic and biotic drivers are essential (Cairns and Moen, 2004; Holtmeier and Broll, 2005; Sturm et al., 2005; Aune et al., 2011). Furthermore, tree limits are also affected by anthropogenic factors such as herbivore activity by domestic animals and pastoral economy which may inhibit the climatic responses (Callaghan et al., 2002; Holtmeier & Broll, 2005; Post and Pedersen, 2008; Olofsson et al., 2009; Hofgaard et al., 2010; Aune et al., 2011). Changes in the tree layer of the forest-tundra ecotone may also be expected to influence the biodiversity, landscape characteristics, biomass, and carbon pools of vegetation zones adjacent to the forest-tundra ecotone, i.e., mountain forest and tundra. For example, more biomass in the forest-tundra ecotone provides better protection for the mountain forest which may improve growth conditions.

The United Nations Framework Convention on Climate Change and the Kyoto protocol involve reporting on greenhouse gas emission and amongst others land use change in respect of deforestation, afforestation and reforestation (UNFCCC, 2008). Hence, there is an important need for data acquisition in low biomass areas with regard to carbon accounting. However, the National Forest Inventory (NFI) in Norway or other monitoring systems do commonly not prioritise the forest-tundra ecotone and sample plots are established on a

sparser grid than in forested areas. In forest-tundra ecotone areas, the measurement costs are often high relative to the importance of these areas for timber resource assessment which has been the main motivation for the Norwegian NFI. However, as carbon reporting and monitoring of the extent of forests areas under climate change become important, assessments of the current state and monitoring of changes in the forest-tundra ecotone become important (Callaghan et al., 2002). In conjunction with climate change, land use change and carbon accounting, the focus in monitoring the forest-tundra ecotone is on tree establishment, i.e., regeneration, growth, mortality, and colonisation of former treeless areas. This requires efficient monitoring systems that have the capability to both cover vast areas and to detect changes at small scales.

Different remote sensing techniques can provide objective wall-to-wall data for large areas. Air- or spaceborne optical sensors have frequently been used for assessments of land cover. However, the limited spatial resolutions of optical remote sensing instruments are not sufficient to detect small-sized trees or changes in their biophysical properties and spatial distribution. With an assumed height growth of 1 to 10 cm per year for such small trees depending on locality and the prevailing microclimate, a remote sensing technique as airborne laser scanning (ALS) is required that has the capability to observe subtle changes in growth and colonisation patterns. The ability of ALS to detect small single trees in the forest tundra ecotone was verified by Næsset and Nelson (2007), Rees (2007), and Thieme et al. (2011) using different laser point densities. Rees (2007) discriminated individual trees with a minimum tree height of 2 m over vast areas covering hundreds of square kilometres using ALS data with a point density of ~ 0.25 m^{-2}. By employing high-density ALS data with point densities ranging between 6.8 m^{-2} and 8.5 m^{-2}, Næsset and Nelson (2007) and Thieme et al. (2011) successfully detected small trees irrespective of tree height. Both studies used positive laser height values inside field-measured tree crown polygons as criterion for successful tree detection, reporting success rates of over 90% for coniferous and at least 84% for mountain birch trees, provided a tree height exceeding 1 m (Næsset and Nelson, 2007; Thieme et al., 2011). This indicated an adequate reliability of the detection method for trees with heights greater than 1 m. The success rates for trees lower than 1 m when merely utilising positive laser height values as criterion for tree detection were significantly lower because of severe commission errors (Næsset and Nelson, 2007; Næsset, 2009). Næsset and Nelson (2007) observed commission errors up to 490% in their study employing a dataset based on a terrain model that was computed with commonly adopted smoothing criteria. The magnitude of positive laser height values emerging from non-tree objects is not just depending on the

occurrence of for instance rocks, hummocks, and other terrain structures, but also on the properties of the terrain model, the sensor, and the flight settings (Næsset, 2009). Thus, the reliability of tree detection solely using positive laser height values is highly affected by such commission errors, especially with regard to tree heights lower than 1 m. In terms of monitoring, however, high rates of commission errors may not be considered a serious concern because of the multi-temporal context in which only trees will change in size and number over time while terrain and terrain objects will remain stable.

Other approaches using different types of ALS-derived variables have also been used to identify small trees in the forest-tundra ecotone. By employing generalised linear models and support vector machines, individual laser echoes were classified into two classes (tree/non-tree) based on preset decision rules as received by the utilisation of training data to characterise the respective classes. These supervised classification techniques employed various types of discriminators such as laser height and intensity values as well as the terrain variable slope (Thieme et al., 2012a), and geostatistical and statistical measures such as the mean semivariances, the arithmetic means, and the standard deviations derived from laser height and intensity values (Thieme et al., 2012b). Such parametric decision rules in linear and nonlinear modelling techniques are not provided in unsupervised classification methods that generally embody a cluster analysis. Classes are built without the usage of training data and without any previous knowledge of the thematic content, but by an aggregation of elements into clusters where each cluster represents a homogeneous class. An unsupervised classification technique utilising the presence and height information of individual laser echoes on different scales may be useful for automatic detection of trees since ALS datasets involve a huge amount of data depending on the laser point density. A dataset covering vast areas such as the forest-tundra ecotone may consist of millions of laser echoes that are challenging to handle and require effective data processing techniques that are able to handle a huge amount of data without the support of field data for calibration. With regard to the classification of laser echoes into tree and non-tree echoes and thus the potential detection of trees, this method represents a yet unutilised approach with an unknown potential for inventory and monitoring purposes in vegetation zones such as the forest-tundra ecotone.

The main objective of this study was to assess the potential of an unsupervised classification for the automatic detection of small single trees in the forest-tundra ecotone using high-density ALS data. For this purpose, a concept for a raster-based algorithm was developed for the classification into tree and non-tree raster. Different raster cell sizes as well as varying laser height thresholds for the laser echoes included were employed in a modified

region quadtree approach. Finally, the accuracy of the classification as well as its suitability for monitoring purposes in a forest-tundra ecotone environment was assessed by evaluating the rate of detection and commission error.

2. Study area and data

2.1 Study area

The study was conducted along a 1,500 km long and approximately 180 m wide north-south transect that encompasses hundreds of mountain forest and alpine elevation gradients. The transect stretches from Tromsø in the northern part of Norway (69°3' N 17°5' E) to Tvedestrand in the southern part of the country (58°3' N 9°0' E) (Figure 1). It covers sample plots in the transition between the mountain forest and the alpine zone, where the terrain is often characterised by rounded forms, but also occurrences of hummocks, rocks and boulders, and steep slopes. The prevalent tree species were Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and mountain birch (*Betula pubescens* ssp *czerepanovii*).

2.2 Field data

The field work was carried out in summer 2008 to provide *in situ* tree data from 35 different field sites allocated along the transect. At each field site, two to four sample plots with a radius of 25 m were laid out to cover the range of the forest-tundra ecotone. The width of the forest-tundra ecotone varies for different locations and therewith the number of sample plots was adapted visually according to the altitudinal range of the ecotone at the specific site. To avoid overlap, sample plots were established with an interdistance of 50 m within field sites. In total, 111 sample plots were laid out at the 35 sites.

For the precise navigation and positioning with real-time kinematic differential Global Navigation Satellite Systems (dGNSS), two Topcon Legacy E+ 20-channel dual-frequency receivers observing pseudo range and carrier phase of both Global Positioning System and Global Navigation Satellite System satellites were used as base and rover receivers. A base station was established at the closest suitable reference point of the Norwegian Mapping Authority for each field site. The expected accuracy of the reference points was 3 cm, whereas the expected horizontal accuracy of the field recordings relative to the base station was about 2 cm. Thus, the expected accuracy of the sample plots centre points was 3–4 cm.

Figure 1 - Overview of the study area with the 35 specific field sites (black dots). The 1,500 km long transect (black line) stretches from to 69°3' N 17°5' E to 58°3' N 9°0' E.

Individual sample trees were selected for measurement within each plot. Three mutually exclusive tree height classes were used for the tree selection: (1) lower than 1 m, (2) between 1 m and 2 m, and (3) taller than 2 m. A modified version of the point-centred quarter sampling method (PCQ) (Cottam and Curtis, 1956; Warde and Petranka, 1981) was used for selection. In the process, each sample plot was divided into four quadrants defined by the cardinal directions from the sample plot centre using a Suunto compass. In each quadrant the trees that were closest to the sample plot centre in the respective tree height classes were sampled independent of tree species and with a maximum search distance of 25 m. In cases of doubt, the maximum search limit and the closest tree were determined using a surveyor's tape measure.

Several tree parameters were recorded individually for each sample tree. Stem diameter was callipered at root collar and tree height was measured with a steel tape measure or a Vertex III hypsometer for tall trees. Crown diameters were measured in the cardinal directions using a steel tape measure and tree species was determined. For each sample tree, the precise position was captured using dGNSS.

In this study, a total of 744 trees were measured. However, ten trees were regarded as invalid for the analyses because of their tree crown areas being completely overlapped by tree crown areas of taller trees. These ten trees were discarded from the dataset which resulted in a total number of 734, i.e. 614 mountain birch, 67 Norway spruce, and 53 Scots pine, included in this study. Tree heights ranged between 0.02 m and 7.80 m. Tree crown areas were computed as the ellipse defined by the crown diameters as major and minor axes and ranged from 0.001 to 19.54 m^2. A summary of the tree parameters is given in Table 1.

Table 1 - Summary of field measurements of trees.

Tree species	Characteristics	*n*	Mean	Min.	Max.
Mountain birch	Height (m)	614	1.27	0.02	7.80
	Diameter (cm)	613[a]	3.65	0.10	34.00
	Crown area (m²)	614	0.91	0.001	19.54
Norway spruce	Height (m)	67	1.67	0.07	7.00
	Diameter (cm)	65[a]	6.54	0.20	19.10
	Crown area (m²)	67	1.45	0.006	5.69
Scots pine	Height (m)	53	1.33	0.10	5.10
	Diameter (cm)	53	5.00	0.30	18.90
	Crown area (m²)	53	0.81	0.002	7.28

Note: [a] Missing diameter measurements due to field conditions.

2.3 Laser data

Airborne laser scanner data were collected in two separate acquisitions because of the large geographical extent of the study area and difficult weather conditions. The first acquisition was carried out in southern and central Norway on 23 and 24 July 2006 using an Optech ALTM 3100C laser scanning system. The second acquisition in northern Norway was conducted on 1 July 2007 with a Gemini upgraded version of the Optech ALTM 3100C laser scanner system, denoted as ALTM Gemini. An overlap zone in the county of Nordland (65°53' N 13°27' E) was scanned with both systems to provide approximately 80 km ALS data for comparison of the two systems. A Piper PA-31 Navajo aircraft carried both laser scanning systems at an average flying altitude of 800 m a.g.l. with a flight speed of

approximately 75 ms^{-1}. The scan frequency was 70 Hz, maximum half scan angle was 7°, and the average footprint diameter was estimated to 20 cm in both acquisitions. Pulse repetition frequency (PRF) was 100 kHz for the ALTM 3100C laser scanner system resulting in a mean pulse density of 6.8 m^{-2}. To obtain laser point clouds as similar as possible for the two acquisitions, the PRF was set to 125 kHz for the ALTM Gemini system, as suggested by a test flight in May 2007 conducted in another area. This resulted in a mean pulse density of 8.5 m^{-2} for the acquisition in northern Norway in 2007. To keep the flying altitude above the terrain and hence the pulse density as constant as possible, the 1,500 km long transect was split into 147 individual flight lines.

Pre-processing of the laser data was accomplished by a contractor (Blom Geomatics Norway), computing planimetric coordinates (x and y) and ellipsoidal height values for all laser echoes.

Laser echoes labelled "last-of-many" and "single", hereafter denoted as LAST, were used for the derivation of the terrain model. Ground echoes were classified from the planimetric coordinates and the corresponding height values of the LAST echoes to derive a triangulated irregular network (TIN) based on an iteration distance of 1.0 m and an iteration angle of 9° using the TerraScan software (Terrasolid, 2011). Furthermore, laser echoes labelled as "first-of-many" and "single", hereafter denoted as FIRST, were used for the analysis in the current study. For this purpose, the FIRST echoes were projected onto the TIN surface and corresponding terrain height on these locations was interpolated. The height differences between the FIRST echo heights and the corresponding interpolated terrain height values were computed and stored. For the present analysis, only FIRST echoes with height values greater than zero were used because this criterion represents the sole indicator for the presence of objects on the terrain surface.

Both the ALTM 3100C and the ALTM Gemini instrument record up to four echoes per laser pulse with a minimum vertical distance of 2.1 m between two subsequent echoes of an individual laser pulse for the ALTM 3100C. Because of the pulse width influencing the vertical resolution, the minimum vertical distance is assumed to be larger for the ALTM Gemini (cf. Baltsavias, 1999). However, in combination with low vegetation, this instrument property involves that very few pulses have more than a single echo. Therefore, the LAST and FIRST datasets will be identical for many of the sample plots.

3. Methods

3.1 Concept

The background of the present study was the development of an algorithm for automatic detection of small single trees in the forest-tundra ecotone using high-density ALS data based on an unsupervised classification approach classifying tree and non-tree raster cells. In general, unsupervised classification methods involve cluster analysis where elements are aggregated to homogeneous classes without any reference or training data. Based on statistical parameters, each element is assigned to a specific class without any information on the thematic content or affiliation of the respective element.

For this purpose, a concept for a raster-based algorithm was developed employing grids with decreasing raster cell sizes as provided by a region quadtree approach (Figure 2). In general, region quadtrees are used as a systematic procedure to display homogeneous parts of an image (Samet, 1984). Based on a chosen criterion, an "image array is successively subdivided into quadrant, subquadrants, etc. until homogeneous blocks are obtained" (Samet, 1984) and a regular decomposition of the image is obtained. Here, the main idea was to overlay the ALS data with the different grids in order to receive a binary raster where raster cells containing at least one laser echo are assigned the value 1 and empty raster cells the value 0. In an algorithmic context, the classification would start using a grid with a relatively large raster cell size that would be intersected with the ALS data and in case of a positive response, i.e. the presence of at least one laser echo, the raster cell would be quartered and again overlaid with the laser data. This procedure would be repeated until the raster cell is assigned the value 0 or until the minimum cell size is reached. Finally, using the raster with the minimum raster cell size, all cells with value 1 would be aggregated to tree clusters using a commonly adopted clustering analysis. Figure 2 illustrates the entire process.

3.2 Computations

In order to assess the accuracy of an unsupervised classification for automatic detection of small single trees in the forest-tundra ecotone, tree and non-tree polygons had to be computed.

For this purpose, the field-measured crown diameters were used to estimate elliptical tree crown polygons. The contractor reported a positioning error of the laser data of up to 0.5 m. Therefore, trees with a crown diameter value less than 1.0 m in at least one cardinal direction were assigned a tree crown polygon with a constant radius of 0.5 m.

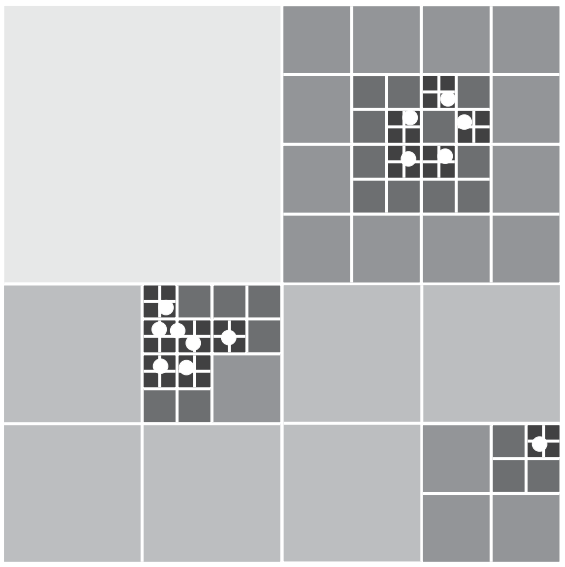

Figure 2 - Concept for a raster-based algorithm using a modified region quadtree. The white dots indicate laser echoes.

Furthermore, non-tree polygons were generated utilising the basic properties of the PCQ method resulting in full control of some of the areas without any trees. For each plot, the sampling design of the PCQ method led to a maximum of three sampled trees per quadrant. The tree closest to the respective plot centre was selected irrespective of tree height class in each of the four quadrants. Using the areas between the selected tree in a quadrant and the respective plot centres, non-tree polygons were computed. In this process, the tree crown polygons of the selected trees were erased from the non-tree polygons to obtain full control over the treeless areas.

3.3 Analysis

To assess the capability of an unsupervised classification to distinguish between tree and non-tree raster cells, parts of the raster-based algorithm concerning the utilisation of grids with different raster cell sizes were tested. For this purpose, different raster cell sizes were adapted to the size of the sample plots starting with the radius of the sample plot as the initial raster cell side length. Raster cell side lengths ranged from 25 m to 39 cm which resulted in raster cell sizes of 625 m^2 and down to 0.153 m^2. Table 2 gives an overview over the different raster cell sizes used in the current analysis. Furthermore, the suitability of different height thresholds for the laser echoes included to detect tree raster cells in an unsupervised

classification approach was assessed: 0 cm, 10 cm, 20 cm, 30 cm, 40 cm, and 50 cm. Albeit different ALS-derived variables have been employed for the supervised classification of tree and non-tree laser echoes, laser height represents the strongest indicator for the presence of trees. Thus, only laser height was used in the current analysis to assess the capability of an unsupervised classification method for the classification of tree and non-tree raster cells.

Table 2 - Summary of raster cell sizes at each location.

Label for cell size	Number of cells	Cell side length (m)	Cell size (m^2)
1	4	25	625.000
2	16	12.5	156.250
3	64	6.25	39.063
4	256	3.125	9.766
5	1024	1.5625	2.441
6	4096	0.78125	0.610
7	16384	0.390625	0.153

For all classifications, grids were computed with the seven different raster cell sizes. Each grid was overlaid with ALS data using the six different laser height thresholds. Raster cells containing at least one laser echo were assigned the value 1, and empty raster cells the value 0. These grids were subsequently intersected with the tree crown polygons in order to assess the classification performance. In the first classification (*I*), all tree crown polygons were included irrespective of their tree height. For the second classification (*II*), only tree crown polygons with a tree height equal to or higher than the laser height thresholds were used. Furthermore, results of studies conducted by Næsset and Nelson (2007) and Thieme et al. (2011) suggested that when trees reach a height taller than 1 m they have a large potential for successful detection by an unsupervised raster-based classification, given a high laser point density. Therefore, a third classification (*III*) was assessed using only tree crown polygons with tree heights exceeding 1 m. Previous studies on the detection of small individual trees reported an underestimation of laser-derived tree heights compared to the corresponding tree heights measured in field (Næsset and Nelson, 2007; Næsset, 2009; Thieme et al., 2011). To investigate a potential effect of the underestimation of the laser-derived tree heights, classifications for the laser height thresholds of 20 cm and 30 cm were evaluated for tree crown polygons with field-measured tree heights larger than the respective thresholds. For the present dataset, Thieme et al. (2011) reported mean underestimations of 20 cm for mountain birch, 29 cm for Norway spruce, and 47 cm for Scots pine, respectively.

Because of the unbalanced tree species composition in the dataset, classifications were evaluated for tree crown polygons with tree heights of 30 cm, 40 cm and 50 cm for the laser height threshold of 20 cm (classification *IV*), and 40 cm and 50 cm for the laser height thresholds 30 cm (classification *V*). Commission errors were investigated by the intersection of the different grids with the non-tree polygons.

4. Results and discussion

4.1 Raster cell sizes

The detection success rates for the different raster cell sizes demonstrated their different usages and suitability in an algorithmic context. Starting with an initial value of 625 m^2, which corresponds to approximately a quarter of the sample plot size in this study, the three largest raster cell sizes *1* with 625 m^2, *2* with 156 m^2, and *3* with 39.1 m^2 (Table 2) were well suited for the exclusion of areas without any trees in all classifications. However, these raster cell sizes were too large for reliable tree raster cell detection since the description of the positioning of the small trees is very imprecise using such large raster cell sizes.

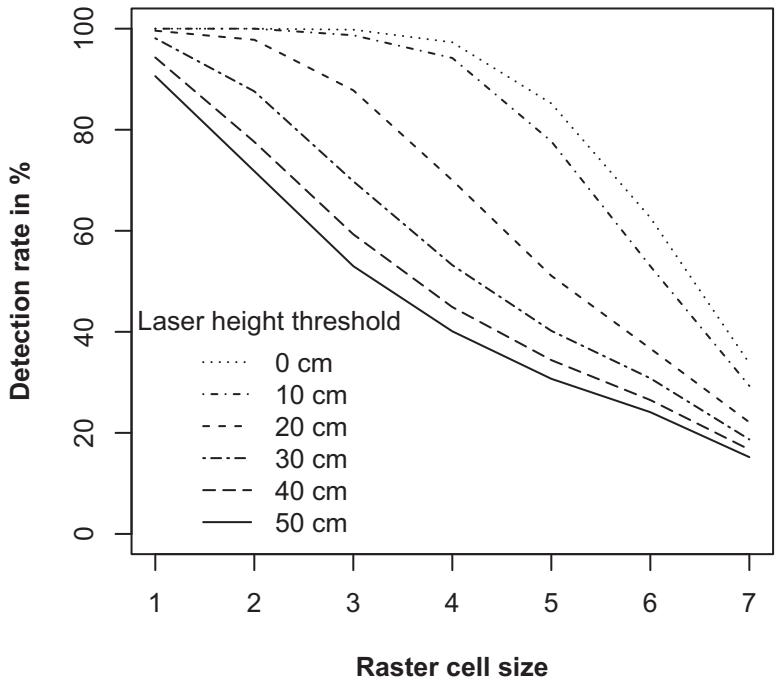

Figure 3 - Detection rate for different raster cell sizes for classifications *I*.

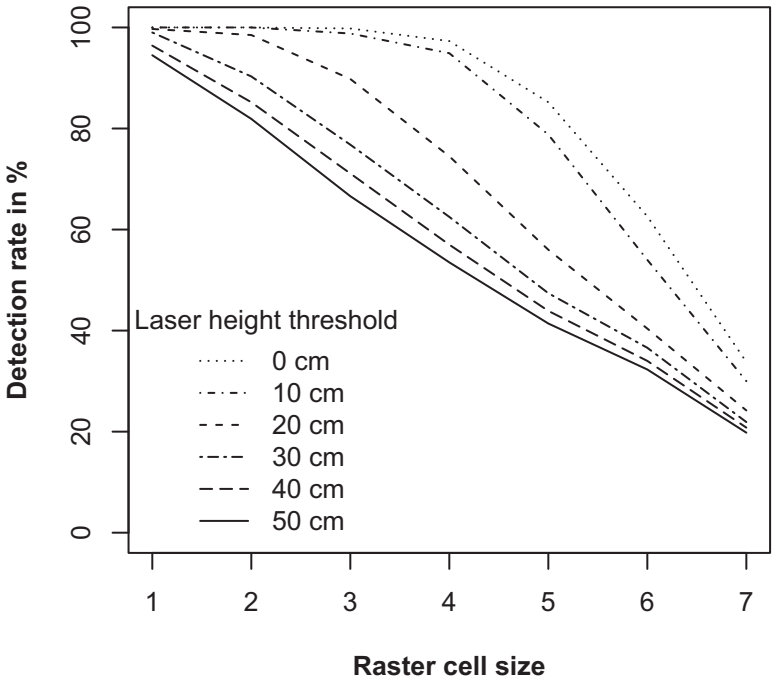

Figure 4 - Detection rate for different raster cell sizes for classifications *II*.

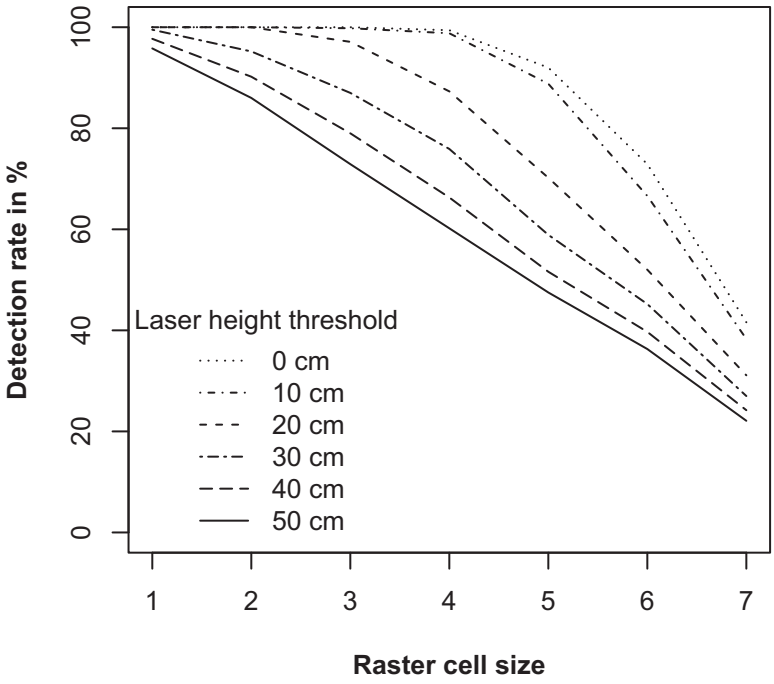

Figure 5 - Detection rate for different raster cell sizes for classifications *III*.

Raster cell size *6* with 0.61 m^2 (Table 2), that is approximately half the size of the mean tree crown area, was identified as the optimal raster cell size in terms of a small raster cell size providing a relatively precise positioning of the small trees and still satisfying detection success rates. In all classifications for raster cell size *6*, 24.1 to 72.9% of the tree raster cells were classified correctly depending on the laser height threshold (Figures 3-7). For raster cell size *7* with 0.15 m^2 (Table 2), however, a significant decrease in the success rate of the tree raster cell detection was found (Figures 3-7). Therewith, raster cell sizes that are considerably smaller than half of the mean tree crown area seem to be too small and hence inapplicable for the detection of tree raster cells.

The three classifications *I*, *II*, and *III* revealed that large raster cell sizes involved high success rates for the detection of tree raster cells and that a decrease in raster cell size was accompanied by a decreasing success rate (Figures 3-5). The success rates were also inversely proportional to the laser height thresholds, i.e. the lower the threshold, the higher the success rate (Figures 3-5). Furthermore, a decrease in raster cell size resulted in a decrease in differences of the success rates between the laser height thresholds. Also for classifications *IV* and *V*, a decrease in raster cell size was accompanied by a decreasing success rate (Figures 6 and 7).

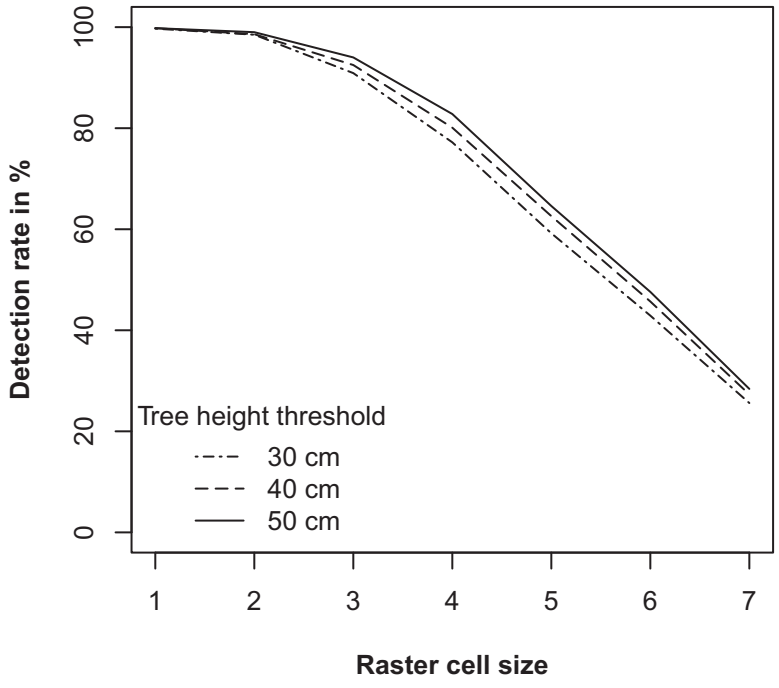

Figure 6 - Detection rate for different raster cell sizes for classifications *IV*.

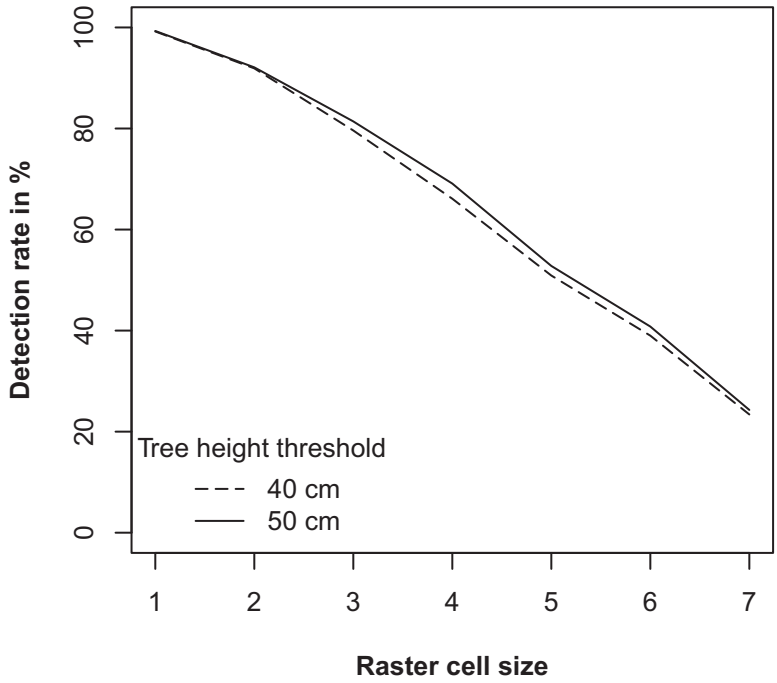

Figure 7 - Detection rate for different raster cell sizes for classifications *V*.

4.2 Laser height thresholds

For non-tree raster cells classified as tree raster cells, rates ranged between 0.01% and 37.3% depending on the laser height threshold (Figure 8). Lower laser height thresholds resulted in higher rates of commission errors, for some raster cell sizes up to 100% (Figure 8), that decreased with increasing laser height threshold. For the classifications *I*, *II*, and *III*, the laser height thresholds of 20 cm and 30 cm represented turning points concerning the rates of commission error which was emphasised by a significant decrease of non-tree raster cells classified as tree raster cells for the respective raster cell sizes *3* to *6* (Figure 8). This distinct decrease in commission errors for the two laser height thresholds suggested a diminution in laser data noise in the range of 20 cm to 30 cm and supported an assumption of such an upper limit as observed during data processing. For higher laser height thresholds, a decrease in the number of non-tree raster cells classified as tree raster cells was almost non-existent for the raster cell sizes mentioned earlier (Figure 8).

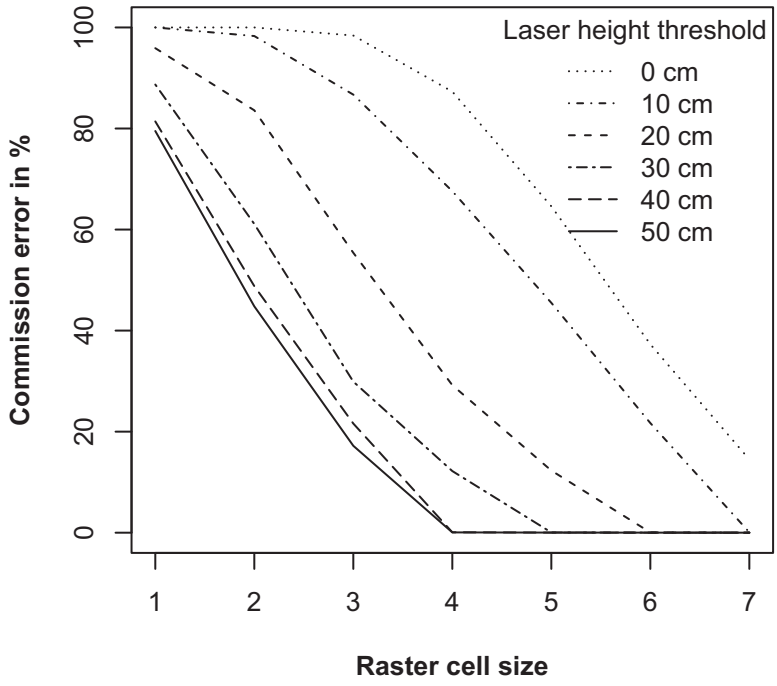

Figure 8 - Commission errors for different laser height thresholds.

4.3 Classification for grid cell size 6

Because of its identification as the optimal raster cell size the following subsection presenting and discussing the results of the different classifications refers to raster cell size *6*.

For classification *I* where the accuracy of the classification was assessed using all field-measured tree data without any restriction of the tree height to the laser height thresholds, at least 24.1% and up to 62.6% of the tree raster cells were found depending on the laser height threshold (Figure 3). By restricting the tree heights to the laser height threshold for the echoes included in the classification (classification *II*), the accuracy of the classification ranged between 32.3% and 62.6% depending on the laser height threshold (Figure 4). For tree heights exceeding a height of 1 m (classification *III*), the rates of successful detection ranged between 36.3% and 72.9% for the different thresholds (Figure 5).

Results from classification *III* where the accuracy was assessed using trees with tree heights larger than 1 m revealed slightly lower detection rates than found by previous studies on small individual tree detection in the forest-tundra ecotone (Næsset and Nelson, 2007; Næsset, 2009; Thieme et al., 2011). However, in this raster-based approach, the detection

rates apply to the tree raster cells and not the individual trees themselves which may have an influence on the detection rates in the present study.

For the classifications with laser height thresholds of 20 cm (*IV*) and 30 cm (*V*), success rates for the detection of tree raster cells ranged between 42.9% and 47.6% (Figure 6), and 39.0% and 40.8% (Figure 7), respectively. Both classifications revealed higher success rates for higher tree height thresholds. However, the differences between the success rates for the different tree height thresholds were low for classification *IV* and almost equal to zero for classification *V* (Figures 6 and 7). These results revealed an increase in successful detection of tree pixels by increasing the height of trees included in the classification. This behaviour reflects the influence of a potential underestimation of real tree height using ALS data as demonstrated by Næsset and Nelson (2007), Næsset (2009), and Thieme et al. (2011).

4.4 Suitability for monitoring purposes

The results from the different classifications revealed that the parameters for raster cell sizes, laser height threshold of the echoes included, as well as a potential lower tree height limit, have to be chosen carefully ensuring a justifiable trade-off between detection success rates and commission errors. For monitoring purposes also additional challenges represented by the usage of different sensors and acquisition settings over time (Næsset, 2009) have to be met.

Varying laser point densities caused by the usage of different instruments have to be tested for their comparability. In general, the probability of a tree for being hit by at least one laser pulse is a function of the laser point density. Low point densities may therewith not be capable to detect trees with sizes that are typical in the forest-tundra ecotone. Næsset and Nelson (2007), as well as Thieme et al. (2011) reported that almost all trees exceeding 1 m in height were hit by at least one laser pulse using high-density ALS data with point densities ranging from 6.8-8.5 m^{-2}. However, in this dataset high point densities involved a relatively large proportion of data noise for laser echoes with heights lower than 20-30 cm as revealed by the sudden decrease of commission errors for these thresholds in the present analysis. In the forest-tundra ecotone, the smallest trees and other typical vegetation such as shrubs are often equal in height which limits their distinguishability considerably. Thus, severe commission errors may occur using an unsupervised classification technique only employing laser height values.

Furthermore, data acquisitions with a sufficient time span are essential to detect regeneration and mortality of small individual trees. Tree height growth is detectable over

relatively short time spans as 2 to 5 years both using high- and low-density ALS data (Næsset and Gobakken, 2005; Yu et al., 2006). However, for small individual trees located in the forest-tundra ecotone, tree height growth is strongly depending on local climatic and topographic conditions. Based on an assumed height growth of 1 to 10 cm per year for small trees, longer time spans may be required for such a raster-based automatic detection algorithm, especially with regard to tree establishment.

5. Conclusion

To conclude, the present study demonstrated the potential of an unsupervised classification approach for the automatic detection of small individual trees in the forest-tundra ecotone based on high-density ALS data. By employing different raster cell sizes, suitable initial values for the exclusion of large areas without any laser echoes reflected from trees could be recognised providing an efficient tool for data processing. Furthermore, a lower limit for raster cell sizes was determined providing a relatively precise positioning of the small trees and still ensuring a satisfying rate of successfully detected tree raster cells.

With regard to the laser height thresholds for the laser echoes included in the respective classifications, the thresholds of 20 cm and 30 cm turned out to be the turning points where the rate of non-tree raster cells classified as tree raster cells decreased significantly accompanied by a still satisfying rate of successfully detected tree raster cells.

ALS has already been adopted for NFI and monitoring programmes, e.g. the Land Use and Carbon Analysis System in New Zealand (Beets et al., 2010). In context of a national monitoring program covering such vast areas as the forest-tundra ecotone in Norway, it is advisable to identify a laser point density that both is high enough to detect the small objects of interest and low enough to keep data noise emerging from other vegetation as low as possible. Then, provided a sufficient time span and an adequate selection of raster cell sizes and laser height threshold of the echoes included, an unsupervised classification technique may be useful to detect regeneration and mortality of small individual trees. The raster grid cells may further build the basis for map products presenting variation of tree presence over time.

Acknowledgments

This research has been funded by the Research Council of Norway (project #184636/S30). We wish to thank Blom Geomatics AS, Norway, for collection and processing

of the airborne laser scanner data. Thanks also appertain to Mr. Vegard Lien at the Norwegian University of Life Sciences, who was responsible for the fieldwork.

References

ACIA. 2004. *Impacts of a warming Arctic: Arctic Climate Impact Assessment*. Cambridge University Press, Cambridge, UK. 146 pp.

Aune, S., Hofgaard, A., and Söderström, L. 2011. Contrasting climate- and land-use-driven tree encroachment patterns of subarctic tundra in northern Norway and the Kola Peninsula. *Canadian Journal of Forest Research*, Vol. 41, pp. 437–449.

Baltsavias, E.P. 1999. Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 54, pp. 199–214.

Battlori E., and Gutiérrez E. 2008. Regional tree line dynamics in response to global change in the Pyrenees. *Journal of Ecology*, Vol. 96, No. 6, pp. 1275–1288.

Beets, P.N., Brandon, A., Fraser, B.V., Goulding, C.J., Lane, P.M., and Stephens, P.R. 2010. National Forest Inventories: New Zealand. In *National Forest Inventories - Pathways for Common Reporting*. Edited by Tomppo, E., Gschwantner, T., Lawrence, M., and McRoberts, R.E. Springer, Dordrecht, Netherlands. pp. 391–410.

Cairns, D.M., and Moen, J. 2004. Herbivory influences tree lines. *Journal of Ecology*, Vol. 92, No. 6, pp. 1019–1024.

Callaghan, T.V., Werkman B.R., and Crawford, R.M.M. 2002. The tundra-taiga interface and its dynamics: Concepts and applications. *Ambio Special Report*, Vol. 12, pp. 6–14.

Clements, F.E. 1905. *Research methods in ecology*. University Publishing Company, Lincoln, Nebraska. 334 pp.

Cottam, G., and Curtis, J.T. 1956. The use of distance measures in phytosociological sampling. *Ecology*, Vol. 37, No. 3, pp. 451–460.

Danby, R.K., and Hik, D.S. 2007. Variability, contingency and rapid change in recent subarctic alpine tree line dynamics. *Journal of Ecology*, Vol. 95, pp. 352– 363.

Harper, K.A., Danby, R.K., De Fields, D.L., Lewis, K.P., Trant, A.J., Starzomski, B.M., Savidge, R., and Hermanutz, L. 2011. Tree spatial pattern within the forest-tundra ecotone: a comparison of sites across Canada. *Canadian Journal of Forest Research*, Vol. 41, pp. 479–489.

Hofgaard, A., Løkken, J.O., Dalen, L., and Hytteborn, H. 2010. Comparing warming and grazing effects on birch growth in an alpine environment – a 10-year experiment. *Plant Ecology & Diversity*, Vol. 3, No. 1, pp. 19–27.

Holtmeier, F.-K., and Broll, G. (2005). Sensitivity and response of northern hemisphere altitudinal and polar treelines to environmental changes at landscape and local scales. *Global Ecology and Biogeography*, *14*, 395–410.

Hyyppä, J., Kelle, O., Lehikoinen, M., and Inkinen, M. 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 5, pp. 969–975.

Kullman, L. 2002. Rapid recent range-margin rise of tree and shrub species in the Swedish Scandes.

Kullman, L., and Öberg, L. 2009. Post-Little Ice Age tree line rise and climate warming in the Swedish Scandes: a landscape ecological perspective. *Journal of Ecology*, Vol. 97, No. 3, pp. 415–429.

Næsset, E. 2009. Influence of terrain model smoothing and flight and sensor configurations on detection of small pioneer trees in the boreal-alpine transition zone utilizing height metrics derived from airborne scanning lasers. *Remote Sensing of Environment*, Vol. 113, pp. 2210–2223.

Næsset, E., and Gobakken, T. 2005. Estimating forest growth using canopy metrics derived from airborne laser scanner data. *Remote Sensing of Environment*, Vol. 96, pp. 453–465.

Næsset, E., and Nelson, R. 2007. Using airborne laser scanning to monitor tree migration in the boreal-alpine transition zone. *Remote Sensing of Environment*, Vol. 110, pp. 357–369.

Olofsson, J., Oksanen, L., Callaghan, T., Hulme, P.E., Oksanen, T., and Suominen, O. 2009. Herbivores inhibit climate-driven shrub expansion on the tundra. *Global Change Biology*, Vol. 15, No. 11, pp. 2681–2693.

Ørka, H.O., Wulder, M.A., Gobakken, T., and Næsset, E. 2012. Subalpine zone delineation using LiDAR and Landsat imagery. *Remote Sensing of Environment*, Vol. 119, pp. 11–20.

Persson, Å., Holmgren, J., and Söderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, Vol. 68, No. 9, pp. 925–932.

Post, E., and Pedersen, C. 2008. Opposing plant community responses to warming with and without herbivores. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 105, No. 34, pp. 12353–12358.

R Development Core Team. 2007. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rees, W.G. 2007. Characterisation of arctic treelines by LiDAR and multispectral imagery. *Polar Record*, Vol. 43, No. 227, pp. 345–352.

Samet, H. 1984. The quadtree and related hierarchical data structures. *Computing Surveys*, Vol. 16, No. 2, pp. 187–260.

Solberg, S., Næsset, E., and Bollandsås, O.M. 2006. Single tree segmentation using airborne laser scanner data in a heterogeneous spruce forest. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 12, pp. 1369–1378.

Sturm, M., Schimel, J., Michaelson, G., Welker, J.M., Oberbauer, S.F., Liston, G.E., Fahnestock, J., and Romanovsky, V.E. 2005. Winter biological processes could help convert arctic tundra to shrubland. *Bioscience*, Vol. 55, No. 1, pp. 17–26.

Terrasolid. 2011. *TerraScan User's Guide*. Terrasolid Ltd., Jyväskylä, Finland. 311 pp. Available from www.terrasolid.fi [accessed 26 September 2011].

Thieme, N., Bollandsås, O.M., Gobakken, T., and Næsset E. 2011. Detection of small single trees in the forest-tundra ecotone using height values from airborne laser scanning. *Canadian Journal of Remote Sensing*, Vol. 37, No. 3, pp. 264–274.

Thieme, N., Ørka, H. O., Bollandsås O. M., Gobakken, T. & Næsset E. (2012a). Classifying tree and non-tree echoes from airborne laser scanning in the forest-tundra ecotone. Manuscript submitted for publication to *Canadian Journal of Remote Sensing*.

Thieme, N., Ørka, H. O., Bollandsås O. M., Gobakken, T. & Næsset E. (2012b). *Improving classification of airborne laser scanning echoes in the forest-tundra ecotone using geostatistical and statistical measures*. Unpublished manuscript.

UNFCCC. 2008. *Kyoto protocol reference manual on accounting of emissions and assigned amount.*

Warde, W., and Petranka, J.W. 1981. A correction factor table for missing point-center quarter data. *Ecology*, Vol. 62, No. 2, pp. 491–494.

Yu, X., Hyyppä, J., Kukko, A., Maltamo, M., and Kaartinen, H. 2006. Change detection techniques for canopy height growth measurements using airborne laser scanner data. *Photogrammetric Engineering and Remote Sensing*, Vol. 72, No. 12, pp. 1339–1348.

ISBN 978-82-575-1081-7
ISSN 1503-1667

Norwegian University of Life Sciences
NO–1432 Ås, Norway
Phone +47 64 96 50 00
www.umb.no, e-mail: postmottak@umb.no