Dissertationes Forestales 365

Assessing the structural biodiversity of forests with airborne laser scanning and optical data

Janne Toivonen

School of Forest Sciences
Faculty of Science, Forestry and Technology
University of Eastern Finland

Academic dissertation

To be presented, with the permission of the Faculty of Science, Forestry and Technology of the University of Eastern Finland, for public criticism in the auditorium M100 of the University of Eastern Finland, Yliopistokatu 7, Joensuu, on 25 of April 2025, at 12 o' clock noon.

Title of dissertation: Assessing the structural biodiversity of forests with airborne laser scanning and optical data
Author: Janne Toivonen

Dissertationes Forestales 365
https://doi.org/10.14214/df.365

Thesis Supervisors:
Professor Petteri Packalen
Natural Resources Institute Finland, Helsinki, Finland

Professor Annika Kangas
Natural Resources Institute Finland, Joensuu, Finland

Professor Matti Maltamo
School of Forest Sciences, University of Eastern Finland, Joensuu, Finland

Dr. Mikko Kukkonen
Natural Resources Institute Finland, Joensuu, Finland

Pre-examiners:
Associate Professor Hans Ole Ørka
Norwegian University of Life Sciences, Ås, Norway

Research Professor Petteri Vihervaara
Finnish Environment Institute, Helsinki, Finland

Opponent:
Senior Lecturer Eva Lindberg
Swedish University of Agricultural Sciences, Umeå, Sweden

ISSN 1795-7389 (online)
ISBN 978-951-651-820-9 (pdf)

Publishers:
Finnish Society of Forest Science
Faculty of Agriculture and Forestry of the University of Helsinki
School of Forest Sciences of the University of Eastern Finland

Editorial Office:
Finnish Society of Forest Science
Viikinkaari 6, FI-00790 Helsinki, Finland
http://www.dissertationesforestales.fi

Toivonen, J. (2025) Assessing the structural biodiversity of forests with airborne laser scanning and optical data. Dissertationes Forestales 365. 54 p. https://doi.org/10.14214/df.365

ABSTRACT

Forests play a significant role in biodiversity-related decision-making as they support approximately 80% of the world's terrestrial biodiversity. Forest vegetation structure and its complexity affect local biodiversity by modifying microclimatic conditions, providing shelter and breeding sites, and affect the distribution and availability of resources and niches. In assessment of forest vegetation structure, remote sensing data, such as airborne laser scanning (ALS) data and optical data, are widely used. The objective of this dissertation was to examine the potential of ALS data in the assessment of biological and structural diversity of forests.

First, the utilisation of ALS data in the assessment of biological and structural diversity of forests was reviewed. The most studied topics and geographical regions of the study areas were reported, and the most useful and common ALS metrics were listed. Second, the performance of ALS and aerial images for the mapping of ecologically valuable large European aspen (*Populus tremula* L.) trees was assessed. The ecological importance of European aspen has been highlighted by the large number of Red-listed species that are dependent on it. Remote sensing-based mapping of aspen is known to be difficult as the species is mixed with other deciduous tree species in a forest stand and its occurrence can be sparse. To account for the rarity of large aspen trees and to balance the training data, the synthetic minority oversampling technique (SMOTE) was tested. Third, the performance of ALS and Sentinel-2 data in the prediction of plot-level forest age was analysed. Remote sensing metrics were combined with field data-based categorical variables that describe site conditions. For the prediction of forest age, linear mixed effects modelling (LME) and tree boosting with random effects (GPBoost) were compared. Some of the field plots contained so-called hold over (seed and retention) trees from the previous generation, which hampered age predictions in these plots. This was addressed by testing an alternative prediction method that included a classification step to identify the hold-over plots.

The results showed that most of the research to date with regard to ALS-based assessment of forest biological and structural diversity has been clustered in Europe and North America. Animal ecology, dead trees and tree species diversity measures have been the most frequently studied topics. The ALS data were usually fused with other remote sensing data, especially aerial or satellite images, which was highly advantageous in studies where tree species were considered. There was no single ALS metric that was suitable for all assessments of forest biological or structural diversity. However, the most often utilised and powerful ALS metrics were standard deviation, the mean and the coefficient of variation (COV) of vegetation heights, which were widely utilised across the studied topics.

The classification of large aspen trees, when SMOTE data augmentation was utilised, improved classification accuracy at both the tree- and plot-levels. For the classification of large aspen trees, aerial image metrics were found to be more important than ALS metrics. In particular, the near-infrared band and its ratios with other spectral bands were important. Results suggest that the detection of large aspen trees in genuine populations is still difficult.

The results presented in this dissertation showed that GPBoost was superior to LME in the prediction of plot-level forest age, and that the addition of categorical variables as random group effects led to a clear decrease in the prediction error. Inclusion of categorical variables improved the root mean square error (RMSE) values for LME more than for the GPBoost model. The best modelling strategy was found to include an initial hold-over plot classification before age prediction.

This dissertation demonstrated that ALS data can provide valuable information for the assessment of forest biodiversity at both fine and broad spatial scales. It also showed that it is important to assess the performance of the method with data that provide a realistic picture of the population. Further research on functional diversity, which has received less attention to date, is needed to cover other aspects of forest diversity. Also, the application of the GPBoost model should be further tested for forest attributes other than forest age.

Keywords: airborne laser scanning, aspen, biodiversity, forest age, forest structure, remote sensing

Toivonen, J. (2025) Assessing the structural biodiversity of forests with airborne laser scanning and optical data. [Metsien rakenteellisen monimuotoisuuden arviointi lentolaserkeilauksen ja optisen datan avulla]. Dissertationes Forestales 365. 54 p. https://doi.org/10.14214/df.365

TIIVISTELMÄ

Metsillä on erityinen rooli luonnon monimuotoisuuteen liittyvässä päätöksenteossa, sillä ne kattavat noin 80 % maaekosysteemien monimuotoisuudesta maailmanlaajuisesti. Metsäkasvillisuuden rakenne ja sen moninaisuus vaikuttavat paikalliseen monimuotoisuuteen muokkaamalla mikroilmastollisia olosuhteita, tarjoamalla suojaa ja lisääntymispaikkoja, sekä vaikuttamalla resurssien ja ekologisten lokeroiden jakautumiseen ja saatavuuteen. Metsien kasvillisuuden rakenteen arvioinnissa kaukokartoitusdataa, kuten lentolaserkeilausdataa (ALS; Airborne Laser Scanning) ja optista dataa (esim. ilma- ja satelliittikuvat), hyödynnetään laajalti. Tämän väitöskirjan tavoitteena oli tarkastella ALS-datan käyttömahdollisuuksia metsien biologisen ja rakenteellisen monimuotoisuuden arvioinnissa.

Ensiksi tehtiin katsaus ALS-datan hyödyntämisestä metsien biologisen ja rakenteellisen monimuotoisuuden arvioinnissa. Katsauksessa raportoitiin yleisimmin tutkitut aiheet ja yleisimmät tutkimusalueiden maantieteelliset sijainnit, sekä listattiin kaikista eniten käytetyt ja hyödyllisimmät ALS-metriikat. Toiseksi arvioitiin ALS-datan ja ilmakuvien yhteiskäyttöä ekologisesti arvokkaiden metsähaapojen tunnistuksessa. Haavan ekologista tärkeyttä alleviivaa se, että lukuisat haavasta riippuvat lajit ovat Punaisen listan lajeja. Kaukokartoitukseen perustuva haavan kartoitus on tunnetusti haastavaa, sillä haavat sekoittuvat eniten muiden lehtipuiden kanssa, mutta myös siksi, että haapoja esiintyy vain harvakseltaan. Haapojen harvalukuisuus otettiin huomioon tasapainottamalla opetusaineistoa niin kutsutulla SMOTE-menetelmällä (Synthetic Minority Oversampling TEchnique). Kolmanneksi arvioitiin ALS-datan ja Sentinel 2-satelliittikuvien yhteiskäyttöä metsikkökoealojen iän ennustamisessa. Kaukokartoitusmuuttujien lisäksi laskettiin maastoaineistosta kategorisia selittäjiä, joilla kuvattiin koealan kasvuolosuhteista. Koealojen iän ennustamisessa verrattiin lineaarista sekamallia (LME) ja tehostettua päätöspuumenetelmää, joka hyödyntää satunnaisvaikutuksia (GPBoost). Joillakin koealoilla oli edellisen puusukupolven niin kutsuttuja ylispuita (siemen- ja jättöpuita), jotka vaikeuttivat iän ennustamista näillä koealoilla. Ylispuut otettiin huomioon testaamalla vaihtoehtoista ennustusmenetelmää, joka sisälsi ylispuukoealojen luokituksen ennen iän ennustamista.

Tulokset osoittivat, että suurin osa ALS-perustaisesta metsien biologisen ja rakenteellisen monimuotisuuden tutkimisesta on tähän saakka tapahtunut Euroopassa ja Pohjois-Amerikassa. Eläinekologia, kuollut puusto ja puulajien monimuotoisuusindeksit olivat eniten tutkittuja aihealueita. ALS-dataa käytettiin usein yhdessä muiden kaukokartoitusaineistojen, kuten ilma- ja satelliittikuvien kanssa, mikä oli erityisen hyödyllistä, kun puulajeja käsiteltiin suorasti tai epäsuorasti. Katsauksen perusteella ei löydetty yhtä selvää ALS-selittäjää, joka olisi hyödyllinen kaikenlaisessa metsien biologisen ja rakenteellisen monimuotoisuuden arvioinnissa. Kasvillisuuden korkeuden keskihajonta, keskiarvo ja variaatiokerroin olivat eniten hyödynnettyjä ja useimmiten hyödyllisiksi osoittautuneita ALS-selittäjiä.

Kookkaiden haapojen puu- ja koealatasojen luokittelutarkkuus parani, kun SMOTE-menetelmää hyödynnettiin harvinaisten haapahavaintojen augmentoinnissa. Ilmakuvaselittäjät osoittautuivat ALS-selittäjiä tärkeämmiksi kookkaiden haapojen tunnistamisessa. Eritoten lähi-infrakanava ja sen suhteet muiden ilmakuvakanavien kanssa olivat tärkeitä selittäjiä. Tulokset osoittavat, että kookkaiden haapojen tunnistaminen aidoissa populaatioissa on edelleen haasteellista.

Koealatason iän ennustamisessa GPBoost-menetelmä oli LME-menetelmää parempi, ja luokka-asteikollisten selittäjien mukaan ottaminen satunnaisvaikutuksina johti selvään ennustevirheen pienentymiseen. Ennustevirheen pieneneminen oli LME-malleissa suurempaa kuin GPBoost-malleissa. Kaikista parhaat tulokset saatiin, kun ylispuukoealojen luokitus tehtiin ennen iän ennustamista.

Tämä väitöskirja osoitti, että ALS-data tarjoaa arvokasta informaatiota metsäluonnon monimuotoisuuden arviointiin niin pienessä kuin suuressakin mittakaavassa. Se myös osoitti, että on tärkeää arvioida menetelmän tehokkuutta aineistolla, joka antaa realistisimman kuvan tarkasteltavasta populaatiosta. Tulevaisuudessa tarvitaan enemmän tutkimusta vähemmän tutkituista aiheista, kuten funktionaalisesta monimuotoisuudesta. Lisäksi GPBoost-menetelmää tulisi testata myös muiden metsää kuvaavien ominaisuuksien kuin iän ennustamisessa.

Avainsanat: laserkeilaus, metsähaapa, biodiversiteetti, metsän ikä, metsän rakenne, kaukokartoitus

ACKNOWLEDGEMENTS

Although there is only one name on the cover of this dissertation, there are multiple people that have assisted me in this endeavour. I would like to express my gratitude to the School of Forest Sciences, University of Eastern Finland (UEF) and its Forest and Bioresources (FORES) Doctoral Programme as the enabling organisation that allowed me to conduct the research in this thesis. Also, great gratitude must be expressed to the Natural Resources Institute Finland, which made my research possible and to the Kone Foundation, which mostly funded this doctoral dissertation. The first thoughts and meetings with my main supervisor, Professor Petteri Packalen, with regard to the possibility of conducting this doctoral dissertation were held in late 2021. This was the time when I started working as research scientist at Natural Resources Institute Finland and started the preparation of the first paper of this dissertation. In early 2022, my research plan was finalised, the line-up of the supervisors was agreed, and I received a position in the FORES Doctoral Programme at UEF. I am most grateful to my supervisors Professor Petteri Packalen, Professor Annika Kangas, Professor Matti Maltamo and Dr. Mikko Kukkonen for the valuable education, guidance and expertise that they offered during this period of writing research papers and for this thesis. Also, I would like to thank my co-authors Dr. Mari Myllymäki and Dr. Timo P. Pitkänen for the ideas and knowledge that they provided to support this work, especially the third paper of the thesis. I am also grateful to Dr. Joel Kostensalo and other colleagues for all their advice and assistance that they sincerely offered to me.

The main research data for the thesis was provided by National Land Survey of Finland, Natural Resources Institute Finland and Finnish Forest Centre. I would like to express my gratitude to these organisations and their professional staff for the pre-processed data that they made available for the use of forest science.

This thesis was pre-examined by Associate Professor Hans Ole Ørka and Research Professor Petteri Vihervaara. I am grateful to my pre-examiners for their valuable feedback and suggestions. I would also like to thank Senior Lecturer Eva Lindberg who will be my opponent in the public examination.

I would like to thank my dad, Risto, and my mom, Riitta, who always encouraged me to study and work hard, but to keep life in balance and do whatever drives you forward. Unfortunately, my mom passed away from cancer in late 2023, and could not see the end of my journey to become a doctor. She would absolutely have loved to see my public defence, although I know that she is watching me every day from the edge of the clouds — I will always remember all the love and care that you gave me. Finally, I would like to express my deepest gratitude to my partner, Riina, who has closely watched my journey since my bachelor days, has always given support whenever needed, and has showed me endless unconditional love.

Joensuu, March 2025

Janne Toivonen

LIST OF ORIGINAL ARTICLES

This thesis is based on data presented in the following articles, referred to by the Roman Numerals **I-III**.

I Toivonen, J., Kangas, A., Maltamo, M., Kukkonen, M., Packalen, P. (2023). Assessing biodiversity using forest structure indicators based on airborne laser scanning data. Forest Ecology and Management, 546, 121376-. https://doi.org/10.1016/j.foreco.2023.121376.

II Toivonen, J., Kangas, A., Maltamo, M., Kukkonen, M., Packalen, P. (2024). Mapping large European aspen (Populus tremula L.) in Finland using airborne lidar and image data. Canadian Journal of Forest Research, 54(7), 762–773. https://doi.org/10.1139/cjfr-2023-0271.

III Toivonen, J., Kangas, A., Pitkänen, T. P., Myllymäki, M., Maltamo, M., Kukkonen, M., Packalen, P. (2025). Tree Boosting with Linear Forest-Type Effects Improves the Performance of Forest Age Predictions Using Airborne Laser Scanning and Satellite Data. Manuscript.

Janne Toivonen was the corresponding author in all three articles. The corresponding author had overall responsibility for the studies, i.e. responsibility for data processing, modelling, analysis, writing and editing. Petteri Packalen, Annika Kangas, Matti Maltamo and Mikko Kukkonen offered the original research ideas and planned the study design with the corresponding author for studies **I**, **II** and **III**. Petteri Packalen assisted the corresponding author in the collection of the research data for studies **II** and **III**. Timo P. Pitkänen contributed with research ideas and a prepared satellite image mosaic in study **III**. Mari Myllymäki assisted in the selection of age model structures in study **III**. All co-authors participated by commenting on and editing the manuscripts.

TABLE OF CONTENTS

INTRODUCTION

Biodiversity and forests

The term ‘biodiversity’ covers all the variation in lifeforms found at different scales of biological organisation on Earth. It is divided into genetic, species and ecosystem diversity components (United Nations, 1992). Biodiversity is highly important for the global system as it enables the various ecosystem services from which society obtains direct or indirect benefits (Duffy 2009). However, human-induced changes on ecosystems have placed much of the global biodiversity in danger (Jetz et al., 2007). This is why conservation and future- and current-state assessment of biodiversity have received both scientific and political interest in recent decades (Millennium Ecosystem Assessment, 2005).

The role of forests in biodiversity-related decision-making is fundamental because they support 80% of global terrestrial biodiversity. Forests provide multiple ecosystem services, such as genetic, species and ecosystem diversity, as well as provisioning services, e.g. timber and recreation (Balvanera et al., 2014). Continental-scale variation of biodiversity is perceived to be determined by climatic conditions, but at smaller scales (e.g. forests), vegetation structure influences diversity and species distribution (Zellweger et al., 2013). Forest structure and its complexity modifies microclimatic conditions, provides breeding sites, affects the distribution and availability of resources and niches, and provides shelter from predators (Brokaw et al., 1999; MacArthur and MacArthur 1961; Melin et al., 2014). A greater availability of niches and resources (i.e. more structural diversity) will likely result in greater diversity of species than a forest with less of these features (Stein et al., 2014). Also, more structurally diverse forests will be able to better cope with changing environmental conditions. In addition to vegetation structure, the variation in terrain elevation influences species diversity at the local-scale (Vogeler et al., 2014; Zhou et al., 2015). The large-scale measurement of habitat structure and the reconstruction of 3D vegetation characteristics using ground sampling requires a lot of time and resources, which is why more efficient methods of data collection are needed.

As direct measurement of biodiversity is not possible, so called ‘biodiversity indicators’ are utilised. These measures transform the complex nature of biodiversity information into a simple and quantifiable form (Biodiversity Indicators Partnership 2011). For example, the change in forest area can be interpreted as an indicator of the progress in forest conservation. In this dissertation, studies **II** and **III** focused on two biodiversity indicators of forests: European aspen (*Populus tremula* L.) and forest age.

European aspen as a biodiversity indicator

European aspen is a keystone species and pioneer in coniferous-dominated boreal forests. Many forest-dwelling species, including birds, fungi, invertebrates and mammals are dependent on both living and dead aspen trees (Kouki et al. 2004; Kivinen et al. 2020). Numerous secondary hole-nesters, such as flying squirrel (*Pteromys volans*) and tits (*Paridae*) also utilise aspen trees (Baroni et al. 2020). The ecological importance of the species is highlighted by the fact that many aspen-dependent species are also Red-listed species (Jonsell et al. 1998; Tikkanen et al. 2006), which means that the species itself is in danger of extinction (International Union for Conservation of Nature 2023). Old-growth, large-diameter (>20 or 25 cm) aspen trees are especially valuable from a biodiversity perspective (Latva-Karjanmaa et al. 2007; Maltamo and Packalen 2014).

As a keystone species, there is a need for information on the abundance and occurrence of these trees at the landscape-level. This information is valuable in the planning and implementation of sustainable forest management and conservation. Also, a time series of aspen occurrence would provide valuable information on landscape health and integrity (Kay 1997). However, the number of aspen trees and other deciduous trees in boreal forests is limited due to traditional forest silvicultural practices that have favoured coniferous trees, and also because natural disturbances rarely occur (Esseen et al. 1997; Kuuluvainen 2002). To date, information on aspen occurrence and abundance in northern Europe is limited.

Forest age as a biodiversity indicator

Forest age is an extremely important attribute for sustainable forest management and planning (Franklin et al., 2018; Rogers et al., 2022), for the assessment of forest biodiversity and habitats (Costanza et al., 1998; Pan et a., 2011), as a predictor variable for various forest attribute models (Eerikäinen et al., 2002; Racine et al., 2014), and in regional or national level reporting (Gillis et al., 2005). Age is usually known in even-aged plantation forests (Packalén et al. 2011), but not in managed boreal forests. In Finland, for example, forest age information is not always available because of the long rotation period of the main tree species and the natural regeneration of minor tree species, such as aspen, during the rotation period (Maltamo et al., 2020).

Identification of old-growth forests is an important aspect related to forest age. In addition, old-growth forests are characterised by the large amount and variation of dead wood, by multilayered canopies, and by a substantial variation in age structure and tree sizes (European Commission 2023; FAO 2022). In this context, terms such as natural forest and virgin forest are also used, sometimes interchangeably (Sabatini et al., 2018). Approximately 34% of global forests are old-growth and are mostly fragmented and small in extent (FAO 2022). Their ecological importance is highlighted by the considerable provisioning potential of ecosystem services and the fact that they provide habitats for multiple endemic and endangered species (Hyvärinen et al., 2019).

Airborne Laser Scanning and other remote sensing technologies

Remote sensing data, such as optical aerial and satellite images and light detection and ranging (lidar) data, are widely utilised in operative forest inventories (Maltamo and Packalen 2014; Næsset 2014). Optical images contain information related to the reflected intensity of electromagnetic radiation. The use of optical remote sensing data for biodiversity assessment has been studied extensively in recent decades and offers relevant data for the classification of vegetation types or for the definition of the horizontal structure of the landscape (Turner et al., 2003). However, many applications require information on the vertical structure, which cannot be directly measured using optical data alone (Vierling et al., 2008). One such technique to overcome such restriction is lidar as it can be used to map both the horizontal and vertical vegetation structures (Bergen et al., 2009).

Lidar is an active remote sensing technology that can be spaceborne, airborne or terrestrial. Airborne laser scanning (ALS or airborne lidar) is often utilised in 3D ecosystem assessments because these surveys aim to provide valuable 3D information about the vegetation over large areas (Bakx et al. 2019). Different statistical metrics can be computed from ALS data that can be used to quantify the structural attributes of forest habitats (Bakx et al., 2019; Davies and Asner, 2014; Hill et al., 2013). In general, they are statistics associated with the height distribution of the ALS echoes, which can be linked to the attributes of interest. These metrics can be roughly categorised into height, density, distribution and intensity metrics.

Forest inventory methods that use ALS data can be divided into two categories: an area-based approach (Næsset 2002) and individual tree detection (Hyyppä et al. 2001). In the former, point cloud metrics are calculated at the plot- or raster cell-level and are used as predictor variables for forest attributes, while in the latter, the derivation of tree- or stand attributes is based on the prediction of tree dimensions from trees that are individually segmented from the ALS data. The required scale of information will determine the method that is employed (Maltamo et al. 2014); for example, it may be more desirable to detect large aspen trees at the tree-level rather than predicting the characteristics of aspen at the plot- and raster cell levels.

Clearly, ALS and other remote sensing technologies have considerable advantages: they can cover large areas, which allows for the generation of wall-to-wall maps, and their use can also reduce logistical and other costs. Therefore, it is evident why ALS technologies have become increasingly popular in the fields of ecology, biodiversity and conservation in recent years. Numerous studies have suggested that remote sensing information can serve as a relevant proxy for biodiversity and ecosystem assessments in vertically complex ecosystems, such as forests (Clawges et al. 2008; Bergen et al. 2009; Vihervaara et al. 2015). However, to-date, reviews of the utilisation of ALS data as the main remote sensing data source in the assessment of forest structural biodiversity and in the prediction of forest biodiversity indices are non-existent.

Earlier studies have highlighted the potential of 3D data for the mapping of aspen (Li et al. 2013; Alonzo et al. 2018). However, the detection of large aspen by remote sensing is reported to be difficult because of their low quantity and scattered occurrence (Maltamo et al. 2015; Viinikka et al. 2020). Moreover, the overlapping spectral response of aspen with other broadleaved trees, such as birch (*Betula* spp.), has been shown to be problematic (see Korpela et al. 2010; Pippuri et al. 2013; Hovi et al. 2017). Also, the overlapping ALS-intensity properties of aspen and spruce have been reported (Ørka et al. 2007).

Remote sensing data have also been reported to be useful in the prediction of forest age. For this purpose, spectral data (Jensen et al., 1999; Gillis et al., 2005; Dye et al., 2012), ALS data (Racine et al., 2014; Maltamo et al., 2020) or a combination of these datasets (Straub and Koch 2011; Schumacher et al., 2020) have been tested. Earlier studies have pointed out that age predictions for forests older than approximately 100 years is a challenge (Maltamo et al., 2020; Wylie et al., 2019). For example, Wylie et al. (2019) reported that the stand age-height relationship in boreal forests in Canada weakened after 120 years.

Objectives

The main objective of this dissertation was to assess the potential of ALS data in determining the biological and structural diversity of forests. The objectives of the three original articles were:

I To provide a review of the use of ALS data in forest biological and structural diversity assessments and describe the ALS metrics that were used in those studies.

II To evaluate the performance of the combined use of ALS data and aerial images for the detection of large European aspen trees.

III To analyse the performance of ALS and Sentinel-2 data for forest age prediction when combined with categorical site condition variables available from each unit of prediction.

MATERIALS

Field data

European aspen

In study **II**, field data consisted of National Forest Inventory (NFI) plots and so-called "treemap" plots acquired from the Mikkeli study site (3,082 km^2). The NFI plot data were provided by Natural Resources Institute Finland (Luke) and the treemap plots were provided by the Finnish Forest Centre. Treemap plots were measured in 2020 and 2021, whereas the NFI plots were measured between 2018 and 2021. To account for the time difference in the NFI data, trees from 2018, 2019 and 2020 were "grown" to 2021 using internal (Luke) growth models (unpublished). The sampling design of the Finnish NFI is explained in detail in Korhonen et al. (2021). How the locations of the trees were determined differed between the NFI and treemap plots: In the former, locations were recorded as bearing and distance from the individual tree to the plot centre, while for the latter, the locations were determined with the TerraHärp system implemented with Masser ExCaliper II callipers (Kostensalo et al. 2023). Treemap plots were subjectively placed within the study area as part of the operational stand-level forest management inventory. Treemap plots were irregular in their shape and size (i.e. neither round nor rectangular), whereas NFI plots were always round with variable radii.

In total, there were 33,195 field-measured trees located across 701 plots, including both NFI and treemap sub-plots. The number of deciduous trees (other than aspen) was 9,079, which corresponded to 27.4% of the trees in the field dataset. There were only 167 aspen trees in the dataset, which corresponded to 0.5% of all field-measured trees. Of those, 35 aspen trees had diameter at breast height (DBH) values ≥ 22 cm, which equated to 1.96 large (DBH ≥ 22 cm) aspen per hectare. The corresponding mean estimate in the NFI for the larger Etelä-Savo region that surrounds the study site was 2.21 large aspen (DBH ≥ 22 cm) per hectare.

Forest age

In study **III**, field data consisted of NFI plots from 12 ALS data acquisition areas in northern Finland (Fig. 1). Plots were measured between 2019 and 2022. For the main tree species in the dominant tree layer, age was always defined by coring or by the calculation of the number of whorl of branches. This age was used as the *observed* plot age in study **III**. Plot ages were updated to match the year of ALS data acquisition by adding or subtracting the difference between the field measurement year and the ALS data acquisition year.

Tree-level age is not always the attribute of interest. In many cases, the mean age of trees within a plot or stand is often of interest, although accurate determination of the plot- or stand-level age can be difficult. For example, ages in mixed forests often vary between tree species, and many old-growth forests have canopy gaps caused by wind or by other disturbances that lead to a substantial variation in tree ages in the gaps and adjoining areas. Also, the systematic sampling design utilised in the NFI means that plots can be located either completely inside a forest stand or in between multiple forest stands of different ages. In this dissertation, only NFI plots that were completely located within one forest stand were selected, as the definition of plot age is more straightforward on a one-stand plot than on a multi-stand plot.

One aspect of study **III** was to identify plots that contained hold-over tree(s) (hereafter, hold-over plots). This issue was addressed because plot age is determined by the age of the dominant stratum, and the dominant stratum on hold-over plots is composed of seedlings. In this situation, employing a common age model would be suboptimal, which is why two separate age models were fitted: one for hold-over plots and one for the other plots. A plot was labelled as a hold-over plot if it contained at least one hold-over tree and its recorded age was ≤ 50 years. Hereafter plots other than hold-over plots are referred to as “other plots”.

In total, there were 870 plots across the ALS data blocks. From these, 50 plots were observed as hold-over plots (5.7% of all plots). The plot age ranged between 0 and 300 years and the most common age class was 70–80 years (Fig. 1). The mean and median plot ages were 87 and 79 years, respectively. The number of plots ≥150 years old was 83 (out of a total of 870).

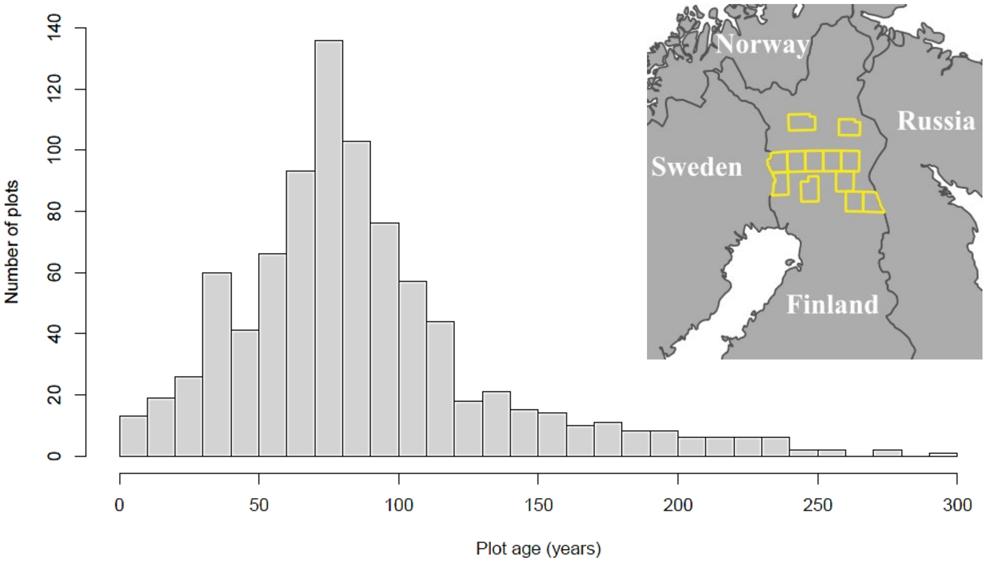

Figure 1. Histogram for the age distribution of plots (left) and map (upper right) showing the locations of airborne laser scanning (ALS) data blocks (n=12) used in study **III**. Each bar in the histogram equates to a 10-year age class.

Remote sensing data

European aspen

Remote sensing data for study **II** consisted of ALS data and aerial images acquired during leaf-on conditions in June 2020 (Tables 1 & 2). The ALS echoes were classified as ground and non-ground, as proposed by Axelsson (2000). The original echo heights in the N2000 vertical coordinate reference system were normalised to above ground level (a.g.l) using a digital terrain model interpolated from ground echoes by Delaunay triangulation. Resulting negative echo heights were set to zero.

Multispectral bands (Level-2) from a Vexcel UltraCam Eagle camera were used, without pan-sharpening or orthorectification. *First of many* and *only* echoes were projected to unrectified colour bands using the internal and external orientation of the camera at the time of exposure (see details in Packalén et al. 2009). External orientations of images were determined using the bundle block adjustment technique with tie points, control points and global navigation satellite system (GNSS) and inertial measurement unit (IMU) values as ancillary observations (Mikhail et al. 2001). The pixel values of the colour bands (red, green, blue and near-infrared (NIR)) were then assigned to the *first of many* and *only* ALS echoes.

Table 1. Metadata of the utilised airborne laser scanning (ALS) data used in studies **II** and **III**.

	Study **II**	Study **III**
ALS device	RIEGL VQ-1560i	RIEGL VQ-780i RIEGL VQ-780 II RIEGL VQ-780 II-S RIEGL VQ-1560 II RIEGL VQ-1560 II-S
Flying altitude	1,525 m a.g.l.	855–2100 m a.g.l.
Scanning frequency	148 Hz	131–230 Hz
Pulse repetition frequency	1400 kHz	700–1 620 kHz
Point density	7.2 points/m^2	5.1–9.5 p/m^2
Maximum scanning angle	20°	20–23°
Side overlap	20.8%	20–24%

Table 2. Metadata of the aerial images used in aspen detection in study **II**.

Multispectral Camera	Vexcel UltraCam Eagle Mk. 1
Image Format, mm	68.016 × 104.052
Image Format, pixels	4360 × 6670
Pixel size (in CCD)	15.600 μm × 15.600 μm
Focal length	100.5 mm
Flying altitude	7,700 m a.g.l.
Ground Sampling Distance of camera	120 cm

[1] Full width at half maximum.

Forest age

Remote sensing data for study **III** consisted of ALS data and Sentinel-2 satellite imagery. The ALS data were acquired during leaf-on conditions in the summers of 2020, 2021 and 2022, depending on the ALS block (12 in total) (Table 1, second column). The ALS data were processed in the PALUS remote processing platform operated by the CSC–IT Center for Science. The Sentinel-2 mosaic was generated in the Google Earth Engine using the percentile-based method described in Pitkänen et al. (2024). The mosaic included all bands where the 20 m pixels were resampled to a 10 m pixel size using the nearest neighbour method. The mosaicking process started with selection of all the summertime (between June 15 and August 15) Level-2A images between 2020–2022 that had a maximum overall cloud coverage of 25%. Then, these initial images were masked to remove the majority of the remaining clouds by using a 30% threshold, based on the S2_CLOUD_PROBABILITY layers (Zupanc, 2017). The final output was calculated using the band-wise 40th reflectance percentile values, derived from all the unmasked pixels. The resulting mosaic was also manually checked and verified to confirm that there were no clouds or shadows.

METHODS

Literature analysis (I)

Literature searches for study **I** were conducted in July 2023 using Scopus, Google Scholar and UEF Primo (search service for electronic materials at the University of Eastern Finland). The used keywords were carefully chosen to minimise the potential of excluding relevant studies. The selected keywords were “ALS”, “airborne lidar” or “airborne laser scanning” for the lidar component, “forest” or “forest structure” for the vegetation component, and either “biodiversity”, “forest type”, “deadwood”, “fragmentation”, “species diversity”, “species richness”, “successional stage”, “understorey” or “animal ecology” as a third component. These keywords were used in all possible combinations for the vegetation, lidar and third components. After the search, the number of articles was reduced by screening the articles – first by title and then by abstract. Lastly, the selected articles were read entirely. To limit the length and complexity of the review, studies where the primary data were acquired with terrestrial, spaceborne or profiling lidar systems, and studies that used photogrammetric point clouds derived from aerial and satellite image data, were excluded. Furthermore, studies where the main interest was the assessment of leaf-area index (LAI), canopy cover, primate habitats or were conducted in aquatic ecosystems, were excluded from the study.

In total, 182 scientific articles were included in study **I** and all articles were published between 2003 and 2022. Most of the research was conducted between 2009 and 2018, with a peak occurring in 2016 when 17 studies were published. The least frequent years for published studies were 2003 and 2006 when only one paper was published per annum.

The most common domain for publications were remote sensing journals, which accounted for 45% of the studies. Approximately 75% of studies only utilised ALS data, and approximately one quarter of studies used optical images and ALS data. Both ALS and terrestrial laser scanning data were utilised in four studies. There was a clear trend in the type of ALS data used: discrete return (DR) ALS were used in 93% of studies, whereas FW-ALS was only used in 7% of the studies. Clear trends with regard to the use of FW-ALS data were not found. Each study was classified by its spatial extent, defined here as the spatial unit on which the study was conducted. Approximately 45% and 35% of the studies were conducted at landscape- and plot-levels, respectively.

Most of the study areas were located in Europe and North America (Fig. 2). Note that there are more study areas than study articles, since some studies utilised data from multiple locations. Approximately 42% of the study areas were located in temperate broadleaved and mixed forests. Boreal and temperate coniferous forests were hosts for approximately 20% of the study areas. The remainder were mainly located in Mediterranean and tropical forests – both accounting for < 10% of the study areas.

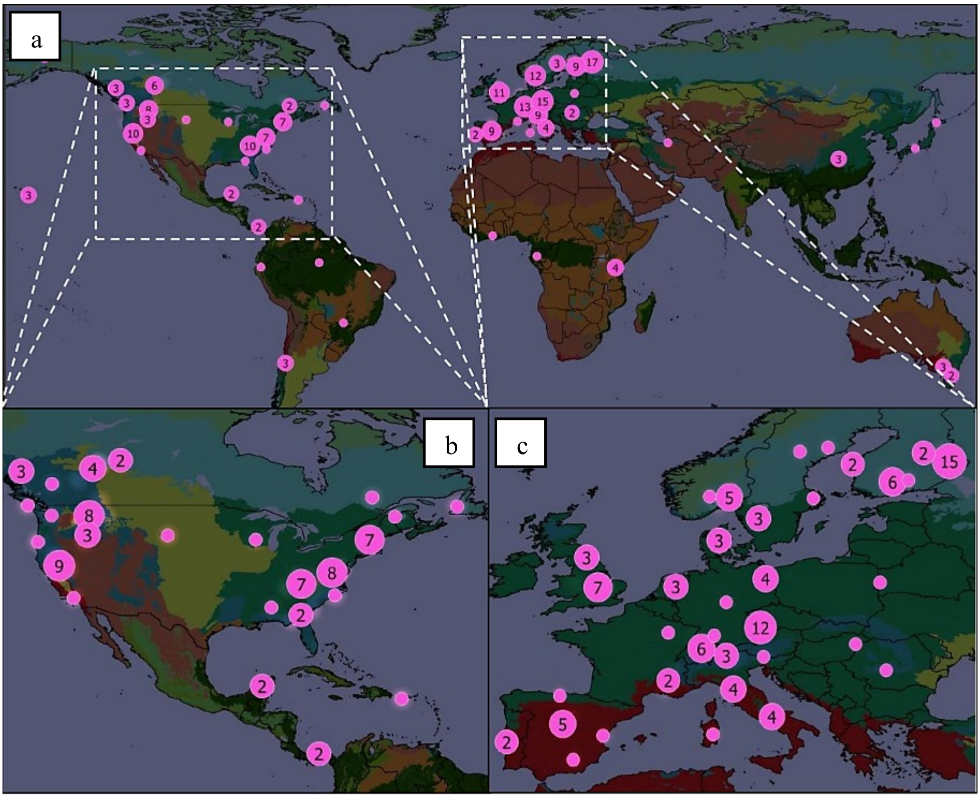

Figure 2. Locations of study areas (a) globally, (b) in North and Central America, and (c) in Europe. Number of clustered studies is shown inside the purple circles. Colouring of the global map (a) is based on the global terrestrial biome classification as presented by Olson et al. (2001). Map: WWF (2021)

Mapping large European aspen in Finland (II)

For the identification of large aspen, single trees were identified from a canopy height model (CHM) interpolated from the ground normalised ALS echoes. Treetops were identified from the low-pass filtered CHM with the assumption that the local maximum in the 8-neighbourhood corresponded to a tree. Trees were delineated from the CHM using marker-controlled watershed segmentation with identified treetops as seeds. Field measured trees were linked to remotely detected trees in order to attach the information related to a field tree (here DBH and tree species) to the detected tree (Fig. 3). As a linking criterion, a 2.5 m limit in the X and Y directions, and a 3 m limit in the Z direction were used. Linking resulted in 12,189 linked trees.

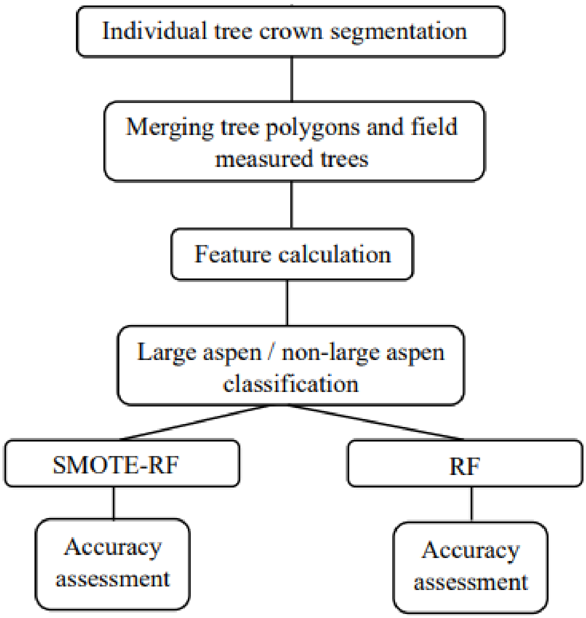

Figure 3. Workflow for the classification of large aspen trees.

The ALS and image metrics were calculated for the CHM segmented tree polygons (Fig. 3). The ALS metrics were calculated for the *first*, *intermediate*, *last* and *all* echo categories and commonly included height and intensity percentiles (5th, 10th, …100th). In addition, the proportion of the different echo categories, as well as mean, standard deviation, skewness and kurtosis values for height and intensity were calculated. Image metrics of mean and standard deviation values for each band, ratios between the bands (e.g. Red/Blue) and Normalised Difference Vegetation Index (NDVI), were computed for the trees using the pixel values linked to the echoes. Only echoes with a height $> 0.5 \times$ the predicted height of the tree (e.g. the identified treetop) were selected. Tree height was predicted for detected trees using a linear model fitted with all the detected and correctly linked trees. In the model, the maximum pixel value within the tree polygon (*hPoly*) explained the *measured*/*predicted* height of the tree (*hTree*).

The Random Forest (RF) classifier was utilised to separate large aspen from other remotely detected trees (Fig. 3). This is a well-established, supervised learning algorithm in data science and has been widely used in remote sensing applications (Belgiu and Drăguț 2016). More specifically, the “fast implementation of random forest algorithm” from package ranger (Wright and Ziegler 2017) in the R environment (R Core Team 2022) was selected.

There was a substantial imbalance in the field data as the proportion of large aspen was very small compared to other trees (e.g. 35 aspen with DBH > 22 cm from a total of 33,195 trees). In addition, the number of aspen trees was small compared to other deciduous trees; approximately only 1.8% of all deciduous trees were aspen. To account for the imbalance in model construction, the Synthetic Minority Oversampling Technique (SMOTE) (Chawla et al., 2011) was adopted. In SMOTE, new instances of observations are generated by combining features of the target observation and its k-nearest neighbours in the feature space. In the analysis, values for user-defined parameters were calculated in such a way that the numbers of aspen and other tree species in the training data were approximately equal. SMOTE was implemented in the R-package *performanceEstimation* (Torgo 2016).

Initially, there were 179 ALS and 15 image metrics. Prior to fitting the RF model, ALS metrics were removed until the absolute correlations between the remaining metrics were < 0.8. All image metrics were used as there were significantly more ALS metrics. The response class "large aspen" was assigned to those observations that were aspen and where DBH values were ≥ a range of DBH limits (18, 20, 22, 24 and 26 cm). The definition of large aspen here may be ambiguous, as earlier studies have used DBH values between 20 and 25 cm as a criterion (Latva-Karjanmaa et al., 2007; Maltamo et al., 2015).

Classification was divided into two separate tests. In the first classification (SC1 II), balanced (SMOTE-RF; SRF) and imbalanced RF classifications were compared using the fixed 22 cm DBH limit for large aspen. The second classification (SC2 II) compared SMOTE-RF classifications across a range of DBH limits for large aspen (18, 20, 22, 24 and 26 cm). In SC1 II, the data were sub-divided into A, B and C datasets for both balanced and imbalanced classifications. Dataset A only included linked trees (i.e. trees for which tree polygons were assigned) from the aspen plots. Here, "aspen plot" indicates that the plot contained at least one large aspen tree. In dataset B, unlinked field measured trees from the aspen plots were added to the dataset, and in dataset C, all trees from all plots were considered in the classification. Dataset C was considered as representative of the population, while A and B were unrepresentative.

The leave-one-plot-out cross-validation (CV) technique was utilised, which means that for each iteration, trees from one plot were used as the test data and trees from all the other plots were used as the training data. The following accuracy statistics were reported: precision, recall and the F1-score (Sasaki 2007). In SC1 II, tree-level accuracies for aspen plots (A, B) and all plots (C) were reported. In SC2 II, only accuracies for all plots are shown, but the results are reported at both the tree- and plot-levels.

Finally, the most important metrics used in the SMOTE-RF classification at the 22 cm DBH limit when data from aspen plots and all plots were utilised, are reported. The importance measure employed here was the Gini importance. It is calculated as the number of times that a metric was used to split a tree node divided by the number of all trees. Initially, the importance values across RF runs with different datasets are at different scales, so to make them comparable (between the aspen plots and all plots), importance was normalised by dividing it by its overall mean value.

Forest age prediction using tree boosting and remote sensing data (III)

In study **III**, ALS, Sentinel-2 and categorical metrics were calculated at the plot-level. The ALS metrics were calculated separately for the *first* and *last* echo categories using the *lascanopy* program (LAStools, 2024). Height metrics included commonly used height percentiles (20th, 40th, …), height bincentiles (a.k.a. deciles), mean, standard deviation, skewness and kurtosis values for height, and a vertical complexity index. The ALS metrics also included two terrain height metrics. These were the terrain height at the plot location above sea level and the relative terrain height calculated as the terrain height at the plot location (field measured in the NFI) divided by the mean terrain height in the 3 km radius around the plot. The Sentinel-2 metrics contained the mean values for each band, the ratios between the bands (e.g. B2/B11) and NDVI. Plot-level Categorical predictors (ALS block, WoodProdRestr and MainType) were derived from the NFI data. As a 12-level factor, the ALS block addresses the effect of an inventory area, while WoodProdRestr describes wood production restrictions of a forest and has three levels: “no restrictions”, “protected forests” and “protected mires”. The MainType variable describes the main type of the forest plot and has three levels: “mineral soil”, “drained peatland” and “undrained peatland”. The inclusion of categorical variables, which are indicative of site conditions, was based on the fact that these variables had previously been reported to be important in the modelling of forest age (Maltamo et al., 2020).

Age predictions between tree boosting with random effects (GPBoost) and the linear mixed effects model (LME) for three scenarios (SC1 III, SC2 III and SC3 III) were compared (Fig. 4). In the first scenario (SC1 III), plot age was predicted using only remote sensing metrics. In the second scenario (SC2 III), three categorical variables were added to the configuration of SC1 III. In the third scenario (SC3 III), a hold-over plot classification was added to the configuration of SC2 III. Age predictions for hold-over and other plots differed by the structure of the random component of the age models. Categorical variables were treated as random effects for other plots, whereas in the models for the hold-over plot (SC3 III only), they were not included in the models at all. This was done because hold-over plots only contained a few observations in most categories. All the analyses in study **III** were implemented in the R environment (R Core Team 2024).

Predictor variable computation resulted in 42 ALS and 56 Sentinel-2 metrics. Prior to the fitting of the age models, predictor variables were removed until the correlations between the remaining metrics were < 0.9. From the remaining metrics, the 10 metrics that had the strongest correlation with plot age were selected. This was done within each of the 10-fold CV schemes described below. The resulting 10 metrics in each fold were included in the age models in each scenario.

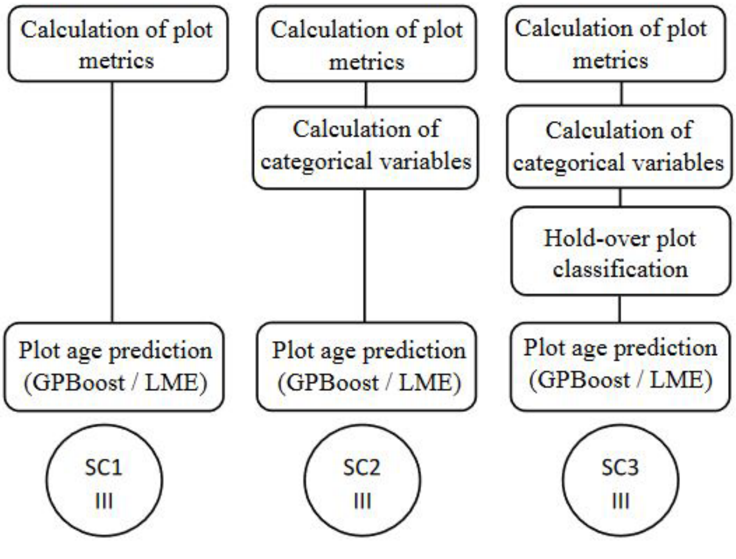

Figure 4. Workflow of age predictions on National Forest Inventory (NFI) plots for the three scenarios SC1 III, SC2 III and SC3 III.

GPBoost was the first tested prediction method. It is a relatively new software library used to combine tree boosting with the Gaussian process and grouped random effects models (Sigrist et al., 2021; Sigrist 2022). Tree boosting is a machine learning technique that trains models in a sequential fashion. It is based on the premise that the best possible next model can be acquired when correcting the previous model. In study **III**, only tree boosting with random effects was used, but the Gaussian process part of the library was not included.

In this study, LME was the second method to be tested for age prediction. It is an extension of common simple linear models as it has both fixed and random effects, and LME models are highly usable when dealing with hierarchical data. The LME method is simpler than GPBoost and can be considered as a sort of a baseline here. First, the ‘maximum model’ was formulated with *lmer* function (Bates et al., 2015), which contains all the predictors that were used in the corresponding GPBoost model. In SC1 III, only fixed effects (i.e. ALS and Sentinel-2 predictor variables) were used in the model. In SC2 III and SC3 III, categorical variables were included in the mixed effects model by adding random intercepts for the categorical variables. After model formulation, *buildmer* function (Voeten 2021) finds the maximal feasible model by starting with an “empty model” and adding terms to this model until the model can no longer converge. After the maximal feasible model was found, stepwise elimination was applied. In SC2 III and SC3 III, ALS block, MainType and WoodProdRestr were included in the random part of the models. This means that the “empty” *lmer* model was not entirely empty but contained the pre-defined variables in the random component of the model. Eventually, the best model of *buildmer* was fitted with *lmer*.

Hold-over plot classification of SC3 III was implemented with logistic regression using the *glm* function from the stats package. The stepwise model selection was carried out using the *stepAIC* function from the MASS package. The candidate set of predictor variables was the same as for age prediction, although categorical variables were included as the fixed effects in the model formulation. All computations were carried out using 10-fold CV, which means that the data were first split into 10 subsamples, i.e. folds. Then, each plot fold (~ 10%) was, in turn, used as the test data, while the remainder of the plots (~ 90%) were used as training data. For each fold, plots were sampled from each of the 12 inventory areas without replacement. The same folds were used for both prediction methods (GPBoost and LME) and for all scenarios.

The data were split 50 times to minimise the effect of randomness due to random selection of observations into the 10-fold CV. All the results from the three scenarios are reported based on the 50 repetitions. This means that age was predicted 50 times for each plot and the mean of these 50 values was used as the predicted value. The relative root mean square error (RMSE) (Eq. 1) and relative bias (Eq. 2) values are reported as measures of prediction accuracy:

$$\%\mathrm{RMSE} = \frac{100}{\bar{y}} \times \sqrt{\sum_{i=1}^{n} \frac{(\hat{y_i} - y_i)^2}{n}} \tag{1}$$

$$\%\mathrm{BIAS} = 100 \times \frac{\sum_{i=1}^{n} (\hat{y_i} - y_i)}{\sum_{i=1}^{n} y_i} \tag{2}$$

where $\hat{y_i}$ is the predicted age on plot i, y_i is the observed age on field plot i, $\bar{y}$ is the mean of observed ages on plot i, and n is the number of plots.

Finally, the most important predictors of the GPBoost models of SC3 and the hold-over classification were analysed. For GPBoost, the importance measure *Gain* (Sigrist et al., 2021) was utilised as the mean values over the 50 repetitions. The importance metric for the hold-over plot classification was represented as the number of times the predictor variable was chosen in the logistic regression models divided by 500 (the value of 500 is the total number of logistic regression models in each 10-fold over 50 repetitions).

RESULTS

ALS-based assessment of forest biodiversity (I)

The most studied topics in the biodiversity-related assessment of forests with ALS data were highlighted in study **I**. In addition, the most common ALS metrics and the best performing combinations with other remote sensing data were reviewed. The division of sections below represent the most important findings from the review conducted in study **I**.

Forest habitat type and successional stage

The distribution of habitat types and forest successional stages across large areas is reflective of potential species richness and helps to understand broad-scale biodiversity, but mapping to date has proved a challenge.

Classification of land cover and the assessment of successional stages have been mainly studied in temperate forests, often located in national parks or protected sites. Classification schemes for land cover classification have varied greatly and were most often named after the main tree species or the species group that was most representative of the landscape (Dalponte et al., 2008; Simonson et al., 2013). In successional stage studies, the stage to be classified was determined based either on an existing classification scheme utilised in forest management (Alberti et al. 2013) or was tailored by the authors of the study (Martin and Valeria 2022). In both topics, classification was most often implemented using supervised machine learning algorithms, such as RF and Support Vector Machine (SVM). The overall accuracy was reported to depend on the number of habitat types in question. For example, Sverdrup-Thygeson et al. (2016) obtained an overall accuracy of approximately 94% in a 2-class scenario, whereas Álvarez-Martínez et al. (2018) mapped 11 forest-related habitat types with an overall accuracy of approximately 66%. For the successional stages, the classification accuracy was greater in older successional stages (Alberti et al., 2013; Torresan et al., 2016).

The most important metrics for both types of studies were those that described the vertical canopy structure. For example, average/mean values (Álvarez-Martínez et al., 2018; Bottalico et al., 2014), standard deviation (Fuhr et al., 2022), median absolute deviation (Martinuzzi et al., 2013) and average absolute deviation of vegetation height (Valbuena et al., 2016) were found to perform well with ALS metrics. The addition of spectral variables (Dalponte et al., 2008) and/or ALS intensity features (Pippuri et al., 2016) were found to be favourable when broadleaved trees were used in habitat type classification. In successional stage studies, classification performance was reported to be less successful with Landsat metrics than ALS metrics (Martinuzzi et al., 2013).

In some habitat type studies, poor class-level accuracies were attributed to the small number of training samples for a specific class (Álvarez-Martínez et al., 2018; Pippuri et al., 2016) and to incorrect categorisation of observations in the training data (Vehmas et al., 2009). In successional stage studies, errors were most notable in forests that exhibited significant structural variability: forests with multi-layered stages (Alberti et al., 2013) and stages with a distinct understorey (Falkowski et al., 2009).

Canopy assessment

The studies in this section either attempted to assign the forest into classes that described height, openness or layering of the canopy, and or were studies that described the structure of the forest understorey (i.e. shrubs and suppressed trees under a dominant canopy). Studies were most often conducted in temperate broadleaved and coniferous forests. For dominant canopy, the most common attribute of interest was forest canopy layering, whether single- or multi-storey (Leiterer et al., 2015; Wilkes et al., 2016). For the understorey assessment, the aims of the studies were generally twofold: some studies predicted the existence of an understorey or its trees (Hill and Broughton 2009; Miura and Jones 2010) while other studies predicted the forest attributes that describe the understorey, such as volume or height (Dees et al., 2012; Lindberg et al., 2012). Unsupervised classification, such as k-means (Guo et al., 2017; Pascual et al., 2008), was used in most of the studies.

Studies on canopy structure reported that the classification accuracy (overall 66.9–97%) was dependent on the number of classes that were classified, i.e. a smaller number of classes resulted in greater overall accuracy (Leiterer et al., 2015), which is to be expected. Some studies reported that the volume of the understorey layer was predicted more accurately (R^2: 0.88–0.95) than height features of the understorey (R^2: 0.76–0.96). The most powerful ALS metrics represented the vertical distribution of the canopy material. Such metrics included, for example, the median and standard deviation (Guo et al., 2017) and COV (Zimble et al., 2003) associated with vegetation height. The inclusion of aerial imagery metrics was favourable when coniferous and deciduous canopies were distinguished (Jayathunga et al., 2018). Full-waveform ALS data were found to be more powerful in the estimation of vegetation volume for lower canopy layers than discrete return ALS data (Lindberg et al. 2012).

Misclassification of canopy structure was mostly exhibited in classes with an open canopy structure and multi-layered canopy (Adnan et al., 2019; Guo et al., 2017). It was observed that broadleaved canopies tended to be classified less accurately than evergreen canopy types (Morsdorf et al., 2010). In some cases, the heights of the understorey canopy layers were overestimated when the understorey layer intersected the dominant canopy layer (Sumnall et al., 2017).

Measures of tree species diversity

Assessment of tree species diversity measures is of major global interest as they contribute to ecosystem (here, forest) health. This section also covers studies that assessed understorey plant species and their richness and/or diversity indices, ecologically valuable trees, and the structural diversity of forests.

Studies on tree species diversity, ecologically valuable trees and structural diversity were most often conducted in boreal coniferous and temperate broadleaved forests. Understorey plant species were assessed in various terrestrial biomes, from mixed mountainous to broadleaved Mediterranean forests. Tree and plant species richness (i.e. the number of species within a defined region) or diversity (e.g. Shannon diversity index) were most often assessed based only on the trees (George-Chacon et al., 2019), although some studies also considered shrub and herb species (Leutner et al., 2012). Only a small number of studies were interested in the assessment of ecologically valuable tree species, such as aspen (Maltamo et al., 2015; Sankey, 2012). Structural diversity of forests was commonly assessed with the Gini Coefficient (GC) of tree size inequality with respect to DBH or tree height (Adhikari et al., 2020; Valbuena et al., 2014). Also, standard deviation values associated with DBH and tree height were also utilised (Mura et al., 2015). Studies that were assigned to this category mainly utilised regression techniques, such as linear regression (Ceballos et al., 2015) to predict the attributes of interest.

In the review, studies reported widely varying model fits for both tree species richness and diversity (R^2: 0.30–0.89), and understorey richness and diversity (R^2: 0.2–0.6). Both tree species and understorey species richness could be predicted more accurately than diversity when ALS data were used (Leutner et al., 2012; Mauya, 2015). Model fits for structural diversity studies also exhibited a varying trend (e.g. R^2 for GC_{DBH}: 0.33–0.89). One of the most important ALS metrics for tree species richness/diversity and structural diversity predictions was the standard deviation of vegetation height (Fricker et al., 2015; Mohammadi et al., 2020). The ALS metrics related to heterogeneity in the lower parts of the vegetation layers were considered the most important for the assessment of understorey richness/diversity, particularly the low percentiles of vegetation heights (Bourgouin et al., 2022). The combination of ALS data and optical data was not reported to significantly increase the performance of tree or understorey species richness/diversity models (Singh et al., 2015).

The poorest fits for the structural diversity models were reported for forests with a very diverse structure (Adhikari et al., 2020). Mapping of ecologically valuable aspen was reported to be difficult as the species shared similar ALS intensity metric values with spruce and birch (Korpela et al., 2010).

Dead wood

Dead wood plays an important role in nutrient cycling in forests and also provides a habitat and resources for multiple species. Large-diameter dead wood is especially of interest, as the increment of dead wood diameter positively correlates with forest biodiversity and the naturalness of forested areas (Jönsson and Jonsson 2007).

Studies in this topic were mostly conducted within protected forests in the temperate and boreal ecoregions, which naturally contain a large amount of dead wood. Some studies assessed fallen (Heinaro et al., 2021; Lindberg et al., 2013) and standing dead trees (Amiri et al., 2019, Hardenbol et al., 2022) separately, while other studies took both types into consideration (Pesonen et al., 2008). The objective was often to identify single standing trees or fallen dead trees (Blanchard et al., 2011), although characteristics of the dead wood, such as volume (Chirici et al., 2018), diameter (Stitt et al., 2022) and the proportion of standing dead tree (decay) classes (Bater et al., 2009), were also predicted.

There was a clear variation in the detection accuracies for both standing and fallen dead trees across studies (overall accuracy: ~40% to >90%). Detection accuracy clearly increased with increasing diameter and length of the stem (Hardenbol et al., 2022; Heinaro et al., 2021) but there were no clear differences in detection accuracies between fallen or standing trees. The most powerful ALS metrics in dead wood assessment were related to canopy heterogeneity, such as COV of vegetation heights (Martinuzzi et al., 2009), and to the geometrical structure of the canopy surface (Hardenbol et al., 2022). Also, waveform information, such as return width and amplitude, increased detection accuracy (Mücke et al., 2012). Incorrect detection of dead trees was often due to clusters of fallen trees and dense understorey vegetation (Blanchard et al., 2011), but roads, stone walls and ditches were also problematic (Lindberg et al., 2013).

Fragmentation

Forest fragmentation affects the structure, function and biodiversity of forests (Laurance et al., 2017). The microclimatic conditions on the forest edge are different from the interior, as the air is drier and wind effects are more prominent (Didham and Lawton 1999).

In the review, the number of studies that investigated forest fragmentation was low. Study areas were mostly located in tropical and temperate broadleaved forests. The aims of the reviewed studies could be roughly designated as studies that focused on within-fragment properties (Almeida et al., 2019; Vaughn et al., 2014) and studies that investigated inter-fragment connections (Guo et al., 2018). Two studies investigated the long-term effects of fragmentation on canopy structure (Almeida et al., 2019; Vaughn et al., 2014) and one study investigated the effects of fragmentation and fragment area on bird species richness (Flaspohler et al., 2010). In general, ALS information was reported to be suitable to expand our understanding of fragmentation effects on forest structure and function (Vaughn et al., 2014). Almeida et al. (2019) reported that ALS information on mean canopy height showed clear differences between near-edge and inner forests.

Animal ecology

Habitat structural heterogeneity has long been recognised as the main driver of local biodiversity (MacArthur and MacArthur 1961). In forests, plant communities drive such heterogeneity, which further affects the diversity and richness of various animals.

Studies on faunal species richness (α-diversity) and species diversity indices were mainly conducted in temperate and boreal forests. In contrast, habitat suitability studies were conducted in various terrestrial biomes that ranged from mountainous ecosystems to riparian forests and wetlands. In habitat suitability studies, the main objective was to map the occurrence of old-growth forest species and to use their occurrence as indicators of suitable habitats for a specific animal or animals (Vogeler et al., 2013). Other studies also predicted habitat quality classes for specific species through the mean body mass of birds, for example (Hinsley et al., 2006). To date, most of the research has focused on bird species (Melin et al., 2018; Herniman et al., 2020), but some studies have shown interest in butterflies (Zellweger et al., 2016; Zellweger et al., 2017; de Vries et al., 2021), spiders (Vierling et al., 2011) and other forest-dwelling species (Lindberg et al., 2015; Vogeler et al., 2022).

In general, faunal species richness, diversity indices and habitat suitability have been modelled using regression techniques. The R^2 values for all taxa ranged between 0.18–0.96 and 0–0.34 for species richness and diversity indices, respectively. Group-specific accuracies across bird species varied widely and were in some cases much greater (R^2: ~ 0.50) than total bird species richness. Spider species richness could be predicted with similar accuracy as birds, however, the R^2 values for forest beetle- and butterfly- richness models were generally lower (Lindberg et al., 2015; Zellweger et al., 2016). Taxa of the modelled species or forest terrestrial biome did not have any notable effect on model performance for species diversity indices. The most powerful ALS metrics were similar across faunal species richness, diversity and habitat suitability studies: lower vegetation height and density metrics, and metrics that describe the vertical heterogeneity in the canopy. Climatic variables (e.g. temperature, precipitation) were found to have a greater explanatory power than ALS metrics for bird species richness (Carrasco et al., 2019; Zellweger et al., 2016) and diversity indices (Zellweger et al., 2017), especially at the country-scale. In the diversity index models, the ALS metrics were reported to be superior to, or at least as accurate as, field-derived explanatory variables (Müller and Brandl 2009; Sasaki et al., 2016).

The inclusion of aerial or satellite imagery variables in the diversity index modelling had little effect on accuracy (Melin et al., 2019). In the habitat suitability studies, the addition of optical image data was reported to lead to only slight improvements in model performance, with the improvement more pronounced for species connected to deciduous trees (Eldegard et al., 2014).

Mapping of large European aspen in Finland (II)

The results of study **II** highlighted the difficulties in the detection of large European aspen trees in a population that realistically portrays the imbalance in class representations. The utilisation of SMOTE data augmentation was found to improve detection performance. Confusion matrices for the first classification (SC1 II) are presented in Table 3 and the F1-score values for SC2 II are shown in Fig. 5.

The number of large aspen trees in the dataset strongly declined when the DBH limit was increased. For example, the number of aspen with DBH values ≥18 cm and ≥26 cm was 64 and 22, respectively. In SC1 II, 29 large aspen were linked to remotely detected tree crowns. In addition, a link between a remotely detected tree and 6 large aspen was not possible given the distance restrictions of the linking procedure. This resulted in a total number of 35 large aspen trees in SC1 II. In the aspen plots, the RF method (Table 3, RF-A) was used to identify 10 large aspen out of 29, and the SMOTE-RF method (Table 3, SRF-A) 16 out of 29. When all 701 plots were used, the number of correctly identified large aspen trees was 1 and 7 (out of 35) with the RF-C and SMOTE-RF methods, respectively.

The F1-scores for dataset A were similar between the RF (0.42) and SMOTE-RF methods (0.44) because a considerable number of large aspen were incorrectly predicted in the latter. The SMOTE-RF method obtained smaller precision (0.36 vs. 0.56) and better recall (0.55 vs. 0.34) values than RF, which indicated a greater proportion of incorrectly predicted aspen. The recall values for dataset B with RF and SMOTE-RF classifications were 0.05 and 0.09 smaller, respectively, than in dataset A. The F1-scores for both classifications were 0.04 smaller than dataset A. The difference in the statistics between datasets A and B remained relatively modest because most of the field-measured trees were not aspen and most of the field measured large aspen were linked to a remotely detected tree.

Table 3. Confusion matrices associated with first classification (SC1 II) where the random forest (RF) algorithm and Synthetic Minority Oversampling Technique augmented RF classification (SMOTE-RF) were compared across different datasets. Labels "Aspen" and "Other" denote large aspen and other trees, respectively. Suffix "A" denotes the dataset where only linked trees from aspen plots were considered, suffix "B" denotes the dataset where unlinked field-measured trees from aspen plots were added to previous results, and suffix "C" denotes the dataset where all trees from all plots were used in the classification.

Observed class

Predicted class

RF-A	Other	Aspen	Sum
Other	**430**	19	449
Aspen	8	**10**	18
Sum	438	29	**467**

SRF-A	Other	Aspen	Sum
Other	**409**	13	422
Aspen	29	**16**	45
Sum	438	29	**467**

RF-B	Other	Aspen	Sum
Other	**1371**	25	1396
Aspen	8	**10**	18
Sum	1379	35	**1414**

SRF-B	Other	Aspen	Sum
Other	**1350**	19	1369
Aspen	29	**16**	45
Sum	1379	35	**1414**

RF-C	Other	Aspen	Sum
Other	**33159**	34	33193
Aspen	1	**1**	2
Sum	33160	35	**33195**

SRF-C	Other	Aspen	Sum
Other	**33136**	28	33164
Aspen	24	**7**	31
Sum	33160	35	**33195**

The advantage of SMOTE-RF was clear when all plots were utilised (dataset C) as it resulted in a greater F1-score than RF (0.21 vs. 0.06). Moreover, SMOTE-RF also predicted a greater number of large aspen, although many were incorrectly predicted (precision 0.23). Further inspection on the tree species distribution of false positives in SRF-C revealed that most of the false positives (20 trees out of 24) were birch trees (*B. pendula* Roth and *B. pubescens* Ehrh.), three were spruce (*Picea abies* (L.) H. Karst.) and one was pine (*Pinus sylvestris* L.).

In SC2 II, tree-level F1-scores increased as the DBH limit increased (Fig. 5). The number of correctly classified aspen at the tree-level varied between 5 and 7 across a range of DBH limits. At the plot-level, the number of correctly classified aspen plots remained approximately similar (7–9), although the number of aspen plots (i.e. plots with ≥ 1 large aspen) declined from the 18 cm DBH limit (n = 33) to the 26 cm limit (n = 17). Overall, F1-scores increased with increasing DBH, although the 24 cm limit exhibited greater precision than the 26 cm limit (Fig. 5), while F1-scores at the plot-level were greater than their equivalents at the tree-level.

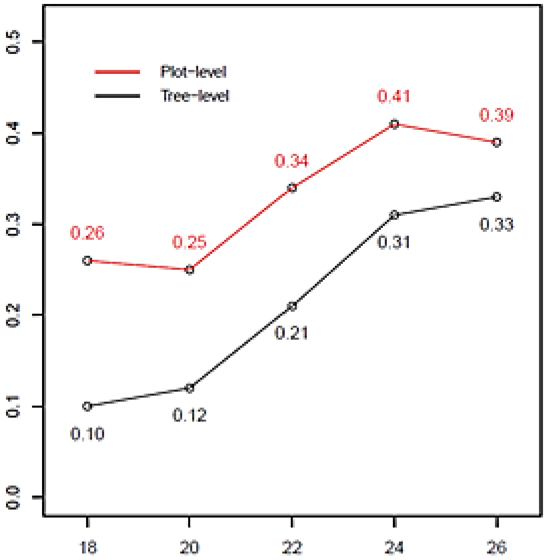

Figure 5. Plot- and tree-level F1-scores associated with second classification (SC2 II) where the Synthetic Minority Oversampling Technique augmented RF classification (SMOTE-RF) was used to identify large aspen trees across a range of diameter at breast height (DBH) limits at the tree- and plot-levels.

In SC1 II, we analysed the most important remote sensing metrics for SMOTE-RF when all trees from aspen plots (SRF-B) and all trees from all plots (SRF-C) were used. The most important remote sensing metrics yielded similar results between these datasets. In general, image metrics were more important predictors than ALS metrics. In particular, metrics that contained the near-infrared (NIR) band were important. Examples of these included NIR/Blue, NIR/Red and standard deviation of NIR. The green band appeared to be the second most important spectral band. The most important ALS metric was ranked 8th in importance with the aspen plots and 12th with all plots. The most important ALS metrics were related to lidar intensity (e.g. standard deviation of intensity for the *last* and *only* echoes)

Forest age prediction using tree boosting with linear forest-type effects (III)

The most essential findings of the study **III** were the superior performance of GPBoost compared to LME, and that the inclusion of categorical variables clearly decreased the prediction error for forest plot age. Predicted vs. observed age scatter plots of the three scenarios are presented in Figure 6.

In SC1 III, RMSE values were large for both GPBoost and LME (Fig. 6). The RMSE value associated with the LME model was more than 4 percentage points larger than the value associated with GPBoost. Neither of the methods was able to accurately predict the ages of plots that had an observed value close to zero. The LME model clearly overpredicted the age of young plots, while the age of old plots was underpredicted.

Overall, the predictions of GPBoost were closer to the observed values than the predictions from the LME model (Fig. 6).

In SC2 III, accuracies were greater in both methods compared to SC1 III (Fig. 6). The RMSE value associated with LME decreased more than the RMSE value associated with GPBoost. Also, GPBoost in SC2 III was clearly the better alternative of the two methods. As in SC1 III, the age of plots with observed ages close to zero was difficult to predict accurately in both methods, but this tendency was stronger in LME. For both methods, the predicted values in SC2 III better corresponded to the observed values than in SC1 III. Moreover, the overprediction of young ages and underprediction of old ages in SC2 III was clearly reduced in LME compared to SC1 III. This was observed as a much wider range of predictions in SC2 III than in SC1 III.

In SC3 III, a hold-over plot classification was added prior to age prediction. Its purpose was to separate age predictions for hold-over and other plots. Hold-over plot classification with a logistic regression model provided an overall accuracy value of 94.7% and a Kappa coefficient of 0.58 (as the mean of 50 repetitions). The performance of age predictions for both methods was further improved compared to SC2 III (Fig. 6). Biases were clearly larger in this alternative compared to SC1 III and SC2 III (where biases were marginal). Larger bias originated from the splitting of age modelling between hold-over plots and other plots. However, plots with observed ages close to zero were clearly predicted more accurately than in the previous scenarios. Poor age predictions for the remaining single hold-over plots close to zero could be mostly accounted for by the fact that the hold-over classification failed in those plots.

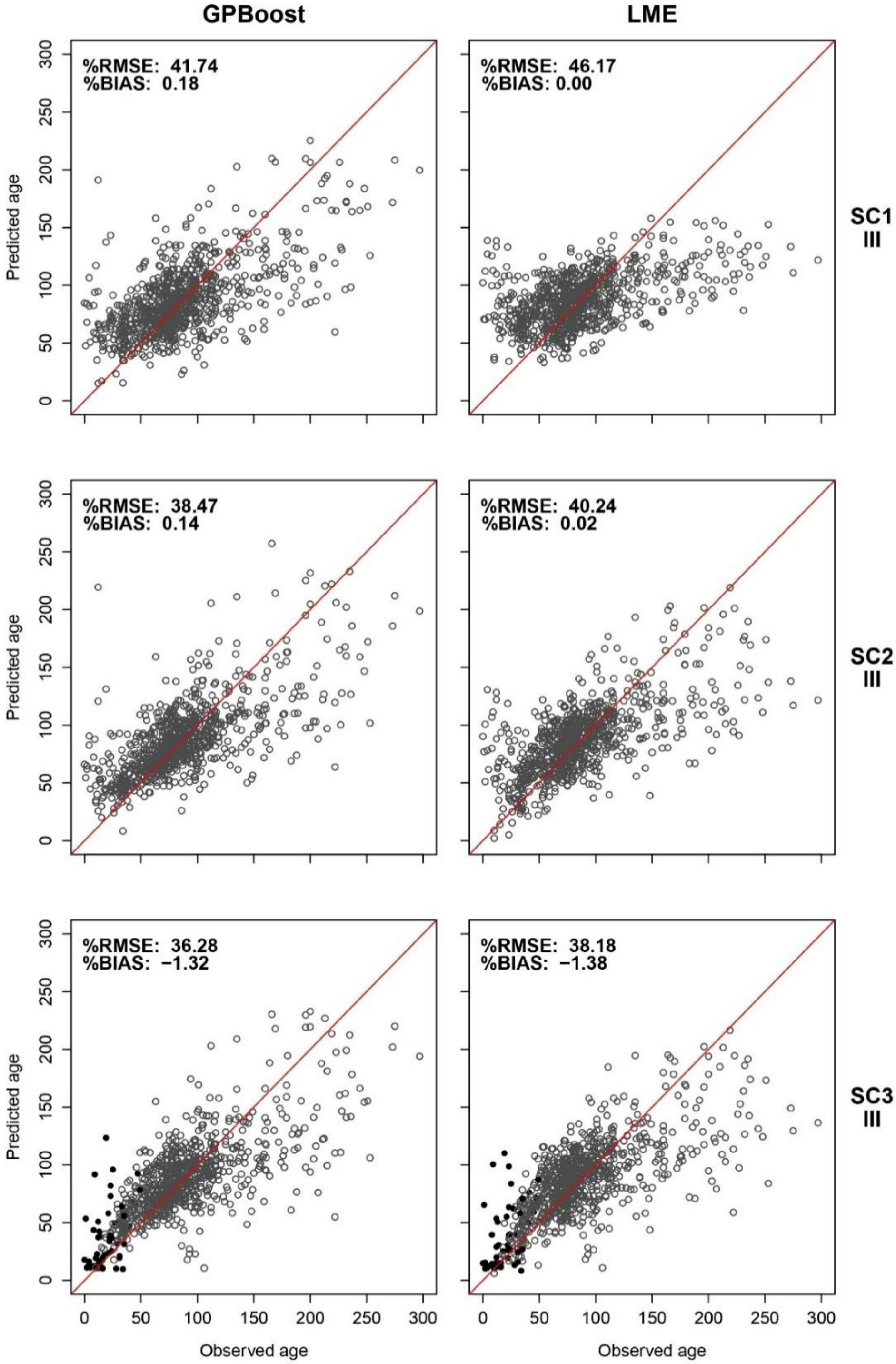

Figure 6. Predicted vs. observed age for tree boosting with random effects (GPBoost) and the linear mixed effects (LME) model in the first scenario (SC1 III), second scenario (SC2 III) and third scenario (SC3 III). The observed hold-over plots are highlighted in bold in the SC3 III scatter plots.

In study **III**, the proportion of plots predicted as ≥150, ≥175 and ≥200 years old from the number of observed plots in these age classes, was analysed. One trend was obvious: GPBoost predicted older forests more accurately than LME. For example, in SC1 III, LME did not predict any plots for age classes 175 and 200, while in contrast, GPBoost predicted 29.1% and 23.5% of the number of observed plots in these classes, respectively. In SC1 III, GPBoost predicted 43.3% and LME 8.4% of the number of observed plots in the 150 year old class. Another trend was that the proportion of predicted plots observed in each class declined from age class 150 to age class 200, especially for LME. In SC2 III, the difference in percentage values between methods was not as large as in SC1 III, but GPBoost still yielded larger values, especially in age class 200 (41.2% vs. 11.8%). In SC3 III, both methods yielded identical percentage values for age class 150 (44.6%), but the greatest difference was again in age class 200 in favour of GPBoost (38.2% vs. 8.8%). It is also worth noting that SC2 III yielded slightly larger proportions in age classes 150 and 200 for both LME and GPBoost.

The proportion of plots predicted as 150, 175 and 200 age classes did not provide any information as to whether the predictions were false positives, i.e. the actual age was less than the age limit. A clear trend was observed that showed that the proportion of false positives increased with an increased age limit. In age class 150, approximately one quarter of plots were false positives. In age class 200, almost half of the plots were younger than 200 years. Furthermore, GPBoost predicted a lower rate of false positives than LME.

The random effects of the SC3 III models for “other plots” are shown in Figure 7. The effect of the ALS block variable was minor compared to other categorical variables and is not shown here. The "Protected forests” category yielded the strongest random effect with an increase in the predicted age in GPBoost by 40 years and by 30 years in the LME model. The second strongest random effect was in the “undrained peatlands” category, which increased the predicted age by 25 years in GPBoost and by 20 years in the LME model. For the “drained peatlands” category, the increase in age was 10 years in GPBoost and 5 years in the LME model. The “Mineral soil” category yielded the strongest negative random effect. It decreased the predicted age by 10 years in GPBoost and by 25 years in the LME model. For plots with “no wood production restrictions”, the predicted age decreased by approximately 5 years in GPBoost and by 25 years in LME. The random effect for the “protected mires” category in GPBoost increased the predicted ages by 20 years, although for the LME model, it decreased predicted ages by 5 years. The effect of the “protected mires” category was negative in LME but positive in GPBoost, which is likely caused by the different prediction levels (between GPBoost and LME) in the fixed parts of the models.

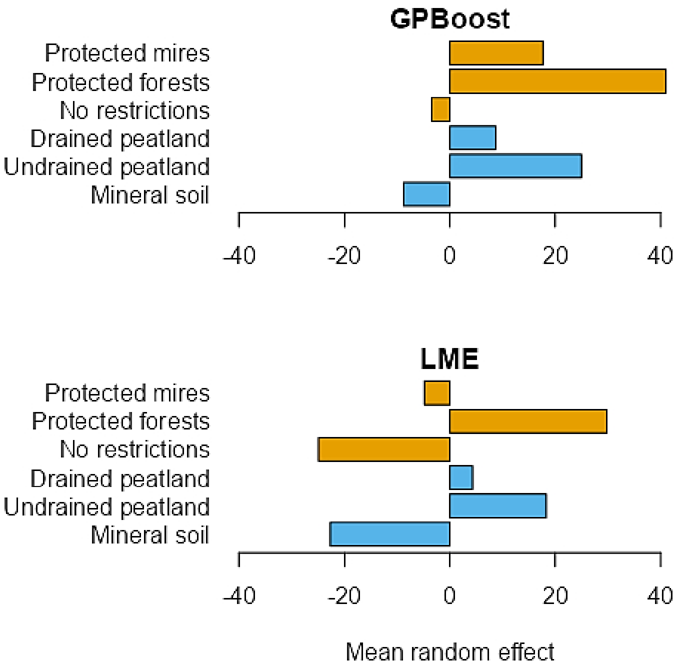

Figure 7. Mean random effects by grouping variables in the “other plots” models of scenario 3 (SC3 III). Effects are presented separately for tree boosting with random effects (GPBoost) and the linear mixed effects (LME) model. Categorical variables are presented in the following colours: WoodProdRestr (levels 2,1,0) as orange and MainType (levels 2,1,0) as blue.

The most important metrics were observed in the GPBoost models in SC3 III for age models on hold-over and other plots, and in the hold-over plot classification. Overall, ALS metrics were more important predictors for age prediction than Sentinel-2 metrics. In the case of hold-over plot classification, the 10 most important predictors included all categorical variables (n=3) and four Sentinel-2 metrics. The most important ALS metrics for hold-over and other plots were partly the same: relative density of the *first* and *last* ALS echoes between 5 and 10 m, and the 60th bincentile of vegetation heights were included in the 10 most important variables in both cases. For other plots, field-measured height of a plot (TH) was ranked as the second most important metric, but was not in the ten most important metrics for hold-over plots. Another difference between the most important metrics in the age models for hold-over and other plots was that the most important Sentinel-2 metric (B3) was ranked as second for hold-over plots, but for the other plots, the most important Sentinel-2 metric (B8/B11) was only the fifth most important. Also, the most important Sentinel-2 metric for hold-over plots (B3) was not ranked in the ten most important metrics for other plots. For both hold-over plots and other plots, half of the 10 most important metrics were Sentinel-2 and ALS metrics, respectively. The ratio of some Sentinel-2 bands (e.g. B8/B11) was often the more important metric than the Sentinel-2 band *per se* (e.g. B8).

DISCUSSION

Role of ALS data in the assessment of forest biodiversity

Heterogeneity of the vegetation structure affects the patterns in biological diversity (Heidrich et al., 2023). The more heterogenous areas provide greater niche space than areas with low heterogeneity, and so allow more species to coexist (Stein et al., 2014). Also, the more heterogenous areas provide shelter from hazardous environmental conditions, which should promote species resilience.

At this point in time, ALS is a well-established technology used to assess such heterogeneity in forested ecosystems (Adhikari et al., 2020). Moreover, quantification of forest 3-D structure via ALS data has key advantages over field-based approaches: ALS-based approaches are less labour-intensive and can cover larger spatial areas (50–100 km^2) during a single flight.

Study **I** reviewed 182 scientific articles that utilised ALS data as the main RS data to assess various topics on biological diversity in forested ecosystems. Most of the study areas were located in Europe and North America, and temperate and boreal ecoregions were the most commonly assessed regions. The reasons for such clustering are the high level of expertise in this domain, the information needs of decision-makers and the fact that ALS data are widely available from these areas. The review showed that the most studied topics were animal ecology, tree species diversity/richness and dead wood.

The clear majority of reviewed studies in study **I** utilised DR-ALS data; FW-ALS data were used only in a small number of studies. The ALS data were often combined with other remote sensing data, especially with optical satellite or aerial data; the premise here is to utilise complementary data sources that describe both the spectral and structural features of the forest. This fusion is especially beneficial for the assessment of tree species diversity measures and forest land cover classification, when tree species information is needed (George-Chacon et al., 2019; Jayathunga et al., 2018). The assessment of dead wood (Pesonen et al., 2010), forest understorey (Bouvier et al., 2017) and successional stages (Martinuzzi et al., 2013) did not appear to benefit from the addition of optical data. One probable reason for this is that optical data does not describe below-canopy conditions as its information is limited to the upper-most canopy layer (i.e. visible from above).

Based on the results from study **I**, there is no single ALS metric that would be suitable for all forest biodiversity assessments. The ALS metrics most often utilised in the assessment of forest biodiversity describes the central tendency and dispersion of ALS vegetation heights. Especially, standard deviation, mean and COV of vegetation heights were employed across reviewed topics, while ALS-intensity was reported to be useful in studies that considered tree species (e.g. the classification of canopy types or forest habitat types). Another often utilised ALS-based variable group is the terrain variables that describe local topography (e.g. slope and altitude), which performed well in the assessment of non-flying mammals and arthropods (Zellweger et al., 2017). This is because terrain variables are associated with solar radiation and hydrography, which further influence these taxonomic groups. It is worth noting that the derived metrics, especially ALS intensity metrics, are dependent on the ALS sensor, acquisition parameters and forest type, which means that the relevant metrics and observed relationships in one study are not necessarily universal.

Currently, ALS data provides meaningful measures for the classification of forested areas based on the assessment of canopy structure (e.g. layering) and for the assessment of structural diversity (**I**). Despite the small number of studies that have examined forest fragmentation, ALS data were found to be relevant to expand our understanding of fragmentation effects on forest structure and function (Almeida et al., 2019). Research has also shown that ALS data is highly usable in the assessment of animal-habitat relationships (Davies and Asner 2014). In particular, bird populations have received considerable attention in ALS-based animal ecology, which is understandable given the implicit 3-D nature of avian ecology and the strong suitability of ALS data to map forested environments (Davies and Asner 2014).

However, ALS data alone is not always sufficient when tree species are assessed directly or indirectly with regard to a specific biodiversity indicator. Currently, forest biodiversity indicators that are difficult to assess with ALS data include diversity indices of forest fauna (Wallis et al., 2016) and understorey flora (Barber et al., 2016), and fallen dead trees (Hardenbol et al., 2022). One issue often reported is the low point density of the utilised ALS data (here <10 pt./m^2). With greater point density, lower canopy structures could be better described and new approaches enabled to retrieve information from the ALS data. In the case of fallen dead trees, the difficulties are manifold: clusters of fallen dead trees, dense understorey vegetation, ditches, roads and stone walls have all been reported to cause incorrect detection. Also, detection of later decay stage fallen dead trees has been shown to be more difficult than the detection of earlier decay stage trees because the height of the stem objects above ground is reduced (Mücke et al., 2013). In assessment of bird species richness, climatic variables (e.g. temperature and precipitation) were sometimes reported to be even more important than ALS metrics (Zellweger et al., 2016) and diversity (Zellweger et al., 2017).

As study **I** included scientific articles from 2003–2022, a brief review of articles from 2023 until the writing of this dissertation was carried out detect new trends in the domain. From this additional review, it was found that the number of studies in 2023 (n=22) and 2024 (n=17) had increased from the peak years observed in study **I**. Since 2022, the most studied topics were somewhat similar to those identified in study **I**: dead wood and tree species diversity, although animal ecology and forest undergrowth were studied to a lesser extent than previously. The greatest difference to the trends noted in study **I** was the clustering of study areas: since 2022, more than 75% of the studies were conducted in Europe. Also, there was a slight increase in the number of ALS-biodiversity studies conducted in Asia. One potential cause of such a high number of (a) dead wood studies and (b) studies in Europe, are the goals of both the EU Forest Strategy and the EU Biodiversity Strategy (European Commission 2021), which aim to protect all remaining primary and old-growth forests, and one of the indicators of such forests is the amount/presence of deadwood (European Commission 2023).

To date, the utilisation of ALS data to map rare but ecologically valuable tree species has not been studied to any great extent. In the case of European aspen in boreal forests, this is mainly due to overlapping ALS intensity metrics between aspen and spruce (Korpela et al., 2010) and the overlapping spectral response between aspen and birch (Viinikka et al., 2020), but also to the naturally rare occurrence of aspen trees in the prevailing tree population in managed forests (Maltamo et al., 2015). Study **II** provides a realistic picture of the mapping of rare phenomena in true populations: in the field data, the proportion of aspen with respect to stem number was 0.49%, while the Finnish NFI estimate for the larger geographical area around the study area was 0.5%. The proportion of large aspen was even smaller: approximately 0.1% for aspen with DBH ≥ 22 cm. The rarity aspect of aspen in the population has not been adequately considered in earlier studies. For example, in the studies by Viinikka et al. (2020) and Kuzmin et al. (2021), the proportion of aspen were estimated at approximately 11% and 27%, respectively. From a biodiversity perspective, it is meaningful to identify large aspen trees, whether clustered or few and far between. This is why the basic unit of interest in study **II** was the tree – not the quantity. In addition, the classification accuracy (absence or presence of large aspen) was validated at the plot-level, which was based on the predictions at the tree-level in each plot.

The low accuracy values reported in study **II** highlight the inherent challenges in remote sensing-based detection of large aspen trees. For example, the highest F1-scores were reported for the least representative dataset of the population, while the lowest F1-scores were reported for the most representative dataset. SMOTE, which balances the training data by (synthetically) oversampling the minority class, was clearly beneficial, even though the accuracies of large aspen detection were still low. The greatest benefit of SMOTE was its ability to increase the number of correctly predicted large aspen compared to the method that did not utilise SMOTE. When the SMOTE data augmentation was utilised, the number of predicted large aspen was almost identical to the number of observed large aspen, but a large proportion were false positives and truly large aspen were not identified. It is evident that SMOTE improves the recall at the price of reduced precision but can still be considered useful for imbalanced classification scenarios, such as large aspen. Minority class oversampling by the duplication of minority class observations would lead to overfitting of the model. Instead, SMOTE combines the features of the k-nearest neighbours in the dataset, which results in completely new datapoints in the training data.

A comparison of the results in study **II** to other studies is difficult as the proportion of aspen compared to other deciduous trees/other trees varies considerably across studies. Indeed, the number of tree species and differences in overall vegetation structure has been found to differ widely across studies.

The accuracy statistics reported in study **II** were low compared to many earlier aspen tree detection studies. The reason for this anomaly most probably lies in our field dataset, which realistically described the low proportion of aspen in the forest area reported in study **II**. Other studies that have reported high accuracy values either contained a large proportion of aspen trees among deciduous trees (> 50%) or the proportion of aspen was large overall (> 30%). For example, Li et al. (2013) classified individual tree crowns into four species, one of which was quaking aspen (*P. tremuloides* Michx.). The proportion of aspen in their field data was >30% of all trees and approximately 63% of deciduous trees. For comparison, the proportion of aspen of all deciduous trees in study **II** was only 1.8%.

Historically, the prediction of forest age with ALS data (Wylie et al., 2019; Maltamo et al., 2020) and optical remote sensing data (Cohen and Spies 1992) has been a difficult task. In study **III**, we investigated the issue by comparing two alternative methods used for the prediction of forest age with ALS and aerial image data: tree boosting combined with random effects (GPBoost) and linear mixed effects modelling (LME). It was shown that the additional variables that describe site conditions (e.g. MainType of plot) improved the prediction accuracy for both GPBoost and LME. Their high importance originates from the fact that these variables described the growing conditions that the ALS data could not explain. For example, forests located on undrained peatlands are less productive compared to drained peatlands, which causes the trees in the former to be older than trees in the latter provided the spectral response and 3-D structure are assumed to be similar.

In study **III**, the predictions of the GPBoost algorithm were superior to LME in each tested scenario. The greatest differences occurred in the age predictions when the categorical variables that describe site conditions were not utilised in the models. The age prediction methods considered were fundamentally different: GPBoost is a non-parametric machine learning method, whereas LME is a parametric statistical method and assumes that the target variable (here, age of a plot) can be expressed as a linear combination of the explanatory variables. In contrast, GPBoost as a gradient-boosted tree method attempts to approximate functions for any type of relationship, and it can be used in both classification and regression tasks. However, LME does have a number of advantages: it is faster to converge and it is generally more interpretable than GPBoost. On the other hand, if the data contains non-linearities, GPBoost will generally fit the data better than LME, which results in greater prediction accuracy. Also, GPBoost can handle high-dimensional data and is more robust when there are non-linear relationships between variables.

There are major challenges in remote sensing-based prediction of ages in forests that are older than approximately 100 years due to the diminishing link between spectral information and age, and a weakened height-age relationship (Cohen and Spies 1992; Maltamo et al., 2020). This is why some earlier studies have ignored prediction of older plots (Maltamo et al., 2020) or have fitted a separate model for the older plots (Wylie et al., 2019). In study **III**, it was evident that most of the variation in age predictions occurred after 150 years, although predicted ages closely followed observed plot ages up to 250 years old (Figure 6, SC3 III, GPBoost). The weakened age-height relationship can partly be accounted for by the reduction in height growth. In the later successional stages, tree height growth will eventually cease, but the tree diameter continue to grow and this causes significant variation in age for trees of similar heights.

A comparison of the results in study **III** to earlier studies that have predicted forest age with remote sensing is not straightforward. For instance, there are considerable variations in how the ALS and image data are utilised, and the forests vary in many ways. Studies that have a narrower age range in the field data generally report smaller prediction errors than in study **III**. This was the case with Racine et al. (2014) who reported an age range between 11 and 94 years in their dataset, whereas the age range for field plots in study **III** was between 0 and 300 years. The RMSE value reported by Racine et al. (2014) was clearly lower than the smallest value reported in study **III** (19% vs. 36%). In Finland, Maltamo et al. (2020) predicted the forest plot age for plots with ages between 10 and 100 years and the lowest RMSE value in their study was 32.9%, which was similar to the value reported in study **III**.

One of the key challenges in study **III** was the age definition of utilised NFI data. Ages for hold-over plots were defined based on the dominant tree stratum (i.e. seedling), although tall hold-over trees left from the previous tree generation caused substantial errors in the modelling of plot age. Also, there were plots that had observed ages near 100 years old and predicted ages between 0–30 years. These plots were found to have been recently harvested and the seed trees were left from the previous tree generation. On these plots, however, age was defined based on seed trees as the seedlings had not yet emerged. If the ages of plots were defined from a biodiversity perspective (i.e. based on hold-over trees) rather than from a forest management perspective, the error rate of age prediction would have been lower.

There are biodiversity indicators that have not yet been assessed in-depth with ALS data (study **I**). One such is functional diversity, which quantifies the range of functional variation in the population. Instead, species taxonomic diversity studies, especially related to species richness, have received considerable attention. However, there is a drawback with species richness, as it does not give information on the abundance of species: there can be two sites that have same number of species, but they are not necessary the same species. Functional diversity is considered to be a stronger predictor of ecosystem functioning than species richness (Petchey et al., 2004; van der Plas 2019). Studies on functional diversity have shown that greater functional diversity leads to more productive and resilient ecosystems (Schmitt et al., 2019; Schneider et al., 2023). Functional diversity studies that have utilised ALS data are still rare, but have been conducted in steadily increasing numbers in recent years (Kamoske et al., 2022; Schneider et al., 2023).

In the mapping of large aspen trees, one of the greatest issues observed was the naturally low quantity of aspen trees in the population (study **II**). The use of SMOTE clearly increased the detection accuracy of large aspen trees, but the accuracy statistics still yielded low values due to an increased number of false positives. Another issue to be considered is the crown shape of aspen trees, specifically the assumption that one local maximum at the top of a tree is not always realistic. The increasing point densities in ALS datasets enable more precise determination of the canopy, which could enable better discrimination between aspen and other deciduous trees.

In the assessment of forest age (study **III**), the issue centred mainly on the determination of plot age for different types of field plots. A more uniform determination of plot age could have decreased prediction errors. However, any relevant approach to define plot- or stand-level age may differ depending on silvicultural or ecological viewpoints. Another issue was the fact that the height growth of a tree eventually ceases in later successional stages so age can vary significantly for trees of similar heights as both diameter and crown continue to grow for a longer period than tree height. The increased point density in new ALS datasets could open up novel possibilities for a more detailed assessment of the tree crown. For example, the prediction of tree/plot/stand age could make use of crown biomass predictions to better discriminate trees of different ages.

CONCLUSIONS

Based on the conducted literature review, research on ALS-based assessment of forest biodiversity has been shown to have clustered on the European and North American continents, and most of the research currently focuses on animal ecology, tree species richness/diversity measures and the assessment of dead wood. However, there are no ALS metrics that would suit all the various assessments of forest biodiversity, although some ALS metrics have been used more often than others. These are mainly related to dispersion and the central tendency of vegetation height; standard deviation, mean and COV, for example.

It is common that the biodiversity indicator of interest is rare in a population. Data augmentation using the SMOTE algorithm was found to be beneficial when the aim was to detect rare tree species and an increased number of false positives is not an issue. However, the mixing of aspen and surrounding broadleaved trees is still an issue in remote sensing-based detection of aspen. Also, the crown structure of large aspen complicates the detection of these trees with remote sensing, which is why a more dedicated approach for the automatic detection and delineation of large aspen needs to be developed. Future studies that deal with the issue of imbalanced data on forest biodiversity indicators should use field data that provide a realistic picture of the population. The utilisation of unrepresentative data of a population provides an overly optimistic picture of the reliability of the utilised method.

Forest age is a highly important attribute in the assessment of forest biodiversity. In the prediction of plot-level forest age, tree boosting with random effects (GPBoost) showed overall better prediction performance than linear mixed effects modelling (LME). The inclusion of categorical site variables clearly increased the prediction accuracy for both GPBoost and LME. The application of tree boosting combined with random effects (GPBoost) in the forest remote sensing arena can also be recommended for the modelling of attributes other than forest age.

This dissertation demonstrated that ALS data provide valuable information on horizontal and vertical vegetation structures of a forest and can assist in the assessment of forest biodiversity indicators, at both fine and broad spatial scales. It is highly advisable to further extend ALS-based forest diversity research to investigate functional diversity.

REFERENCES

Adhikari, H., Valbuena, R., Pellikka, P., Heiskanen, J. 2020. Mapping forest structural heterogeneity of tropical montane forest remnants from airborne laser scanning and Landsat time series. Ecol. Indic., 108, 105739–. https://doi.org/10.1016/j.ecolind.2019.105739.

Adnan, S., Maltamo, M., Coomes, D., García-Abril, A., Malhi, Y., Manzanera, J., Butt, N., Morecroft, M., Valbuena, R., 2019. A simple approach to forest structure classification using airborne laser scanning that can be adopted across bioregions. For. Ecol. Manage. 433, 111–121. https://doi.org/10.1016/j.foreco.2018.10.057.

Alberti, G., Boscutti, F., Pirotti, F., Bertacco, C., de Simon, G., Sigura, M., Cazorzi, F., Bonfanti, P., 2013. A LiDAR-based approach for a multi-purpose characterization of alpine forests: An Italian case study. Iforest (Viterbo) 6 (1), 156–168. https://doi.org/10.3832/ifor0876-006.

Almeida, D. R. A., Stark, S. C., Schietti, J., Camargo, J. L. C., Amazonas, N. T., Gorgens, E. B., Rosa, D. M., Smith, M. N., Valbuena, R., Saleska, S., Andrade, A., Mesquita, R., Laurance, S. G., Laurance, W. F., Lovejoy, T. E., Broadbent, E. N., Shimabukuro, Y. E., Parker, G. G., Lefsky, M., … Brancalion, P. H. S. 2019. Persistent effects of fragmentation on tropical rainforest canopy structure after 20 yr of isolation. Ecological Applications 29(6), e01952–n/a. https://doi.org/10.1002/eap.1952.

Álvarez-Martínez, J.M., Jim´enez-Alfaro, B., Barquín, J., Ondiviela, B., Recio, M., Silió-Calzada, A., Juanes, J.A., Isaac, N., 2018. Modelling the area of occupancy of habitat types with remote sensing. Methods Ecol. Evol. 9 (3), 580–593. https://doi.org/10.1111/2041-210X.12925.

Amiri, N., Krzystek, P., Heurich, M., Skidmore, A. 2019. Classification of tree species aswell as standing dead trees using triplewavelength als in a temperate forest. Remote Sens. (Basel, Switzerland) 11(22). https://doi.org/10.3390/rs11222614.

Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. Int. Arch. Photogramm. Remote Sens. 33: 110–117.

Balvanera, P., Siddique, I., Dee, L., Paquette, A., Isbell, F., Gonzalez, A., Byrnes, J., O'Connor, M.I., Hungate, B.A., Griffin, J.N., 2014. Linking biodiversity and ecosystem services: current uncertainties and the necessary next steps. Bioscience 64 (1), 49–57. https://doi.org/10.1093/biosci/bit003.

Bakx, T.R.M., Koma, Z., Seijmonsbergen, A.C., and Kissling, W.D. 2019. Use and categorization of light detection and ranging vegetation metrics in avian diversity and species distribution research. Divers. Distrib. 25(7): 1045–1059. http://dx.doi.org/10.1111/ddi.12915.

Barber, Q.E., Bater, C.W., Braid, A.C.R., Coops, N.C., Tompalski, P., Nielsen, S.E., 2016. Airborne laser scanning for modelling understory shrub abundance and productivity. For. Ecol. Manage. 377, 46–54. https://doi.org/10.1016/j.foreco.2016.06.037.

Baroni, D., Korpimäki, E., Selonen, V., and Laaksonen, T., 2020. Tree cavity abundance and beyond: nesting and food storing sites of the pygmy owl in managed boreal forests. For. Ecol. Manage. 460: 117818. http://dx.doi.org/10.1016/j.foreco.2019.117818.

Bater, C.W., Coops, N.C., Gergel, S.E., LeMay, V., Collins, D., 2009. Estimation of standing dead tree class distributions in northwest coastal forests using lidar remote sensing. Can. J. For. Res. 39 (6), 1080–1091. https://doi.org/10.1139/X09-030.

Bergen, K. M., Goetz, S. J., Dubayah, R. O., Henebry, G. M., Hunsaker, C. T., Imhoff, M. L., Nelson, R. F., Parker, G. G., Radeloff, V. C. 2009. Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. J. Geophys. Res.: Biogeosci. 114(G2). https://doi.org/10.1029/2008JG000883.

Biodiversity Indicators Partnership. 2011. Guidance for national biodiversity indicator development and use. UNEP World Conservation Monitoring Centre, Cambridge, UK. 40pp. Available from https://www.cbd.int/doc/meetings/ind/ahteg-sp-ind-01/other/ahteg-sp-ind-01-bipnational-en.pdf.

Blanchard, S.D., Jakubowski, M.K., Kelly, M., 2011. Object-based image analysis of downed logs in disturbed forested landscapes using lidar. Remote Sensing (Basel, Switzerland) 3 (11), 2420–2439. https://doi.org/10.3390/rs3112420.

Bottalico, F., Travaglini, D., Chirici, G., Marchetti, M., Marchi, E., Nocentini, S., Corona, P., 2014. Classifying silvicultural systems (coppices vs. high forests) in Mediterranean oak forests by Airborne Laser Scanning data. Eur. J. Remote Sens. 47 (1), 437–460. https://doi.org/10.5721/EuJRS20144725.

Bourgouin, M., Valeria, O., Fenton, N. J. 2022. Predictive mapping of bryophyte diversity associated with mature forests using LiDAR-derived indices in a strongly managed landscape. Ecol. Indic., 136, 108585–. https://doi.org/10.1016/j.ecolind.2022.108585.

Bouvier, M., Durrieu, S., Gosselin, F., Herpigny, B. 2017. Use of airborne lidar data to improve plant species richness and diversity monitoring in lowland and mountain forests. PloS One 12(9), e0184524–e0184524. https://doi.org/10.1371/journal.pone.0184524.

Brokaw, N. V. L., Lent, R. A. 1999. Vertical structure. In I. Hunter, & L. Malcom (Eds.), Maintaining biodiversity in forest ecosystems (pp. 373-399). Cambridge: Cambridge University Press.

Carrasco, L., Giam, X., Papeş, M., Sheldon, K. S. 2019. Metrics of lidar-derived 3D vegetation structure reveal contrasting effects of horizontal and vertical forest heterogeneity on bird species richness. Remote Sens. (Basel, Switzerland), 11(7), 743–. https://doi.org/10.3390/rs11070743.

Ceballos, A., Hernández, J., Corvalán, P., Galleguillos, M., 2015. Comparison of airborne LiDAR and satellite hyperspectral remote sensing to estimate vascular plant richness in deciduous mediterranean forests of Central Chile. Remote Sens. (Basel, Switzerland) 7 (3), 2692–2714. https://doi.org/10.3390/rs70302692.

Chirici, G., Bottalico, F., Giannetti, F., Del Perugia, B., Travaglini, D., Nocentini, S., Kutchartt, E., Marchi, E., Foderi, C., Fioravanti, M., Fattorini, L., Bottai, L., McRoberts, R.E., Næsset, E., Corona, P., Gozzini, B., 2018. Assessing forest windthrow damage using single-date, post-event airborne laser scanning data. Forestry (London) 91 (1), 27–37. https://doi.org/10.1093/forestry/cpx029.

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R. V., Paruelo, J., Raskin, R. G., Sutton, P., van den Belt, M. 1998. The value of the world's ecosystem services and natural capital. Ecological Economics, 25(1), 3–15. https://doi.org/10.1016/S0921-8009(98)00020-2.

Dalponte, M., Bruzzone, L., Gianelle, D., 2008. Fusion of Hyperspectral and LIDAR Remote Sensing Data for Classification of Complex Forest Areas. IEEE Trans. Geosci. Remote Sens. 46 (5), 1416–1427. https://doi.org/10.1109/TGRS.2008.916480

Davies, A.B., and Asner, G.P. 2014. Advances in animal ecology from 3D-LiDAR ecosystem mapping. Trends Ecol. Evol. (Amsterdam), 29 (12): 681–691. http://dx.doi.org/10.1016/j.tree.2014.10.005.

Dees, M., Straub, C., Koch, B., 2012. Can biodiversity study benefit from information on the vertical structure of forests? Utility of LiDAR remote sensing. Curr. Sci. 102 (8), 1181–1187.

de Vries, J.P.R., Koma, Z., WallisDeVries, M.F., Kissling, W.D., Tingley, R., 2021. Identifying fine-scale habitat preferences of threatened butterflies using airborne laser scanning. Divers. Distrib. 27 (7), 1251–1264. https://doi.org/10.1111/ddi.13272.

Didham, R. K., Lawton, J. H. 1999. Edge structure determines the magnitude of changes in microclimate and vegetation structure in tropical forest fragments. Biotropica, *31*(1), 17–30. https://doi.org/10.2307/2663956.

Duffy, J.E., 2009. Why biodiversity is important to the functioning of real-world ecosystems. Front. Ecol. Environ. 7 (8), 437–444. https://doi.org/10.1890/070195.

Dye, M., Mutanga, O., Ismail, R. 2012. Combining spectral and textural remote sensing variables using random forests: predicting the age of Pinus patula forests in KwaZulu-Natal, South Africa. Journal of Spatial Science, 57(2), 193–211. https://doi.org/10.1080/14498596.2012.733620.

Eerikäinen, K., Mabvurira, D., Nshubemuki, L., Saramaki, J. 2002. A calibrateable site index model for Pinus kesiya plantations in southeastern Africa. Canadian Journal of Forest Research, 32(11), 1916–1928. https://doi.org/10.1139/x02-106.

Eldegard, K., Dirksen, J.W., Ørka, H.O., Halvorsen, R., Næsset, E., Gobakken, T., Ohlson, M., 2014. Modelling bird richness and bird species presence in a boreal forest reserve using airborne laser-scanning and aerial images. Bird Study 61 (2), 204–219. https://doi.org/10.1080/00063657.2014.885492.

Esseen, P.A., Ehnström, B., Ericson, L., Sjöberg, K. 1997. Boreal forests. Ecol. Bull. 46: 16–47. Available from https://www.jstor.org/stable/20113207.

European Commission, Directorate-General for Environment. 2021. EU biodiversity strategy for 2030: bringing nature back into our lives. Publications Office of the European Union. https://data.europa.eu/doi/10.2779/677548

European Commission, Directorate-General for Environment. 2023. Commission guidelines for defining, mapping, monitoring and strictly protecting EU primary and old-growth forests, Publications Office of the European Union. https://data.europa.eu/doi/10.2779/481811

Falkowski, M.J., Evans, J.S., Martinuzzi, S., Gessler, P.E., Hudak, A.T., 2009. Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. Remote Sens. Environ. 113 (5), 946–956. https://doi.org/10.1016/j.rse.2009.01.003.

Flaspohler, D.J., Giardina, C.P., Asner, G.P., Hart, P., Price, J., Lyons, C.K., Castaneda, X. 2010. Long-term effects of fragmentation and fragment properties on bird species richness in Hawaiian forests. Biol. Conserv. 143 (2), 280–288. https://doi.org/10.1016/j.biocon.2009.10.009.

Franklin, J.F., Johnson, K.N., Johnson, D.L. 2018. Ecological forest management. Waveland Press Inc, Long Grove, Illinois.

Fricker, G.A., Wolf, J.A., Saatchi, S.S., Gillespie, T.W. 2015. Predicting spatial variations of tree species richness in tropical forests from high-resolution remote sensing. Ecol. Appl. 25 (7), 1776–1789. https://doi.org/10.1890/14-1593.1.

Fuhr, M., Lalechère, E., Monnet, J., Bergès, L., Disney, M., Hernandez-Clemente, R. 2022. Detecting overmature forests with airborne laser scanning (ALS). Remote Sens. Ecol. Conserv. 8 (5), 731–743. https://doi.org/10.1002/rse2.274.

George-Chacon, S.P., Dupuy, J.M., Peduzzi, A., Hernandez-Stefanoni, J.L. 2019. Combining high resolution satellite imagery and lidar data to model woody species diversity of tropical dry forests. Ecol. Ind. 101, 975–984. https://doi.org/10.1016/j.ecolind.2019.02.015.

Gillis, M. D., Omule, A. Y., Brierley, T. 2005. Monitoring Canada's forests: The National Forest Inventory. Forestry Chronicle, 81(2), 214–221. https://doi.org/10.5558/tfc81214-2.

Guo, X., Coops, N., Tompalski, P., Nielsen, S., Bater, C., John Stadt, J., 2017. Regional mapping of vegetation structure for biodiversity monitoring using airborne lidar data. Eco. Inform. 38, 50–61. https://doi.org/10.1016/j.ecoinf.2017.01.005.

Guo, X., Coops, N.C., Gergel, S.E., Bater, C.W., Nielsen, S.E., Stadt, J.J., Drever, M., 2018. Integrating airborne lidar and satellite imagery to model habitat connectivity dynamics for spatial conservation prioritization. Landsc. Ecol. 33 (3), 491–511. https://doi.org/10.1007/s10980-018-0609-0.

Hardenbol, A.A., Korhonen, L., Kukkonen, M., Maltamo, M., 2022. Detection of standing retention trees in boreal forests with airborne laser scanning point clouds and multispectral imagery. Methods Ecol. Evol. 14 (7), 1610–1622. https://doi.org/10.1111/2041-210X.13995.

Heidrich, L., Brandl, R., Ammer, C., Bae, S., Bässler, C., Doerfler, I., Fischer, M., Gossner, M. M., Heurich, M., Heibl, C., Jung, K., Krzystek, P., Levick, S., Magdon, P., Schall, P., Schulze, E. D., Seibold, S., Simons, N. K., Thorn, S., … Müller, J. 2023. Effects of heterogeneity on the ecological diversity and redundancy of forest fauna. Basic and Applied Ecology, 73, 72–79. https://doi.org/10.1016/j.baae.2023.10.005.

Heinaro, E., Tanhuanpää, T., Yrttimaa, T., Holopainen, M., Vastaranta, M. 2021. Airborne laser scanning reveals large tree trunks on forest floor. For. Ecol. Manage., 491, 119225–. https://doi.org/10.1016/j.foreco.2021.119225.

Herniman, S., Coops, N., Martin, K., Thomas, P., Luther, J., van Lier, O. 2020. Modelling avian habitat suitability in boreal forest using structural and spectral remote sensing data. Remote Sens. Appl., 19, 100344–. https://doi.org/10.1016/j.rsase.2020.100344.

Hill, R.A., Broughton, R.K. 2009. Mapping the understorey of deciduous woodland from leaf-on and leaf-off airborne LiDAR data: A case study in lowland Britain. ISPRS J. Photogramm. Remote Sens. 64 (2), 223–233. https://doi.org/10.1016/j.isprsjprs.2008.12.004.

Hill, R. A., Hinsley, S. A., Broughton, R. K. 2013. Assessing habitats and organism-habitat relationships by airborne laser scanning. In: Forestry Applications of Airborne Laser Scanning, 335–356. Springer Netherlands. https://doi.org/10.1007/978-94-017-8663-8_17.

Hinsley, S.A., Hill, R.A., Bellamy, P.E., Balzter, H., 2006. The application of lidar in woodland bird ecology: Climate, canopy structure, and habitat quality. Photogramm. Eng. Remote Sens. 72 (12), 1399–1406. https://doi.org/10.14358/PERS.72.12.1399.

Hyyppä, J., Kelle, O., Lehikoinen, M., Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. IEEE Trans. Geosci. Remote Sens. 39 (5), 969–975. https://doi.org/10.1109/36.921414.

International Union for Conservation of Nature. 2023. IUCN red list of threatened species. Available from https://www.iucn.org/resources/conservation-tool/iucn-red-list-threatened-species [assessed 12 August 2024].

Jayathunga, S., Owari, T., Tsuyuki, S., 2018. Analysis of forest structural complexity using airborne LiDAR data and aerial photography in a mixed conifer–broadleaf forest in northern Japan. J. For. Res. 29 (2), 479–493. https://doi.org/10.1007/s11676-017-0441-4.

Jensen, J. R., Qiu, F., Ji, M. 1999. Predictive modelling of coniferous forest age using statistical and artificial neural network approaches applied to remote sensor data. International Journal of Remote Sensing, 20(14), 2805–2822. https://doi.org/10.1080/014311699211804.

Jetz, W., D. S. Wilcove, A. P. Dobson. 2007. Projected impacts of climate and land-use change on the global diversity of birds, PLoS Biol., 5, 1211 – 1219, https://doi:10.1371/journal.pbio.0050157

Jonsell, M., Weslien, J., Ehnstrom, B. 1998. Substrate requirements of red-listed saproxylic invertebrates in Sweden. Biodivers. Conserv. 7(6): 749–764. http://dx.doi.org/10.1023/A:1008888319031.

Jönsson, M.T., Jonsson, B.G., 2007. Assessing coarse woody debris in Swedish woodland key habitats: Implications for conservation and management. For. Ecol. Manage. 242 (2), 363–373. https://doi.org/10.1016/j.foreco.2007.01.054.

Kamoske, A.G., Dahlin, K.M., Read, Q.D., Record, S., Stark, S.C., Serbin, S.P., Zarnetske, P.L., Dornelas, M., 2022. Towards mapping biodiversity from above: Can fusing lidar and hyperspectral remote sensing predict taxonomic, functional, and phylogenetic tree diversity in temperate forests? Glob. Ecol. Biogeogr. 31 (7), 1440–1460. https://doi.org/10.1111/geb.13516.

Kay, C.E. 1997. Is Aspen doomed? J. For. 95(5): 4–11.

Kivinen, S., Koivisto, E., Keski-Saari, S., Poikolainen, L., Tanhuanpää, T., Kuzmin, A., et al. 2020. A keystone species, European aspen (Populus tremula L.), in boreal forests: ecological role, knowledge needs and mapping using remote sensing. For. Ecol. Manage. 462: 118008–. http://dx.doi.org/10.1016/j.foreco.2020.118008.

Koivuniemi, J., Korhonen, K.T. 2006. Inventory by compartments. In: Kangas, A., Maltamo, M. (eds) Forest inventory. Methodology and applications. Managing forest ecosystems, vol 10. Springer, Dordrecht, pp 271–278.

Korhonen, K.T., Ahola, A., Heikkinen, J., Henttonen, H.M., Hotanen, J.-P., Ihalainen, A., et al. 2021. Forests of Finland 2014–2018 and their development 1921–2018. Silva Fenn. 55(5). http://dx.doi.org/10.14214/sf.10662.

Korpela, I., Ørka, H. O., Maltamo, M., Tokola, T., 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 44(2), 319–339. https://doi.org/10.14214/sf.156.

Kouki, J., Arnold, K., and Martikainen, P. 2004. Long-term persistence of aspen – a key host for many threatened species – is endangered in old-growth conservation areas in Finland. J. Nat. Conserv. 12(1): 41–52. http://dx.doi.org/10.1016/j.jnc.2003.08.002 .

Kuuluvainen, T. 2002. Natural variability of forests as a reference for restoring and managing biological diversity in boreal Fennoscandia. Silva Fenn. 36: 97–125. http://dx.doi.org/10.14214/sf.552.

Kuzmin, A., Korhonen, L., Kivinen, S., Hurskainen, P., Korpelainen, P., Tanhuanpää, T., Maltamo, M, Vihervaara, P., Kumpula, T. 2021. Detection of European aspen (Populus tremula L.) based on an unmanned aerial vehicle approach in boreal forests. Remote Sens. 13: 1723. https://doi.org/10.3390/rs13091723.

Latva-Karjanmaa, T., Penttilä, R., and Siitonen, J. 2007. Demographic structure of European aspen (Populus tremula) populations in managed and old-growth boreal forests in eastern Finland. Can. J. For. Res. 37(6): 1070–1081. http://dx.doi.org/10.1139/X06-289.

Laurance, W.F., Camargo, J.L.C., Fearnside, P.M., Lovejoy, T.E., Williamson, G.B., Mesquita, R.C.G., Laurance, S.G.W., 2017. An Amazonian rainforest and its fragments as a laboratory of global change. Biol. Rev. https://doi.org/10.1111/brv.12343.

Leiterer, R., Furrer, R., Schaepman, M., Morsdorf, F., 2015. Forest canopy-structure characterization: A data-driven approach. For. Ecol. Manage. 358, 48–61. https://doi.org/10.1016/j.foreco.2015.09.003.

Leutner, B.F., Reineking, B., Müller, J., Bachmann, M., Beierkuhnlein, C., Dech, S., Wegmann, M., 2012. Modelling forest α-diversity and floristic composition — On the added value of LiDAR plus hyperspectral remote sensing. Remote Sens. (Basel, Switzerland) 4 (9), 2818–2845. https://doi.org/10.3390/rs4092818.

Lindberg, E., Olofsson, K., Holmgren, J., Olsson, H., 2012. Estimation of 3D vegetation structure from waveform and discrete return airborne laser scanning data. Remote Sens. Environ. 118, 151–161. https://doi.org/10.1016/j.rse.2011.11.015.

Lindberg, E., Hollaus, M., Mücke, W., Fransson, J., Pfeifer, N. 2013. Detection of lying tree stems from airborne laser scanning data using a line template matching algorithm. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. II-5/W2, 169–174. https://doi.org/10.5194/isprsannals-ii-5-w2-169-2013.

Lindberg, E., Roberge, J.-M., Johansson, T., Hjältén, J., 2015. Can airborne laser scanning (ALS) and forest estimates derived from satellite images be used to predict abundance and species richness of birds and beetles in boreal forest? Remote Sens. (Basel, Switzerland) 7 (4), 4233–4252. https://doi.org/10.3390/rs70404233.

MacArthur, R., MacArthur, J.W., 1961. On bird species diversity. Ecology 42, 594–598. https://doi.org/10.2307/1932254.

Maltamo, M., Packalén, P., Suvanto, A., Korhonen, K. T., Mehtätalo, L., Hyvönen, P. 2009. Combining ALS and NFI training data for forest management planning: a case study in Kuortane, Western Finland. European Journal of Forest Research, 128(3), 305–317. https://doi.org/10.1007/s10342-009-0266-6.

Maltamo, M., Næsset, E., Vauhkonen, J. 2014. Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies (2014th ed., Vol. 27). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-8663-8.

Maltamo, M., Packalen, P. 2014. Species-specific management inventory in Finland. In Forestry applications of airborne laser scanning. Dordrecht: Springer Netherlands. pp. 241–252. http://dx.doi.org/10.1007/978-94-017-8663-8_12.

Maltamo, M., Pesonen, A., Korhonen, L., Kouki, J., Vehmas, M., Eerikäinen, K. 2015. Inventory of aspen trees in spruce dominated stands in conservation area. For. Ecosyst. 2, 12. https://doi.org/10.1186/s40663-015-0037-4.

Maltamo, M., Kinnunen, H., Kangas, A., Korhonen, L. 2020. Predicting stand age in managed forests using National Forest Inventory field data and airborne laser scanning. Forest Ecosystems, 7(1), 1–11. https://doi.org/10.1186/s40663-020-00254-z.

Martin, M., Valeria, O. 2022. “Old” is not precise enough: Airborne laser scanning reveals age-related structural diversity within old-growth forests. Remote Sens. Environ., 278, 113098–.

Martinuzzi, S., Vierling, L.A., Gould, W.A., Falkowski, M.J., Evans, J.S., Hudak, A.T., Vierling, K.T., 2009. Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. Remote Sens. Environ. 113, 2533–2546. https://doi.org/10.1016/j.rse.2009.07.002.

Martinuzzi, S., Gould, W.A., Vierling, L.A., Hudak, A.T., Nelson, R.F., Evans, J.S. 2013. Quantifying tropical dry forest type and succession: substantial improvement with LiDAR. Biotropica 45 (2), 135–146. https://doi.org/10.1111/j.1744-7429.2012.00904.x.

Mauya, E. 2015. Methods for Estimating Volume, Biomass and Tree Species Diversity Using Field Inventory and Airborne Laser Scanning in the Tropical Forests of Tanzania. Ph.D. Thesis, Norwegian University of Life Sciences, Ås, Norway, 2015; p. 54.

Melin, M., Matala, J., Mehtätalo, L., Tiilikainen, R., Tikkanen, O.-P., Maltamo, M., Pusenius, J., Packalen, P., 2014. Moose (Alces alces) reacts to high summer temperatures by utilizing thermal shelters in boreal forests – an analysis based on airborne laser scanning of the canopy structure at moose locations. Glob. Chang. Biol. 20 (4), 1115–1125. https://doi.org/10.1111/gcb.12405.

Melin, M., Hill, R.A., Bellamy, P.E., Hinsley, S.A., 2019. On bird species diversity and remote sensing-utilizing lidar and hyperspectral data to assess the role of vegetation structure and foliage characteristics as drivers of avian diversity. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12 (7), 2270–2278. https://doi.org/10.1109/JSTARS.2019.2906940.

Mikhail, E., Bethel, J., and McGlone, J., 2001. Introduction to modern photogrammetry. John Wiley & Sons, New York, pp. 479.

Millenium Ecosystem Assessment, 2005. Ecosystems and Human Well-Being: Synthesis. Island Press, Washington, DC.

Miura, N., Jones, S., 2010. Characterizing forest ecological structure using pulse types and heights of airborne laser scanning. Remote Sens. Environ. 114 (5), 1069–1076. https://doi.org/10.1016/j.rse.2009.12.017.

Mohammadi, J., Shataee, S., Næsset, E. 2020. Modeling tree species diversity by combining ALS data and digital aerial photogrammetry. Sci. Remote Sens. 2, 100011–. https://doi.org/10.1016/j.srs.2020.100011.

Morsdorf, F., Marell, A., Koetz, B., Cassagne, N., Pimont, F., Rigolot, E., Allgower, B. 2010. Discrimination of vegetation strata in a multi-layered Mediterranean forest ecosystem using height and intensity information derived from airborne laser scanning. Remote Sens Environ 114, 1403–1415. https://doi.org/10.1016/j.rse.2010.01.023.

Mücke, W., Hollaus, M., Pfeifer, N. 2012. Identification of dead trees using small footprint fullwaveform airborne laser scanning data. Silvilaser 2012.

Mücke, W., Deák, B., Schroiff, A., Hollaus, M., Pfeifer, N., 2013. Detection of fallen trees in forested areas using small footprint airborne laser scanning data. Can. J. Remote. Sens. 39 (sup1), S32–S40. https://doi.org/10.5589/m13-013.

Müller, J., Brandl, R., 2009. Assessing Biodiversity by Remote Sensing in Mountainous Terrain: The Potential of LiDAR to Predict Forest Beetle Assemblages. J. Appl. Ecol. 46 (4), 897–905. https://doi.org/10.1111/j.1365-2664.2009.01677.x.

Mura, M., McRoberts, R., Chirici, G., Marchetti, M. 2015. Estimating and mapping forest structural diversity using airborne laser scanning data. Remote Sens. Environ. 170, 133–142. https://doi.org/10.1016/j.rse.2015.09.016.

Myllymäki, M., Tuominen, S., Kuronen, M., Packalen, P., Kangas, A. 2023. The relationship between forest structure and naturalness in the Finnish national forest inventory. Forestry (London). https://doi.org/10.1093/forestry/cpad053.

Naesset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. 80 (1), 88–99. https://doi.org/10.1016/S0034-4257(01)00290-5.

Næsset, E. 2014. Area-based inventory in Norway – from innovation to an operational reality. In Forestry applications of airborne laser scanning. Springer Netherlands, Dordrecht. pp. 215–240. http://dx.doi.org/10.1007/978-94-017-8663-8_11.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D’Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., Kassem, K. R. 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. Bioscience 51(11), 933–938. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Packalén, P., Suvanto, A., Maltamo, M. 2009. A two stage method to estimate species-specific growing stock by combining ALS data and aerial photographs of known orientation parameters. Photogramm. Eng. Remote Sens. 75(12): 1451–1460. http://dx.doi.org/10.14358/PERS.75.12.1451.

Packalen, P., Heinonen, T., Pukkala, T., Vauhkonen, J., Maltamo, M. 2011. Dynamic treatment units in eucalyptus plantation. *Forest Science*, *57*(5), 416–426. https://doi.org/10.1093/forestscience/57.5.416.

Pan, Y., Chen, J.M., Birdsey, R., McCullough, K., He, L., Deng, F. 2011. Age structure and disturbance legacy of north american forests. Biogeosciences 8, 715–732. https://doi.org/10.5194/bg-8-715-2011.

Pascual, C., García-Abril, A., García-Montero, L.G., Martín-Fernández, S., Cohen, W.B. 2008. Object-based semi-automatic approach for forest structure characterization using lidar data in heterogeneous Pinus sylvestris stands. For. Ecol. Manage. 255 (11), 3677–3685. https://doi.org/10.1016/j.foreco.2008.02.055.

Pesonen, A., Maltamo, M., Eerikäinen, K., Packalén, P., 2008. Airborne laser scanning-based prediction of coarse woody debris volumes in a conservation area. For. Ecol. Manage. 255 (8), 3288–3296. https://doi.org/10.1016/j.foreco.2008.02.017.

Pesonen, A., Kangas, A., Maltamo, M., Packalén, P., 2010. Effects of auxiliary data source and inventory unit size on the efficiency of sample-based coarse woody debris inventory. For. Ecol. Manage. 259 (10), 1890–1899. https://doi.org/10.1016/j.foreco.2010.02.001.

Petchey, O. L., Hector, A., Gaston, K. J. 2004. How do different measures of functional diversity perform? Ecology, 85(3), 847–857. https://doi.org/10.1890/03-0226.

Pippuri, I., Suvanto, A., Maltamo, M., Korhonen, K.T., Pitkänen, J., Packalen, P., 2016. Classification of forest land attributes using multi-source remotely sensed data. Int. J. Appl. Earth Obs. Geoinf. 44, 11–22. https://doi.org/10.1016/j.jag.2015.07.002.

Pitkänen, T. P., Balazs, A., Tuominen, S. 2024. Automatized Sentinel-2 mosaicking for large area forest mapping. International Journal of Applied Earth Observation and Geoinformation, 127, 103659-. https://doi.org/10.1016/j.jag.2024.103659.

Racine, E. B., Coops, N. C., St-Onge, B., Begin, J. 2014. Estimating forest stand age from LiDAR-derived predictors and nearest neighbor imputation. Forest Science, 60(1), 128–136. https://doi.org/10.5849/forsci.12-088.

Rogers, B. M., Mackey, B., Shestakova, T. A., Keith, H., Young, V., Kormos, C. F., DellaSala, D. A., Dean, J., Birdsey, R., Bush, G., Houghton, R. A., Moomaw, W. R. 2022. Using ecosystem integrity to maximize climate mitigation and minimize risk in international forest policy. Frontiers in Forests and Global Change, 5. https://doi.org/10.3389/ffgc.2022.929281.

Sankey, T., 2012. Decadal-scale aspen changes: evidence in remote sensing and tree ring data: Decadal-scale aspen changes. Appl. Veg. Sci. 15 (1), 84–98. https://doi.org/10.1111/j.1654-109X.2011.01141.x.

Sabatini, F. M., Bluhm, H., Kun, Z., Aksenov, D., Atauri, J. A., Buchwald, E., Burrascano, S.,
Cateau, E., Diku, A., Duarte, I. M., et al. 2020. European Primary Forest Database (EPFD) v2.0.
bioRxiv. https://doi.org/10.1038/s41597-021-00988-7.

Sasaki, T., Imanishi, J., Fukui, W., Morimoto, Y., 2016. Fine-scale characterization of bird habitat using airborne LiDAR in an urban park in Japan. Urban For. Urban Green. 17, 16–22. https://doi.org/10.1016/j.ufug.2016.03.007.

Schmitt, S., Marechaux, I., Chave, J., Fischer, F., Piponiot, C., Traissac, S., Herault, B. 2019. Functional diversity improves tropical forest resilience: Insights from a long-term virtual experiment. Journal of Ecology, 108(3), 831–843. https://doi.org/10.1111/1365-2745.13320.

Schneider, F. D., Longo, M., Paul-Limoges, E., Scholl, V. M., Schmid, B., Morsdorf, F., Pavlick, R. P., Schimel, D. S., Schaepman, M. E., Moorcroft, P. R. 2023. Remote Sensing-Based Forest Modeling Reveals Positive Effects of Functional Diversity on Productivity at Local Spatial Scale. Journal of Geophysical Research. Biogeosciences, 128(6). https://doi.org/10.1029/2023JG007421.

Schumacher, J., Hauglin, M., Astrup, R., Breidenbach, J. 2020. Mapping forest age using National Forest Inventory, airborne laser scanning, and Sentinel-2 data. Forest Ecosystems, 7(1), 1–14. https://doi.org/10.1186/s40663-020-00274-9.

Simonson, W.D., Allen, H.D., Coomes, D.A., 2013. Remotely sensed indicators of forest conservation status: Case study from a Natura 2000 site in southern Portugal. Ecol. Ind. 24, 636–647. https://doi.org/10.1016/j.ecolind.2012.08.024.

Singh, K.K., Davis, A.J., Meentemeyer, R.K. 2015. Detecting understory plant invasion in urban forests using LiDAR. Int. J. Appl. Earth Obs. Geoinf. 38, 267–279. https://doi.org/10.1016/j.jag.2015.01.012.

Stein, A., Gerstner, K., Kreft, H., Arita, H. 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. Ecology Letters, 17 (7), 866–880. https://doi.org/10.1111/ele.12277.

Stitt, J. M., Hudak, A. T., Silva, C. A., Vierling, L. A., Vierling, K. T. 2022. Evaluating the use of lidar to discern snag characteristics important for wildlife. Remote Sens. (Basel, Switzerland), 14(3), 720–. https://doi.org/10.3390/rs14030720.

Straub, C., Koch, B. 2011. Enhancement of bioenergy estimations within forests using airborne laser scanning and multispectral line scanner data. Biomass & Bioenergy, 35(8), 3561–3574. https://doi.org/10.1016/j.biombioe.2011.05.017.

Sumnall, M., Fox, T.R., Wynne, R.H., Thomas, V.A. 2017. Mapping the height and spatial cover of features beneath the forest canopy at small-scales using airborne scanning discrete return Lidar. ISPRS J. Photogramm. Remote Sens. 133, 186–200. https://doi.org/10.1016/j.isprsjprs.2017.10.002.

Sverdrup-Thygeson, A., Ørka, H.O., Gobakken, T., Næsset, E. 2016. Can airborne laser scanning assist in mapping and monitoring natural forests? For. Ecol. Manage. 369, 116–125. https://doi.org/10.1016/j.foreco.2016.03.035.

Tikkanen, O.-P., Martikainen, P., Hyvärinen, E., Junninen, K., Kouki, J. 2006. Red-listed boreal forest species of Finland: associations with forest structure, tree species, and decaying wood. Ann. Zool. Fenn. 43(4): 373–383.

Torresan, C., Corona, P., Scrinzi, G., Marsal, J. V. 2016. Using classification trees to predict forest structure types from LiDAR data. Ann. For. Res. 59(2), 281–298. https://doi.org/10.15287/afr.2016.423.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., Steininger, M., 2003. Remote sensing for biodiversity science and conservation. Trends Ecol. Evol. (Amsterdam) 18 (6), 306–314. https://doi.org/10.1016/S0169-5347(03)00070-3.

United Nations. 1992. Convention on Biological Diversity. Retrieved from: www.cbd.int/convention/text/.

Valbuena, R., Vauhkonen, J., Packalen, P., Pitkänen, J., Maltamo, M. 2014. Comparison of airborne laser scanning methods for estimating forest structure indicators based on Lorenz curves. ISPRS J. Photogramm. Remote Sens. 95, 23–33. https://doi.org/10.1016/j.isprsjprs.2014.06.002.

Valbuena, R., Maltamo, M., Packalen, P. 2016. Classification of multilayered forest development classes from low-density national airborne lidar datasets. Forestry (London) 89 (4), 392–401. https://doi.org/10.1093/forestry/cpw010.

van der Plas, F. 2019. Biodiversity and ecosystem functioning in naturally assembled communities. Biological Reviews, 94(4). https://doi.org/10.1111/brv.12499.

Vaughn, N.R., Asner, G.P., Giardina, C.P. 2014. Centennial impacts of fragmentation on the canopy structure of tropical montane forest. Ecol. Appl. 24 (7), 1638–1650. https://doi.org/10.1890/13-1568.1.

Vehmas, M., Eerikäinen, K., Peuhkurinen, J., Packalén, P., Maltamo, M. 2009. Identification of boreal forest stands with high herbaceous plant diversity using airborne laser scanning. For. Ecol. Manage. 257 (1), 46–53. https://doi.org/10.1016/j.foreco.2008.08.016.

Vierling, K.T., Vierling, L.A., Gould, W.A., Martinuzzi, S., Clawges, R.M. 2008. Lidar: shedding new light on habitat characterization and modeling. Front. Ecol. Environ. 6 (2), 90–98. https://doi.org/10.1890/070001.

Vierling, K.T., Vierling, L.A., Bässler, C., Brandl, R., Weiß, I., Müller, J. 2011. Spinning a laser web: predicting spider distributions using LiDAR. Ecol. Appl. 21 (2), 577–588. https://doi.org/10.1890/09-2155.1.

Viinikka, A., Hurskainen, P., Keski-Saari, S., Kivinen, S., Tanhuanpää, T., Mäyrä, J., Poikolainen, L., Vihervaara, P., Kumpula, T. 2020. Detecting European aspen (Populus tremula L.) in boreal forests using airborne hyperspectral and airborne laser scanning data. Remote Sens. 12: 2610. https://doi.org/10.3390/RS12162610.

Vogeler, J.C., Hudak, A.T., Vierling, L.A., Vierling, K.T., 2013. Lidar-derived canopy architecture predicts brown creeper occupancy of two western coniferous forests. Condor 115 (3), 614–622. https://doi.org/10.1525/cond.2013.110082.

Vogeler, J., Hudak, A., Vierling, L., Evans, J., Green, P., Vierling, K., 2014. Terrain and vegetation structural influences on local avian species richness in two mixed-conifer forests. Remote Sens. Environ. 147, 13–22. https://doi.org/10.1016/j.rse.2014.02.006.

Vogeler, A.B., Otte, I., Ferger, S., Helbig-Bonitz, M., Hemp, A., Nauss, T., Böhning-Gaese, K., Schleuning, M., Tschapka, M., Albrecht, J., 2022. Associations of bird and bat species richness with temperature and remote sensing-based vegetation structure on a tropical mountain. Biotropica 54 (1), 135–145. https://doi.org/10.1111/btp.13037.

Wallis, C.I.B., Paulsch, D., Zeilinger, J., Silva, B., Curatola Fernández, G.F., Brandl, R., Farwig, N., Bendix, J., 2016. Contrasting performance of Lidar and optical texture models in predicting avian diversity in a tropical mountain forest. Remote Sens. Environ. 174, 223–232. https://doi.org/10.1016/j.rse.2015.12.019.

Wilkes, P., Jones, S.D., Suarez, L., Haywood, A., Mellor, A., Woodgate, W., Soto-Berelov, M., Skidmore, A.K., McMahon, S. 2016. Using discrete-return airborne laser scanning to quantify number of canopy strata across diverse forest types. Methods Ecol. Evol. 7 (6), 700–712. https://doi.org/10.1111/2041-210X.12510.

Wing, B., Ritchie, M., Boston, K., Cohen, W., Olsen, M., 2015. Individual snag detection using neighborhood attribute filtered airborne lidar data. Remote Sens. Environ. 163, 165–179. https://doi.org/10.1016/j.rse.2015.03.013.

WWF – World Wide Fund for Nature. 2021. Terrestrial Ecoregions of the World. Retrieved from: https://www.worldwildlife.org/publications/terrestrial-ecoregions-of-the-world.

Zellweger, F., Braunisch, V., Baltensweiler, A., Bollmann, K., 2013. Remotely sensed forest structural complexity predicts multi species occurrence at the landscape scale. For. Ecol. Manage. 307, 303–312. https://doi.org/10.1016/j.foreco.2013.07.023.

Zellweger, F., Baltensweiler, A., Ginzler, C., Roth, T., Braunisch, V., Bugmann, H., Bollmann, K., 2016. Environmental predictors of species richness in forest landscapes: abiotic factors versus vegetation structure. J. Biogeogr. 43 (6), 1080–1090. https://doi.org/10.1111/jbi.12696.

Zellweger, F., Roth, T., Bugmann, H., Bollmann, K., Kerr, J., 2017. Beta diversity of plants, birds and butterflies is closely associated with climate and habitat structure. Glob. Ecol. Biogeogr. 26 (8), 898–906. https://doi.org/10.1111/geb.12598.

Zhou, T., Chen, B. M., Liu, G., Huang, F. F., Liu, J. G., Liao, W. B., Wang, Y. Y., Ren, S. J., Chen, C. Q., Peng, S. L. 2015. Biodiversity of Jinggangshan Mountain: The importance of topography and geographical location in supporting higher biodiversity. PloS One, 10(3), e0120208–e0120208. https://doi.org/10.1371/journal.pone.0120208.

Zimble, D.A., Evans, D.L., Carlson, G.C., Parker, R.C., Grado, S.C., Gerard, P.D. 2003. Characterizing vertical forest structure using small-footprint airborne LiDAR. Remote Sens. Environ. 87 (2), 171–182. https://doi.org/10.1016/S0034-4257(03)00139-1.

Zupanc, A. 2017. Improving Cloud Detection with Machine Learning. Sentinel Hub Blog. https://medium.com/sentinel-hub/improving-cloud-detection-with-machine-learning-c09dc5d7cf13 (accessed 14 August 2024).